

# **An Introduction to SAS<sup>®</sup> Visual Analytics**

**How to Explore Numbers, Design Reports,  
and Gain Insight into Your Data**

**Tricia Aanderud · Rob Collum · Ryan Kumpfmiller**

The correct bibliographic citation for this manual is as follows: Aanderud, Tricia, Rob Collum, and Ryan Kumpfmiller. 2017. *An Introduction to SAS® Visual Analytics: How to Explore Numbers, Design Reports, and Gain Insight into Your Data*. Cary, NC: SAS Institute Inc.

**An Introduction to SAS® Visual Analytics: How to Explore Numbers, Design Reports, and Gain Insight into Your Data**

Copyright © 2017, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-62960-291-2 (Hard copy)

ISBN 978-1-63526-042-7 (EPUB)

ISBN 978-1-63526-043-4 (MOBI)

ISBN 978-1-63526-044-1 (PDF)

All Rights Reserved. Produced in the United States of America.

**For a hard copy book:** No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

**For a web download or e-book:** Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

**U.S. Government License Rights; Restricted Rights:** The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication, or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a), and DFAR 227.7202-4, and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

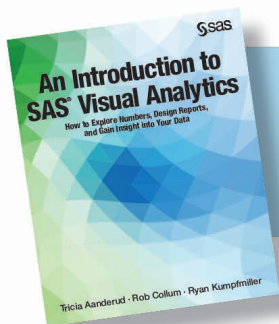
SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

March 2017

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.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to <http://support.sas.com/thirdpartylicenses>.



An Introduction to SAS® Visual Analytics: How to Explore Numbers, Design Reports, and Gain Insight into Your Data. Full book available for purchase [here](#).

## contents

about this book.....	ix
about these authors.....	xiii
acknowledgements .....	xv
introduction .....	xvii
<b>part one: getting started .....</b>	<b>1</b>
<b>chapter one: accessing content.....</b>	<b>3</b>
Methods of accessing content .....	3
Accessing content with a web browser .....	3
Accessing content through the public portal .....	3
Accessing content with the mobile bi app.....	4
Understanding roles .....	4
Accessing SAS Visual Analytics.....	5
Transformation of the homepage .....	5
Understanding SAS home .....	7
Opening a report.....	7
Creating a shortcut .....	8
Creating a collection or content tile .....	9
Using the report viewer .....	11
Navigating a report.....	12
References.....	14
<b>chapter two: building your first report .....</b>	<b>15</b>
Accessing the designer .....	15
Introducing the designer layout .....	16
Building your first report .....	21
Adding a data source .....	22
Creating new data items.....	25
Populating your objects .....	27
Improving the data object appearance .....	28
Adding data to other objects.....	29
Working with data objects.....	31
Adding object interactions .....	33

Saving the report.....	35
Reviewing the report .....	36
References.....	36
<b>chapter three: building your first dashboard.....</b>	<b>37</b>
Dashboard building process.....	38
Understanding your customer .....	38
Establishing objectives .....	38
Tips for more useable dashboards.....	42
Building the dashboard .....	43
Adding the data objects.....	44
Creating the layout.....	52
Working with data objects.....	56
Linking to another section.....	65
Other dashboard enhancements.....	66
Adding text boxes.....	67
Adding artwork .....	67
Embedding a stored process .....	68
Summary .....	69
References.....	70
<b>chapter four: using the data builder .....</b>	<b>71</b>
Using the Data Builder.....	71
Creating a data query.....	72
Opening the Data Builder .....	73
Filtering the data.....	80
Create a summary data query .....	82
Updating the code .....	84
Scheduling a query.....	84
References.....	86
<b>part two: customizing your data visualizations .....</b>	<b>87</b>
<b>chapter five: visualizing your data .....</b>	<b>89</b>
Elements of an effective data visualization.....	89
Your message: know your point .....	89
Your audience: know who is listening.....	90
Your technique: follow the KISS principle .....	90
Line charts .....	91
Interpreting the results .....	91
Line charts: guidelines .....	91
Line charts: tips and tricks .....	95
Bar charts .....	98
Interpreting the results .....	99
Bar charts: guidelines .....	99
Bar charts: tips and tricks .....	101
Pie and donut charts.....	103
Interpreting the results .....	103

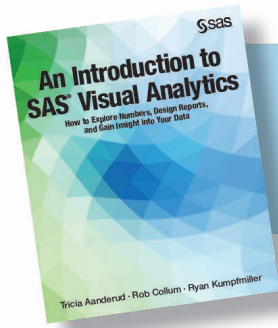
Pie and donut charts: guidelines .....	104
Pie and donut charts: tips and tricks.....	106
Treemaps .....	109
Interpreting the results .....	109
Treemaps: guidelines .....	110
Treemaps: tips and tricks .....	111
Waterfall charts .....	113
Interpreting the results .....	113
Waterfall charts: guidelines for use .....	114
Waterfall charts: tips and tricks .....	115
Gauges .....	118
Interpreting results.....	118
Gauges: Guidelines.....	119
Gauges: tips and tricks .....	120
Tip 1: Use display rules .....	120
Tip 2: Add a shared rule .....	121
Tables and cross tabs .....	122
Interpreting the results .....	122
Tables and crosstabs: guidelines for use .....	123
Tables and crosstabs: tips and tricks .....	124
Bubble plots.....	127
Interpreting the results .....	127
Bubble plots: guidelines.....	128
Bubble plots: tips and tricks.....	129
References.....	130
<b>chapter six: the where of data .....</b>	<b>131</b>
Using geospatial data effectively .....	131
When location is not part of the data story .....	131
When location is the data story .....	132
Preparing data for geospatial visualizations.....	133
Creating a predefined geographic data item.....	133
Creating a predefined geographic data item.....	134
Creating a custom geospatial data item .....	135
Creating a custom geographic data item .....	136
Displaying geospatial objects.....	137
Get to the point with geo coordinate data objects .....	138
Compare area with geo regional data objects .....	140
Show overall trends with bubble plots data objects .....	143
Expanding location intelligence .....	144
Understanding details about mapping technologies .....	145
References.....	146
<b>chapter seven: approachable analytics .....</b>	<b>147</b>
About the Explorer .....	147
Automatic chart feature .....	149

Box plots .....	150
Interpreting the results .....	150
Adding more data items.....	152
When to use box plots .....	154
Histograms .....	154
Changing objects in a visualization .....	154
Histogram options .....	155
Using a correlation matrix.....	156
Calculating a correlation.....	156
Understanding the matrix .....	157
Interpreting a correlation value.....	158
Forecasting .....	159
Working with the forecasting option .....	159
Using the scenario analysis.....	161
Word clouds.....	163
Loading social media data.....	163
Setting up the word cloud .....	164
Scatter plot .....	167
Data analysis .....	168
Adding categories .....	170
Heat map.....	171
Data analysis .....	171
Using a category.....	173
Other tips when using the Explorer .....	173
Include and exclude .....	173
Moving visualizations to the Designer .....	175
References.....	177
<b>part three: administration and data loading .....</b>	<b>179</b>
<b>chapter eight: loading data.....</b>	<b>181</b>
In-memory is different .....	181
It's about speed.....	181
Understanding the non-distributed deployment .....	182
Understanding the distributed deployment.....	182
Loading data to LASR from HDFS.....	183
Enabling support for SASHDAT files .....	184
The exception to the rule .....	185
SASHDAT does not require SAS/ACCESS.....	185
Loading data to LASR from Base SAS.....	186
Loading data to LASR with SAS In-Database technology .....	188
Loading data to LASR from a different LASR Analytic Server .....	189
Loading data into LASR automatically .....	190
SAS Autoload to LASR facility.....	190
LASR Reload-on-Start feature .....	191

References.....	192
<b>chapter nine: LASR administration .....</b>	<b>193</b>
Administration overview.....	193
Administration tools.....	193
SAS Management Console.....	194
SAS Visual Analytics Administrator .....	194
SAS Environment Manager.....	195
SAS Program Code .....	196
Other tools.....	197
Interesting LASR Administration Tasks.....	198
The role of SAS metadata.....	198
Defining new LASR Analytic Servers.....	198
Defining new LASR libraries.....	202
Managing LASR Analytic Servers with code .....	202
Working with the Autoloader Facility .....	205
Monitoring resources used by LASR.....	206
References.....	212
<b>chapter ten: performance considerations .....</b>	<b>213</b>
LASR performance.....	213
Non-Distributed LASR (SMP) .....	214
Distributed LASR (MPP) .....	215
Load balancing by data distribution.....	215
High-volume access to smaller tables .....	216
Fast loading of data to distributed LASR Analytic Server .....	218
LASR and a remote data provider (asymmetric) .....	219
LASR symmetrically co-located with HDFS .....	220
LASR co-located with dedicated HDFS and loading data from remote HDFS .....	220
References.....	221
<b>part four: SAS Visual Analytics 8.1 .....</b>	<b>223</b>
<b>chapter eleven: introducing the SAS Viya platform .....</b>	<b>225</b>
Overview of the SAS Viya platform .....	225
Understanding the CAS In-Memory Analytics Server .....	226
Introducing massively parallel analytics .....	226
Adding persistence.....	227
Providing more flexibility .....	227
SAS Viya and SAS 9.4 together.....	228
Managing the SAS Viya environment.....	228
Managing users and groups.....	229
Managing data .....	230
Managing content.....	231
References.....	232
<b>chapter twelve: wrangling your data.....</b>	<b>233</b>
Introducing a modern user interface.....	233

New features .....	241
References.....	248
<b>chapter thirteen: visualizing and exploring your data.....</b>	<b>249</b>
Introducing the new layout.....	249
Starting a new report .....	251
All-in-one application .....	256
Additional features .....	261
References.....	263
<b>index .....</b>	<b>265</b>





An Introduction to SAS® Visual Analytics: How to Explore Numbers, Design Reports, and Gain Insight into Your Data. Full book available for purchase [here](#).

## chapter six

# the where of data

Traditional business intelligence systems have focused on answering the *who*, *what*, and *when* questions, but organizations often need to know the *where* of data as well. Businesses want to plan sales territories based on existing customers. School districts want to understand where students live in relation to the school. Companies want to know whether equipment failures are related to specific locations. All of these issues are related to the *where* of the data.

SAS Visual Analytics makes it easy to plot geospatial data, which can add a completely new element to your data visualizations and analysis. In a tabular report, multiple columns might represent customers, competitors, and demographic information. The tabular report might not reveal anything useful. But if you can geocode the data and overlay it on a map, you quickly see where the better customers are, where they are in relation to competitors, and the regions that provide the most market potential based on underlying demographics.

SAS Visual Analytics geo mapping capabilities are based on integration with two mapping technologies: OpenStreetMap and ESRI ArcGIS. The examples in this chapter use OpenStreetMap with SAS Visual Analytics 7.3 unless otherwise noted. In this chapter, you will learn how to create geographic data items and geospatial objects.

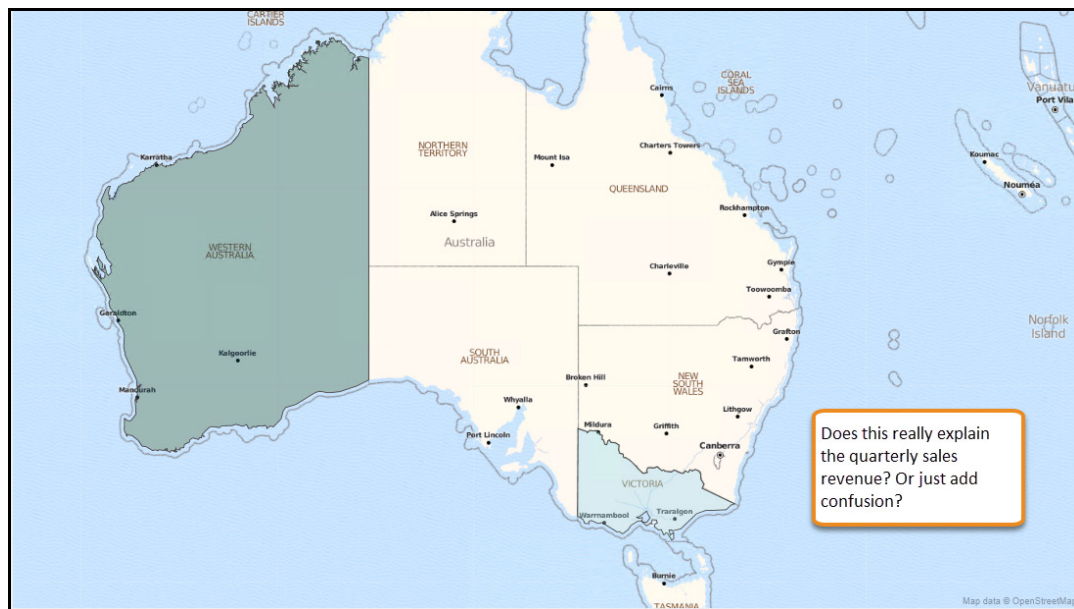
## Using geospatial data effectively

SAS Visual Analytics makes plotting geospatial data effortless. This is exciting! It is very easy to just use geospatial data objects for everything. Before traveling that path, though, consider if you really have a location-based data story. You want to use geospatial objects thoughtfully. In this topic, we will discuss an ineffective and an effective use of location analysis.

### When location is not part of the data story

In her book, *The Wall Street Guide to Information Graphics*, Dona Wong suggests that there are times when geography is not part of the story, so it doesn't make sense to force it to be. In her example, she shows two sales regions where sales were higher in one. In the following figure, the example was re-created using Australian states.

Figure 6.1 Location is not part of this data story



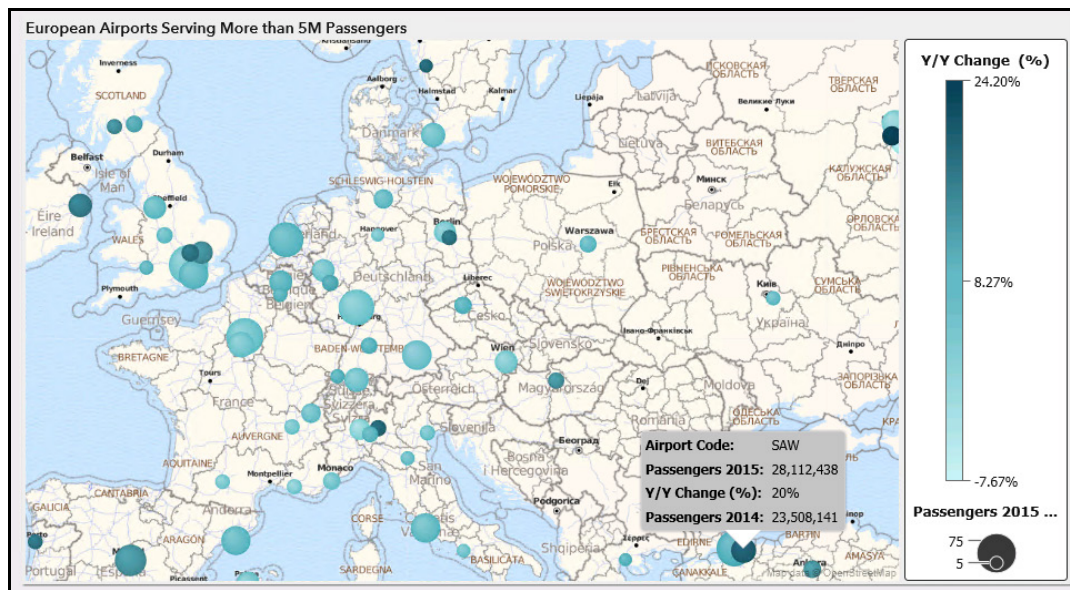
Because the regions are so disproportionate in size, comparing the sales revenue is not helpful. It does not lead to any conclusion except that Western Australia generated more revenue than Victoria. Your viewer does not have any useful takeaway, because the conclusion might be expected. It doesn't seem relevant to a data story in which the focus was really sales revenue. With only two values to display, would a list table or even a pie chart have been a better data visualization choice? The point of this example is to understand that even if you can show a cool geo object, you should ask yourself if it makes sense for your data story.

### When location is the data story

Europe is an international business center and a leading tourist destination. If you want to tell a data story about how popular a location it is, you could start by exploring the airport traffic. The Anna Aero website (<http://www.anna.aero>) contains data that details trends from most of the world's airports. You can use the airport codes and passenger counts to start a data story about the most and least popular cities.

In the following figure, the location of each airport is shown along with the passenger count and the difference from the previous year. In this example, location enhances the story. Instead of fancy calculations, the viewer can simply use their eyes to search for patterns.

Figure 6.2 Location matters in this story



There are multiple observations. You might notice the darker circles that show increased passenger counts. Perhaps you notice where there are multiple airports within a close radius, and you are curious why one has a higher passenger count. When used effectively, geospatial data can reveal previously unknown patterns or assist with confirming suspicions.

## Preparing data for geospatial visualizations

Before creating a geospatial visualization, you must have a geographic data item. If a data item contains a location, such as a country or state, then it is considered a geographic data item. Common location examples are customer addresses, store locations, or sales regions. You can use the SAS predefined geographic data elements or create custom geographic data items. This topic describes how to create each type.

### Creating a predefined geographic data item

To keep things easy, SAS Visual Analytics has predefined geographic data elements ranging from general values such as country names, to specific values such as ISO country codes. If your geographic data item contains a country name, then it can be matched to an internal table so that the location can be plotted on a map. Starting with SAS Visual Analytics 7.1, geographic data items can be country name, ISO 2-Letter codes, ISO Numeric Codes, or SAS Map ID values. You select the predefined method when you create the geographic data item.

These geographic data elements represent the center of the area. If you are showing France, it can be shown with a country outline or at the center of the country. The following table contains the available predefined geographical data elements provided by default and examples of how the incoming data items are expected to appear. The table shows the data values expected from three countries.

Geographic data element	Examples from Australia	Examples from the Netherlands	Examples from the United States
Country or Region Name	Australia	Netherlands	United States
Country or Region ISO 2-Letter Codes	AU	NL	US
Country or Region ISO Numeric Codes	036	528	840
Country or Region SAS Map ID Values	AU	NL	US
Subdivision (State, Province Names)	Queensland	Noord-Nederland	North Carolina
Subdivision (State, Province) SAS Map ID Values	AU-3	NL-01	US-37
US State Names	—	—	North Carolina
US State Abbreviations	—	—	NC
US ZIP Codes	—	—	27513

On the SAS Support site, there is a Geographic Lookup Values for SAS Visual Analytics (at <http://support.sas.com/rnd/datavisualization/vageo/71/VA71LookupValues.html>). This page contains a list of these values to help you understand your specific location. The tables at this site list the countries and the associated ISO numeric codes.

SAS Visual Analytics uses internal tables in the MAPSGFK library that is shipped with the product. You can review the tables in this library to ensure that your data matches the expected name by using an application such as SAS Studio. For additional assistance in creating geo data, you can use the GEOCODE procedure that is available with the SAS/GRAPH software.

#### What is an ISO code?

There is an international standard called ISO 3166 published by the International Organization for Standardization. This standard applies numeric values to countries and regions that everyone can use. There are several advantages to using numeric references, particularly in the data world.

If a programmer is using a non-Latin based language, such as Chinese or Hebrew, the number makes it easy to look up values. Also, when new countries form, a new number can be assigned while maintaining the older number for historical purposes.

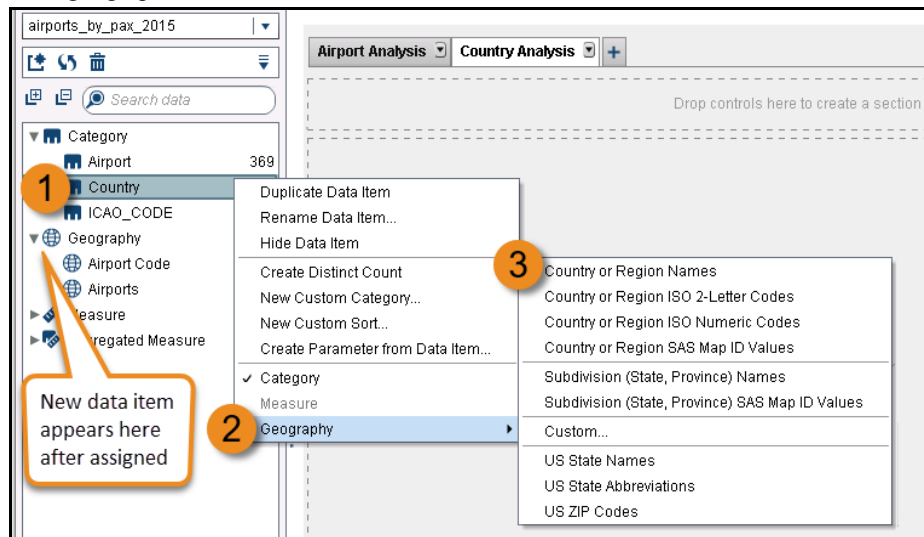
## Creating a predefined geographic data item

After importing data into SAS Visual Analytics, you must assign the data item to a Geography role before it can be used with any of the geo objects. You can create a geographic data item from an existing character or numeric data item. To create a geographic data item:

- 1 Right-click the data item that contains the geographic element that matches the predefined role. In this example, the Country data item contains the country names, such as Australia or Brazil.

**Note:** Some users prefer to duplicate the data item before assigning it to this role.

- 2 Select Geography ► Country or Region Names. Your data item is moved under the Geography section. You can use it with geographic roles.
- 3 Choose a geographic role.



## Dealing with location accuracy

If SAS Visual Analytics cannot plot your country data item, you might need to convert a country's common name to its official name. For example, Russia, United States of America, and Great Britain could be in your data set, but SAS Visual Analytics cannot plot them. When you search the country names in the MAPSGFK.WORLD data set, you learn that these countries use a different IDNAME. Often it is easier to convert the values to the ISO numeric code rather than using names.

Figure 6.3 MAPSGFK world data set values

	ID	IDNAME	ISO
1	AE	United Arab Emirates	784
2	GB	United Kingdom	826
3	RU	Russian Federation	643
4	TZ	Tanzania, United Republic of	834
5	US	United States	840

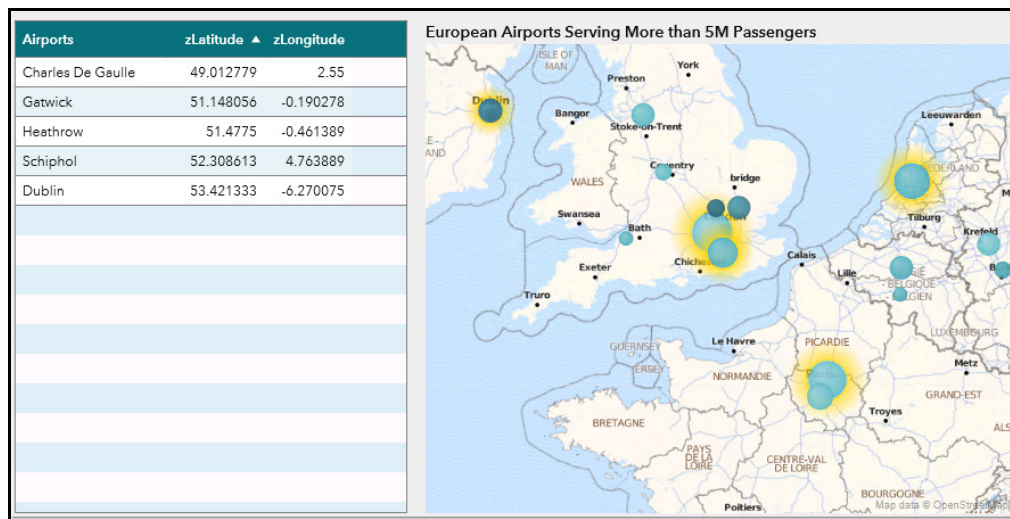
## Creating a custom geospatial data item

All geospatial data items represent a location on the planet Earth. A specific point or address has a set of coordinates, which are called latitude and longitude. You might recall from elementary school when you studied the globe and learned how it was divided by imaginary parallel lines that circle the globe from east to west (called latitude) and from north to south (called longitude).

When you provide a location's latitude and longitude, you are referencing these lines. If you think about the world's airports, it's possible to describe the geospatial location with just latitude and longitude coordinates. In the following figure, the airports are highlighted on the map, and the table on the left shows the airport name with its coordinates. Notice that the latitude numbers are similar because these airports are in a similar eastern European location. There is some variation in the longitude numbers as the airport is further south. Compare the Charles De Gaulle (Paris, France) coordinates to the Dublin (Dublin, Ireland) coordinates to better understand the values.



Figure 6.4 Airports with latitude and longitude



## Creating a custom geographic data item

To create a custom geographic item, you must have the latitude and longitude coordinates available in the data set. The coordinates can be based on the World Geodetic System (WGS84), Web Mercator, and the British National Grid (OSGB36). The default is the World Geodetic System (WGS84).

### What is a coordinate system?

There are three coordinate systems available for custom data points. These standards were developed for diverse purposes but are now commonly used.

- WGS84 was developed by the United States military for satellite-positioning systems.
- Web Mercator is a web standard. It was first used with Google Maps.
- OSGB36 is a British-developed system that is heavily used in British-based maps.

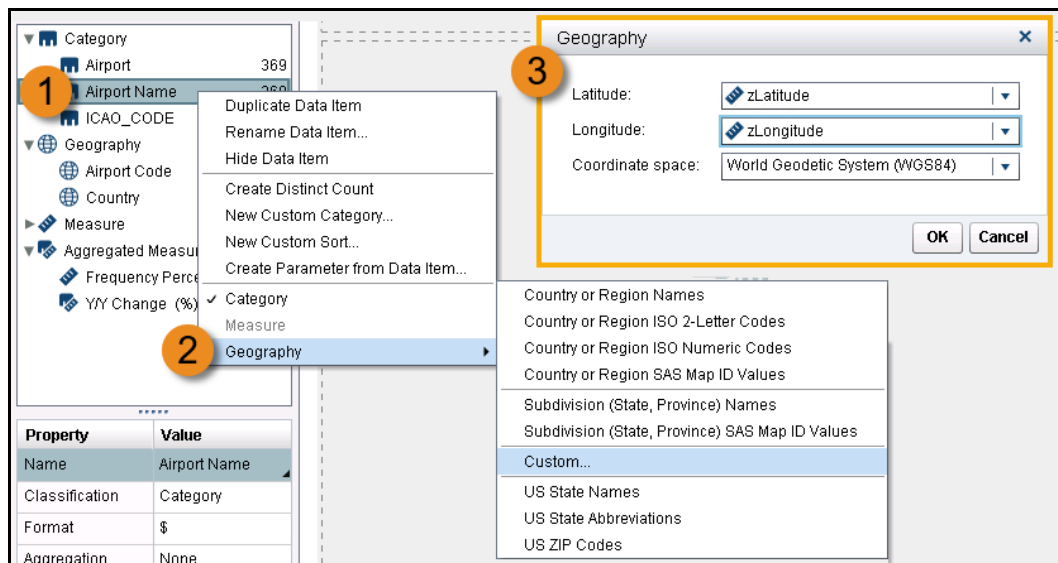
You should choose the system that works best for your specific location or geo data. In most cases the WGS84 system works.

Let's use the airport coordinates as the basis for the new geographic item. We use the airport name to create this data item, but other data items such as Airport Code would also work.

To create a custom geography data item:

- 1 Duplicate the Airport data item and name the new item Airport Name.
- 2 Right-click the new data item, and then select **Geography ► Custom**. A Geography window appears.
- 3 Select your data items for latitude and longitude in the appropriate fields. Your new data item appears in the Geography area.

Figure 6.5 Adding a custom data point



### Finding geo coordinates

If your data set does not have the geo coordinates available, you can get them through several sources.

- The SAS MAPSGFK library contains multiple countries and regions.
- There are open-source databases available that you can find with a web search.
- Google has an API that you can query through code. The free service has a daily access limit, but you can subscribe to their service or other commercial services.

## Displaying geospatial objects

There are three ways to display geo data in SAS Visual Analytics:

- Coordinate – Pinpoints an exact location on the map using a custom geographical item
- Regional – Outlines a regional area
- Bubble – Combines a bubble plot to show a value at the location

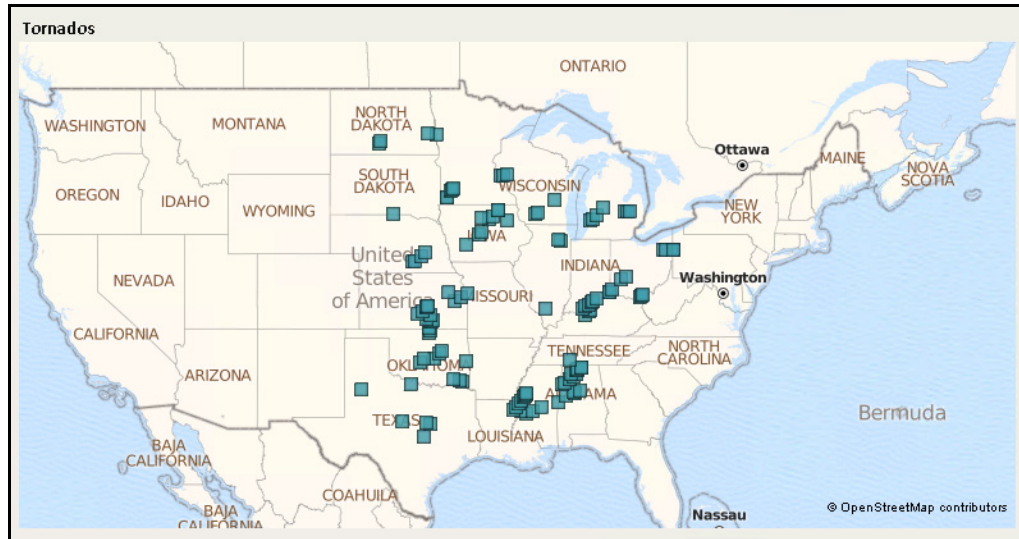
These objects enable you to highlight your geospatial data in different ways and for different stories. Let's explore the different ways that these data objects are used.

For the remaining topics in this chapter, the examples are created using the storm events data set from the United States National Climatic Data Center website. This database contains US storm events (such as tornadoes and thunderstorms) since 1950. The data set contains other facts such as the number of deaths or injuries and the estimated property damage. The tornadoes are rated by their intensity on the Fujita scale from F0 to F5 with F5 being the most destructive. In 2007, the Enhanced Fujita scale was introduced and tornadoes were categorized as EF0-EF5.

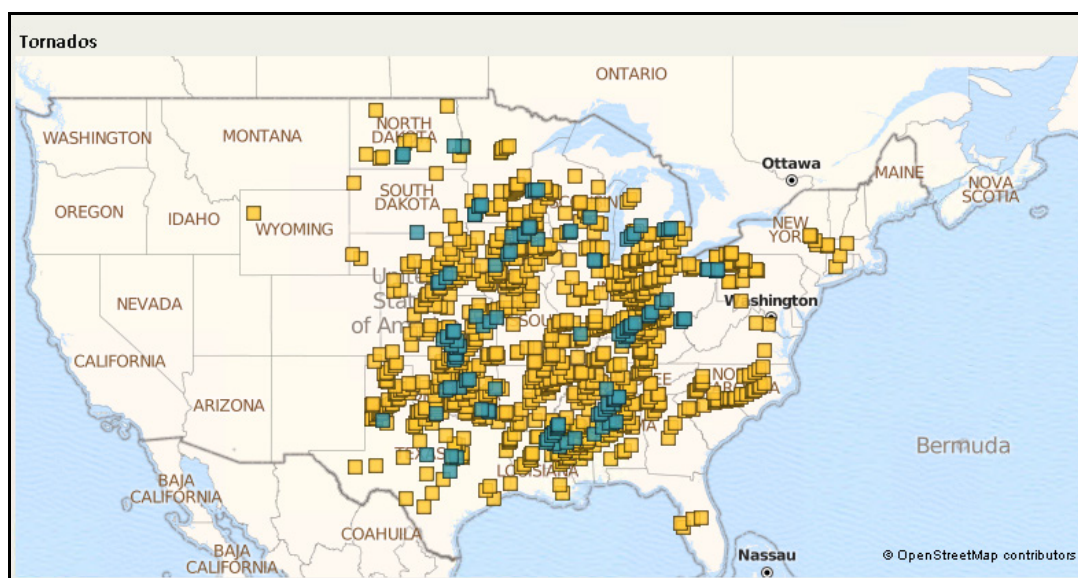
## Get to the point with geo coordinate data objects

Perhaps you've heard people talk about tornado alley—it's an area down the middle of the United States where tornadoes occur more frequently. Tornadoes are powerful and scary storms that produce wind speeds capable of sending a wood board through a metal car door. These storm events are responsible for massive property damage and loss of lives. Geo coordinate maps are excellent at showing exactly where an event occurred. In the following figure, the teal markers indicate where tornadoes with EF5/F5 strength of 230 mph+ (370 kph) winds arose in the past 50 years.

Figure 6.6 F5/EF5 tornado locations



To make this geo coordinate object a little more interesting, let's add the EF4/F4 tornado touchdown points for the same time period. By contrasting the teal and gold markers, the viewer sees that an EF5/F5 tornado is less common.



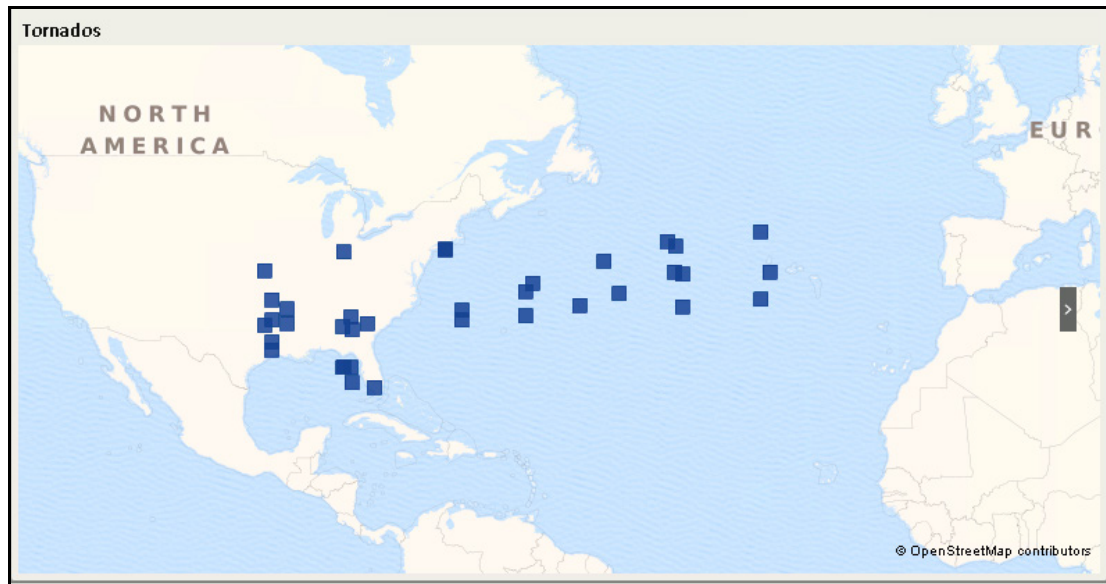
While the data points are chaotic, it's clear where severe tornadoes occurred. This data would not have had the same impression if we had plotted it as a line chart or even a pie chart. The touchdown points help you realize why those particular states have a higher disaster recovery budget.



### Tip 1: Dealing with odd locations

This data object uses a custom geography data item that is based on supplied latitude and longitude values. If the coordinates are incorrect, then the map might show your data in the middle of the ocean. In our sample data, some of the coordinates were entered incorrectly. This resulted in tornados appearing in the Atlantic Ocean. To correct this situation, the latitude and longitude values in the data set would have to be edited or filtered.

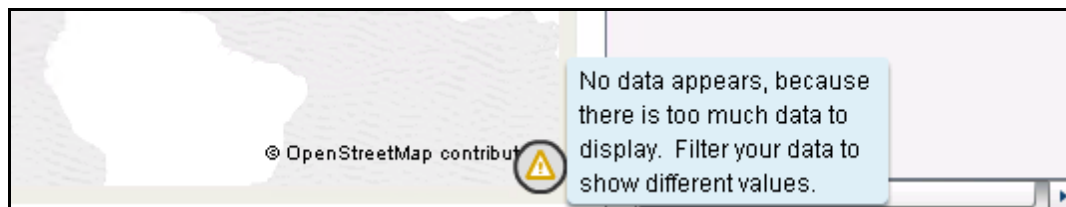
Figure 6.7 Tornados in the ocean



### Tip 2: Controlling the data

When you have too much data to display, SAS Visual Analytics issues a yellow icon and warns you to add some filters to your data. With custom geographic data items, it is more likely to happen. The solution is to control how much data appears at once by setting filters.

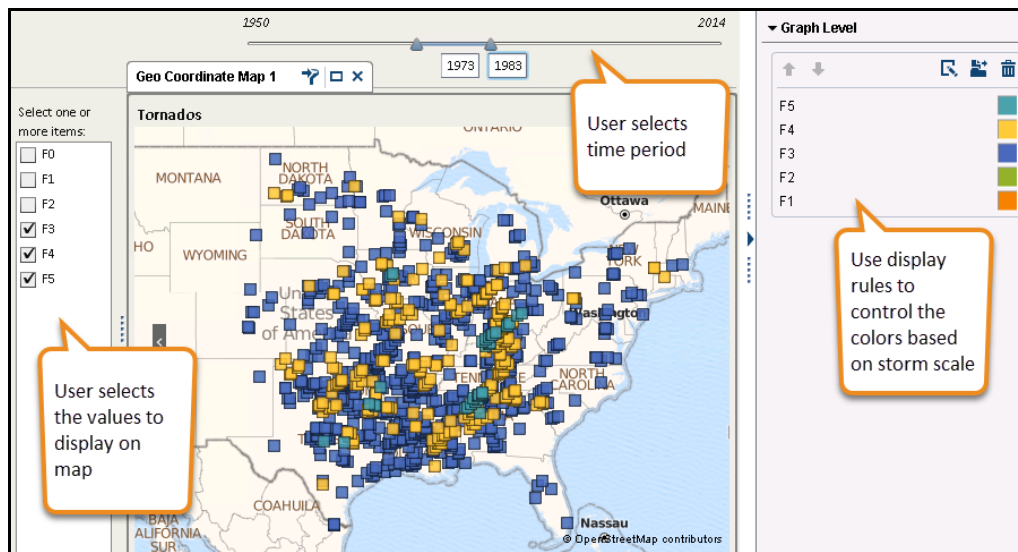
Figure 6.8 There is too much data at one time!



Here are a few suggested filters:

- Add a date range slider to compare events along a time scale. Adding Event Year to the slider enables the user to compare which years might have had more active storm seasons.
- Split the data item categories. Use the display rules to assign the tornado scale to a different color so that each level is clearer. Then add a List filter and assign the Tornado F/EF Scale to the list. Users can select which tornado scale they want to compare.

Figure 6.9 Add filters to keep data visualization manageable

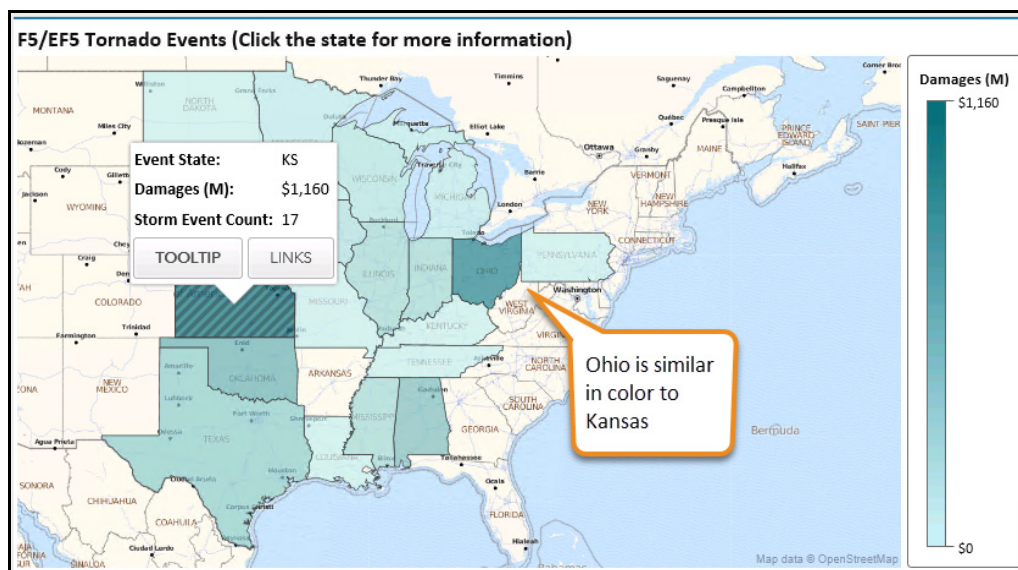


## Compare area with geo regional data objects

Use the geo regional data object when you need to introduce a subject about location. This geospatial object helps a viewer understand where to focus their attention or understand how much variation occurs for a value. These objects are also called *choropleth* maps, which is Greek for multitude of areas.

When you start thinking about dangerous storm events, you can imagine that these events cause considerable property damage. States more prone to severe tornadoes will plan larger disaster recovery budgets. It would be interesting to compare the damage costs by state. Using a geo regional map, you can place a value over an entire region, such as a country or a state. Color is then applied over the regions to indicate the intensity of the value.

Figure 6.10 Understanding regional events



In the preceding figure, you can see the associated property damage cost for the tornadoes across the areas. The darker the color, the costlier the storm damage. Use an average or percentage to make the values comparable or normalized. By exploring the visualization, you can easily see the areas of most damage, but it's harder to understand where there is the least damage. Be sure to use a legend so that the user understands the color range.

When you position your pointer over each state, a data tip appears that contains the assigned data items values. Since Ohio and Kansas are similar in color, viewers might be interested to learn more. Most of Kansas is farm land and rural areas, while Ohio is more densely populated and industrial. Being in tornado alley, Kansas probably experiences more tornados and thus more crop damage. With the larger population, it might be costlier for Ohio when there is an extreme storm event.

### Tip 1: Improving your geo regional map

There are a few settings that can make a geo regional data object a nicer user experience.

- Add data tips to provide more information when the user positions the pointer over content.  
You can add as many as you like, but make sure that the data items enhance instead of confuse the viewer. For the preceding example, we added the Storm Event Count as a data tip.
- Adjust the color transparency for the overlay so that the user can see the underlying values.  
If the underlying values are masked, it might cause confusion. For this visualization, the transparency was adjusted to 25%. It was just enough to maintain the color while still allowing the underlying value to peek through.
- Adjust the gradient color to ensure enough contrast.  
The ocean is a light blue, so a contrasting color that does not appear too similar to the landscape features is required. In Figure 6.9, a single color for the Gradient value is used. It is easier to decode a value when the color intensity increases as the value increases.

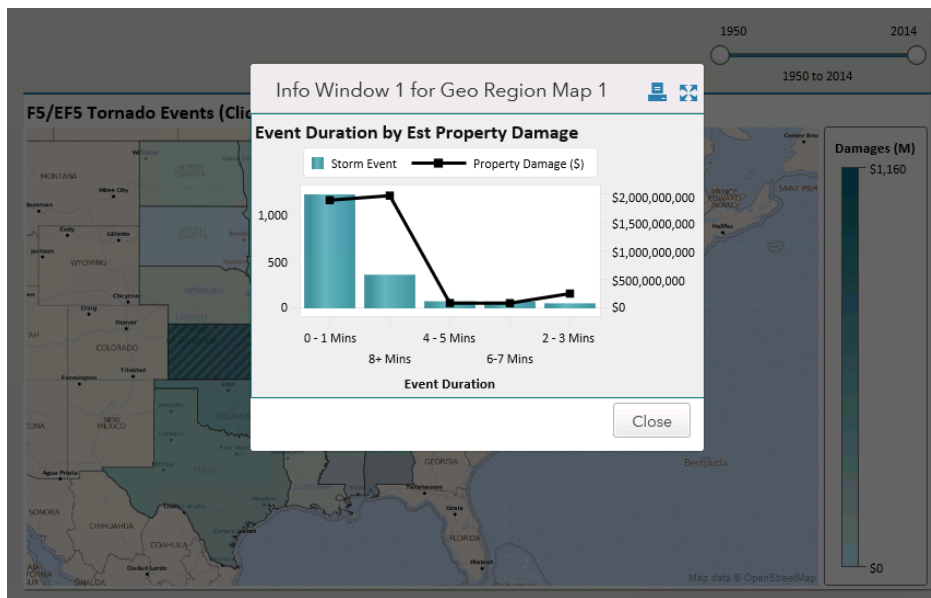
#### Choose lighter colors

In *Envisioning Information*, Edward Tufte has a fascinating discussion of color with maps. His suggestion is to use colors that are found in nature. He encourages using a color palette on the lighter side and provides several examples used across several centuries.

### Tip 2: Adding rich details for exploration

The geo region data object is excellent for getting the user to focus on specific areas. It leads to more questions about the storm events, so it might be convenient to use an info window to provide more details. This info window shows the storms by duration with estimated damage. A quick storm can result in as much damage as a longer one, although this probably depends on where the tornado touches down.

Figure 6.11 Use a pop-up window to provide more details



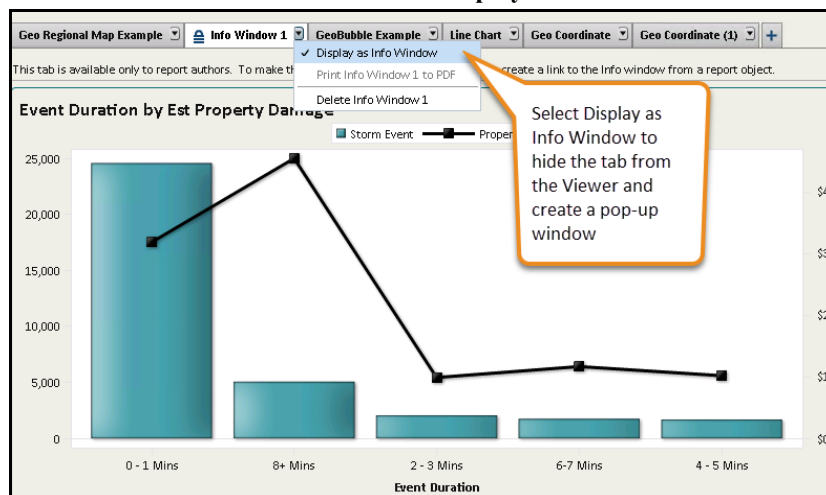
When the user clicks on the state, the info window appears with additional information. This example uses a bar-line chart, but it can be anything you can create in a section. This map is a good way to start a story. It provides an overview and helps the viewer understand where to focus their attention. In this case, it was Kansas and Ohio.

The only pitfall to an info window is that the viewer might not recall the values from the previous pop-up. Use this technique for data discovery or as a way to entice someone into your story. This data story is completely about the location and comparing how the events affected the states.

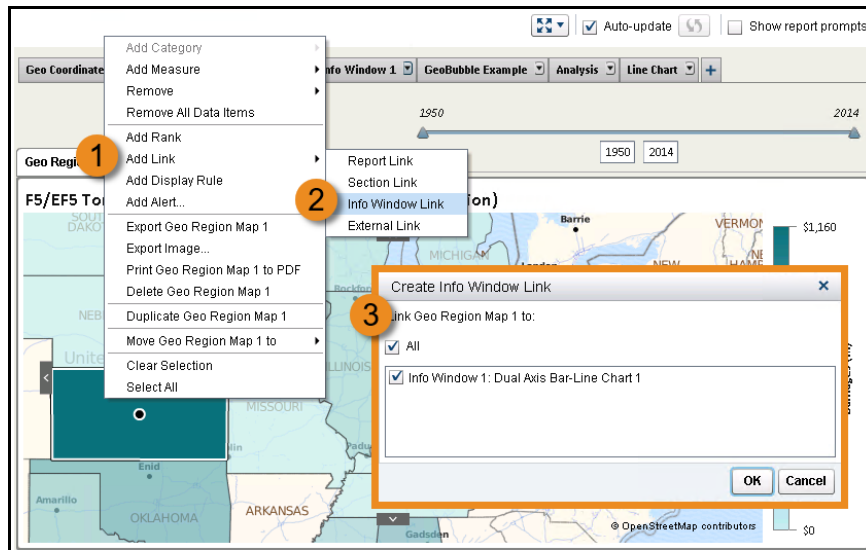
### Adding an info window to your map

Info windows are pages that you can link to from another page. Use the following steps to add an information window to your map.

1. Create a tab. In this example, the geo regional map was created.
2. Create another tab with the data objects of your choice. For this example, a bar-line data object was used to show the event duration and estimated property damage.
3. Select the down arrow next to the title and select **Display as Info Window**.



- Return to the page that you created in step 1. Right-click on the map, and then select Add Link > Info Window Link. In the window that appears, select which info window you want.



Once you turn the tab into an info window, it does not appear to the viewer. You can use an info window in other situations to provide information about the tab.

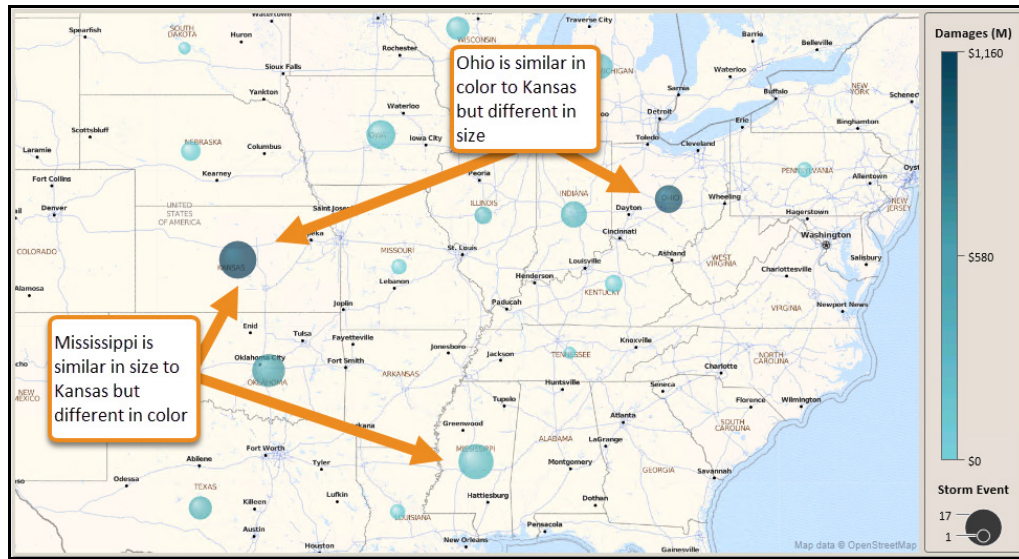
## Show overall trends with bubble plots data objects

Bubble plots receive a lot criticism for being difficult to understand. These charts can pack a lot of data into a few variables. A layperson might spend more time trying to understand a bubble plot, but this doesn't seem true for the geo bubble maps. Possibly it's because the user sees the map and understands that it is related to location.

In the previous topic, we created a geo regional map to show the average damage cost from F5 tornadoes for each state. One issue with the method was that users had to position the pointer over each state to see how many storm events were associated with each event. If a user wants details, it is a little awkward. A geo bubble map resolves this issue..

A geo bubble plot places a bubble on the geographic location and enables you to control two aspects of the bubble: its size and color. In the following example, the bubble size is the event count (the number of tornadoes) while the color shows the estimated property damages (shown with the scale). Now it is more apparent that Kansas endured a similar number of events as Mississippi, but the price tag was a little larger. However, it also shows that Ohio had a similar cost but fewer events than Kansas.





### Tip 1: Ensure that the legend is visible

When you use bubbles to encode data, you are asking the user to compare the bubble size and the bubble color. The legend ensures that the user has some visual cues to assist with understanding. You can place the legend anywhere around the object. In the preceding example, the legend is placed on the right.

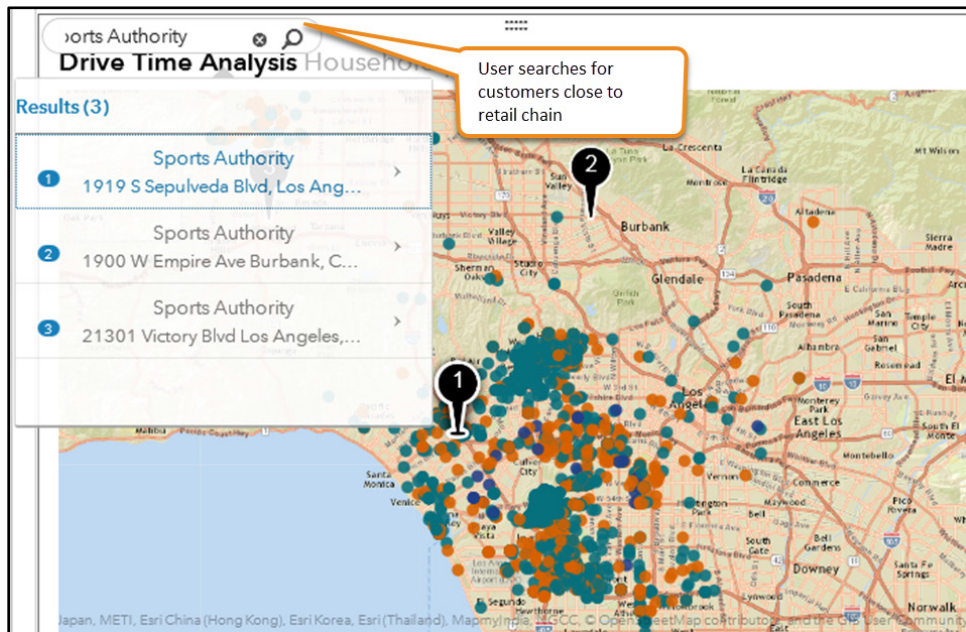
### Tip 2: Watch the default colors

By default, the geo bubble object uses a gradient scale of red to blue. This scale is acceptable when working with performance data and is commonly referred to as trafficliting. The color mimics the traffic signals where red means stop and green means go. However, we have a logic problem in this instance. The red indicates the least amount of damage and the blue indicates the most. Technically, any property damage is bad. (After all, we are not measuring how well the tornado was at damaging property!)

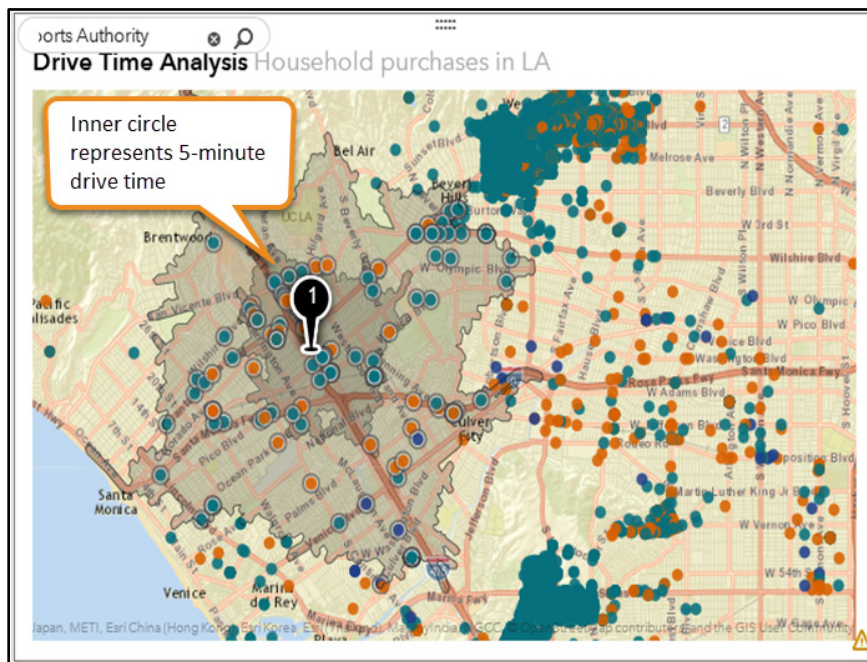
The gradient scale was changed to teal in our chart. The bubbles are not as close to the ocean and provide enough contrast with the circle. However, notice that the bubble over Tennessee is barely visible. The bubbles were set to 30% transparency to make the state names visible. Perhaps another color would be more suitable? You can experiment with your data object and decide.

## Expanding location intelligence

Starting with SAS Visual Analytics 8.1, users have unlimited access to the ESRI base maps from within SAS Visual Analytics. This provides geo search functionality and ad hoc selection of data points on a map. Here's an example of a geo search where the user was looking for how close Los Angeles, CA customers were to the retail chain Sports Authority.



For users who want additional functionality, they can subscribe to the ESRI premium features. The premium service offers drive-time analysis, drive-by-distance analysis, and the ability to create custom shapes. In this example, the user was looking for customers within a 5- to 10-minute driving distance of the store location. The darker inner area is the 5-minute drive, while the lighter inner band in the 10-minute drive.



## Understanding details about mapping technologies

SAS Visual Analytics geo mapping capabilities are based on integration with two mapping technologies: OpenStreetMap and ESRI ArcGIS. SAS Visual Analytics enables its users to view their enterprise data mapped across the various locations on the map.

**OpenStreetMap** This is an open-source project, where a worldwide user community maintains the data about roads, boundaries, trails, and much more.

**ESRI ArcGIS Maps** This advanced mapping platform uses highly interactive and informative geographical maps. The maps are maintained by ESRI, a SAS partner.

The SAS Visual Analytics environment must be configured to point to one of these mapping technologies. An OpenStreetMap server is hosted by SAS and is available as part of the default configuration. Organizations might want to host and maintain their own OpenStreetMap server. Organizations can also use the ESRI server (ArcGIS for Server, version 10.1 or higher) for access to maps. Refer to the *SAS Visual Analytics: Administration Guide* for your release for more configuration details.

Many SAS Visual Analytics users are concerned about what information from their data must be shared in order to retrieve map tiles from OpenStreetMap or ESRI ArcGIS Maps. After all, if the data is confidential to their enterprise, it needs to be kept secure. Fortunately, none of your actual data is leaked outside of the environment. SAS Visual Analytics simply requests the specific map tiles necessary to render the selected geographic area. The highlighted regions, bubble plots, and all are created within the SAS Visual Analytics application.

## References

- Aanderud, Tricia. 2016. "Where in the World is SAS Visual Analytics?" Available at <https://www.zencos.com/blog/review-geoplot-in-sas-visual-analytics/>.
- Massengill, Darrell. 2016. "The GEOCODE Procedure and SAS Visual Analytics." *Proceedings of the SAS Global Forum 2016 Conference*. Paper SAS3480-2016. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings16/SAS3480-2016.pdf>.
- Nori, Murali, and Himesh Patel. 2016. "Location, Location, Location—Analytics with SAS Visual Analytics and ESRI." *Proceedings of the SAS Global Forum 2016 Conference*. Paper SAS4060-2016. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings16/SAS4060-2016.pdf>.
- Schulz, Falko, and Anand Chitale. 2014. "More Than a Map: Location Intelligence with SAS Visual Analytics." *Proceedings of the SAS Global Forum 2014 Conference*. Paper SAS021-2014. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings14/SAS021-2014.pdf>.
- Tufte, Edward R. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press.
- US NOAA Data. Storm Events Database. Accessed 2016. See <https://www.ncdc.noaa.gov/stormevents/>.
- Wong, Dona M. 2013. *The Wall Street Journal Guide to Information Graphics: the Dos and Don'ts of Presenting Data, Facts, and Figures*. New York, NY: W. W. Norton & Company, Inc.



# about this book

SAS Visual Analytics is a data visualization and analytics application that enables organizations to understand and analyze their most important asset: their data! There is a wealth of information freely available to users on [support.sas.com](http://support.sas.com), everything from simple blog tips to in-depth user documentation. All of this information can actually make it more difficult for users to understand how to get the most power from the application, or even know where to start!

Our purpose in writing this book is to provide a single quick-start guide that moves you swiftly through the applications; we aim to teach you the basics so that you can then easily expand that knowledge with some intermediate and advanced techniques. We want to ensure that users understand how to get the most from the application for their organizations.

## Is this book for you?

New users or those with some familiarity with SAS Visual Analytics will find this book useful. This book provides a tutorial for new users. It also contains in-depth explanations of data visualizations and how to use them effectively.

If you have the administrative privileges for deploying, configuring, and maintaining your SAS Visual Analytics environment, then we've got you covered here as well. Our goal is to enhance your understanding of how the software works so that you can readily implement proven practices for efficient operation.

## Prerequisites

You do not need any previous experience with the Visual Analytics toolset or SAS to use this book. If you are familiar with SAS coding techniques and SAS data handling, you might find some areas of the product easier to understand.

## Scope of this book

This book covers the following topics:

- SAS Visual Analytics 7.3 (and later)
- SAS Visual Analytics 8.1

## About the examples

### Software used to develop the book's content

Reports and explorations for this book were created using the following software:

- SAS Visual Analytics 7.3 (and later)
- SAS Visual Analytics 8.1

SAS Enterprise Guide 7.1 and SAS Studio was used for some data preparation in this book. Most of the data sets were taken from existing sample libraries and no preparation was required. These sources are identified at first use.

The primary references used to write this book were the following sections of the online documentation for SAS Visual Analytics:

- SAS Visual Analytics User Guide, 7.3
- SAS Visual Analytics Administration Manual, 7.3

### Example data and reports

You can access an import package for the *Building Your First Report* and *Building Your First Dashboard* chapters in this book by linking to its author page at <http://support.sas.com/publishing/authors>. Select the name of the author. Then, look for the cover thumbnail of this book, and select Example Code and Data to display the SAS programs that are included in this book.

If you are unable to access the code through the website, email [saspress@sas.com](mailto:saspress@sas.com).

### We Want to Hear from You

SAS Press books are written *by* SAS Users *for* SAS Users. We welcome your participation in their development and your feedback on SAS Press books that you are using. Please visit <https://support.sas.com/publishing> to do the following:

- Sign up to review a book
- Recommend a topic
- Request information on how to become a SAS Press author
- Provide feedback on a book

Do you have questions about a SAS Press book that you are reading? Contact the author through [saspress@sas.com](mailto:saspress@sas.com) or [https://support.sas.com/author\\_feedback](https://support.sas.com/author_feedback).

SAS has many resources to help you find answers and expand your knowledge. If you need additional help, see our list of resources: <https://support.sas.com/publishing>.

## **Subscribe to the SAS Learning Report**

Receive up-to-date information about SAS training, certification, and publications via email by subscribing to the SAS Learning Report monthly eNewsletter. Read the archives and subscribe today at <http://support.sas.com/community/newsletters/training>!

## **Publish with SAS**

SAS is recruiting authors! Are you interested in writing a book? Visit <http://support.sas.com/saspress> for more information.



## about these authors



Tricia Aanderud, Director of the Data Visualization Practice at Zencos Consulting, provides SAS consulting services to organizations that need assistance understanding how to transform their data into meaningful reports and dashboards. She has been a SAS user since 2002, is a frequent speaker at SAS Global Forum and SAS regional user groups, and is the author of two SAS Press books and one self-published book. She has a background in technical communications, process engineering, and customer service. She has a BA in Mass Communications from Eastern Kentucky University.



Rob Collum is a Principal Technical Architect in the Professional Services Division at SAS. For the past twenty years, Rob has enabled the delivery of high-performance solutions that provide substantial value and meaningful impact to SAS customers around the world. He currently works alongside a team of professionals who partner with other divisions to identify, create, and standardize architectural practices for the newest SAS technologies. Rob received a Bachelor's degree in Computer Science from North Carolina State University, is a regular contributor to SAS Global Forum, and has coauthored several SAS certification exams.



Ryan Kumpfmiller works at Zencos Consulting as a SAS and Data Visualization Consultant where he provides consulting services to companies that need support understanding their data and how to get actionable information from it. He has been a SAS user since 2014 and has presented papers at the SouthEast SAS Users Group (SESUG), as well as at SAS Global Forum conferences. Ryan has a BS in Computer Information Systems and an MS in Competitive Intelligence Systems from Robert Morris University.

Learn more about these authors by visiting their author pages, where you can download free book excerpts, access example code and data, read the latest reviews, get updates, and more:

<http://support.sas.com/publishing/authors/aanderud.html>

<http://support.sas.com/publishing/authors/kumpfmiller.html>

<http://support.sas.com/publishing/authors/collum.html>



# Ready to take your SAS<sup>®</sup> and JMP<sup>®</sup> skills up a notch?



Be among the first to know about new books,  
special events, and exclusive discounts.

**[support.sas.com/newbooks](https://support.sas.com/newbooks)**

Share your expertise. Write a book with SAS.

**[support.sas.com/publish](https://support.sas.com/publish)**

 [sas.com/books](https://sas.com/books)  
for additional books and resources.

  
THE POWER TO KNOW.®

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. © 2017 SAS Institute Inc. All rights reserved. M1588358 US.0217