



Centrus™ Desktop User Guide

Version 4.0

Copyright 2001-2002, All rights reserved, Sagent Technology, Inc.
National Zip+4 Directory Copyright 2001-2002 United States Postal Service.

Trademark Notices

ZIP Code and ZIP+4 are registered trademarks of the U.S. Postal Service.

Within this manual, other trademarked names may be used. In every occurrence, the names are used editorially and to the benefit of the trademark owner, with no intent of infringement on the trademark.

USPS Notice

Sagent Technology, Inc. holds a nonexclusive license to publish and sell ZIP+4 databases on optical and magnetic media. The price of the Sagent product is neither established, controlled, nor approved by the U.S. Postal Service.

Corporate Office

800 W. El Camino Real, Suite 300
Mountain View, CA 94040
Phone: 650-493-7100
Fax: 650-815-3500

Development Office

4750 Walnut, Suite 200
Boulder, CO 80301-2538
Phone: 303-442-2838
Fax: 303-440-3523

Contents

Preface	ix
About Centrus Desktop	ix
Centrus Desktop Components	x
Address Coding Module	x
Data Locator Module	xi
Demographic Coding Module	xi
Point-in-Polygon Module	xi
Closest Site Module	xii
Geographic Determination Module	xii
Primary Market Area Function	xii
About This Guide	xii
Typographic Conventions	xiii
Assumptions	xiii
Chapter Descriptions	xiii
Contacting Technical Support	xv
Chapter 1: Installation	1
Installation Requirements	1
Local Versus Server Installation	2
Licensing	2
Demonstration License	2
License Renewal	3
Installation Procedure	3
Step 1 – Install Adobe Acrobat Reader	3
Step 2 – Install the Product	3
Step 3 – Install the Data Files	4
Chapter 2: Configuration	7
Specifying the Search Path for the Data Files	7
Data Files for the Address Coding Module	8
Data Files for the Data Locator Module	10
Data Files for the Demographics Module	11
Data File for the Geographic Determination Module	11
Setting Processing Options	12
Address Coding Options Tab	12
Data Locator Options Tab	21
Point-in-Polygon Tab	23
Closest Site Tab	27
Geographic Determination Tab	31
Verifying Your Settings	37
Using Command Line Syntax	37

Working with Tasks	38
Creating a New Task	38
Saving a Task	38
Opening a Task	39
Editing a Task	39
Renaming a Task	39
Chapter 3: Centrus Desktop Tutorial	41
Starting Centrus Desktop	42
Setting Up a Task.	43
Specify Input and Output Tables	43
Specify Address Coding Inputs and Outputs	44
Specify Demographic Inputs and Outputs	46
Specify Point-in-Polygon Inputs and Outputs	47
Specify Closest Site Inputs and Outputs	48
Specify Data Locator Inputs and Outputs	51
Specify Geographic Determination Inputs and Outputs.	52
Processing a Task	55
Creating a Primary Market Area (PMA)	57
Chapter 4: Centrus Desktop Tools	61
QuickFind	61
Data Locator Mode	62
Standard Mode	63
Multiline Mode	64
Address Wizard Mode	65
The Query Function	67
Data Viewer	74
Map Viewer	75
Map Viewer Buttons	75
Manual Geocoding with the Map Viewer	77
Working with Layers.	78
Increasing Centrus Desktop Mapper Performance	81
Shape Layers	81
Importing Geography Files	81
Exporting Geography Files	82
Chapter 5: File Specification	85
File Types Used in Centrus Desktop	85
Preparing Data Files for Processing	86
Access, dBASE, and FoxPro	86
ASCII Text Files	86
Excel Files	86
Specifying Input and Output Tables	87
Specifying an Input Table	87
Specifying an Output Table	88
Specifying a Reject File	89
Report Files	90

Chapter 6:	Primary Market Area Function	93
	Creating Store-Based PMAs	94
	Creating Customer-Based PMAs	96
Chapter 7:	The Address Coding Module	101
	About the Address Coding Module	101
	U.S. Address Processing	102
	Canadian Address Processing	103
	Z4 Change Processing Option	103
	Address Line Two Preferences	104
	Specifying Address Input Fields	104
	Using Different Input Formats	105
	Multiline Input	105
	General Guidelines for Using Multiline Mode	106
	Process if Unmatched or Field Is Blank	106
	Optimization Sort	107
	Assigning Address Output Fields	107
	Assigning Address Elements	107
	Creating New Fields	108
	Unassign	108
	Standardization Defaults	108
	Geocode Defaults	109
	Blank if Unmatched	109
	Available Address Data Elements	110
Chapter 8:	The Data Locator Module	115
	About the Data Locator Module	115
	How Data Locator Works	116
	Specifying Data Locator Input Fields	116
	Process if Field Is Blank	117
	Assigning Data Locator Output Fields	118
	Assigning Field Names	118
	Tips for Using Data Locator	119
	Decide Which Information You Want to Extract	119
	Define a Task to Extract the Information	119
	Evaluate Results and Redefine the Task	120
	Using the Custom Dictionary	120
	Using Multiple Passes and Process if Field Is Blank	120
	Available Data Locator Elements	121
Chapter 9:	The Demographic Coding Module	125
	Specifying Demographic Input Fields	126
	Process if Field Is Blank	126
	Assigning Demographic Output Fields	126
	Assigning Field Names	127
	Available Demographics Elements	129

Chapter 10: The Point-in-Polygon Module	131
Object Files and Layers	131
Point-in-Polygon Analysis	132
Specifying Point-in-Polygon Input Fields	133
Buffer Width (Feet)	133
Process if Field Is Blank	133
Assigning Point-in-Polygon Output Fields	134
Available Point-in-Polygon Elements	134
Assigning Field Names	135
Creating New Fields	136
Unassign	136
Blank if Unmatched	136
Chapter 11: The Closest Site Module	137
Object Files and Layers	137
Closest Site Analysis	139
Specifying Closest Site Input Fields	139
Maximum Distance (Miles)	139
Process if Field Is Blank	140
Assigning Closest Site Output Fields	140
Available Closest Site Elements	140
Assigning Field Names	141
Creating New Fields	141
Unassign	141
Blank if Unmatched	142
Chapter 12: The Geographic Determination Module	143
Using the Geographic Determination Module	144
Search Distance (feet)	144
Buffer Width (feet)	145
Assigning Output Fields	146
Assigning Fields	146
Creating New Fields	147
Unassign	147
Blank if Unmatched	147
Available Geographic Determination Elements	148
Chapter 13: File Processing	149
About File Processing	149
Interactive Processing	150
Batch Processing	150
Process Dialog Box	151
Possible Matches Dialog Box	154
Appendix A: System Messages and Codes	155
Match Codes	155
Match Codes for the United States	156

Match Codes for Canada	158
Location Codes	160
Address Location Codes	161
ZIP+4 Centroid Location Codes	161
GeoStan Canada Location Codes	165
.	165
Appendix B: Early Warning System Data	167
Appendix C: Reference	169
Customizing the Data Locator Module	169
Regular Expressions	170
More About Regular Expressions	171
Custom Dictionaries in Data Locator	172
Premium Demographics	175
Claritas Demographics Data	175
Claritas PRIZM Codes	180
WinSplit	187
Centrus Data Formatter and ASCII Text Files	187
About ASCII Text Files	188
Using Centrus Data Formatter	188
CASS Overview	195
Using CASS to Reduce USPS Bulk Mailing Rates	195
The CASS 3553 Report	196
Where to Get Help	198
Audit and Log Report Samples	198
Audit Report	199
Log Report	200
Canadian Address Accuracy Program	203
Statement of Accuracy	203
Valid and Correctable Addresses	204
Where to Get Help	204
Glossary	205
Index	209

Preface

This preface provides an introduction to Centrus Desktop and the *Centrus Desktop User Guide*. It contains the following information:

- About Centrus Desktop
- Centrus Desktop Components
- About This Guide
- Contacting Technical Support

About Centrus Desktop

Your organization spends thousands of dollars each year collecting the names and addresses of your customers and prospects. Unfortunately, collecting this information is only part of the battle. In today's aggressive marketplace you must know more about your customers to do a better job of targeting them. Centrus Desktop helps you learn more about your customers by properly fielding and cleansing your data and adding valuable information, such as:

- Age range
- Income range
- Ethnicity
- Gender
- Presence of children
- Latitude and longitude

- Census codes
- Postal codes
- Distance to user-defined locations

Centrus Desktop correctly fields and parses, cleans, standardizes, geocodes, spatially codes, and demographically codes your data *in a single pass*. It can perform either interactive or high speed batch processing, so there is no need to tie up your mainframe.

Centrus Desktop works natively with a variety of data formats, including text (delimited and fixed-length ASCII), SQL Server, Oracle, dBase, Access, Excel, and FoxPro.

Centrus Desktop Components

Centrus Desktop uses a flexible, modular design that lets you license any or all of the components your organization needs. You can choose among the following modules:

- Demographic Coding Module
- Data Locator Module
- Demographic Coding Module
- Point-in-Polygon Module
- Closest Site Module
- Geographic Determination Module

Centrus Desktop also includes a Primary Market Area Function.

Address Coding Module

The Address Coding module cleanses addresses to the highest level, exceeding the U.S. Postal Service's Coding Accuracy Support System (CASS) and Canada Post Corporation's Software Evaluation and Recognition Program (SERP) standards. Because Centrus Desktop is CASS and SERP certified, you are assured that mailings will be correctly delivered. With over 20 million addresses changing every year, Sagent's bimonthly updates ensure the integrity and effectiveness of your valuable address data.

As Centrus Desktop corrects addresses, it incorporates a comprehensive address-level geographic coding function specifically designed to spatially enable your data. This process adds key components to each database record including latitude, longitude, Census Tract, Block Group, and more.

For more information about the Address Coding Module, see Chapter 7, "The Address Coding Module."

Data Locator Module

The Data Locator module scans your data and identifies name elements, address elements, and a user-defined expression (such as e-mail address, phone number, customer number, or Social Security number). It then quickly and accurately outputs the information to fields of your choice.

The Data Locator also identifies ethnicity and gender, and can generate appropriate salutations for your mailings. It includes a user-extensible dictionary that allows you to add additional name information, tailoring the Data Locator to your specific needs.

For more information about the Data Locator Module, see Chapter 8, “The Data Locator Module.”

Demographic Coding Module

The Demographic Coding module appends valuable demographic variables and cluster codes to your existing databases. Knowing your customers’ habits and lifestyles is critical to understanding them. Once you identify who your customers are, you can use this information to target your marketing efforts to similar groups or households, increasing your chances of reaching the correct audience.

Standard 2000 Census Demographics are included as a basic feature of the Demographic Coding module. Sagent also offers current year and five-year projected Block Group demographics and PRIZM™ and MicroVision™ cluster codes at a ZIP+4 level from Claritas. Cluster codes provide a simple way to identify, understand, and target consumers. They are widely used for strategic planning, customer profiling, direct mail, site analysis, media planning, and many other functions. Centrus Desktop appends demographic codes based on each address’s geographic location, using either Census Block Group or ZIP+4.

For more information about the Demographic Coding Module, see Chapter 9, “The Demographic Coding Module.”

Point-in-Polygon Module

The Point-in Polygon module lets you accurately determine which areas a given point falls within. For example, retail, insurance, and telecommunication industries use Point-in-Polygon to identify key markets, insurance risk areas, and telephone service areas.

For more information about the Point-in-Polygon Module, see Chapter 10, “The Point-in-Polygon Module.”

Closest Site Module

The Closest Site module provides a distance test between two sets of points. For example, the healthcare, banking, and insurance industries use this type of analysis to match doctors and patients, find the nearest bank branch to a customer, and calculate the distance between a home and a fire station.

For more information about the Closest Site Module, see Chapter 11, “The Closest Site Module.”

Geographic Determination Module

The Geographic Determination module eliminates any uncertainty (in the calculations of the initial geocode) by providing a confidence factor and associating it with the geocode so that additional rules may be applied. This confidence factor can be described as either a distance value (between two points and/or lines) or as a percentage representing the amount of overlap between two polygons.

Geographic determination describes both the spatial relationship between two entities and the degree of confidence in that description. High confidence determinations, based on any level of geocode (address, ZIP+4, ZIP+2, or ZIP code) are critical for many geographically sensitive business decisions.

For more information about the Geographic Determination Module, see Chapter 12, “The Geographic Determination Module.”

Primary Market Area Function

The Primary Market Area (PMA) function works with Spatial Coding tools to simply and quickly generate geographic market areas based on store locations, customer addresses, and other values. You can use your existing database with PMA to create a precise polygonal boundary file using the outermost address points of customers or prospects surrounding your store or office location.

For more information about the PMA function, see Chapter 6, “Primary Market Area Function.”

About This Guide

This section provides the following information to help you use this guide:

- Typographic Conventions
- Assumptions
- Chapter Descriptions

Centrus Desktop uses several data files, which Sagent provides on the *Data Products Suite CDs*. This user guide does not provide the locations of the data files on the *Data Products Suite CDs*. For descriptions and locations of these files, refer to the *Release Notes for the Centrus Data Products*.

Typographic Conventions

To distinguish text that appears in Centrus Desktop's user interface, this guide uses the following typographic conventions:

- Menu items and buttons are shown in **bold**, as are functions and dialog box options.
- Sections of dialog boxes are shown in *italics*.

The notation, **Process > Options**, indicates that you should click the **Process** menu and choose **Options** from the resulting drop-down menu.

Assumptions

This guide assumes that you are familiar with Windows and that you know how to use a mouse and choose dialog box options. If you are not comfortable with Windows, you might want to take the introductory tour:

1. Click **Start** on the Taskbar. Windows opens the start menu.
2. Click **Help**. The Windows help file opens.
3. If the **Contents** tab is not selected, click it now.
4. Double-click either **Tour: Ten Minutes to Using Windows** (Windows 98) or **Introducing Windows NT** (Windows NT).

Chapter Descriptions

This guide contains the following chapters:

- Chapter 1, "Installation," describes system requirements and how to install Centrus Desktop.
- Chapter 2, "Configuration," describes how to configure Centrus Desktop for processing with each of the available modules. It details which data files each module requires, as well as optional files for extended functionality.
- Chapter 3, "Centrus Desktop Tutorial," takes you step-by-step through the process of creating and processing a sample task that uses each module.
- Chapter 4, "Centrus Desktop Tools," describes how to use Quick Find, the Data Viewer, and the Map Viewer, Centrus Desktop's tools for finding records and manipulating processed data.
- Chapter 5, "File Specification," describes how to prepare and specify input, output, reject, and report files.

- Chapter 6, “Primary Market Area Function,” describes how to use the Primary Market Area (PMA) function to generate geographic market areas to identify market overlap, areas of “cannibalizing,” and pockets of untapped opportunity.
- Chapter 8, “The Data Locator Module,” describes how to use the Data Locator Module to analyze databases, identify and extract name elements, and add a variety of name-related information.
- Chapter 7, “The Address Coding Module,” describes how to use the Address Coding Module to standardize and geocode U.S. and Canadian addresses, thereby reducing bulk mailing costs and providing valuable information about your customers.
- Chapter 9, “The Demographic Coding Module,” describes how to use the Demographic Coding Module to append demographic data to the information in your database, allowing you to target marketing campaigns, forecast sales, determine market penetration, and assess market potential.
- Chapter 10, “The Point-in-Polygon Module,” describes how to use the Point-in-Polygon Module to determine in which area a given point falls, enabling you to identify key markets, assess risk, and determine service availability.
- Chapter 11, “The Closest Site Module,” describes how to use the Closest Site Module to determine the distance between an address and a site point, such as a bank branch or fire station, in order to identify the nearest site location to a customer or resident.
- Chapter 12, “The Geographic Determination Module,” describes how to use the Geographic Determination Module to determine the spatial relationship between two entities and describe the degree of confidence in that determination, essential information in geographically sensitive business decisions.
- Chapter 13, “File Processing,” describes how to process a task and how to view the results.
- Appendix A, “System Messages and Codes,” describes how to interpret the codes Centrus Desktop generates to provide detailed information about the matches retrieved, such as whether or not a match was found, information about the type of match found, information about why no match was found, and information about the geocode assigned.
- Appendix B, “Early Warning System Data,” provides information about the Early Warning System data the USPS provides to prevent matching errors due to the age of the address data in the Use.gsd and Usw.gsd files.
- Appendix C, “Reference,” provides reference information about customizing the Data Locator Module, using premium demographics, using WinSplit, using Centrus Data Formatter, and satisfying CASS and SERP requirements.

This guide also contains a Glossary and an Index.

Contacting Technical Support

If you have any technical questions regarding the use of Centrus Desktop, please contact Centrus Technical Support:

Sagent Technology, Inc.
4750 Walnut Street, Suite 200
Boulder CO 80301-2538
(303) 442-2838 (voice)
(303) 440-3523 (fax)
support@centrus.com
www.centrus.com

When contacting Centrus Technical Support, please provide the following information so that Sagent can be more effective in helping you:

- Product name
- Company name under which the product is registered
- Company name with which you are affiliated
- Telephone number at which you may be contacted
- Development environment (if applicable)
- Thorough description of the problem

Visit *update.centrus.com* for patches, FAQs, product manuals, and a comprehensive industry glossary.

Downloads and technical documentation are password protected on *update.centrus.com*. The password changes bimonthly. The current password is published in every *Release Notes* document, including the *Release Notes for the Centrus Data Products*.

Chapter 1

Installation

This chapter describes system requirements and how to install Centrus Desktop on the Windows platform. It includes the following information:

- Installation Requirements
- Local Versus Server Installation
- Licensing
- Installation Procedure

Installation Requirements

Centrus Desktop has the following minimum system requirements:

- A computer running Windows 98/2000/NT/XP
- Pentium processor or higher
- 16 megabytes of memory
- A hard disk with a minimum of 10 megabytes of free disk space (additional disk space required if copying database files from the CD to your hard drive)
- 3 1/2" floppy disk drive

- CD-ROM drive
- Mouse or other pointing device

Note: Hard disk space requirements vary with the Centrus Desktop modules and data files licensed. A fast disk drive, fast processor, and more memory improves performance.

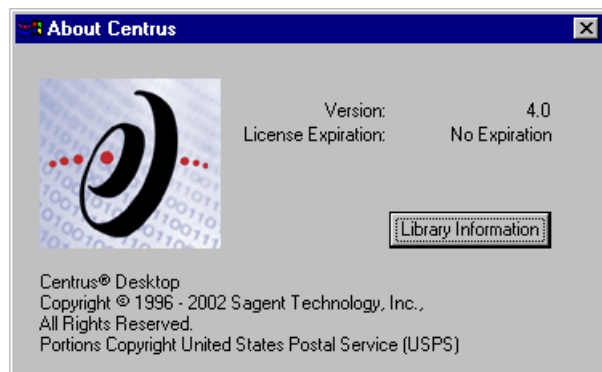
Local Versus Server Installation

Currently, executing Centrus Desktop in a client-server configuration is unsupported. Centrus Desktop is designed to be executed locally only, and should not be installed on a server.

Licensing

Centrus Desktop uses several data files, which Sagent provides on the *Data Products Suite* CDs. The data files you are licensed to use are installed during the data installation process. Refer to the *Release Notes for the Centrus Data Products* for data installation instructions.

To see version, license expiration, and data file library information, click **Help > About Centrus** from the main Centrus Desktop menu.



Click **Library Information** to see additional information about which data files are currently loaded. The **Library Info** window displays library version, release date, and a list of data files and their expiration dates.

Demonstration License

To examine modules for which you have no license, you may run Centrus Desktop in demonstration mode. In demo mode, all modules are fully accessible except for the Geographic Determination module, which is not available in demo mode. Processing is limited to interactive processing of 25 records per job, and you cannot produce a CASS report. A demo mode icon is

installed with Centrus Desktop and appears on your Windows **Start** menu in the Centrus Desktop program group. Click the **Centrus Desktop Demo** icon to start Centrus Desktop in demo mode. If you prefer to run Centrus Desktop from the command line, add the switch “/d” to invoke demo mode.

License Renewal

License renewal is automatic; you do not need to take any action to renew your license. You must, however, provide written notice to Sagent if you do not wish to renew. Written notice of your intent not to renew must be received by Sagent no less than 60 days prior to the expiration of your current term.

Installation Procedure

The installation procedure depends on whether you are installing Centrus Desktop for the first time, or you have a previous version already installed on your machine. If you have a previous version of Centrus Desktop, the installation program detects the previously installed version. The installation program then gives you the option of updating the existing files or installing the new version into a different location.

The procedure for installing Centrus Desktop includes:

- Step 1 – Install Adobe Acrobat Reader
- Step 2 – Install the Product
- Step 3 – Install the Data Files

Step 1 – Install Adobe Acrobat Reader

To view and print the product manuals, you must have Acrobat Reader installed on your machine. Visit Adobe’s Web site (www.adobe.com) to download the appropriate reader.

Step 2 – Install the Product

1. Close all Windows programs.
2. Place the *Software Installation* CD in the CD-ROM reader.
The installation program starts automatically if you have AutoPlay enabled. If the installation program does not start automatically, double-click QMSstart.exe in the root directory of the CD-ROM to start the main installation application.
3. Select **Centrus Desktop** in the **Centrus Products - Setup** dialog box.
4. Press the **Install** button to begin the installation process.
5. Click **Next** in the **Welcome** dialog box if all Windows programs are closed. Click **Cancel** if you need to quit Setup and close programs.

6. Read the license agreement and click **I Agree**.
7. Read the Readme file and click **Next**.
8. Click **Next** on the **Choose Destination Location** dialog box to accept the default destination folder. Click **Browse** to choose a different folder; then click **Next**.
9. Click **Next** to install the files.
10. For Windows 98 or NT systems, a prompt to install Microsoft Data Access Components appears. Select **Yes** to update, or **Cancel** to not update. We recommend that you update the Microsoft Data Access Components.
11. Click **Finish** on the **Installation Complete** dialog box.

Step 3 – Install the Data Files

Note: If you plan to use WinSplit, you do not need to complete the steps provided here. The *Sagent Utilities Manual* provides all the steps necessary to split and install the data files.

The data files required for processing must reside on your machine, or the CD-ROMs must be directly accessible to your machine. For descriptions of the data files you may need to complete the desired processing, click **Data Products Information** on the **Centrus Products - Setup** screen. The *Release Notes for the Centrus Data Products* included in the installation package also describe the data files.

Note: See the *Release Notes for the Centrus Data Products* for instructions on downloading and installing optional data files for Early Warning System (EWS). See Appendix B, "Early Warning System Data," for additional information about EWS.

To copy data to your local drive:

1. Delete previous copies of .gsd, .gsu, .gsl, .gsi, .gsz, .z9, .dir, .las, and .los files if you have installed Centrus Desktop before.
2. Select **Data Products** on the **Centrus Products - Setup** dialog box.
3. When prompted, place *Data Products Suite, Disc A*, in your CD-ROM drive and click **OK**.
4. Click **Next** on the **Welcome** dialog box.
5. Read the license agreement and click **I Agree**.
6. Read the data products Readme file and click **Next**.
7. Click **Next** to accept the default installation directory or browse to a different directory.
8. Select the check boxes for the products you want to install.

Note: Click **Data Info** or **HTML Docs** for information about which data sets to install.
9. Click **Next**.

10. Click **Next** to start installing the data.
Insert the appropriate data CDs when prompted.
11. If you selected the GDL Data Set in Step 8, the **Choose Destination Location** dialog box displays. Click **Next** to accept the default installation directory or browse to a different directory. The **Installation Complete** dialog box displays.
12. Click **Finish**.
Note: If you copy any files to directories other than the Centrus Desktop directories, you need to specify the location of those files in the **Configuration** dialog box. Refer to “Specifying the Search Path for the Data Files” on page 7 for information about specifying the location of the data files.
13. If you chose the GDL Data Set in Step 8, the encrypted GDL file (**Zip5.gs_**) is decrypted and installed on your machine. You will have an opportunity to change the data installation directory if you rather not accept the default data install directory.
To install the **Zip5.gs_** file:
 - a. Click **Next** if you accept the default destination directory in the **Choose Destination Location** dialog.
 - b. Enter the password when prompted.
The password is provided in your product letter and is case sensitive. The Zip5.gs_ file is renamed to Zip5.gsb and ready for use with Centrus Desktop GDL.

Chapter 2

Configuration

To configure Centrus Desktop for processing, you need to specify the search path to the data files, set your processing options, verify your settings, and set up a task. The following sections explain how to complete these tasks:

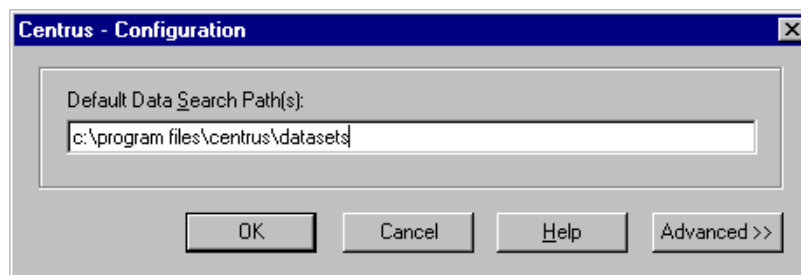
- Specifying the Search Path for the Data Files
- Setting Processing Options
- Verifying Your Settings
- Working with Tasks

If you prefer to run Desktop from the command line, see “Using Command Line Syntax” on page 37.

Specifying the Search Path for the Data Files

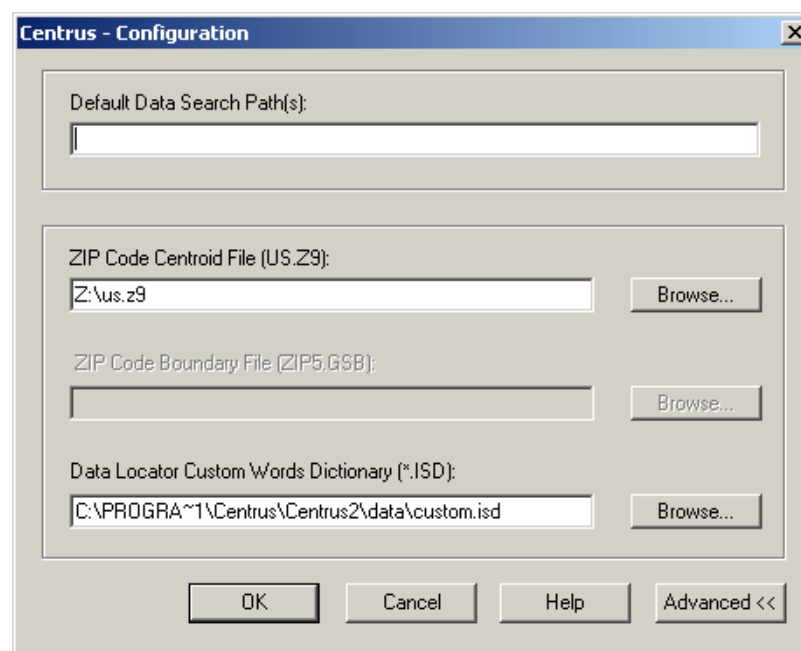
Centrus Desktop requires certain data files for different types of processing. The path to these files must be specified in the **Configuration** dialog box (shown below). During installation, the path to these files is automatically configured and is displayed in the **Default Data Search Path(s)** text box. If you must change the search path, manually type it into the text box, as there is no **Browse** button for specifying this search path. The following sections list which files are required for each function.

- Data Files for the Address Coding Module
- Data Files for the Data Locator Module
- Data Files for the Demographics Module
- Data File for the Geographic Determination Module



If you will be accessing all Centrus Desktop data files from the Centrus Desktop program directory or CD-ROM, you may use the default search path(s) as configured during installation.

Note: If you need to change the search path to the Us.z9, Zip5.gsb, or the Custom.isd files, select the **Advanced** button to display their search directories in the **Configuration** dialog box the first time you start Centrus Desktop.



Data Files for the Address Coding Module

The following files are required for all processing with Centrus Desktop:

- Ctyst.dir
- Msac.dir
- Parse.dir
- Us.z9
- Use.gsd and/or Usw.gsd

During installation, the path to these files is automatically configured and is displayed in the **Default Data Search Path(s)** text box. If you must change the search path, manually type it into the text box, as there is no **Browse** button for specifying this search path.

Certain functions of the Address Coding module use additional files. The following sections list which files are required by these functions.

- Data Files for CASS Compliance
- Data Files for Processing with Extended Alias Information
- Data Files for Canadian Address Processing
- Data Files for Intersection Matching
- Data Files for NAD Conversion

Data Files for CASS Compliance

The following files are required for CASS compliance and to generate a CASS report. They are not, however, required for processing in CASS mode.

- Us.gsz
- Use.gsd *and* Usw.gsd
- Zip9e.gsu *and* Zip9w.gsu

Data Files for Processing with Extended Alias Information

The following optional files, which are licensed separately, provide extended alias information:

- Us_pe.gsi (use with Use.gsd)
- Us_pw.gsi (use with Usw.gsd)
- Us_pse.gsi (use with Use.gsd and Uste.gsd)
- Us_psw.gsi (use with Usw.gsd and Ustw.gsd)

Data Files for Canadian Address Processing

The following files are required for processing Canadian addresses:

- Datafile.dat
- Datatab.dat
- Gscgcodes.cp9

These files are installed in the executable directory and must reside in that location for Centrus Desktop.

Data Files for Intersection Matching

The following files are used for intersection matching with the Address Coding module. (They also can be used for manual geocoding and mapping.)

- Uste.gsd
- Ustw.gsd

The following optional files, which are licensed separately, allow intersection matching without a complete last line:

- Use.gsi
- Usw.gsi

Data Files for NAD Conversion

The following files are used for NAD conversion with the Address Coding module:

- Alaska.las
- Alaska.los
- Conus.las
- Conus.los
- Prvi.las
- Prvi.los

For information about NAD conversion, see “Return Coordinates in NAD 27 or NAD 83” on page 17.

Data Files for the Data Locator Module

The Custom.isd file is the *Data Locator Custom Words Dictionary*, a user-defined custom dictionary. Specify the path to this file, including the file name, in the **Data Locator Custom Words Dictionary** text box. You can browse to the file or type the path manually.

For more information, see “Custom Dictionaries in Data Locator” on page 172.

The Words.isd file is installed in the */centrus2/data* directory and must reside at that location for Centrus Desktop.

Data Files for the Demographics Module

The following files are used for demographics processing with the Demographics module:

- Census2k.dld
- Claritas.dld
- Prizmbg.dld
- Prizmz4.dld

The Claritas.dld, Prizmz4.dld, and Prizmbg.dld are licensed separately.

Data File for the Geographic Determination Module


The Zip5.gsb file is required for processing with the Geographic Determination module.

This file is only available if you are licensed for GDL and if you specified that it be installed during the data installation process.

If the Zip5.gsb file is not located in the default search path, you may type in the directory to this file by using the **Advanced** option to display its search directory in the **Configuration** dialog box.

Setting Processing Options

You can define the behavior of each Centrus Desktop module in the **Options** dialog box.

To set options, click the **Options** button  or select **Process > Options**. The **Options** dialog box appears with a tab for each module. You access a module's options by clicking its tab. When Centrus Desktop is started, these options are set to the most recently used settings.

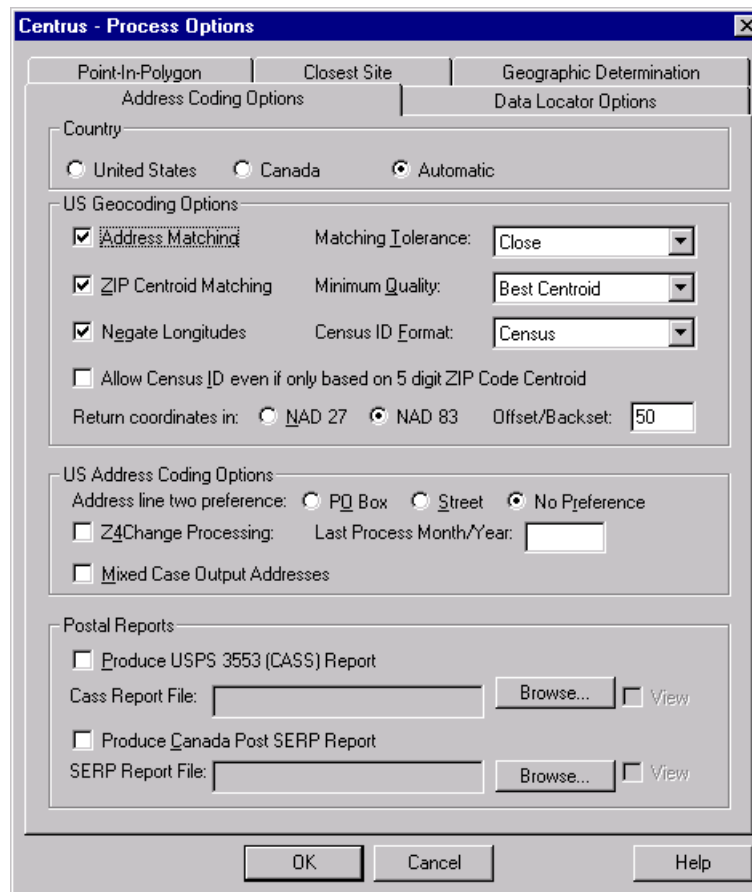
The **Options** dialog box includes the following tabs:

- Address Coding Options Tab
- Data Locator Options Tab
- Point-in-Polygon Tab
- Closest Site Tab
- Geographic Determination Tab

Address Coding Options Tab

The **Address Coding Options** tab contains settings that determine how Centrus Desktop matches addresses, and the format of the returned information. These settings can often remain the same for processing different address files. The **Address Coding Options** tab controls the following options:

- Country Option
- U.S. Geocoding Options
- U.S. Address Coding Options
- Postal Reports Option



The image shows the 'Centrus - Process Options' dialog box. It has a title bar with a close button. The dialog is divided into several sections: 'Point-In-Polygon', 'Closest Site', and 'Geographic Determination'. Below these are 'Address Coding Options' and 'Data Locator Options'. The 'Country' section has three radio buttons: 'United States', 'Canada', and 'Automatic'. The 'US Geocoding Options' section includes checkboxes for 'Address Matching', 'ZIP Centroid Matching', and 'Negate Longitudes', along with dropdown menus for 'Matching Tolerance', 'Minimum Quality', and 'Census ID Format'. There is also a checkbox for 'Allow Census ID even if only based on 5 digit ZIP Code Centroid' and a section for 'Return coordinates in' with radio buttons for 'NAD 27' and 'NAD 83', and a text field for 'Offset/Backset'. The 'US Address Coding Options' section includes a section for 'Address line two preference' with radio buttons for 'PO Box', 'Street', and 'No Preference', and checkboxes for 'Z4Change Processing' and 'Mixed Case Output Addresses'. The 'Postal Reports' section includes checkboxes for 'Produce USPS 3553 (CASS) Report' and 'Produce Canada Post SERP Report', each with a text field for the report file, a 'Browse...' button, and a 'View' checkbox. At the bottom are 'OK', 'Cancel', and 'Help' buttons.

Country Option

The **Country** option allows you to select United States, Canada, or Automatic processing.



The image shows a close-up of the 'Country' section of the dialog. It has a label 'Country' and three radio buttons: 'United States', 'Canada', and 'Automatic'. The 'United States' radio button is selected.

When you select **United States**, Centrus Desktop performs standard or multiline processing using selected U.S. Geocoding and U.S. Address Coding options.

When you select **Canada**, Centrus Desktop performs standard processing only.

When you select **Automatic**, Centrus Desktop determines if the address is U.S. or Canadian and processes it accordingly.

U.S. Geocoding Options

The geocoding options determine the quality of matches and levels of geocodes to return, as well as how the information is to be presented. While these settings generally can remain the same from task to task, different uses may require different settings. Available options are:

- Address Matching
- ZIP Centroid Matching
- Negate Longitudes
- Census ID Format
- Allow Census ID Even if Only Based on 5-digit ZIP Code Centroid
- Return Coordinates in NAD 27 or NAD 83
- Offset/Backset

Address Matching

When the **Address Matching** check box is selected, Centrus Desktop performs address standardization and address geocoding. The setting indicates the type of match that is required.

US Geocoding Options

☒ Address Matching Matching Tolerance: Close

☐ ZIP Centroid Matching Minimum Quality: Best Centroid

☐ Negate Longitudes Census ID Format: Census

☐ Allow Census ID even if only based on 5 digit ZIP Code Centroid

Return coordinates in: ☒ NAD 27 ☐ NAD 83 Offset/Backset: 50

The possible settings for **Matching Tolerance** are:

- **Tight** – This setting requires that addresses are virtually an exact match. Sagent recommends using Tight only when the addresses have already been standardized to USPS standards. This is the fastest option.
- **Close** – This is the setting we recommend for most uses. It allows minor misspellings, as well as incorrect or missing directionals or street types. This is the default setting.

- **Extended** – This setting is similar to Close, except that the street name search looks at all streets that have the same first letter, rather than using pattern matching. This setting also performs the widest search possible, using the USPS Finance Area. We recommend that you review all matches made by this setting in which the street name was modified. You can identify these matches via the Match Code. (See Appendix A, “System Messages and Codes” for more information about Match Codes.) This is the slowest setting.
- **CASS** – This setting is used to ensure CASS compliance and is required when producing a CASS report. With this setting, some addresses that could be successfully coded using the Close or Extended settings may not be successfully matched.

Note: This function is not supported for Canadian addresses.

ZIP Centroid Matching

When the **ZIP Centroid Matching** check box is selected, Centrus Desktop provides a ZIP+4, ZIP+2 or ZIP Code centroid geocode when an address geocode is not available. The **Minimum Quality** setting indicates which type of centroids are used. The best possible centroid is always used if multiple centroids are available.

The possible settings are:

- **9 Digit Best Location** – This setting uses the most positionally accurate 9-digit centroids. These centroids are almost as accurate as an address geocode, in that they are accurate to a single block face in most cases. Census accuracy varies, but most centroids in this class are accurate to the Block Group level.
- **9 Digit Good Location** – This setting uses 9-digit centroids with very good positional accuracy. These centroids are accurate to a single block face in most cases. Census accuracy varies, but most centroids in this class are accurate to the Block Group or Census Tract level.

- **9 or 7 Digit Location** – This setting uses all 9- or 7-digit centroids. A 7-digit centroid is positionally accurate to within several blocks in most areas. Census accuracy varies. This setting yields all but 5-digit centroids.
- **Block Group Accuracy** – This setting uses centroids that are accurate to the Block Group level. Positional accuracy varies, but is normally accurate to the city block on which the address is located.
- **Census Tract Accuracy** – This setting uses centroids that are accurate to the Census Tract level or better. Positional accuracy varies, but is normally accurate to within several city blocks of where the address is located.
- **5 Digit Only** – This setting uses only 5-digit centroids. This is the least accurate geocode, but it is the fastest to assign.
- **Best Centroid** – This setting uses the best centroid, but all centroids are available. This is the default setting.

See Appendix A, “System Messages and Codes” for a complete listing of the location codes and settings assigned to each centroid type.

Note: This function is not supported for Canadian addresses.

Negate Longitudes

When the **Negate Longitudes** check box is selected (the default setting), longitudes are returned as negative numbers for western hemisphere locations. When the check box is cleared, longitudes are returned as positive numbers. Latitudes are always returned as positive numbers.

US Geocoding Options

☒ Address Matching Matching Tolerance: Close

☐ ZIP Centroid Matching Minimum Quality: Best Centroid

☒ Negate Longitudes Census ID Format: Census

☐ Allow Census ID even if only based on 5 digit ZIP Code Centroid

Return coordinates in: ☒ NAD 27 ☐ NAD 83 Offset/Backset: 50

This option is available to support those applications that require positive longitudes. Most mapping and GIS applications today use negative longitudes for western hemisphere locations.

Note: This function is not supported for Canadian addresses.

Census ID Format

The **Census ID Format** option determines the format of the Census FIPS Code. The default setting is **Census**. The possible formats are listed below with an example of how a Block Group would be displayed:

Delimited	08 013 0002.01-1
Census	080130002.011
Digits Only	080130002011

Note: This function is not supported for Canadian addresses.

Allow Census ID Even if Only Based on 5-digit ZIP Code Centroid

When selected, this check box permits the assignment of the Census FIPS Code even when only 5-digit ZIP Code centroids are available. The default setting for this check box is cleared (not selected).

US Geocoding Options

☒ Address Matching Matching Tolerance: Close

☐ ZIP Centroid Matching Minimum Quality: Best Centroid

☐ Negate Longitudes Census ID Format: Census

☒ Allow Census ID even if only based on 5 digit ZIP Code Centroid

Return coordinates in: ☒ NAD 27 ☐ NAD 83 Offset/Backset: 50

Note: This function is not supported for Canadian addresses.

Return Coordinates in NAD 27 or NAD 83

This option lets you select which datum Centrus Desktop uses to return the latitude and longitude of an address. A datum is the mathematical model of the Earth used to calculate the coordinates on any map, chart, or survey system. Surveyors take an ellipsoid model of the Earth and fix it to a base point. The North American Datum (NAD) is the official reference ellipsoid used for the primary geodetic network in North America. Two NADs have been developed:

- NAD27 (U.S. addresses only) has its origin at Meades Ranch, Kansas. NAD27 does not include Alaskan islands or Hawaii. Latitudes and longitudes that are surveyed in the NAD27 system are valid only in reference to NAD27 and do not tie to any maps outside the U.S.
- NAD83 (default for U.S. and Canadian addresses) is earth-centered and defined with satellite and terrestrial data. NAD83 is compatible with the World Geodetic System 1984 (WGS84), the terrestrial reference frame associated with the NAVSTAR Global Positioning System (GPS) now used extensively for navigation and surveying.

The coordinates for a point on the Earth's surface in one datum do not match the coordinates from another datum for that same point. A grid shift exists between datums because each datum has a different origin. For example, according to NAD27, the center of the intersection of Baseline Road and County Line Road near Boulder, Colorado is exactly 40 N. Latitude, 105 W. Longitude. But according to NAD83, that point is located at 39deg 59min 59.97sec N, 105deg 00min 01.93sec W. This is four feet south and fifty feet west from 40 N, 105 W.

This option affects only the numeric coordinates returned for the latitude and longitude of an address. QuickFind's mapping function and the spatial analysis functions chart found address locations based on the selected NAD27/NAD83 option, and chart other layers (such as primary and secondary streets) based on the database used.



Offset/Backset

If an offset distance is used, that distance is calculated perpendicular to the portion of the street segment in which it lands. This distance calculation yields the best visual representation for any mapping package, as well as be the most accurate location able to be imputed from the geographic data. Block data is returned for the side of the segment on which a match was found.

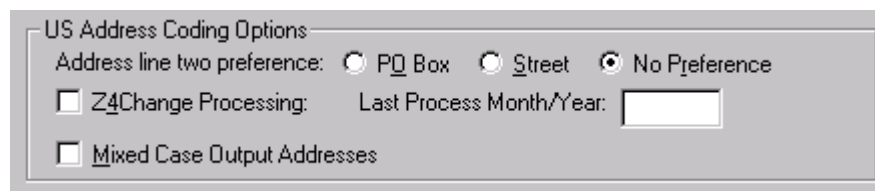
This option lets you override the default offset distance of 50 feet. The backset value is the same as the offset value or half the length of the segment, whichever is smaller.

The offset/backset value must be an integer value between 0 and 999.

U.S. Address Coding Options

In the **U.S. Address Coding Options** area, you may control how an address is returned, or turn **Z4 Change Processing** on or off. The following options are available:

- Address Line Two Preference
- Z4 Change Processing
- Mixed Case Output Addresses



Note: This function is not supported for Canadian addresses.

Address Line Two Preference

Two-line addressing allows you to pass two different address lines to Centrus Desktop. Centrus Desktop scans these two lines and extract and standardize a two-line address, if possible. You can set Centrus Desktop to accept a P.O. Box or a street address as the preferred address. If no option is specified or the two lines contain similar information, line one is preferred.

If both street address and P.O. Box are present on the two address lines, in any order, the checked preference determines which type of address (street address or P.O. Box) is returned for purposes of geocoding and standardization. For example, consider an address containing this data for the first two address lines:

123 Main Street
P.O. Box 24

If the “Street” option is selected, then “123 Main” is returned as the address and “P.O. Box 24” is returned as the second address line. If the “PO Box” option is selected, then “P.O. Box 24” is returned as the address and “123 Main” is returned as the second address line. If “No Preference” is selected, line one is preferred.

If a unit designator is present on one address line and a valid address appears on the other line, the unit information is appended to the address.

Note: This function is not supported for Canadian addresses.

Z4 Change Processing

See “Z4 Change Processing Option” on page 103 for an explanation of this feature.

When you select the **Z4Change Processing** option, type the date that the list records were last processed in the “Last Process Month/Year” text box. Once **Z4 Change Processing** is activated, only records that have changed since the last time the list was processed are re-processed.

Note: This function is not supported for Canadian addresses.

Mixed Case Output Addresses

Normally, addresses are returned in uppercase letters, according to USPS and Canada Post Corporation (CPC) recommendations. If the **Mixed Case Output Addresses** check box is selected, addresses are returned with the first letter of each word capitalized.

Postal Reports Option

The Postal Reports option allows you to generate a CASS report for the USPS or a SERP report for the CPC. The following options are available:

- Produce USPS 3553 (CASS) Report
- Produce Canada Post SERP Report

Produce USPS 3553 (CASS) Report

Commonly referred to as the CASS report, this report is filed with bulk mailings to receive postal discounts.

When the **Produce USPS 3553 (CASS) Report** check box is selected, Centrus Desktop produces a text file containing the report information. Specify the name and location of this file by clicking the **Browse** button and entering the path and name in the file selection dialog box.

The address file is processed in strict accordance with CASS guidelines. In order to be CASS-compliant, the following settings must be present:

- Address fields must be updated during processing or alternate address output fields must be specified.
- Output fields for **Carrier Route**, **DPBC**, and **Check Digit** must be present.
- **Address Matching** is checked and the **Matching Tolerance** setting is set to CASS. (This is done automatically when you check the **Produce USPS 3553 (CASS) Report** option.)
- The **Unmatched Records Only** check box on the **Address Coding** tab is ignored. All records must be processed in order to derive accurate coding counts for the reports.

The **View** checkbox allows you to choose whether or not the CASS report is displayed when processing is complete. If the **View** check box is selected, the CASS report displays in Windows Notepad, if Notepad is available.

If the **View** check box is not selected, the CASS report does not display. This is a useful option if you are processing in batch mode.

See “CASS Overview” on page 195 for more information on the USPS CASS program.

Produce Canada Post SERP Report

When the **Produce Canada Post SERP Report** check box is selected, Centrus Desktop produces a text file containing the report information. Specify the name and location of this file by clicking the **Browse** button and entering the path and name in the file selection dialog box.

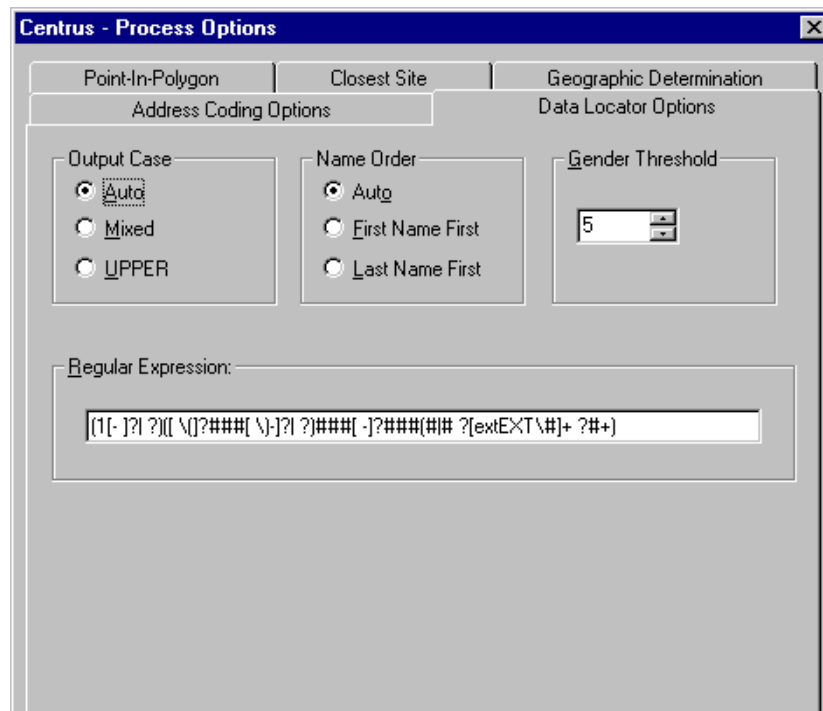
The **View** checkbox allows you to choose whether or not the SERP report is displayed when processing is complete. If the **View** check box is selected, the SERP report displays in Windows Notepad, if Notepad is available.

If the **View** checkbox is not selected, the SERP report does not display. This is a useful option if you are processing in batch mode.

See “Canadian Address Accuracy Program” on page 203 for more information on CPC’s Address Accuracy Program and SERP.

Data Locator Options Tab

The **Data Locator Options** tab contains settings that determine how the Data Locator module assigns gender information to names, as well as the format in which the names are returned.



The following options are available:

- Case
- Name Order Options
- Gender Threshold Option
- Regular Expression

Case

This option determines how names are capitalized in output.

- **Auto** – If mixed case is found in the original, the exact original capitalization is preserved in the output. Otherwise, Data Locator uses mixed case, capitalizing the first letter of each word. This is the default setting.
- **Mixed** – Converts all names to mixed case, capitalizing the first letter of each word.
- **UPPER** – Converts all names to upper case. For CASS certified mailings, names must be in upper case letters.

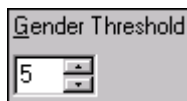
Name Order Options

This option determines the order in which the Name Coding module expects to see names input.

- **Auto** – The Data Locator module determines word order based on context. This is the default setting.
- **First Name First** – Assumes that names are input first name first.
- **Last Name First** – Assumes that names are input last name first.

Gender Threshold Option

This setting represents the error threshold for the assignment of gender to a name.



It is measured as the accuracy tolerance away from 100 percent female or 100 percent male. For example, a threshold of 5 (the default setting) allows a name that is male 95 percent of the time to be assigned a “Mr.”

Regular Expression

A regular expression is a pattern that describes a set of ASCII strings. It can be used in the Data Locator module to identify information such as e-mail addresses, Social Security numbers, and telephone numbers (the default setting).

Regular Expression:

`([!- ?] ?)([\[\]#### \-]? ?)####[-]?####(## ?(extEXT\#)+ ?#+)`

For more information about regular expressions, see “Customizing the Data Locator Module” on page 169.

Point-in-Polygon Tab

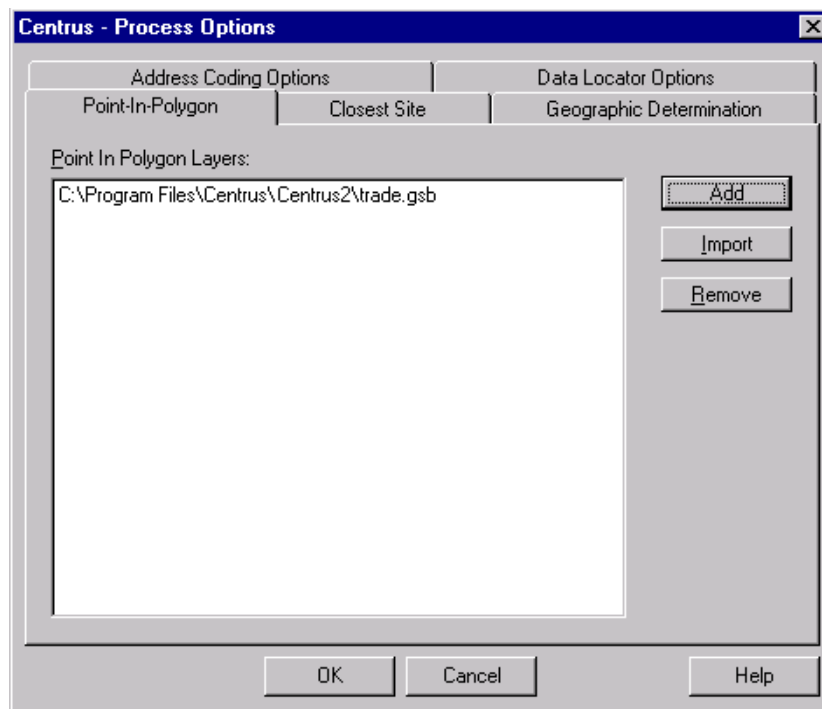
Use the **Point-in-Polygon** tab to determine which data layers are available to the Point-in-Polygon module. You can add data layers that are already in .gsb format, such as the States.gsb and Counties.gsb files. You can also import layers from MapInfo's .mif/.mid files, ESRI's Atlas GIS .bna files, or ESRI's .shp files. Note that data layers must be built with buffers enabled for use with the Point-in-Polygon function.

Note: The **Point-in-Polygon** tab displays only a **Configure...** button until a .gsb file is specified in the Point-in-Polygon page of the **Process Options** dialog box.

These operations are associated with Point-in-Polygon layers:

- Adding Point-in-Polygon Layers
- Importing Point-in-Polygon Layers
- Removing Point-in-Polygon Layers

For more information about buffering, see “Point-in-Polygon Analysis” on page 132.



Adding Point-in-Polygon Layers

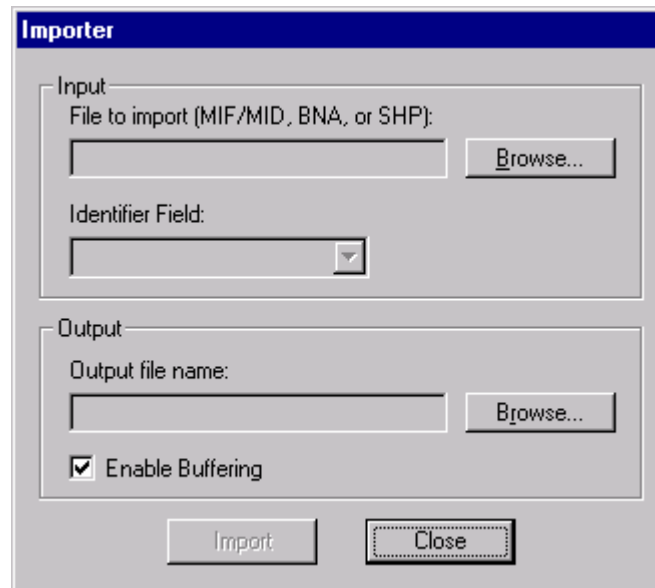
To add layers:

1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Point-in-Polygon** tab.
2. Click the **Add** button. A file selection dialog box appears.
3. Select the desired file name, then click **OK**.

Importing Point-in-Polygon Layers

To import layers:

1. From the main Centrus Desktop menu, select **Process > Options**, then click the Point-in-Polygon tab.
2. Click the **Import** button. The **Import Layers** dialog box appears.



3. Click the **Browse** buttons to select the path and name of the .mif/ .mid, .bna, or .shp file to import and the object file (.gsb) to create.
4. When importing, you can determine which field is used as the object's identifier. Click the drop-down list box to select which **Field to use as identifier**.
5. Be sure that the **Enable Buffering** box is checked.

If this check box is not selected, you do not receive in border or in buffer polygon location and distance-to-border values.

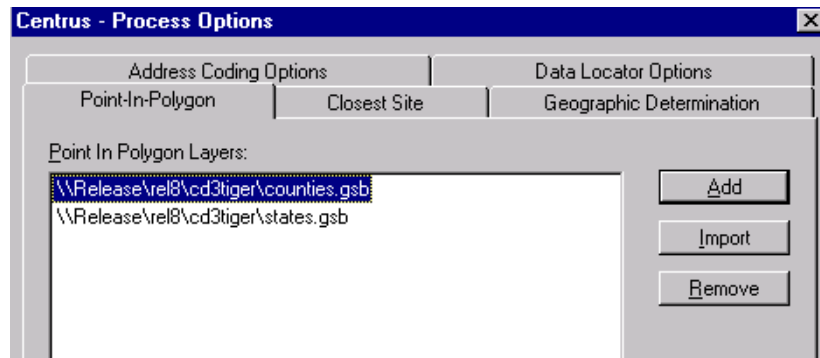
6. Click **Import** to finish.

Note: While western hemisphere locations are correctly expressed as negative longitudes, some mapping and GIS applications still return longitudes as positive numbers. Be sure that the longitude data in your imported layer is consistent with that of your address file.

Removing Point-in-Polygon Layers

To remove layers:

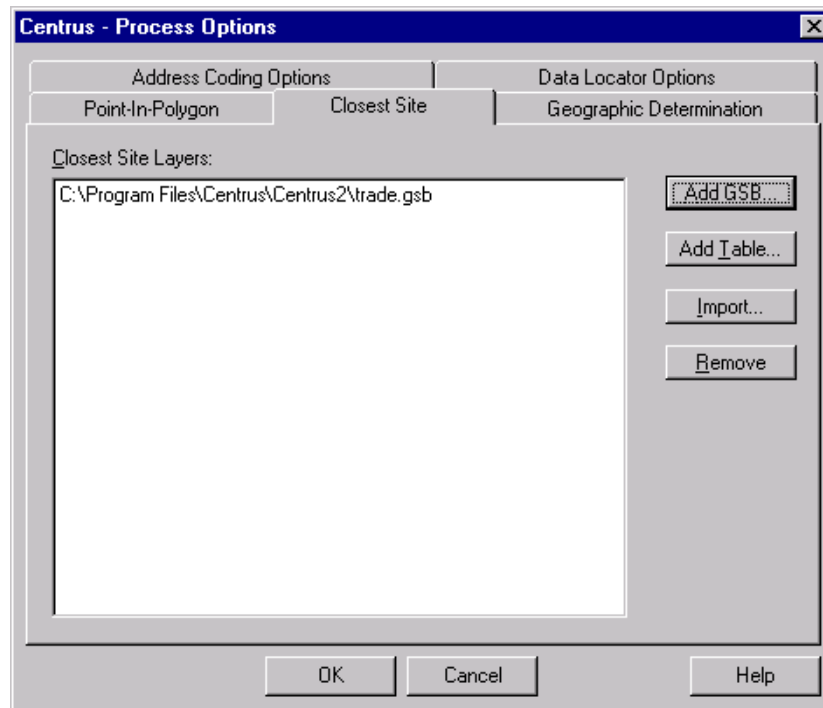
1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Point-in-Polygon** tab.
2. Select the file you want to remove from the **Point-in-Polygon Layers** list, then click **Remove**.



Closest Site Tab

The **Closest Site** options tab is where you determine which data layers are available to the Closest Site module. You can add data layers that are already in .gsb format, such as the States.gsb and Counties.gsb files. You can also import geocoded tables from database files, or import layers from MapInfo's .mif/.mid files, ESRI's Atlas GIS .bna files, or ESRI's .shp files.

Note: The **Closest Site** tab displays only a **Configure...** button until a file is specified in the Closest Site page of the **Process Options** dialog box.



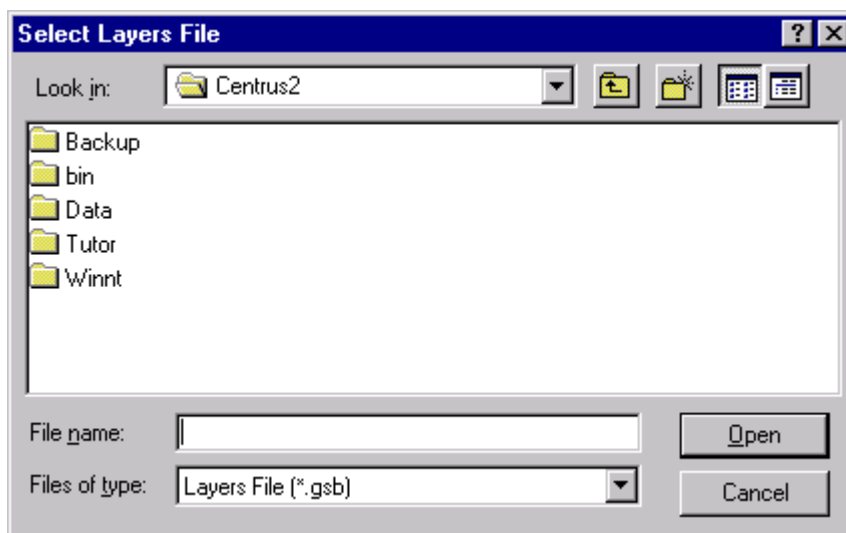
These operations are controlled from the **Closest Site** tab:

- Adding Closest Site Layers from a .gsb File
- Adding Closest Site Layers from a Database Table
- Importing Closest Site Layers
- Removing Closest Site Layers

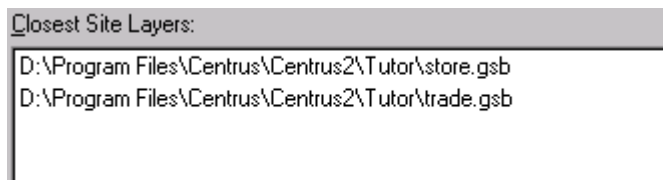
Adding Closest Site Layers from a .gsb File

To add layers from a .gsb file:

1. From the main Centrus Desktop screen, click **Process**, then **Options**, then the **Closest Site** tab.
2. Click the **Add GSB** button. A file selection dialog box appears.



3. Select the desired file name, then click **Open**. The added file appears on the list of Closest Site Layers.



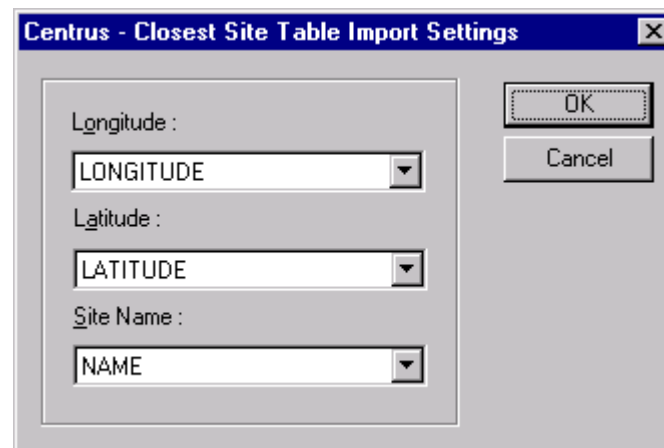
Adding Closest Site Layers from a Database Table

To add layers from a geocoded database table:

1. From the main Centrus Desktop screen, click **Process**, then **Options**, and then the **Closest Site** tab.
2. Click the **Add Table** button. A file selection dialog box appears.



3. Select the desired file name, then click **Open**. The **Closest Site Table Import Settings** dialog box appears.

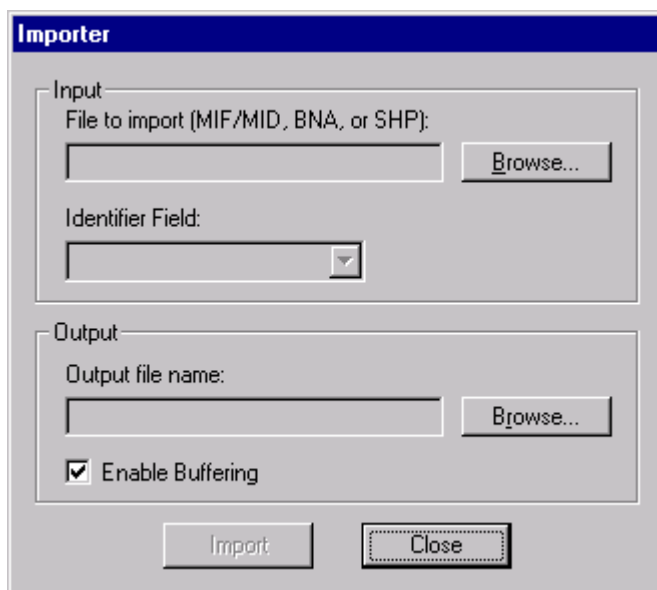


4. Check that **Latitude** and **Longitude** fields are correctly specified. Then select the field you want to use as **Site Name** from the drop-down list and click **OK**.

Importing Closest Site Layers

To import layers:

1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Closest Site** tab.
2. Click the **Import** button. The **Import Layers** dialog box appears.



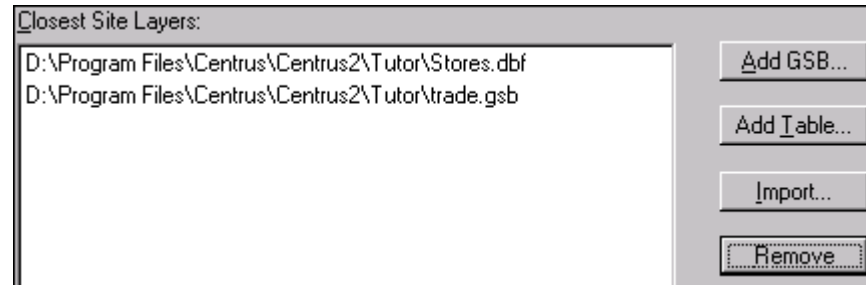
3. Click the **Browse** buttons to select the path and name of the .mif/.mid, .bna, or .shp file to import and the object file (.gsb) to create.
4. When importing, you can determine which field is used as the object's identifier. Click the drop-down list box to select which **Field to use as identifier**.
5. Click the **Import** button to finish.

Note: While western hemisphere locations are correctly expressed as negative longitudes, some mapping and GIS applications still return longitudes as positive numbers. Be sure that the longitude data in your imported layer is consistent with that of your address file.

Removing Closest Site Layers

To remove layers:

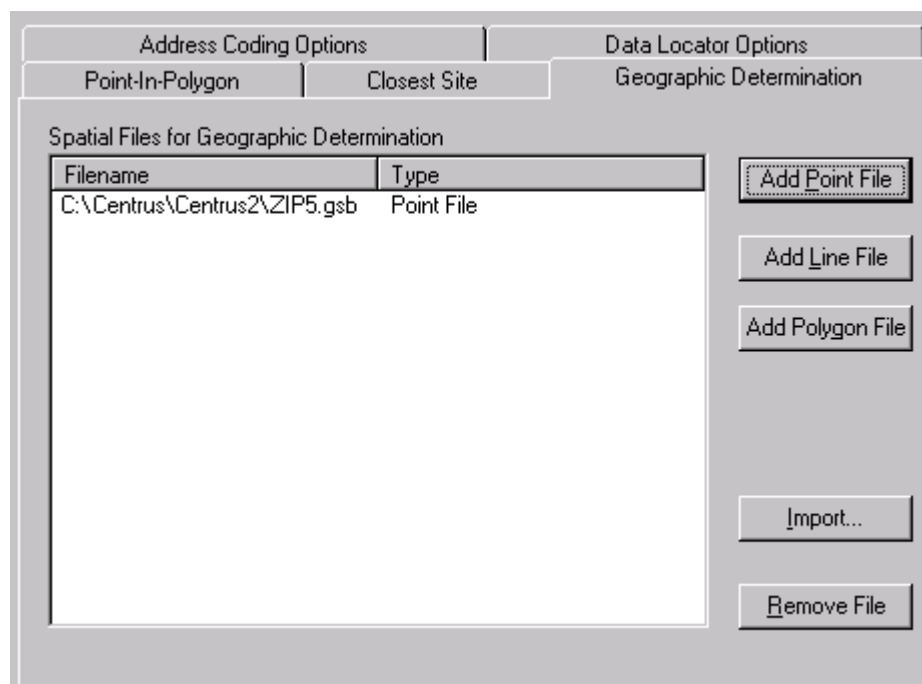
1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Closest Site** tab.
2. Select the file you want to remove from the **Closest Site Layers** list, then click **Remove**.



Geographic Determination Tab

The **Geographic Determination** options tab is where you determine which data layers are available to the Geographic Determination module. You can add data layers that are already in .gsb format, such as the States.gsb and Counties.gsb files. You can also import geocoded tables from database files, or import layers from MapInfo's .mif/.mid files, ESRI's Atlas GIS .bna files, or ESRI's .shp files.

Note: The **Geographic Determination** tab displays only a **Configure...** button until a file is specified in the Geographic Determination page of the **Process Options** dialog box.



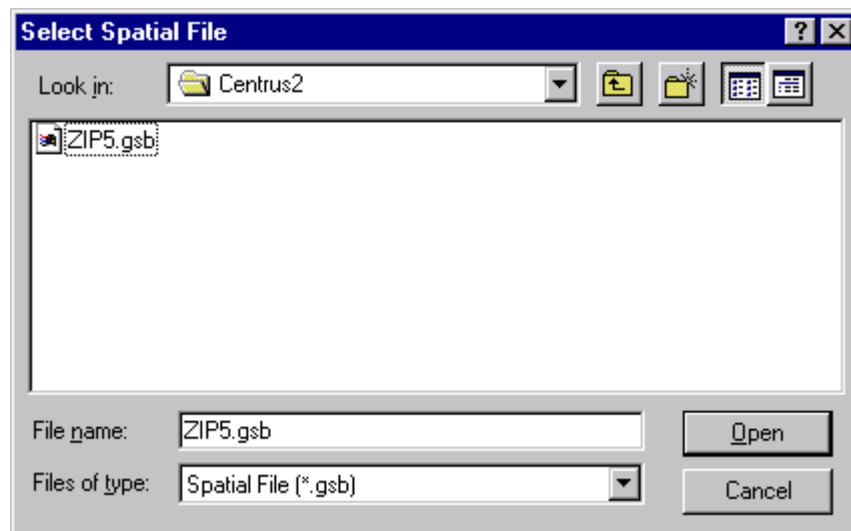
These operations are controlled from the **Geographic Determination** tab:

- Adding a Geographic Determination Point File
- Adding a Geographic Determination Line File
- Adding a Geographic Determination Polygon File
- Importing a Geographic Determination File
- Removing Files

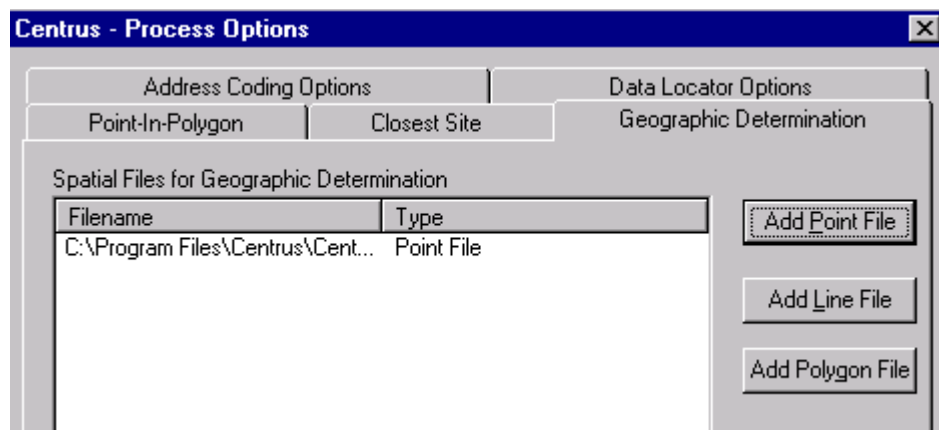
Adding a Geographic Determination Point File

To add a layer:

1. From the main Centrus Desktop screen, click **Process** on the menu bar, then select **Options**, then click the **Geographic Determination** tab.
2. Click the **Add Point File** button. A file selection dialog box appears.



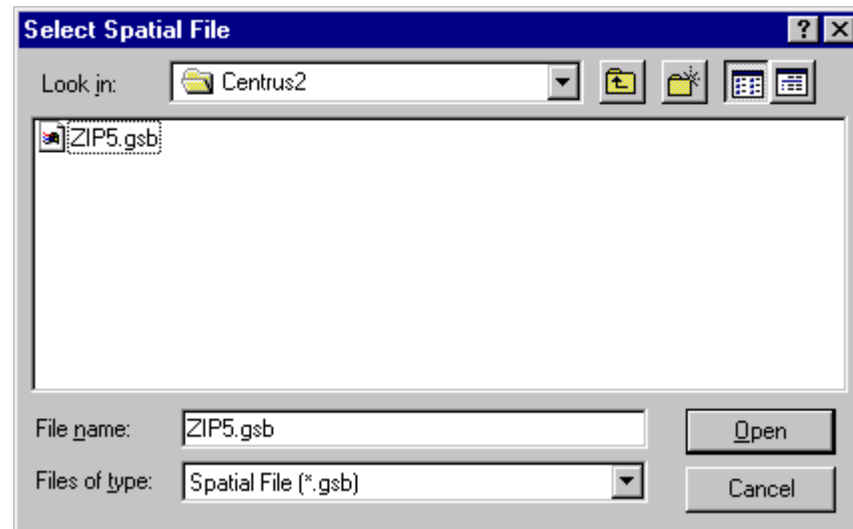
3. Select the desired file name, then click **Open**. The added file appears on the **Spatial Files for Geographic Determination** list.



Adding a Geographic Determination Line File

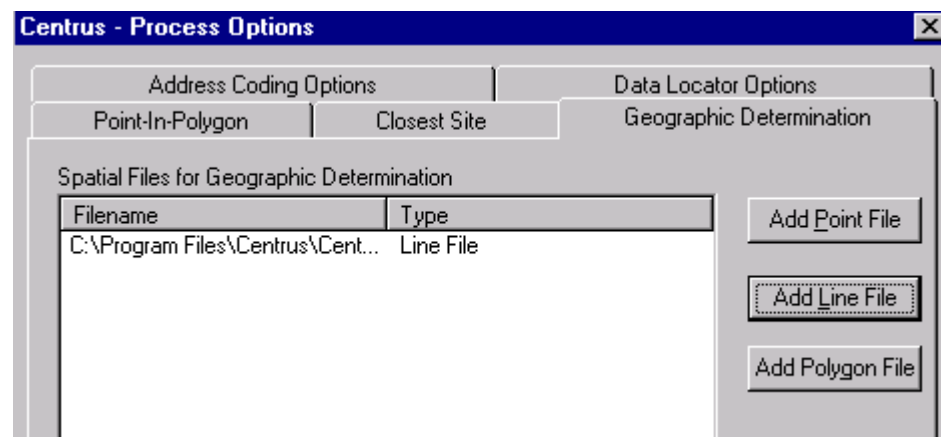
To add a layer:

1. From the main Centrus Desktop screen, click **Process** on the menu bar, then select **Options**, then click the **Geographic Determination** tab.
2. Click the **Add Line File** button. A file selection dialog box appears.



3. Select the desired file name, then click **Open**. The added file appears on the **Spatial Files for Geographic Determination** list.

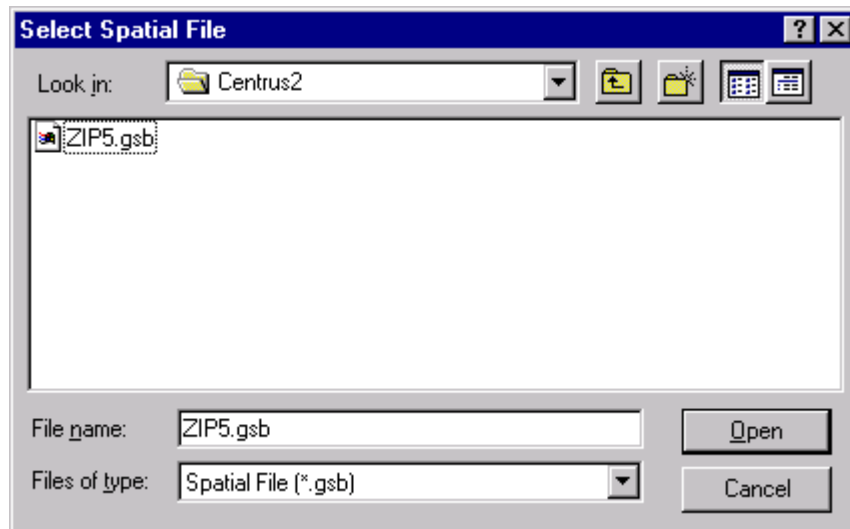
Selecting files from the list that do not have objects for their specified type, generates an error. Either change the object type or browse to a different .gsb file.



Adding a Geographic Determination Polygon File

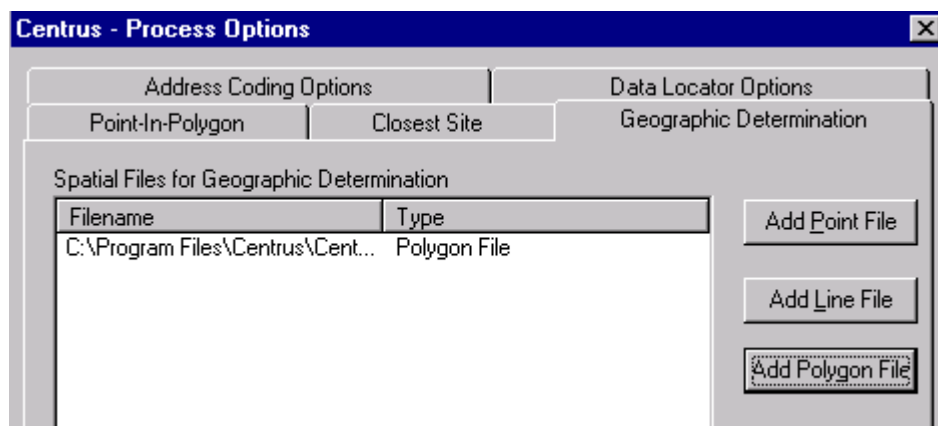
To add a layer:

1. From the main Centrus Desktop screen, click **Process** on the menu bar, then select **Options**, then click the **Geographic Determination** tab.
2. Click the **Add Polygon** button. A file selection dialog box appears.



3. Select the desired file name, then click **Open**. The added file appears on the **Spatial Files for Geographic Determination** list.

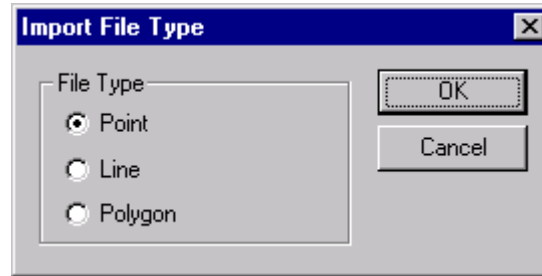
Selecting files from the list that do not have objects for their specified type, generates an error. Either change the object type or browse to a different .gsb file.



Importing a Geographic Determination File

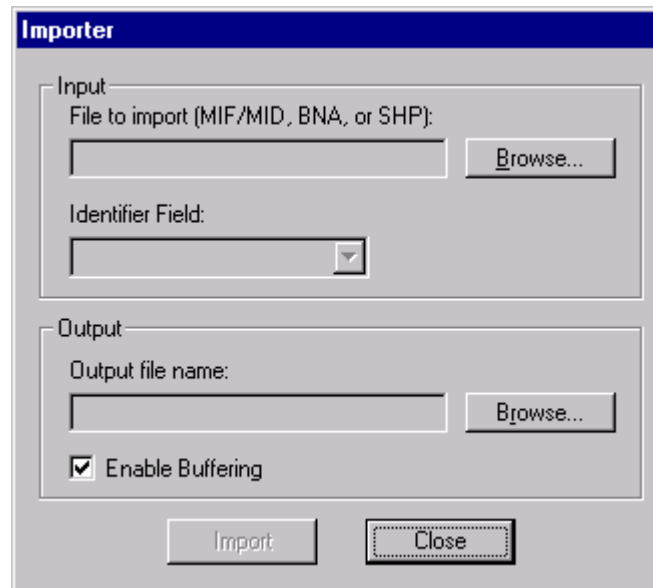
To import a file:

1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Geographic Determination** tab.
2. Click the **Import** button. The **Import File Type** dialog box appears.



Geographic Determination requires that all objects in the file be of the same type.

3. Click the corresponding radio button of the file you want to create. The **Import Layers** dialog box appears.



4. Click the **Browse** buttons to select the path and name of the .mif/.mid, .bna, or .shp file to import and the object file (.gsb) to create.
5. When importing, you can determine which field is used as the object's identifier. Click the drop-down list to select which **Field to use as identifier**.

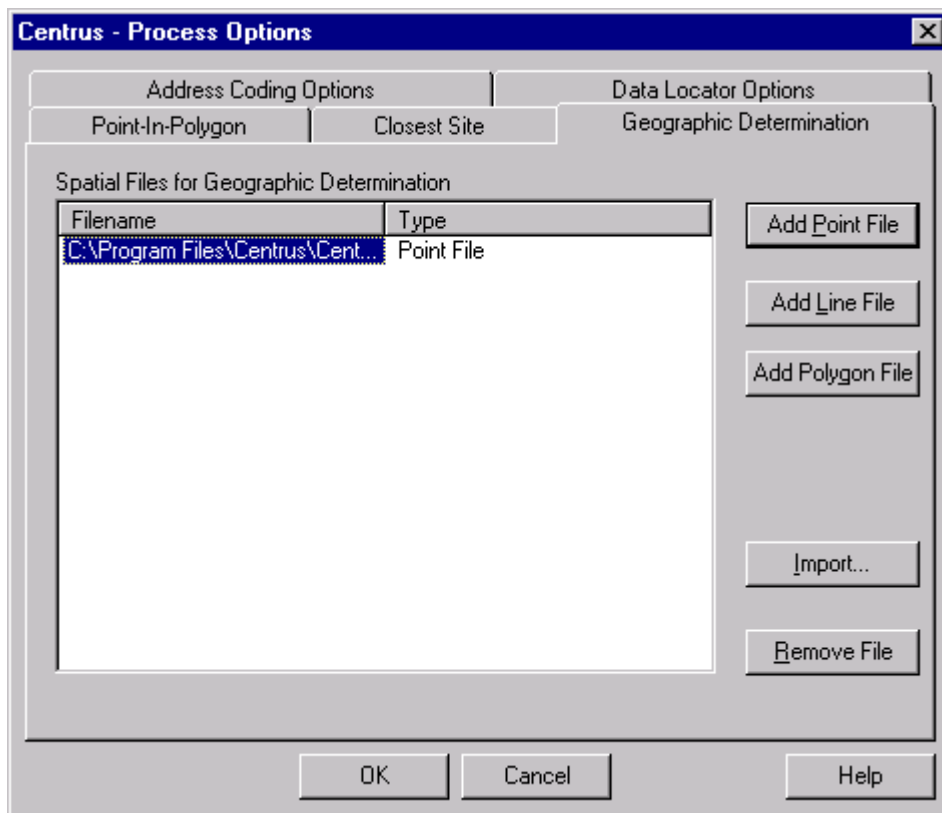
6. Check the **Enable Buffering** check box to return any distance values.
7. Click the **Import** button to finish.

Note: While western hemisphere locations are correctly expressed as negative longitudes, some mapping and GIS applications still return longitudes as positive numbers. Be sure that the longitude data in your imported layer is consistent with that of your address file.

Removing Files

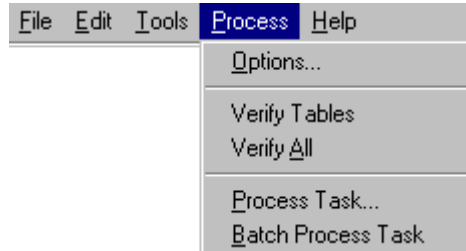
To remove files:


1. From the main Centrus Desktop menu, select **Process > Options**, then click the **Geographic Determination** tab.
2. Select the file you want to remove from the **Spatial Files for Geographic Determination** list, then click **Remove**.



Verifying Your Settings

When you select **Process** from the Centrus Desktop main menu, you see two choices for verification: **Verify All**, and (depending on which module is currently active) **Verify Tables** (or **Data Locator**, **Address Coding**, **Demographics**, **Point-in-Polygon**, **Closest Site**, **Geographic Determination**). These items let you verify that you've correctly specified your tasks. As their names imply, you can verify the settings of the currently visible module, or verify all module settings at once.



Click the **Verify Current** button  to verify the settings within the currently selected module.

Using Command Line Syntax

If you prefer to run Centrus Desktop from the command line, several command line options are available. The syntax is:

```
Centrus FileName.qmi /q /b /d /n /e
```

These work as follows:

<i>FileName.qmi</i>	Specifies a saved task to be immediately loaded.
/q	Starts Centrus Desktop in QuickFind, suppressing the splash screen.
/b	Starts Centrus Desktop in batch mode. This switch should be used with a <i>FileName.qmi</i> argument. No splash screen is displayed; the task file is opened and batch processing begins immediately. Note that batch processing is disabled when you run Centrus Desktop in demo mode.
/d	Starts Centrus Desktop in demo mode. All modules are available, but there is a 25 record processing limit and no CASS or batch processing is permitted.


/n	Starts Centrus Desktop with no splash screen.
/e	Causes the application to exit upon completion of processing. Typically used in conjunction with the /b option and qmi function. Note: if you open Centrus with the /e option only, and then define a new task to process, once the task is done batch processing Centrus Desktop closes.

Working with Tasks

Once you have configured Centrus Desktop for your system, you're ready to set up tasks. Note that there are seven tabs on the Centrus Desktop main window: one tab for each Centrus Desktop component, plus a **Tables** tab for specifying input and output files and tables. Each component's tab allows you to define how you want that module to process your data. Collectively, these settings constitute a *task*. Tasks are to Centrus Desktop what documents are to your word processing software. You can create them, save them, edit them, and rename them.


By default, your tasks are stored in the Centrus Desktop program directory. They are required to have the file name extension .qmi. The default settings for all Centrus Desktop modules are stored in the file Centrus Desktop.qmi, located in your Windows directory. You can set new defaults by overwriting this file with your own task file.

Creating a New Task

If you want to create a new task, click the **New Task** button  or select **File > New**. If you already have an unsaved task open, you are prompted to save it.

An "empty" new task is automatically created when you open Centrus Desktop.

Saving a Task

To save a task, click the **Save Task** button  or select **File > Save**. You'll be prompted to assign a name to the task.

Opening a Task

To open an existing task, click the **Open Task** button  or select **File > Open**. Select the desired task using the file selection dialog box.


To open a task you've used recently, display the **File** menu and select the desired file. The most recently used files display at the bottom of the **File** menu.

Editing a Task

It is often easier to edit a previously defined task than to define a new one. To do this, simply open an existing task and change the settings to suit your preferences. If you specify a new input table within a previously defined task, Centrus Desktop attempts to map output fields to the new input fields.

If you prefer to redefine some or all of your output fields, use the **Edit** menu. You can choose to:

- Clear output fields within the current module.
- Clear outputs fields in all modules.
- Clear outputs fields in all modules *except* the current module.

Note: Click the **Clear Current Outputs** button  to clear previously defined outputs within the current module.

Renaming a Task

To rename a task, select **File > Save As**. You are prompted to assign a new name to the task.

Chapter 3

Centrus Desktop Tutorial

This tutorial is designed to show you how to use each of the modules in Centrus Desktop. You will set up a *task* specifying how you want each module to process the data, and then you will process the task and view the results.

Before continuing with the tutorial, make sure Centrus Desktop is installed and configured as described in “Installation” on page 1 and “Configuration” on page 7. Then start Centrus Desktop using the instructions in “Starting Centrus Desktop” on page 42.

The first part of the tutorial, “Setting Up a Task,” shows you how to:

- Specify Input and Output Tables
- Specify Address Coding Inputs and Outputs
- Specify Demographic Inputs and Outputs
- Specify Point-in-Polygon Inputs and Outputs
- Specify Closest Site Inputs and Outputs
- Specify Data Locator Inputs and Outputs
- Specify Geographic Determination Inputs and Outputs

All users should complete the first two sections in “Setting Up a Task.” If you will be using only the Address Coding module, you can then skip to “Processing a Task.” If you will be using other modules as well, you may choose to complete the corresponding sections before moving on to “Processing a Task.”

In addition, you may choose to complete the section, “Creating a Primary Market Area (PMA).”

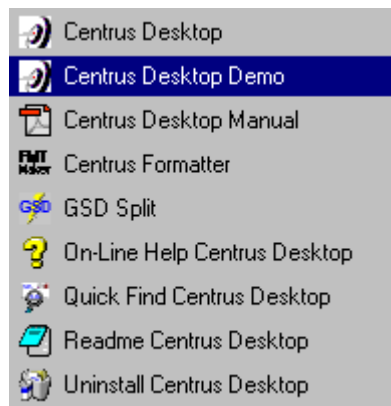
When Centrus Desktop was installed, sample data files were copied to a directory on your computer's hard drive, usually C:\Program Files\Centrus\Centrus2\Tutor\. These files are:

- trade.bna
- trade.gsb
- Stores.dbf
- Customer.dbf

You will use these files during the tutorial.

Starting Centrus Desktop

From the Windows **Start** menu, select **Programs > Centrus > Centrus Desktop Demo**. This option allows you to use all of the modules to process up to 25 records, even if your license does not include all modules.



Setting Up a Task

Tasks are to Centrus Desktop what documents are to your word processing software. You can create them, save them, edit them, and rename them. There are seven tabs on the Centrus Desktop main screen: one tab for each component, plus a **Tables** tab for specifying input and output files and tables. Each component's tab allows you to define how you want that module to process your data. Collectively, these settings constitute a *task*. To set up a task, complete the following steps:

- Specify Input and Output Tables
- Specify Address Coding Inputs and Outputs
- Specify Demographic Inputs and Outputs
- Specify Point-in-Polygon Inputs and Outputs
- Specify Closest Site Inputs and Outputs
- Specify Data Locator Inputs and Outputs
- Specify Geographic Determination Inputs and Outputs

For information about saving, editing, and renaming tasks, see “Working with Tasks” on page 38

Specify Input and Output Tables

The **Tables** tab is where you select the input file and (if desired) the output, reject, and report files. Your input file determines which fields are displayed as choices on the other tabs.

To specify input and output tables:

1. From the Centrus Desktop main screen, click the **Tables** tab.
2. In the **Input Tables** section, click **Browse** and select the file Customer.dbf in the Tutor directory.
3. In the **Output Tables** section, click **Browse**.
4. Type “Q_Cust.dbf” in the **File Name** field and click **Save** to create an output table.

The screenshot shows the 'Tables' tab interface in Centrus Desktop. It is divided into two main sections: 'Input Table' and 'Output Table(s)'. The 'Input Table' section includes an 'Input:' text field, a 'Browse...' button, and an 'In-Place Update' checkbox. The 'Output Table(s)' section includes an 'Output:' text field, a 'Browse...' button, and a 'Reject Records That Are Not:' checkbox. Below this checkbox are six checkboxes: 'Parsed', 'Geocoded', 'Point In Polygon Coded', 'Standardized', 'Demographic Coded', and 'Closest Site Coded'. At the bottom of the 'Output Table(s)' section is a 'Rejects:' text field and a 'Browse...' button.

Specify Address Coding Inputs and Outputs

The Address Coding module is a complete address standardization and geocoding solution that you can use to enhance any database containing address information. In this example, each record's address line and lastline information is standardized, the corrected addresses geocoded, and USPS delivery fields added, Census Block codes, match codes, and location codes.

To specify address coding inputs and outputs:

1. From the Centrus Desktop main screen, click the **Address Coding** tab.

The Address Coding module defines the input fields for you.

2. Check **Optimization Sort**.

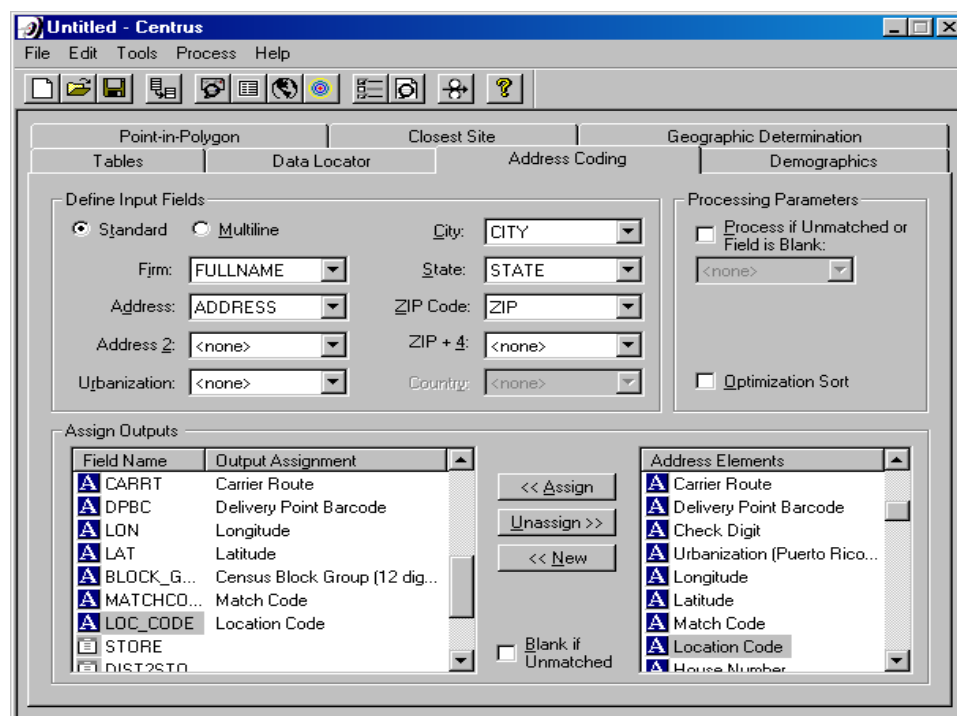
For most files, Optimization Sort significantly speeds processing by sorting the file in ZIP Code sequence prior to processing.

Note: Because the addresses in the example file are limited to the Boulder, Colorado area, Optimization Sort has little effect on this particular example.

3. In the **Assign Outputs** section of the **Address Coding** tab, make the following assignments:

Field Name		Address Element
Q_ADDRESS	←	Address Line
Q_LASTLINE	←	Last Line
CARRT	←	Carrier Route
DPBC	←	Delivery Point Barcode
LON	←	Longitude
LAT	←	Latitude
BLOCK_GRP	←	Census Block Group (12 digits)
MATCHCODE	←	Match Code
LOC_CODE	←	Location Code

You need to scroll down to find some of the items.



To make an assignment:

1. Highlight a field name in the **Field Name** column on the left.
2. Highlight an item in the **Address Elements** list on the right.
3. Click the **<<Assign** button to assign the address element to the proper field name. The assignment appears to the right of the field name, in the **Output Assignment** column.

At this point, you can skip to “Processing a Task” if you will not be using the other modules. If you *will* be using other modules, you may choose to complete their sections before moving on to “Processing a Task.”

Specify Demographic Inputs and Outputs

The Demographics module allows you to append valuable demographic information to your own databases. In this example, you will add measures of median household income and median housing value to each record.

Note: To use this module, the file Census2k.dld must be in your search path (as specified in the **File > Configuration** dialog box).

To specify demographic inputs and outputs:

1. From the Centrus Desktop main screen, click the **Demographics** tab.
If your paths are configured correctly, the **Demographics** tab shows the input field name, **BLOCK_GRP**, selected in the **Block Group** menu.
2. In the **Assign Outputs** section of the **Demographics** tab, make the following assignments:

Field Name	Available Demographics
MEDHHI00	← 2000 MEDIAN HOUSEHOLD INCOME
MEDHOMEV00	← 2000 MEDIAN HOUSING VALUE

To make an assignment:

1. Highlight a field name in the **Field Name** column on the left.
2. Highlight an item in the **Available Demographics** list on the right.
3. Click the <<**Assign** button to assign the demographic type to the proper field name. The assignment appears to the right of the field name, in the **Output Assignment** column.

Specify Point-in-Polygon Inputs and Outputs

Point-in-Polygon analysis determines within which object, or objects, a point falls. The Point-in-Polygon module can perform analyses on polygons, as well as on buffered lines, buffered points, and buffered polygons. In this example, we'll use Point-in-Polygon analysis to determine whether addresses fall within store trade areas.

To use the Point-in-Polygon module, you need to add *layers*—spatial data files containing information about specific geographic features and locations.

To add a point-in-polygon layer:

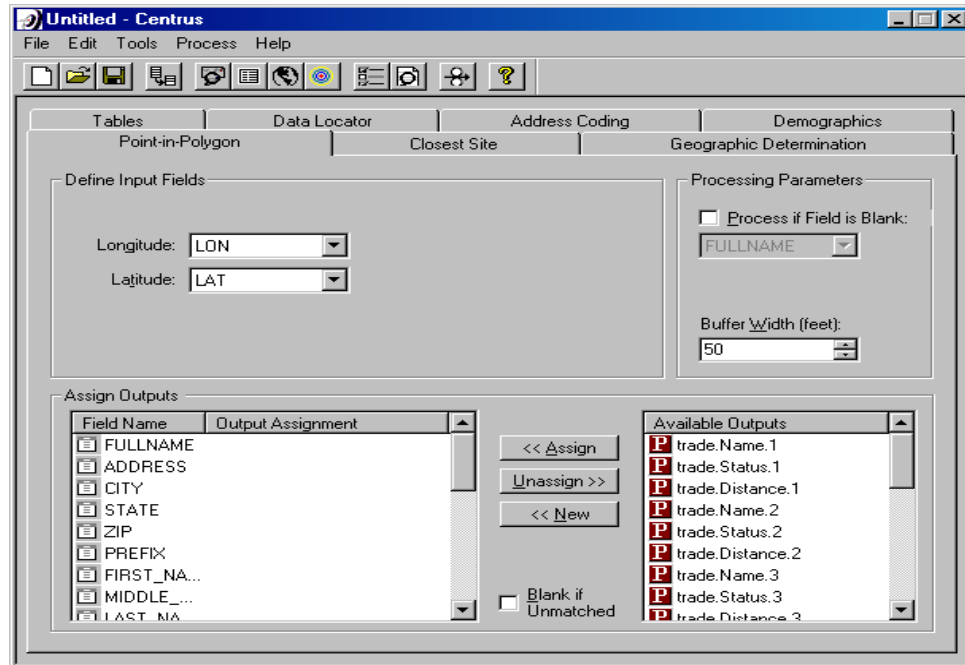
1. From the menu bar, choose **Process > Options**. The **Process Options** dialog box appears.
2. Click the **Point-in-Polygon** tab.
3. Click the **Add** button. The **Select Layers File** dialog box appears.
4. Select the Trade.gsb file from the Tutor directory and click **Open**. The **Centrus -Process Options** dialog box reappears with the file displayed. Click **OK** to return to the main Centrus Desktop screen.

Note: If you attempt to add a layer that is already listed, an error message displays. Click **OK** to clear the error message and continue.

To specify point-in-polygon inputs and outputs:

1. From the Centrus Desktop main screen, click the **Point-in-Polygon** tab.
Centrus Desktop automatically fills in the **Latitude** and **Longitude** input fields for you, based on the assignments you made in the **Address Coding** tab.
2. In the **Outputs** section of the **Point-in-Polygon** tab, make the following assignment:

Field Name	Available Outputs
TRADE_AREA	← trade.Name.1



To make an assignment:

1. Highlight a field name in the **Field Name** column on the left.
2. Highlight an Available Output in the **Available Outputs** list on the right.
3. Click the **<<Assign** button to assign the output to the proper field name. The assignment appears to the right of the field name, in the **Output Assignment** column.

Specify Closest Site Inputs and Outputs

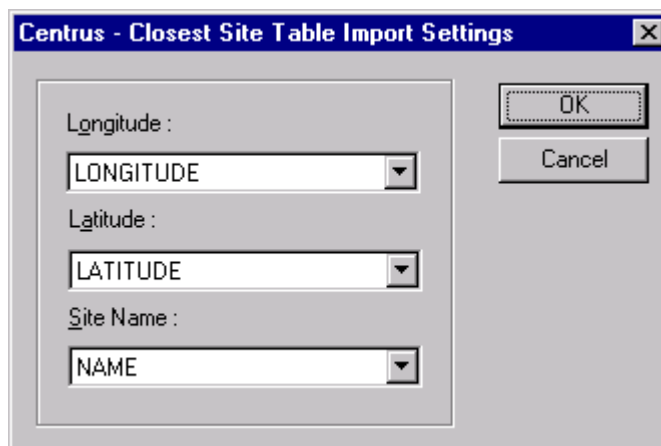
Closest site analysis compares the geocoded addresses (or other points) in your input file to the geocoded site layer specified in the **Closest Site** tab in the **Process Options** dialog box. Centrus Desktop can identify the closest sites for each point in the input file. The site's identifier and the straight-line distance to the site can be returned for each site found. In this example, we'll use Closest Site analysis to find store locations near our customers.

To use the Closest Site module, you need to import *layers*—spatial data files containing information about specific geographic features and locations.

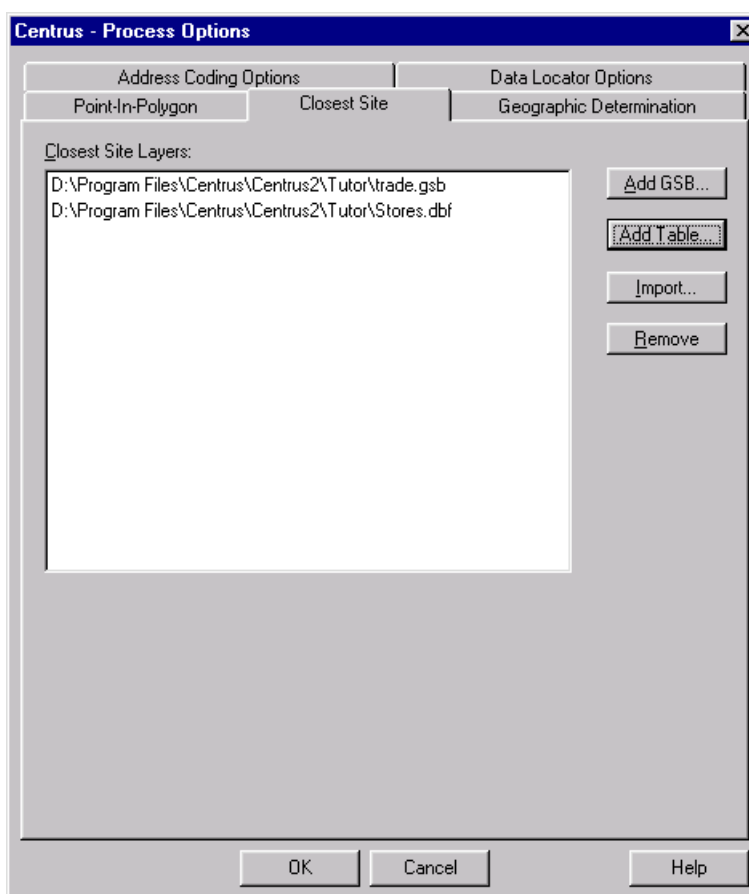
To Import Closest Site Layers:

1. From the menu bar, choose **Process > Options**. The **Process Options** dialog box appears.
2. Click the **Closest Site** tab.
3. Click the **Add GSB** button. The **Select Layers File** dialog box appears. Select the file Trade.gsb from the Tutor directory and click **Open**.

4. Click the **Add Table** button. The **Select File to Process** dialog box appears. Select the geocoded table called Stores.dbf from the Tutor directory and click **Open**. The **Closest Site Table Import Settings** dialog box appears.
5. You should see Longitude, Latitude, and Name listed in the appropriate field boxes. Click **OK** to import the table.



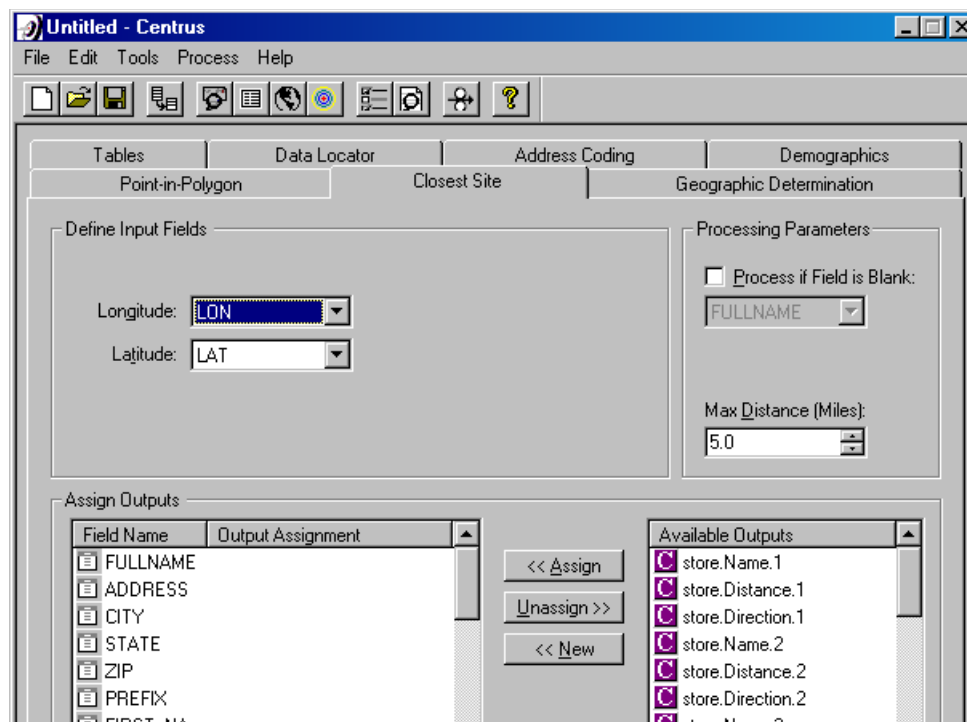
6. The added layers should now appear on the list of Closest Site layers. Click **OK** to return to the main screen.



To specify closest site inputs and outputs:

1. From the Centrus Desktop main screen, click the **Closest Site** tab.
Centrus Desktop automatically fills in the **Latitude** and **Longitude** input fields for you, based on the assignments you made in the **Address Coding** tab.
2. Set the **Max Distance** to 50. **Max Distance** is the distance to search from each record in the customer file. If Max Distance is set to 0, you won't find any store locations near your customers.
3. In the **Assign Outputs** section of the **Closest Site** tab, make the following assignments:

Field Name	Available Outputs
STORE	← Stores.Name.1
DIST2STORE	← Stores.Distance.1



Available Outputs from Trade.gsb are listed first. Scroll down to the Stores.dbf fields.

To make an assignment:

1. Highlight a field name in the **Field Name** column on the left.
2. Highlight an Available Output in the **Available Outputs** list on the right.
3. Click the <<**Assign** button to assign the output to the proper field name. The assignment appears to the right of the field name, in the **Output Assignment** column.

Specify Data Locator Inputs and Outputs

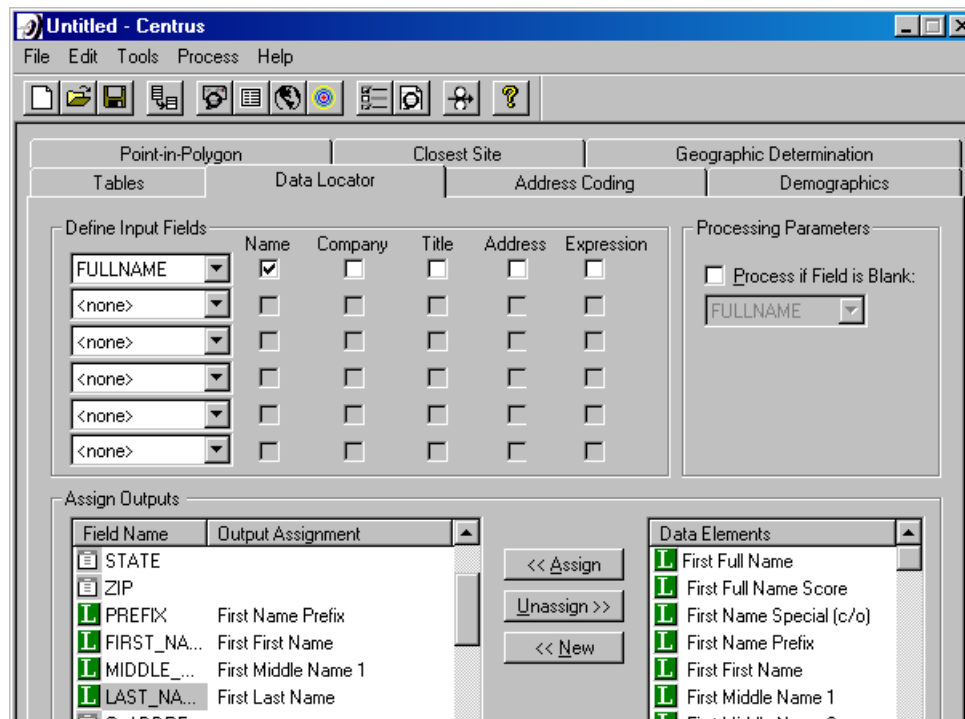
The Data Locator module enables you to take full advantage of the name information in your databases. It uses Sagent's proprietary name parsing technology to analyze databases, identify and extract name elements, and add a variety of name-related information.

In this example, you identify name components for each record and generate appropriate greetings.

To specify data locator inputs and outputs:

1. From the Centrus Desktop main screen, click the **Data Locator** tab.
2. In the **Define Input Fields** section of the **Data Locator** tab, click the arrow next to the first menu to see available input fields. Select the input field called **FULLNAME**, then check the **Name** box.
Note: If no input field has been specified on the **Tables** tab, the only choice available under **Define Input Fields** is **<none>**. Specify an input file on the **Tables** tab before continuing.
3. In the **Assign Outputs** section on the left, select **PREFIX** in the **Field Name** column. In the **Data Elements** list on the right select **First Name Prefix**. Click the **Assign** button to create the assignment.

4. Continuing in the same manner, assign **First First Name** to **FIRST_NAME**, **First Middle Name 1** to **MIDDLE_NAME**, and **First Last Name** to **LAST_NAME**.



Specify Geographic Determination Inputs and Outputs

Note: You must be licensed for the Geographic Determination module to complete this section of the tutorial. Geographic Determination is not available with the demonstration license installed with Desktop.

The Geographic Determination module can report a confidence value for a geocoded address. This confidence factor can be described as either a distance value (between two points and/or lines) or as a percentage representing the amount of overlap between two polygons. The Geographic Determination module compares a confidence surface surrounding a geocoded address (or other points) in your input file to the geocoded site layer you specify in the **Geographic Determination** tab of the **Process Options** dialog box.

To use the Geographic Determination module, you need to add *layers* — spatial data files containing information about specific geographic features and locations.

To add a Geographic Determination layer:

1. From the menu bar, choose **Process > Options**. The **Process Options** dialog box appears.
2. Click the **Geographic Determination** tab.
3. Click the **Add Point File** button. The **Select Spatial File** dialog box appears.
4. Select the Store.gsb file from the Tutor directory and click **Open**. The **Process Options** dialog box reappears with the file displayed.
5. Click **OK** to return to the main Centrus Desktop screen.

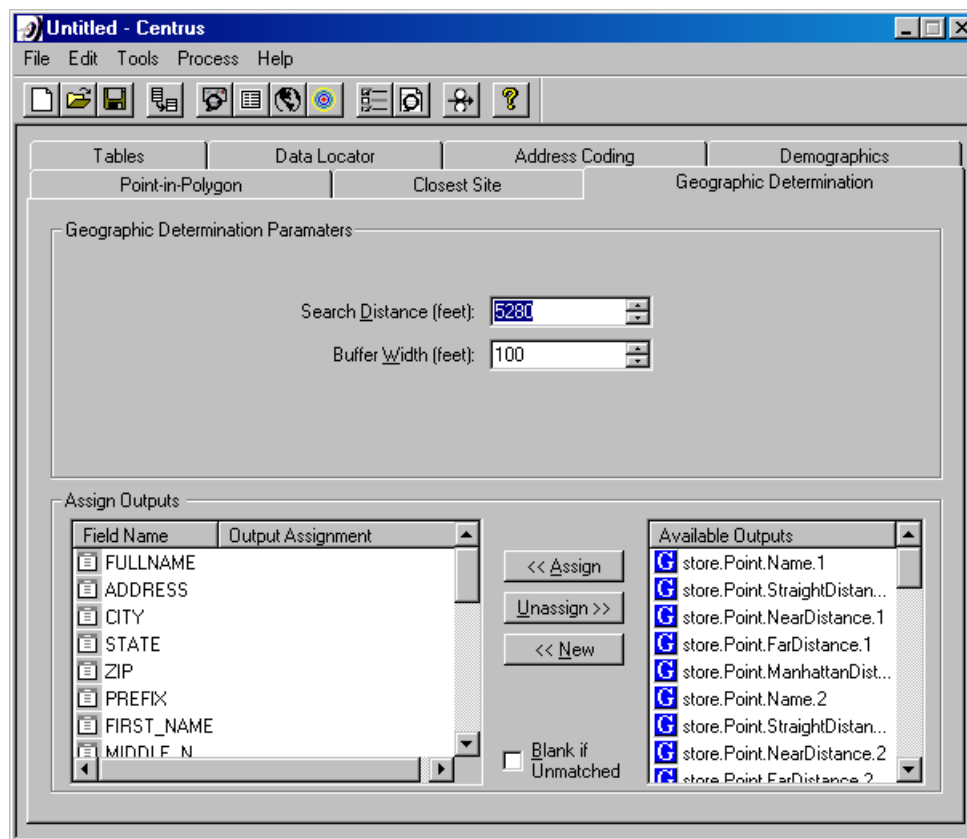
For this example, we are interested in knowing the maximum and minimum possible distance between a customer and a store site. This information provides a confidence range for the distance between the customer and the store site.

To specify Geographic Determination inputs and outputs:

1. From the Centrus Desktop main screen, click the **Geographic Determination** tab.

Centrus Desktop automatically fills in the **Search Distance** and **Buffer Width** input fields for you. **Search Distance** (in feet) is the distance to search from each record in the customer file. (If **Search Distance** is set to 0, the Geographic Determination module does not find any store locations near your customers.) **Buffer Width** (in feet) is the polygon (overlap) around the point.

2. In the **Assign Outputs** section of the Geographic Determination tab, add the following new fields:
 - Name
 - Near distance
 - Far distance



To add new fields:

1. Highlight an Available Output in the **Available Outputs** column on the right (for example, store.point.name.1, store.point.neardistance.1, and store.point.fardistance.1).
 2. Click <<New.
- A dialog box appears with a suggested column name and length.
3. Click **OK** to accept the default values.

The new field is created and is appended to the end of the output field list.

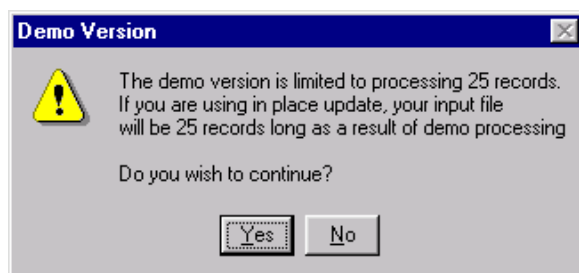
Processing a Task

Before processing, check that you have:

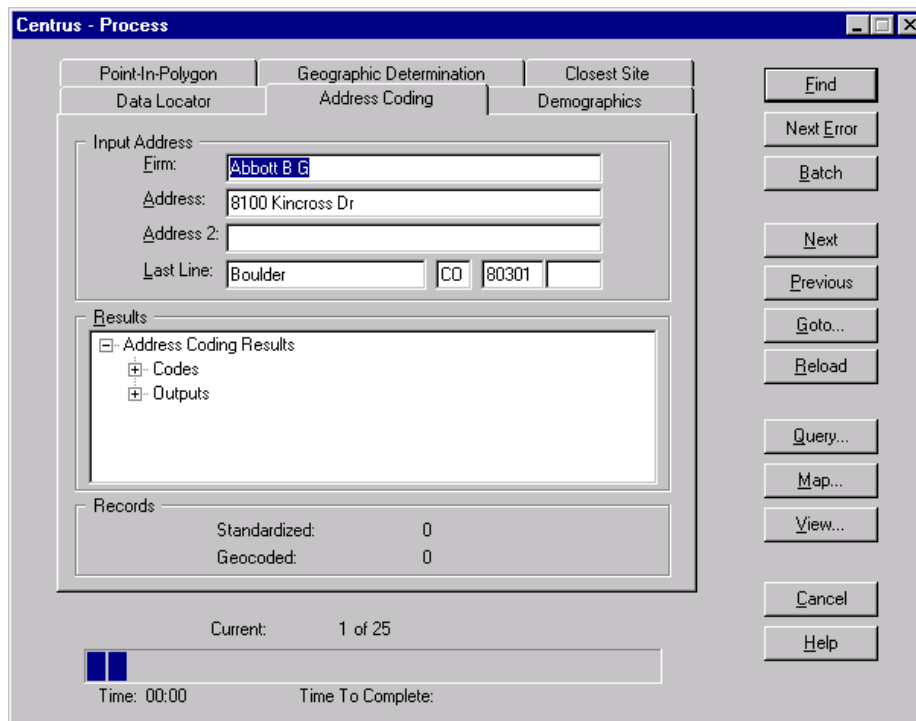
- Selected an input and, if desired, an output file (see “Specify Input and Output Tables” on page 43)
- Set the processing options if necessary (in this tutorial, processing options were described under “Specify Point-in-Polygon Inputs and Outputs” on page 47 and “Specify Closest Site Inputs and Outputs” on page 48)
- Specified the input and output fields for the modules you are using (see the “Specify Inputs and Outputs” section for each module you are using)

To process a task:

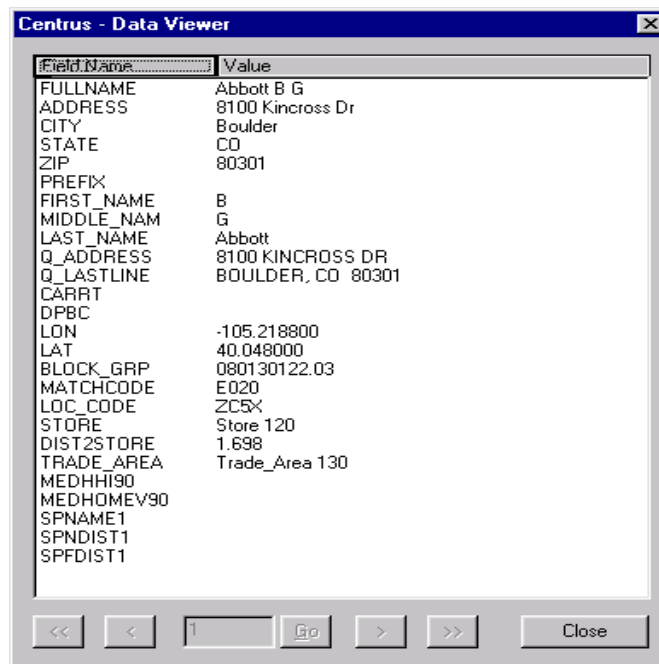
1. From the main Centrus Desktop menu, select **Process > Process Task**. If you see the demo warning message, click **Yes** to continue.



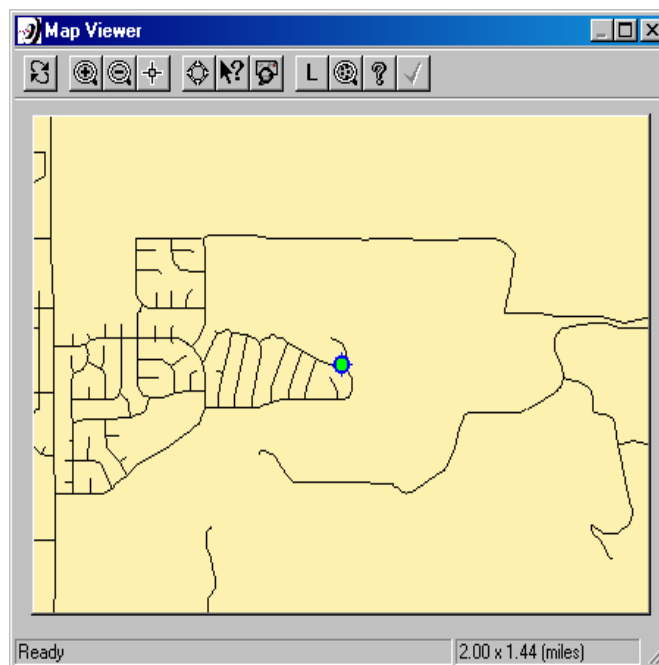
2. You now see the first record in your file. Click each of the tabs to see the results of each of the processes you have set up. The **Address Coding** tab is shown below.





- Click the **View** button to see all processing results for the current record.

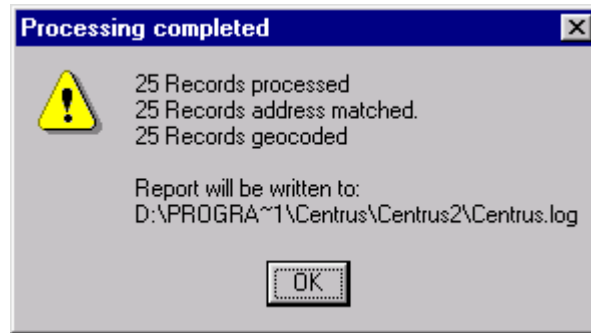


- Click the **Map** button to see the first record plotted on the a street network.



- In the **Map Viewer**, click the **Zoom In** button  and click on the map to zoom in for a better view of the relationship between the record being processed, store trade areas, and the store locations.

6. Click the **OK** button  to close the **Map Viewer**, then click the **Batch** button to process the rest of the file. When all records have been processed, a message appears telling you that processing has completed. Select the **X** button at the upper right corner of the dialog box. This will close the map.



7. Click **OK**, then close the **Process** dialog box by clicking the **Done** button.
8. If you wish to save this task, on the menu bar select **File > Save Task**. Give the task a name, and then click the **Save** button.

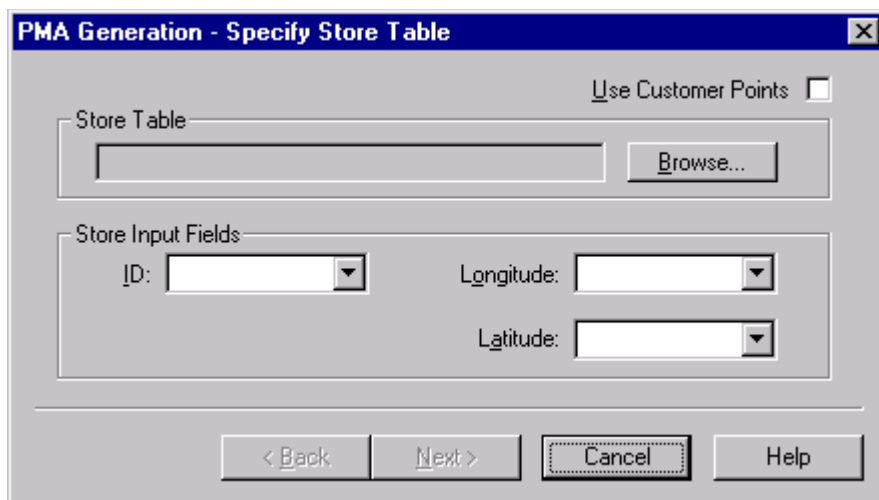
Congratulations, you have successfully processed a task with Centrus Desktop! Continue with “Creating a Primary Market Area (PMA)” to learn how to use Centrus Desktop’s PMA function.

Creating a Primary Market Area (PMA)

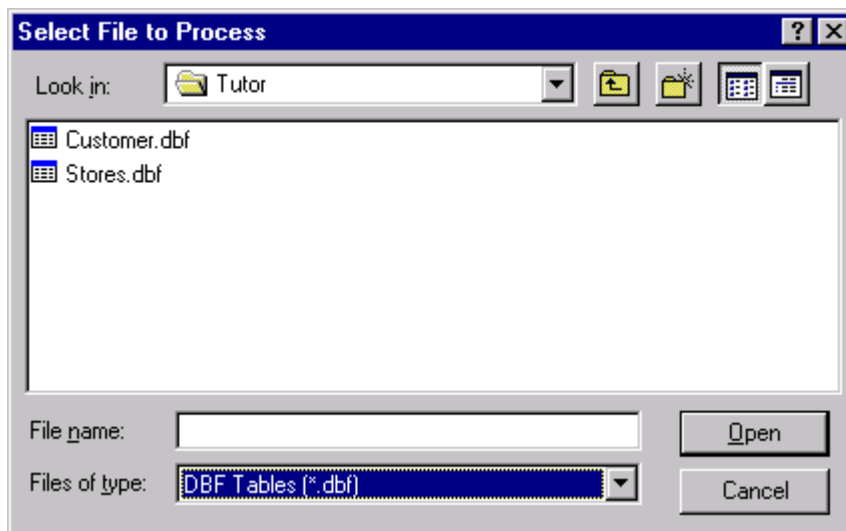
In this part of the tutorial, you use the Primary Market Area (PMA) function to create a store-based PMA from a sample stores database file. You create a PMA of non-overlapping rings drawn in concentric circles at a distance of 2 miles from a store location. The PMA you create is saved as a data layer, which contains spatial data about a geographical location, in Centrus Desktop’s GSB format.

To create a primary market area:

1. From the menu, select **Tools > Shape Layers > PMA Generate**. The **PMA Generation - Specify Store Table** dialog box appears.



2. Click the **Browse** button to display the **Select File to Process** dialog box. Find and open the Tutor directory. (If you accepted the default installation directory, the Tutor directory is located in C:\Program Files\Centrus\Centrus2\Tutor.)

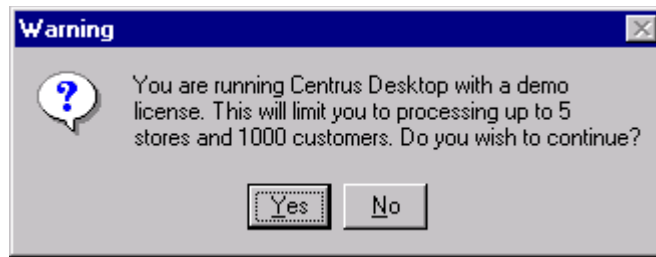


3. Select the Stores.dbf file from the Tutor directory. This file contains store locations. Click **Open**. The **PMA Generation - Specify Store Table** dialog box reappears with the file displayed.

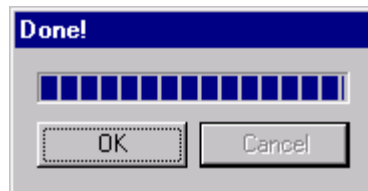
- Under **Store Input Fields**, select an ID of **NAME**. Click **Next >**. The **PMA Generation - Define Output Files** dialog box appears.

- Under **Type**, select **Non-Overlapping Rings** and **Inclusive**. Under **Primary Radius**, enter 2. Click the **Browse** button to specify an output, or GSB, file to contain your PMA information. The **Save As** dialog box appears.
- Select the Tutor directory from the **Save in** pull-down list. Under **File name**, specify an output file named Trade.gsb. (If the file already exists, select that file. When asked if you wish to replace it, answer **Yes**.) Click the **Save** button. The **PMA Generation - Define Output Files** dialog box reappears.
- Click **Next**. The **PMA Generation - Layer Options** dialog box appears. Select **Add to PIP Layer**, **Show In Map**, **Show Stores**, and **Show Roads**.

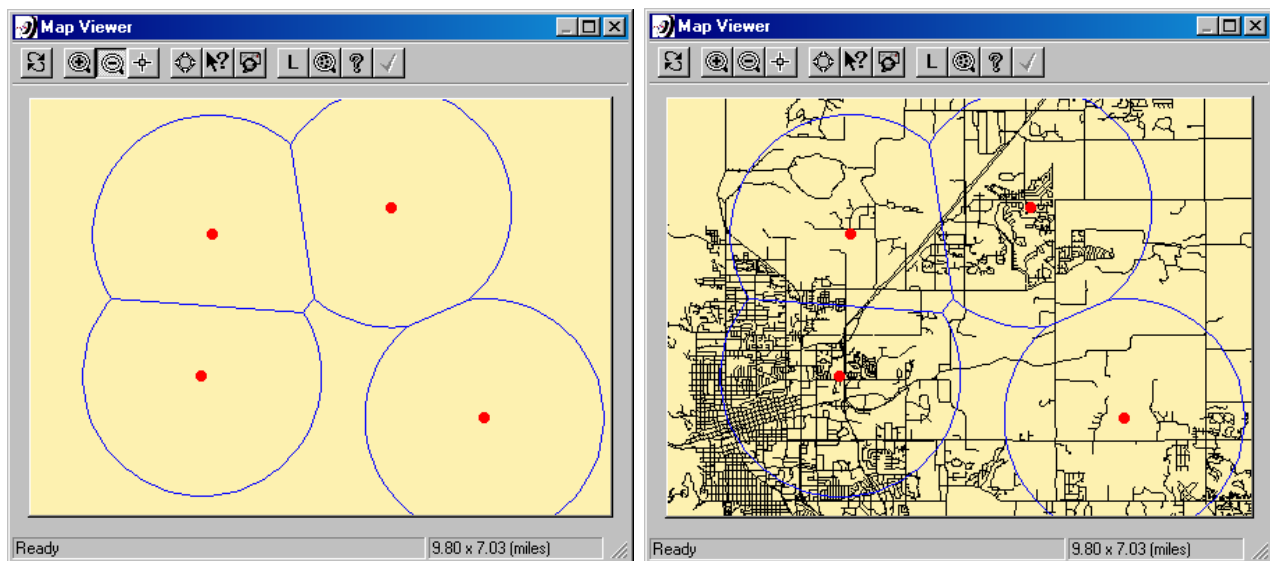
- Click the **Finish** button. If you are running Centrus Desktop in demo mode, the following message displays:



- Click **Yes** to continue. The Trade.gsb file is created, and a confirmation message is displayed.



- Click **OK** to continue. The **Map Viewer** window is displayed. At left you see PMA-created market areas. The points within those market areas are store locations specified in the file Stores.dbf. The picture at right shows the same market areas drawn on top of a street grid.



The information you see on your screen was determined by your choices in the **PMA Generation - Layer Options** screen. To change the display, click the **Layers** button on the toolbar. See "Working with Layers" on page 78 for more information on layers and screen display.

Close the **Map Viewer** window to return to the main Centrus Desktop screen.

Chapter 4


Centrus Desktop Tools

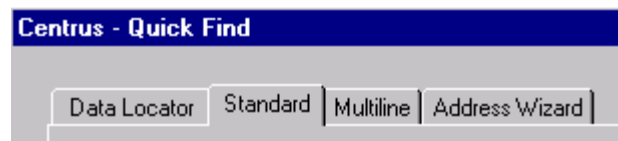
Centrus Desktop includes tools to help you find and analyze processed records. These tools include:

- QuickFind
- Data Viewer
- Map Viewer
- Shape Layers

QuickFind

With QuickFind, you can access a variety of information about an address.

Click the **QuickFind** button  or select **Tools > QuickFind** to start QuickFind.



QuickFind has four modes:

- Data Locator Mode
- Standard Mode
- Multiline Mode

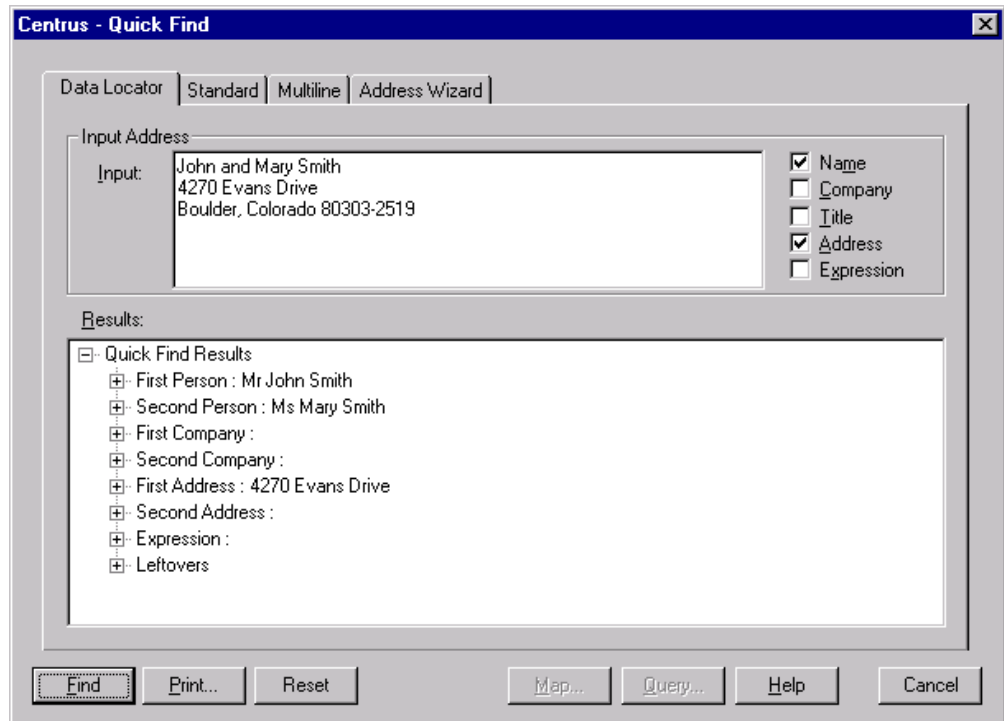
■ Address Wizard Mode

All use the **Negate Longitudes**, **Census ID Content**, **Census ID Format**, and **NAD** settings specified in the **Options** dialog box.

QuickFind also includes a comprehensive Query function, which provides complete access to the address standardization and geocoding database.

Data Locator Mode

In Data Locator Mode, you enter a standard or free-form name and address (or other information) into a single edit box. This address can either be typed in or pasted in from other applications. Click the check boxes to the right of the input box to select the type of information you want Data Locator to extract. Once the information is entered, click the **Find** button to process the address.

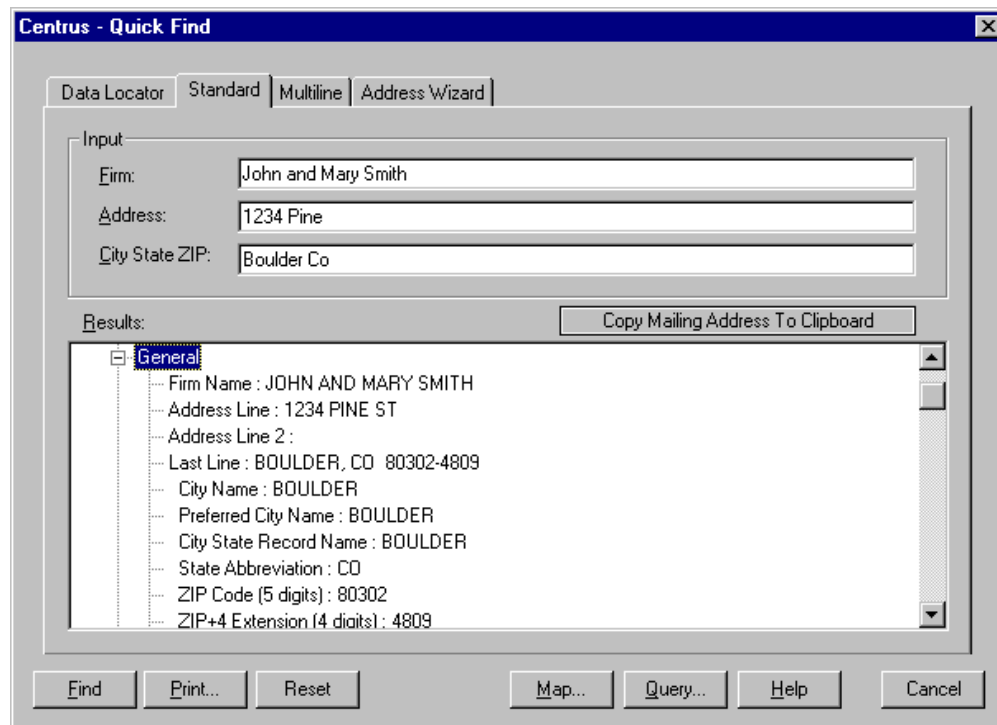


The processed information is shown in the **Results** section of the dialog box. The data is presented as a hierarchical tree. Double-click on an item to expand or contract the information displayed.

Note: Unlike Standard, Multiline, and Address Wizard mode, the Data Locator mode does not standardize address results.

Standard Mode

In Standard mode, you enter a standard three-part address into three text boxes: **Firm**, **Address**, and **City State Zip**. Once the information is entered, click the **Find** button to process the address.



The processed information is shown in the **Results** section of the dialog box. The data is presented as a hierarchical tree. Double-click on an item to expand or contract the information displayed.

If a match is made, use the Map function to map the current address.

If a match is not made, the Query function can be accessed by clicking the **Query** button.

The **Print** button displays the output in Notepad.

The **Copy Mailing Address To Clipboard** button copies the Firm Name (if available), Address Line, Address Line 2 (if available), and Last Line to the Clipboard.

Note: The information copied to the Clipboard is from the results tree, rather than the input boxes. Hence, the information is only available after a successful match has been found.

Multiline Mode

In Multiline mode, you enter a free-form address into a single text box. This address can either be typed in or pasted in from other applications. The street address information (such as 123 Main St) should be listed before last line information (city, state, and ZIP). If a firm name is given, it should be listed first. Centrus Desktop ignores any lines of extraneous information. Once the information is entered, click the **Find** button to process the address.

The processed information is shown in the **Results** section of the dialog box. The data is presented as a hierarchical tree. Double-click on an item to expand or contract the information displayed.

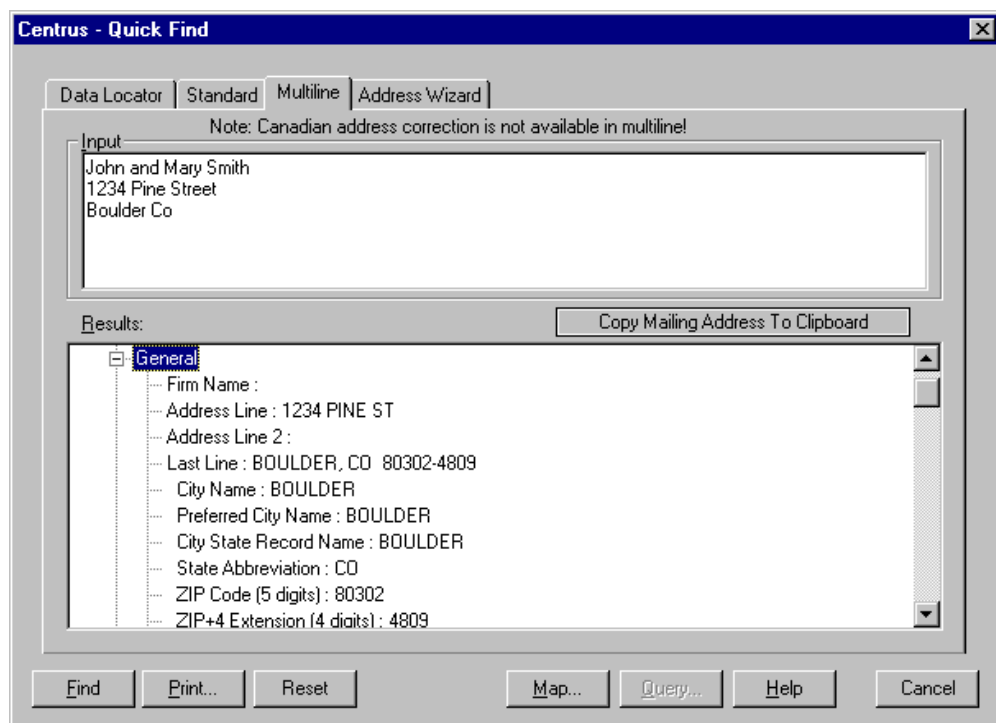
If a match is made, the Map function can be used to map the current address.

The Query function is unavailable in Multiline mode.

The **Print** button displays the output in Notepad.

The **Copy Mailing Address To Clipboard** button copies the Firm Name (if available), Address Line, Address Line 2 (if available), and Last Line to the Clipboard.

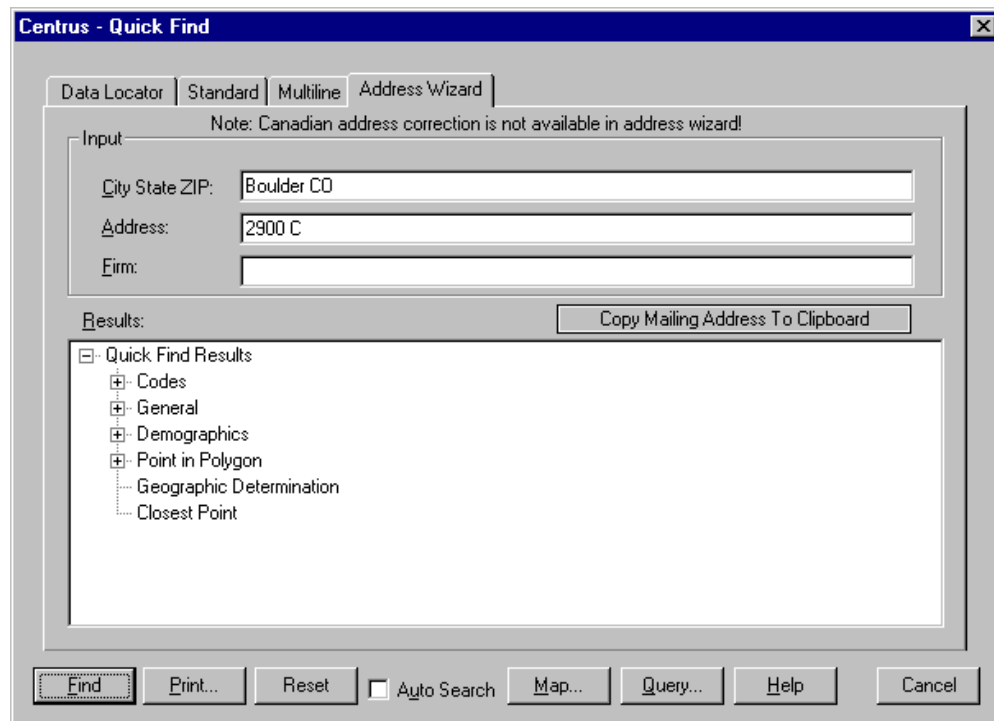
Note: In Multiline mode, the Private Mail Box Designator and Private Mail Box Number output address elements are not available.



Note: This function is not supported for Canadian addresses.

Address Wizard Mode

Address Wizard mode demonstrates the capabilities of the Address Wizard development library. This library is used in call-center and other applications where address information is being entered on a real-time basis.

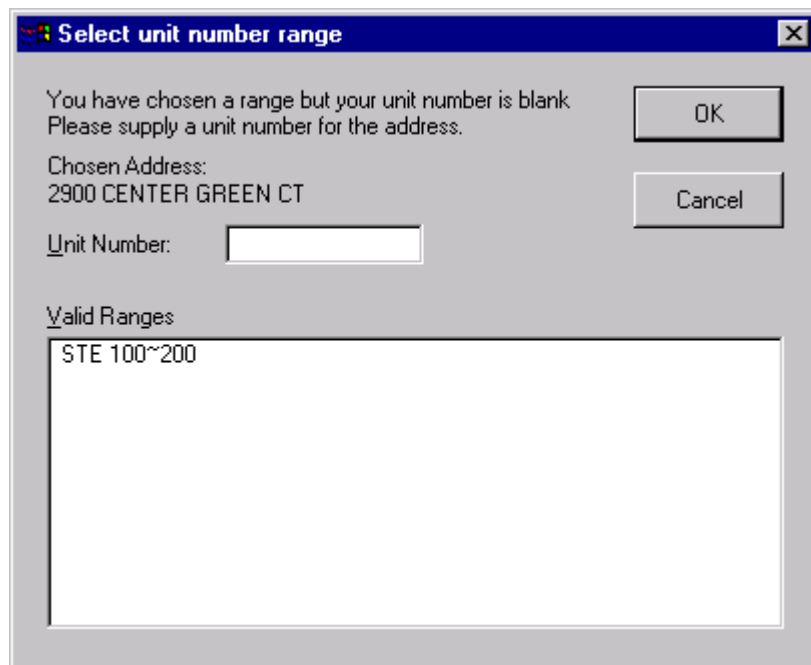


Address Wizard verifies that correct and complete information is being entered. Using Address Wizard, all elements of an address are corrected. The processed information is shown in the **Results** section of the dialog box. The data is presented as a hierarchical tree. Double-click an item to expand or contract the information displayed.

The **Print** button displays the output in Notepad.

The **Copy Mailing Address To Clipboard** button copies the Firm Name (if available), Address Line, Address Line 2 (if available), and Last Line to the Clipboard.

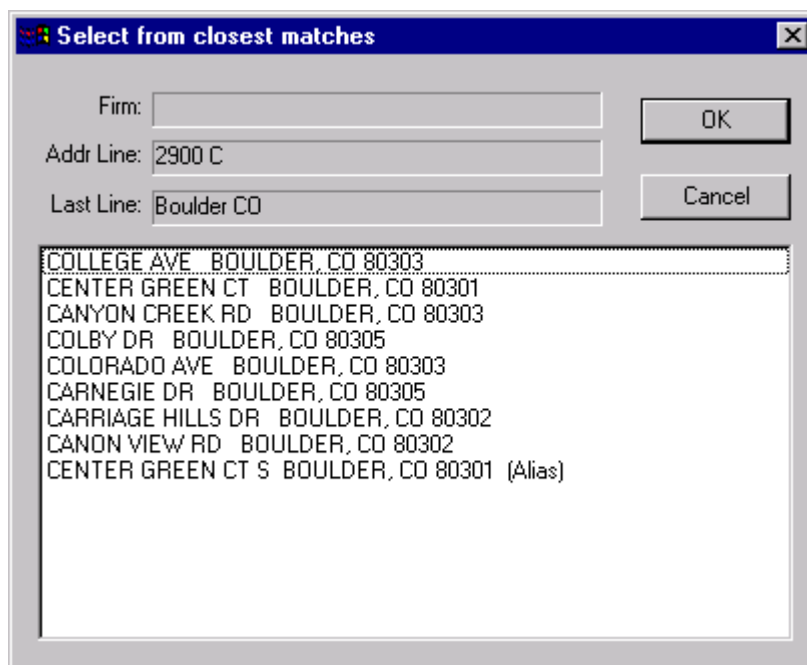
If an address is entered that is known to have apartments or multiple units, a dialog box displays with the possible unit numbers.



The **Auto Search** option at the bottom of the **Address Wizard** window tells Address Wizard to begin searching after two seconds of keyboard inactivity. Address Wizard also attempts a match on partial addresses. In a busy call-center, this is a real time saver.

Note: If **Auto Search** is not checked, you must click the **Find** button or press **Enter** to perform the address lookup.

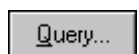
If the address has more than one possible match, a dialog appears displaying the possible matches. Select the appropriate street name and click **OK**.



Note: Address Wizard Mode is not supported for Canadian addresses.

The Query Function

The Query function is available through the Centrus Desktop **QuickFind** and **Process** dialog boxes.



It provides complete access to the address standardization and geocoding database. For example, **Query** can be used to display all streets that start with the letter 'K' in the city of Washington, DC. The query can be defined and refined in a variety of ways, and results from **Query** are virtually instantaneous.

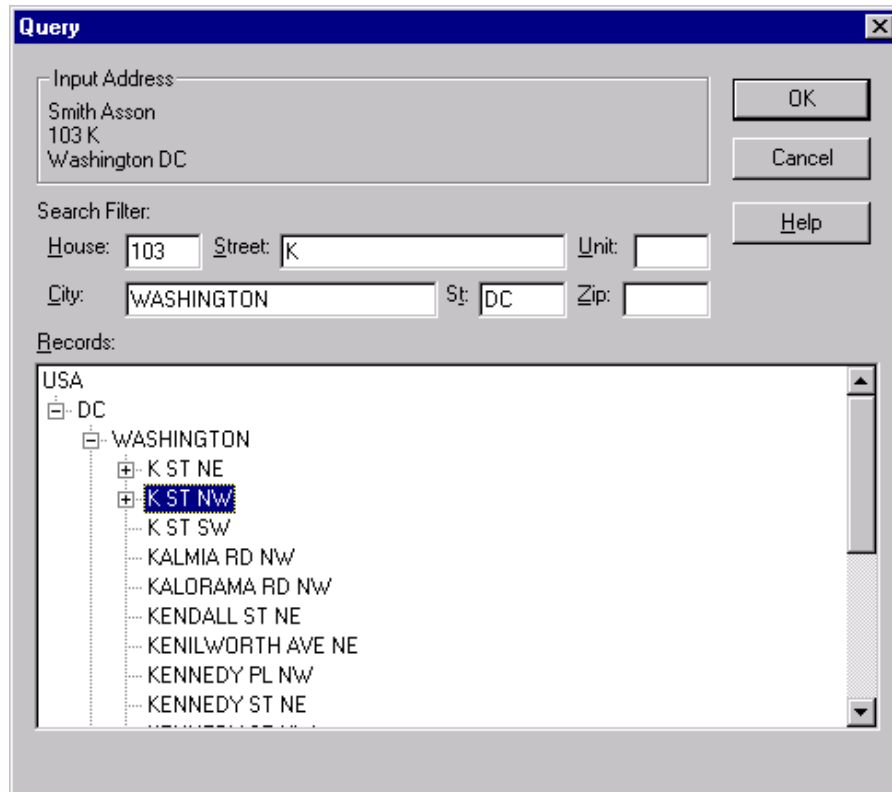
This section covers the following topics:

- The Query Dialog Box
- Specifying a Search Filter
- The Query Tree
- Navigating the Query Tree
- Selecting a Match Using Query

Note: This function is not supported for Canadian addresses.

The Query Dialog Box

When you click the **Query** button in the **QuickFind** or **Process** dialog boxes, the current address record is loaded into the **Query** dialog box and displayed. Query processes the input record and expands as many levels as possible in the **Records** section of the dialog box.



The **Query** dialog box consists of three different sections. The top section lists the **Input Address**. This is a fixed, non-editable area that displays the record as entered in the **Input Address** section of the **Process** or **QuickFind** dialog boxes. It is displayed as a reference to the original address.

The middle section of the **Query** dialog box is the **Search Filter**. This section consists of fields for **House**, **Street**, **Unit**, **City**, **State**, and **Zip**. Only street records matching all of these fields display in the **Records** section.

The bottom section of the **Query** dialog box is the **Records** section. This section displays the information from the address standardization and geocoding database. It is hierarchical in nature in that it displays only the requested level of detail. The level detail can be increased or decreased.

Specifying a Search Filter

A query can display all streets that start with 'K' in Washington, but queries can also be much more refined. For example, you could have a query that specified "show all streets in ZIP Code 80302 that begin with APPLE and have a house number range that includes "4300."

Search Filter:

House: 4300 Street: APPLE Unit:

City: BOULDER St: CO Zip: 00000

The **Search Filter** section of the **Query** dialog box is used to specify the limits of the query. In general, the more information entered in the search filter, the more refined the search results. This means that fewer possible matches display in the **Records** section.

The search filter requires the user to enter a city and state, or a ZIP Code, or both. If the **City** and **State** entry are for one state but the **ZIP** is for a different state, the two states are displayed in the **Records** section. You must choose which state to search. Centrus Desktop does not list more than two states in the **Records** section.

The **City** entry needs only the first letter of the cities that are to be displayed. Centrus Desktop displays all the cities within the given state which match the letters in the **City** entry. The state abbreviation, however, must be a recognized state abbreviation or the entries for both **City** and **State** are ignored.

The **ZIP** entry is used solely to generate the list of possible cities. It does not determine which entries are displayed in the **Records** section. The **Zip** entry must be a valid five-digit or three-digit ZIP Code. Three-digit ZIP Codes are commonly referred to as *Sectional Centers*. A Sectional Center is comprised of all five-digit ZIP Codes that begin with those three digits.

The **Street** entry may have zero or more characters entered. If no street data is entered all streets are displayed. If the **Street** entry was "APPLEW", streets with the name of "Applewood" or "Appleworm" would both match, but a street with the name "Apple" would not match and would not be displayed.

Note: The **Street** entry should be a base street name only; it should not contain pre-directionals, post-directionals, or street types. For example, "Vine" would be a valid entry for Boulder, CO, but "Vine PL" would be invalid because of the additional street type ("PL").

The **House** entry specifies a house number. The **House** entry is used to constrain the search to show only those streets that have a block on which that house number would fall. If the entry for **House** was "1000", a house range of "200 to 300" would not match and would not be displayed. This entry may be blank, in which case all ranges would match and the search would not be constrained by house number.

The Unit entry is very similar to the House entry, except that it contains the unit number, such as “12” or “E”. House ranges that do not contain the unit number entered do not display.

Note: The U. S. Postal Service does not list separate unit numbers for all buildings. Due to this limitation, house range entries that do not have any unit numbers match any unit number entered. This is consistent with USPS CASS requirements.

The Query Tree

Queries are displayed in a hierarchical nature in the Records section of the Query dialog box. Each level of the hierarchical tree is indented below the rest. All entries are sorted alphabetically, including Street Block and House Range. The levels of the tree are:

```

Country
    State
        City
            Street Name
                Street Block
                    House Range
  
```

Country

Always USA. Centrus Desktop’s area is the 50 states and Washington, D.C., plus Puerto Rico and other U.S. Protectorates.

State

Displays the two-letter state abbreviations that match the search filter criteria. If the ZIP Code is for a different state, two state abbreviations display.

City

Displays the city names that match the search filter. If the City entry in the filter was “C”, then all city names beginning with “C” are displayed. Some city names are displayed with a different name in parentheses. The name in parentheses is the city name that should be used when mailing to that location. The name outside of the parentheses is a name that the USPS recognizes and may not be a city name at all. These names might be buildings, military installations, or even large corporations. For example, a search for

“READERS DIGEST” in the city field and “NY” in the state field results in “READERS DIGEST (PLEASANTVILLE)”. When a house range record is chosen, the proper USPS City name is displayed in the bottom left corner of the dialog box.

Street Name

Contains the full street name including directionals and street type. For example, there are separate entries for “N Main St” and “S Main St”. The entries are sorted by street name, then street type, then predirectional, then postdirectional. A street may also appear with (*Alias*) next to the street name, which indicates a different name was used to match street name with the search filter.

Street Block

Contains one “block” of house ranges. A Street Block often is the same as an actual city block, but not always. In some instances, a Street Block may represent a partial city block. Less frequently, a Street Block represents several city blocks. Street blocks list a range of all house numbers that might appear on that block. In most instances, all house numbers within the range shown are valid. On some blocks, however, there may be gaps within the range where a house number is invalid. For example, the Street Block might indicate “100-199 Main St,” but inspection of the House Range records shows records for “100-120 Main St” and for “150-199 Main St.” In this instance, a house number of 140 would be invalid.

House Range

Displays the actual range information contained in the USPS files. An example would be “1000-1098 Kearney St NE, 20017-4526 (S,E).” The first part of this line is the house range and full street name. The next entry is the nine digit ZIP Code for that house range. In parentheses are displayed the Record Type and Record Parity, described below:

Record Type

Indicators are as follows:

- “S” for Street
- “R” for Rural Route/highway contract
- “F” for Firm
- “H” for high-rise, building, or apartment
- “P” for P.O. Boxes and General Delivery
- “G” for General Delivery records

A Street record consists of a house number range and can include any unit ranges given. A Firm record has a firm name associated with it. This record should only be chosen if the firm name is a correct part of the address being searched for. A High-rise record usually denotes an apartment or office building and has a range of unit numbers. A High-rise record should only be chosen if the unit number of the input address is within the unit range of the High-rise record. P.O. Box records simply list all available box numbers or the General Delivery record. When a Firm record is chosen, the Firm name is displayed in the bottom right corner of the dialog box.

Should there not be a match on firm name or unit number, both Firm and High-rise records should always have a “default” record that is either for the building as a whole, or for the entire block. Assigning a non-matching firm or high-rise record may result in incorrect ZIP+4 or carrier route information being assigned.

Range Parity

Indicates whether the House Range contains (O)dd, (E)ven or (B)oth types of house numbers. If the parity of the record does not match the input record, it should not be chosen, as incorrect ZIP+4 and carrier route information may be assigned.

Navigating the Query Tree

When a query result is first displayed, the *Records* section displays Country and State. Double-click on State to expand the query tree. Double-clicking on any entry expands or contracts the tree, unless the selected entry is at the House Range level of the tree. If the selected entry is at the House Range level, double-clicking the entry accepts it as the correct record and the information is pasted into the *Results* section of the Process dialog box. If the selected entry is the last entry in the tree, and it is not a House Range entry, the tree is expanded, showing the next lowest level. If the selected entry is in the middle of the tree, the tree is collapsed to that point.

For example, let's assume that the search filter is for all streets that start with "KEN" in Washington DC, and have a possible house number of 1000. The following query tree is displayed.

The screenshot shows the QuickFind search interface. The Search Filter section includes fields for House (1000), Street (KEN), Unit (empty), City (WASHINGTON), St (DC), and Zip (00000). A Help button is located to the right. Below the Search Filter is the Records section, which displays a query tree. The tree starts with USA, followed by DC, and then WASHINGTON. Under WASHINGTON, several entries are listed: WASHINGTON (ANACOSTIA ANNEX), WASHINGTON (WASHINGTON NAVY YARD), WASHINGTON DC POST OFFICE (WASHINGTON), WASHINGTON GAS (WASHINGTON), WASHINGTON INTEL BUR (WASHINGTON), WASHINGTON NAVY YARD, and WASHINGTON POST (WASHINGTON). The WASHINGTON entry is highlighted with a blue box.

If we were to expand the WASHINGTON branch and select the KENILWORTH entry, the tree would then display.

The screenshot shows the QuickFind search interface with the same Search Filter as the previous image. The Records section displays a query tree that has been expanded to show the KENILWORTH entry. The tree structure is: USA > DC > WASHINGTON > KENDALL ST NE > KENILWORTH AVE NE. The KENILWORTH AVE NE entry is highlighted with a blue box. Below this entry, several sub-entries are listed: KENILWORTH AVE NE, 1000-1048 KENILWORTH AVE NE, 901-1099 KENILWORTH AVE NE, KENNEDY PL NW, KENNEDY ST NW, KENYON ST NW, and WASHINGTON (ANACOSTIA ANNEX). A vertical scrollbar is visible on the right side of the Records section.

Any adjustment of the Search Filter collapses the query tree to the proper point. If House, Street, or Unit is modified, the tree is collapsed to the City level. If City, State, or ZIP is modified, the tree is collapsed to the State or Country level, depending upon the modification.


Selecting a Match Using Query

You select a match within the Query dialog box by selecting a House Range entry from the query tree. If there are entries in the House and Unit section of the Search Filter, then the selected entry's information is simply pasted to the Process or QuickFind dialog boxes in the *Results* section.

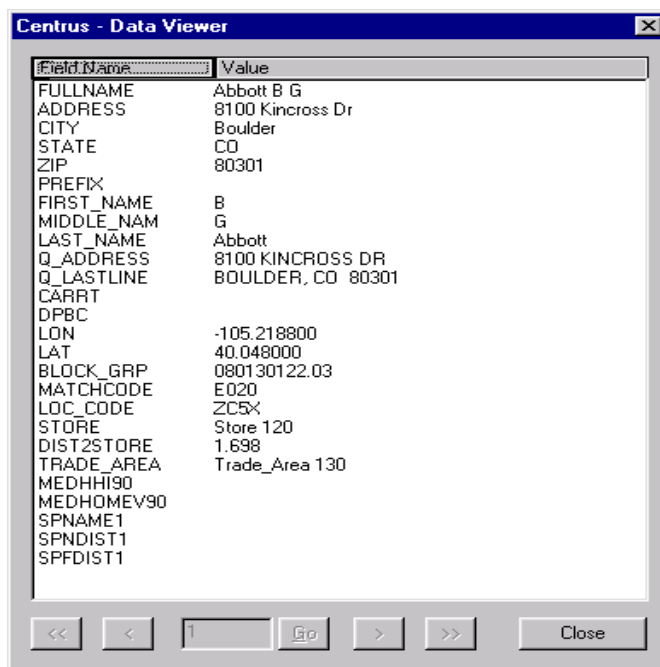
If either the House entry was blank, or if the Unit entry was blank but the House Range entry chosen contained unit numbers, Centrus Desktop displays a dialog box allowing you to select House and Unit information before returning to the QuickFind or Process dialog boxes.

Data Viewer

The Data Viewer allows you to view the contents of the current input file.

To start Data Viewer, click the Data Viewer button  or select **Tools > Data Viewer**.


Field names are shown in the left column of the Data Viewer window; field values appear in the right column. Click the arrow buttons to navigate forwards and backwards through the file, or type a record number into the text box and click the **Go** button to view a particular record.

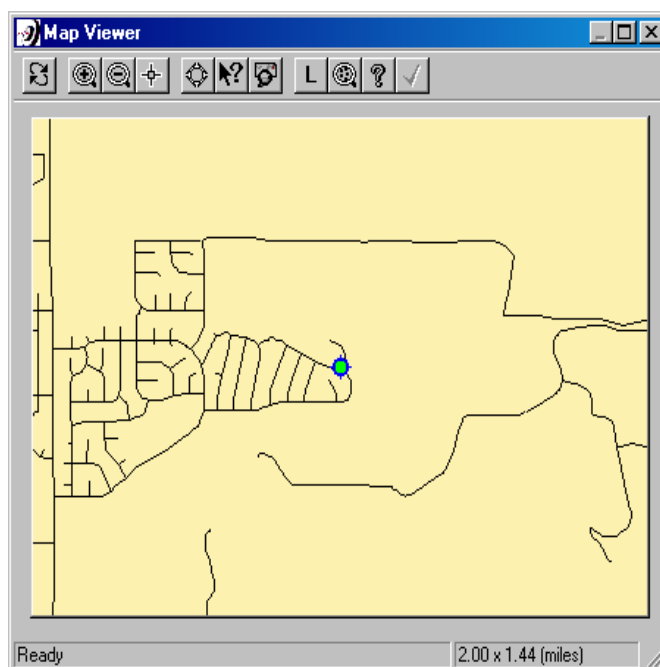


Map Viewer

The Map Viewer displays streets and county boundaries for the city currently entered. The currently geocoded point, if any, is shown when the map is first displayed. The point can be moved or created manually on the map, allowing users with local knowledge to place points for locations that Centrus Desktop failed to match.

Note: There are no street segments available for mapping for Canada.

To start the Map Viewer, click the Map Viewer button  on the main toolbar, or click the **Map** button within the QuickFind or Process dialog boxes. You can also select **Tools > Map Viewer**.




This section covers these topics:


- Map Viewer Buttons
- Manual Geocoding with the Map Viewer
- Working with Layers
- Increasing Centrus Desktop Mapper Performance


Map Viewer Buttons

Redraw Button


Clicking the Redraw button  causes the map to be redrawn at the same scale.

Zoom In Button


The Zoom In button  “zooms in” the current map view, increasing the magnification while showing a smaller area. To zoom in, click the button, then click the map.

You can also define an area to be magnified. Simply click the Zoom In button , then click and drag the pointer over the map window to define the area to be zoomed.


Zoom Out Button

The Zoom Out button  “zooms out” the current map view, decreasing the magnification while showing a larger area. To zoom out, click the button, then click the map.


Zoom To Extent Button

The Zoom To Extent button  allows you to specify one or more spatial layers and zooms the map window to show all features within the specified layer(s).


Center Button

The Center button  redraws the map with the selected point at the center of the map window. To center the map on a point, click the button, then click the map at the new point.


Select Button

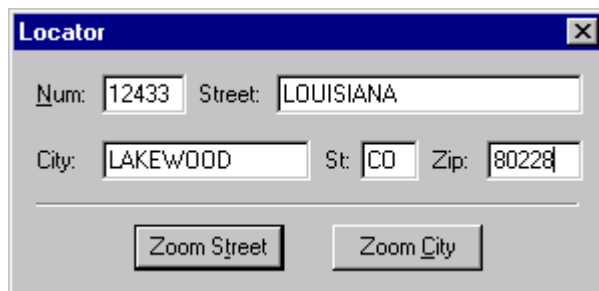
The Select button  is used to manually place a point on the map. See “Manual Geocoding with the Map Viewer” on page 77 for more information.

Info Button

When you click the Info button , the pointer becomes a small square. Position this square over a mapped street and click the left mouse button to view information about that street segment.

Locator Button

The Locator button  opens the Locator dialog box, where you can enter an address or city to locate on the map.



The Locator dialog box is titled "Locator" and contains the following fields and buttons:


- Num:** 12433
- Street:** LOUISIANA
- City:** LAKEWOOD
- St:** CO
- Zip:** 80228
- Buttons:** Zoom Street, Zoom City

To locate a street, enter the street address, city, state, and ZIP Code (if known). Click the **Zoom Street** button to display all street segments in the specified city with a matching street name.


To locate a city, type either the city name and the state abbreviation, or a ZIP Code. Click the **Zoom City** button to display the streets in the city or ZIP Code within the map window.

Note: The Zoom Street function does not work correctly if street input includes street type or pre-/post-directionals. Enter only the street name in the Street box. For example, "West Louisiana Avenue" should be entered as just "Louisiana".


Layers Button

The Layers button  opens the Layers dialog box. See "Working with Layers" on page 78 for more information.

Help Button



Click the Help button  to view Centrus Desktop Help.

OK Button

Click the OK button  to accept the most recently geocoded location and close the Map Viewer. If you manually geocoded a point, the Census ID, latitude, and longitude are copied to the Results section of the QuickFind or Process dialog box.


Manual Geocoding with the Map Viewer

If Centrus Desktop is unable to match an address, you can assign information to the point manually. This is most useful if you have knowledge of the area or maps to which you can refer.

Using the **Zoom**  and **Info**  tools, locate the spot where the point should be placed, as well as a street whose Census ID and coordinates you wish to assign to the new location. Click the **Select** button, then place the pointer over the site to be geocoded and click the left mouse button. This “anchors” the point. Drag the pointer to the street segment whose Census ID you wish to assign to the new location, creating an “attach line” between the point to be geocoded and the existing street segment. The Census ID and coordinates of the selected street segment are assigned to the new location. In the message area at the bottom of the screen, the Census ID is displayed, as well as the side of the street that the ID was taken from.

An address that has been manually geocoded displays a location code of “M”.

Working with Layers

Click the **Layers** button  in the Map dialog box to display the Layers dialog box. This is where you control the map’s appearance. You can add features, change colors, change drawing priority, and more.

Centrus Desktop comes with two default layers: *Primary Streets* and *Secondary Streets*. Two additional layers (COUNTY.GSB and STATE.GSB) come with the Centrus Data Products Suite CDs. Other layers can be added by clicking the **Add** button.

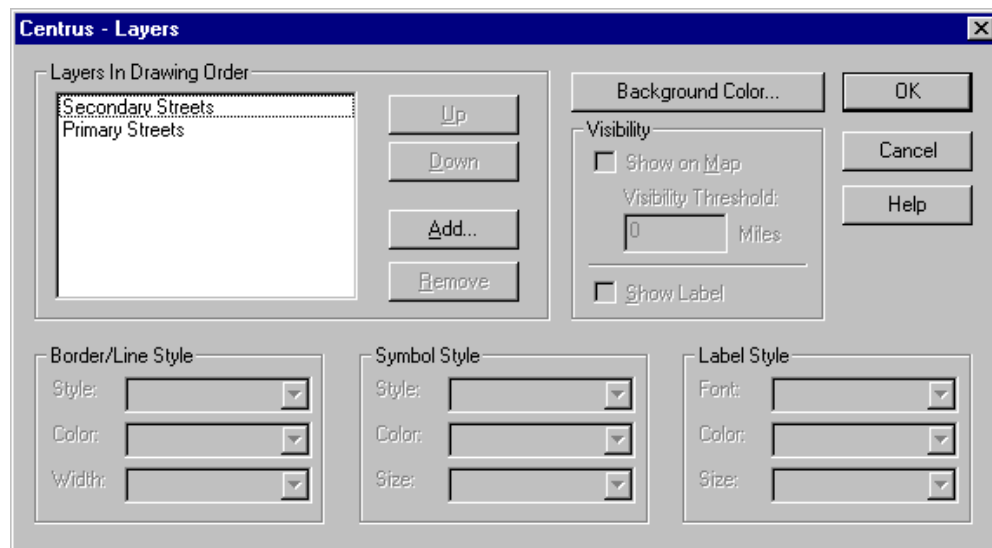
The Primary and Secondary Street layers control the display of the road network. Primary streets are major roads and highways. Each layer can have a different look, in order to customize the presentation of the map.

The Layers dialog box consists of several sections. At the top left is a list of layers. Layers are drawn in the order they are listed. You can change the order in which layers are drawn by moving layers up or down in the list. You can also add or remove layers.

Note: You can only remove layers that have been added through the Map Viewer. To add or remove layers from the Point-in-Polygon, Closest Site, or Geographic Determination modules, click the Point-in-Polygon, Closest Site, or Geographic Determination tab. Highlight the name of the layer you wish to remove, then click **Remove**.

Visibility options control whether – and how – a layer is drawn. The bottom portion of the dialog box has three sections for defining how lines, points, and labels appear on the map.

The dialog box shown below lists three layers: the two standard layers, plus the counties layer included with Centrus Desktop. Since layers are drawn in the order listed, Secondary Streets are drawn first, followed by Primary Streets and finally Counties. Note that streets are always drawn together, based upon the first street layer listed.



The Counties layer is selected, so all option settings display the current information about this layer.

Layers in Drawing Order

Up – This button moves the currently selected layer up one position towards the top of the list, causing the layer to be drawn earlier in the drawing sequence.

Down – This button moves the currently selected layer down one position towards the bottom of the list, causing the layer to be drawn later in the drawing sequence.

Add – This button allows you to add a .gsb file or other geocoded table to the list of layers, or import a .mif/.mid, .bna, or .shp file. Layers added in this manner are available only within the Map Viewer – they are not accessible by the Point-in-Polygon, Closest Site, or Geographic Determination modules.

Remove – This button removes the current layer from the list of layers. Primary Streets and Secondary Streets cannot be removed. You can use the **Visible** check box to “turn off” these layers if you do not want them drawn.

Background Color – This button allows you to change the background color of the map.

Note: You can only remove layers you have added within the Map Viewer. To add or remove layers from the Point-in-Polygon, Closest Point, or Geographic Determination modules select **Process > Options**. Then click the Point-in-Polygon, Closest Site, or Geographic Determination tab. Highlight the name of the layer you wish to remove, then click **Remove**.

Visibility

Show on Map – This check box controls whether a layer is displayed. If it is not checked, then no elements in the current layer are displayed.

Visibility Threshold – Each time you **Zoom Out**, the map is redrawn more slowly as the Map Viewer must draw more features. At very low levels of magnification, features appear as a solid area, and redrawing can be extremely slow. Setting a **Visibility Threshold** appropriate to your data lets you focus on the big picture. The Visibility Threshold is expressed in miles. The default setting is zero (no threshold).

Show Label – This check box controls whether the objects in the current layer are labeled. Primary and Secondary Streets cannot be labeled – use the **Info** tool in the Map window to get complete street segment information.

Layer Drawing Styles


Border/Line Style – These settings determine how lines are drawn in the current layer. Line style, color and width can all be specified.

Symbol Style – These settings determine how symbols are drawn in the current layer. Symbol style, color, and size can all be specified. Please note that symbols are drawn with the Wingdings font which comes with all versions of Windows. You may see characters other than those specified in this dialog box if the Wingdings font has been removed.

Label Style – These settings determine how labels are drawn in the current layer. Font, color, and size can all be specified.

Importing Layers

To import layers into the Map Viewer:

1. In the Map Viewer, click the Layers button . The Layers dialog box appears.
2. Click the **Add** button. The Add Type dialog box appears.
3. Choose **Add GSB File**, **Import Layer**, or **Add Table**, and then click **OK**. A file selection dialog box appears.
4. Click the **Browse** buttons to select the path and name of the file to import, and (if necessary) the object file (.GSB) to create. When importing layers or tables, you can determine which field is used as the object's identifier. Click the drop-down list box to select which **Field to use as identifier**.
5. Click the **Import** button to finish.

Note: Layers you import in this manner are available only to the Map Viewer. Use the Options dialog box to import layers for use with the Point-in-Polygon, Closest Site, or Geographic Determination modules.

Increasing Centrus Desktop Mapper Performance

When the Centrus Desktop mapper is asked to draw a region, it first determines which finance areas are located within the region. It then builds an index file for each of those finance areas, if one does not already exist. These files are named for their six-digit finance area and have a .gsx extension (e.g. 100134.gsx). Centrus Desktop typically builds and looks for these files in its “Data” subdirectory.

The Centrus Desktop mapper spends most of its time generating the finance area index files it needs. If you are doing a large amount of mapping, the mapping time may be greatly decreased by using the batchind.exe utility (installed with Centrus Desktop) to build all indexes for the U.S, and placing the resulting .gsx files in Centrus Desktop’s “Data” subdirectory. Centrus Desktop uses these .gsx files directly.

For information on executing batchind.exe, refer to the *Sagent Utilities Reference Manual*.

Shape Layers

Shape Layers allows you to create and export .gsb files, and to generate PMAs.

To access the Shape Layers function, select **Tools > Shape Layers**.

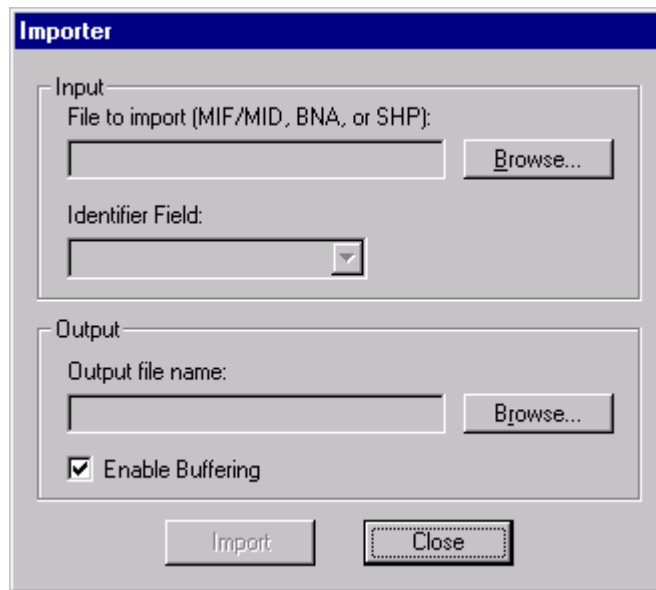
To generate a PMA, refer to the chapter, “Primary Market Area Function”.

Importing Geography Files

To import layers:

1. From the main Centrus Desktop menu, select **Tools > Shape Layers > Import**.

- Click the **Import** button. The **Import Layers** dialog box appears.



- Click the **Browse** buttons to select the path and name of the .mif/.mid, .bna, or .shp file to import and the object file (.gsb) to create.
- When importing, you can determine which field is used as the object's identifier. Click the drop-down list box to select which **Field to use as identifier**.
- Be sure that the **Enable Buffering** box is checked.
If this check box is not selected, you do not receive in border or in buffer polygon location and distance-to-border values.
- Click **Import** to finish.

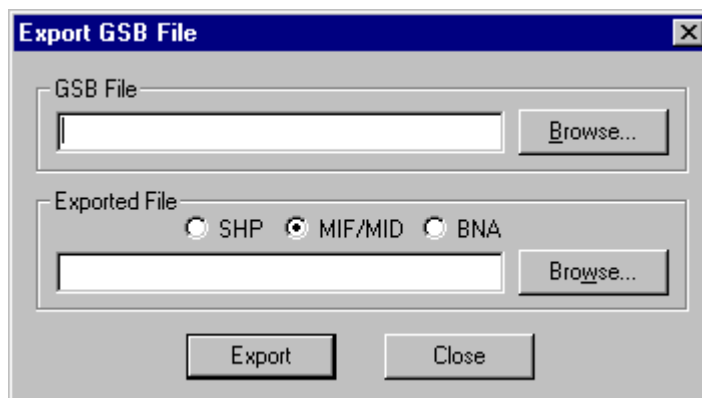
Note: While western hemisphere locations are correctly expressed as negative longitudes, some mapping and GIS applications still return longitudes as positive numbers. Be sure that the longitude data in your imported layer is consistent with that of your address file.

Exporting Geography Files

You can export geography files created in Centrus Desktop (GSB files) to ESRI's .bna or .shp format, or MapInfo's .mif/.mid format.

To export geographies:

1. Select **Tools > Shape Layers > Export**. The **Export GSB File** dialog box appears.



2. Specify the name of the GSB file you wish to export, and the name and type (**SHP**, **MIF/MID**, or **BNA**) of the exported file you wish to create, or click **Browse...** to search for the files.

Click the **Export** button to finish.

Chapter 5

File Specification

This chapter describes how to prepare and specify files used in processing. It includes the following information:

- File Types Used in Centrus Desktop
- Preparing Data Files for Processing
- Specifying Input and Output Tables

File Types Used in Centrus Desktop

Centrus Desktop natively supports a wide variety of database file formats, including:

- dBase III, IV, V and FoxPro 2.0, 2.5, 2.6 (*.dbf)
- Oracle
- SQL Server
- Excel 5.0 (95), Excel 97 (*.xls), and Excel 2000
- Access (*.mdb)
- FoxPro 3.0 (*.dft)
- Fixed-field and delimited ASCII text files

The Centrus Desktop Point-in-Polygon and Closest Site modules use Sagent's native .gsb format, and can import layers from .mif/.mid, .bna, and .shp file formats. The Closest Site module can also import layers from geocoded database tables.

Preparing Data Files for Processing

Supported input table fields are placed in the output table; fields that are potentially of an unsupported type causes Centrus Desktop to generate an warning message. For more information regarding supported input table field types, refer to the section “Specifying an Input Table” on page 87.

You can protect the data in the unsupported field in either of the following ways:

- Change all the fields in your file to text or numeric.
- Cut the supported fields out of the file and paste them into a new file. Then process the new file as normal. After the file is processed, append the unsupported fields back into the resulting file.

Access, dBASE, and FoxPro

Access, dBASE, and FoxPro files require no additional preparation. Field labels, however, should be no longer than ten characters, as labels exceeding that length are truncated. Duplicate labels are numbered to render them unique.

ASCII Text Files

ASCII text files may require a special “format file,” which can be created using the Centrus Data Formatter utility program. If you select an ASCII file for processing and Centrus Desktop finds no format file, the **Data Formatter** dialog box appears.

See “Centrus Data Formatter and ASCII Text Files” on page 187 for more information on creating format files.

Excel Files

When using Excel as an input source for Centrus Desktop, follow these guidelines:

- The first worksheet row should contain field names. If it does not, insert a row at the top of the worksheet and enter field names.
- Be sure to close the input and output files in Excel before specifying them in Centrus Desktop.
- Ensure the correct version of the workbook is selected for the input type. Trying to output to an incorrect version results in an error.

Specifying Input and Output Tables

Use the **Tables** tab to select the input file and (if desired) the output, reject, and report files.

Note: You cannot type file names directly into the file name boxes. Instead, click the **Browse** button to select the path and file names in file selection dialog boxes.

Specifying an Input Table

In the **Input Tables** section you specify the file name and path of an input file containing a table of addresses to be processed. By default, the file selection dialog box for input files shows files with the .dbf extension, or the extension of the file last opened. If you want to view other file types, select the appropriate option from the **Files of type** list box in the file selection dialog box. *Be sure to specify the correct file type and version!*

Note: The **Files of Type** list box contains the type **All Text Files (*.*)**. Selecting this type shows all files of all types. However, any file you select opens as a text file. This file type is designed to allow you to access text files that have extensions other than ".txt". Use this selection *only* to open ASCII text files.

Input table fields of the following types are placed in the output table: integer, short integer, long integer, unsigned integer, unsigned short, unsigned long, boolean, floating point, date, date time, and string. Input table fields of the following types generate a warning noting that they are to be converted to string: currency, time, decimal, and numeric (unlimited precision). All other types generate a warning noting that the data is not placed in the output table (no string conversion available).

Specifying an Output Table

Use the **Output Tables** box to tell Centrus Desktop where you want to send the output. Centrus Desktop can either write updated data back to the input file, or create a new output file. To create an output file, enter the name of the output file to create. (If the Output Tables section appears dimmed, be sure the *In-Place Update* option is not checked.)

Note: If a file already exists with the same name as the specified output file, that file is overwritten.

When you use an output file, you can set which records are written to the output file using the radio buttons listed below the output file name.

- **All** – This selection writes every record from the input file to the output file, even records that are not processed, standardized, or geocoded.
- **Only Records That Are** – This selection lets you specify the conditions that a record must meet before being written to an output file. Options are:
 - Parsed
 - Standardized
 - Geocoded
 - Demographic Coded
 - Point-in-Polygon Coded
 - Closest Site Coded

The output file has the same field names as the input file, along with any fields added during processing.

If an existing output table is being modified with unsupported data types (Update in Place is specified), processing halts.

Specifying a Reject File

If you choose the **Reject Records That Are Not** option on the **Tables** tab, you can specify a reject file to contain all the records that did not meet the specified conditions. To use a reject file, make sure the **Reject Records That Are Not** option is selected and conditions are specified. Then click **Browse** next to the **Rejects:** text box, and specify the path and file name of the reject file in the file selection dialog box.

The reject file has the same fields as the input and output files, along with any fields added during processing.

Report Files

The **Report Files** section of the **Tables** tab allows you to create two types of reports: an *Audit Report* and a *Log Report*.

The Audit report gives you a “snapshot” of the processing performed by Centrus Desktop. You select the frequency of the snapshot, such as every 1,000 records. For example, if you have a 50,000 record file, and choose a sample frequency of 1,000, you would get 50 records in your Audit report. If you select a frequency of 1, every record in the input file would be sent to the Audit file. (This is not recommended, as the file will be quite large.)

The Audit report lists the record as it was input, and as it was processed by Centrus Desktop, along with the Match and Location codes. This presents a quick view of what Centrus Desktop accomplished on a certain set of records.

The Log report lists extensive detail regarding the file and options used to geocode, as well as the results of the processing. This report can be displayed in Windows Notepad (if Notepad is available) by checking the **View log file when processing is completed** check box.

Examples of these reports can be found in “Audit and Log Report Samples” on page 198.

The screenshot shows the 'Tables' tab in the Centrus Desktop interface. The 'Report Files' section is expanded, showing options for generating an Audit report and a Log report.

Input Table: C:\Program Files\Centrus\Customer.dbf (Browse... button). ☐ In-Place Update. Fields... button.

Output Table(s): C:\Program Files\Centrus\customer_output.dbf (Browse... button).

☒ **Reject Records That Are Not:**

- ☐ Parsed
- ☐ Geocoded
- ☐ Point In Polygon Coded
- ☒ Standardized
- ☐ Demographic Coded
- ☐ Closest Site Coded

Rejects: C:\Program Files\Centrus\customer_nonstd_out.dbf (Browse... button)

Report Files:

- ☒ **Audit:** C:\Program Files\Centrus\Centrus2\audit.log (Browse... button). Sample frequency: 100 (spinner).
- Log:** D:\Data\Projects\SQA\Centrus2\Centrus.log (Browse... button).
- ☐ **View log file when processing is completed**

Specifying an Audit Report

To specify an audit report, be sure the check box next to the **Audit** text box is selected, then click the **Browse** button and specify the path and file name in the file selection dialog box.

Specifying a Log Report

To specify a log report, click the **Browse** button next to the **Log** text box, and specify the path and file name in the file selection dialog box.

Chapter 6

Primary Market Area Function

The Centrus Desktop Primary Market Area (PMA) function simply and quickly generates geographic market areas based on store locations, customer addresses, and other values.

If you have a location-dependent business, PMA can help you identify market overlap, areas of “cannibalizing,” and pockets of untapped opportunity. You can use your existing database with PMA to create a precise polygonal boundary file using the outermost address points of customers or prospects surrounding your store or office location.

To ensure that you’re working with the most compact trade area possible, you can specify parameters that tighten the size of the boundary. For instance, you can set parameters to include only the closest group of customers representing 80 percent of your sales.

Even if you do not have any customer data, PMA lets you create trade areas based on simple or non-overlapping rings. With PMA, you can also export geography files created in Centrus Desktop (GSB files) to ESRI’s .bna or .shp format, or MapInfo’s .mif/.mid format.

To start PMA, click the **Generate PMA** button , or from the menu bar select **Tools > Shape Layers > Generate PMA**.

This chapter includes the following information:

- Creating Store-Based PMAs
- Creating Customer-Based PMAs

Creating Store-Based PMAs

Store-based PMAs consist of one or more simple or non-overlapping rings drawn at distances that you specify from a store location. Both simple and non-overlapping rings may be inclusive (concentric circles) or exclusive (“donuts”). The resulting PMAs are saved as data *layers* in Centrus Desktop’s GSB format.

To create store-based PMAs, you must have a data file containing geocoded store locations with Latitude and Longitude fields and unique store identifiers. If you need to geocode the addresses in your files, use the Centrus Desktop Address Coding module for quick, accurate geocode assignments. For more details on geocoding, see “The Address Coding Module” on page 101.

To generate store-based PMAs:

1. Select **Tools > Shape Layers > Generate PMA**. The **PMA Generation - Specify Store Table** dialog box appears.

2. Specify the name of the geocoded **Store Table** you wish to use to generate PMAs.

Note: You cannot type file names directly into the text boxes. Instead, click the **Browse** button to select the path and file names in file selection dialog boxes.

- Specify the **Store Input Fields**. Centrus Desktop automatically fills in **Latitude** and **Longitude**, if possible. Check that these field names are correct, and select an **ID** field from the drop-down list box. Remember that the ID field must contain unique values.

The dialog box titled "PMA Generation - Specify Store Table" has a close button (X) in the top right. It contains a "Store Table" section with a text box showing "D:\Program Files\Centrus\Centrus2\Tutor\Stores.db" and a "Browse..." button. To the right is a checkbox labeled "Use Customer Points" which is unchecked. Below is the "Store Input Fields" section with three dropdown menus: "ID:" (showing "NAME"), "Longitude:" (showing "LONGITUDE"), and "Latitude:" (showing "LATITUDE"). At the bottom are four buttons: "< Back", "Next >", "Cancel", and "Help".

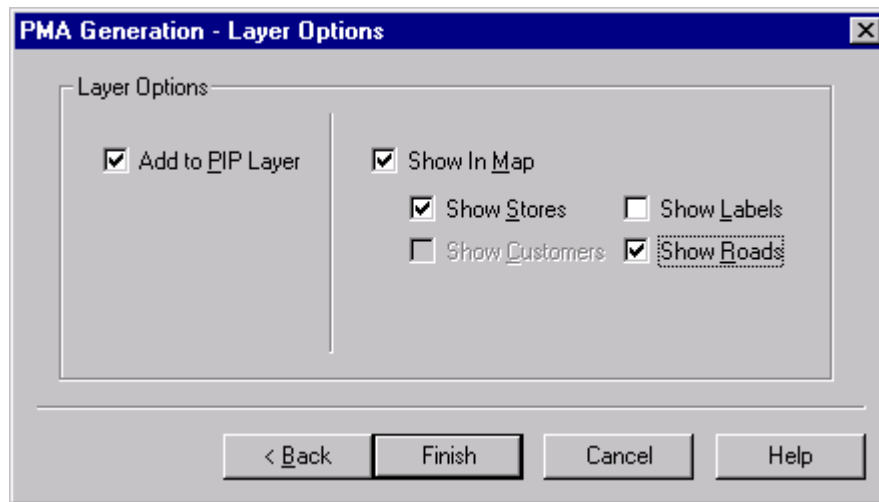
- Click **Next >** to continue. The **PMA Generation - Define Output Files** dialog box appears.

The dialog box titled "PMA Generation - Define Output Files" has a close button (X) in the top right. It features a "Type" section with three radio buttons: "Simple Rings" (selected), "Non-Overlapping Rings", and "Inclusive" (selected), followed by "Exclusive". Below is the "Outputs" section with a table-like structure. It has columns for "Radius" and "GSB File". There are three rows: "Primary:", "Secondary:", and "Tertiary:". The "Primary" row has a radius of "3" mi and a GSB file path "D:\Program Files\Centrus\Centr". Each row has a "Browse..." button next to the GSB file field. At the bottom are four buttons: "< Back", "Next >", "Cancel", and "Help".

- Specify the type of PMA(s) to generate. In the **Type** section, select **Simple Rings** or **Non-Overlapping Rings**, and **Inclusive** or **Exclusive**. In the **Outputs** section, specify the radius for each ring and a file name for the layer. You can specify up to three layers.

Note: Remember, you cannot type file names directly into the file name boxes.

6. Click the **Browse** button to select the path and file names in file selection dialog boxes. Click **Next >** to continue. The **PMA Generation - Layer Options** dialog box appears.



7. Specify what you want to do with the resulting PMA layer(s). The **Add to PIP Layer** option adds the PMA output files to the current list of data layers available to the Point-in-Polygon module. The **Show In Map** option starts the Map Viewer and displays the PMA(s) after the PMA is created. You can optionally **Show Stores**, **Show Labels**, and **Show Roads** in the map display.
8. Click **Finish**.

Creating Customer-Based PMAs

Customer-based PMAs are polygonal boundary files which are defined using the outermost address points of customers or prospects that surround one or more points — typically stores or office locations. The boundaries can reflect the distribution of customers, or be weighted to reflect some other variable, such as sales. The resulting PMAs are saved as data layers in Centrus Desktop's GSB format.

To create customer-based PMAs, you must have a data files containing geocoded store and customer locations with Latitude and Longitude fields and unique identifiers with common values in each table. If you need to geocode the addresses in your files, you can use the Centrus Desktop Address Coding module for quick, accurate geocode assignments. For more details on geocoding, see "The Address Coding Module" on page 101.

To generate customer-based PMAs:

1. Select **Tools > Shape Layers > Generate PMA**. The **PMA Generation - Specify Store Table** dialog box appears.

2. Select the **Use Customer Points** check box.
3. Specify the name of the geocoded **Store Table** you wish to use to generate PMAs.

Note: You cannot type file names directly into the file name boxes. Instead, click the **Browse** button to select the path and file names in file selection dialog boxes.

4. Specify the **Store Input Fields**. Centrus Desktop automatically fills in **Latitude** and **Longitude**, if possible. Check that these field names are correct, and select an **ID** field from the drop-down list box. Remember that the ID field must contain unique values and should correspond to values in the customer input table.
5. Click **Next >** to continue. The **PMA Generation - Specify Customer Table** dialog box appears.

6. Specify the name of the geocoded **Customer Table** you want to use to generate PMAs.
7. Specify the **Customer Input Fields**. Centrus Desktop automatically fills in **Latitude** and **Longitude**, if possible. Check that these field names are correct, and select an **ID** field from the drop-down list. The ID field must contain values that correspond to the ID field in the store input table.
8. Select a **Weight** field from the drop-down list, if desired. Click **Next >** to continue. The **PMA Generation - Define Output Files** dialog box appears.

PMA Generation - Define Output Files

Type: ☒ Inclusive ☐ Exclusive

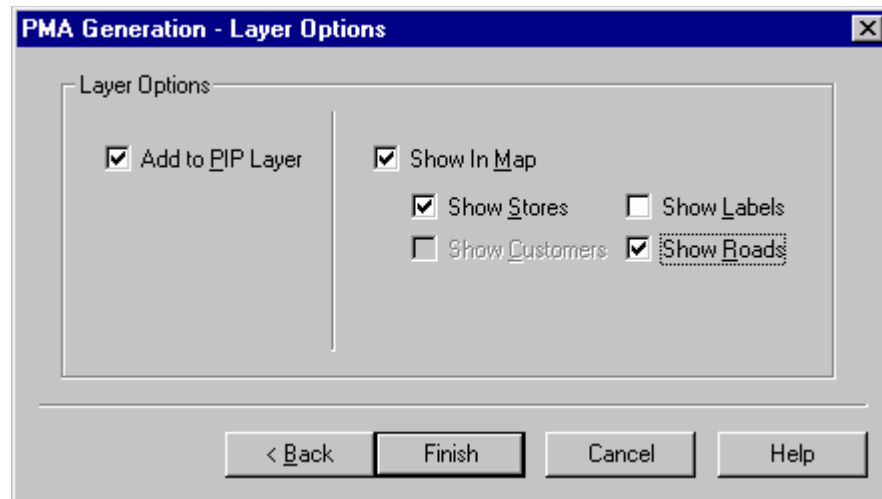
	Percent	GSB File	
Primary:	40 %	D:\Program Files\Centrus\Centr	Browse...
Secondary:	%		Browse...
Tertiary:	%		Browse...

< Back **Next >** Cancel Help

9. Specify the type of PMA(s) to generate. In the **Type** section, select **Inclusive** or **Exclusive**. In the **Outputs** section, specify the percentage of the value you would like to include for each PMA (for example, percent of customers, sales, etc.). Then specify a file name for each layer. You can specify up to three layers.

Note: Remember, you cannot type file names directly into the file name boxes. Click the **Browse** button to select the path and file names in file selection dialog boxes.

10. Click **Next >** to continue. The **PMA Generation - Layer Options** dialog box appears.



11. Specify what you want to do with the resulting PMA layer(s). The **Add to PIP Layer** option adds the PMA output files to the current list of data layers available to the Point-in-Polygon module. The **Show In Map** option starts the Map Viewer and displays the PMA(s) after the PMA is created. You can optionally **Show Stores**, **Show Labels**, **Show Customers**, and **Show Roads** in the map display.
12. Click **Finish**.

Chapter 7

The Address Coding Module

This chapter describes how to use the Address Coding Module to standardize and geocode U.S. and Canadian addresses. It includes the following information:

- About the Address Coding Module
- Specifying Address Input Fields
- Assigning Address Output Fields
- Available Address Data Elements

About the Address Coding Module

The Address Coding module is licensed by country, state, region, or nationally. The license file that is shipped with Centrus Desktop informs the system which areas are available for coding. If there are addresses outside of the coverage area in a file being processed, these addresses are flagged with a match code indicating an “out of coverage” address. Your licensed coverage area can be expanded by contacting your Sagent sales representative.

When an address file is processed, each address is standardized to USPS or CPC standards. If you want to “clean” your addresses and help ensure deliverability, you may choose to have the standardized output returned.

When standardizing U.S. addresses, the address is also geocoded to either the address or ZIP+4 level. To perform address standardization and geocoding, you can either use fields containing address and last line information (city, state, and ZIP), or simply give Centrus Desktop a list of fields. Centrus Desktop then finds and corrects the address information in those fields.

The Address Coding module compares the street addresses from the address file to the records in the USPS ZIP+4 Directory and the enhanced street network files. If the address is located in the USPS files, the address is standardized and a ZIP+4 (and all other USPS information) can be returned. If the address is also located within the street network files, Centrus Desktop can determine a very accurate latitude and longitude for the location. If the address was not found in the enhanced street network files, location and census information is then taken from the ZIP+4, ZIP+2, or ZIP Code centroid.

The overall standardization and geocoding rates are determined by the quality of the addresses in the file. Centrus Desktop can correct most minor misspellings as well as missing or incorrect directionals, street types, and ZIP Codes. However, if an address has an incorrect street number, or if the address contains a number of errors throughout, Centrus Desktop may not be able to make a successful match.

For best results, Sagent recommends that addresses be entered into address files following the USPS guidelines outlined in “Publication 28, Postal Addressing Standards.” This document is available free of charge from the USPS. For more information, contact the USPS National Customer Support Center in Memphis, TN at 1-800-238-3150.

When standardizing Canadian addresses, the Address Coding module compares the street addresses from the address file to the records in the CPC data files. Address-level geocoding is not available for Canadian addresses.

U.S. Address Processing

When you select **United States** on the **Process Options** dialog box, the Centrus Desktop Address Coding module is a complete address standardization and geocoding solution that you can use to enhance any database containing U.S. address information. The Address Coding module:

- Standardizes, corrects, and appends information about your addresses.
- Standardizes addresses to USPS standards. This reduces mailing costs and increases delivery speed by supplying correct USPS information. The Address Coding Module also identifies poorly formed or undeliverable addresses.
- Produces a USPS CASS report. This report allows even greater savings on postage for your mailings.
- Using address geocoding, assigns Latitude, Longitude, and Census ID information to the Block level.
- Using ZIP+4 centroid geocoding, assigns a geocode when address geocoding is unavailable. ZIP+4 centroids return Census ID information to the Block Group level.

When you select the **Automatic** option, Centrus Desktop determines if the address is U.S. or Canadian, processes it for you, and produces the associated report.

For more information on the USPS CASS program or bulk mail discounts, see “CASS Overview” on page 195.

Canadian Address Processing

When you select **Canada** on the **Process Options** dialog box, the Centrus Desktop Address Coding module is a complete address standardization and geocoding solution that you can use to enhance any database containing Canadian address information. The Address Coding module:

- Standardizes, corrects, and appends information about your addresses
- Produces a Canadian Post Corporation (CPC) SERP report
- Performs Postal Code Level geocoding

Note: To use this option, you must have a license. Contact your Sagent sales representative for more information.

When you select the **Automatic** option, Centrus Desktop determines if the address is U.S. or Canadian, processes it for you, and produces the associated report.

For more information on the CPC Address Accuracy Program and SERP, see “Canadian Address Accuracy Program” on page 203.

Z4 Change Processing Option

Note: The Us.gsl file contains the data required for Z4 Change Processing. This file must be located in the default search path.

Z4 Change Processing is an option that greatly enhances processing speed when you are using large databases that frequently require reprocessing. Z4 Change Processing uses USPS data that details which records have changed in the past 12 months. Centrus Desktop looks at the last date on which a particular list was processed, and compares each address to a Z4 Change Directory to see if it has changed in any way. Only changed addresses are reprocessed rather than the entire database—making processing speed three to five times faster than using the Address Coding Module alone.

Note: This function is not supported for Canadian addresses.

Address Line Two Preferences

Two-line addressing allows you to pass two different address lines to Centrus Desktop. Centrus Desktop scans these two lines and extracts and standardizes a two-line address, if possible. You can set Centrus Desktop to accept a P.O. Box or a street address as the preferred address. If no option is specified or the two lines contain similar information, line one is preferred.

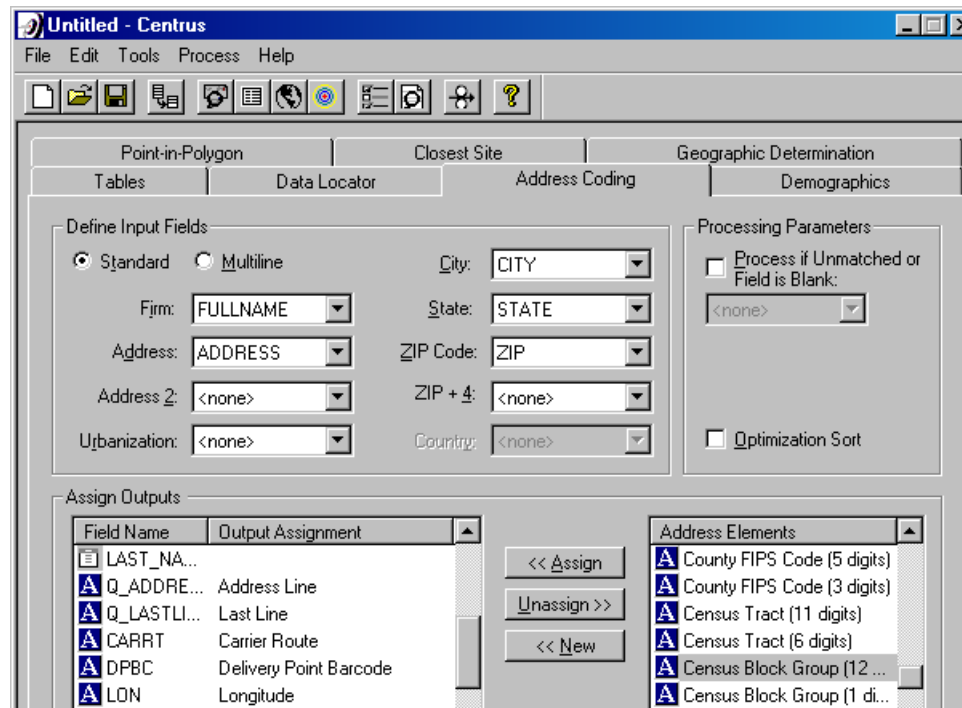
Note: This function is not supported for Canadian addresses.

Specifying Address Input Fields

Before using the Address Coding module, make sure you have:

- Set processing options for the module as described in “Setting Processing Options” on page 12
- Selected a file for processing as described in “File Specification” on page 85

Once you have selected a file to process, the input fields are filled in automatically, wherever possible, by looking at the input file’s field names. If any input fields were not filled in, or were not filled with the correct field, click the drop-down arrows to select from address file fields. If you wish to remove a field, select the special field **<none>** at the top of the list box of input field names.



Note: If you are using Z4 Change Processing, you must select address file fields for ZIP and ZIP+4.

Using Different Input Formats

In many files, the address information is contained in discrete fields that correspond to Firm, Address, City, State, ZIP, and possibly ZIP+4 or Country.

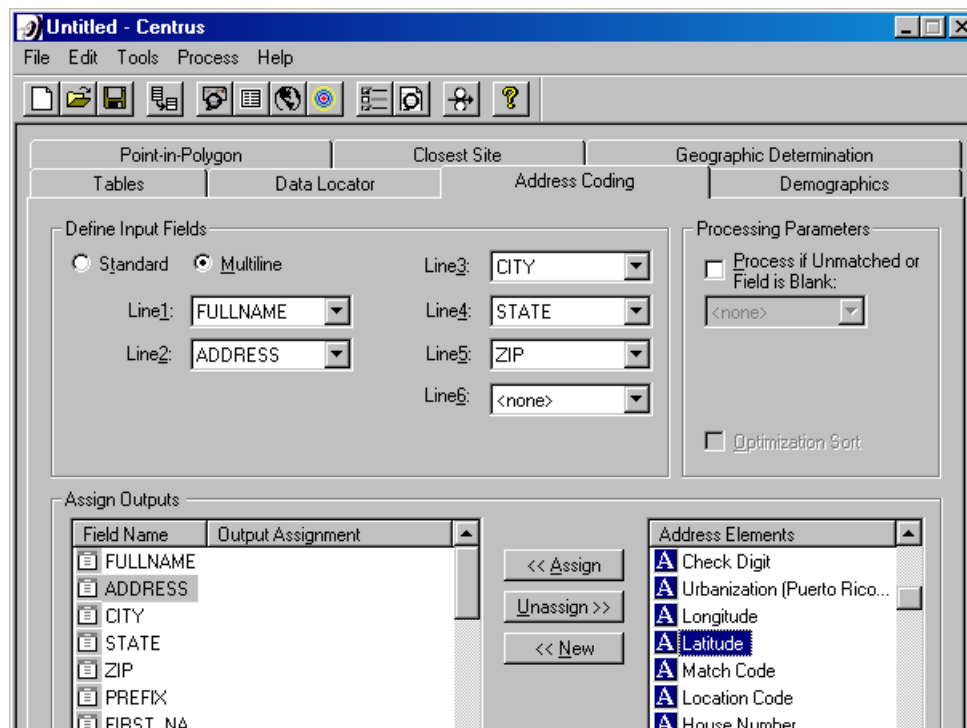
Some files, however, do not have discrete fields for last line information (City, State, and ZIP). In the Address Coding module, you can assign the field containing the Last Line information to the City input field.

If the file to be processed has multiple address lines or contains addresses in different fields, the Address Coding module provides a **Multiline Input** option.

Multiline Input

Note: This function is not supported for Canadian addresses.

If you select the **Multiline Input** check box, the field identifiers for the input Fields change to **Line 1** through **Line 6**, and the **Optimization Sort** check box disappears. When using this option, you may specify any six fields that Centrus Desktop should search through in order to find an address. Centrus Desktop can determine, for each record, which fields contain address information, and which fields do not. Fields containing no address information are ignored.



Fields should be specified in logical address order. Fields that could contain the street address information should be listed before fields that could contain last line information (city, state, and ZIP).

For example, Centrus Desktop can handle the following address:

Suite A
2900 Center Green Court
Boulder
CO

but not:

Boulder, CO
2900 Center Green Court, Ste A

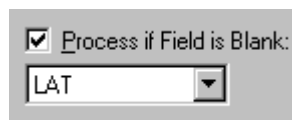
Note: When using the **Multiline Input** option, the Private Mail Box Designator and Private Mail Box Number output address elements are not available.

General Guidelines for Using Multiline Mode

- You can specify from two to six input fields. Note that the Address Coding module cannot detect the address and last line information when all of the information is in one field. If you need this feature, use the Data Locator module to parse your data into discrete fields.
- Firm names should appear before the street address line.
- Suite or apartment numbers can be listed on the street address line, above the street address line, or below the street address line. However, they must appear *before* the last line information.
- Last line information must come after the street address information. You can specify different input fields for City, State, and ZIP Code, but they should be listed in that order. The Address Coding module cannot detect a city name that is listed after a state.

Process if Unmatched or Field Is Blank

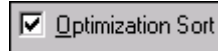
The **Process if Field is Blank** option lets you process only those records with missing data in a particular field, which you specify using the drop-down list. This is useful for “cleaning up” a previously processed database.

A screenshot of a software interface showing a checkbox labeled 'Process if Field is Blank:' which is checked. Below the checkbox is a text input field containing the text 'LAT' and a small downward-pointing arrow to its right, indicating a drop-down menu.

Note: This option is disabled when producing a CASS report—all records must be processed.

Optimization Sort

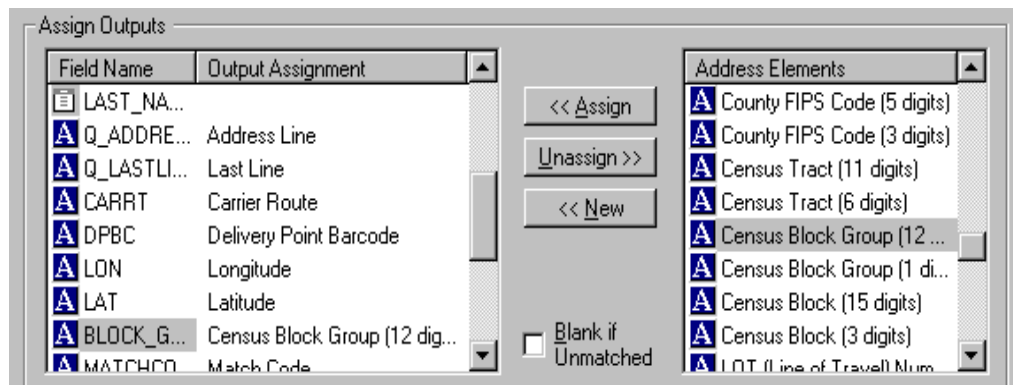
Selecting **Optimization Sort** causes the file to be processed in ZIP Code sequence. In most cases this significantly speeds processing. Optimization Sort requires a ZIP input field, and is limited to processing non-text files.



Assigning Address Output Fields

In the **Assign Outputs** section of the **Address Coding** dialog box you associate address data elements with output field names. By default, the list of field names is the same as the input fields in the **Define Input Fields** section of the dialog box. You can assign address data elements to existing field names, or create new fields to receive the data.

If you are using Z4 Change Processing, make sure that you select the **Z4ChangeResult** address element in this dialog box.



By default, outputs are ordered alphabetically by field name. Click **Output Assignment** to view outputs ordered alphabetically by output assignment.

Assigning Address Elements

To associate an address data element with an output field:

1. Select a field name by clicking it.
2. Select the desired address data element by clicking it.
3. Click the <<Assign button.

The selected address element appears in the **Output Assignment** column next to the field name to which it is assigned. An "A" icon appears before the address element, signifying that the data is being assigned by the Address Coding module.

Creating New Fields

To create a new field:

1. Select the desired address data element by clicking it.
2. Click the <<New button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected address element appears in the **Output Assignment** column next to the field name to which it is assigned. An “A” icon appears before the address element, signifying that the data is being assigned by the Address Coding module.

Unassign

If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Note: You can unassign an output assignment from any tab.

Standardization Defaults

Centrus Desktop can automatically assign the most commonly used address standardization data elements to their corresponding field names, or create new fields for these data elements. From the main Centrus Desktop menu, select **Edit**, then **Standardization Defaults**. Select **Existing Fields** to have Centrus Desktop try to assign the standardization defaults to your existing fields. Select **New Fields** to have Centrus Desktop create new fields for any standardization defaults you have not yet assigned. The standardization default data elements are:

- Firm Name
- Address Line
- Last Line
- Delivery Point Barcode
- Check Digit
- Carrier Route
- Urbanization (Puerto Rico)
- Match Code

Geocode Defaults

Note: This function is not supported for Canadian addresses.

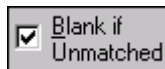
Centrus Desktop can automatically assign the most commonly used geocode data elements to their corresponding field names, or create new fields for these data elements. From the main Centrus Desktop menu, select **Edit**, then **Geocode Defaults**. Select **Existing Fields** to have Centrus Desktop try to assign the standardization defaults to your existing fields. Select **New Fields** to have Centrus Desktop create new fields for any geocode defaults you have not yet assigned. The geocode default data elements are:

- Latitude
- Longitude
- Match Code
- Location Code
- Census Block Group (12 Digits)

Blank if Unmatched

If you check the **Blank if Unmatched** check box, *all* output fields defined within the Address Coding module is cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Address Coding elements to new fields rather than input fields, or be sure the **Blank if Unmatched** check box is not selected.

Note: Do not use this option if you wish to retain unmatched input addresses.



Available Address Data Elements

This section details all information about the address data available within the Address Coding module. Each address element is listed, and the maximum length of the information returned is given in parentheses. For example, Firm Name (40) indicates "Firm Name" data, with a maximum length of 40 characters. If the assigned field has less than 40 characters, the firm information may be truncated.

Address Element (size)	Explanation
Address Line (60)*	Returns the full address line (e.g., 1920 MAIN ST W APT 12).
Address Line 2 (60)	Returns the second address line (e.g., Suite 200).
Carrier Route (4)	Returns Carrier Route ID number.
Carrier Route Sortation (1)	Indicates whether a discount is provided for letter-sized carrier route sorted mail in current ZIP Code.
Census Block (15)	Returns the full block code (e.g., 080130122052031).
Census Block (3)	Returns the block code only (e.g., 031).
Census Block Group (1)	Returns the Block Group code only (e.g., 2).
Census Block Group (12)	Returns the full Block Group code (e.g., 080130122052).
Census Tract (11)	Returns the full Census Tract code (e.g., 08013012205).
Census Tract (6)	Returns the Census Tract code only (e.g., 012205).
Check Digit (1)	Used with the DPBC to ensure that the bar code printed on the mailing piece is correct.
City Delivery (1)	Indicates whether a Post Office has city-delivery carrier routes.
City Name (28)*	Returns the valid USPS city name (e.g., BOULDER).
City State Record Name (28)	Returns the city name for the matched address from the City State record.
CMSA Name (81)	Returns the Consolidated Metropolitan Statistical Area name.
CMSA Number (4)	Returns the Consolidated Metropolitan Statistical Area number.
Country (50)*	Returns the country name.
County FIPS Code (3)	Returns the county code only (e.g., 013).

Address Element (size)	Explanation
County FIPS Code (5)	Returns the full state and county code (e.g., 08013).
County Name (44)	Returns the county name.
Default Match (1)	Returns the following values: Y = Either High-Rise Default or Rural Route Default returned Y. Blank = Both High-Rise Default and Rural Route Default returned N or Blank
Delivery Point Barcode (2)	This two-digit field, when appended to the end of the 9-digit ZIP+4 Code, creates the Delivery Point Bar Code, which is then printed in the address section on the mailing piece to assist with automated sorting.
EWS Match (1)	Returns a letter indicating whether the record matched to EWS data: Y = The match was denied because it matched to EWS data Blank = Input record did not match to EWS data
Extra Line 1-6 (103)	Used to capture data contained in unassigned address lines.
Extra/Mail Stop (60)	Returns address information appearing after mail stop designator words: MSC, MS, MAILSTOP, MAIL STOP, ATTN, ATTENTION.
Firm Name (40)*	Returns the firm name, as known by the USPS or CPC, or as entered. Firm Name is blank if the USPS or CPC does not know what firm is at that location.
High Rise Default (1)	Returns the following values: N = Matched to an exact high rise record or a street record Y = An exact record was not found. Matched to the USPS default high rise record or a street record. The input address should be checked for accuracy and completeness. Blank = The flag does not apply to the input address (e.g., P.O. Boxes and General Delivery addresses) or no match was found
House Number (11)*	Returns the house number in the address (e.g., 123).
Intersection Flag (1)	Returns T if an intersection match was made or F if a normal address match was made.

Address Element (size)	Explanation
LACS Status (1)	Locatable Address Conversion Service Status Indicator: L = old (usually rural route) address which has been converted for the LACS system blank = not applicable
Last Line (60)*	Returns the full last line (e.g., BOULDER CO 80301-1234).
Latitude (11)	Returns the latitude coordinate in decimal degrees to 6 decimal places (e.g., 123.234234).
Location Code (4)	Reports the locational accuracy of the match. See "Location Codes" on page 160 for details.
Longitude (11)	Returns the longitude coordinate in decimal degrees to 6 decimal places (e.g., 123.234234). The number is positive or negative depending on the setting of the Negate Longitudes check box in the Options dialog box.
LOT (Line of Travel) Number (4)	The numeric LOT code used for presort.
LOT Direction Flag (1)	LOT direction: A = Ascending D = Descending
Match Code (4)*	Details which components of an address were modified if a match was successful. If a match was not successful, explains why the match could not be made. See "Match Codes" on page 155 for details.
MSA Name (66)	Returns the Metropolitan Statistical Area name.
MSA Number (4)	Returns the Metropolitan Statistical Area FIPS number.
Postal Code (10)	Returns the Canadian postal code.
Preferred City Name (28)	Returns the preferred city name for the output ZIP Code of the matched address.
Primary Post-Directional (2)	Returns the post-directional (e.g., E).
Primary Pre-Directional (2)	Returns the pre-directional (e.g., NW).
Primary Street Name (40)*	Returns the street name (e.g., MAIN).
Primary Street Suffix (4)*	Returns the street suffix, or type (e.g., ST).
Private Mail Box (PMB) Designator (5)	Returns the private mail box designator.
Private Mail Box (PMB) Number (9)	Returns the private mail box number.

Address Element (size)	Explanation
Range Record Type (1)	For non-intersection matches, returns a letter denoting the USPS Range type, where: G = General delivery record H = High-rise record F = Firm record S = Street record P = P.O. Box record R = Rural route/highway contract record
Rural Route Default (1)	Returns the following values: N = Matched to an exact rural route record Y = An exact record was not found. Matched to the USPS default rural route record. The input address should be checked for accuracy and completeness. Blank = The flag does not apply to the input address (for example, PO Boxes and General Delivery addresses) or no match was found
Second Post-Directional (2)	Returns the second post directional (e.g., W).
Second Pre-Directional (2)	Returns the second pre-directional (e.g., N).
Second Street Name (40)	Returns the second street name (e.g., PEARL).
Second Street Suffix (4)	Returns the second street suffix or type (e.g., ST).
State Abbreviation (2)*	Returns the 2-letter state (or Province) abbreviation (e.g., CO).
State FIPS Code (2)	The state code (e.g., 08).
Unit Designator (4)*	Returns the unit designator (e.g., STE).
Unit Designator 2 (4)	Returns the second unit designator parsed from the address line. This element is supported only in CASS mode.
Unit Number (11)*	Returns the unit or apartment number (e.g., 2A).
Unit Number 2 (11)	Returns the second unit number parsed from the address line. This element is supported only in CASS mode.
Urbanization (30)	Returns the urbanization code for the address. Used for Puerto Rican addresses only.

Address Element (size)	Explanation
Z4ChangeResult (2)	Returns the following values: N = There have been no changes. Your address records match the current USPS Z4 data. Y = There have been changes to current USPS Z4 data. Re-standardize your address records.
ZIP Classification (1)	Describes type of area that a 5-digit ZIP Code serves: M = military ZIP Code P = P.O. Boxes Only U = unique ZIP Code (single organization) blank = standard ZIP Code
ZIP Code (5)	Returns the ZIP Code (e.g., 80301).
ZIP Facility (1)	Returns the USPS City State Name Facility Code: A=Airport Mail Facility (AMF) B=Branch C=Community Post Office (CPO) D=Area Distribution Center (ADC) E=Sectional Center Facility (SCF) F=Delivery Distribution Center (DDC) G=General Mail Facility (GMF) K=Bulk Mail Center (BMC) M=Money Order Unit N=Non-postal community name P=Post Office S=Station U=Urbanization
ZIP+4 Extension (4)	Returns the +4 code (e.g., 1234).
ZIP10 (10)	Returns the complete ZIP+4 with hyphen (e.g., 80301-1234).
ZIP9 (9)	Returns the complete ZIP+4 (e.g., 803011234).

* Address elements for U.S. and Canadian addresses.

Latitude and Longitude must be used together, or not at all.

In an xBASE file, if the field for Latitude and Longitude are numeric and have 0 decimal places, Latitude and Longitude are returned in millionths of degrees. ASCII files always receive decimal degrees.

The postal bar code printed on a mailing label represents the full nine-digit ZIP Code with the +4 extension, plus the two-digit Delivery Point Barcode, plus the one-digit Check Digit. The resulting 12-digit bar code can be printed using a Postnet bar code font.

Chapter 8

The Data Locator Module

This chapter describes how to use the Data Locator Module to analyze databases, identify and extract name elements, and add a variety of name-related information. It includes the following information:

- About the Data Locator Module
- Specifying Data Locator Input Fields
- Assigning Data Locator Output Fields
- Tips for Using Data Locator
- Available Data Locator Elements

About the Data Locator Module

The Data Locator Module uses Sagent's proprietary name parsing technology to analyze databases, identify and extract name elements, and add a variety of name-related information. Data Locator can identify and correctly process first and last names, nicknames, and more. Using Data Locator, you can analyze names and assign probable gender and ethnicity. You can also add gender-appropriate salutations in a variety of styles (formal, informal, casual, family). With Data Locator, you can:

- Identify up to two people in a record.
- Assign name elements (first name, last name, middle name, title or prefix) to their proper fields.
- Format name order and capitalization preferences.
- Identify nicknames and aliases—useful for identifying duplicate records.

- Extract company name and job title (if available) from records.
- Assign gender according to variable confidence levels.
- Analyze ethnicity using first names, last names, or both.

How Data Locator Works

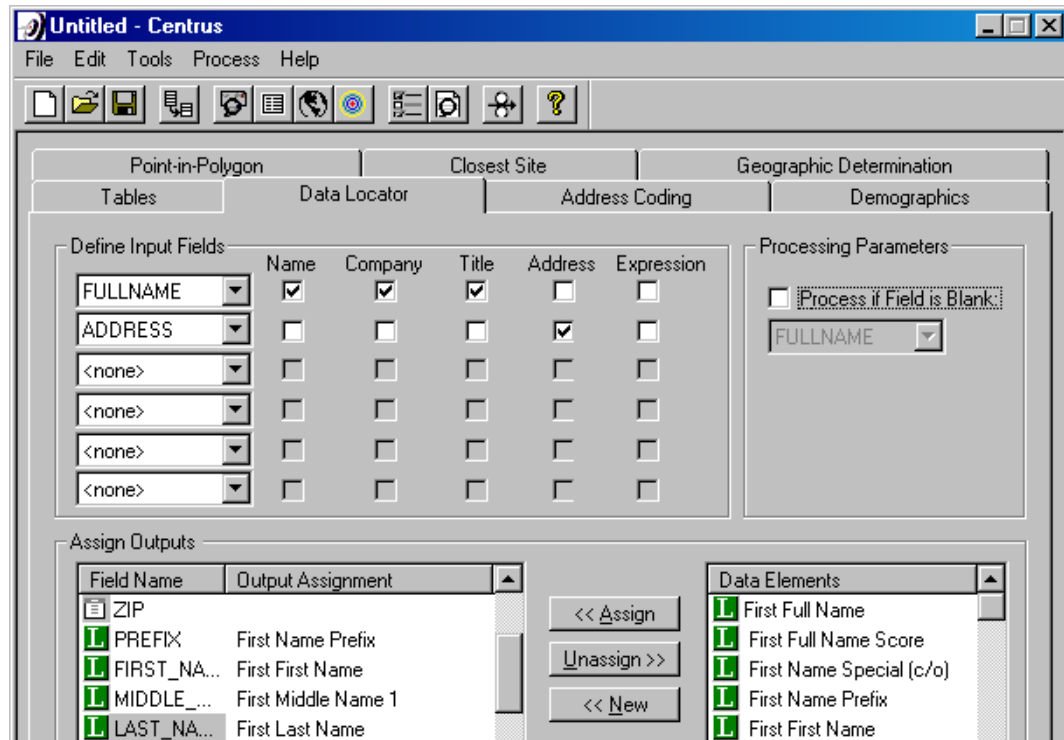
Data Locator examines your input data fields and searches for the components you identify in the inputs section of the Data Locator tab. If you specify more than one component for an input field, Data Locator matches each of the requested components against the input data and assigns a score. The highest scoring component is returned as the result.

For Names, Addresses, Companies, and Titles you have the option of outputting the result score. This is a number between 0 and 100 which gives a “confidence level” for the match. A score of 0 is poor; 100 is excellent.

Specifying Data Locator Input Fields

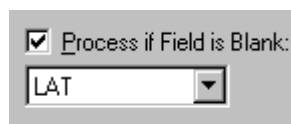
Once you have specified an input file in the **Tables** tab, you can select the input fields you want to analyze with the Data Locator module. Click the arrows next to the list boxes to see available input fields, then click on the ones you wish to select. You can analyze up to six fields.

Next, choose the type of information you wish to extract from each field. The five general categories are Name, Address, Company, Title, and Expression. (*Expression* is a special user-defined category. For more information, see “Customizing the Data Locator Module” on page 169.) Select or clear the appropriate check boxes by clicking them. You may select as many categories as you like.



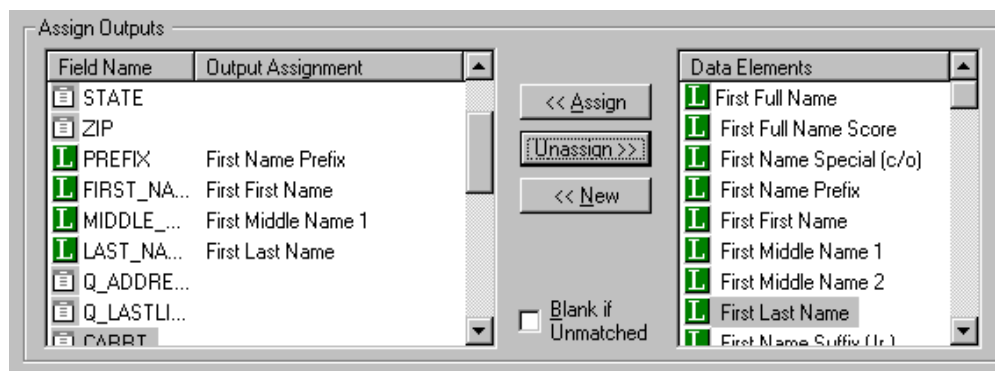
Process if Field Is Blank

The **Process if Field is Blank** option lets you process only those records with missing data in a particular field, which you specify using the drop-down list. This is useful for “cleaning up” a previously processed database.



Assigning Data Locator Output Fields

The **Assign Outputs** section of the **Data Locator** dialog box is where you associate Data Locator elements with output field names. By default, the list of field Names is the same as the input fields in the **Define Input Fields** section of the dialog box. You can assign Data Locator elements to existing field names, or create new fields to receive the data.



Assigning Field Names

To associate a data element with an output field:

1. Select a field name by clicking it.
2. Select the desired data element by clicking it.
3. Click the **<<Assign** button.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. An "L" icon appears before the data element, signifying that the data is being assigned by the Data Locator module.

Creating New Fields

To create a new field:

1. Select the desired data element by clicking it.
2. Click the **<<New** button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. An "L" icon appears before the data element, signifying that the data is being assigned by the Data Locator module.

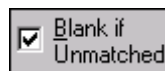
Unassign

If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Blank if Unmatched

If you select the **Blank if Unmatched** check box, *all* output fields defined within the Data Locator module are cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Data Locator elements to new fields rather than input fields, or be sure the **Blank if Unmatched** check box is cleared.

Note: Do not use this option if you wish to retain unmatched input addresses.



Tips for Using Data Locator

Decide Which Information You Want to Extract

Before you begin setting up a Data Locator task, decide which information you want to extract from your data. Data Locator can identify more than 50 different data elements. (See “Available Data Locator Elements” on page 121 for the complete list.) If you are especially interested in address components, you should use the Address Coding Module to standardize your list *before* processing it with the Data Locator. Data Locator matches standardized addresses but may give inaccurate results for poorly-formed addresses.

Define a Task to Extract the Information

Examine your input file. If the input file is large, you may want to create a small test file containing a subset of the data. List your fields, describing what you think each field contains. Your list might look something like this:

Field1	names or companies
Field2	names, companies, or titles
Field3	companies, titles, or addresses
Field4	titles or addresses
Field5	addresses

Using your list as a guide, set up input fields, then assign the desired outputs in the outputs section. (You may wish to output scores to aid in evaluating the results.) Then process your data.

Evaluate Results and Redefine the Task

If you are processing interactively, you can examine results as you process. If you processed in batch mode, open your output file to look at the results.

Ambiguous or incorrect results can often be avoided by changing the components being requested. For example, *Smith Barney* is identified as a company if *Company* is the only requested component for a field. But if both *Name* and *Company* are requested, *Smith Barney* scores higher as a name than as a company. *Smith Barney Inc.* always scores highest as a company, regardless of the components requested.

Similarly, requesting both Title and Company components can produce ambiguous results. *Quality Assurance Specialist* scores highest as a title, but *Quality Assurance Specialists* scores highest as a company. Each has a fairly high score for both company and title.

You may find it useful to request a component even if you do not want it as an output field. For example, you might have a field which could contain names, company names, or addresses. You want to extract only company names from the field. If the field contains the address “11 West 42nd Street” and you’ve asked Data Locator to find Name, Company, and Address components, Address scores higher than Name and Company. However, if you requested only Company components, the field scores highest as a company. Over-specifying components avoids the problem of mistaken identity.

Using the Custom Dictionary

If you find that one or more words are often incorrectly assigned, you may want to create a custom dictionary file. Once you specify the name and location of this file, Data Locator adds your custom dictionary entries to its own list of terms and parses input accordingly. For example, the word *Quality* appears frequently in both Company names and Titles. If you add *Quality* to the custom dictionary as a Company word, *Quality Consultants* is likely to be identified as a company. If *Quality* is added as a Title word, *Quality Consultants* is likely to be identified as a title. For more information, see “Custom Dictionaries in Data Locator” on page 172.

Using Multiple Passes and Process if Field Is Blank

You may find that several passes over a database are required to get the answers you desire. For example, you might find when reviewing your initial results that many of the blank Title fields contain the word *Mgmnt* in the corresponding input field. You could use the custom dictionary to define *Mgmnt* as a Title word, then process the original output file using the custom dictionary. You’d set **Process if Field is Blank** to **Title** and assign only the title field as output. Data Locator would update the Title field with additional results while retaining the initial results for other fields.

Available Data Locator Elements

The following table describes Data Locator elements.

Field Name	Explanation
First Full Name Score	Accuracy estimate for identification of the first full name.
First Name Special (c/o)	Extra information associated with first full name. This can include the words ATTN, C/O, DEAR, ESTATE, FAMILY, HONORABLE, HUSBAND, SPOUSE, THE, TRUST, WIFE.
First Name Prefix	Mr., Ms., Mrs.
First First Name	First name of first person.
First Middle Name 1	First middle name of first person.
First Middle Name 2	Second middle name of first person.
First Last Name	Last name of first person.
First Name Suffix (Jr.)	Name suffix of first person.
First Formal Greeting	Formal greeting for first person.
First Informal Greeting	Informal greeting for first person.
First Casual Greeting	Casual greeting for first person.
First All Aliases	List of all aliases and diminutives for first name of first person.
First First Name Gender	Gender of first person, based on first name.
First Middle Name Gender	Gender of first person, based on middle name.
First First Name Ethnicity	Ethnicity of first person, based on first name.
First Name Ethnicity	Ethnicity of first person, based on first name (next best estimate).
First Next Best First Last Name Ethnicity	Ethnicity of first person, based on last name.
First Next Best Last Name Ethnicity	Ethnicity of first person, based on last name (next best estimate).
First Person Gender	Gender of first person, all factors.
Second Full Name	Full name of the second person identified.
Second Full Name Score	Accuracy estimate for identification of the second full name.

Field Name	Explanation
Second Name Special (c/o)	Extra information associated with second full name. This can include the words ATTN, C/O, DEAR, ESTATE, FAMILY, HONORABLE, HUSBAND, SPOUSE, THE, TRUST, WIFE.
Second Name Prefix	Mr., Ms, Mrs.
Second First Name	First name of second person.
Second Middle Name 1	First middle name of second person.
Second Middle Name 2	Second middle name of second person.
Second Last Name	Last name of second person.
Second Name Suffix (Jr.)	Name suffix of second person
Second Formal Greeting	Formal greeting for second person.
Second Informal Greeting	Informal greeting for second person.
Second Casual Greeting	Casual greeting for second person.
Second All Aliases	List of all aliases and diminutives for first name of second person.
Second First Name Gender	Gender of second person, based on first name.
Second Middle Name Gender	Gender of second person, based on middle name.
Second First Name Ethnicity	Ethnicity of second person, based on first name.
Second Next Best First Name Ethnicity	Ethnicity of second person, based on first name (next best estimate).
Second Last Name Ethnicity	Ethnicity of second person, based on last name.
Second Next Best Last Name Ethnicity	Ethnicity of second person, based on last name (next best estimate).
Second Person Gender	Gender of second person, all factors.
First Company Name	First company name identified.
First Company Score	Accuracy estimate for identification of the first company name.
Second Company Name	Second company name identified.
Second Company Score	Accuracy estimate for identification of the second company name.
First Title	Title of the first person identified.
First Title Score	Accuracy estimate for identification of the first title.

Field Name	Explanation
Second Title	Title of the second person identified.
Second Title Score	Accuracy estimate for identification of the second title.
First Address	First address identified.
First Address Score	Accuracy estimate for identification of the first address.
First Unit	First unit identified.
First Lastline	First lastline identified.
First Lastline Score	Accuracy estimate for identification of the first lastline.
First City	First city identified.
First State	First state identified.
First ZIP	First ZIP identified.
Second Unit	Second unit identified.
Second Address	Second address identified.
Second Lastline	Second lastline identified.
Second Lastline Score	Accuracy estimate for identification of the second lastline.
Second City	Second city identified.
Second State	Second state identified.
Second ZIP	Second ZIP identified.
Expression	Regular expression to be matched
Extra1	"Leftovers"-unmatched strings.
Extra2	"Leftovers"-unmatched strings.

Chapter 9

The Demographic Coding Module

Corporations now use demographic information to target their marketing campaigns, forecast sales, determine market penetration, and assess market potential. Demographic coding of existing databases is an important part of this process. Sagent's Demographic Coding module allows you to append valuable demographics to your own databases.

This chapter describes how to use the Demographic Coding Module and includes the following information:

- Specifying Demographic Input Fields
- Assigning Demographic Output Fields
- Available Demographics Elements

Specifying Demographic Input Fields

Once you have selected a file to process, the input fields are filled in automatically, if possible, by looking at the input file's field names and any fields assigned in the Address Coding module. Only fields required for demographic coding, such as **Block Group**, are enabled. If any required input fields were not filled in, or were not filled with the correct field, click the drop-down arrows to select the correct field. If you wish to remove a field, select the special field **<none>** at the top of the list box of input field names.

Process if Field Is Blank

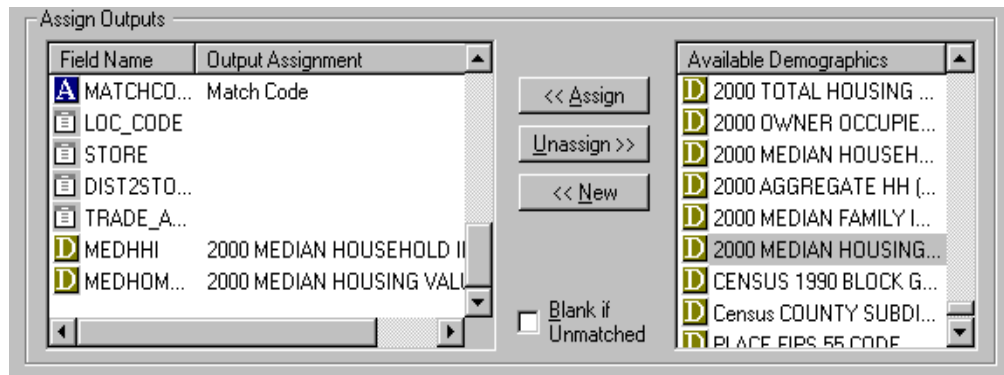
The **Process if Field is Blank** option lets you process only those records with missing data in a particular field, which you specify using the drop-down list. This is useful for “cleaning up” a previously processed database.

Assigning Demographic Output Fields

The **Assign Outputs** section of the **Demographics** dialog box is where you associate address data elements with output field names. By default, the list of field names is the same as the input fields in the **Define Input Fields** section of the dialog box. You can assign demographics to existing field names, or create new fields to receive the data.

The available output data elements are dependent on which demographics data files you have licensed. All users can access the 2000 Census data, contained in the file Census2k.dld (if installed). If no data elements are listed under **Available Outputs**, be sure that the file Census2k.dld is in the

default path specified in the **Default Data Search Path(s)** dialog box. You may need to insert one of the *Data Products Suite* CDs, or copy the file from CD to the Centrus Desktop directory on your hard drive. (See “Specifying the Search Path for the Data Files” on page 7 for more details.)



Assigning Field Names

To associate a data element with an output field:

1. Select a field name by clicking it.
2. Select the desired demographic data element by clicking it.
3. Click the **<<Assign** button.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “D” icon appears before the data element, signifying that the data is being assigned by the Demographics module.

Creating New Fields

To create a new field:

1. Select the desired data element by clicking it.
2. Click the **<<New** button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “D” icon appears before the data element, signifying that the data is being assigned by the Demographics module.

Unassign

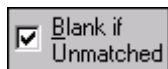
If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Note: You can unassign a field from any tab.

Blank if Unmatched

If you select the **Blank if Unmatched** check box, *all* output fields defined within the Demographics module are cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Demographics elements to new fields rather than input fields, or be sure the **Blank if Unmatched** box is not selected.

Note: Do not use this option if you wish to retain unmatched input addresses.



Available Demographics Elements

All Centrus Desktop users can access 2000 Census data using the following data elements.

- 2000 Total Population
- 2000 Total Female Population
- 2000 Households
- 2000 Median Age
- 2000 Median Age Female
- 2000 Median Age Male
- 2000 Total Housing Units
- 2000 Families
- 2000 Average Household Size
- 2000 Aggregate Household (\$000s)
- 2000 Owner Occupied Housing Units
- 2000 Median Household Income
- 2000 Median Family Income
- 2000 Median Housing Value
- 2000 Population in Group Quarters
- 2000 % White Population
- 2000 % Black Population
- 2000 % Asian/Pacific Islander Population
- 2000 % American Indian/Eskimo Population
- 2000 Total Households with Children under 18
- 2000 Total Households with Adults over 60
- 2000 Householder Age 15-24
- 2000 Householder Age 25-34
- 2000 Householder Age 35-44
- 2000 Householder Age 45-54
- 2000 Householder Age 55-64
- 2000 Householder Age 65-74
- 2000 Householder Age 75-84
- 2000 Householder Age 85+
- Census 1990 Block Group
- Census County Subdivision FIPS 55 Code
- Place FIPS 55 Code

Note: If you have licensed premium demographics, additional data elements are available. See “Premium Demographics” on page 175 for more information.

Chapter 10

The Point-in-Polygon Module

Point-in-Polygon analysis lets you accurately determine which areas a given point falls within. Retail, insurance, and telecommunication industries use Point-in-Polygon to identify key markets, insurance risk areas, and telephone service areas.

This chapter includes information about:

- Object Files and Layers
- Point-in-Polygon Analysis
- Specifying Point-in-Polygon Input Fields
- Assigning Point-in-Polygon Output Fields

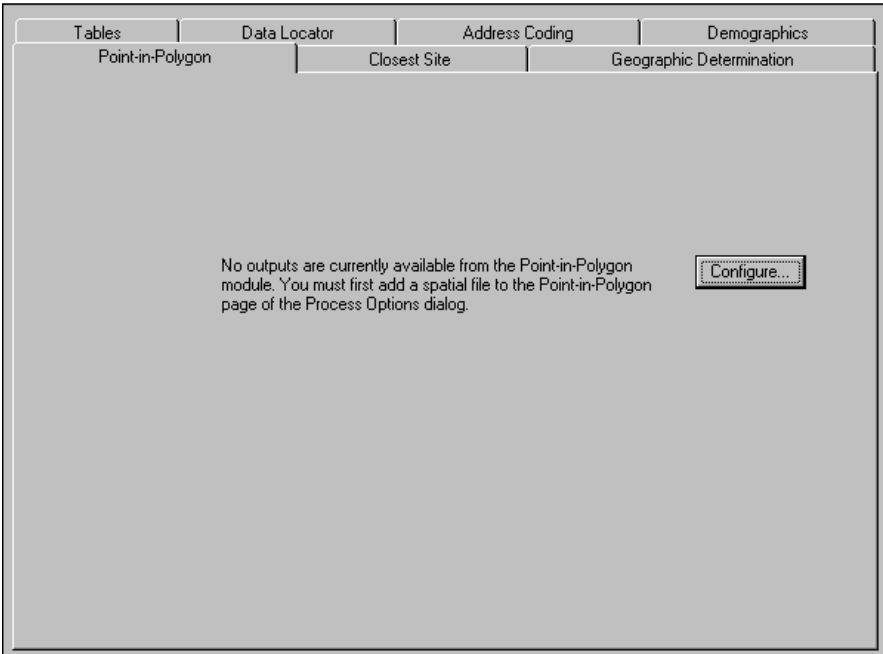
Object Files and Layers

The Point-in-Polygon module requires geocoded addresses with Latitude/Longitude coordinates as input. If you need to geocode the addresses in your files, we recommend that you use the Centrus Desktop Address Coding module, for the quickest, most accurate geocode assignments. You can even geocode and perform spatial analyses in the same task.

The available outputs are determined by the contents of the “object” files or *layers* you have specified in the **Process Options** dialog box. Geographic features (points, lines, or polygons) are imported into the native Centrus Desktop format which allows very fast searches. To get your data into our .gsb format, we support MapInfo’s .mif/.mid format, as well as ESRI’s Atlas GIS .bna and .shp formats.

For more information about specifying or importing layers, see “Point-in-Polygon Tab” on page 23.

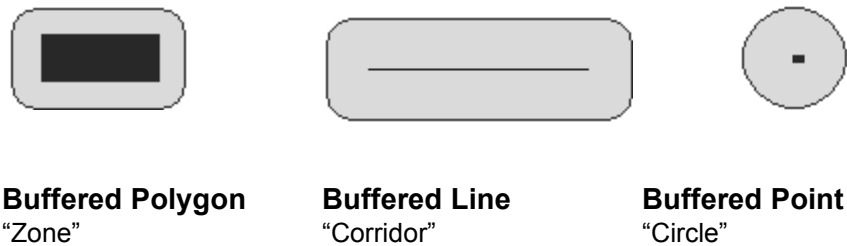
Note: The **Point-in-Polygon** tab displays only a **Configure...** button until a .gsb file is specified in the Point-in-Polygon page of the **Process Options** dialog box.



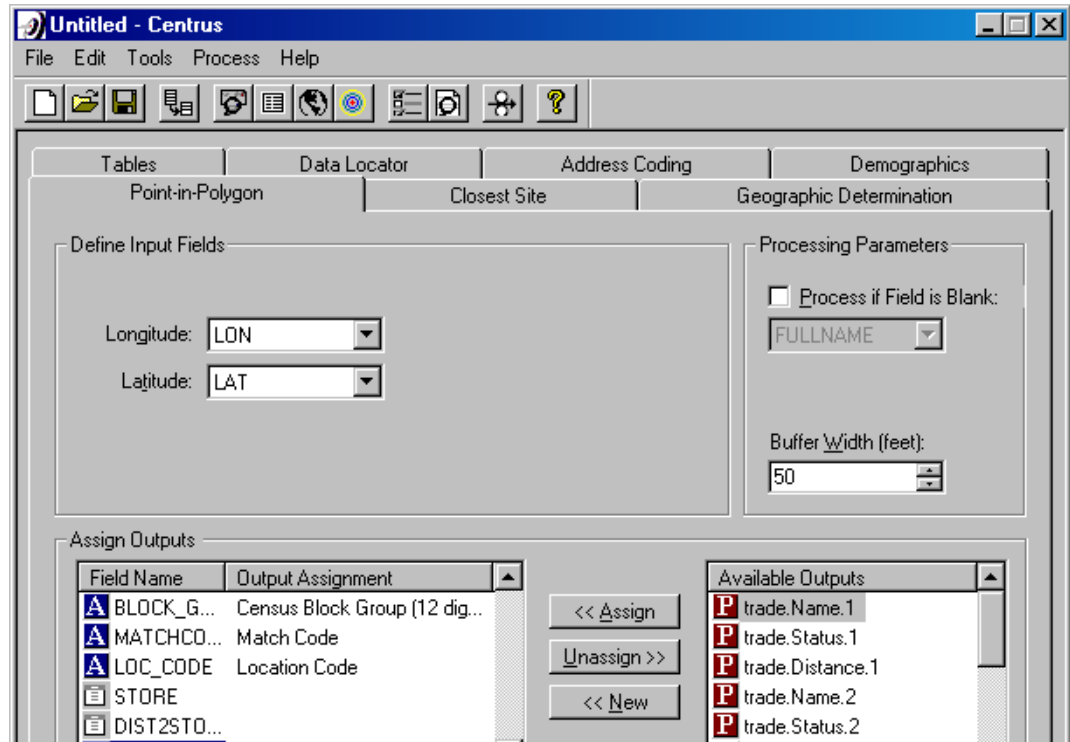
Point-in-Polygon Analysis

Point-in-Polygon (or PiP) analysis determines in which object, or objects, a point falls. This analysis requires a buffered object file, as well as a geocoded address file containing latitude and longitude fields. The Point-in-Polygon module can perform analyses on polygons, as well as buffered lines, buffered points, and buffered polygons, referred to as corridors, circles, and zones, respectively. Collectively, all of these items are referred to as *objects*. Centrus Desktop can process an entire file of points against an object file very rapidly.

The following diagrams illustrate the different types of buffering.



Buffered objects allow you to determine if a point (or points) is within 500 feet of a shoreline, or within 2 miles of a store, or within 1.3 miles of a sales territory. The buffering in Centrus Desktop is done during the actual test, and therefore is completely dynamic. The only requirement is that the object file (.gsb) be built with buffering enabled.



Specifying Point-in-Polygon Input Fields

The Point-in-Polygon module requires only two input fields: **Latitude** and **Longitude**. Once you have selected a file to process, the input fields are filled in automatically, if possible, by looking at the input file's field names and any fields assigned in the Address Coding module. If any required input fields were not filled in, or were not filled with the correct field, click the drop-down arrows to select the correct field.

Buffer Width (Feet)

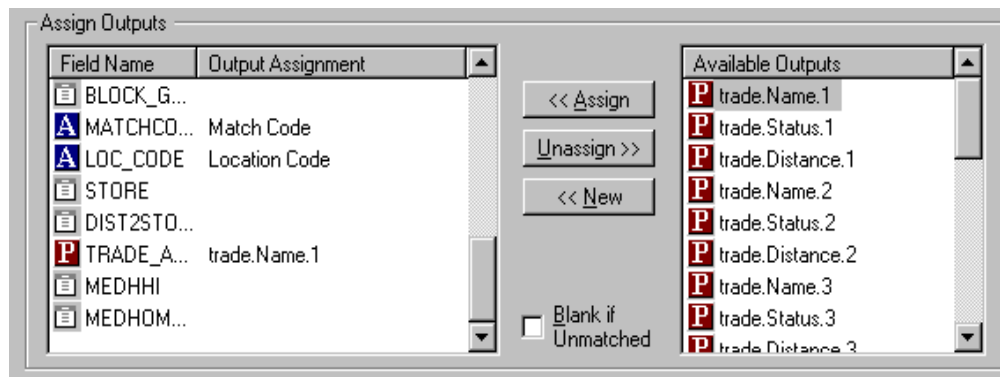
Enter the size of the buffer to use (in feet) in this box, or "0" (zero) if no buffer is desired.

Process if Field Is Blank

The **Process if Field is Blank** option lets you process only those records with missing data in a particular field, which you specify using the drop-down list. This is useful for "cleaning up" a previously processed database.

Assigning Point-in-Polygon Output Fields

The **Assign Outputs** section of the **Point-in-Polygon** dialog box is where you associate available outputs with output field names. You can assign outputs to existing field names, or create new fields to receive the data.



By default, the list of field names is the same as the input file's fields, plus any fields added during processing by other Centrus Desktop modules.

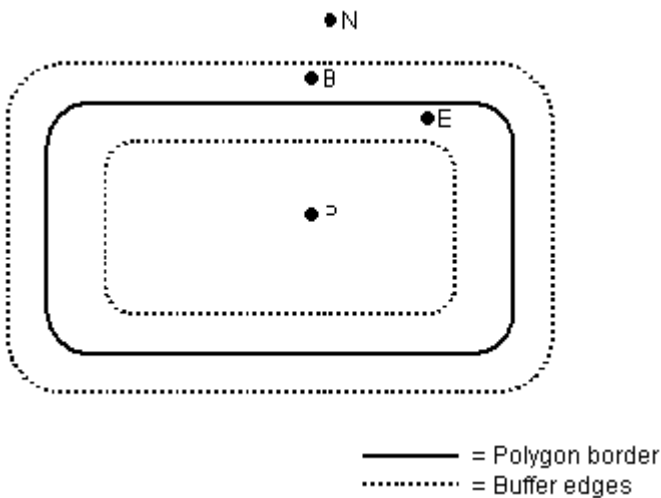
Available Point-in-Polygon Elements

The output data elements are dependent on which layers you have specified in the **Process Options** dialog box. If no data elements are listed under **Available Outputs**, be sure you have correctly specified your layers in the **Process Options** dialog box.

Each feature within a PIP layer is associated with three output data elements:

Field Name	Explanation
Status	<p>Defines where an address lies in relation to the polygon. Status codes are:</p> <ul style="list-style-type: none"> P = In polygon E = Edge; in buffer, inside polygon B = In buffer, outside polygon N = Not found <p>See Figure 1 for an illustration of the status codes.</p>
Distance	<p>The distance from the address to the edge of the polygon. The GSB file must be built with buffering enabled to return a distance value.</p>
Name	<p>The name of the feature as it is defined in the layer.</p>

Figure 1: Point-in-Polygon Status Codes

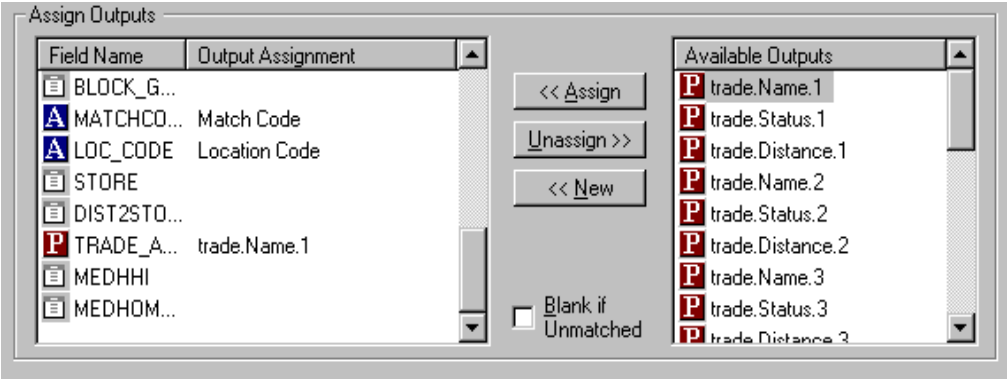


Assigning Field Names

To associate an output data element with an output field name:

1. Select a field name by clicking it.
2. Select the desired output data element by clicking it.
3. Click the <<Assign button.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “P” icon appears before the data element, signifying that the data is being assigned by the Point-in-Polygon module.



Creating New Fields

To create a new field:

1. Select the desired output data element by clicking it.
2. Click the <<New button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “P” icon appears before the data element, signifying that the data is being assigned by the Point-in-Polygon module.

Unassign

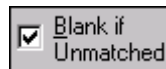
If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Note: You can unassign an output assignment from any tab.

Blank if Unmatched

If you check the **Blank if Unmatched** box, *all* output fields defined within the Point-in-Polygon module are cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Point-in-Polygon elements to new fields rather than input fields, or be sure the **Blank if Unmatched** box is not checked.

Note: Do not use this option if you wish to retain unmatched input addresses.



Chapter 11

The Closest Site Module

Closest Site analysis is a distance test between two sets of points. Health care, banking, and insurance industries use Closest Site analysis to match up doctors and patients, find the nearest bank branch to a customer, and to calculate the distance between a home and the closest fire station. Closest Site calculations return the distance from the address to the site point, as well as the name of that site point.

This chapter includes information about:

- Object Files and Layers
- Closest Site Analysis
- Specifying Closest Site Input Fields
- Assigning Closest Site Output Fields

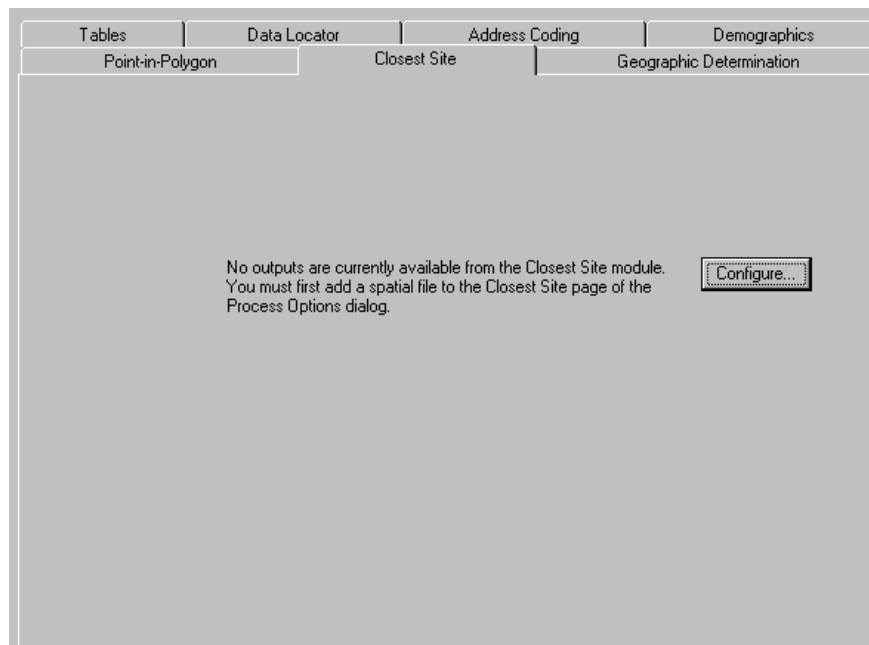
Object Files and Layers

The Closest Site module requires geocoded addresses with Latitude/Longitude coordinates as input. If you need to geocode the addresses in your files, we recommend that you use the Centrus Desktop Address Coding module, for the quickest, most accurate geocode assignments. You can even geocode and perform spatial analyses in the same task.

The available spatial outputs are determined by the contents of the “object” files or *layers* you have specified in the **Process Options** dialog box. Geographic features (points, lines, or polygons) are imported into the native Centrus Desktop format which allows very fast searches. To get your data into our .gsb format, we support MapInfo’s .mif/.mid format, as well as ESRI’s Atlas GIS .bna and .shp formats.

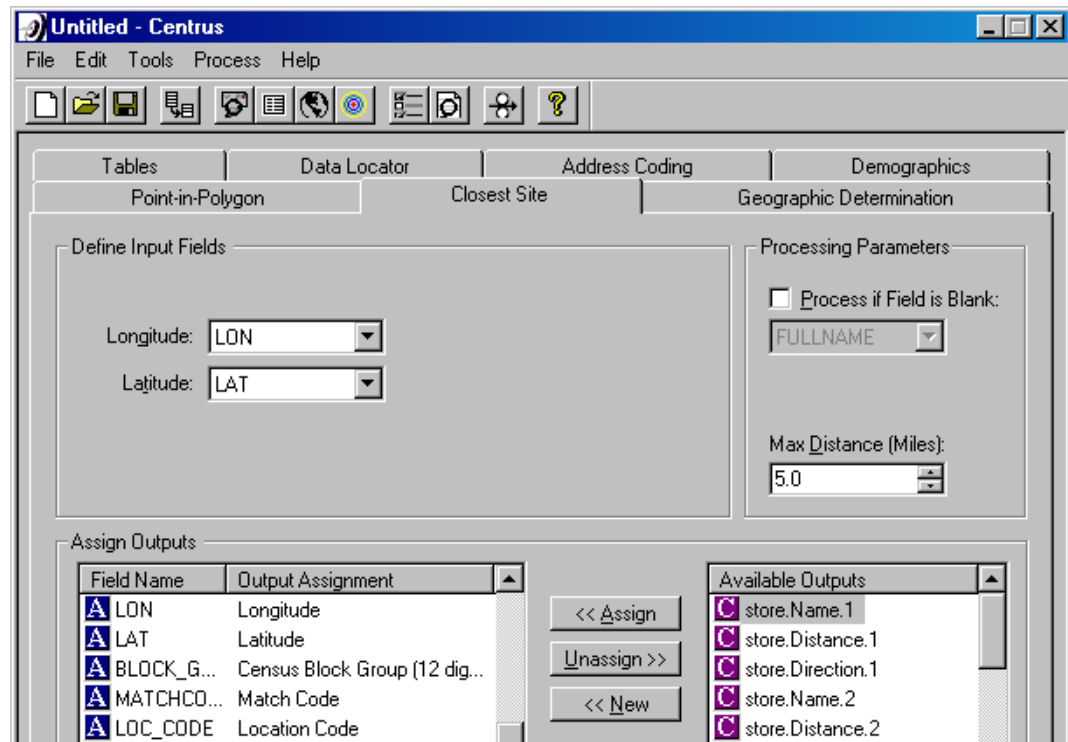
For more information about specifying or importing layers, see “Point-in-Polygon Tab” on page 23 and “Closest Site Tab” on page 27.

Note: The **Closest Site** tab displays only a **Configure...** button until a closest site layer is specified in the Closest Site page of the **Process Options** dialog box.



Closest Site Analysis

Closest Site analysis compares the geocoded addresses (or other points) in your input file to a geocoded site layer you specify in the **Closest Site** options tab. Centrus Desktop can identify the closest sites for each point in the input file. The site's identifier and the straight-line distance to the site can be returned for each site found.



Specifying Closest Site Input Fields

The spatial analysis module requires only two input fields: **Latitude** and **Longitude**. Once you have selected a file to process, the input fields are filled in automatically, whenever possible, by looking at the input file's field names and any fields assigned in the Address Coding module. If any required input fields were not filled in, or were not filled with the correct field, click the drop-down arrows to select the correct field.

Maximum Distance (Miles)

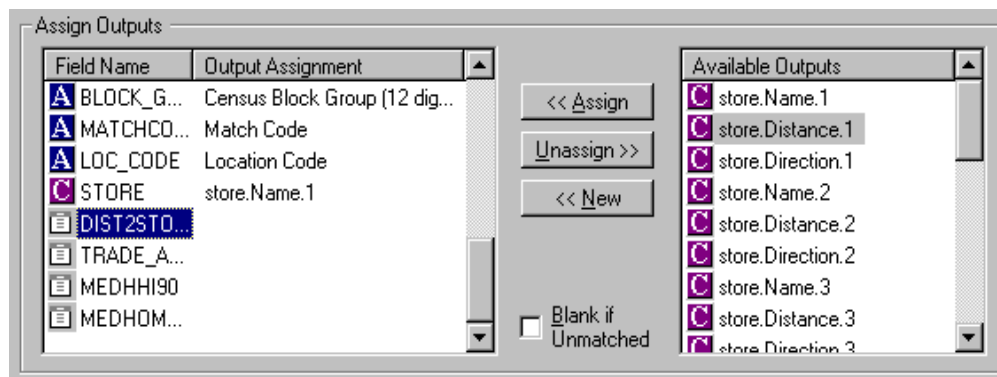
This setting tells Centrus Desktop how far to look for a site. For example, if you are interested only in those sites within 50 miles of an address, you should enter 50 in the **Max Distance (Miles)** box.

Process if Field Is Blank

The **Process if Field is Blank** option lets you process only those records with missing data in a particular field, which you specify using the drop-down list. This is useful for “cleaning up” a previously processed database.

Assigning Closest Site Output Fields

The **Assign Outputs** section of the **Closest Site** dialog box is where you associate available outputs with output field names. You can assign outputs to existing field names, or create new fields to receive the data.



By default, the list of field names is the same as the input file’s fields, plus any fields added during processing by other Centrus Desktop modules.

Available Closest Site Elements

The output data elements are dependent on which layers you have specified in the **Process Options** dialog box. If no data elements are listed under **Available Outputs**, be sure you have correctly specified your layers in the **Process Options** dialog box.

Each feature within a Closest Site layer is associated with three output data elements:

Field Name	Explanation
Name	The name of the feature as it is defined in the layer.
Distance	The distance from the address to the closest site in the layer. The GSB file must be built with buffering enabled to return a distance value.
Direction	The direction from the address to the closest site in the layer, measured in degrees, with north at 0/360 degrees.

Assigning Field Names

To associate an output data element with an output field name:

1. Select a field name by clicking it.
2. Select the desired output data element by clicking it.
3. Click the <<Assign button.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “C” icon appears before the data element, signifying that the data is being assigned by the Closest Site module.

Creating New Fields

To create a new field:

1. Select the desired output data element by clicking it.
2. Click the <<New button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “C” icon appears before the data element, signifying that the data is being assigned by the Closest Site module.

Unassign

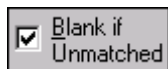
If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Note: You can unassign a field from any tab.

Blank if Unmatched

If you check the **Blank if Unmatched** box, *all* output fields defined within the Closest Site module are cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Closest Site elements to new fields rather than input fields, or be sure the **Blank if Unmatched** box is not checked.

Note: Do not use this option if you wish to retain unmatched input addresses.



Chapter 12

The Geographic Determination Module

The Geographic Determination module eliminates any uncertainty (in the calculations of the initial geocode) by providing a confidence factor and associating it with the geocode so that additional rules may be applied. This confidence factor can be described as either a distance value (between two points and/or lines) or as a percentage representing the amount of overlap between two polygons.

Geographic determination describes both the spatial relationship between two entities and the degree of confidence in that description. High confidence determinations, based on any level of geocode (address, ZIP+4, ZIP+2, or ZIP code) are critical for many geographically sensitive business decisions.

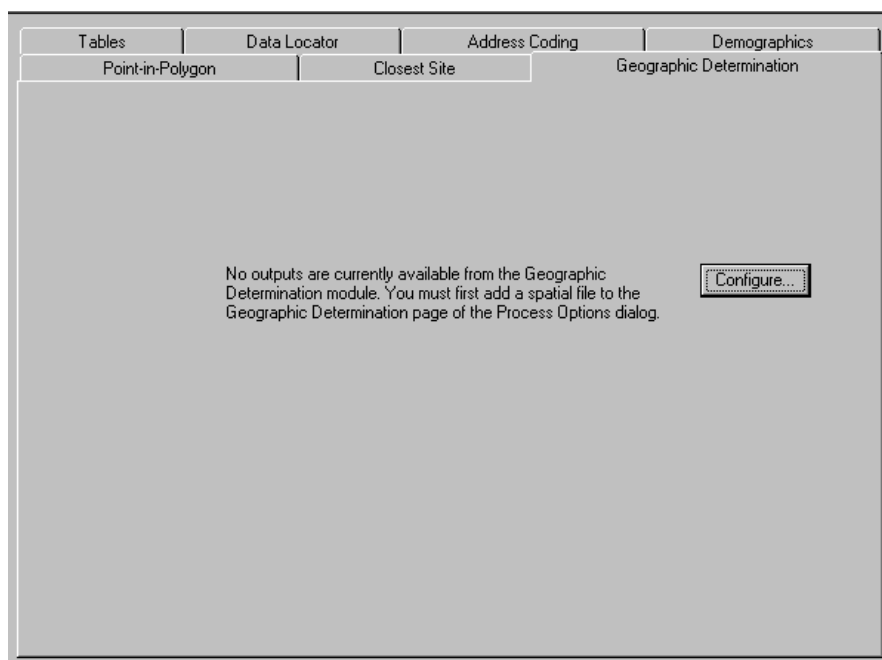
This chapter describes how to use the Geographic Determination Module and includes information about:

- Using the Geographic Determination Module
- Assigning Output Fields
- Available Geographic Determination Elements

Using the Geographic Determination Module

“Setting Processing Options” on page 12 explains how to configure the Geographic Determination module to meet your requirements. “File Specification” on page 85 explains how to select a file for processing. Once you’ve performed these steps, you can specify the search distance and buffer width, assign outputs, then click **Process** on the task bar to process the file or configure a different module.

Note: The **Geographic Determination** tab displays only a **Configure...** button until a .gsb file is specified in the Geographic Determination page of the **Process Options** dialog box.

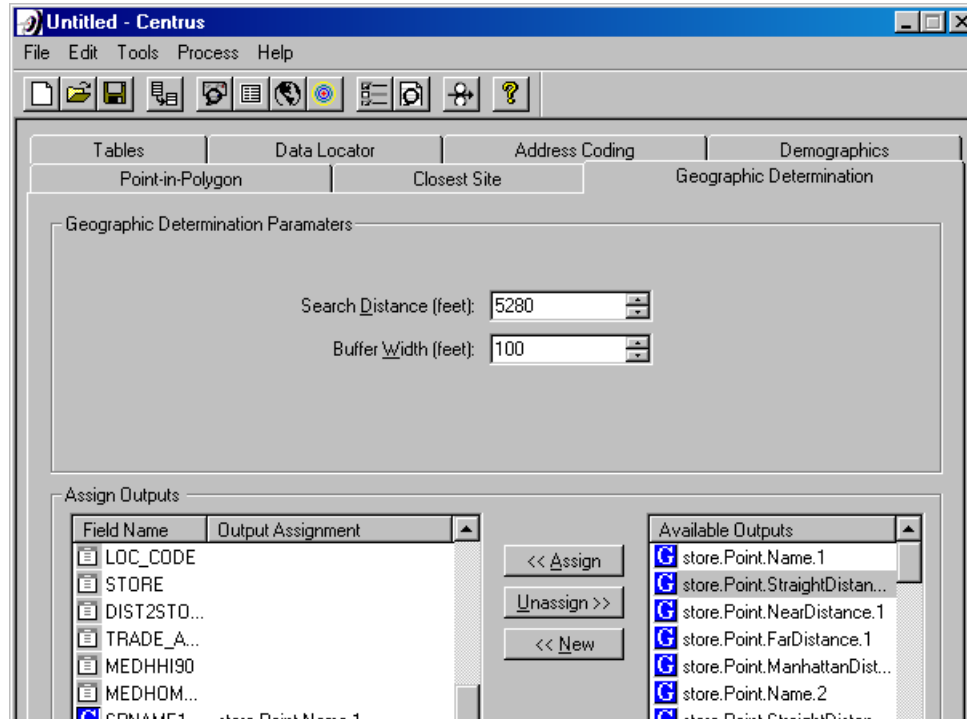


Search Distance (feet)

This setting tells Centrus Desktop how far to look for a site. The default search distance is 5280. For example, Centrus Desktop finds the distance within 5280 feet of the point and assigns a confidence level.

Buffer Width (feet)

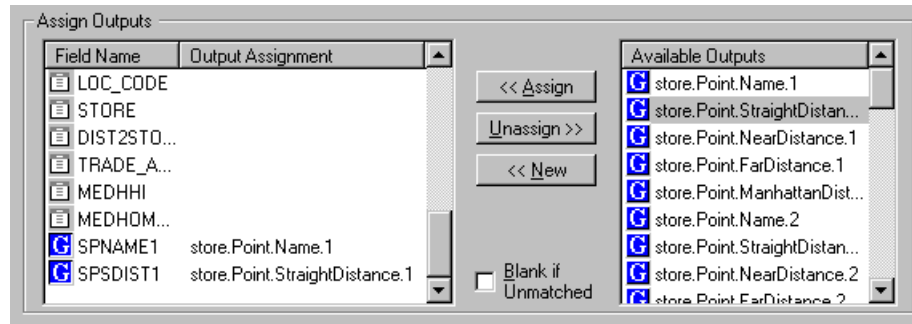
This setting tells Centrus Desktop how far to map the polygon. The default is 100 feet. For example, Centrus Desktop finds the distance within 100 feet of the boundary and creates a polygon around it.



Assigning Output Fields

In the **Outputs** section of the **Geographic Determination** dialog box, you associate data elements with output field names. You can assign data elements to existing field names, or create new fields to receive the data.

By default, outputs are ordered alphabetically by field name. Click **Output Assignment** to view outputs ordered alphabetically by output assignment.



Assigning Fields

To associate a field name with an output field:

1. Select a field name by clicking it.
2. Select the desired data element by clicking it.
3. Click the **<<Assign** button.

The selected data element appears in the **Output Assignment** column next to the field name to which it is assigned. A “G” icon appears before the address element, signifying that the data is being assigned by the Geographic Determination module.

Creating New Fields

To create a new field:

1. Select the desired data element by clicking it.
2. Click the **<<New** button. A **New Field** dialog box appears.
3. Specify the name, type, width, and (if numeric) decimal places desired, or click **OK** to accept the default values.

The selected address element appears in the **Output Assignment** column next to the field name to which it is assigned. A “G” icon appears before the address element, signifying that the data is being assigned by the Geographic Determination module.

Unassign

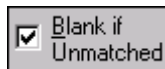
If the currently selected field is an input field, clicking the **Unassign>>** button removes its output assignment. If the currently selected field is a new field, clicking the **Unassign>>** button deletes it.

Note: You can unassign a field from any tab.

Blank if Unmatched

If you select the **Blank if Unmatched** check box, *all* output fields defined within the Geographic Determination module are cleared if the record cannot be matched. If you want to retain unmatched input addresses, either assign Geographic Determination elements to new fields rather than input fields, or be sure the **Blank if Unmatched** check box is cleared.

Note: Do not use this option if you wish to retain unmatched input addresses.



Available Geographic Determination Elements

The outputs available for the Geographic Determination module vary according to the type of file (point, line, or polygon) specified in the **Process Options** dialog box. The name of each available output consists of the name of the GSB file, the type of file, the data element, and a number indicating a distance ranking from closest (1) to farthest (8). For example:

`store.Line.Name.1`

is the *Name* data element that is associated with the closest object in the GSB file, and the data element exists in a *line* file called *store.gsb*. The number of distance rankings for each element is specified in the input file. The following table lists the Geographic Determination elements available for use with each type of file.

Field Name	Explanation	Available With Files of Type
Name	Returns the name of the object in the file.	Point, Line, Polygon
StraightDistance	Returns the straight line distance between two points.	Point
NearDistance	Returns the <i>minimum</i> distance that exists between an error surface or a geocoded point and another point or line.	Point, Line
FarDistance	Returns the distance between the farthest point of the error surface and the nearest point of the feature (point, line or polygon).	Point, Line
Manhattan Distance	Returns the Manhattan distance between two points.	Point
Overlap	Returns the percentage of overlap between the error surface and the input polygon.	Polygon

Chapter 13

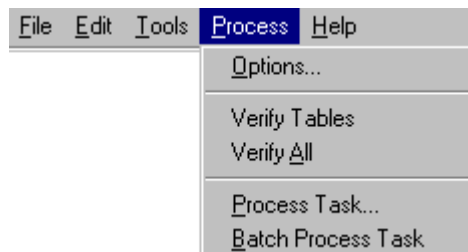
File Processing

This chapter describes how to process a task and how to view the results. It includes the following information:

- About File Processing
- Interactive Processing
- Batch Processing
- Process Dialog Box
- Possible Matches Dialog Box


About File Processing

Once you have selected an input file, set the processing options, and specified the input and output fields for the modules you are using, Centrus Desktop is ready to begin processing. From the main Centrus Desktop menu, select **Process**, then choose either **Process Task** for interactive processing, or **Batch Process Task** for batch processing.



Different input files may require different approaches to matching. If you are processing a small number of addresses, and every address is important, Centrus Desktop allows you to select processing options to maximize the number of matches while making the best use of your time. If you are processing many thousands, or even millions of addresses, Centrus Desktop allows you to process in batch mode for unattended matching.

Interactive Processing

If you select **Process > Process Task** or click the **Process** button , the **Process dialog** box appears. The first record in the address file is always loaded and processed, regardless of the **Process if Field is Blank** settings in the individual Centrus Desktop modules. Each Centrus Desktop module displays its current status on its own **Process** dialog box, accessed by clicking the appropriate tab. Modules not currently in use display neither inputs nor results.

Note: Many features of interactive processing are unavailable if you are processing in CASS mode. CASS certification requires that all input records be processed.

Batch Processing

If you select **Process > Batch Process Task**, the entire file is processed in batch mode, letting you configure a different task or perform other work while processing takes place.

Process Dialog Box

Each module's **Process** dialog box has an **Inputs** section, a **Results** section, and a **Records** section.

The **Inputs** section of a dialog box displays the record as it exists in the input file, or input data generated by other Centrus Desktop modules. Inputs are directly editable in the **Data Locator** and **Address Coding** dialog boxes, but not in the **Demographics**, **Point-in-Polygon**, or **Closest Site** dialog boxes.

The processed information is shown in the **Results** section of the dialog boxes. The data is presented as a hierarchical tree. You can double-click an item to expand or contract the information display.

The **Records** section contains the processing statistics. It shows the total records processed and the number of records standardized and geocoded, as well as the estimated completion time.

Within the **Process** dialog box, you can perform several different functions using the buttons on the right side of the dialog box. A description of these functions follows.

Note: When using Centrus Desktop in CASS mode, all buttons except **Cancel** are disabled. At this time, the USPS does not allow interactive coding when producing a CASS report. If you need to use both interactive coding and produce a CASS report, you need to process twice.

Find

Click the **Find** button to process the current address as it appears in the **Inputs** section. The results of the process are displayed in the **Results** section of the dialog box. The address may be edited and geocoded as often as necessary in order to receive a proper match.

Next Error

Click the **Next Error** button to process the address file until the next unmatched record is encountered. The unmatched record displays in the **Inputs** and **Results** sections, along with the reason that no match could be made. The address can be edited to correct any errors or other flaws in the address. Click the **Next**, **Next Error**, or **Batch** buttons to continue processing.

This feature is useful if you wish to maximize the number of matches that can be received.

Batch

Click the **Batch** button to begin processing records without interruption. Centrus Desktop processes the entire file, displaying statistics for the number of records processed. The **Inputs** and **Results** areas of the dialog box do not update during batch processing.

Once batch processing begins, all buttons except **Cancel** appear dimmed. Click the **Cancel** button to interrupt processing on the current record and display the same options that were available before batch processing began. The current record is displayed in the **Inputs** and **Results** sections.

Next

Click the **Next** button to save the current information in the **Results** section to the output file and load the next address for processing. This function is useful for interactive processing of a file, or if you wish to verify the results of every match. If you are processing unmatched records only, clicking this button loads the next unmatched record.

Previous

This button is not available if you are processing an ASCII file or writing to a separate output file. Clicking the **Previous** button saves the current information in the **Results** section to the output file, and loads the previous address for processing. If you are writing a Reject file or Audit report, you are warned that they will be aborted if you load the previous record. If you are writing a Log report, the statistics section at the bottom of the report is omitted.

Goto...

Click the **Goto** button to jump to an absolute record number. If you are processing an ASCII file or writing to a separate output file, you can only jump forward in the file. If you are writing a Reject file or Audit report, and you choose to jump to a previous record, you are warned that these files will be aborted if you perform the jump. If you are writing a Log report and jump backward in the file, the statistics section at the bottom of the report is omitted.

Reload

Click the **Reload** button to load the original data from the address file into the **Inputs** and **Results** sections. To geocode the record, click the **Find** button. This button is useful if you have made a number of edits to an address in an attempt to receive a geocode, and would like to return to the original address in the address file.

Query...

Note: This function is not supported for Canadian addresses.

If you are unable to produce a match by editing an address and using the **Find** button, click the **Query** button to display the **Query** dialog box. See “The Query Function” on page 67 for more information.

Map...

Note: This function is not supported for Canadian addresses.

Click the **Map** button to display streets and county boundaries for the city entered. The currently geocoded point is shown when the map is first displayed. The map function can be used to place unmatchable addresses, or simply to view streets in an area. See “Map Viewer” on page 75 for more information.

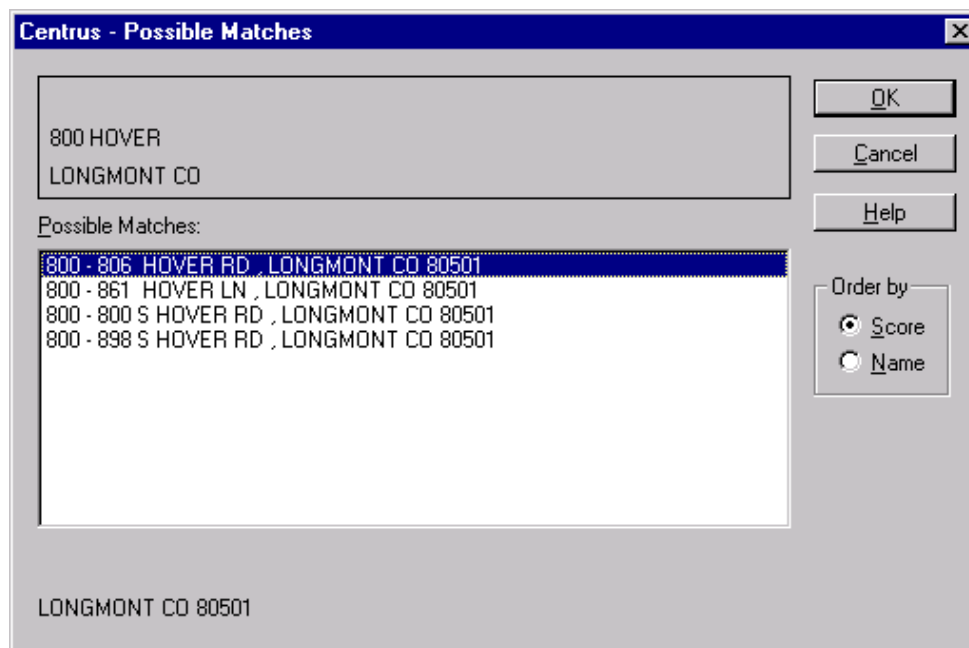
Cancel

Click the **Cancel** button to abort the current geocoding process. If you are updating a file in place, the process stops, and records that have been processed remain coded. If you are creating an output file, you are asked if you want to “*Save geocodes from current process?*”. The possible choices are:

- **Yes**—Skip the remaining records. If you are writing all records to an output file, then the remaining records are copied to the output file, which could take some time. If you are not storing all records to the output file, then the remaining records are written to the Reject file, if specified, which can also take some time.
- **No**—This option erases any output and Reject files.
- **Cancel**—This option continues processing and returns to the **Process** dialog box.

Possible Matches Dialog Box

If processing was initiated by using the **Find**, **Next**, or **Next Error** buttons, Centrus Desktop may display a **Possible Matches** dialog box. This dialog box appears if Centrus Desktop could not determine on its own which record matches the input address. When this dialog box is displayed, you can either select one of the addresses displayed, or click the **Cancel** button if none of the displayed addresses is an appropriate match. This dialog box can also appear if you are using QuickFind.



You can adjust the way possible matches are displayed using the radio buttons in the **Order by** section. **Order by Score** sorts them from the “most likely” candidate to the “least likely” candidate. **Order by Name** does a simple sort based upon the street name.

Appendix A

System Messages and Codes

Centrus Desktop uses match codes and location codes to provide detailed information about the matches retrieved. Match codes provide information about the match itself, and location codes provide information about the assigned geocode.

The following topics describe the two types of codes and explain how to interpret them:

- Match Codes
- Location Codes

Match Codes

When you use Centrus Desktop to perform address standardization, a match code is returned in the MatchCode output field. The match code is an alpha-numeric code that encapsulates information about the address standardization process—including whether or not a match was found, information about the type of match found (when applicable), and information about why no match was found (when applicable). The following sections provide the information you need to interpret these codes:

- Match Codes for the United States
- Match Codes for Canada

Match Codes for the United States

Table 1 and Table 2 describe the codes returned in `MatchCode` when a match is found. The first character, an alphabetic element, describes the type of match found. The two- or three-digit numeric (or hexadecimal) element of the code provides detailed information about the match. The meaning of the numeric elements are enumerated, by position, in Table 2.

Table 3 describes the codes returned in `MatchCode` when a match is not found.

Table 1: Values returned in MatchCode when a match is found

Code	Description
Shh	Indicates a match found in USPS data. This is considered the best address match, because it was matched directly against the USPS list of addresses. See below for the interpretation of the hex digits.
Ahh	Same as Shh, but indicates match to an alias name record or an alternate record.
D00	Match is a small town with P.O. Box and/or General Delivery only.
Uhh	Indicates match found in USPS data but the ZIP+4 code cannot be resolved without the firm name or other information. See below for the interpretation of the hex digits. In CASS mode, an E023 (multiple match) error code will be returned.
Qhh	Indicates a match to USPS range records with unique ZIP Codes. This match code was introduced for clarity because CASS rules prohibit altering an input ZIP if it matches a unique ZIP Code value.
Thh	Indicates a match to a street segment record. Street segment records do not contain ZIP Code information. If an input ZIP is entered, the input ZIP is returned. If the input city and state has only one ZIP Code, that ZIP Code is returned.
Xhhh	Match found was for an intersection of two streets, for example, "Clay St & Michigan Ave." Please note that the USPS does not allow intersections as a valid deliverable address. The first hex digit refers to the last line information, the second hex digit refers to the first street in the intersection, and the third hex digit refers to the second street in the intersection. To decode the hex digits, see the following list.
Yhhh	Same as Xhhh, but indicates that an alias name record for one or both streets was used. The first hex digit refers to the last line information, the second hex digit refers to the first street in the intersection, and the third hex digit refers to the second street in the intersection. To decode the hex digits, see the following list.
Z	No address was given, but the ZIP Code was verified as valid.

Table 2: Value of hexadecimal digits in "match found" return codes

Code	In first hex position means:	In second and third hex position means:
0	No change in last line.	No change in address line.
1	ZIP Code was changed.	Street type was changed.
2	City was changed.	Pre-directional was changed.

Table 2: Value of hexadecimal digits in “match found” return codes (continued)

Code	In first hex position means:	In second and third hex position means:
3	City and ZIP Code were changed.	Street type and Pre-directional were changed.
4	State was changed.	Post-directional was changed.
5	State and ZIP Code were changed.	Street type and Post-directional were changed.
6	State and City were changed.	Pre-directional and Post-directional were changed.
7	State, City, and ZIP Code were changed.	Street type, Pre-directional, and Post-directional were changed.
8	ZIP+4 was changed.	Street name was changed.
9	ZIP and ZIP+4 were changed.	Street name and Street type were changed.
A	City and ZIP+4 were changed.	Street name and Pre-directional were changed.
B	City, ZIP, and ZIP+4 were changed.	Street name, Street type, and Pre-directional were changed.
C	State and ZIP+4 were changed.	Street name and Post-directional were changed.
D	State, ZIP, and ZIP+4 were changed.	Street name, Street type, and Post-directional were changed.
E	State, City, and ZIP+4 were changed.	Street name, Pre-directional, and Post-directional were changed.
F	State, City, ZIP, and ZIP+4 were changed.	Street name, Street type, Pre-directional, and Post-directional were changed.

Table 3: Values returned in MatchCode when no match is found

Code	Description
Ennn	Indicates an error, or no match. This can occur when the address entered either did not exist in the GeoStan Directory, or the address was badly malformed and could not be parsed correctly. The last three digits of an error code (below) indicate which parts of an address were unable to be matched to the GeoStan Directory.
nnn = 000	No address found by Address Wizard.
nnn = 001	Low level error. Use <code>GsErrorGet</code> to query.
nnn = 002	Could not find GSD file.
nnn = 003	Incorrect GSD file signature or version ID.
nnn = 004	GSD file out of date. Can only occur when CASS Mode is on.
nnn = 010	No city+state or ZIP Code was found.
nnn = 011	Input ZIP was not in the directory.

Table 3: Values returned in MatchCode when no match is found (continued)

Code	Description
nnn = 012	Input city was not in the directory.
nnn = 013	Input city was not unique in the directory.
nnn = 014	Out of licensed area. Can only occur if using Sagent licensing technology.
nnn = 015	GeoStan record count has been depleted and license has expired.
nnn = 020	No matching streets found in directory.
nnn = 021	No matching cross streets for an intersection match.
nnn = 022	No matching ranges. Note: GeoStan EPI does not return this error code, instead it returns 020.
nnn = 023	Match is unresolved.
nnn = 024	No matching ranges. (Same as 022.)
nnn = 025	Too many possible cross streets for intersection matching.
nnn = 026	No address found when attempting a multi-line match.
nnn = 027	Invalid directional attempted. Forbidden by CASS.
nnn = 028	Record also matched Early Warning System (EWS) data, therefore the match was denied.

Match Codes for Canada

Table 4 and Table 5 describe the codes returned in MatchCode when a match is found. The first character, an alphabetic element, describes the type of match found. The two- or three-digit numeric (or hexadecimal) element of the code provides detailed information about the match. The meaning of the numeric elements are enumerated, by position, in Table 5.

Table 6 describes the codes returned in MatchCode when a match is not found.

Table 4: Values returned in MatchCode when a match is found

Return code	Explanation
Chhh	Indicates a match found in CPC data; changes were made to the address to make it deliverable.
Vhhh	Indicates a match found in CPC data; the input address is valid and no changes were made.

The returned address is the best address because it was matched directly against the CPC list of deliverable addresses. See below for the interpretation of the hex digits.

Table 5: Value of hexadecimal digits in "match found" return codes

Code	First hex position indicates	Second and third hex position indicates
0	No change in last line.	No change in street type, direction, number, or name.
1	Postal Code was changed.	Street type was changed.
2	Municipality was changed.	Postfix direction was changed.
3	Municipality and Postal Code were changed.	Street type and Postfix direction were changed.
4	Province was changed.	House number was changed.
5	Province and Postal Code were changed.	Street type and House number were changed.
6	Province and Municipality were changed.	House number and Postfix direction were changed.
7	Province, Municipality, and Postal Code were changed.	Street type, Postfix direction, and House number were changed.
8	Reserved for future use.	Street name was changed.
9	Reserved for future use.	Street name and type were changed.
A	Reserved for future use.	Street name and Postfix direction were changed.
B	Reserved for future use.	Street name, Street type, and Postfix direction were changed.
C	Reserved for future use.	Street name and House number were changed.
D	Reserved for future use.	Street name, Street type, and House number were changed.
E	Reserved for future use.	Street name, House number and Postfix direction were changed.
F	Reserved for future use.	Street name, Street type, House number and Postfix direction were changed.

Table 6: Values returned in MatchCode when no match is found

Code	Explanation
Ecnn	Indicates an error or no match. This can occur when the address entered either did not exist in the GeoStan Canada Directory, or the address was badly malformed and could not be passed correctly.
nn = 01	Internal error.
nn = 10	No Municipality+Province or Postal Code found.
nn = 20	No matching addresses.
nn = 22	Missing or wrong street name.
nn = 23	Could not resolve address.
nn = 25	Inconsistent address.
nn = 26	Missing or wrong box range.
nn = 27	Missing or wrong unit range.
nn = 40	Address vs. Postal Code conflict, SERP rule prevents correction.
nn = 41	Postal Code has multiple street names, SERP rule prevents correction.
nn = 42	Change of delivery mode attempted, SERP rule prevents correction.
nn = 43	Total number of changed address elements exceeds maximum.
nn = 44	Change of address type is not allowed for current setting.
nn = 50	Minor error.

Location Codes

Location codes indicate the accuracy of the assigned geocode. There are two types of geocodes—*Address* and *ZIP+4* centroids.

Address geocodes are simple to interpret because they indicate a geocode made directly to a street segment (or two segments in the case of an intersection). *ZIP+4* centroids, however, have a range of “confidence” depending upon how the *ZIP+4* centroid was determined. An “E” code indicates that no geocode was possible. The following sections describe the location codes:

- Address Location Codes
- ZIP+4 Centroid Location Codes
- GeoStan Canada Location Codes

Address Location Codes

Address location codes detail the known qualities about the geocode. An address location code has three characters. The first character is always an “A,” indicating an address location. The second character is either an “S,” indicating a location on a street range, or an “X,” indicating a location on an intersection of two streets. The third character is a digit, indicating other qualities about the location. Address codes are defined in detail below.

Table 7: Address location code descriptions

Code	Description
AIn	Indicates the correct segment is inferred from the candidate records at match time.
ASn	Indicates a house range address geocode. This is the most accurate geocode available.
Note: Both AIn and ASn share the same qualities for “n” as follows:	
n = 0	Best location.
n = 1	Street side is unknown. The Census FIPS Block ID from the left side is assigned; however, no offset is assigned and the point is placed directly on the street.
n = 2	May indicate one or both cases: Address was interpolated onto a TIGER segment that did not contain address ranges initially. The original segment had its name changed to match the US Postal Service spelling. This specifically refers to street type, predir, and postdir (any or all). Note: Because segment range interpolation is done only for TIGER data, only the second case is valid for non-TIGER data (e.g., GDT).
n = 3	Both 1 and 2.
n = 7	Placeholder. A placeholder is used when starting and ending points of segments contain the same value and no shape data is available.
AXn	Indicates an intersection geocode. The digit at the end indicates the following:
n = 3	Standard single-point intersection computed from the center lines of street segments.
n = 8	Indicates an interpolated (divided-road) intersection geocode. An attempt to return a centroid for the intersection will be made.

ZIP+4 Centroid Location Codes

ZIP centroid codes actually indicate the quality of two location attributes: Census ID accuracy and positional accuracy. The first character is always a “Z,” indicating a location derived from a ZIP centroid. The second character indicates Census ID accuracy and the third character indicates location type.

The last character indicates how the location and Census ID was defined. This last code is provided for completeness, and may not be useful for most applications. If this is the case, you can set the width of the Location Code field to three and this last identifier is not returned.

2nd Character (Census ID Accuracy)

Table 8: ZIP+4 centroid location code description – 2nd character

Code	Description
B	Indicates Block Group accuracy (most accurate).
T	Indicates Census Tract accuracy.
C	Indicates unclassified Census accuracy. Normally accurate to at least the County level.

3rd Character (Positional Accuracy)

Table 9: ZIP+4 centroid location code description – 3rd character

Code	Description
9	Indicates location based upon a ZIP+4 centroid. These are the most accurate centroids and normally place the location on the correct block face. For a small number of records, the location may be the middle of the entire street on which the ZIP+4 may fall. See the 4th character for a precise indication of locational accuracy.
7	Indicates location based upon a ZIP+2 centroid. These locations can represent a several block area in urban locations, or slightly larger in rural areas.
5	Indicates the location of the Post Office that delivers mail to that address, or a 5-digit ZIP Code centroid, or a location based upon locale (city). See the 4th character for a precise indication of locational accuracy.

4th Character (Methodology)

In practice, there are 23 possible combinations of location codes, one for each methodology type. Therefore, in this section we list the entire location code at the end of the description for each methodology. The offset distance is 50 feet (default), or the value specified in the Address Coding tab of the **Process Options** dialog box (for more information about setting the offset distance, see “Offset/Backset” on page 18).

Table 10: ZIP+4 centroid location code description – 4th character

Code	Description
A	Created via address matching to a single segment. Location is assigned in the middle of the matched street segment, offset to the proper side of the street.
a	Created via address matching to a single segment, but the correct side of the street is unknown. Location is assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
B	Created via address matching to multiple segments, all segments have the same Block Group. Location is assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location is offset to the proper side of the street.
b	Same as methodology B except the correct side of the street is unknown. Location is assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
C	Created via address matching to multiple segments, all segments have the same Census Tract. The Block Group that represents the most households in this ZIP+4 is returned. Location is assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location is offset to the proper side of the street.
c	Same as methodology C except the correct side of the street is unknown. Location is assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
D	Created via address matching to multiple segments, all segments have the same County. The Block Group that represents the most households in this ZIP+4 is returned. Location is assigned to the middle of the matched street segment with the most house number ranges within this ZIP+4. Location is offset to the proper side of the street.
d	Same as methodology D except the correct side of the street is unknown. Location is assigned in the middle of the matched street segment, offset to the left side of the street, as address ranges increase.
E	Created via street name matching (no house ranges available). All matched segments have the same Block Group. Location placed on the segment that is closest to the center of matched segments. In almost all cases, this is on the mid-point of the entire street.
F	Created via street name matching (no house ranges available). All matched segments have the same Census Tract. Location placed on the segment that is closest to the center of matched segments. In almost all cases, this is on the mid-point of the entire street.
G	Created via street name matching (no house ranges available). All matched segments have the same County. Location placed on the segment that is closest to the center of matched segments. In almost all cases, this is on the mid-point of the entire street.
H	Same as methodology G, but some segments were not in the same County. This methodology is used for less than .05 percent of the centroids.

Table 10: ZIP+4 centroid location code description – 4th character (continued)

Code	Description
I	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, and b. All centroids in this ZIP+2 cluster have the same Block Group. Location is assigned to the ZIP+2 centroid.
J	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, b, C, and c. All centroids in this ZIP+2 cluster have the same Census Tract. Location is assigned to the ZIP+2 centroid.
K	Created ZIP+2 cluster centroid as defined by methodologies A, a, B, b, C, c, D, and d. Location is assigned to the ZIP+2 centroid.
L	Created ZIP+2 cluster centroid as defined by methodology E. All centroids in this ZIP+2 cluster have the same Block Group. Location is assigned to the ZIP+2 centroid.
M	Created ZIP+2 cluster centroid as defined by methodology E and F. All centroids in this ZIP+2 cluster have the same Census Tract. Location is assigned to the ZIP+2 centroid.
N	Created ZIP+2 cluster centroid as defined by methodology E, F, G, and H. Location is assigned to the ZIP+2 centroid.
V	Over 95 percent of addresses in this ZIP Code are in a single Census Tract. Location is assigned to the ZIP Code centroid.
W	Over 80 percent of addresses in this ZIP Code are in a single Census Tract. Reasonable Census Tract accuracy. Location is assigned to the ZIP Code centroid.
X	Less than 80 percent of addresses in this ZIP Code are in a single Census Tract. Census ID is uncertain. Location is assigned to the ZIP Code centroid.
Y	Generally indicates a rural or sparsely populated area. Census code is uncertain. Location is based upon the USGS places file.
Z	Generally assigned for P.O. Box or General Delivery addresses. Census code is uncertain. Location is based upon the Post Office location that delivers the mail to that address.

ZIP Centroid Matches

In the **Options** dialog box, there are seven choices for **Minimum Quality** when selecting ZIP Centroid Matching. The following indicates which methodologies (the fourth character in the location code) are used for each **Quality** selection:

9-Digit Best Location	A, a, B, b, C, c, D, d
9-Digit Good Location	A, a, B, b, C, c, D, d, E, F, G, H
9- or 7-Digit Location	A, a, B, b, C, c, D, d, E, F, G, H, I, J, K, L, M, N
Block Group Accuracy	A, a, B, b, E, I, L
Census Tract Accuracy	A, a, B, b, E, I, L, C, c, F, J, M, V
5-Digit Only	V, W, X, Y, Z
All Centroids	Can return all Methodology codes.

Note: This function is not supported for Canadian addresses.

GeoStan Canada Location Codes

If a valid Postal Code centroid is found, one of the following location codes is returned (Table 11).

Table 11: GeoStan Canada location code descriptions

Code	Description
CAN6	Postal Code level geocode.
EC	Indicates that a geocode is unavailable.

Appendix B

Early Warning System Data

Early Warning System (EWS) data is a free data file the USPS provides to prevent matching errors due to the age of the address data in the Use.gsd and Usw.gsd files.

You can use the Use.gsd and Usw.gsd files that Sagent provides on the *Data Products Suite* CDs for 105 days for CASS processing and 135 days for other processing modes. However, during that time, the USPS may add new addresses to the Address Management System (AMS), from which the USPS address data is extracted. Sagent then adds this new data to the Use.gsd and Usw.gsd file. However, any new addresses activated after the creation of the Use.gsd and Usw.gsd files are not accessible by your Centrus products until you receive and install new data. Therefore, new addresses may be matched to a record in the current Use.gsd and Usw.gsd files that may not be the best match when the updated USPS address information is used, or the address may not match a record at all.

The USPS creates the EWS data set by examining their address database for new records that could match incorrectly to an existing record, and for addresses not present in the most recent USPS data product. The USPS creates a new EWS data set for download each week. See the *Release Notes for the Centrus Data Products* for information about downloading and installing the EWS data set.

By downloading the new EWS data on a regular basis, you can ensure more accurate address matching and standardization. If the EWS data file is present, you receive match code E028 if an input address matches to a record in the EWS data and no match is made.

Note: The USPS does not require the EWS file for CASS processing with the CASS 2002-2003 cycle.

Appendix C

Reference

This chapter provides reference information on the following topics:

- Customizing the Data Locator Module
- Premium Demographics
- WinSplit
- Centrus Data Formatter and ASCII Text Files
- CASS Overview
- Audit and Log Report Samples
- Canadian Address Accuracy Program

Customizing the Data Locator Module

There are several ways to customize the Data Locator Module. This section includes information on the following topics:

- Regular Expressions
- Custom Dictionaries in Data Locator

Regular Expressions

Regular expressions comprise a powerful pattern matching language. They are commonly used in the UNIX world on the command line, within text editors, and in the context of pattern matching programs such as *grep*.

A regular expression is a pattern that describes a set of ASCII strings. The fundamental building blocks are the regular expressions that match a single character. Most characters, including all letters and digits, are regular expressions that match themselves. Certain characters have special meanings. In the Centrus Desktop Data Locator module, these special characters are:

+ * ? . # % [] ^ \$ () . \

Symbol	Explanation
\	"Escapes" other special characters to match the literal character.
.	Matches any character.
#	Matches a digit, same as [0-9].
%	Matches a single letter, same as [a-zA-Z].
~	Matches any single digit or letter, same as [A-Za-z0-9].
[]	Matches any single character of set.
^	Matches any single character not in set.
()	Groups regular expressions together into subexpressions that can be treated as a single unit.
*	x* matches any number of x's.
+	x+ matches 1 or more x's.
?	x? matches 0 or 1 x's, where x is any single character.
	Boolean OR character; matches one of a choice of expressions.
^string	Anchors match to the beginning of a string.
string\$	Anchors match to the end of a string.

These characters can be combined in various ways to broadly or narrowly specify the strings to be matched. For example:

E-Mail:

+@(+\.)+(com|COM|edu|EDU|gov|GOV|org|ORG|net|NET)

Phone Number:

(1[-]?| ?)([\(\)?###[\)]-]?| ?)###[-]?###(##|
? [extEXT\#]+ ?#)

Social Security Number:

###[-]##[-]####

More About Regular Expressions

The Centrus Desktop Data Locator module uses and extends the standard UNIX regular expression set, adding special uses for the following three characters:

- The hash or pound character (#) matches any single digit.
- The percent sign (%) matches any single letter.
- The tilde (~) matches any single digit or letter. Non-alphanumeric characters, such as punctuation, symbols, or ASCII “control” characters, are *not* matched.

A backslash (\) followed by any special character “escapes” the special character, matching the literal character itself. For example, to match the pound character (#), you’d use the expression \#.

The period (.) matches any single ASCII character. Thus, **.umpty** matches either “Humpty” or “Dumpty”.

A range of ASCII characters may be specified by giving the first and last characters, separated by a hyphen (-).

A set of characters enclosed within square brackets ([and]) matches any single character in that list. For example, the regular expression **[0123456789]** matches any single digit. If the first character of the list is the caret (^), it matches any character *except* those in the set. It does *not* match an empty string. The caret loses this special meaning if it is not the first character of a set enclosed by brackets.

The Boolean OR character (|) allows a choice between two regular expressions. For example, **jell(y|ies)** matches either “jelly” or “jellies”.

The following expressions are equivalent:

```
0|1|2|3|4|5|6|7|8|9
[0123456789]
[11987654321111011]
0-9
```

Parentheses (()) group parts of regular expressions together into subexpressions that can be treated as a single unit. For example, **(ha) +** matches one or more instances of “ha”.

A regular expression matching a single character may be followed by one of several repetition operators: **?**, *****, **+**. These work as follows:

string?	The string element is optional and matched no more than once. Thus, xy?z matches either “xyz” or “xz”.
string*	The string element will be matched zero or more times.
string+	The string element will be matched one or more times.

The following two expressions are equivalent:

```
[0-9] +
[0-9] [0-9] *
```

All or part of a regular expression can be “anchored” to either the beginning or end of the search string. A regular expression that starts with a caret (^) only matches if it occurs at the beginning of the input string. A regular expression that ends with dollar (\$) only matches at the end of the input string. Combining these two anchors constrains a regular expression to match the entire input string.

Custom Dictionaries in Data Locator

The Data Locator comes with an extensive dictionary file. This file contains:

- 48,300 first names
- 11,400 first name “aliases”
- 312,000 last names
- 36,700 company names
- 43 common name prefixes
- Gender scores for first names
- Ethnicity assignments for first and last names

You can define additional dictionary items in a *custom dictionary file*. The *Data Locator Custom Words Dictionary* is a user-defined custom dictionary file, **Custom.isd**. It is usually located in the Centrus Desktop data directory below the main Centrus Desktop program directory. From the **Configuration** dialog box, click the **Browse** button to display a file selection dialog box, then select **Custom.isd**.

Once you specify the name and location of this file, Data Locator adds your custom dictionary entries to its own list of terms and parses input accordingly.

Note: The file **Words.isd** is the Data Locator’s internal dictionary. Do not specify this file as the Data Locator Custom Words Dictionary.

Creating a Custom Dictionary File

To create a custom dictionary file, you must first define the dictionary entries in a comma-delimited ASCII text file where each line is of the general format:

Type, Word[, [Parameter,] [Parameter,] [Parameter]

Type must be one of the following: **FirstName**, **LastName**, **NickName**, **Title**, **CompanyWord**, **CompanyName**, **CityAlias**, **StateName**, **Suppress**, or **Junk**. The type is immediately followed by the word to be identified. Words are not case sensitive, and they may not have spaces except as noted. Depending on the type, the word may be followed by one or more parameters. These are explained in the section “Custom Dictionary Types” on page 173.

Some examples of custom dictionary entries are:

```

FirstName, Grit, 0, English
LastName, Vanbiesbrouck, German, English
NickName, Grit, Margaret
Title, Representative
Title, Secretary
CompanyWord, Software
CompanyWord, Restaurant
CompanyName, Sagent
CityAlias, laf, Lafayette
StateName, fla, FL
Suppress, darn
Junk, NA

```

Once you have defined the dictionary entries and saved them as an ASCII text file, use the `bldcstm.exe` program to compile the text file into an InfoStan dictionary. This command-line utility is located in your Centrus Desktop program directory, and has the following usage:

```
Bldcstm Infile Outfile Logfile
```

where:

- **Bldcstm** starts the program,
- *Infile* is the name of your ASCII text source file,
- *Outfile* is the name to assign to the custom dictionary file, and
- *Logfile* is the name to assign to the program's processing log.

Custom Dictionary Types

The following table provides details on creating custom dictionary entries.

Type and syntax	Description
FirstName , Word,	FirstName types may list a gender score [GenderScore] of 0-100, where 0 represents a name that is [Ethnicity1], feminine in 100% of cases, and 100 represents a [Ethnicity2] name that is masculine in 100% of cases. The gender score may be followed by up to two ethnic identifications. (Valid ethnicities are listed in the next table.)
LastName , Word,	LastName types may be followed by up to [Ethnicity1], [Ethnicity2] two ethnic identifications. (Valid ethnicities are listed in the next table.)
NickName , Word,	Adds a nickname to Data Locator's global <code>RootName</code> list of nicknames.
Title , Word,	Adds a title to Data locator's global list of titles.

Type and syntax	Description
CompanyWord , Word,	Indicates words that are likely to indicate the name of a business, for example, "Ltd." or "Restaurant".
CompanyName , Word,	Used to add the names of specific companies to Data locator's list of known company names. Spaces are permitted.
CityAlias , Word,	Spaces are permitted in the real name but <code>RealName</code> not in the word (alias). For example, "LA" is an acceptable alias word for "Los Angeles", but "L A" is not permitted.
StateName , Word,	A state name or variant state abbreviation is <code>Abbrev</code> followed by the standard abbreviation. Spaces are permitted in the word. Abbreviations must be in the standard two-letter USPS form.
Suppress , Word,	"Suppressed" words are dropped from all output. Spaces are permitted in the word.
Junk , Word,	"Junk" words are words that are not assigned to any of the standard parsing categories (First Name, Last Name, Title, etc.) but can still be returned. Spaces are permitted.

Valid Ethnicities

African	Arabic	Armenian
Asian	Basque	Belgian
Brazilian	Bulgarian	Croatian
Czech	Danish	Dutch
EastEurope	EastIndian	English
Estonian	Finnish	French
German	Greek	Hawaiian
Hispanic	Hungarian	Hungary
Icelandic	Iranian	Irish
Italian	Japanese	Lithuanian
NativeAmerican	Norwegian	Pakistani
Persian	Polish	Portuguese
Rumanian	Russian	Scandinavian
Scottish	Serbian	Slavic
Swedish	Swiss	Turkish
Ukrainian	Vietnamese	Welsh
Many		

Premium Demographics

Sagent provides Census demographics with the standard Centrus Desktop license. Sagent also provides the following premium demographic data sets through separate licensing:

- Claritas Demographics Data
- Claritas PRIZM Codes

Claritas Demographics Data

Table 12 lists the Claritas Demographics data elements, which Sagent provides through a separate license:

Table 12: Claritas demographics data elements.¹

DMA Designated Market Areas
Census Place Code
Census County Subdivision
Place FIPS Code
Minor Civil Division
Current Year Estimated Population
Current Year Estimated Total Female Population
Current Year Estimated Housing Units
Current Year Estimated Median Age Total Population
Current Year Estimated Median Age Female Population
Current Year Estimated Median Age Male Population
Current Year Estimated Households
Current Year Estimated Families
Current Year Estimated Average Household Size
Current Year Estimated Percentage Of Householders Between The Age Of 15 & 24
Current Year Estimated Percentage Of Householders Between The Age Of 25 & 34
Current Year Estimated Percentage Of Householders Between The Age Of 35 & 44
Current Year Estimated Percentage Of Householders Between The Age Of 45 & 54
Current Year Estimated Percentage Of Householders Between The Age Of 55 & 64
Current Year Estimated Percentage Of Householders Between The Age Of 65 & 74
Current Year Estimated Percentage Of Householders Age 75 And Over
Current Year Estimated Average Household Income
Current Year Estimated Median Household Income
Current Year Estimated Median Family Income

Table 12: Claritas demographics data elements.¹ (continued)

Current Year Estimated Percentage Of Householders With Income Between \$0 & \$10,000
Current Year Estimated Percentage Of Householders With Income Between \$10,000 & \$15,000
Current Year Estimated Percentage Of Householders With Income Between \$15,000 & \$25,000
Current Year Estimated Percentage Of Householders With Income Between \$25,000 & \$35,000
Current Year Estimated Percentage Of Householders With Income Between \$35,000 & \$50,000
Current Year Estimated Percentage Of Householders With Income Between \$50,000 & \$75,000
Current Year Estimated Percentage Of Householders With Income Between \$75,000 & \$100,000
Current Year Estimated Percentage Of Householders With Income Over \$100,000
Current Year Estimated Percentage Of Householders With Income Under \$35,000
Current Year Estimated Percentage Of Householders With Income Over \$50,000
Current Year Estimated Median Value Owner-Occupied Housing Units
Current Year Estimated Population In Group Quarters
Current Year Estimated Percentage Of White Population
Current Year Estimated Percentage Of Black Population
Current Year Estimated Percentage Of Asian Population
Current Year Estimated Percentage Of Other Population
Current Year Estimated Percentage Of Hispanic Population
Current Year Estimated Avg Hhld Expend: Grocery Stores
Current Year Estimated Avg Hhld Expend: Eating Places
Current Year Estimated Avg Hhld Expend: Drinking Places
Current Year Estimated Avg Hhld Expend: Drug Stores (Including Proprietary Stores)
Current Year Estimated Avg Hhld Expend: Building Materials And Supply Stores
Current Year Estimated Avg Hhld Expend: Hardware Stores
Current Year Estimated Avg Hhld Expend: Retail Nurseries, Lawn & Garden Supply Stores
Current Year Estimated Avg Hhld Expend: Furniture And Home Furnishing Stores
Current Year Estimated Avg Hhld Expend: Household Appliance Stores
Current Year Estimated Avg Hhld Expend: Radio, Television, And Music Stores
Current Year Estimated Avg Hhld Expend: Department Stores
Current Year Estimated Avg Hhld Expend: Apparel Stores
Current Year Estimated Avg Hhld Expend: Shoe Stores
Current Year Estimated Avg Hhld Expend: Woman's Accessory Stores
Current Year Estimated Avg Hhld Expend: Auto Supply Stores
Current Year Estimated Avg Hhld Expend: Gasoline Service Stations
Current Year Estimated Avg Hhld Expend: Jewelry Stores
Current Year Estimated Avg Hhld Expend: Book Stores

Table 12: Claritas demographics data elements.¹ (continued)

Current Year Estimated Avg HHld Expend: Stationary Stores
Current Year Estimated Avg HHld Expend: Gift, Novelty, And Souvenir Shops
Current Year Estimated Avg HHld Expend: Florist
Current Year Estimated Avg HHld Expend: Hobby, Toy, And Game Shops
Current Year Estimated Avg HHld Expend: Sporting Goods Stores And Bicycle Shops
Current Year Estimated Avg HHld Expend: Camera And Photographic Supply Stores
Current Year Estimated Avg HHld Expend: Luggage Stores
Current Year Estimated Percentage Of Married Couples/Families
Current Year Estimated Percentage Of Married Couples/Families With Related Children
Current Year Estimated Percentage Of Married Couples/Families With Related Children Under Age 5
Current Year Estimated Percentage Of Married Couple/Families With Related Children Age 5-17
Current Year Estimated Percentage Of Population 25+ With Education Below 9th Grade
Current Year Estimated Percentage Of Population 25+ With Education Between 9th & 12th Grades With No Diploma
Current Year Estimated Percentage Of Population 25+ Diploma As Highest Education (Includes Equivalency)
Current Year Estimated Percentage Of Population 25+ Completing 2 Years Or Less Of College
Current Year Estimated Percentage Of Population 25+ With Bachelor's Degree As Highest Education
Current Year Estimated Percentage Of Population 25+ With Graduate Or Professional Degree
Current Year Estimated Persons In Occupied Units In Buildings With 2 To 4 U.I.S.
Current Year Estimated Persons In Occupied Housing Units With 5 To 9 Units In Structure
Current Year Estimated Persons In Occupied Units In Buildings With 10-19 U.I.S.
Current Year Estimated Persons In Occupied Units In Buildings With 20-49 U.I.S.
Current Year Estimated Persons In Occupied Housing Units with 50+ Units In Structure
Current Year Estimated Persons In Occupied Units Considered Mobile Homes Or Trailers
5-Year Projected Population
5-Year Projected Total Female Pop
5-Year Projected Housing Units
5-Year Projected Median Age Total Pop
5-Year Projected Median Age Female Pop
5-Year Projected Median Age Male Pop
5-Year Projected Households
5-Year Projected Families
5-Year Projected Average Household Size

Table 12: Claritas demographics data elements.¹ (continued)

5-Year Projected Percentage Of Householders Between The Age Of 15 & 24
5-Year Projected Percentage Of Householders Between The Age Of 25 & 34
5-Year Projected Percentage Of Householders Between The Age Of 35 & 44
5-Year Projected Percentage Of Householders Between The Age Of 45 & 54
5-Year Projected Percentage Of Householders Between The Age Of 55 & 64
5-Year Projected Percentage Of Householders Between The Age Of 65 & 74
5-Year Projected Percentage Of Householders Age 75 And Over
5-Year Projected Average Household Income
5-Year Projected Median Household Income
5-Year Projected Median Family Income
5-Year Projected Percentage Of Householders With Income Between \$0 & \$10,000
5-Year Projected Percentage Of Householders With Income Between \$10,000 & \$15,000
5-Year Projected Percentage Of Householders With Income Between \$15,000 & \$25,000
5-Year Projected Percentage Of Householders With Income Between \$25,000 & \$35,000
5-Year Projected Percentage Of Householders With Income Between \$35,000 & \$50,000
5-Year Projected Percentage Of Householders With Income Between \$50,000 & \$75,000
5-Year Projected Percentage Of Householders With Income Between \$75,000 & \$100,00
5-Year Projected Percentage Of Householders With Income Over \$100,000
5-Year Projected Percentage Of Householders With Income Under \$35,000
5-Year Projected Percentage Of Householders With Income Over \$50,000
5-Year Projected Median Value Owner-Occupied HUS
5-Year Projected Population In Group Quarters
5-Year Projected Percentage Of White Population
5-Year Projected Percentage Of Black Population
5-Year Projected Percentage Of Asian Population
5-Year Projected Percentage Of Other Population
5-Year Projected Percentage Of Hispanic Population
5-Year Projected Avg Hhld Expend: Grocery Stores
5-Year Projected Avg Hhld Expend: Eating Places
5-Year Projected Avg Hhld Expend: Drinking Places
5-Year Projected Avg Hhld Expend: Drug Stores (Including Proprietary Stores)
5-Year Projected Avg Hhld Expend: Building Materials And Supply Stores
5-Year Projected Avg Hhld Expend: Hardware Stores
5-Year Projected Avg Hhld Expend: Retail Nurseries, Lawn & Garden Supply Stores
5-Year Projected Avg Hhld Expend: Furniture And Home furnishing Stores

Table 12: Claritas demographics data elements.¹ (continued)

5-Year Projected Avg Hhld Expend: Household Appliance Stores
5-Year Projected Avg Hhld Expend: Radio, Television, And Music Stores
5-Year Projected Avg Hhld Expend: Department Stores
5-Year Projected Avg Hhld Expend: Apparel Stores
5-Year Projected Avg Hhld Expend: Shoe Stores
5-Year Projected Avg Hhld Expend: Woman's Accessory Stores
5-Year Projected Avg Hhld Expend: Auto Supply Stores
5-Year Projected Avg Hhld Expend: Gasoline Service Stations
5-Year Projected Avg Hhld Expend: Jewelry Stores
5-Year Projected Avg Hhld Expend: Book Stores
5-Year Projected Avg Hhld Expend: Stationary Stores
5-Year Projected Avg Hhld Expend: Gift, Novelty, And Souvenir Shops
5-Year Projected Avg Hhld Expend: Florists
5-Year Projected Avg Hhld Expend: Hobby, Toy, And Game Shops
5-Year Projected Avg Hhld Expend: Sporting Goods Stores And Bicycle Shops
5-Year Projected Avg Hhld Expend: Camera And Photographic Supply Stores
5-Year Projected Avg Hhld Expend: Luggage Stores
1990 Occupied Single, Detached Housing Units
1990 Occupied Single, Attached Housing Units
Housing Units Built 1985-1990
Housing Units Built 1980-1984
Housing Units Built 1970-1979
Housing Units Built Before 1970
Occupied Housing Units, Hhldr Moved In 1989 To March 1990
Occupied Housing Units, Hhldr Moved In 1985-1988
Occupied Housing Units, Hhldr Moved In 1980-1984
Occupied Housing Units, Hhldr Moved In Before 1980
Percentage Of Population Commuting To Work In 10 Minutes Or Less
Percentage Of Population Commuting To Work Between 10 And 19 Minutes
Percentage Of Population Commuting To Work Between 20 And 29 Minutes
Percentage Of Population Commuting To Work In 30 Or More Minutes
Percentage Of Households With No Vehicle

Table 12: Claritas demographics data elements.¹ (continued)

Percentage Of Households With One Vehicle
Percentage Of Households With 2 Or More Vehicles
2000 % Of Housing Units Built 1939 Or Earlier

1. Information in this table provided by Claritas.

Claritas PRIZM Codes

Table 13 lists the Claritas PRIZM Codes, which Sagent provides through a separate license.

Table 13: Claritas PRIZM Codes.¹

Cluster Number	Cluster Name	Short Description	Detailed Description
01	Blue Blood Estates	Elite Super-Rich Families	America's wealthiest suburbs are populated by established executives, professionals, and heirs to old money. They are accustomed to privilege and live in luxury, often supported by servants. A tenth of this group is a multimillionaire-millionaire. The next affluence level is a sharp drop from this pinnacle.
02	Winner's Circle	Executive Suburban Families	As its name implies, Cluster 2 is second in American affluence. Typified by new money, they live in expensive new mansions in the suburbs of major metros. They are well-educated, mobile executives and professionals with teen-aged families. They are big producers and prolific spenders who enjoy global travel.
03	Executive Suites	Upscale White-Collar Couples	Cluster 3 describes yesterday's Young Influentials, who are en route to becoming tomorrow's Winner's Circle. Many have married, and moved into condos or starter homes. Unique for S1, this Cluster is above average in pre-school kids. Although they rank well below Winners Circle in affluence, they are as well-educated, ambitious, and competent; they're just ten years younger.
04	Pools & Patios	Established Empty Nesters	Older, established couples in executive, professional, sales, and communication fields make up Cluster 4. Since many have reached their golden post-child years, there is a high index for dual incomes that, in turn, support a rich, busy life of travel, leisure activities, and entertainment.
05	Kids & Cul-de-Sacs	Upscale Suburban Families	Close to Executive Suites and Pools and Patios on all affluence measures, Cluster 5 is ranked first of all 62 PRIZM clusters in married couples with children, and large, 4 or more person families. Since family governs its lives and activities, Kids and Cul-de-Sacs is a noisy mix of bikes, dogs, carpools, and sports.
06	Urban Gold Coast	Elite Urban Singles & Couples	Cluster 6 is unique. It is the most densely populated per square mile, with the highest per-capita income, the greatest concentration of singles in multi-unit, high-rise buildings; the lowest incidence of auto ownership, and the fewest children. Cluster 6 is tops in urbanity, and over half of its population lives in New York City.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
07	Money & Brains	Sophisticated Townhouse Couples	Although Cluster 7 closely trails Urban Gold Coast in affluence measurement, it's very different. Upscale homes and condos on the urban fringe are owned by older, married couples who have few children. Since many enjoy dual incomes, they are sophisticated consumers of adult luxuries, travel, and entertainment.
08	Young Literati	Upscale Urban Singles & Couples	Although Cluster 8 is below Cluster 7 in affluence, it leads in education. A younger mix of executives, professionals, and students live in multi-unit apartments, condos, and town-houses near private urban universities. Having few children, these bon vivants are free to pursue their interests in art, fitness, and travel.
09	American Dreams	Established Urban Immigrant Families	Cluster 9 typifies the American Dream. Immigrants and descendants of multi-ancestries populate these multi-racial, multi-lingual neighborhoods. Cluster 9 tends to have big families, which is unique to Group U1. Multiple incomes from trades and public service have raised them to the second affluence level.
10	Bohemian Mix	Bohemian Singles & Couples	Although it's a short trip from the Upper East Side to the Village, the lifestyle and perspective shifts dramatically. Cluster 10 is America's Bohemia, a truly integrated mixture of executives, students, actors, and writers to whom live high-rises. This multi-racial, educated group is dominated by singles, and has the nation's second lowest index for children.
11	Second City Elite	Upscale Executive Families	Cluster 11 describes the movers and shakers of our second cities that are found coast to coast with its archetypical example in the wealthy enclaves of Huntsville, Alabama. Primarily married with teenage children, they give first attention to their families, homes, and clubs, then steal away to play in Europe.
12	Upward Bound	Young Upscale White-Collar Families	Young, college-educated, computer-literate, dual-income, frequent-flying executives and professionals describe those in Upward Bound. Most of this group is married, with pre- and school-aged children, and live in new, owner-occupied single family homes. They are found in over 100 TV markets that cover 75% of the total US population.
13	Gray Power	Affluent Retirees in Sunbelt Cities	Cluster 13 represents over two million senior citizens who have pulled up stakes, moved to the country or the Sunbelt, to retire among their peers. While these neighborhoods are found nationwide, almost half are concentrated in 13 retirement areas. They are health and golf fanatics with fat investment portfolios.
14	Country Squires	Elite Exurban Families	A private island off the coast of Maine; an elegant restored colonial village in the Berkshires; lush fenced-in horse farms in Leesburg, VA; or manicured gardens in Carmel by the Sea.... These are Cluster 14 neighborhoods where the wealthy have escaped urban stress to live in rustic luxury. Fourth in affluence, this group has big bucks in the boondocks.
15	God's Country	Executive Exurban Families	Many educated, upscale, married executives and professionals are raising their large families in the remote exurbs of major metros, the outskirts of second cities, and scenic towns. Multiple incomes support their affluence and life centers around family and outdoor activities. This is Cluster 15, in the second affluence decile.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
16	Big Fish Small Pond	Small Town Executive Families	Although similar to God's Country in size and national distribution, Cluster 16 ranks seven rungs down the affluence ladder. Also married and family-oriented, these older, ore conservative people are often captains of local industry. They enjoy investing in their homes and clubs, and take car vacations in the US.
17	Greenbelt Families	Young, Middle-Class Town Families	A rung below Big Fish, Small Ponds in affluence, Cluster 17 is smaller and more concentrated in our lesser second cities and uplands. This heavily mortgaged group is young, married, with lots of children. Their energies are devoted to family entertainment and outdoor sports.
18	Young Influentials	Upwardly Mobile Singles and Couples	Cluster 18 was hot in the eighties. Dubbed the Young Urban Professionals, these were the educated, high-tech, metropolitan sophisticates, the swingles and childless live-in couples, whose double incomes bought the good life in Boomtown USA. Then they married, and reduced Cluster 18 to half its former size. Here's what's left: The Last of the Yuppies.
19	New Empty Nests	Upscale Suburban Fringe Couples	Only three rungs down the affluence ladder, Cluster 19 is much more conservative than Young Influentials, and skewed to the northeast. Cluster 19 achieved its affluence through education and career accomplishments in numerous professions and industries. Most of them are married, in their post-child years, and have double incomes
20	Boomers & Babies	Young White-Collar Suburban Families	Cluster 20 ranks second of all PRIZM Clusters for married couples with children, and first for total households with children, many of whom are pre-schoolers. Skewed to the West, Cluster 20 is composed of executives and techies working in varied fields. Their relative youth and early careers place them at the bottom of the third affluence decile.
21	Suburban Sprawl	Young Midscale Suburban Couples & Singles	Multi-racial, multi-lingual neighborhoods are typically found in the centers of major metros. Cluster 21 is the exception, showing above average concentrations of native and foreign-born ethnics who have used education to become executives, administrators, and technicians. They have moved to the suburbs and the fourth affluence decile.
22	Blue-Chip Blues	Upscale Blue-Collar Families	For twenty years, Cluster 22 was one of the largest PRIZM Clusters. Dual income, high school-educated parents headed large suburban families, and topped the blue-collar ladder. During these two decades, their kids grew up and left, and blue-collar employment declined sharply. A smaller core remains, centered in the Great Lakes region.
23	Upstarts & Seniors	Middle Income Empty Nesters	Cluster 23 shows that youths and seniors are very similar if they're employable, single, and childless. In Cluster 23, they share average educations and incomes in business, finance, retail, health, and public service. Preferring to live in condos and apartments, Cluster 23 folks like the Sunbelt and the West.
24	New Beginnings	Young Mobile City Singles	Concentrated in the boomtowns of the Southeast, the Southwest, and the Pacific coast, Cluster 24 is a magnet for fresh starts. Populated by well-educated youths, many are minorities. Some are divorced, while many others are solo parents. The majority live in multi-unit rentals, and work in a variety of low-level, white-collar jobs.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
25	Mobility Blues	Young Blue-Collar/Service Families	In most of the same markets, but two deciles down in affluence, Cluster 25 is the blue-collar equivalent of New Beginnings: young, ethnically mixed, and highly mobile. Conversely, this Cluster shows high indices for Hispanics, large families with children. The military, industry, transportation, and public service are the primary employers of these breadwinners.
26	Gray Collars	Aging Couples in Inner Suburbs	For almost two decades, we read about the decline of the Great Lakes industrial Rust Belt. Decimated by foreign takeovers in the steel and automobile industries, the area lost a million jobs. Although most of the kids left, their highly skilled parents stayed, and are now benefiting from a major US industrial resurgence.
27	Urban Achievers	Mid Level, White-Collar Urban Couples	Due to its rank in the third decile of college education, Cluster 27 is the most affluent of the U2 Clusters. Often found near urban, public universities, these neighborhoods are ethnically diverse with a bi-modal, young/old age profile. Single students mix with older professionals in business, finance, and public service.
28	Big City Blend	Middle-Income Immigrant Families	High indices for Asians, Hispanics, and other foreign-born immigrants make Cluster 28 the most ethnically diverse in the U2 group. Skewed to the West, its affluence level drops two deciles from Urban Achievers. They have big families, are employed in an even mix of white- and blue-collar jobs, and live in old, stable, high density, urban rowhouse areas.
29	Old Yankee Rows	Empty-Nest, Middle-Class Families	Magnet neighborhoods for recent Asian and Latin American immigrants and centered in the Northeast, Cluster 29 is the most multi-lingual Cluster in U2. Although it's five affluence rungs below Big City Blend, Cluster 29 has the same white/blue-collar job mix, and tends toward singles living in multi-unit rentals.
30	Mid-City Mix	African-American Singles & Families	Cluster 30 is in the seventh affluence decile, geographically-centered in the Northeast and Great Lakes regions. Similar to all U2's, Cluster 30 shows above-average ethnic diversity and a mix of white- and blue-collar employment. These neighborhoods are two-thirds black, who live in urban row-house fringes with high college enrollments.
31	Latino America	Hispanic Middle-Class Families	Dominated by Latin Americans, with the nation's highest index for foreign-born immigrants, Cluster 31 is a giant step in achievement. They are concentrated in New York, Miami, Chicago and the Southwest in large young families with lots of children. Although they live in rented homes and have blue-collar jobs, they are moving up, and are college bound.
32	Middleburg Managers	Mid-Level White-Collar Couples	They keep the wheels rolling in our second cities: the business executives, bankers, doctors, lawyers, retailers, and city-hall officials. Half are older, married, post-child; half are younger, single, pre-child. Above-average incomes in all dollar brackets allow active leisure pursuits of clubs and sports.
33	Boomtown Singles	Middle Income Young Singles	Cluster 33 plays host to the youth of a hundred, fast-growing second cities in the Southern, Mid-West, and West. They are young professionals and techies in public and private service industries who live in multi-unit rentals, like music, and vacationing in the Caribbean.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
34	Starter Families	Young Middle-Class Families	Bucking recent trends, Starter Families opted for early marriage and parenthood. Here we see a higher index for blue-collar jobs, large families, and solo parents with young children. Many are living in natural beauty with a skew to the Pacific coast, the Rockies, and the northwestern Canadian borderlands.
35	Sunset City Blues	Empty Nests in Aging Industrial Cities	Equal to Starter Families in affluence, Cluster 35 describes older, skilled blue-collar, policemen, firemen, and technicians who have reached the end of their careers. Some retire to the mountains or St. Pete, but most stay home to rock on their porches near the Great Lakes and Mohawk Valley.
36	Towns & Gowns	College Town Singles	Many college towns and university campus neighborhoods are typically mixed with half locals (Towns) and half students (Gowns). Cluster 36 is composed of thousands of penniless 18-24 year olds and highly educated professionals, all with a penchant for prestige products that are beyond their evident means.
37	New Homesteaders	Young Middle-Class Families	Cluster 37 is the only T2 Cluster that shows above average college educations. Executives and professionals work in local service fields such as administration, communications, health, and retail. Most are married; the young have children, the elders do not. Life is homespun with a focus on crafts, camping, and sports.
38	Middle America	Midscale Families in Midsize Towns	Sitting atop the sixth affluence decile of the US median income, Cluster 38 is aptly named. These are family neighborhoods with a high index for married couples with children. They are busy with kids and dogs, and enjoy fast food, sports, fishing, camping, and watching TV. In approximate balance with the US population, they are found coast to coast.
39	Red, White & Blues	Small Town Blue-Collar Families	Just below Middle America in affluence, Cluster 39 is far more industrial and blue-collar, with skilled workers primarily employed in mining, milling, manufacture, and construction. Geo-centered in the Appalachians, Great Lakes industrial region, and Western highlands, these folks love the outdoors.
40	Military Quarters	GIs and Surrounding Off-Base Families	Since Cluster 40 depicts military life with personnel living in group quarters, its demographics are wholly atypical. Located on/or near military bases, Cluster 40 skews toward our principal harbors and defense perimeters. Fully integrated, and with the highest index for adults under 35, Military Quarters likes fast cars, bars, and action sports.
41	Big Sky Families	Midscale Couples, Kids & Farmland	With an average incidence for college educations, Cluster 41 has income levels well above the US median. They are well-paid, skilled craftsmen, machinists, and builders who live in scenic locales from New England to the Tidewater, in the Great Lakes region, and the Rockies. Family-centered lifestyles are devoted to hobbies, hunting, and boating.
42	New Eco-topia	Rural White/Blue-Collar/Farm Families	Found in the Northern Pacific, the Rockies, and northern New England, Cluster 42 is the only R1 Cluster with above-average college educations. New Eco-topia has an even mix of white/blue-collar jobs. A high index for personal computers reflects several new, high-tech industries in these pristine areas.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
43	River City, USA	Middle-Class Rural Families	Cluster 43 sweeps across New England and the Mohawk Valley, through the corn, grain, and dairy belts, to the Pacific orchards. Solid blue-collar citizens, in towns like Utica, NY; Zanesville, OH; and Butte, MT are raising sturdy, Tom-Sawyer-ish children in decent, front-porch houses. Yes, July 4th parades are still a big event in Cluster 43.
44	Shotguns & Pickups	Rural Blue-Collar Workers & Families	In the 8th decile, Cluster 44, the least affluent of the R1 Clusters, is found in the Northeast, the Southeast, and in the Great Lakes and Piedmont industrial regions. They lead the Group in blue-collar jobs; most are married with school-age kids. They are church-goers who also enjoy bowling, hunting, sewing, and attending car races.
45	Single City Blues	Ethnically-Mixed Urban Singles	Cluster 45 is found in most Eastern megacities, in the new West, and is third in the most singles in America. Often found near urban universities, Cluster 45 hosts a fair number of students. With very few children, it's a mixture of races, transients, and night trades, and is best described as a poor man's Bohemia.
46	Hispanic Mix	Urban Hispanic Singles & Families	Cluster 46 collects the nation's bi-lingual, Hispanic barrios, which are chiefly concentrated in the Atlantic metro corridor, Chicago, Miami, Texas, Los Angeles, and the Southwest. These neighborhoods are populated by large families with many small children. They rank second in percent foreign-born, and are first in transient immigration.
47	Inner Cities	Inner City, Solo-Parent Families	Concentrated in large Eastern cities and among America's poorest neighborhoods, Cluster 47 has twice the nation's unemployment; many residents are receiving public assistance. Eight out of ten households are African American and seven in ten households with children are single-parent families.
48	Smalltown Downtown	Older Renters & Young Families	Highly skewed west of the Mississippi, Cluster 48 has gained a flood of Eastern migrants who are mostly young and single. Often found near city colleges, Cluster 48 is populated with students and those looking for fresh starts and first jobs. They are employed as lower-echelon white-collar salespeople, clerks, and technicians.
49	Hometown Retired	Low-Income, Older Singles & Couples	Cluster 49 is three rungs down from Smalltown Downtown at opposite ends of the age range and geography. Except for some hot spots in the West, Cluster 49 lies mostly in the Appalachians and central Florida. It ranks third in singles, second in ages over 65, and first in retirement. They take bus tours, collect stamps, and play cards and chess.
50	Family Scramble	Low-Income Hispanic Families	Although Cluster 50 is found in many markets, it is centered across the Southwest and Pacific. It ranks third in Hispanic population, with an overlay of Native Americans. Ranked last in higher educations, Cluster 50 shows all the scars of poverty, but many are staying ahead with employment in transport, labor, and service jobs.
51	Southside City	African-American Service Workers	Mostly concentrated in the Southeast, the smaller Mississippi delta cities, the Gulf coast, and Atlantic states, Cluster 51 is very poor. Over 70% of its households are black. Although it's 61st in median household income, a low cost of living and a mix of labor and service jobs keep these families afloat.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
52	Golden Ponds	Retirement Town Seniors	Found coast to coast, Cluster 52 is a myriad of rustic towns and villages in scenic coastal, mountain, lake and valley areas, where seniors living in cottages retire among their country neighbors. Not as old, urban, or affluent as other retirees, a few play golf, but most prefer to adopt local customs.
53	Rural Industria	Low-Income, Blue-Collar Families	Cluster 53 is the most industrial of the T3 Clusters. Once dependent on railroads and major markets, 18 wheelers freed light industry to go farther afield to seek low-cost, non-union labor. It's found in Cluster 53 which is comprised of hundreds of blue-collar mill towns on America's rural backroads.
54	Norma Rae-ville	Young Families, Bi-Racial Mill Towns	Cluster 54 is geographically centered in the South, in the Mississippi delta, and in the Gulf coast and Atlantic states, which have become the center for our non-durable industries, such as clothing and home furnishings. With minimal educations, a black/white population mix, and unskilled labor, Cluster 54 falls in the ninth affluence decile.
55	Mines & Mills	Older Families, Mine & Mill Towns	Although equal to Norma Rae-ville in income, Cluster 55 is very different. Down the Appalachians, across the Ozarks to Arizona, and up the Missouri to the coal fields of Montana, Cluster 55 is exactly as its name implies. This older, mostly single population with few children lives in the midst of scenic splendor.
56	Agri-Business	Rural Farm-Town & Ranch Families	In census parlance, this title covers farming, forestry, fishing, ranching, mining, and other rural occupations. Consequently, Cluster 56 is more affluent and more skewed to the greater northwest from Lake Michigan to the Pacific. It is famous for very large families with lots of kids, countless animals, apple pie, and going fishing.
57	Grain Belt	Farm Owners & Tenants	Feeding America and sometimes the world, Cluster 57 is our breadbasket. Centered in the Great Plains and South Central regions, this Cluster shows a high index of Latino migrant workers. Life here is tied to the land, and ruled by the weather. Mostly self-sufficient, family- and home-centered, these families are poor only in money.
58	Blue Highways	Moderate Blue-Collar/Farm Families	On most maps, the interstates are red and the old highways are blue. Cluster 58 follows these remote roads through our mountains and deserts, and along the coasts and lake shores. These are R3's youngest neighborhoods, with its largest families, and the most children. They hunt and fish, love country music, camping, and attending tractor pulls.
59	Rustic Elders	Low-Income, Older, Rural Couples	Cluster 59 is the third most elderly Cluster in America, and has the lowest incidence of children in Group R3. It covers the nation, but is concentrated in the Great Plains and along the West coast. Although the lifestyle is pure country, the high indices for country clubs, powerboats, sailboats, volleyball, and health walks are surprising.
60	Back Country Folks	Remote Rural/Town Families	Cluster 60 is centered in the Eastern uplands along a wide path from the Pennsylvania Poconos to the Arkansas Ozarks. Anyone who visits their playgrounds in Branson, MO or Gatlinburg, TN can attest that these are the most blue-collar neighborhoods in America. Centered in the Bible Belt, many Back Country Folks are hooked on Christian and country music.

Table 13: Claritas PRIZM Codes.¹ (continued)

Cluster Number	Cluster Name	Short Description	Detailed Description
61	Scrub Pine Flats	Older African-American Farm Families	Cluster 61, the most geo-centric of all the Clusters, is mainly in the coastal flatlands of the Atlantic and Gulf states from the James to Mississippi rivers. These humid, sleepy rural communities with a mix of blacks and whites, live in a seemingly timeless, agrarian rhythm.
62	Hard Scrabble	Older Families in Poor Isolated Areas	Hard scrabble means to scratch a living from hard soil. Cluster 62 describes our poorest rural areas that reach from Appalachia to the Colorado Rockies, and from the Texas border to the Dakota badlands. The highest indices for Native Americans, mining occupations, and chewing tobacco are in Hard Scrabble.

1. Information in table provided by Claritas.

WinSplit

WinSplit can be used to split Use.gsd, Usw.gsd, Uste.gsd, Ustw.gsd, and Us.z9 files into smaller files. This is commonly done to speed processing and to save disk space when only a specific geography is needed. WinSplit divides GSD and Z9 files by ZIP Code, SCC (3-digit ZIP), MSA, or State, and remembers the last extract performed, which makes extracting the updates an easier process.

For information on using WinSplit, refer to the *Sagent Utilities Reference Manual*.

Note: This function is not supported for Canadian addresses.

Centrus Data Formatter and ASCII Text Files

Centrus Data Formatter is a 32-bit Microsoft Windows application designed to automate the process of creating and updating format files. Centrus Data Formatter reads field and record information from the existing format file, if there is one, or from the actual ASCII text file. This information is then displayed, along with a section of the first part of the text file so that you can see how the data is interpreted. You can accept Centrus Data Formatter's settings or make changes as necessary. When you're satisfied that your data is correctly described, Centrus Data Formatter updates any existing format file or creates a new one. Your ASCII text file is then ready to process.

See the following sections for more information:

- About ASCII Text Files
- Using Centrus Data Formatter

About ASCII Text Files

ASCII files are document files encoded in a universally recognized text format. This type of file is useful for transferring data between programs and computing platforms that could not otherwise understand each other's file formats. Most data and text processing programs give you the option of exporting files in ASCII text format, usually with a file name extension of **.TXT**.

You can describe an ASCII text file containing addresses or other structured data by answering two questions:

- How are the fields and records defined?
- What, if any, is the end of line character?

Defining Fields

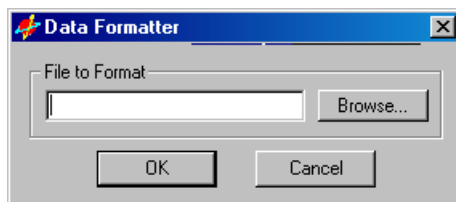
Fields are typically defined in one of two ways. *Delimited files* contain information separated by special characters, called delimiters. For example, the comma and tab characters are commonly used delimiters. The width of the fields defined by the delimiters can vary. *Fixed width* files, by contrast, contain fields that are always a constant width and are arranged in columns. The width of each field must be defined.

Some additional information is needed to complete the description of the file's fields. Each field may contain either *numeric* or *character* data. If the data is numeric, the number of decimal places must be entered. Each fixed-width field should have a name, type (numeric or character) and width. Delimited fields need only name and type. Finally, fixed width or delimited files may have field names rather than data in the first row or record.

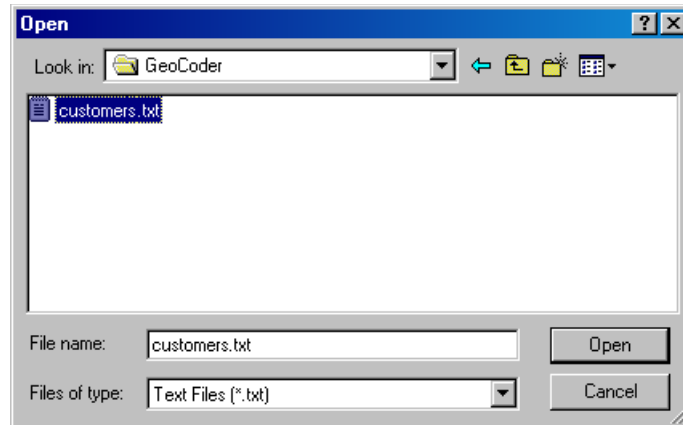
Delimited files also define field widths, which are useful when importing files to another format. Note that the entire field value is always read, regardless of the defined width.

Using Centrus Data Formatter

When you start Centrus Data Formatter, the **Data Formatter** dialog box appears.

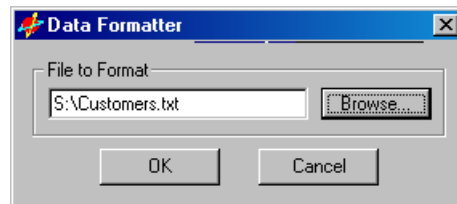


1. Enter the path and file name of the file you want to format or click **Browse...** to search for the file. The **Open** dialog box appears.



2. Select a text file and click **Open**. Centrus Data Formatter looks for the selected text file's format file—a file with the same name as the selected text file, but with an extension of .FMT. If Centrus Data Formatter finds a format file, it uses the information contained in the format file to read the text file. If Centrus Data Formatter does *not* find a format file, it attempts to extract format information directly from the text file.

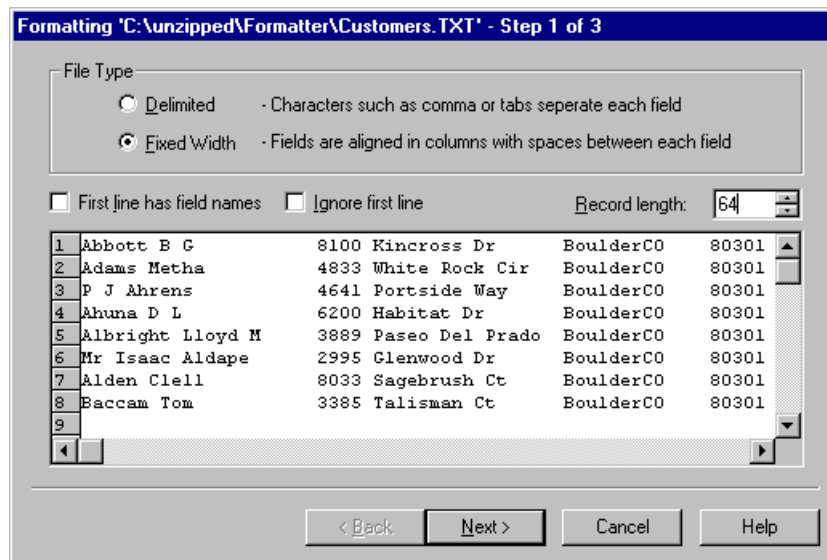
The **Data Formatter** dialog box reappears with the file you selected displayed.



3. Click **OK**.

Step 1

The **Formatting 'filename' – Step 1 of 3** dialog box appears.



When the Data Formatter opens the file, it attempts to determine the **File Type** – either *Fixed* or *Delimited*. The data in the Preview window at the bottom of the dialog is displayed differently, depending on what type of file you are using.

If there are no line terminators in the file, the Data Formatter prompts you to enter the record length. You can either select a record length or have the Data Formatter calculate it for you by searching for the most likely logical breaks in the file.

1. Accept the default **FileType**, or change if necessary.
2. Accept or change the following values:
 - **First line has field names** - the Data Formatter searches the first line of the file for field names.
 - **Ignore first line** - the Data Formatter ignores the first line of the file away without checking it for field names.
 - **Record length (for fixed width records only)** - the length of the records in the file. Valid values are between 1 and 8192 characters.

In a file with no line terminators, the Data Formatter re-processes the file with the new **Record length** and re-displays the records.

In a file with line terminators, the Data Formatter increases or decreases the size of the last field depending on the value in **Record length**.

The Preview window at the bottom of the dialog box shows a portion of the text file, with the first 30 records shown as columns and rows. You can view additional fields and records by scrolling.

In a file with no line terminators, each row is displayed with the exact number of characters as the **Record length** value. You can visually

determine the record length of non-terminated records by looking at the rows displayed.

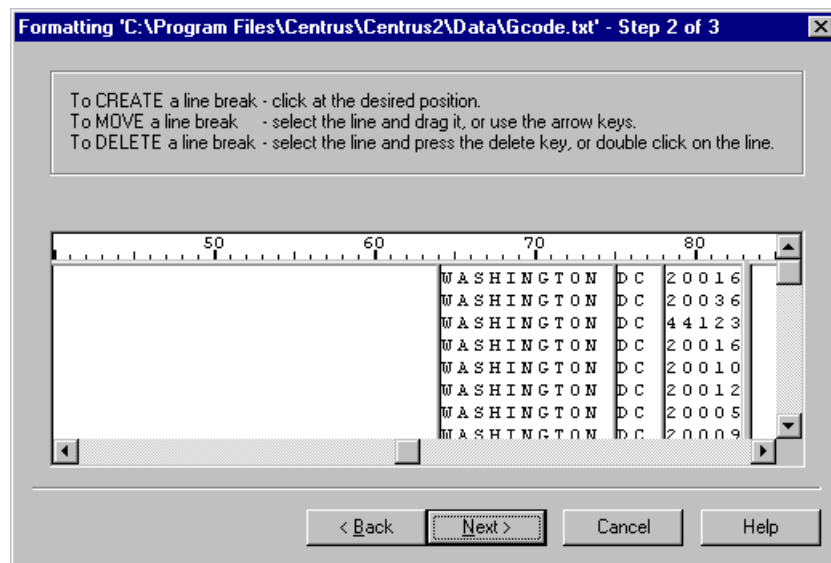
In a file with terminators, records are parsed and the width of the last field is affected by the **Record length** value.

3. Click **Next>**. The **Formatting 'filename' - Step 2 of 3** dialog box appears.

Step 2

Fixed Width

The following dialog box appears if the file's **File Type** is *Fixed Width*.

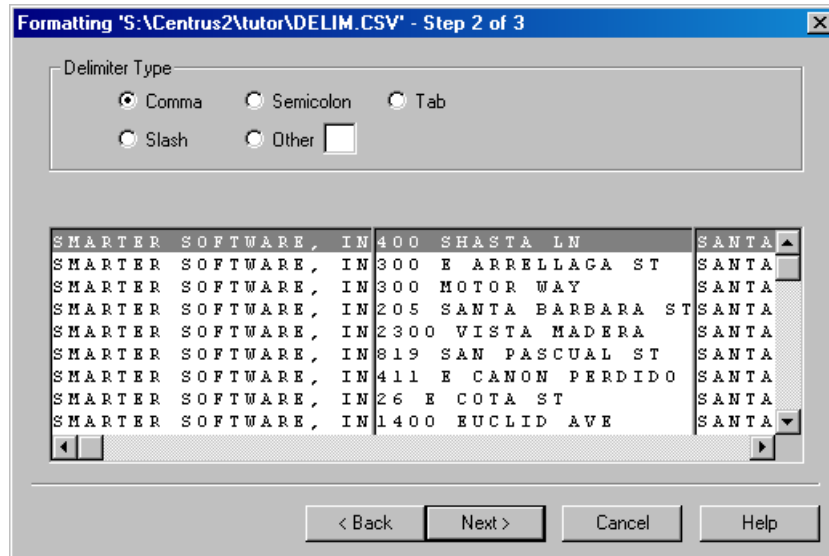


As you can see, the field breaks that Centrus Data Formatter determined are shown in the Preview window of this dialog box.

1. Follow the directions at the top of the dialog box to Create, Move, or Delete a line break.
2. Click **Next>**. The **Formatting 'filename' - Step 3 of 3** dialog box appears.

Delimited

If the file is *Delimited*, the dialog box shown below appears.



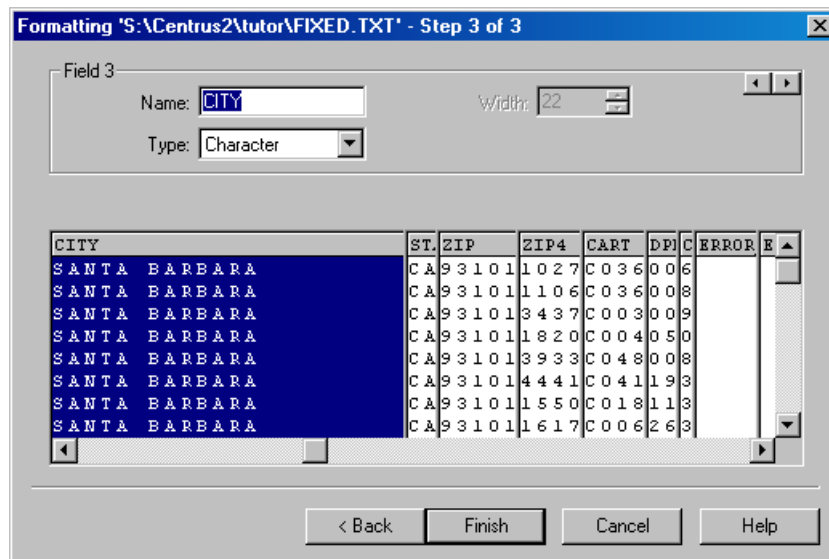
1. Accept the default **Delimiter Type** or change as needed.

You can enter any punctuation character as a delimiter in *Other*, except the characters for space (), quote ("), period (.), minus (-), or underscore (_).

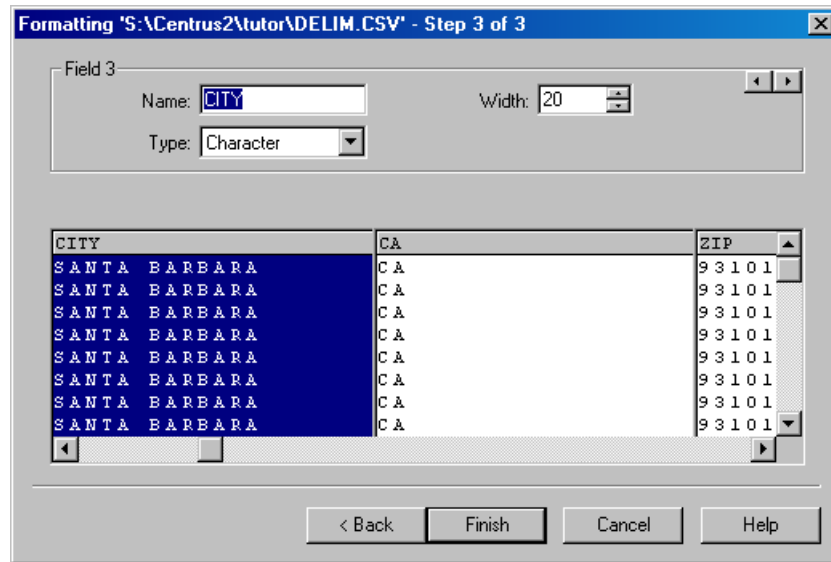
2. Click **Next>**. The **Formatting 'filename' - Step 3 of 3** dialog box appears.

Step 3

The following dialog box appears if the file's **File Type** is *Fixed Width*.



The following dialog box appears if the file's **File Type** is *Delimited*.



1. Select fields to change either in the Preview window or using the left/right arrows at the top of the dialog box.
2. Change the **Field Name** (helpful if you will be mapping fields in Centrus Desktop). Field names must begin with an alphabetic character or underscore, following by alphanumeric or underscore characters.
3. Make the following changes as needed:
 - **Type** - field type, either Character or Numeric.
 - **Decimals (Numeric fields only)** - the number of decimal places in the numeric field.
 - **Width (Delimited files only)** - changes the width of the selected field, padding or truncating the field. Changing this value also changes the **Record length**. Valid values are 1 - 2048.
4. Click **Finish** to write the .FMT file.

Note: If an .FMT file already exists, it will be overwritten.

About Format Files

The format file created by Centrus Data Formatter has the same filename as the ASCII address file, but uses the extension **.FMT** rather than **.TXT**. Since Centrus Data Formatter automates the creation and editing of format files, there's no need for you to "hand edit" format files. But understanding the structure of a format file may aid you in using Centrus Data Formatter. The basic structure is:

```
<0 or more lines of comments, each line beginning with a slash (/) or a
  semicolon (;)>
TYPE=<delimiter type>
EOL=<end of line character>
FLN=<if the first line contains field names> (Line not present otherwise)
<field name>TAB<field type>TAB<field width>.<number of decimal places>
```

Delimited File Example

A sample format file for a delimited text file is shown below.

```
; Comments begin with semicolons (;) or slashes (/).
; Blank lines are ignored.
; Size is not required, as this is a comma delimited
file.
```

```
TYPE=Comma
```

```
EOL=CRLF
```

```
FLN=N
```

First_Name	Character	15
Last_Name	Character	15
Address	Character	20
City	Character	15
State	Character	5
ZIP	Character	5

Sample data records are shown below. Note that all fields are not present for each record, but blank fields between two existing fields are still delimited. Note, too, that each field is enclosed by quotation marks, indicating literal strings. This prevents any commas that may be present within fields from being used as delimiters.

```
"Bob", "Doe", "123 Main St", "Nome", "AK", ""
"Carol", "Jones", "321 28th Ave, Apt. 22", "", "", "80301"
"Ted", "Tyler", "345 Mission St", "Longmont", "CO", "80503"
"Alice", "Brown", "PO BOX 123", "Boulder", "CO", ""
```

Fixed Width File Example

A sample format file for a fixed width text file is shown below.

```
; Comments begin with semicolons (;) or slashes (/).
; Blank lines are ignored.
; Field size IS required; this is a fixed width file.
; Note that numeric fields have decimals defined.
TYPE=Fixed
EOL=CRLF
FLN=N
First_Name      Character      15
Last_Name       Character      15
Address         Character      20
City            Character      15
State           Character      5
ZIP             Character      5
Latitude        Numeric        11.6
Longitude       Numeric        11.6
Block_Grp       Character      10
Match_Cd        Character      10
Loc_Code        Character      10
```

A portion of some sample data records is shown below. Note that all fields are not present for each record. Blank fields between two existing fields are filled with blank spaces, preserving the layout of the data.

Bob	Doe	123	Main St	Nome	AK	
Carol	Jones	321	28th Ave			80301
Ted	Tyler	345	Mission St	Longmont	CO	80503

CASS Overview

CASS stands for the USPS Coding Accuracy Support System. It is a service offered to mailers, service bureaus, and software vendors that improves the accuracy of delivery point codes, ZIP+4 Codes, 5-digit ZIP Codes, and carrier route information on mail. CASS provides a common platform to measure the quality of address matching software and useful diagnostics to correct software problems.

Using CASS to Reduce USPS Bulk Mailing Rates

The United States Postal Service requires that any bulk mailing fit specific guidelines. The *Domestic Mail Manual* (DMM), published quarterly by the USPS, contains the complete list of these guidelines.

In general, the pieces to be mailed must be sorted and arranged according to USPS guidelines. They must be of the same basic size and conform to size and shape requirements. The mail must also have the correct markings to be read by automated sorting equipment. Each piece of mail must also have the correct ZIP Code.

An annual bulk mail permit fee is paid at the post office that receives the mailings. The mailings must also be delivered to that specific post office.


If all the USPS guidelines are followed, you can obtain bulk mail discounts. To receive even greater discounts on postage, the address list can be processed with USPS-certified address standardization software. This process is referred to as CASS certification. To qualify for the CASS certification discount, you must generate a CASS report with the standardization software during processing and submit the completed form with the mailing.

The CASS 3553 Report

USPS-certified address standardization software generates the USPS 3553 report during mailing list processing. This report is commonly referred to as the CASS report. To generate a CASS report, the software compiles statistics during processing regarding the number of 5-digit ZIP Codes, ZIP+4 Codes, and Carrier Routes assigned to records in the address file. The application writes this information, along with other information required by the USPS, to an ASCII text file. This text file is the CASS report. See Figure 2 on page 197 for an example of a blank CASS 3553 report.

Before submitting this report to the USPS along with the mailing, remember to fill out the sections that apply to your organization, such as the name of the company processing the list and the organization doing the mailing.

Figure 2: Representation of a blank ASCII CASS 3553 report

 **UNITED STATES POSTAL SERVICE®**

Coding Accuracy Support System (CASS) Summary Report

This form may be generated as the output of address matching processing using CASS (Coding Accuracy Support System) certified software in conjunction with current USPS Address Database Files. Any facsimile must contain the same information in the same format as this printed form. See DMM A950 for more information.

A. Software			
CASS - A1	1. CASS Certified Company Name	2. CASS Certified Software Name & Version	3. Configuration
	4. Z4 Change Certified Company Name	5. Z4 Change Certified Software Name & Version	6. Configuration
	7. eLOT Certified Company Name	8. eLOT Certified Software Name & Version	9. Configuration
	10. DPC Utility Certified Company Name	11. DPC Utility Software Name & Version	12. Configuration
MASS - A2	1. MASS Certified Company Name	2. MASS Certified Software Name, Version, & Model No.	3. Configuration
			4. MLOCR Serial No.

B. List		
1. List Processor's Name	2. Date List Processed	3. Date of Database Product Used
	a. Master File	a. ZIP + 4 File
	b. Z4 Change	b. Z4 Change
	c. eLOT	c. eLOT
	d. CRIS	d. CRIS
4. List Name or ID No.	5. Number of Lists	6. Total Records Submitted for Processing

C. Output					
Output Rating	1. Total Coded	2. Validation Period		1. Total Coded	2. Validation Period
a. ZIP + 4 Coded ▷		From To	d. 5-Digit Coded ▷		From To
b. Z4 Change Processed ▷			e. CR RT Coded ▷		From To
c. DPBC Assigned ▷		From To	f. eLOT Assigned ▷		From To

D. Mailer		
I certify that the mailing submitted with this form has been coded (as indicated above) using CASS-certified software meeting all of the requirements of <i>Domestic Mail Manual A950</i> .		3. Name & Address of Mailer
1. Mailer's Signature	2. Date Signed	

E. Qualitative Statistical Summary (QSS)						
For informational purposes only: QSS is solely made available for the list processor's review and analysis. This information is not to be considered by the Postal Service personnel in determining rate eligibility under any circumstances. See reverse for a detailed explanation.						
High Rise Default	High Rise Exact	Rural Route Default	Rural Route Exact	LACS	EWS	DPV

PS Form 3553, March 2002 (Page 1 of 2) This form on internet at ribbs.usps.gov/files/cass/

For more information about CASS visit the USPS National Customer Support Center (NCSC) Web site at <http://www.usps.gov/ncsc/> or call the NCSC at 1-800-238-3150.

The USPS ZIPMove file, us.gsz, must be in your search path to generate a CASS report. For information about specifying a search path, see “Specifying the Search Path for the Data Files” on page 7.

Where to Get Help

The USPS has a vested interest in seeing that you get the assistance you need in order to take advantage of bulk mail discounts. By properly following the bulk mail guidelines, you save the USPS a large amount of effort when delivering your mailing.



To receive a *Domestic Mail Manual* (DMM) subscription, call 202-783-3238. This is a quarterly publication and its purpose is to list the specific requirements for ALL types of domestic mailings. A company called WindowBook has an on-line, hypertext version of the DMM, as well as a free DMM utility disk. Call WindowBook at 1-800-370-2410 for more information.

The USPS runs over 100 Postal Business Centers, with at least one center in each state. These offices are staffed exclusively to assist the small mailer and can provide a wealth of information and assistance. Your local post office should have the address and telephone number of the PBC nearest you. The local post office can also provide an assortment of publications targeted towards designing and preparing bulk mailings.

The USPS also has a National Customer Support Center. Their phone number is 1-800-238-3150. The staff at the NCSC are very helpful and can answer, or find answers, to all questions you may have concerning bulk mail.

Finally, many useful publications and mailing-related information can be obtained from the USPS Web site at <http://www.usps.gov>.

Audit and Log Report Samples

The Audit report lists the record as it was input, and as it was processed by Centrus Desktop, along with the Match and Location codes. This presents an overview of what Centrus Desktop accomplished on a certain set of records. You can set the audit frequency to sample a record at defined intervals, for example, every 1000 records.

The Log report returns a complete list of the settings used to process the file. It also contains statistics on how clean the file was coming in and how many records were processed by Centrus Desktop. This report can be displayed in Windows Notepad (if Notepad is available) by checking the **View log file when processing is completed** check box on the **Table** tab.

Audit Report

Below is an example of an audit report. The report frequency was set to 1000. The second entry in the example is a building match. Town and Country Village is at 855 El Camino Real, and 360 is assigned as the unit number.

Record Number: 1000

Input Address

443 ROBIN
WAVERLY OH

Output Results

443 Robin Rd
Waverly, OH 45690-1557

Match Code: S91 ZIP and ZIP+4 were changed.
Street type was changed.

Location Code: AS0 Exact Address Match.

Record Number: 2000

Input Address

360 TOWN AND COUNTRY VLG
PALO ALTO CA 94301

Output Results

855 El Camino Real Ste 360
Palo Alto, CA 94301-2326

Match Code: A89 ZIP+4 was changed.
Street name and Street type were changed.

Location Code: AS0 Exact Address Match.

Log Report

Here is an example of a log report. This report returns a complete list of the settings used to process the file. It also contains statistics on how clean the file was coming in and how many records were processed by Centrus Desktop.

Centrus Report

Wednesday, February 23, 2000

SETUP

```
-----
Centrus Desktop version..... 3.00
Licensed to..... Sagent
Company..... Sagent Technology,
  Inc.
Serial Number..... 2998

Licensed area..... US
Licensed options..... Standardization,
  Geocoding, Demographics, Extended Alias
License Expiration..... No Expiration

Geocoding Data Build Date..... 12/15/1999 (TIGER)
CASS Expiration Date..... 03/29/2000
```

FILES

```
-----
Input File Name..... D:\GSCData\Canydm.txt
In-Place Update..... NO
Output File Name..... D:\GSCData\canout.txt
Reject File Name..... none
Audit File Name..... none

Processing started at..... 02/23/00 15:11:41
Processing finished at..... 02/23/00 15:12:52
Records per hour..... 311983
```

ADDRESS CODING

```
-----
Search Path for Address Geocoding.....
  g:\cd2tiger;g:\cd3tiger;C:\Program
  Files\Centrus\Centrus2\Data;C:\Program
  Files\Centrus\Centrus2
ZIP Code Centroid File..... G:\CD3tiger\Z9\Us.z9
```

Options:

```
Address Matching..... YES
  Matching Tolerance..... Close
ZIP Centroid Matching..... YES
  Minimum Quality..... Best Centroid
Negate Longitudes..... YES
Census ID Content..... Block Group
Census ID Format..... Census
Mixed Case Output Addresses..... NO
Country Preference..... Automatic
Produce SERP Report..... YES
```



```

SERP Report File..... G:\SerpRep.Prn
Produce USPS 3553 (CASS) Report..... NO
CASS Report File..... G:\cass.prn

Process if Field is Blank: NO
Blank Field..... <none>

Input ADDRESS FIELDS:
Firm..... <none>
Address..... Address
Address 2..... <none>
City..... City
State..... State
ZIP Code..... ZIP
ZIP + 4 ..... <none>
COUNTRY ..... Country

Assigned Address Outputs:
Blank if Unmatched..... NO
Country..... N_COUNTRY
Postal Code..... N_PSTCODE

ADDRESS STATISTICS:
Total records in file..... 6153
Records Skipped..... 0
Records Processed..... 6153
  Addresses Matched..... 74 (1.20%)
    Standardized..... 73
      Skipped by Z4Change..... 0
      Records actually processed..... 73

  Intersections..... 0
  Non-USPS..... 1

Unmatched..... 21 (0.34%)
  E001: Low level error..... 0
  E002: State file not found..... 0
  E003: Invalid file..... 0
  E004: File is out of date..... 0
  E010: Unrecognized last line..... 0
  E011: ZIP not found..... 10
  E012: City not found..... 0
  E013: City not unique..... 5
  E014: Out of licensed area..... 0
  E020: No matching streets..... 0
  E021: No cross streets..... 5
  E022: No matching ranges..... 0
  E023: Match is unresolved..... 1
  E025: Too many cross streets..... 0

Standardized Address Quality:
  Original Address Unchanged..... 92.23%
  Original Last Line Unchanged..... 70.95%
  Corrected prefix direction..... 4
  Corrected street name..... 2
  Corrected street suffix..... 16
  Corrected postfix direction..... 1
  Corrected city name..... 5
  Corrected state abbreviation..... 0

```

Corrected ZIP Code.....	8
Corrected ZIP+4 add on.....	73
Geocodes Assigned.....	78 (1.27%)
Address.....	59 (0.96%)
ZIP Centroids.....	19 (0.31%)
Location Accuracy:	
ZIP+4.....	5
ZIP+2.....	1
ZIP Code.....	13
Census Accuracy:	
Block Group.....	6
Census Tract.....	2
County.....	11

Note: All percentages are based on records processed.

CANADIAN ADDRESS STATISTICS:

Addresses Matched.....	4378 (71.15%)
Standardized.....	4378
Unmatched.....	1680 (27.30%)
EC01: Internal error.....	0
EC10: No Lastline info found.....	0
EC20: No matching addresses.....	1
EC22: Missing or wrong street range..	0
EC23: Could not resolve address.....	501
EC25: Inconsistent address.....	64
EC26: Missing or wrong box range.....	508
EC27: Missing or wrong unit range....	286
EC40: Address/Postal Code conflict...	6
EC41: Multiple street ranges.....	193
EC42: Change of delivery mode.....	0
EC43: Exceeded address changes.....	0
EC44: Address type change not allowed	3
EC50: Minor error.....	0
Standardized Address Quality:	
Original Address Unchanged.....	379 (6.16%)
Corrected Postal Code.....	454
Corrected Municipality.....	3478
Corrected Province.....	3391
Corrected street type.....	476
Corrected street postfix.....	444
Corrected house number,,,,,.....	0
Corrected street name.....	829
Corrected for unit, PO Box, General Delivery or Route	
Corrected Keyword.....	954
Corrected Number.....	1161
Corrected Station Type.....	24
Corrected Station Name.....	0
Geocodes Assigned.....	4062 (66.02%)

Note: All percentages are based on records processed.

Canadian Address Accuracy Program

The Address Accuracy Program is a CPC initiative that encourages accurate mail addressing. Customers who adhere to the requirements of the program avoid the increased mailing rates the CPC charges for inaccurately addressed mail. Centrus Desktop processes Canadian addresses according to the requirements of this program. The Software Evaluation and Recognition Program (SERP) is a testing program used to evaluate address matching software packages for their ability to perform address validation and address validation and correction according to the Address Accuracy guidelines. For information about SERP address validation and correction criteria, refer to CPC's document titled, *Software Evaluation and Recognition Program (SERP) Handbook*.

Statement of Accuracy

CPC requests that customers keep a Statement of Accuracy (SOA) on file. Centrus Desktop produces an SOA when you select **Produce Canada Post SERP Report** on the **Address Coding Options** tab as described in "Postal Reports Option" on page 20. The SOA is valid for one year, so your address records must be reprocessed at least yearly. CPC recommends that you generate an SOA as part of your regular process to ensure deliverability. The following figure shows an example of an SOA.

Figure 3: Sample SOA

Address Accuracy Statement	
1.	User's Firm Name Here User's Address Line Here User's Last Line Here
2.	Customer CPC Number: -999999
3.	Total Number of Records Processed: 31621
4.	Address Accuracy Level: 95.2
5.	Processing Date: 2001/08/16
6.	Address Accuracy Expiry Date: 2002/08/16
7.	Effective Date of Postal Code Address Data 2001/02/19
8.	Geostan Canada (Version 21.11)
	Sagent Technology, Inc. - 1999-2001

Valid and Correctable Addresses

According to the *SERP Handbook*, a valid Canadian address meets all of the following criteria:

- All required address elements are present (refer to the preceding tables for the required elements).
- No required element that is found on CPC's Postal Code Address Data Files is missing.
- The address provides a match for one and only one address on CPC's Postal Code Address Data Files.
- All required address elements are presented without ambiguity.

An input address is correctable when both of the following are true:

- There are one or more elements missing or inconsistent from an otherwise valid address.
- Only one address can be derived from the information provided.

Where to Get Help

The Addressing Policy and Programs function at the CPC Head Office is responsible for the development and implementation of the Address Accuracy Program. You can reach Address Accuracy staff by calling 1-800-363-3459. Address Accuracy Coordinators have been appointed within each region. Your CPC representative can provide you with the name of the Address Accuracy Coordinator within your region.

Glossary

A

Address standardization

Address standardization is the process of taking an address and verifying that each component meets USPS guidelines for addresses. For example, “123 Main Avenue” should be abbreviated as “123 Main Ave.” During standardization, minor misspellings, dropped components, and abbreviations are all corrected. The correct city, state, and ZIP are also provided.

B

Buffer

An area drawn around an object. Buffers can be used to test whether points are “near” other objects, such as points, lines, or polygons.

C

CASS

USPS Coding Accuracy Support System. A service offered to mailers, service bureaus, and software vendors that improves the accuracy of delivery point codes, ZIP+4 Codes, 5-digit ZIP Codes, and carrier route information on mail. CASS provides a common platform to measure the quality of address matching software and useful diagnostics to correct software problems. CASS Certified mailings qualify for substantial postage discounts.

Census ID	Also called a Census FIPS Code. A Census ID uniquely identifies each piece of Census geography.
Centroid	The calculated center of an area. The coordinates that define a centroid are the average of the sets of coordinates that describe the area.
Closest site analysis	Also known as Radial Analysis. Determines distances between two sets of points.
Cluster code	Cluster codes define neighborhoods in terms of demographically and behaviorally distinct types, or “clusters.”

D

Datum	A mathematical model of the Earth used to calculate the coordinates on any map, chart, or survey system. Surveyors take an ellipsoid model of the Earth and fix it to a base point. The North American Datum (NAD) is the official reference ellipsoid used for the primary geodetic network in North America.
--------------	--

E

eLOT	The Enhanced Line of Travel (eLOT) Product was developed to provide mailers the ability to sort their mailings in approximate carrier-casing sequence. To aid in mail sorting, eLOT contains an eLOT sequence number field and an ascending/descending code. The eLOT sequence number indicates the first occurrence of delivery made to the add-on range within the carrier route, and the ascending/descending code indicates the approximate delivery order within the sequence number. Mailers can use eLOT processing to qualify for enhanced carrier route presort discounts.
-------------	---

F

Finance Area	A USPS Finance Area is an area defined by the Postal Service to collect cost and statistical data. It is frequently used for area searches, since it covers some or all of the ZIP Code areas in a city or town.
---------------------	--

FIPS code	<p>A Federal Information Processing Standards (FIPS) Code, also called a Census ID, uniquely identifies each piece of Census geography. Centrus Desktop assigns County, Census Tract, and Block Group, all of which are Census geographies. The syntax of the FIPS code is as follows:</p> <p>ssccctttt.ttgbbb</p> <p>where</p> <p>ss is the two-digit State FIPS Code</p> <p>ccc is the three-digit County FIPS Code</p> <p>tttt.tt is the six-digit Census Tract FIPS Code</p> <p>g is the single-digit Block Group FIPS Code</p> <p>bbb is the Block FIPS Code</p>
G	
Geocode	<p>Geocoding is the process of assigning data based upon location information. Centrus Desktop uses the address or ZIP to assign latitude, longitude, and Census FIPS information. Geocoding, coupled with demographic assignment or mapping software, can yield new insights regarding the address information.</p>
GIS	<p>Geographic Information System. A computer-based tool for enhancing geographic data by analyzing both the physical location in space and the set of characteristics associated with a location.</p>
L	
Layers	<p>GIS applications separate the different types of information into data “layers.” For example, store locations might be defined in on one layer and main roads on another. This allows for separate display and processing when necessary but does not prevent cross referencing between data layers during query and analysis. Layers are referenced to a common spatial domain so that they can be scaled and overlain in such a way that any given reference point can be located on any of the layers and the data value extracted.</p>
Location code	<p>Location codes indicate the accuracy of the assigned geocode.</p>

N

NAD27	NAD27 has its origin at Meades Ranch, Kansas. NAD 27 does not include Alaskan islands and Hawaii. Latitudes and longitudes that are surveyed in the NAD27 system are valid only in reference to NAD27 and do not tie to any maps outside the U.S.
NAD83	NAD83 is earth-centered and defined with satellite and terrestrial data. NAD83 is compatible with the World Geodetic System 1984 (WGS84), the terrestrial reference frame associated with the NAVSTAR Global Positioning System (GPS) now used extensively for navigation and surveying.

P

Point-in-Polygon analysis	Point-in-Polygon analysis determines which polygon or polygons a point lies within.
----------------------------------	---

S

SERP	Software Evaluation and Recognition Program. This program was developed by the DMR Group, a contractor for the CPC, to provide a method for verifying the accuracy level of addresses.
SOA	Statement of Accuracy. This report is required by the CPC as verification of the number of records processed and their address accuracy level. The SOA is valid for one year.

Z

ZIP Code	Zone Improvement Plan Code. Established in 1963, the system of 5-digit codes that identifies the individual post office or metropolitan area delivery station associated with an address. ZIP+4 is an enhanced code consisting of the 5-digit ZIP Code and four additional digits that identify a specific range of delivery addresses.
-----------------	---

Index

A

- abbreviations, state 69, 70
- Access files 86
- accuracy
 - location 112
 - score 116
- address
 - available data elements 110
 - cannot detect 106
 - current 68
 - data 107, 110, 126
 - elements 107, 108, 110
 - file 196
 - geocoding 14
 - input fields 104
 - mixed case output 19
 - output fields 107
 - quality 102
 - record 68
 - standardization 14, 196, 205
 - street 77, 102, 106
 - undeliverable 102
 - unmatched 152
- Address Accuracy Program (Canada)
 - correctable addresses 204
 - described 203
 - getting help 204
 - Statement of Accuracy 203
 - valid addresses 204
- address coding
 - CASS mode files 9
 - CASS reports 20
 - data files 8
 - files for Canadian address
 - processing 10
 - geocoding options 14
 - inputs, specifying 44
 - intersection matching files 10
 - NAD conversion files 10
 - options 12
 - outputs, specifying 44
 - overview 101
 - postal reports 20
 - SERP reports 20
 - settings, verifying 37
 - U.S. addresses 18

- Address Coding module
 - overview x
 - using 101
- address location codes 161
- Address Wizard mode 65
- Adobe Acrobat Reader, installing 3
- Alaskan islands, NAD27 208
- analysis
 - closest site 206
 - site 48, 139
 - Spatial 18, 139
- apartment numbers 106
- ASCII
 - characters, range 171
 - files 86, 114
- Atlas GIS 23, 131, 138
- audit report
 - example 199
 - overview 90
 - specifying 91

B

- bar code, delivery point 20, 110, 111
- Batch Indexer utility 81
- batch processing 150
- batchind.exe 81
- best centroid 16
- block
 - city 71
 - code 110
 - group 16, 110
- Block Group
 - code 110
 - ZIP centroid matching accuracy 16
- BNA file format 93
- boundaries, county 75, 153
- buffer, definition 205
- building, office 72
- bulk mail
 - annual permit fee 196
 - discounts 196, 198
 - rates 195

C

- campaigns, marketing 125

- Canada Post SERP Report 20
 - Canadian address processing, data files
 - for 10
 - capitalization
 - in output 22
 - preferences 115
 - CASS
 - certification 22, 150
 - definition 205
 - files 9
 - mode 151
 - report 20, 196
 - Census
 - accuracy 15
 - Block Group 110
 - data 126, 129
 - demographics xi
 - FIPS Code 17
 - ID 17, 78, 102
 - ID, format 17
 - census ID accuracy 162
 - Census Tract, ZIP centroid matching
 - accuracy 16
 - centroid
 - best 16
 - definition 206
 - geocoding 102
 - Centrus Data Formatter, using 187
 - city
 - block 71
 - name 70, 77, 106
 - Claritas Demographics 175
 - closest site
 - analysis 139, 206
 - available elements 140
 - input fields 139
 - inputs, specifying 48
 - layers 137
 - object files 137
 - options 27
 - output fields 140
 - outputs, specifying 48
 - settings, verifying 37
 - closest site layers
 - adding from .gsb files 28
 - adding from database 29
 - importing 30
 - overview 137
 - removing 31
 - Closest Site module
 - overview xii, 137
 - cluster code xi, 206
 - codes
 - address location 161
 - cluster xi, 206
 - FIPS 207
 - GeoStan Canada location 165
 - line of travel 112
 - LOT 112
 - ZIP+4 centroid location 161
 - Coding Accuracy Support System *See* CASS
 - command line syntax 37
 - company
 - components 120
 - name 116
 - company name, extracting 120
 - completion time, estimated 151
 - components x
 - confidence level 116
 - coordinates
 - defining a centroid 206
 - return 17
 - cost
 - data 206
 - mailing 102
 - Countries.gsb file 23, 27
 - county
 - boundaries 75, 153
 - FIPS code 110, 111
 - layer 79
 - name 111
 - coverage area 101
 - current
 - address 68
 - geocoded point 75, 153
 - layer 79, 80
 - map view 76
 - custom dictionary
 - creating 172
 - Custom.isd 10
 - entries 172
 - file 120, 172, 173
 - types 173
 - words 120, 173
 - Custom.isd 10
- ## D
- data
 - address 107, 110, 126
 - cost 206
 - demographic files 126
 - formats x
 - layers 207
 - locator module xi

- statistical 206
 - street 69
 - subset 119
 - data elements
 - address 107, 110
 - data locator 121
 - demographics 129
 - geographic determination 148
 - output field association 107, 118, 126
 - data files
 - address coding 8
 - Canadian address processing 10
 - CASS mode processing 9
 - installing 4
 - intersection matching 10
 - NAD conversion 10
 - search path, specifying 7
 - data locator
 - available elements 121
 - data files 10
 - gender threshold 22
 - input fields 116
 - inputs, specifying 51
 - name order 22
 - options 21
 - output capitalization 22
 - output fields 118
 - outputs, specifying 51
 - regular expression 23
 - settings, verifying 37
 - tips 119
 - Data Locator mode 62
 - Data Locator module
 - custom dictionaries 172
 - customizing 169
 - overview xi, 115
 - regular expressions 170
 - data subset 119
 - Data Viewer
 - starting 74
 - database
 - file formats 85
 - previously processed 106, 117, 126
 - datum
 - definition 17, 206
 - NAD 206, 208
 - North American 206, 208
 - dBASE files 86
 - deliverability, USPS 101
 - Delivery Point Bar Code 20, 110, 111
 - delivery speed 102
 - Demographic Coding module
 - overview xi
 - using 125
 - demographics
 - available elements 129
 - Census xi
 - data files 11, 126
 - input fields 126
 - inputs, specifying 46
 - output fields 126
 - outputs, specifying 46
 - premium 129, 175
 - settings, verifying 37
 - variables xi
 - demonstration license 2
 - demonstration mode 42
 - dictionary
 - See also* custom dictionary
 - file 120, 172, 173
 - InfoStan 173
 - direction, LOT 112
 - directionals, missing 14
 - directory
 - default search 7
 - ZIP+4 102
 - discounts
 - bulk mail 198
 - disk space, requirements 1
 - DMM 198
 - Domestic Mail Manual* 198
 - DPBC 20, 110
 - duplicate records 115
- ## E
- Early Warning System data 167
 - elements
 - address 107, 108, 110
 - data 107, 118, 126
 - name 22, 115
 - eLOT 206
 - ethnicities, valid 174
 - EWS data 167
 - Excel files 86
 - expressions, regular 23
- ## F
- Federal Information Processing
 - Standards code *See* FIPS code
 - fee, annual Bulk Mail permit 196
 - file
 - .gsx 81
 - Access 86
 - address 196
 - ASCII 86, 114
 - audit report 90

- Countries.gsb 23, 27
 - custom dictionary 120, 172, 173
 - database formats 85
 - dBASE 86
 - demographic data 126
 - dictionary 120, 172, 173
 - Excel 86
 - formats 85
 - FoxPro 86
 - name 88
 - object 25, 30, 80, 82
 - output 88, 89
 - processing 149
 - reject 89
 - report 87
 - sample data 42
 - source 173
 - States.gsb 23, 27, 31
 - street 102
 - task 37, 38
 - types 85
 - Us.z9 187
 - finance area 15, 206
 - FIPS code
 - county 111
 - definition 207
 - firm name 106
 - first hex position 156
 - forecast sales 125
 - format
 - data x
 - database files 85
 - file formats 85
 - formatting
 - ASCII files 188
 - FoxPro files 86
 - frequency, select snapshot 90
- G**
- gender 115
 - gender threshold data locator 22
 - geocode
 - address coding 14
 - definition 207
 - geocoded point, current 75, 153
 - quality 14
 - types 160
 - geodetic network, primary 206
 - geographic determination
 - available elements 148
 - data file 11
 - inputs, specifying 52
 - output fields 146
 - outputs, specifying 52
 - settings, verifying 37
 - geographic determination files
 - importing 35
 - removing 36
 - geographic determination line file,
 - adding 33
 - Geographic Determination module
 - overview xii, 143
 - using 144
 - geographic determination point file,
 - adding 32
 - geographic determination polygon file,
 - adding 34
 - Geographic Information System *See* GIS
 - geographic market areas 93
 - geography files, exporting 82, 93
 - GIS
 - Atlas GIS 131, 138
 - definition 207
 - GSX files 81
- H**
- Hawaii, NAD27 208
 - hexadecimal digits, values of 156
 - high speed batch processing x
 - house
 - number range 72
 - possible numbers 73
 - range 70, 71
- I**
- identify
 - duplicate records 115
 - name elements 115
 - InfoStan dictionary 173
 - input table
 - specifying 43, 87
 - inputs
 - address coding 44
 - addresses, unmatched 109, 119, 128
 - closest site, specifying 48
 - data locator, specifying 51
 - demographics, specifying 46
 - file 90
 - geographic determination,
 - specifying 52
 - Point-in-Polygon, specifying 47
 - installing
 - Adobe Acrobat Reader 3
 - Centrus Desktop 3
 - data files 4

- local 2
 - server 2
- interactive processing 150
- intersection matching data files 10
- L**
- last line
 - cannot detect 106
 - multiline mode guidelines 106
 - score 123
 - second 123
- latitude and longitude 102, 112
- layers
 - closest site 137
 - closest site, adding from .gsb files 28
 - closest site, adding from database 29
 - closest site, importing 30
 - closest site, removing 31
 - current 79, 80
 - definition 207
 - overview 131
 - Point-in-Polygon, adding 24
 - Point-in-Polygon, importing 25
 - Point-in-Polygon, removing 26
 - street, primary 78
 - street, secondary 78
- layers, importing 81
- license
 - demonstration 2
 - renewal 3
- line of travel
 - code 112
 - direction 112
- line, last 106, 123
- location accuracy 112
- location codes
 - address 161
 - definition 207
 - ZIP+4 centroid 161
- log report
 - example 200
 - overview 90
 - specifying 91
- longitudes
 - negate 16
 - western hemisphere as negatives 16
- LOT
 - code 112
 - direction 112
- M**
- mail
 - bulk mail 20, 198
 - reduce costs 102
- mailing costs, reduce 102
- manually created point 75
- Map Viewer
 - buttons 75
 - geocoding manually 77
 - layers 78
 - overview 75
 - performance 81
- MapInfo 23
- mapping function 18
- maps
 - current address 63
 - current view 76
 - function 153
 - locating addresses 77
 - manually placing points 76
 - MapInfo 23
 - outside the U.S. 208
 - QuickFind mapping function 18
 - redrawing 75
 - viewing 75
 - visibility 80
 - window 76
 - zoom 76
- market penetration 125
- marketing campaigns 125
- match
 - pattern 15, 170
 - quality 14
 - record 68
- MatchCode values 156
- matching tolerance
 - CASS reports 15
 - close 14
 - extended 15
 - settings 14
 - tight 14
- Microsoft Access files 86
- middle name 121, 122
- minimum system requirements 1
- misspellings 14, 102
- mixed case output, addresses 19
- mode
 - CASS 151
 - demonstration 42
- modules
 - Address Coding x, 101
 - Closest Site xii, 137
 - Data Locator xi, 115
 - Demographic Coding xi, 125
 - Geographic Determination xii, 143
 - Point-in-Polygon xi, 131

Spatial xi, 131
Multiline mode 64

N

NAD conversion, data files 10
NAD27 17, 208
NAD83 17, 208
name
 city 71
 company, extract 116, 120
 county 111
 elements 115
 firm 106
 middle 121, 122
 order 22
 second street 113
 street 71, 112, 113
name order, data locator 22
network, street 102
North American Datum
 NAD27 17, 208
 NAD83 17, 208
number
 house 69, 72, 73, 111
 house range 72
 of records processed 152
 unit 70, 72, 74
 unit range 72

O

object file 25, 30, 80, 82, 131
office
 building 72
optimization sort 107
output capitalization 22
output table
 specifying 43, 88
outputs
 address coding 44
 closest site, specifying 48
 data locator, specifying 51
 demographics, specifying 46
 file 88
 geographic determination,
 specifying 52
 Point-in-Polygon, specifying 47

P

pattern matching 15, 170
performance
 improving 81, 187
permit fee, bulk mail 196

person
 first 121
 second 122

PMA

customer-based 96
exporting geography files 93
overview xii
saved in GSB format 96
store-based 94

point, manually created 75

Point-in-Polygon

analysis 132, 208
buffer width 133
buffering 132
elements 134
field names 135
input fields 133
inputs, specifying 47
objects 132
options 23
output fields 134
outputs, specifying 47
requirements for analysis 132
settings, verifying 37
unassigning fields 136
using the module 131-136

Point-in-Polygon layers

adding 24
importing 25
removing 26

Point-in-Polygon module

overview xi
using 131-136

positional accuracy 162

postage

savings 102

postal bar code 114

postal reports 20

preferences, capitalization 115

premium demographics 129, 175

primary

geodetic network 206
street layer 78

Primary Market Area *See* PMA

processing

batch 150
interactive 150
speed 44, 187
statistics 151

Q

quality
 address 102
 geocodes 14

- matches 14
- query function
 - overview 67
 - search filter 69
 - selecting a match 74
 - tree 70
- QuickFind
 - Address Wizard mode 65
 - Data Locator mode 62
 - Multiline mode 64
 - overview 61
 - query function 67
 - Standard mode 63

R

- range
 - ASCII characters 171
 - house number 70, 71, 72
 - record 71
 - unit numbers 72
- records
 - address 68
 - duplicate 115
 - matching 68
 - number processed 152
- reject file, specifying 89
- remove
 - field 126
 - layer 79
- report
 - audit 91, 199
 - Canada Post SERP 20
 - CASS 20, 196
 - file 87
 - log 200
- report files
 - audit 90
 - log 90
- requirements 1
- return coordinates 17
- roads 78

S

- Sagent Technical Support xv
- sample data files 42
- search paths
 - default 7
 - specifying 7
- second
 - last line score 123
 - person 121, 122, 141, 148
 - street name 113
 - street suffix 113

- second hex position 156
- secondary street layer 78
- SERP (Canada)
 - correctable addresses 204
 - definition 208
 - overview 203
 - report 20
 - reports 20
 - valid addresses 204
- Shape Layers
 - exporting 81
 - Generate PMA 81
 - importing 81
- site
 - analysis 48, 139
 - closest 139
 - identifier 48, 139
- snapshot frequency 90
- SOA
 - definition 208
 - example 203
- Software Evaluation and Recognition Program *See* SERP (Canada)
- sort
 - automated 111
 - optimization 107
 - presort 112
- space free on disk 1
- Spatial
 - adding files from 32, 33, 34
 - analysis 18, 139
- Spatial module
 - overview xi
- speed
 - delivery 102
 - high speed batch processing x
 - processing 187
- Standard mode 63
- standardization
 - address 14, 196
- state abbreviations 69, 70
- Statement of Accuracy
 - definition 208
 - example 203
- States.gsb file 23, 27, 31
- statistical data 206
- statistics 152
- street
 - address 77, 102, 106
 - data 69
 - first 79
 - name 71, 112, 113
 - name, alias 71
 - network files, enhanced 102
 - primary layer 78

- second name 113
- second suffix 113
- secondary layer 78
- segment 76, 80, 160
- suffix 112, 113
- type 14, 71, 102
- suffix
 - second street 113
 - street 112, 113
- syntax, command line 37
- system requirements 1

T

- tasks
 - existing 39
 - file 37
 - new 38
 - opening 39
 - overwriting file defaults 38
 - processing 41
 - setting up 43
 - stored location 38
- technical support, contacting xv
- third hex position 156
- tolerance settings 14
- tools
 - Data Viewer 74
 - Map Viewer 75
 - QuickFind 61
 - Shape Layers 81
- tutorial 41

U

- U.S. address coding options 18
- U.S. Postal Service *See* USPS
- undeliverable address 102
- unit
 - number 70, 72, 74
 - range 72
- UNIX, regular expressions 170
- unmatched
 - input addresses 109, 119, 128
 - input addresses, retain 109, 119, 128
- USPS
 - CASS certification 22, 150
 - CASS report 15, 20, 196
 - deliverability 101
 - file 102
 - finance area 15, 206
 - recommendations 19
 - standards 101, 102

V

- value of hexadecimal digits 156
- values returned in MatchCode 156
- variables, demographic xi

Z

- Z4ChangeResult 114
- ZIP centroid matching
 - 5 digit only 16
 - best location 15
 - Block Group accuracy 16
 - Census Tract accuracy 16
 - good location 15
 - locations 16
 - settings 15
- ZIP Code
 - centroid 15, 102
 - definition 208
 - ZIP+4 102, 160, 196
- ZIP+4 centroid location codes
 - 2nd character 162
 - 3rd character 162
 - 4th character 163