Geocoding and Mapping in SAS®

Shannon J. Moore Sr. Systems Engineer SAS Institute Inc.





- How do you convert addresses into map locations? This is done through the process of geocoding.
- PROC GEOCODE was first included in SAS/GRAPH 9.2 to provide this capability.
- Street-level geocoding for the United States was later added to the third maintenance release of SAS9.2M3
- the new abilities to geocode international cities added in SAS 9.3M2 and Canadian street-level geocoding available in SAS 9.4
- you can now import free postal code data for Great Britain and Australia for geocoding in all releases of PROC GEOCODE
- With 9.4M5 PROC GEOCODE is now part of Base SAS



- The geocoding process depends on lookup data with the necessary information to convert an address to a geographic location
- This data is the key to geocoding. Factors such as age and granularity of the lookup data determine the geocoding results.
- Granularity is another important consideration. Does the location need to be the actual house position, or will a more generalized location be close enough?
- Another factor to consider when choosing a geocoding method is your input data. What address components does it contain?



My SAS environment and a bit about me

- Viya 2023.5
- SAS Studio Engineer*

NOTE: Even though I am using Viya, the code is 9.4. Viya runs SAS 9 code with little to no modification*

- Twenty-six years at SAS specializing in ODS Statistical Graphics, SAS/GRAPH and SAS' Business Analytics interfaces including Enterprise Guide, the Add-In for Microsoft Office, Enterprise BI, Visual Analytics, etc.
- Eight years Geology/Geophysics Oil and Gas
- Five years Geographic Information Systems (GIS)
- SAS Certified in Viya Programming, 9.4
 Programming and as a Visual Business Analyst in Visual Analytics



definition

Geocoding: the process of converting an address to a physical location





Location

What is geocoding?

City **Address** Location



What is geocoding?

City

ZIP Code





What is geocoding?

City

ZIP Code

ZIP+4





What is geocoding?

City

ZIP Code

ZIP+4

Street





What is geocoding?

City

ZIP Code

ZIP+4

Street

IP Address





What is geocoding?

City

ZIP Code

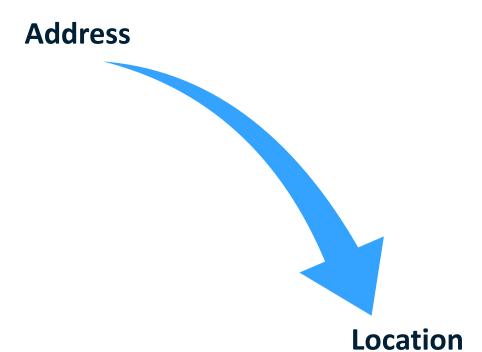
ZIP+4

Street

IP Address

Custom

Region





Mailing Address Component

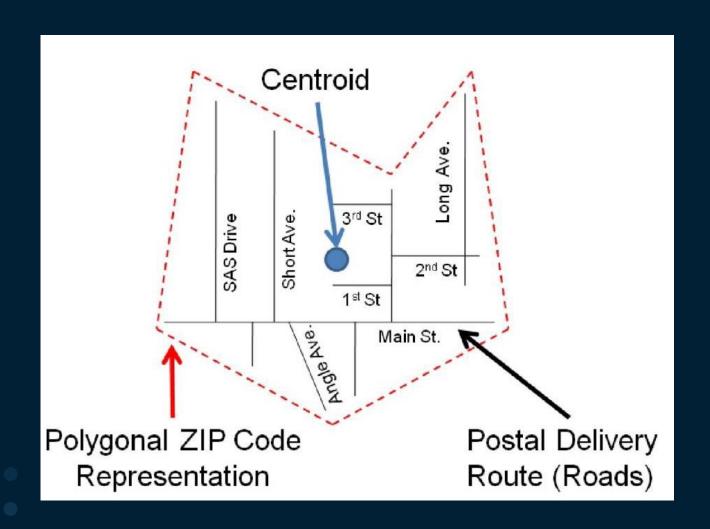
- City Center
- ZIP Code Centroid
- ZIP+4 Center
- Street Level

Other

- IP Address
- Custom Region Centroid



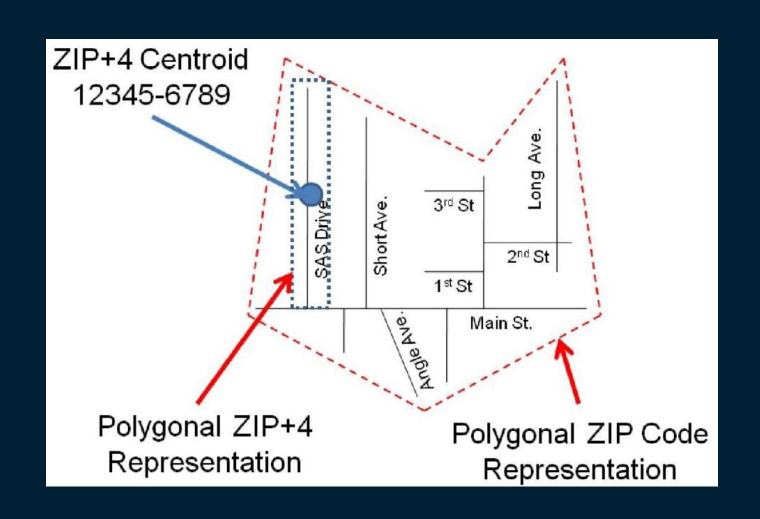
Boundary Wraps Streets with Common ZIP Code



No official ZIP code boundaries

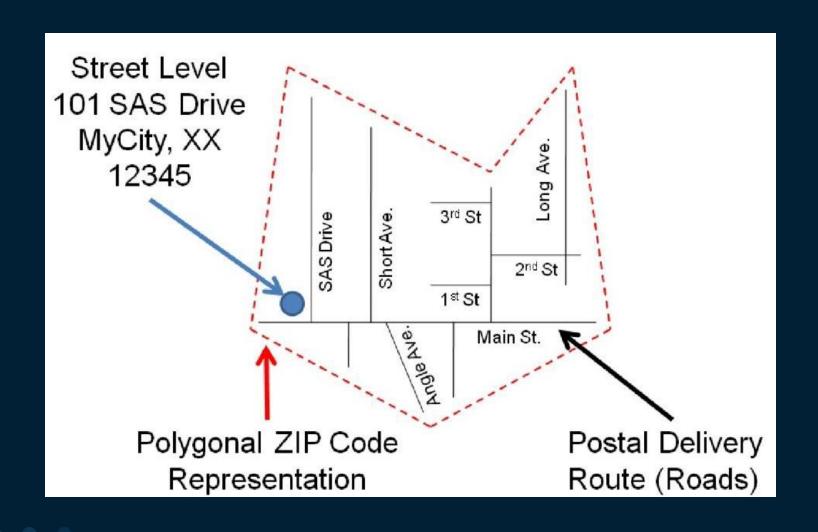


ZIP+4 is a Portion of a ZIP Code



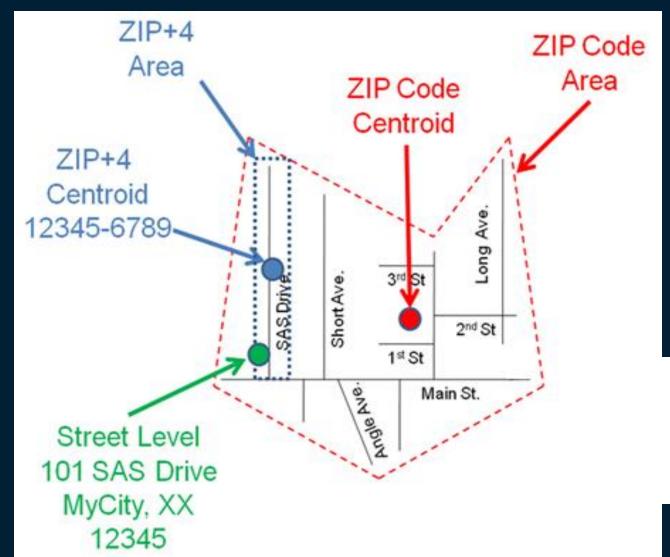


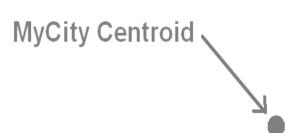
Street-level Geocoding is Most Precise





Relative Precisions







IP Address

- Non-mailing Parameter
- Ranges of IP Addresses are Clustered Geographically
- Precision Limited to Country or City Center
- Also Known as Geolocation



Custom Regions

- Non-mailing Parameter
- Any Data Having Specific Geographic Location
- Examples
 - Sales Regions
 - Telephone Area Code



Geocoding Method Selection Criteria

1. Geographic Extent



- 1. Geographic Extent
- 2. Attribute Values Wanted



- 1. Geographic Extent
- 2. Attribute Values Wanted
- 3. Location Precision



- 1. Geographic Extent
- 2. Attribute Values Wanted
- 3. Location Precision
- 4. Address Components Present



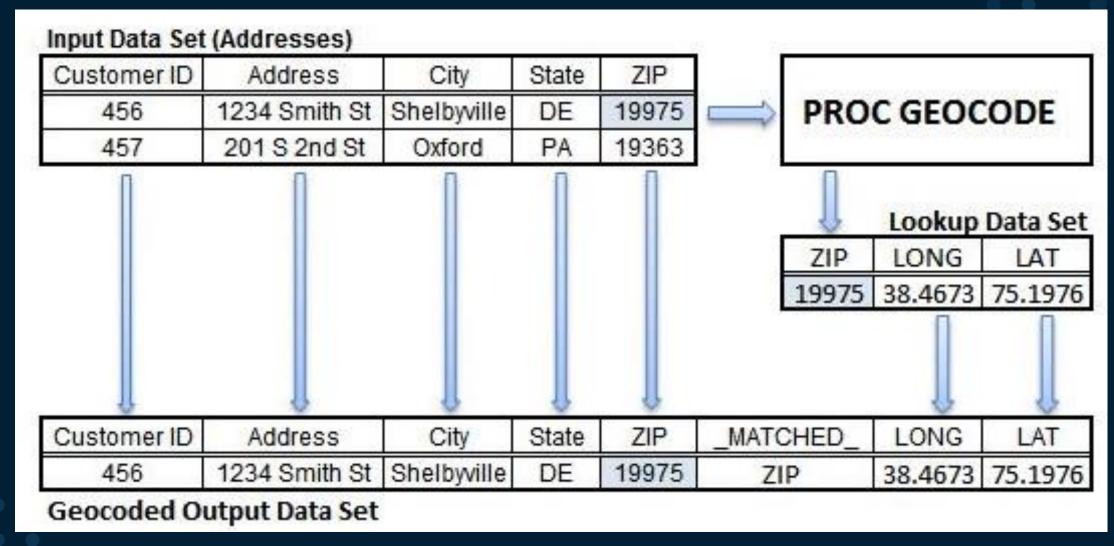
- 1. Geographic Extent
- 2. Attribute Values Wanted
- 3. Location Precision
- 4. Address Components Present
- 5. Lookup Data Availability



Geocoding Method	Input Data	Lookup Data	
		Coverage	Source
CITY	City and state name	United States	MAPSGFK.USCITY_ALL
			SASHELP.ZIPCODE
		World	MAPSGFK.WORLD_CITIES
			WORLD_CITIES_ALL
ZIP	Postal code	United States	SASHELP.ZIPCODE
		Australia	Australian Bureau of Statistics
		England and Scotland	Ordnance Survey of Great Britain
PLUS4	ZIP code with ZIP+4	United States	MapsOnline
			Melissa Data
STREET	Complete mailing address	United States	MapsOnline
			Census TIGER shapefiles
		Canada	GeoBase National Road Network
RANGE	IP address	World	MaxMind
CUSTOM	User-defined region	n/a	User created



Geocoding Process – Example ZIP method



CITY Method: 9.2 – 9.3M1

- Only **U.S. Cities**
- Lookup Data: SASHELP.ZIPCODE
- Installed with SAS
- Updated Each Release
- Quarterly Updates on MapsOnline



CITY Method: 9.3M2 – 9.4

- Worldwide Cities
- Lookup Data: MAPSGFK.USCITY_ALL MAPSGFK.WORLD_CITIES
- Installed with SAS/GRAPH
- Updated Each Release
- Unabridged WORLD_CITIES_ALL Data Set Available on MapsOnline



ZIP Method – available since 9.2

- Geocode by U.S. ZIP Codes
- Lookup Data: SASHELP.ZIPCODE
- Installed with SAS
- Updated Each Release
- Quarterly Updates on MapsOnline



ZIP Method

- Also Geocodes with British Postal Codes
- Free Lookup Data from Ordnance Survey Code-Point Open Product
- Import Program and Documentation on MapsOnline
- Coverage for England and Scotland



ZIP Method

- Also Geocodes with Australian Postal Codes
- Free Lookup Data from Australian Bureau of Statistics' Postal Area (POA) File
- Import Program and Documentation on MapsOnline
- POA Files Include Postcode Polygons



ZIP Method

- Postcode Centroids for Other Countries Available from Vendors
- Examples:
 - www.zipcodedownload.com www.allmapdata.com
- SAS Technical Support can assist importing Third Party Data



ZIP+4 Method

- Geocode by U.S. ZIP+4 Codes
- Free Lookup Data on MapsOnline
- Contains ZIP+4 Values from 2006 (should be 2012 TIGER now)
- Update When ZIP+4 Returns to TIGER Files
- Current ZIP+4 Values Available in **Melissa Data** Geo*Data Product
- Import Using %GCDMEL9 Autocall Macro documented in doc.



STREET Method: 9.2M3 – 9.3M2

- Coverage includes U.S. with Puerto Rico and Virgin Islands
- Free National Lookup Data from TIGER Files
- Download prebuilt nationwide lookup data from MapsOnline
- Updated Each Annual **TIGER** Release
- Macro TIGER Import Program Available from MapsOnline



STREET Method: 9.4

- Added Coverage for Canada
- Free Lookup Data from GeoBase: www.geobase.com
- Download Files for Provinces
- Macro GeoBase Import Program Available from MapsOnline
- No Address Data for Newfoundland and Labrador
- No Postcode Values in GeoBase Data
- Can also use %GEOBASE2GEOCODE macro program to import files for specific provinces downloaded from **GeoBase**.



RANGE Method

- IP Address Geolocation
- Free Lookup Data from **MaxMind**:
 - GeoLite Country
 - GeoLite City
- Import Using %MAXMIND Autocall Macro



CUSTOM Method

- Any Geographic Data Can be Used
- Data Must Have Latitude/Longitude and Third Variable for Lookup Key
- User Builds Lookup Data
- SAS Technical Support can assist



Procedure Syntax

PROC GEOCODE

METHOD =

DATA =

LOOKUP =

options...



Example U.S. Input Address Data

Customer	Address	ZIP	ZIP+4	City	State	Revenue
J. Cheever Loophole	136 E. Water St	19901	3630	Dover	DE	\$985.33
Cuthbert J. Twillie	760 Moose Lodge Road	19934	2220	Camden	DE	\$2,533.25
Kaspar Gutman	4701 Limestone Road	19808	1927	Wilmington		\$33.44



CITY Method Syntax:



CITY Method: 9.4

- Worldwide Cities
- Lookup Data: MAPSGFK.USCITY_ALL MAPSGFK.WORLD_CITIES
- Installed with SAS/GRAPH
- Updated Each Release
- Unabridged WORLD_CITIES_ALL Data Set Available on MapsOnline



U.S. CITY Method Syntax: 9.4

```
proc geocode
  method
                                         /* City method
                = city
                                         /* Address data to geocode
                = work.customers
  data
                = work.geocoded
                                         /* Geocoded output data set
                                                                         */
  out
                = mapsgfk.us_city_all
                                         /* Default U.S. lookup data
                                                                         */
  lookupcity
                                         /* Attribute value to add */
  attribute_var
                = (county_name);
run;
```



World CITY Method Syntax: 9.3M2 – 9.4

```
proc geocode
                                            /* City method
  method
             = city
                                            /* Addresses to geocode
                                                                          */
             = SASoffices
  data
             = work.geocoded
                                            /* Geocoded data set
  out
  lookupcity = mapsgfk.world_cities
                                            /* Default world lookup data
                                                                          */
                                                                          */
  addresscityvar
                                            /* Req'd. city name
                            = city
  addresscountryvar
                            = countryID
                                            /* Req'd. country name
                                                                          */
                                            /* Optional state name
  addressstatevar
                            = state
                                            /* Attribute value to add
  attributevar
                            = (cttype);
run;
```







ZIP Method Syntax: U.S. Postcodes

```
proc geocode
  method
                                              /* ZIP method
                      = zip
                                              /* Addresses to geocode
                      = work.customers
  data
                      = work.geocoded
                                              /* Geocoded data set
  out
                                              /* Default lookup data set
                      = sashelp.zipcode
  lookup
  attribute_var
                      = (msa areacode);
                                              /* Attribute values to add
run;
```



*/

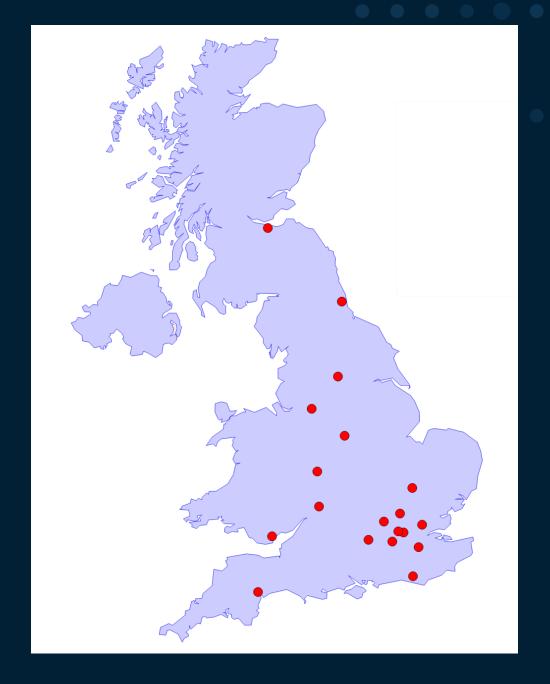
*/

ZIP Method Syntax: British Postcodes

```
proc geocode
  method = zip
                                            /* Method for postcodes
                                            /* Addresses to geocode
          = dealers
  data
                                            /* Geocoded data set
          = geocoded
  out
          = lookup.GBpostcodes
                                            /* Postcode lookup data
  lookup
                                            /* Address postcode var
  addresszipvar
                        = postcode
                                            /* Lookup postcode var
  lookupzipvar
                        = pc
                                            /* Lookup data X-variable
  lookupxvar
                        = Lon WGS84 DD
                                            /* Lookup data Y-variable
                        = Lat WGS84 DD
  lookupyvar
                                            /* Disable CITY method
  nocity;
run;
```



Aston Martin
Dealers
In
Great Britain



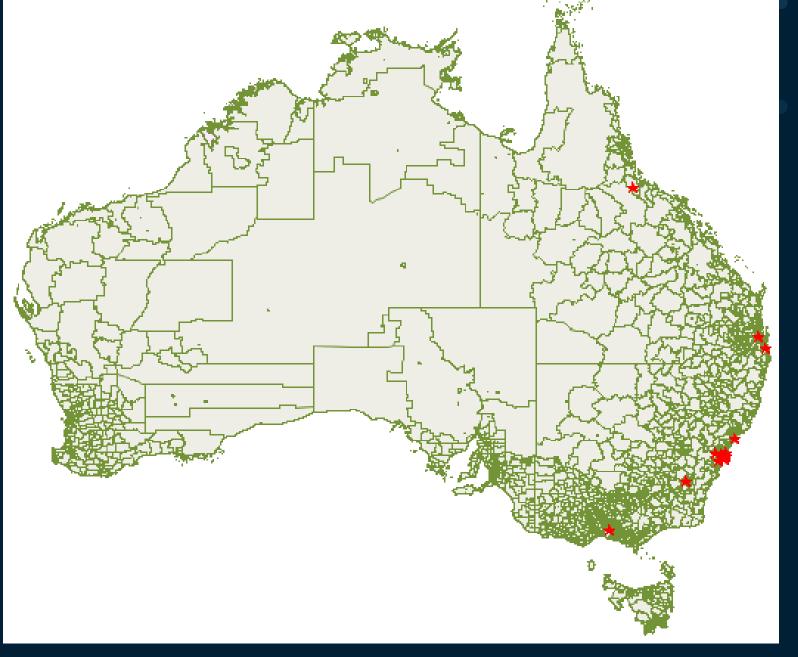


ZIP Method Syntax: Australian Postcodes

```
proc geocode
                                            /* Method for postcodes
  method = zip
                                                                       */
          = NRL_stadiums
                                             /* Addresses to geocode
                                                                      */
  data
                                             /* Geocoded data set
                                                                       */
          = geocoded
  out
          = lookup.postcodes
  lookup
                                             /* Postcode lookup data
                                                                      */
                                             /* Address postcode var
                                                                      */
  addresszipvar
                        = postcode
  lookupzipvar
                        = POA_code
                                             /* Lookup postcode var
                                                                       */
                                             /* Disable CITY method
                                                                      */
  nocity;
run;
```



Australian Postal
Areas
and
National Rugby
League Stadiums





Zip 4 Example

RESOURCES / FOCUS AREAS

FOCUS AREAS

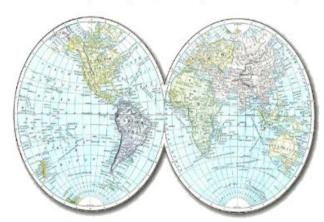
- Analytics Life Cycle
- Base SAS
- Graphics
 - Automatic Graphs
 - Your Graphs
 - Maps
 - Maps Online
 - SAS Bridge For ESRI
- Enterprise Management Integration
- Migration
- Scalability & Performance
- SAS for Containers
- Statistics & Operations Research

Welcome to SAS Maps Online

SAS Maps Online shows maps for areas throughout the world. You can easily locate and identify specific regions in each of the following categories: world maps, continents, countries, and maps of political groups.

SAS users will find:

- · Archived maps from previous releases
- Sample programs
- · Recent Mapping and Geocoding updates
- · Geocoding examples, techniques and look-up data



Click on the World Image to go to MapsOnline



Zip 4 Example

ZIP+4 Geocoding

• Free ZIP+4 Data

A file containing ZIP+4 centers from 2006 is available. These ZIP+4 locations are based on the 2006 Second Edition TIGER/Line files from the Census Bureau. That was the most recent TIGER/Line release which contained ZIP+4 values. This file will be updated when the Census Bureau replaces ZIP+4 values in a future TIGER release. As of the most recent TIGER release, the ZIP+4 values had not yet been restored. Download that 2006 ZIP+4 file: <u>ZIP4_GEOCODE_DATA-2006.zip</u>

Alternate ZIP+4 Data

The GEO*Data product containing current ZIP+4 locations can be purchased from Melissa Data. The autocall macro %GCDMEL9 imports GEO*Data files into lookup data for ZIP+4 geocoding with PROC GEOCODE. See the SAS/GRAPH documentation for details on this macro.

IP Address Geocoding

Worldwide IP address geocoding is supported by all PROC GEOCODE releases. Lookup data is available from MaxMind in their free GeoLite databases. The autocall macro %MAXMIND imports GeoLite files to create lookup data. See PROC GEOCODE in the SAS/GRAPH documentation for details on this macro.



Zip 4 Example

SUPPORT / DOWNLOADS & HOT FIXES

SUPPORT

- Manage Tracks
- SAS Administrators
- Security Bulletins
- License Assistance
- Manage My Software Account
- Downloads & Hot Fixes
 - SAS System Software
 - >SAS DataFlux Software
 - JMP Software
 - ▶ Mobile Device Software
 - ▶Third Party Software
 - > Additional Products
 - ▶ Hot Fixes
- Knowledge Base

Downloads

Maps Online

Updated JANUARY 2019

January 15, 2019 Update - Zipcode and geocode files

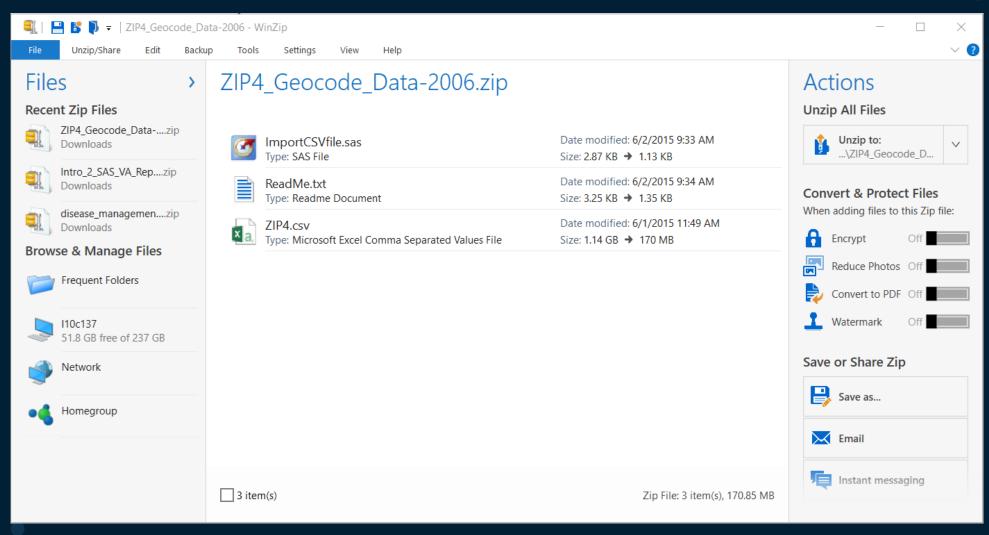
To begin the download process, select the item in the table below which best meets your needs.

Maps Online					
Platform	Description	Request Download	Size	Release Date	
All	SAS Files	ZIP4_Geocode_Data-2006.zip	170.8 MB	2010-01	

For questions you may have about this download, please contact our Product Support Group



Zip 4 Example





Zip 4 Example



ReadMe.txt - Notepad

File Edit Format View Help

Summary

The zip archive 'ZIP4 Geocode Data-2006.zip' contains a comma-separated values (CSV) file to create a SAS data set.

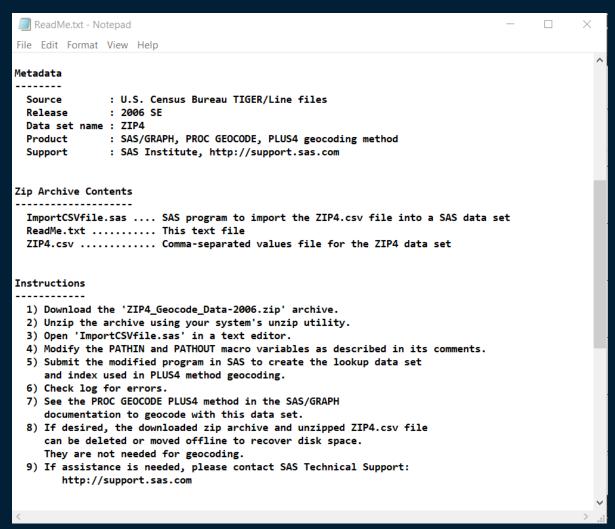
The data set is used by PROC GEOCODE for ZIP+4 geocoding of U.S. mailing addresses. See the PLUS4 geocoding method for PROC GEOCODE in the SAS/GRAPH documentation.

The ZIP+4 values and locations were computed using the 2006 SE TIGER/Line files. In 2007 the U.S. Census Bureau changed from their Record Type (RT) format ASCII files to shapefiles and did not include the ZIP+4 values. Their original intent was to restore the ZIP+4 values in a future TIGER shapefile release, but as of the 2014 TIGER/Line release they had not yet done so.

So be aware that the ZIP+4 values in this CSV file are from 2006 and should be used with caution. Up-to-date ZIP+4 values can be obtained from Melissa Data (www.melissadata.com) in their Geo*Data product. See the PROC GEOCODE documentation on using the GCDMEL9 autocall macro to import Geo*Data files.



Zip 4 Example





ZIP+4 Method Syntax

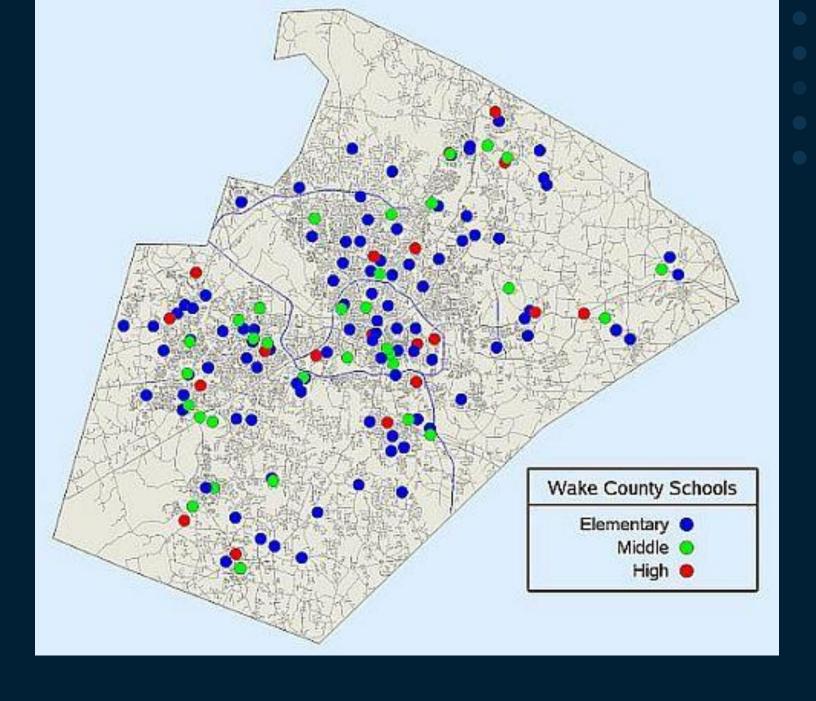
```
proc geocode
  method = plus4
                                     /* ZIP+4 method
  lookup = lookup.zip4
                                     /* Lookup data set
                                     /* Addresses to geocode
                                                                 */
  data = work.customers
                                     /* Geocoded data set
                                                                 */
          = work.geocoded
  out
  attribute_var = (tract)
                                     /* Attribute value to add
                                     /* Disable CITY method
                                                                 */
  nocity;
run;
```



STREET Method Syntax: U.S.

```
proc geocode
  method = street
                                    /* Specify method
                                                                 */
                                    /* Addresses to geocode
  data = work.schools
         = work.geocoded
                                    /* Geocoded data set
                                                                 */
  out
  lookupstreet = sashelp.geoexm
                                    /* Lookup data set
  attribute_var = (tractce00);
                                                                 */
                                    /* Attribute value to add
run;
```







```
proc geocode
   method
                                                      /* Specify method
                 = street
                                                      /* Addresses to geocode
   data
                 = work.inns
                                                                                    */
                 = work.geocoded
                                                      /* Geocoded data
   out
   lookupstreet
                 = lookup.canada_m
                                                      /* Lookup data set
                                                                           mapsOnline*/
                 = lookup.gcdirect_can
                                                      /* Direction names
   direct
                                                                            */
                 = lookup.gctype_can
                                                      /* Street type names
   type
   addresscountryvar= country
                                                      /* Country names
   addressstatevar= province
                                                      /* Province names
   addresscityvar= city
                                                      /* City names*/
                                                      /* Disable method
                                                                            */
   nozip;
run;
```



- Canadian street lookup available on MapsOnline
- JAVA MapApplet with OpenStreetMap background





RANGE Method Syntax

```
proc geocode
                                                         /* Geocoding method */
   method
              = range
                                                         /* IP address input data
   data
              = work.addresses
              = work.geocoded
                                                         /* Geolocated data set */
   out
                                                         /* Address var in input data set  */
   addressvar
                                 = IPaddress
                                 = lookup.cityblocks
                                                         /* Range data set
   rangedata
   beginrangevar
                                 = startipnum
                                                         /* Begin range variable */
                                                         /* End range variable
   endrangevar
                                 = endipnum
   rangekeyvar
                                                         /* Link variable in range data
                                 = locid
                                                         /* Lookup data set from MAXMIND */
   lookup
                                 = lookup.citylocation
   lookupkeyvar
                                                         /* Link variable in lookup data
                                 = locid
                                                         /* X variable in lookup data
   lookupyvar
                                 = latitude
                                                                                         */
   lookupxvar
                                 = longitude
                                                         /* Y variable in lookup data
   attributevar
                                 = (city, country);
                                                         /* Attribute values to add
                                                                                         */
run;
```



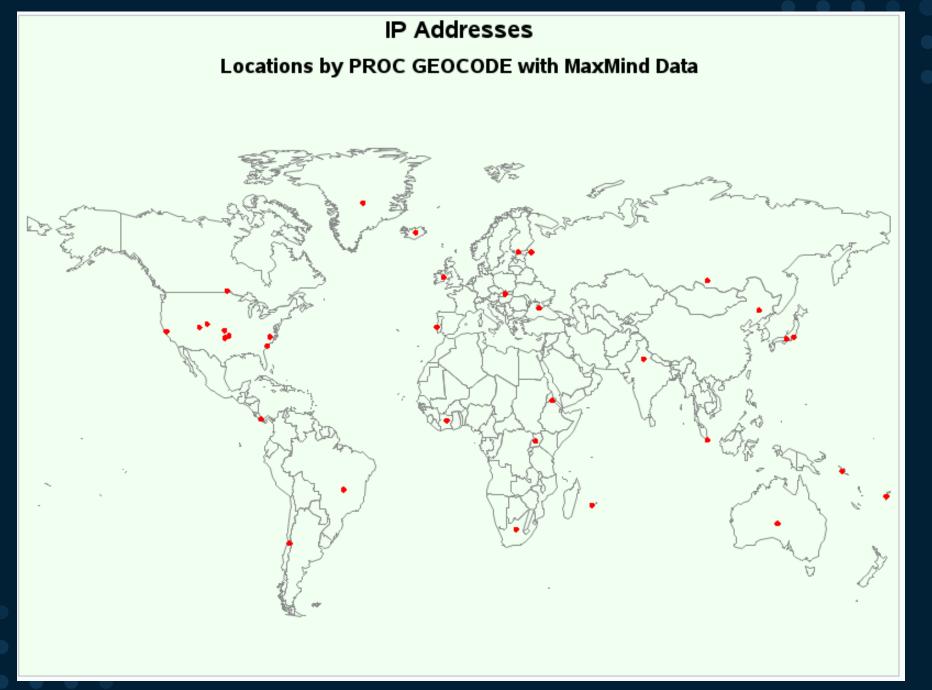






Table of Contents

- Databases
- IP Geolocation
- Autonomous System Numbers
- Support
- License
- Downloads

Links

System Status





GeoLite Legacy Downloadable Databases

New Database Format Available: This page is for our legacy databases. For our latest database format, please see our GeoLite2 Databases.

Databases

IP Geolocation

The GeoLite databases are our free IP geolocation databases. They are updated on the first Tuesday of each month. These databases are offered in the same binary and csv formats as our subscription databases. Any code which can read the subscription databases can also read the GeoLite databases.

IP geolocation is inherently imprecise. Locations are often near the center of the population. Any location provided by a GeoIP database should not be used to identify a particular address or household.

We publish accuracy statistics for GeoLite City.

Autonomous System Numbers

We offer free databases that map IPv4 and IPv6 addresses to Autonomous System Numbers (ASN), including the names of each Autonomous System. The databases are updated every month.

Support

MaxMind does not provide customer support for free GeoLite databases. If you have questions on how to use these databases, we suggest asking on Stack Overflow.

License

The GeoLite databases are distributed under the Creative Commons Attribution-ShareAlike 4.0 International License. The attribution requirement may be met by including the following in all advertising and documentation mentioning features of or use of this database:

1 This product includes GeoLite data created by MaxMind, available from 2 http://www.maxmind.com.

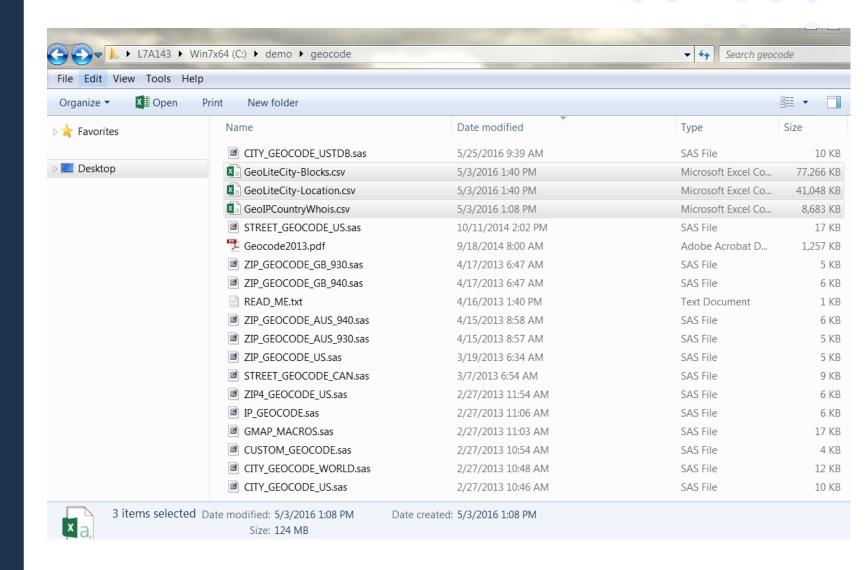
We also offer commercial redistribution licensing.

Downloads

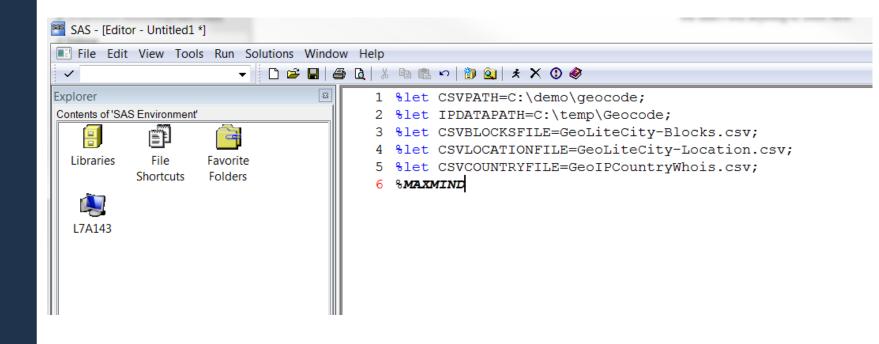
	Download links					
Database	Binary / gzip	Binary / xz	CSV / gzip	CSV / zip	CSV / xz	
GeoLite Country	Download	Gzip only	Zip only	Download	Zip only	
GeoLite Country IPv6	Download	Gzip only	Download	Gzip only	Gzip only	
GeoLite City	Download	Download	Zip and xz only	Download	Download	
GeoLite City IPv6 (Beta)	Download	Gzip only	Download	Gzip only	Gzip only	
GeoLite ASN	Download	Gzip only	Zip only	Download	Zip only	
GeoLite ASN IPv6	Download	Gzip only	Zip only	Download	Zip only	

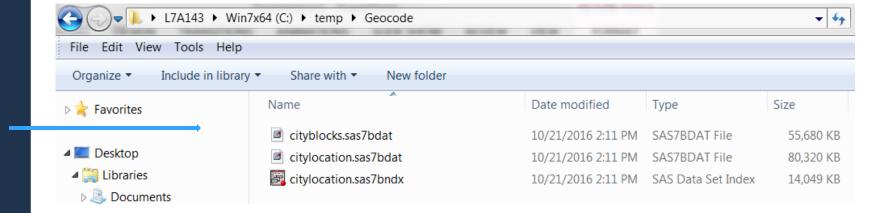
The GeoLite Legacy databases may also be downloaded and updated with our GeoIP Update program.

Downloaded files



From documentation: How to create SAS datasets





CUSTOM Method Syntax

proc geocode
method
data
out
lookup
lookupvar
Addressvar
run;



Summary

- Included with Base SAS
- Determines Location from Address
- Assigns Attribute Data to Locations
- Plot Locations on Map or Use in Calculations





Mapping



Mapping and Geocoding with SAS

- Mapping using SAS procedures
 - GMAP
 - Legacy
 - Part of SAS/GRAPH
 - Lots of examples from Dr Robert Allison's page
 - SGMAP
 - New with 9.4M6
 - Incorporates either OpenStreetMap or ESRI backgrounds to enhance maps
 - Similar but new syntax
 - Examples also on Dr Allison's page and SAS Communities



- New with 9.4M6
- Part of Base SAS ODS Graphics like the other SG procedures
- Can create a tile-based map that uses either the OPENSTREETMAP or ESRIMAP as a background
- Different syntax but straightforward



Procedure

```
PROC SGMAP
         MAPDATA=map-data-set | PLOTDATA=plot-data-set | MAPRESPDATA=response-data-set
<options>;
         STYLEATTRS </options>
         CHOROMAP < response-variable > / </options >;
         ESRIMAP URL='map-service';
         OPENSTREETMAP;
         BUBBLE X=variable Y=variable SIZE=numeric-variable </options>;
         SCATTER X=variable Y=variable </options>;
         SERIES X=variable Y=variable </options>;
         TEXT X=variable Y=variable TEXT=variable </options>;
         GRADLEGEND <"name-1" – "name-n"> </options>;
         KEYLEGEND <"name-1" - "name-n"> </options>;
```



example

```
proc sgmap
    maprespdata=geocoded plotdata=geocoded noautolegend;
    scatter x=x y=y / datalabel=zip
    markerattrs=(color=red size=19 symbol=circlefilled);
    openstreetmap;
run;
```



proc sgmap mapres scatte marker openst

run;

2000 401 North Raleigh ALEIGH-DURHAM Lynn Millbrook Crabtree (2215 2215 27513 Morrisville • 27513 27545 64/264 27513 Cary Knightdale Raleigh 27603 (231) Southeas 1371 Gamer BAGWELL Clayton olly Springs 42 27603 TRIPLE 27526 Fuquay-Varina Wilsons Mills

noautolegend;

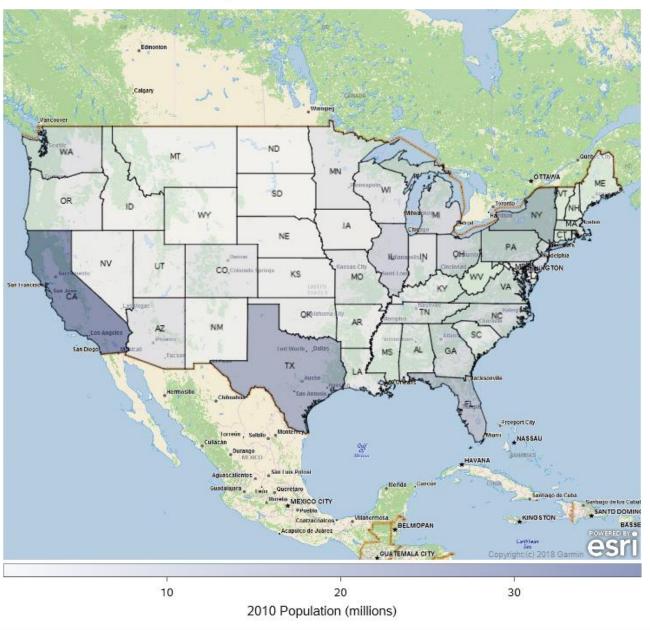
irclefilled);



example



Population from 2010 US Census



Lorme_World_Base_Map';
choro';

So that's all of SAS' map capabilities?

Nope. Visual Analytics has an ever-growing list of capabilities, many from our partnership with ESRI

Be sure to attend the session, "How Can I Create Custom Geo Maps in SAS" Visual Analytics?" on July 25th at 1pm EDT



Mapping and Geocoding with SAS VA/Viya

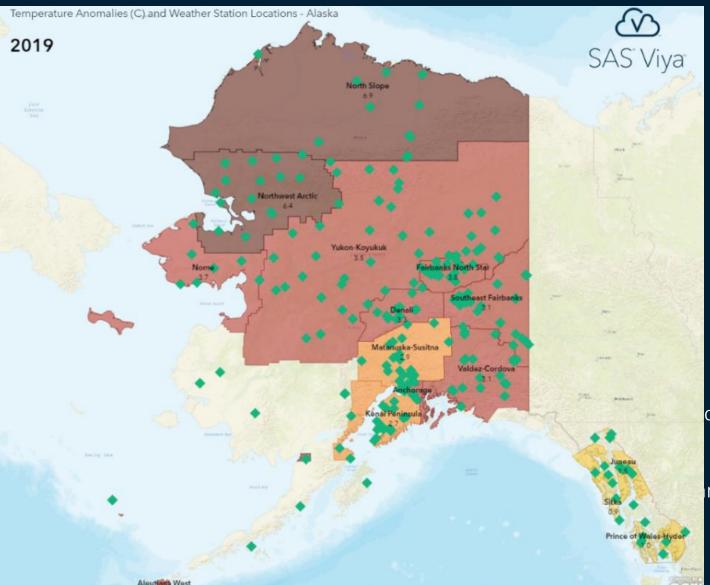
- Mapping using SAS Visual Analytics
 - On 9.4 architecture
 - 7.5
 - Some enhanced capabilities
 - Viya
 - Basic and enhanced features
 - Tighter integration with ESRI
 - Geo-enrich your data by joining it with ESRI demographics data
 - Drop a pin and add geographic selection areas around your pins based on distance or travel time
 - Custom regions
 - Multiple map layers
 - Find optimal routes between points of interest or measure the straight-line distance between
 - Geocoding and direct import of shape files
 - Much more



Mapping and Geocoding with SAS



- On 9.4 arch
 - 7.5
 - Some enha
- Viya
 - Basic and e
 - Tighter inte
 - (-
 - D
 - (
 - N
 - Fi
 - G
 - •

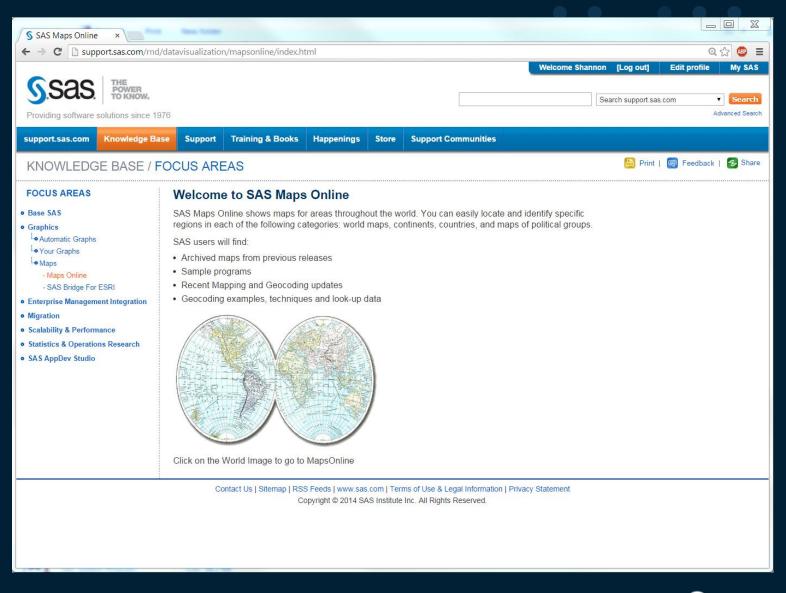


ce or travel time

nce between



www.sas.com/mapsonline





Thank you for your time and for using SAS! Stay safe and well!

sas.com

