

Paper SAS 1801-2018
New Location Analysis and Demographic Data Integration with
SAS® Visual Analytics and Esri

Murali Nor and Himesh Patel, SAS Institute Inc.,

ABSTRACT

SAS® continues to expand the location analytics capabilities that are offered in SAS® Visual Analytics on SAS Viya™. The goal of this paper is to present the new capabilities that can help you use the location information in business data. There are a variety of situations where location data is vital. Often business data can contain a large number of data points, sometimes in the order of 100,000 or a million. The preferred method for visualizing that many points on a geo map is by using geo-clustering. Another situation is when you are analyzing customers in a given location. Adding demographic data will improve insights. If you are performing a task such as analyzing crimes around a location in a city, using travel-time analysis and travel-distance analysis can provide valuable insights. If you need to represent various points of interest on geo maps, using symbols or icons helps make the map clearer.

This paper demonstrates how to use these capabilities in SAS Visual Analytics to solve real use cases for demographic data integration with Esri, geo-clustering, symbols, travel-time analysis, and custom regions.

INTRODUCTION

Location is an important part of business data. Business always happens somewhere. Because smart devices like phones, wearables, and Internet of Things (IOT) devices all produce location data, analyzing where things happen helps to create better understanding of the business.

Often, business data contains some aspects of location information. Taking advantage of this data helps an organization's data transformation and modernization and adds a new perspective of "where" to business analysis.

SAS Visual Analytics can help people of all backgrounds (such as business analysts, report authors, or data scientists) analyze big or small data. Users in SAS Visual Analytics can perform ad hoc data exploration, data discovery, and report creation. With the partnership with Esri, which is a leader in GIS technologies, SAS Visual Analytics is integrated with ArcGIS Online from Esri. The integration of analytically- rich SAS Visual Analytics with the leading capability of GIS ArcGIS Online helps SAS customers to get the best of two technologies.

SAS Visual Analytics supports two mapping technologies. Users can select OpenStreetMap maps or base maps from Esri ArcGIS Online. Many of the location technologies are supported with both of the mapping technologies. Some features require an ArcGIS Online user account. In this presentation we will note the features that require ArcGIS Online account.

GEO-CLUSTERING, a.k.a POINT-CLUSTERING

Geo-clustering is an analytical approach in SAS Visual Analytics that groups large number of data points on a geo map, based on the distance.

Consider a situation where you have 10,000 customers or a population of a million. Representing all these on a geo map will make visualization very difficult. Trying to displaying each point on a geo map,

makes the map almost illegible. If you use symbols on the map, they overlap each other, making it nearly impossible to identify patterns or to identify different measurement-based display rules.

How Does Point-Clustering Work in SAS Visual Analytics?

When you assign the category data and measure data item to a map object in the designer, the points are plotted on the map. Select the **Options** tab on the right side of the map. Expand the Map option and select the Cluster adjacent markers check box. See Figure 1 below.

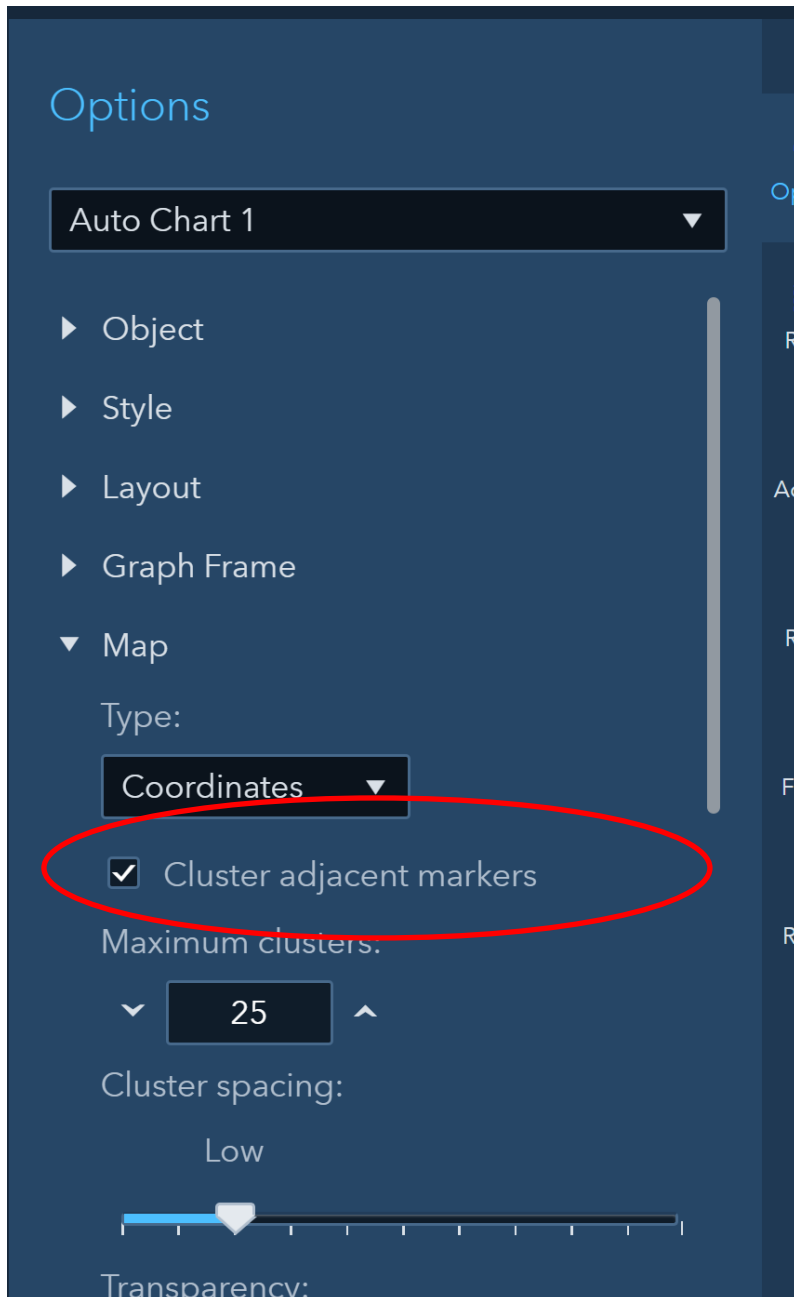


Figure 1: Cluster Option in Designer

After you select the option, SAS Visual Analytics runs the clustering algorithm and groups the points into circles. The size of the circle represents the number of points in the circle. The larger the number of points

that are close together in a geographic location, the bigger the circle size. The number in the center of the circle represents the number of data points in the circle. See Figure 2. The left picture represents the points on the map. The right picture represents the points after the clustering is complete.

Geo clustering in SAS Visual Analytics

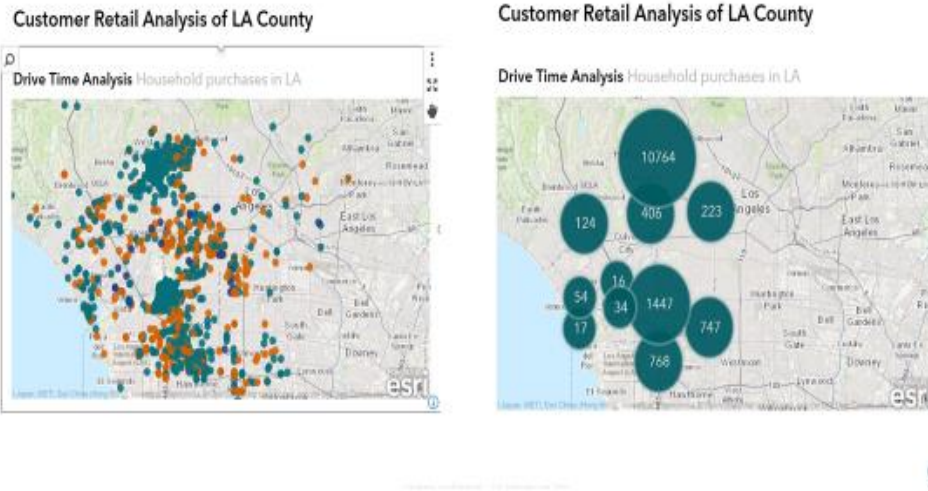


Figure 2: Geo Clustering

When you zoom on the map after the clustering, SAS Visual Analytics re-runs the clustering algorithm and re-draws the circles. The new clusters represent the new zoom level. As you continue to zoom in, the circles will de-aggregate into the individual data points.

DEMOGRAPHIC DATA INTEGRATION WITH ESRI

When you start analyzing business data related to customers or people or populations, you might have more questions to answer which you are difficult to answer using the data in your company data mart.

You might need to add demographic data and lifestyle data to your business data in order to get the insights for the people in a location. For example, you might want to analyze customers and people around a retail location based on ZIP codes or counties, and build insights by including the population income levels, ages, and gender information. These demographic numbers are not in business data and need to be added. The partnership with Esri enables SAS Visual Analytics to directly integrate with the ArcGIS Online demographic data. You can navigate the available data sets and select the fields you need. SAS Visual Analytics retrieves data and merges the data in-memory. You can continue to use this combined data in your analysis. You do not need to do any coding, but can just use the simple user interface.

How does this work?

In SAS Visual Analytics, select **Data** to open the **Open Data Source** window. Select the **Import** tab. When you select the **Esri** menu item, you will see the **Geo Enrichment** option. See Figure 3.

Open Data Source

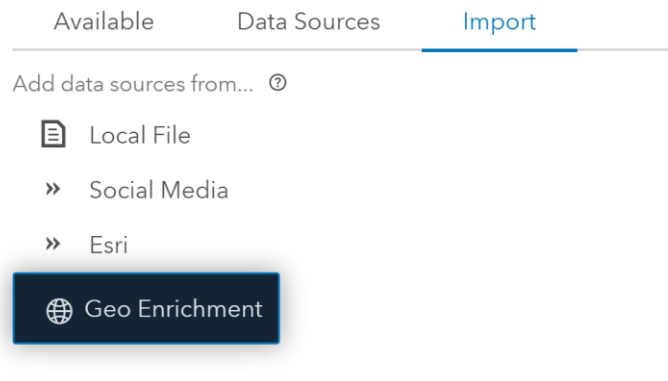


Figure 3: Geo Enrichment Option in Data Import window

Select Geo Enrichment to open the **Select Table** dialog box. Use this dialog box to select the server and library where your source table is located. Select your source table and click **OK** to open the **Esri Import** window, which displays the target table name and destination. See Figure 4.

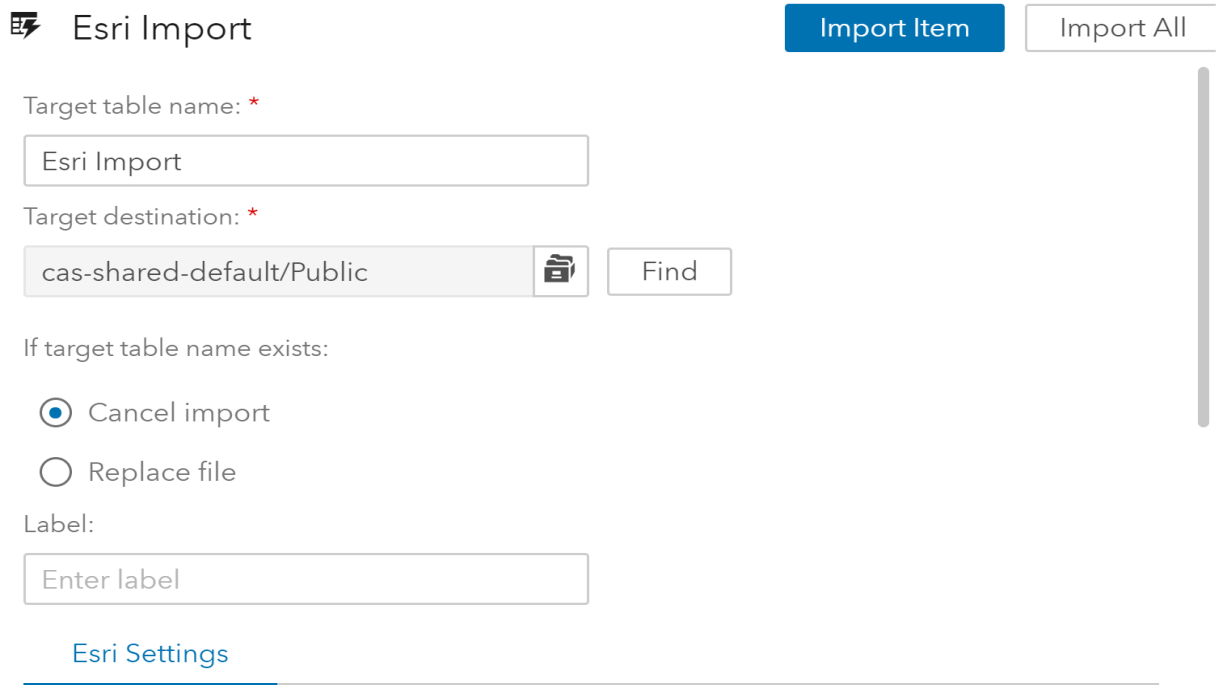


Figure 4: Esri Import Window

In the **Esri Settings** area of the **Esri Import** window, you can select the columns in the source table (for example, the ZIP code column) as well as value for the **Country**, **Level**, and **Category** fields. These fields are provided to filter and navigate the available data sets that you can use. For a given country there could be thousands of tables available. See Figure 5.

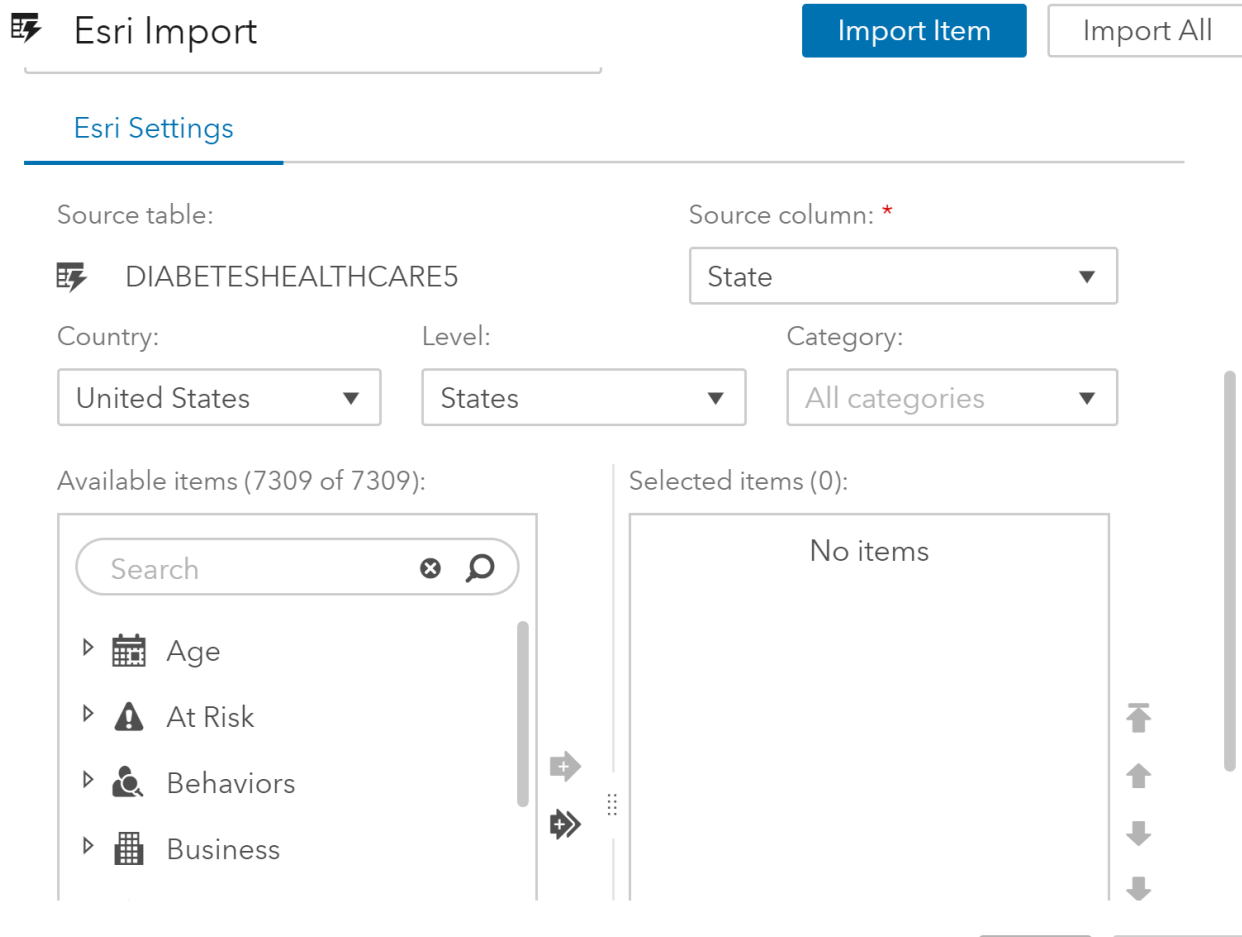


Figure 5: Fields for Filtering and Navigating Tables

Selecting the data for the Esri import enables you to include location data on your geo map. See Figure 6.

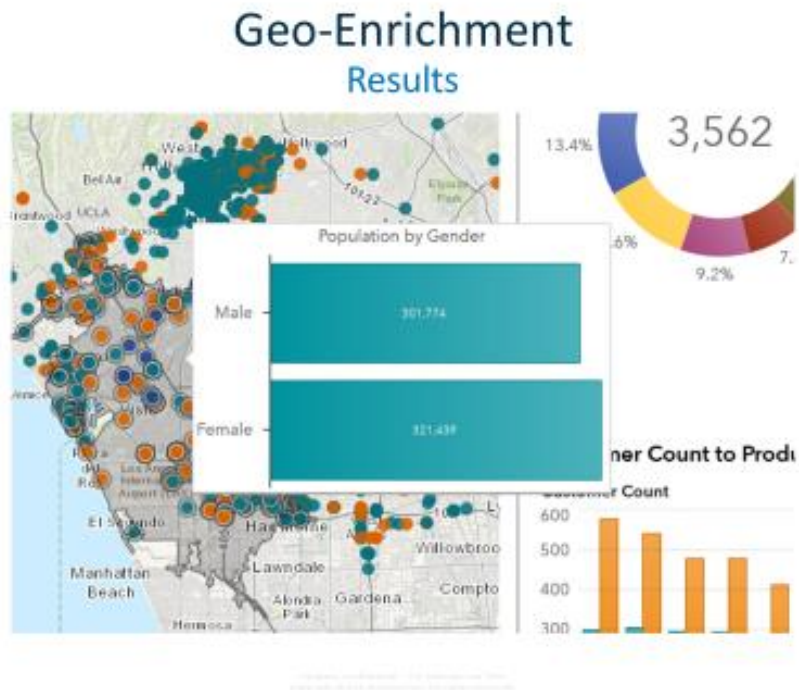


Figure 6. Location Data on a Geo Map

SYMBOLS a.k.a ICONS a.k.a POINTS-OF-INTEREST

Symbols are icons placed on geo maps for specific points of interest. When you display different data points, it is a good practice to use icons to represent the points of interest so that these points are visualized quickly by users. Using icons will greatly enhance the readability and get a quick insight of the data on the geo maps.

For example, if you are working with data related to a city and want to get an insight of the geographic presence of both happy customers and unhappy customers on the map, you can use different icons to represent happy (😊) and unhappy customers (😞). SAS Visual Analytics allows you to select different icons from the icon library.

How Does Selecting Icons Work?

In the designer, select the map object, and assign the roles. After the map is rendered, open the **Display Rules** option on the right pane. Select the geo item that is assigned with the geo map and select a new rule. Select the color for the rule. The dialog box for selecting the color contains two tabs - **Color** and **Icon**. Select the Icon tab to choose an icon for the rule. See Figure 7.

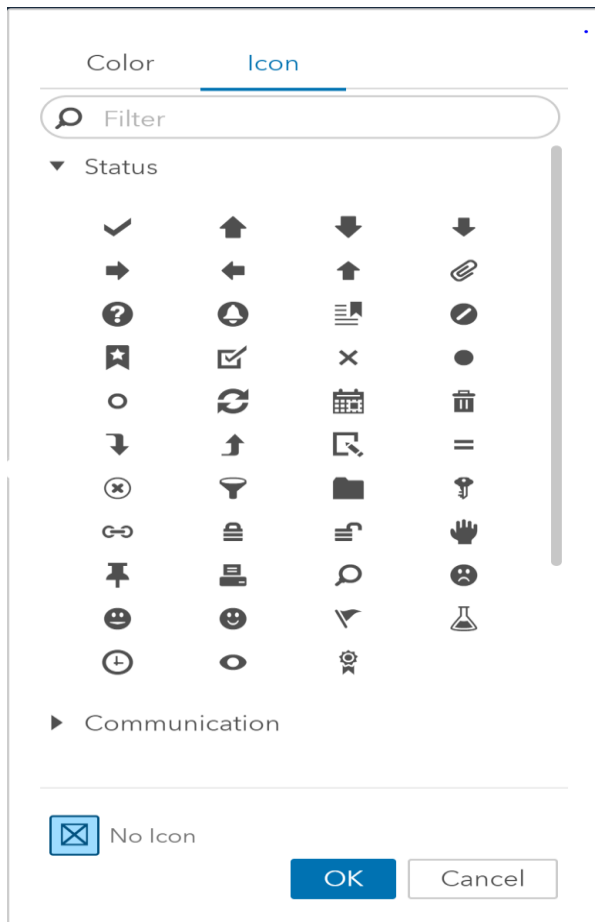


Figure 7: Available Icons as Part of the Display Rule

TRAVEL-TIME AND TRAVEL-DISTANCE ANALYTICS

If you are a business analyst for a retail company, or an analyst working for a state or local government performing a task such as analyzing the crimes around a school location, you might be interested in the area within a certain distance or a certain time to travel from a specific location. Being able to analyze this area gives you the ability to focus on the area for deeper understanding and better decision making.

SAS Visual Analytics support travel-time and travel-distance analysis. Select a point on a geo map and use the pop-up menu to select the analysis you are interested in. For example, if you are interested in analyzing the customers who are within a five-mile driving distance from a given location, use travel-distance analysis. If you are interested in analyzing all customers who are within 10 minutes from the point of interest, use travel-time analysis.

How does Travel-distance and Travel-time Analysis Work?

Select a point on the map. Click the right mouse button, select **Selection**, and select **Create geographic selection**. In the **Create geographic selection** dialog box, select Travel-distance or Travel-time, depending on your situation. See Figure 8.

Customer Retail Analysis of LA County

Product Purchase

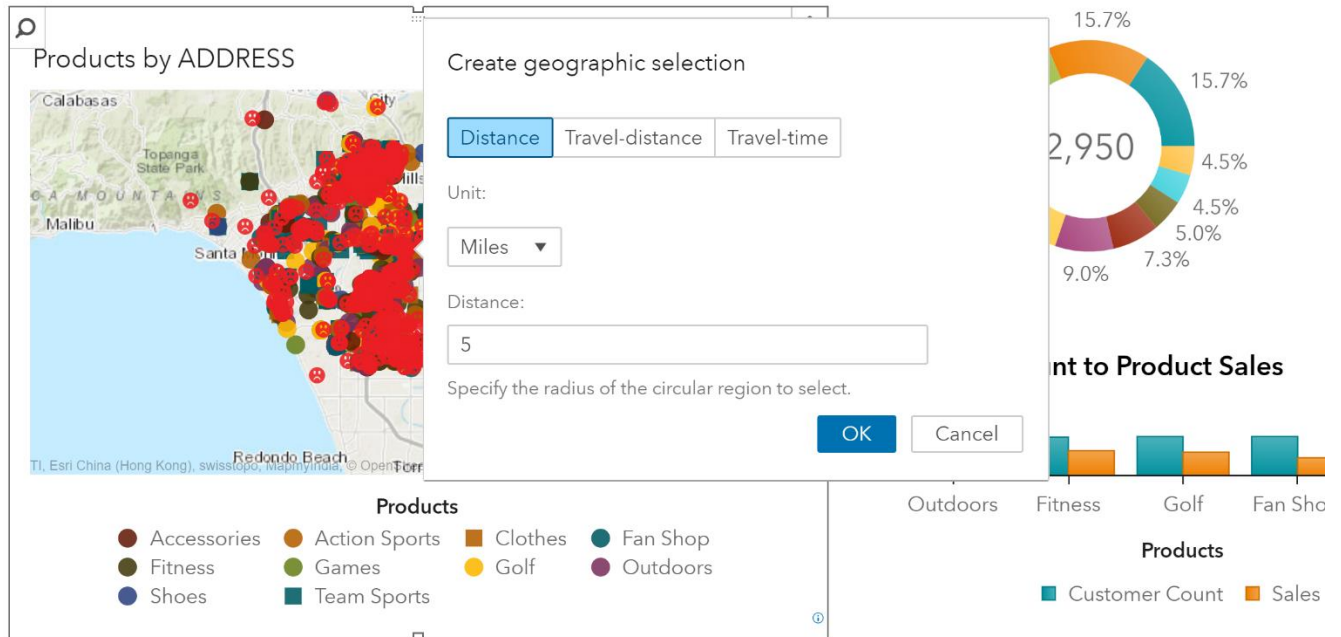


Figure 8: Travel-time and Travel-distance Selection Dialog Box

After the travel time is selected, the algorithm is run using Esri ArcGIS Online. A polygon is drawn on the map based on the distance that can be covered in a given time on ground. If you select multiple travel times (such as five minutes and ten minutes), multiple polygons are drawn, with each one representing a different travel time. See Figure 9.

Drive-Time Analysis with reporting in SAS Visual Analytics

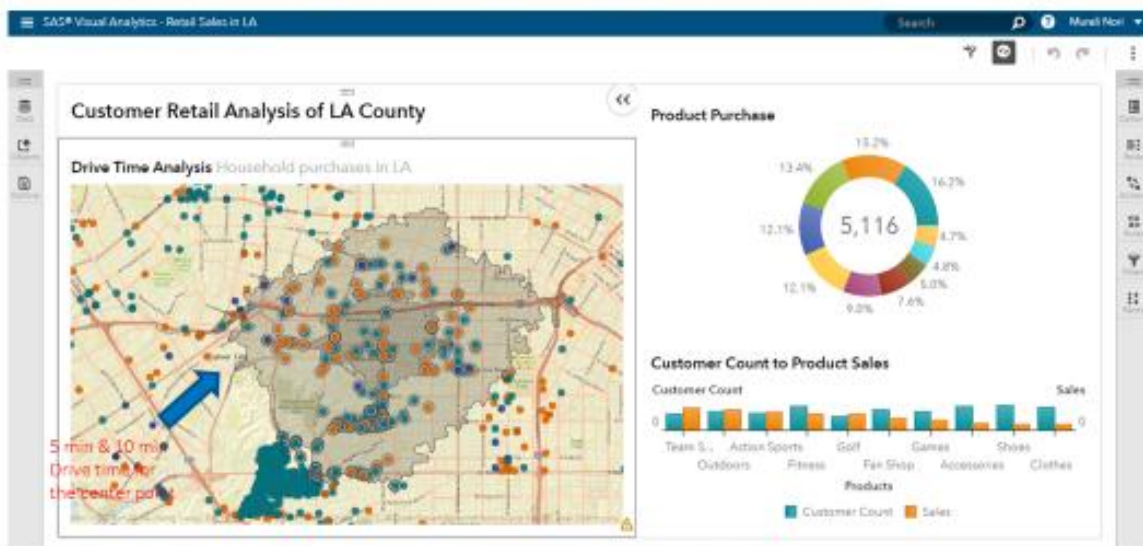


Figure 9: Travel-time Polygons for Five-Minute and Ten-Minute Drive Times

CUSTOM REGIONS ON GEO MAPS

In SAS Visual Analytics, you can render custom regions on a geo map. You can use the user interface in SAS Visual Analytics to import and register custom shape files. See Figure 10.

See SAS Global Forum paper SAS1732-2018, “Leverage Custom Geographical Polygons in SAS® Visual Analytics”, by Falko Schulz.

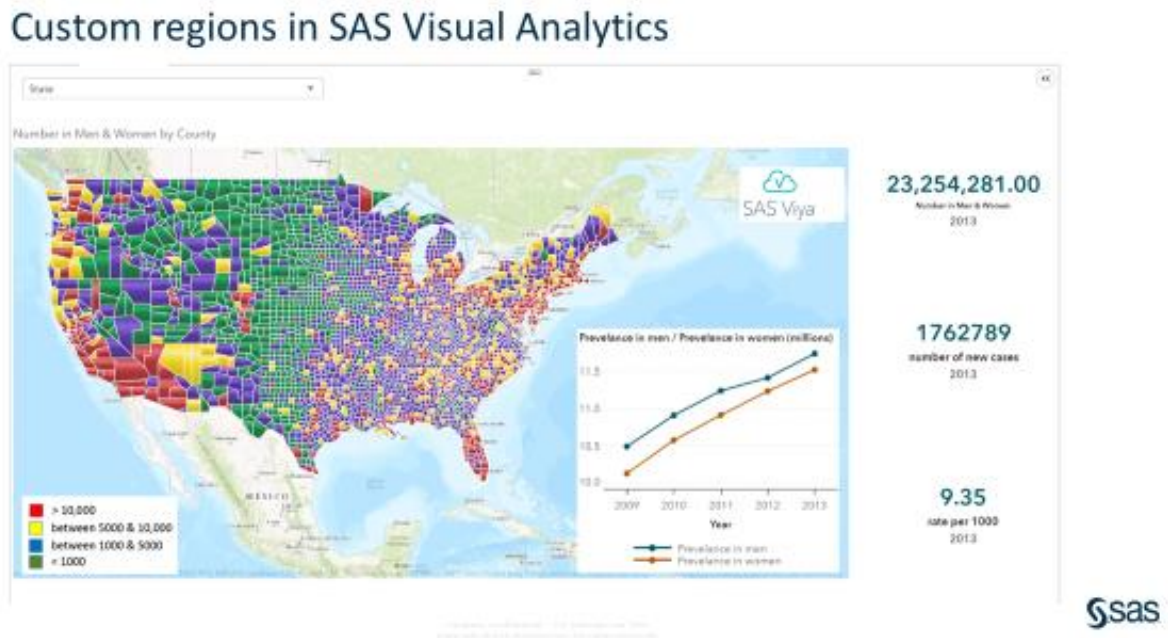


Figure 10. Using Custom Polygons on a Geo Map

GEO CODING

This capability is planned for Visual Analytics in future releases post 8.2 version.

When you need to place points on a geo map, you will need the Latitude and Longitude coordinates for the points. I refer to this process of obtaining these Lat & long coordinates for the data points as Geo coding. For example, when you have the address of different customers to be displayed on a geo map, you want to convert the addresses to Lat & Longs.

There are different approaches to doing the geo coding. One such method is to perform geo coding at a table as part of the Data Preparation. When you do geo coding on the table, the Lat & Long information is added to the table as additional columns and saved with the table. Any one who is try to use this table for analysis, reporting or exploration can use the geo coded data.

CONCLUSION

In today's business world, data is multiplying, and location information is available through IOT devices, smartphones, wearables, and other smart sensors and devices. Because so many of these new location data sources are available, being able to use this data together with other business data will provide better and deeper insights into how and why a business performs in a region.

Location analytics will be an important component of business analysis, and SAS Visual Analytics can help customers use the business data and location data seamlessly. SAS Visual Analytics provides a range of advanced analytics such as forecasting, clustering, correlation, network, and text analytics. Using these analytics together with location-specific analytics will extend the use and reach of analytics in any organization.

REFERENCES

Schulz, Falko. 2018. "Leverage Custom Geographical Polygons in SAS Visual Analytics." Proceedings of the SAS Global Forum 2018 Conference. Cary, NC: SAS Institute Inc. Available <http://support.sas.com/resources/papers/proceedings18>

SAS Institute Inc. 2017. *SAS® Visual Analytics: Working With Report Content*. Cary, NC: SAS Institute Inc. Available: <http://go.documentation.sas.com/?cdclid=vacdc&cdcVersion=8.2&docsetId=vaobj&docsetTarget=n14ezqxeg2a195n1au1q980gbkte.htm&locale=en>

SAS Institute Inc. 2017. *SAS® Visual Analytics: Working with Report Data*. Cary NC: SAS Institute Inc. Available: <http://go.documentation.sas.com/?cdclid=vacdc&cdcVersion=8.2&docsetId=vareportdata&docsetTarget=p031vp9uc5y5iun0zipy3c1trkqn.htm&locale=en#p1a7fw0vj2w6aln18sl77suu45o3>

SAS Institute Inc. 2017. *SAS® Visual Analytics 7.4: Administration Guide*. Cary NC: SAS Institute Inc. Available: <https://support.sas.com/documentation/cdl/en/vaag/69958/PDF/default/vaag.pdf>

SAS Institute Inc. 2017: *SAS® 9.4 Programming, SAS/GRAPH: Mapping Reference*. Cary NC: SAS Institute Inc. Available: <http://go.documentation.sas.com/?cdclid=pgmmvacdc&cdcVersion=9.4&docsetId=grmapref&docsetTarget=p0qsj9oazola04n10usle21dtk67.htm&locale=en>

SAS Institute Inc. 2017: *SAS Viya™ 3.3 Administration: Data*. Cary NC: SAS Institute Inc. Available: <http://pubshelpcenter.unx.sas.com:8080/test/?docsetId=caldatamgmtcas&docsetTarget=p1dwawsidszlpn121j0gllcixp.htm&docsetVersion=3.3&locale=en>

ACKNOWLEDGMENTS

I would like to thank Jeff Phillips and Eric Short for their assistance about details about geo analysis and their critical role in improving the geographical visualization component and associated map algorithms.

RECOMMENDED READING

- *SAS® Viya™ 3.3 Administration*
- [Esri ArcGIS Online](#)

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Murali Nori,
100 SAS Campus Drive
Cary, NC 27513
SAS Institute Inc.
Murali.Nori@sas.com
<http://www.sas.com>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.