SASVIYA Exercise

Forecasting

(3/30/2020)

Sources

James Harroun, Ron Freeze

Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville

SAS® VIYA 8.2 Release V03

Copyright © 2018 *For educational uses only - adapted from sources with permission. 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 from the author/presenter.*

This tutorial walks you through features and functions of SAS Viya (Visual Analytics). As you work through this tutorial, you will create three views in a SAS Viya Report. The steps you'll take is based on a Dillards data set extracted from the University of Arkansas Enterprise System group in the Information Systems Department. The workbook is based on an investigation of the top revenue generation based on: 1) Zip codes – top 10, 2) States – geographic view and 3) Geographic view by state and zip code.

**NOTE**: Use of this tutorial requires registration with Enterprise Systems at the University of Arkansas (<https://walton.uark.edu/enterprise/>) the data used during this workshop is extracted using the file “WCOB\_Dillards Data Extraction-TUNtoExcel-KPIs.pdf” (available on the Enterprise Systems website) and should not be downloaded to your personal drives. The file should remain on the Remote Desktop S: drive provided by the University of Arkansas. This is due to our agreement with the data providers.

ESTIMATED COMPLETION TIME: 20-40 minutes

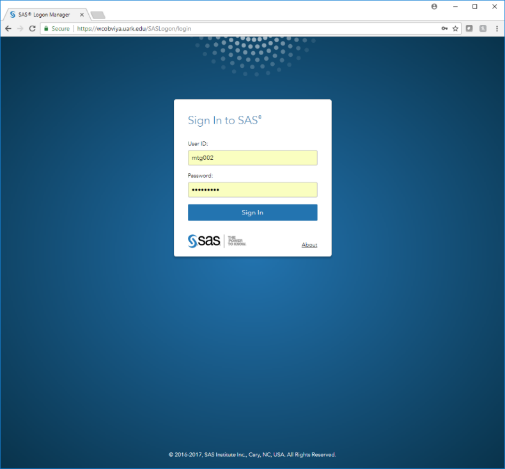
# Here's the story...

Suppose you are an employee for a large retail chain. You have just been hired as an area sales manager and are interested in understanding the sales for your area as well as other areas of the company. You are interested in visualization of the Sales value of the Top 10 Zip Codes, a visualization of the sales by states and a visualization of Sales by Zip code in the states for which you are responsible.

You'll use SAS Viya to build: 1) a column chart of the top 10 Zip Codes, 2) a geographic view by state and 3) a geographic view by Zip Code of your area (Arkansas, Kansas, Missouri and Oklahoma).

# Create Interactive Decision Support Visualizations and Analyses using SAS Visual Analytics on SAS Viya

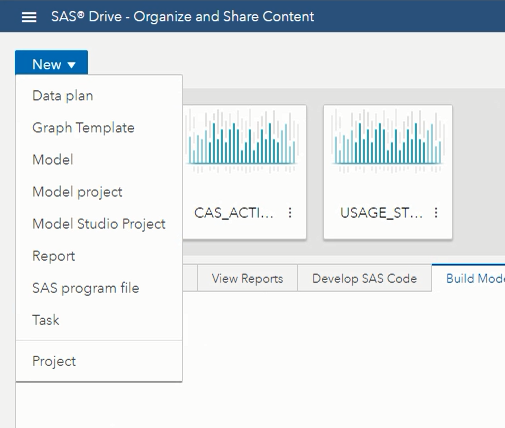
# Connect to your data

The following section is duplicated from the Pre-Conference Preparation.docx document for connecting to your data. The end result is a report file that has the data already identified for this tutorial. If this has been accomplished, please skip to [Top 10 Zip Codes](#_Top_10_Zip). The file you will need is from the extraction file mentioned above and should have been named **WCOB\_DillardsTUN-Export-KPIs.xlsx**.

## Verify SAS Viya Access

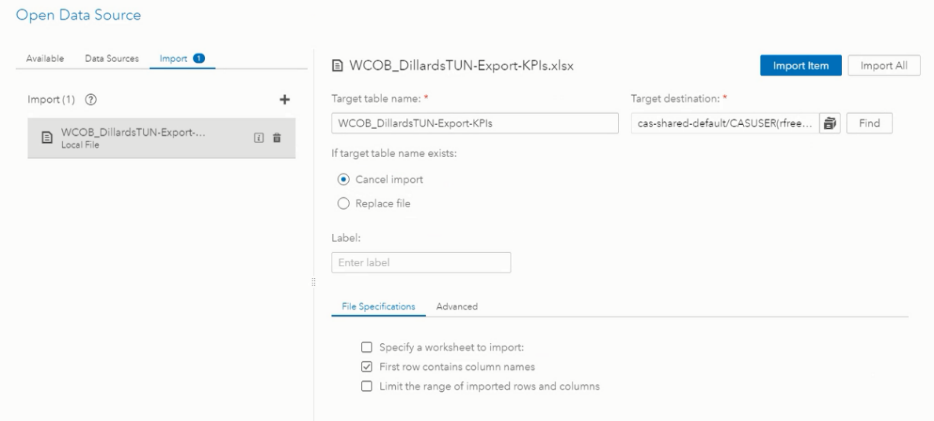
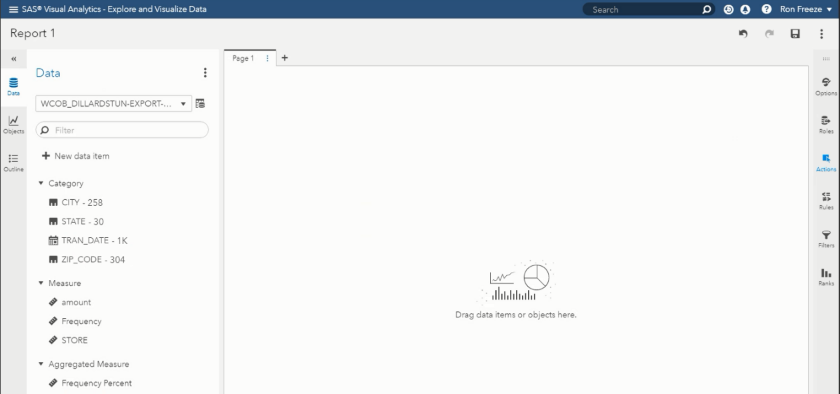
SAS VIYA is accessible from any web browser, but Google Chrome is recommended. For ease, open Google Chrome in the VMware virtual desktop available on the Walton College Enterprise Systems.

1. Use the link to get to SAS Studio: <https://viya.walton.uark.edu>
2. Use the university provided credentials to log on



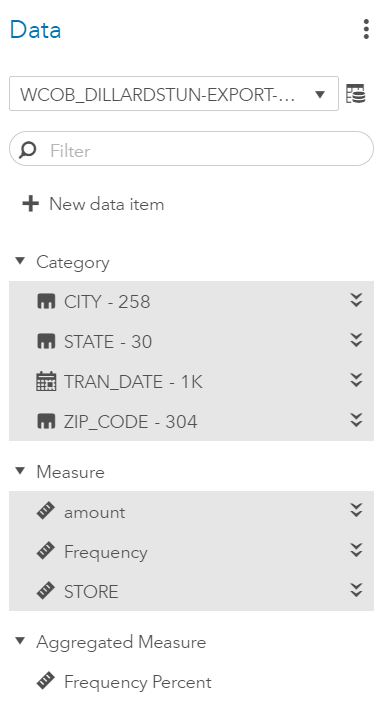
1. Select New -> Report in the upper left corner of the screen



1. Select Data
2. Select “Import”
3. Navigate to your Local File or through the Documents Directory to locate **WCOB\_DillardsTUN-Export-KPIs.xlsx** file
4. Click on your file and then select your **Target destination** (this should be your CASUSER directory)
5. Select **Import Item** and then **OK**
6. You now see the Category – Measure – Aggregated Measure available in the left navigation pane from your Dillard’s data source
7. To save your Report – Select **Save** in the upper right hand corner
8. The default should be your CAS folder labeled **My Folder**
9. Name the file (WCOB\_DillardsTUN-KPI-VA14) and select Save

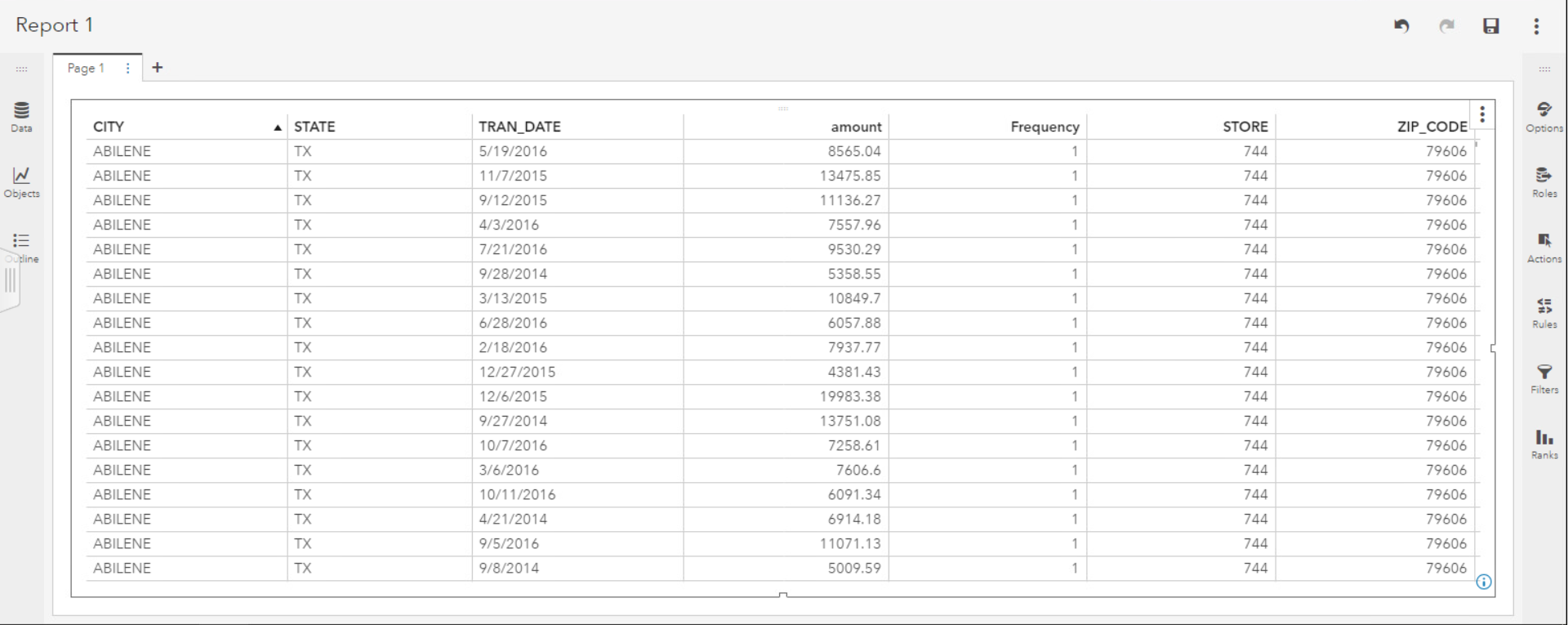
The SAS VIYA project can now be opened at any time and reviewed or updated

# Controls & Histogram

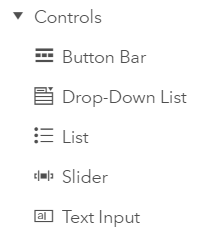
The data elements in the Dillards data set are separated into groups based on measurement level and intended use. On the left you will see the Dillards data you uploaded has been grouped into two basic divisions: Category and Measure. Take a look at the information provided next to the Category items; this number indicates the number of distinct class levels. Let’s begin by building an automatic visualization using all the columns in the data.

1. Left-click CITY
2. Left-click all the Category and Measure items
3. Drag all the data items onto the blank canvas on the right.

The auto-charting function of SAS Visual Analytics on SAS Viya has determined a table object is the best way to represent the data you selected. The table you just created shows a rectangular representation of source data. Let’s turn this into an interactive decision support tool that changes based on user input.



## User Inputs



1. Click the **objects** icon on the left pane and
2. Scroll down until you find the controls section.
3. Drag a **Slider** to the upper left of the report below **the report name.**

Placing the slider in this position makes it a report prompt that filters all report contents.

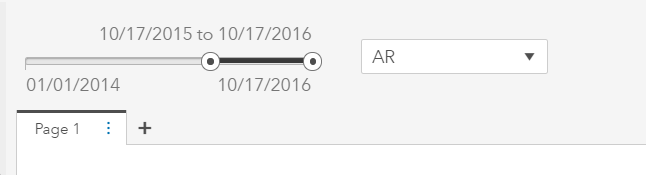
1. Drag a Drop-Down List to the right of the slider

You have created another report prompt that controls displayed content. The report prompts indicate they are ready to accept input data elements.

Let’s assign **TRAN\_DATE** to the slider and **STATE** to the drop-down list.

1. Click **Data** to redisplay the available data elements.
2. Drag and drop **TRAN\_DATE** on the slider
3. Drag and drop / **STATE** on the drop-down list

You have successfully created two interactive filters that allow all report content to be displayed following the selected data at the top of the report. Drag the arrows on the slider and click on the drop-down arrow in the drop-down box to select a date range and a state of your choice.

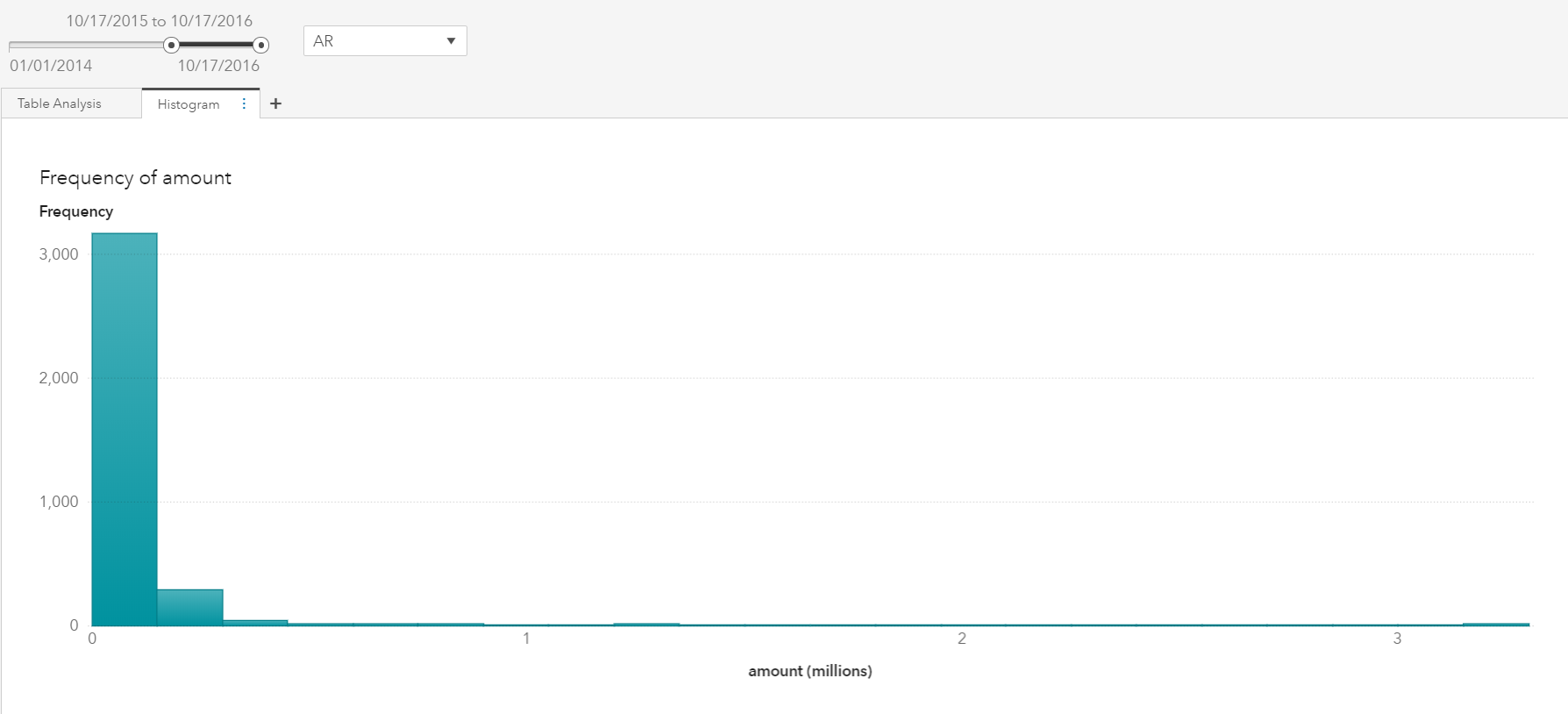
1. Select the date range of **10/17/2015-10/17/2016** for the slider
2. Select the state **AR** for the drop-down

Notice that all content in the table below has been subset to the selected date range and state. If you would like to create a custom date endpoint, click the slider endpoints to choose a date from a calendar pop-up. Control-click columns of the table to sort the columns to your liking. Left-click the floppy diskette icon in the top right to save your report.

1. Double click **Page 1** and rename the page **Table Analysis**

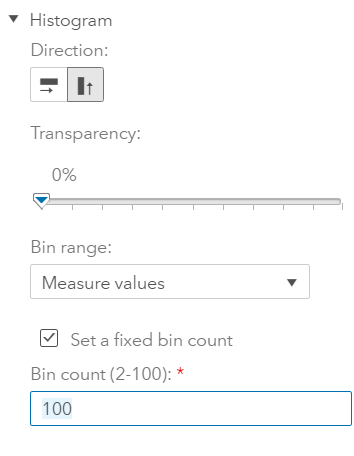
## Histogram

Now let’s create a Histogram of our selected data.

1. Click the **+** to the right of **Table Analysis** to add a new page to the report.
2. Double-click **Page 2** and rename the page **Histogram.**
3. Click **Data** to display the available data elements.
4. Drag and drop **amount** to the canvas.

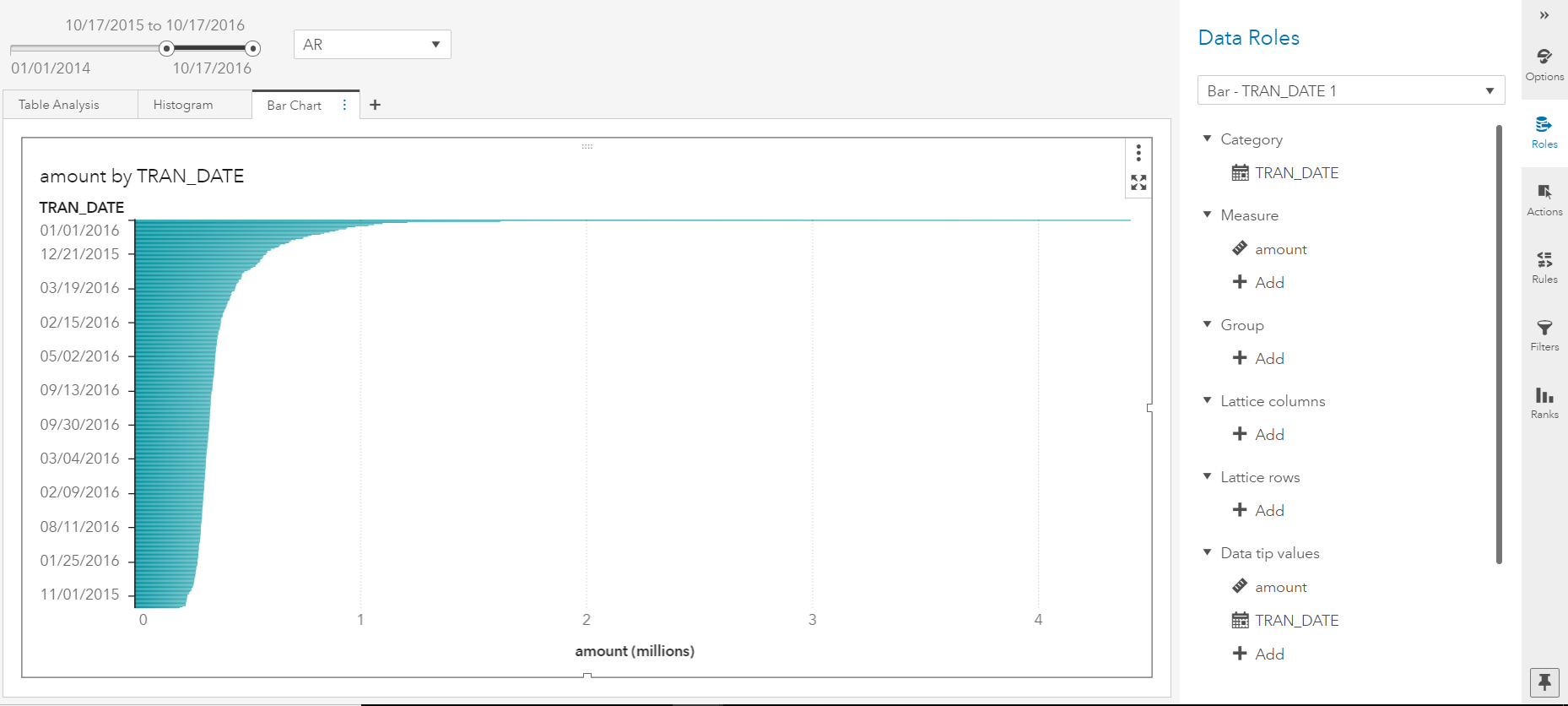
Because **amount** is a measure, SAS Visual Analytics automatically represents the data as a histogram. Because the report is a subset of the values selected at the top of the report, your results might not match the screen capture. Change your slider and drop-down selections and you will see the histogram values change as well. The histogram automatically bins the **Amount** values based on the range of the data. Hover over the bins to view the endpoints of the most frequent purchase amounts.

It would be helpful to view the **amount** values in the histogram with a greater number of bins.

1. Click your histogram to surround it with a dark line and ensure it is the selected visualization
2. On the right pane, click the **Options** icon
3. Scroll down to the **Histogram** section
4. From the **Bin Range** drop-down select **Measure values**
5. Place a check mark in the box to the left of **Set a fixed bin count**
6. Enter **100** into the **Bin count (2-100):** field
7. Press **Enter**
8. Notice the histogram has changed and it is easier to interpret the values at the lower end of the scale.

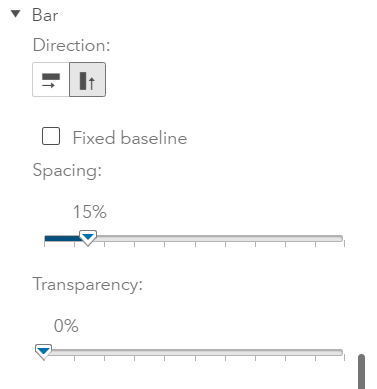
## Bar Chart

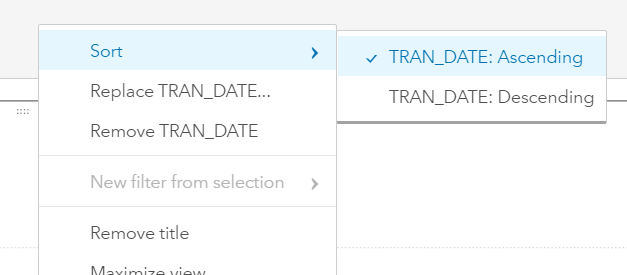
The histogram gives a good understanding of the frequency of amounts reported based on the date range and state selected to filter the report. It would also be useful to understand the **amount**, but it might be a good idea to view total sales by transaction date, so let’s create a bar chart that views the sum of **amount** across time.

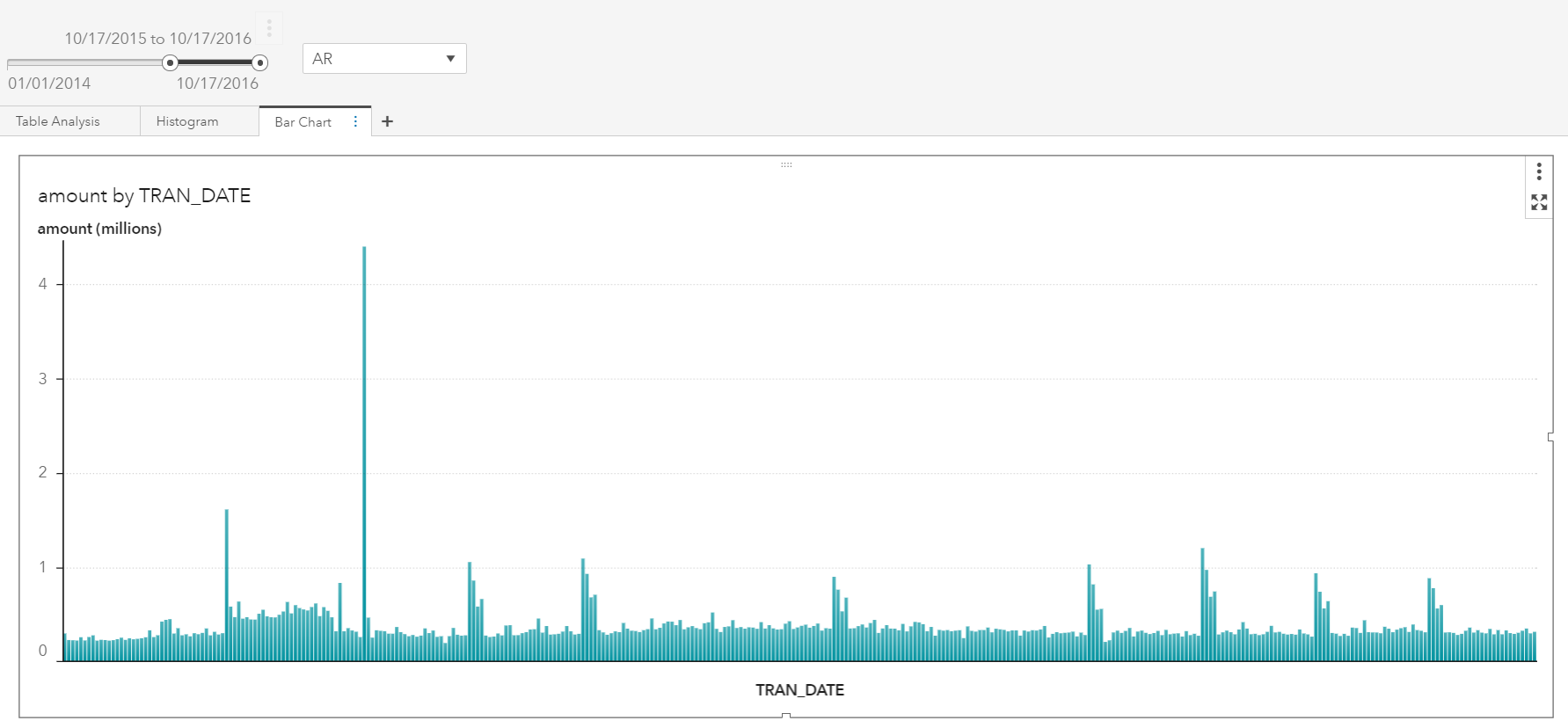
1. Move the right pane out of the way by clicking the right facing arrows  to minimize it.
2. Click the **+** to the right of **Histogram** to create a new page.
3. Double-click **Page 3** and rename the page **Bar Chart.**
4. Left-click the **Objects** icon on the left pane and scroll down to find **Bar Chart**.
5. Drag and drop the **Bar Chart** onto the canvas.
6. On the right pane, left-click **Roles**,
7. Left-click **+ Add** below **Measure,** and select **amount**
8. Left-click **+ Add** below **Category** and select **TRAN\_DATE**

Notice the bar chart is horizontal and the bars are sorted by descending order of frequency. We need to change this to a vertical bar and place the bars in date order.

1. Click the **Options** icon on the right pane

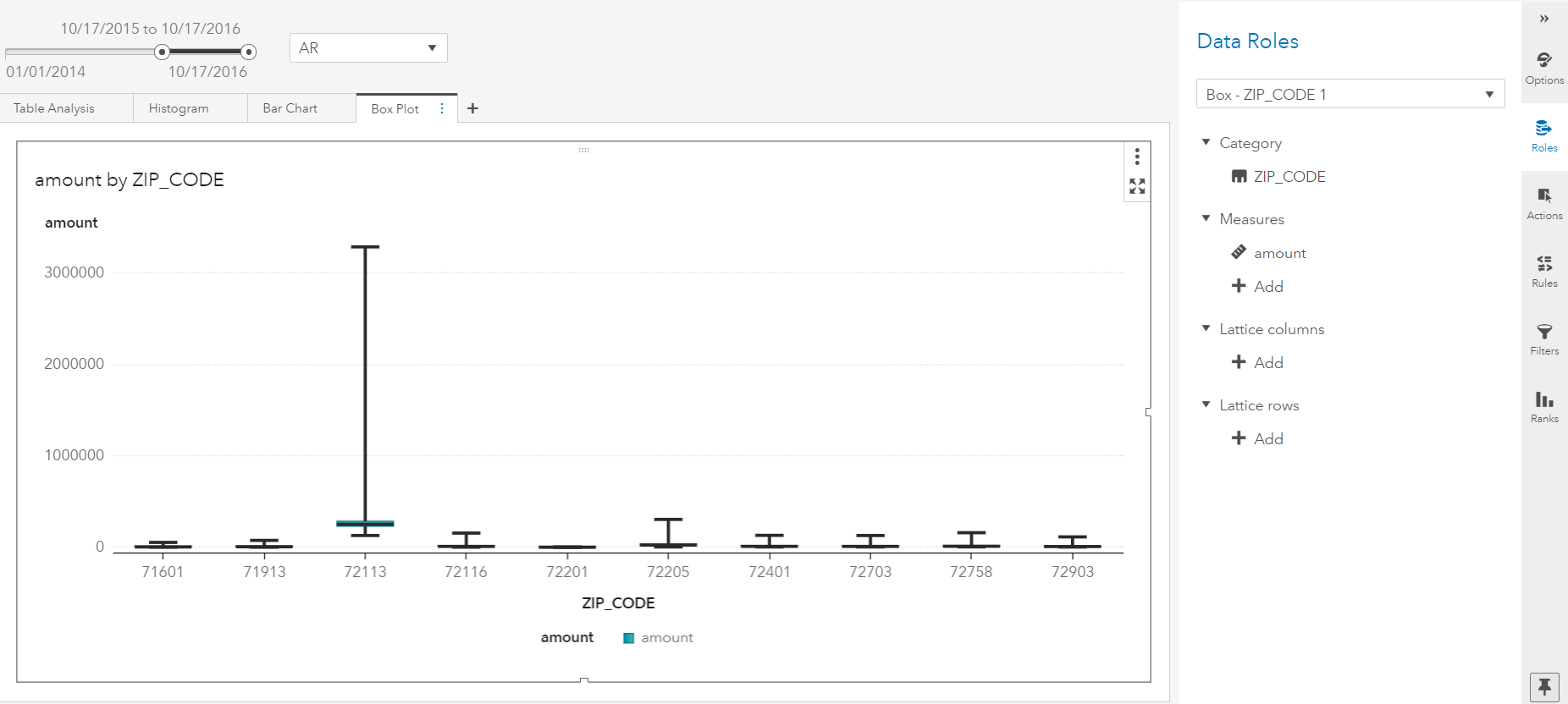


1. Scroll down to the section titled **Bar** and choose the vertical icon
2. To sort the bar chart by **TRANS\_DATE**, right-click **TRANS\_DATE** at the bottom of the bar chart
3. Select **Sort > TRAN\_DATE: Ascending**

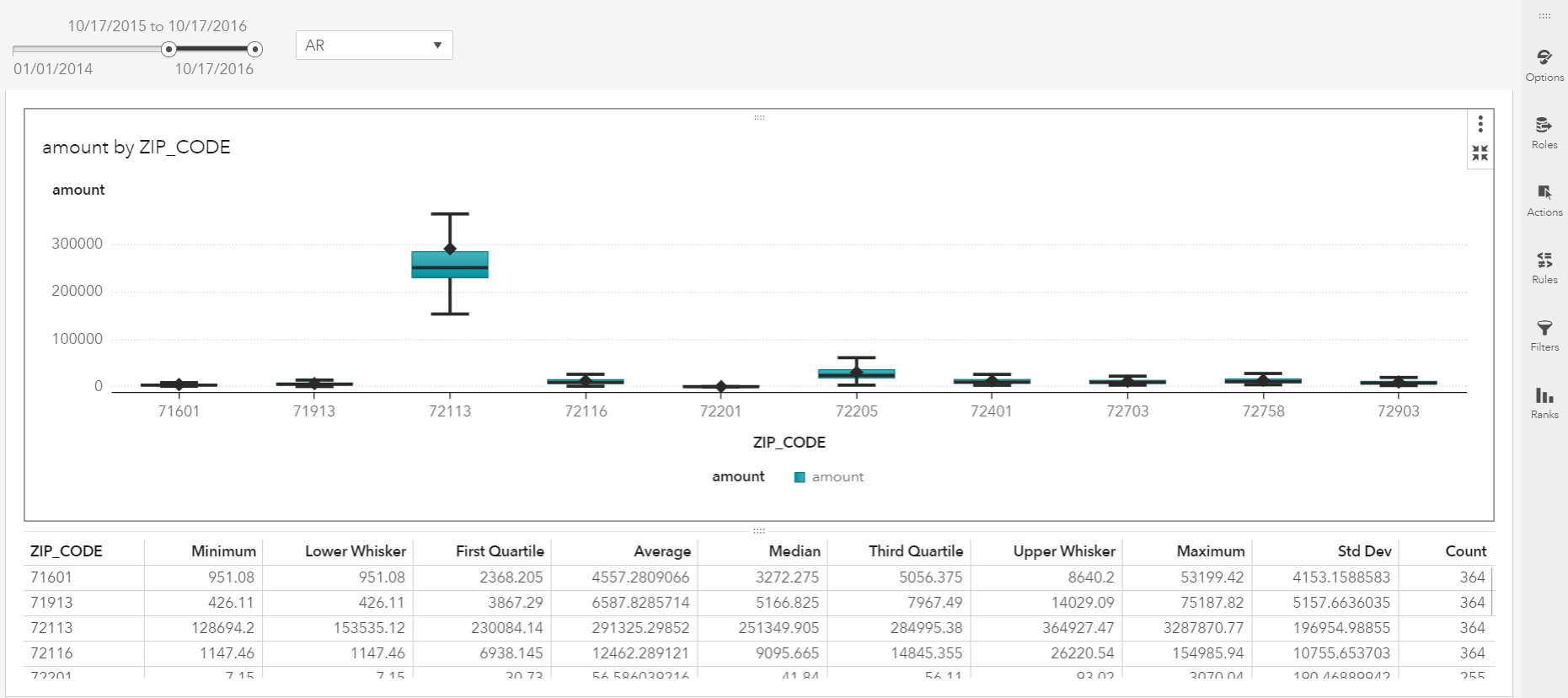
The bar chart will now be arranged by **TRANS\_DATE** on the horizontal axis. Move the right pane out of the way by clicking the right facing arrows . Hover over the peaks in the bar chart to see the dates with the highest total sales. Do you notice anything interesting about these dates?

## Box Plot

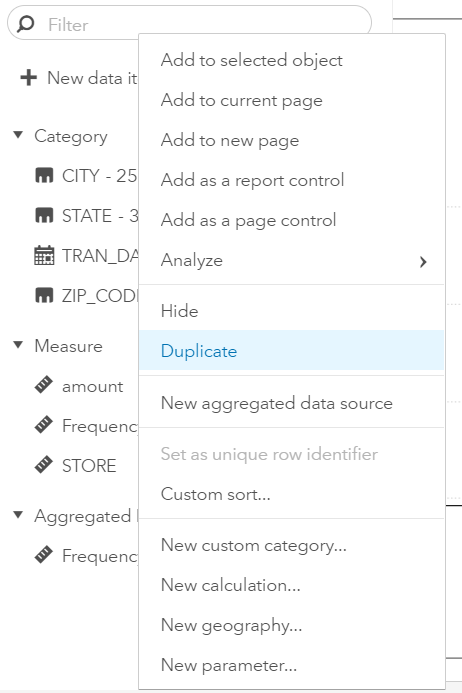
The bar chart gives us a good understanding of how total sales have cycled over time. Let’s create summary statistics to understand averages and quartiles.

1. Click the **+** to the right of **Bar Chart** to create a new page
2. Rename the page **Box Plot**
3. Left-click the **Objects** icon on the left pane and scroll down to find **Box Plot**
4. Drag and drop the **Box Plot** onto the canvas.
5. On the right pane, left-click **Roles**
6. Left-click **+ Add** below **Measure,** and select **amount**
7. Left-click **+ Add** below **Category** and select **ZIP\_CODE.**

Notice the box plot is difficult to understand because outliers are included and should be filtered out.

1. Click the **Options** icon on the right pane
2. Scroll to the **Box Plot** section and select **Ignore Outliers** from the **Outliers:** drop-down list
3. Place a check mark in the box to the right of **Averages**
4. Left-click the four arrows  in the upper right of the box plot to maximize the box plot and display the summary statistics.
5. Click the four arrows  to restore the box plot visualization.

## GeoMap

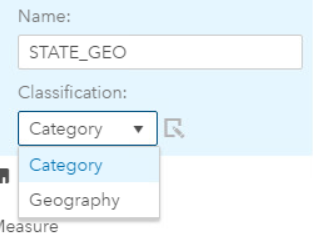


It would be helpful to view the geographic information in map form. While the box plot is convenient, there might be a better way to visually represent data such as **STATE** and **ZIP\_CODE** using geographic visualizations. Let’s duplicate these data elements and specify they are to be used in geographic contexts.

1. Click the **+** to the right of **Box Plot** to create a new page
2. Double-click **Page 5** and rename the page **GeoMap**
3. Click **Data** to display the available data elements
4. Right-click the **STATE** variable
5. Left-click **Duplicate**

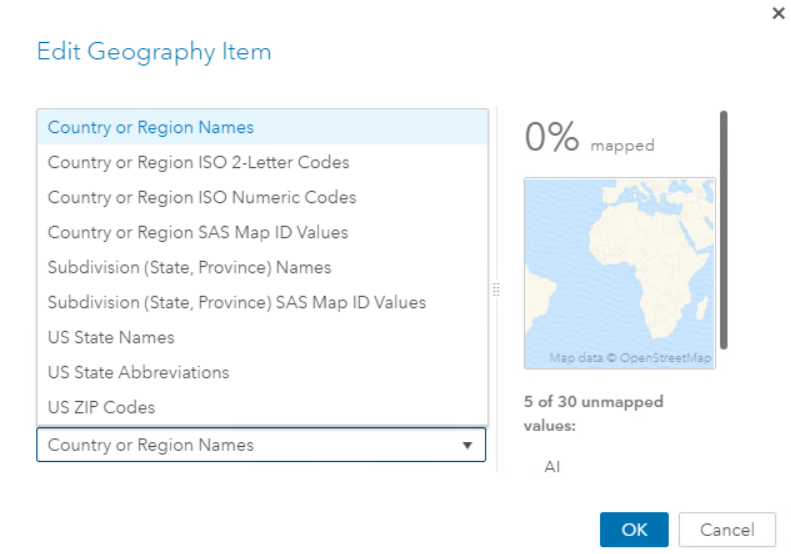
You have just created a duplicate data element named **STATE (1)** in the report. Let’s modify this data element to suit our needs by modifying the metadata properties of the data element.

1. Left-click the two vertical arrows  to the right of **STATE (1)**
2. In the **Name:** field enter **STATE\_GEO**

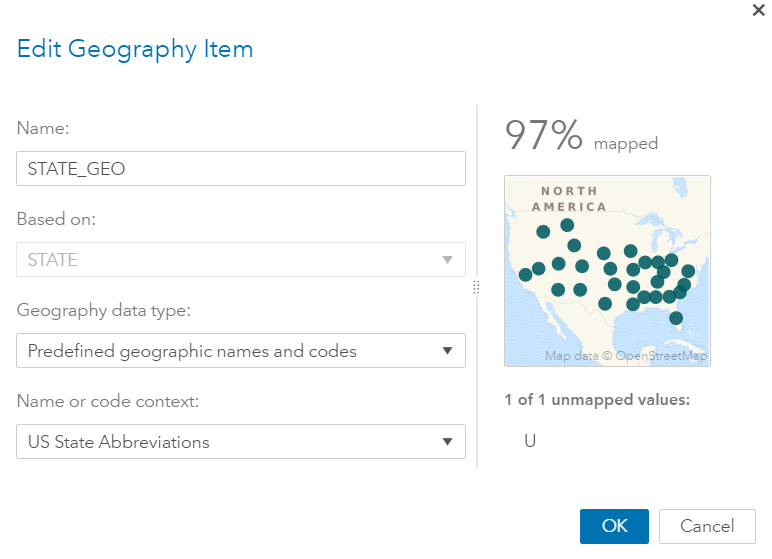


1. Left-click the drop-down arrow below **Classification:** and select **Geography**

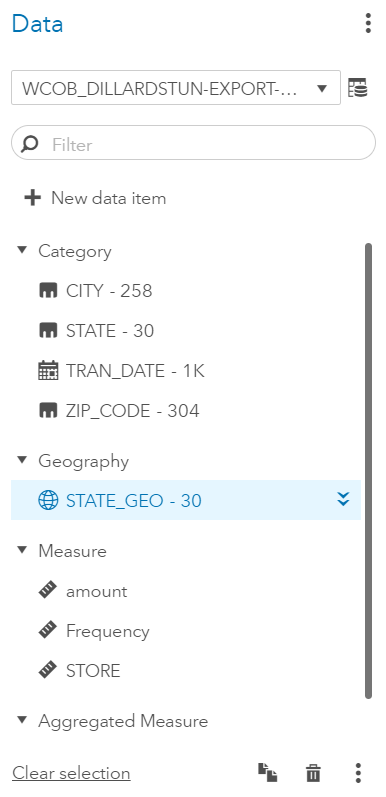
An **Edit Geography Item** dialog appears to facilitate quick and easy geographic data assignment. There are several pre-loaded geographic metadata assignments that allow you to quickly represent data elements as geographic features. SAS Visual Analytics contains pre-loaded state names and abbreviations. From the table object on the canvas, we see state abbreviations are contained in the base data.



1. In the **Edit Geography Item** dialog box, click the triangle in the **Name or code context:** drop-down
2. Select **US State Abbreviations**.

The **Edit Geography Item** dialog displays a preview of the new metadata representation. In this examplle, the data items have mapped 97% successfully and there is one problematic **STATE** value **U**. This information can be sent later to our data stewards for validation and clean-up.

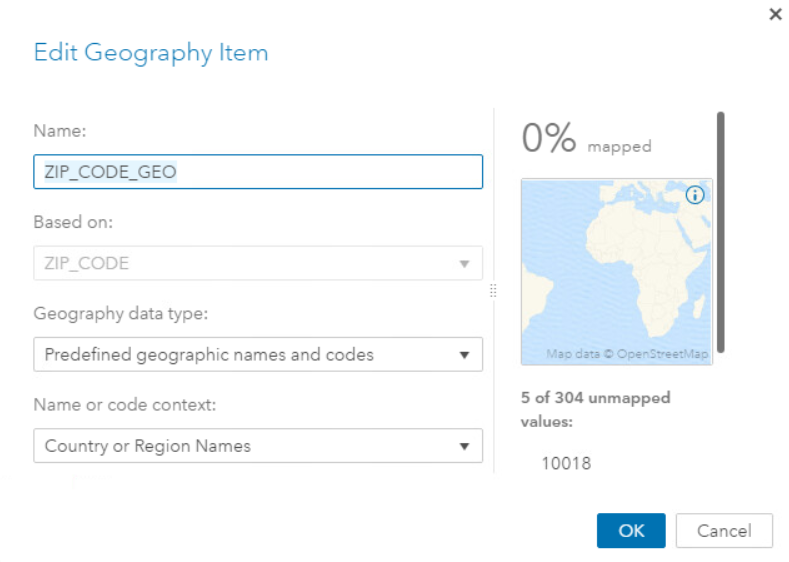
1. Left-click **OK** to close the **Edit Geography Item** dialog.

On the left, under the **Data** pane, you will see you have a new section available: **Geography:**. You will see your new data element **STATE\_GEO**.Let’s prepare the **ZIP\_CODE** data item in a similar manner.

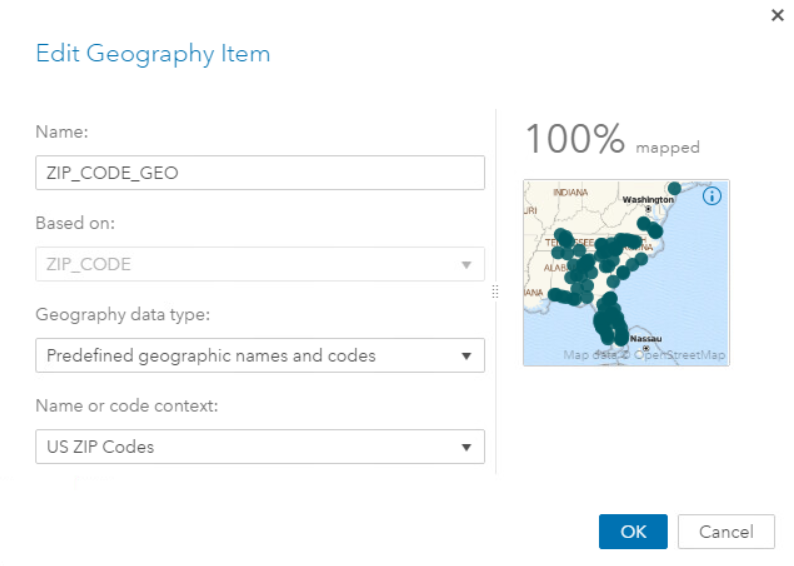
1. Click **Data** to display the available data elements
2. Right-click the two vertical arrows  to the right of **ZIP\_CODE** and
3. Left-click **Duplicate**.

You have just created a duplicate data element named **ZIP\_CODE (1)** in the report. Let’s modify this data element to suit our needs.

1. Left-click the two vertical arrows to the right of **ZIP\_CODE (1)**
2. In the **Name:** field enter **ZIP\_CODE\_GEO**
3. Left-click the drop-down arrow below **Classification:** and select **Geography**.

The **Edit Geography Item** dialog box appears to facilitate quick and easy geographic data assignment. There are several pre-loaded geographic metadata assignments that allow you to quickly represent data elements as geographic features.

SAS Visual Analytics contains pre-loaded state names and abbreviations. From the table object on the canvas, we see state abbreviations are contained in the base data. In the **Edit Geography Item** dialog box find the **Name or code context:** drop-down.



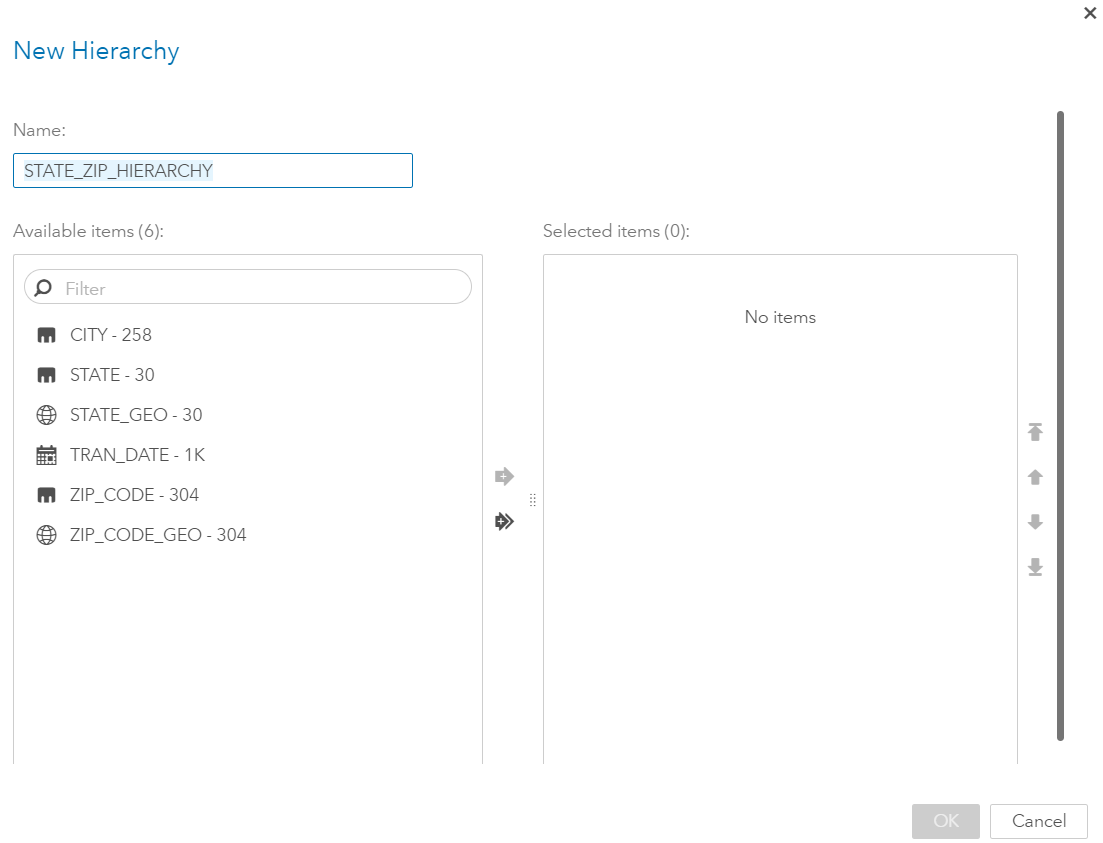
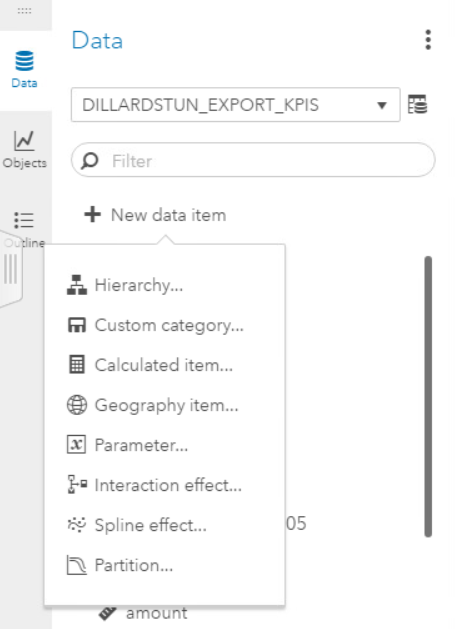
1. Click the triangle in the drop-down and select **US ZIP Codes**.

The **Edit Geography Item** dialog displays a preview of the new metadata representation. You see the data items have mapped 100% successfully.

1. Left-click **OK** to close the **Edit Geography Item** dialog.

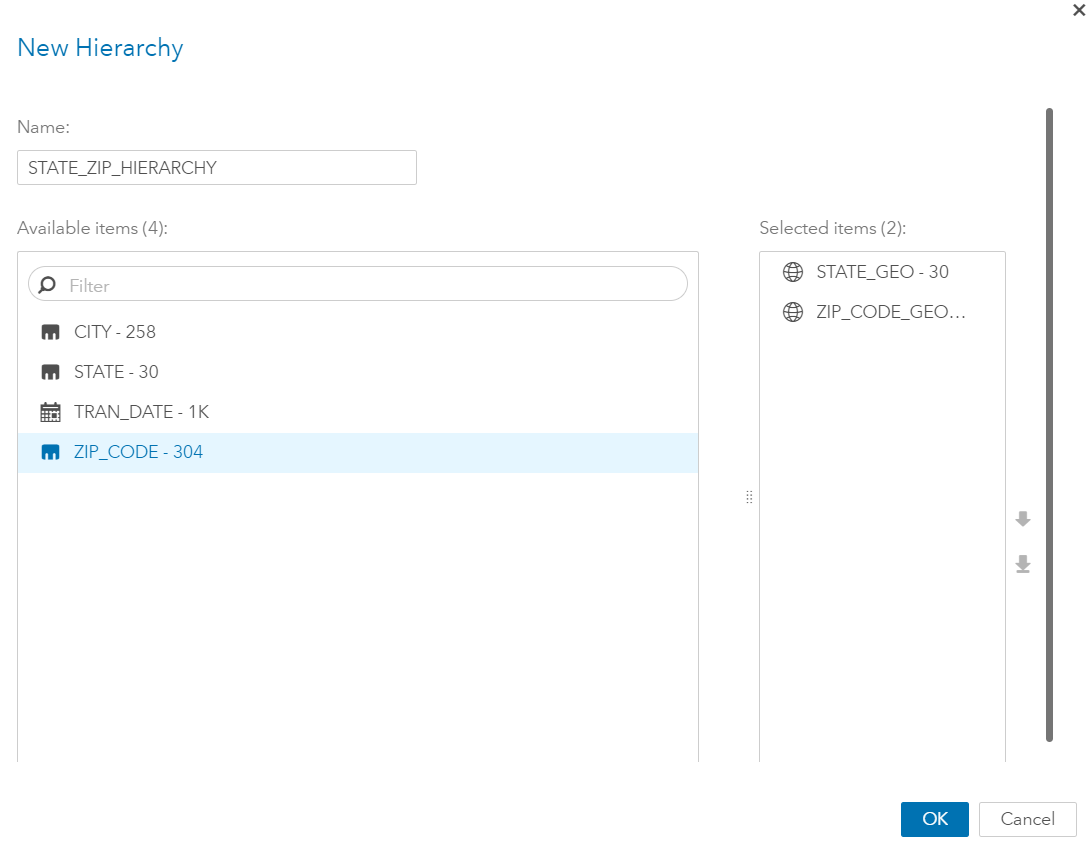
On the left, under the **Data** pane You will see your new data element **ZIP\_CODE\_GEO** listed under the **Geography** section. There are now two levels of granularity of geographic data within this report: **STATE\_GEO** and **ZIP\_CODE\_GEO**. Let’s create an interactive hierarchy that combines both geographic elements.

1. Click **Data** to display the available data elements



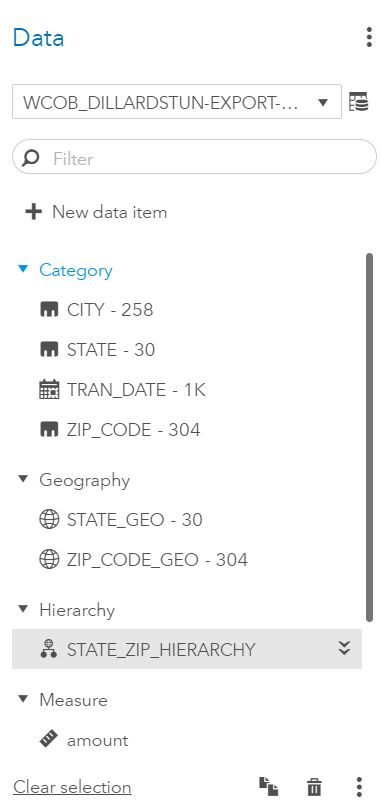
1. Left-click **New** data item and left-click **Hierarchy**…
2. A **New Hierarchy** dialog will open. Enter **STATE\_ZIP\_HIERARCHY** into the **Name:** field.

We need to add the elements to the hierarchy in the correct order, so they are correctly nested together.



1.  Left-click **STATE\_GEO** and left-click the plus arrow in the center of the dialog.
2. Left-click **ZIP\_CODE\_GEO** and left- click the plus arrow in the center of the dialog.

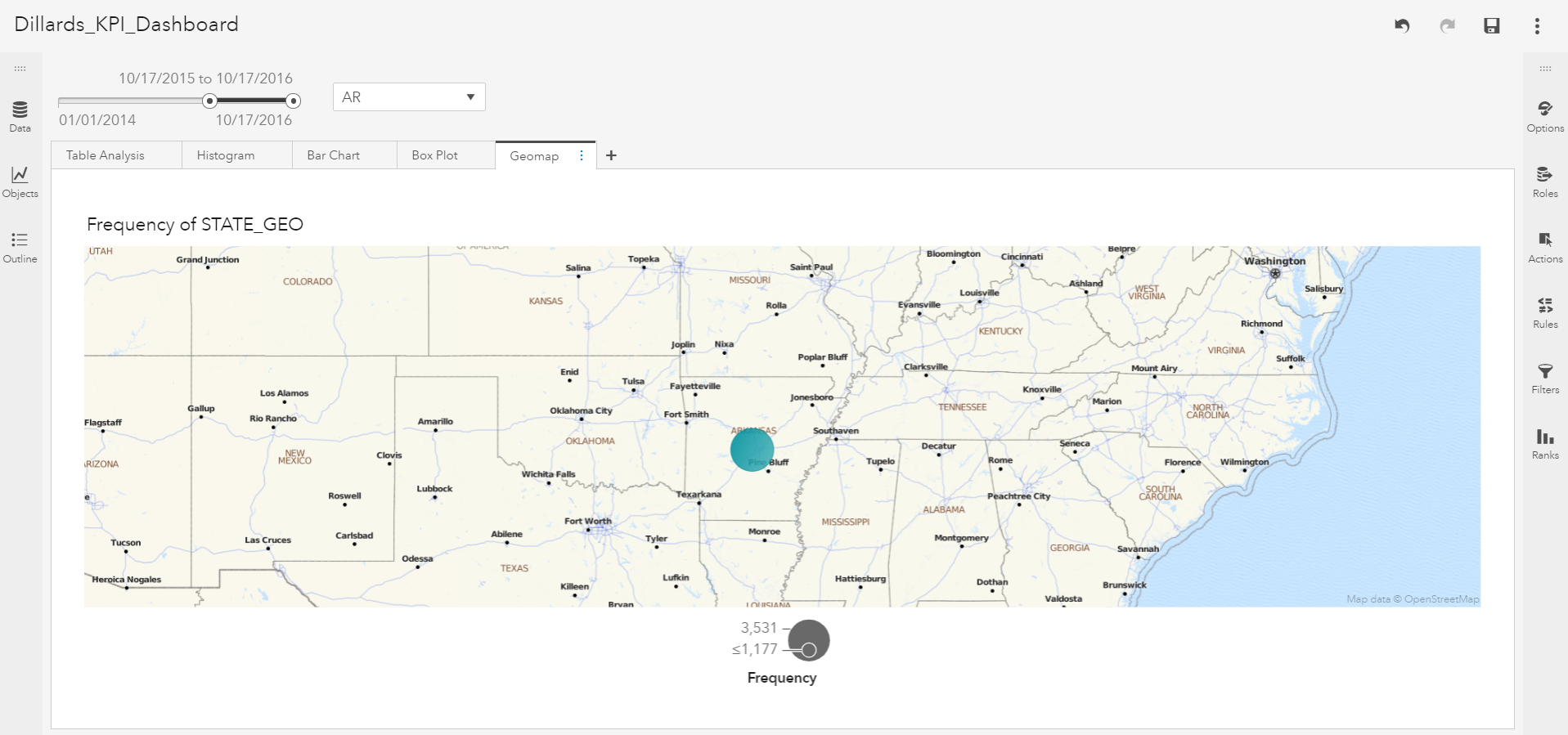
Your hierarchy should consist of **STATE\_GEO** at the top of the hierarchy and **ZIP\_CODE\_GEO** at the bottom of the hierarchy.

1. Left-click **OK** to close the **New Hierarchy** dialog.

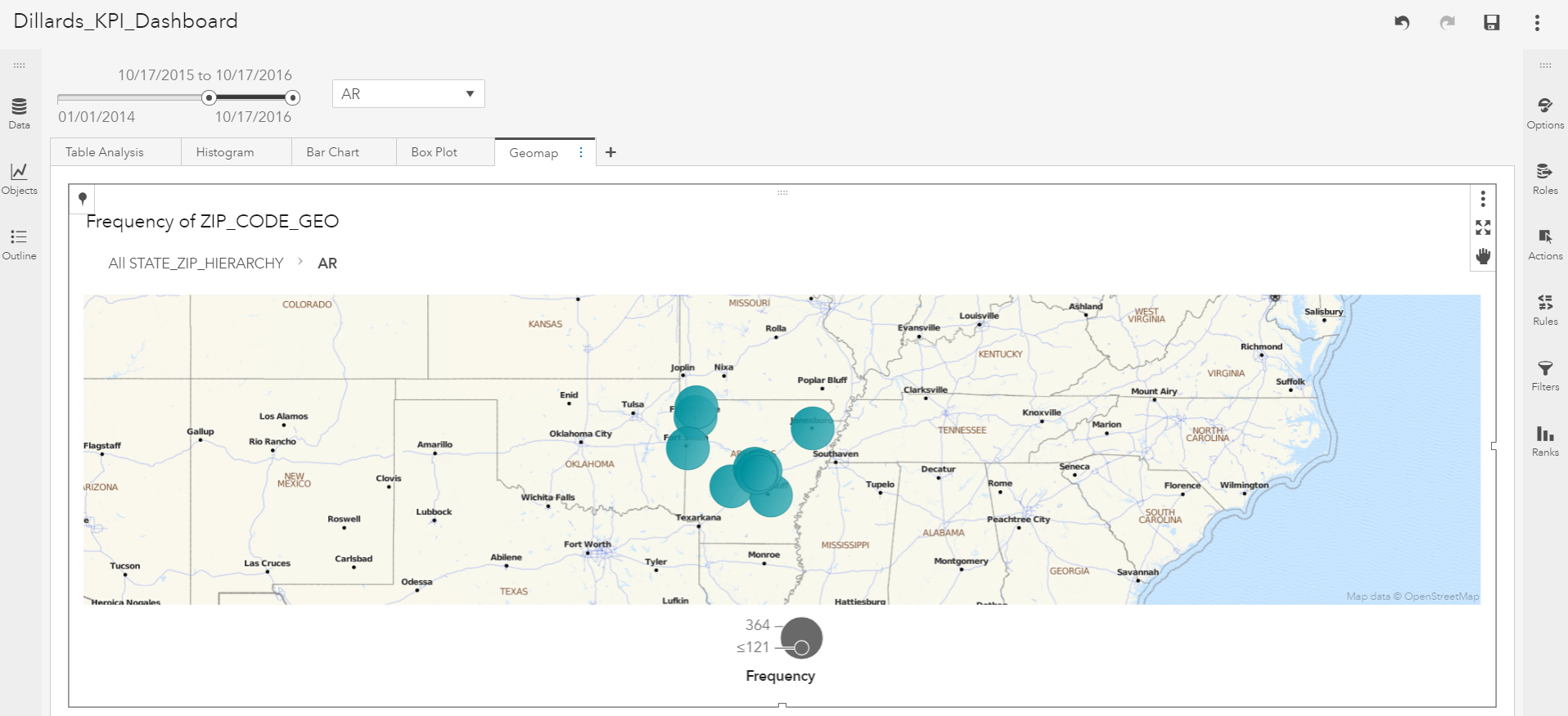
On the left, under the **Data** pane, you will see you have a new section available: **Hierarchy:** You will see your new data element **STATE\_ZIP\_HIERARCHY**.

Create a new visualization from the hierarchy just created.

1. Drag the **STATE\_ZIP\_HIERARCHY** onto the **Geomap** page canvas until **Auto-Chart** is shown.

Because the hierarchy was created from geography variables, the hierarchy is rendered on a geomap. The visualization is displayed at the **STATE\_GEO** level. Remember: the visualization is limited by the report filters at the upper left. You may need to scroll out to zoom to a wider perspective.

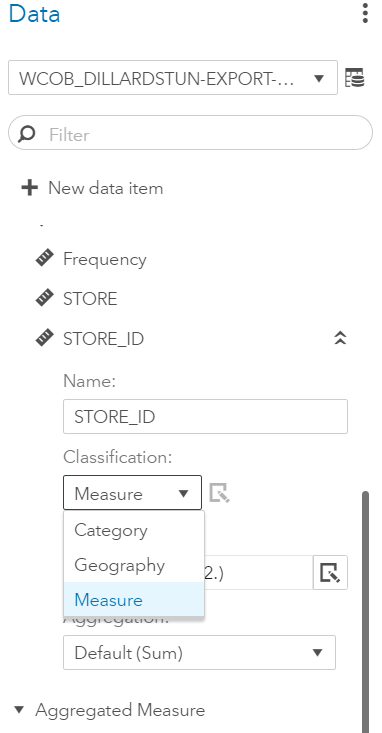
The **STATE\_ZIP\_HIERARCHY** is fully interactive and allows you to click through the hierarchy to different levels of granularity.

1. Hover over Arkansas and double click with your left mouse button. The visualization is immediately refreshed and rendered at the **ZIP\_CODE\_GEO** level.

The visualization is showing frequency of occurrence for each of the values within each level of the hierarchy. Let’s add a data element to change the meaning of what is being measured within the visualization.

1. Click the **Roles** icon on the right pane
2. Below **Size**, hover to the right of **Frequency** until a set of stacked dots appears
3. Click the stacked dots and a **Replace Data Item** dialog will appear
4. Left-click **amount**

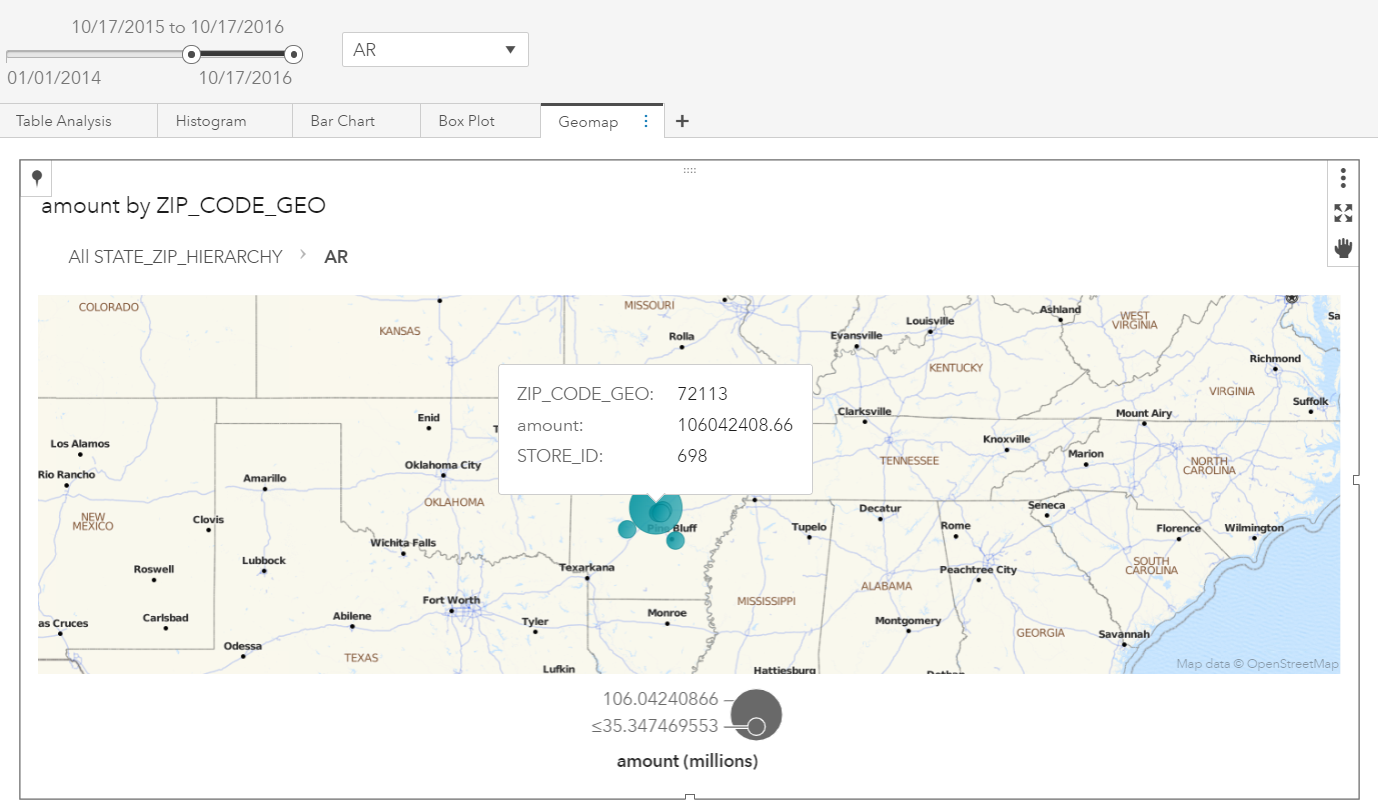
The geomap represents data by state and zip code. Let’s create the store number as an identifier that will help when hovering over the geomap. If your data pane is not displayed on the left,



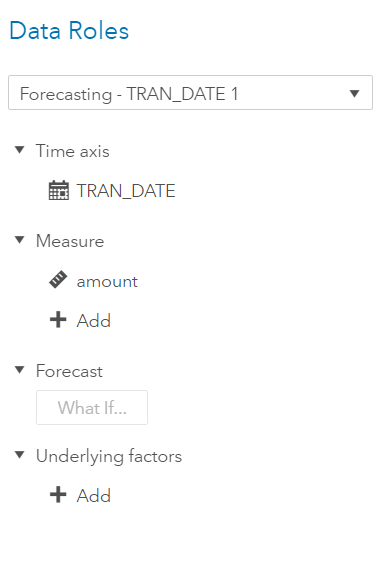
1. Click **Data** to display the available data elements
2. Scroll down to the **Measures** section to find **STORE**
3. Right-click the two vertical arrows  to the right of **STORE** and left-click **Duplicate**
4. Left-click the two vertical arrows  to the right of **STORE (1)**
5. In the box labeled **Name:** Enter **STORE\_ID**
6. Click the drop-down next to **Classification** and select **Category.**
7. Click your geomap to surround it with a dark line and ensure it is the selected visualization
8. Click the **Roles** icon on the right side of the visualization and you will see a listing of the data elements and their roles as they have been assigned to this visualization.
9. If necessary, scroll down to find the **Data tip values** section.
10. Left-click **+ Add**.

An **Add Data Items** dialog will appear and a list of eligible data elements from the data set will be listed.

1. Place a check mark next to **STORE\_ID**.
2. Click **OK**.
3. Move the right pane out of the way by clicking the right facing arrows .

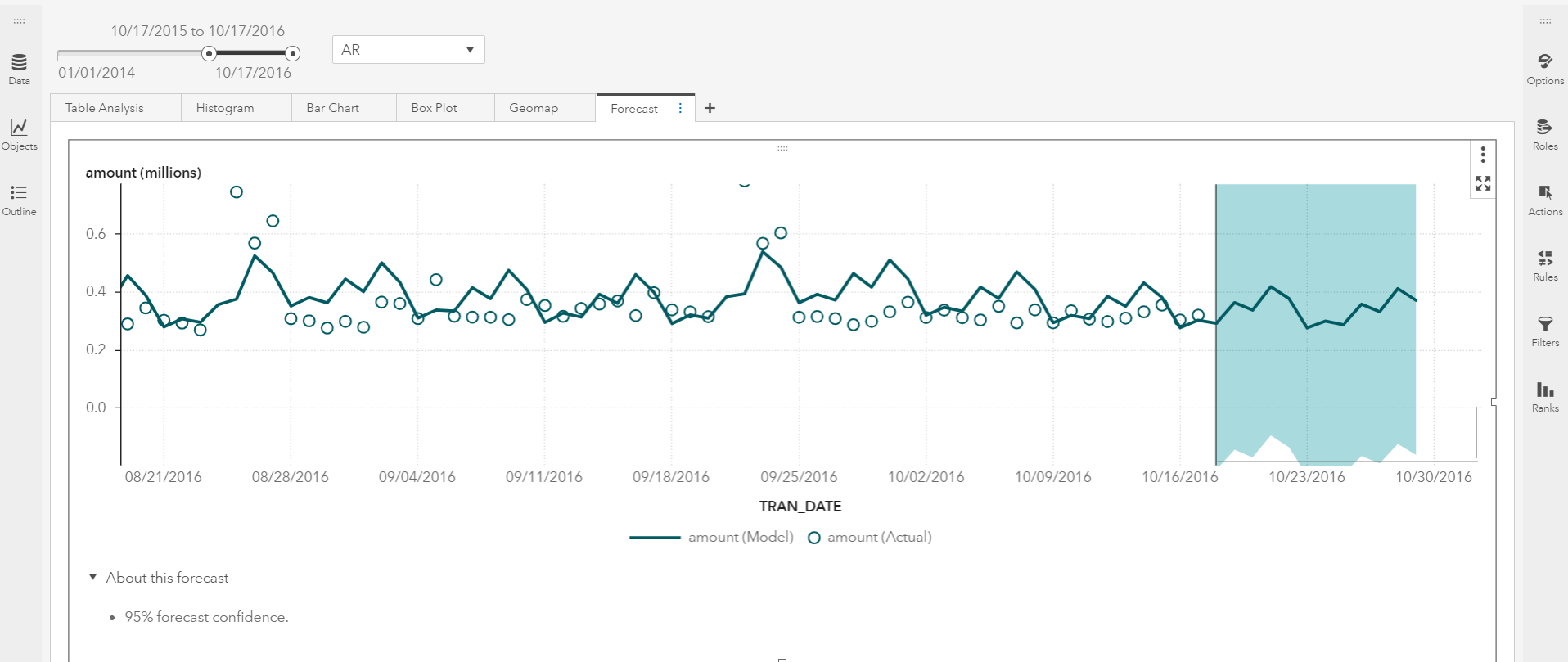
Click or hover over the large bubble in the center of Arkansas. You will now see this large bubble refers to store 698. Let’s continue to add value to this report as a decision support tool by building a forecast that allows us to project future sales**.**

## Build a Forecast

1. Click the **+** to the right of **Geomap** to create a new page
2. Double-click **Page 6** and rename the page **Forecast**
3. Left-click the **Objects** icon on the left pane
4. Scroll down to find **Forecasting**
5. Left-click **Forecasting** and drag it onto the canvas
6. On the right pane, left-click **Roles**
7. Left-click **+ Add** below **Measure,** and select **amount**
8. Left-click **+ Add** below **Time Axis** and select **TRAN\_DATE**

The visualization will create a forecast model with a default of 6 forecast periods. Since our data is represented at the daily level, this means the forecast is created for the six days following the maximum date specified in the slider. We need to extend the value to 30 days following the maximum date.

1. Click the **Object** icon in the right pane
2. Scrolling down to **Forecast**, and enter **30** into the **Forecast horizon** box
3. Press enter on your keyboard and the forecast will update

A forecast will be created for the entire state selected by the drop-down filter at the top of the report. Scroll to zoom into the forecast. Let’s add a filter to this individual visualization to subset the forecast to the individual store level.

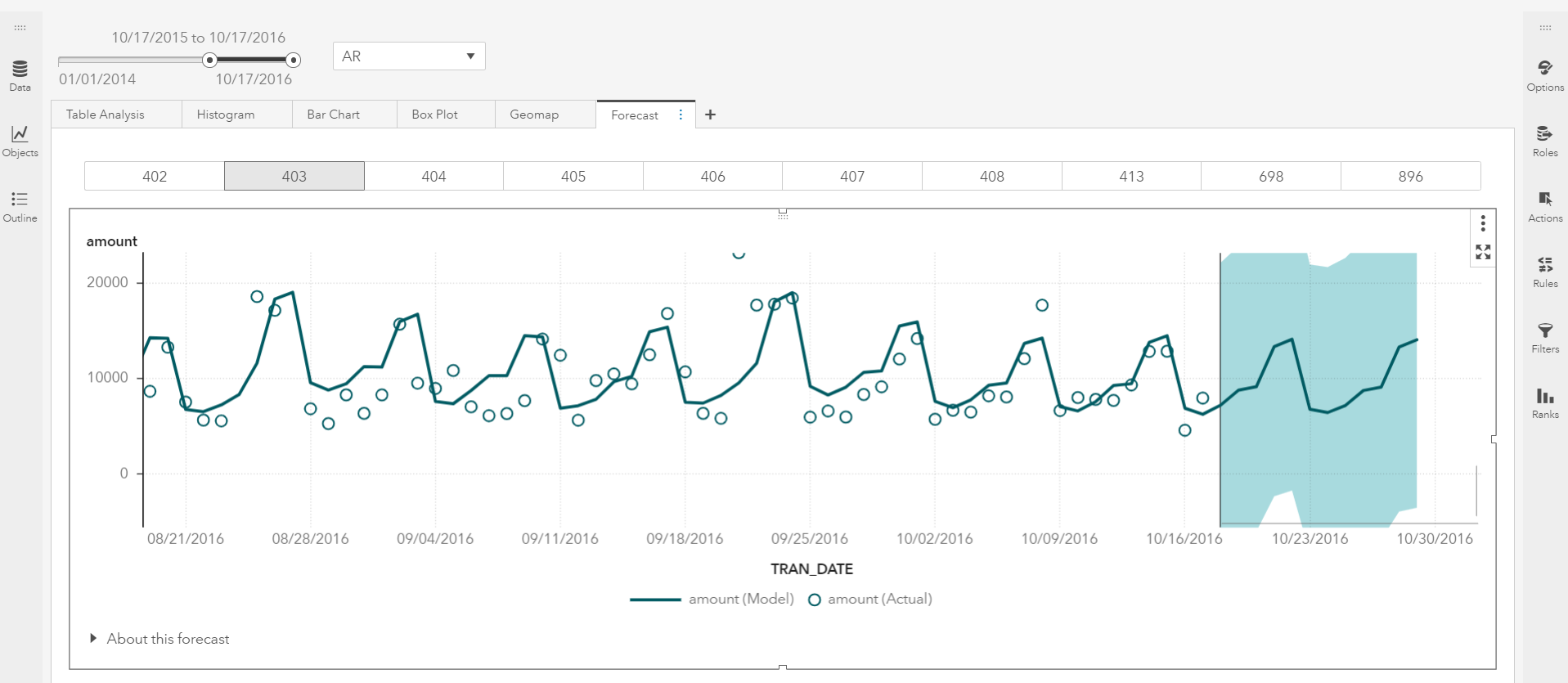
1. Left-click the **Objects** icon on the left pane
2. Scroll down to find **Button Bar**
3. Left-click **Button Bar** and drag it to the center top of the forecast object until a small blue area spans the forecast visualization and **+ Button Bar** is displayed

A button bar has been created and it needs to be assigned a data element.

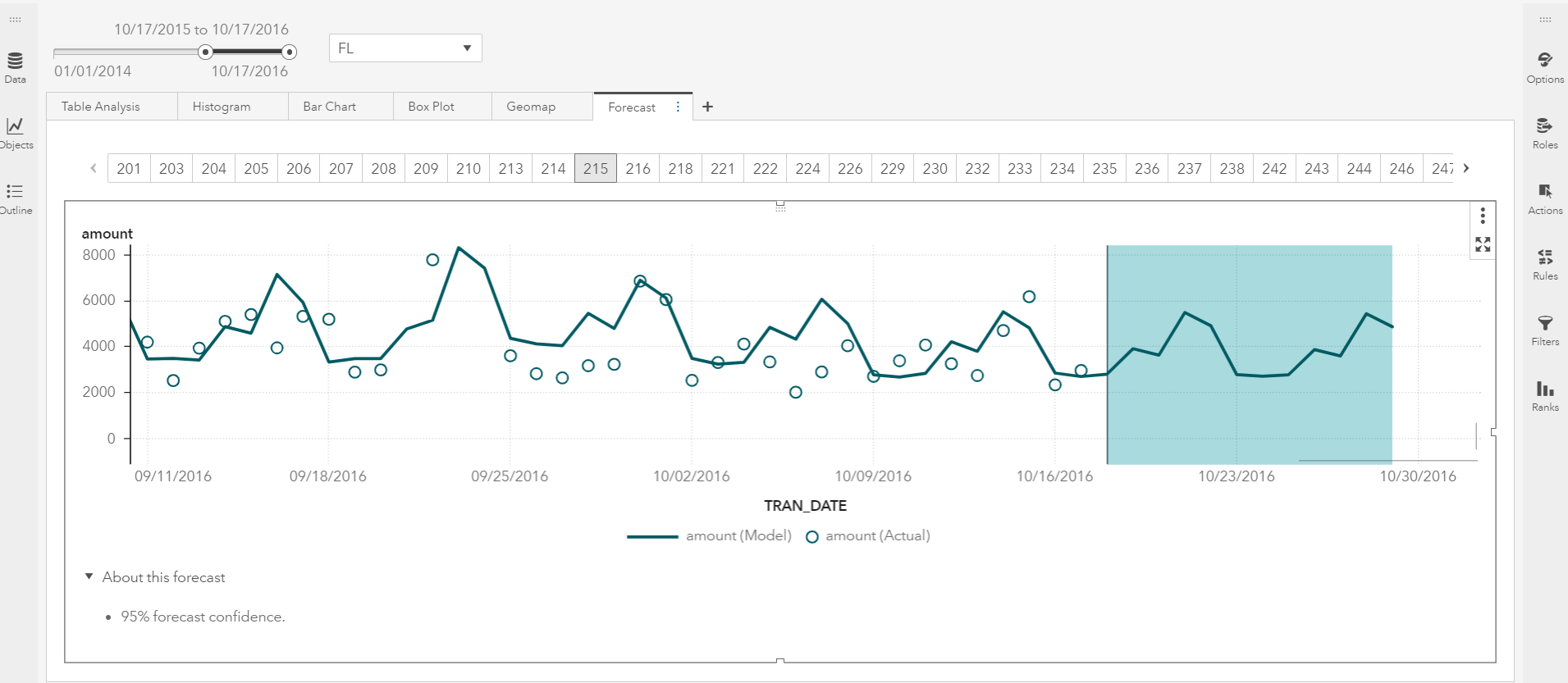
1. Ensure the button bar is selected and click the **Roles** icon in the right pane
2. Click **+ Add** below **Category** and select **STORE\_ID**

The button bar has been added and we need to create an interaction with the forecast visualization on the page.

1. On the right pane, left-click the **Actions** icon and place a check mark next to **Forecasting** in the **Object Links** section
2. Move the right pane out of the way by clicking the right facing arrows .

Experiment by clicking on the store numbers on the button bar and observing how this filters the forecast visualization. To deselect a store, simply click the store number to deselect it.

Return to the top of the report and select a different date range and state. Notice the entire report is subset and all pages have been recreated following your selections.



1. Click the floppy diskette icon at the top of the page to save your report.

Congratulations! In a short period of time, you have used in-memory analytics and SAS Visual Analytics on SAS Viya to explore, visualize, and analyze data. Many more in-memory visual tools are available in this interface, including predictive analytics, data mining, and machine learning.