



## BUS 5100 Term Project Tutorial

**Authors:** Diana Roman; Grecia Hernandez Lugo; Jair Rocha; Melody Gulliver

**Instructor:** Jungwook Woo

**Date:** 04/01/2024

### Lab Tutorial

Group 5

California State University, Los Angeles

04/01/2024

### Spotify 2017 Trend Analysis Using SAP

---

**Objective:** This lab will analyze 2017 Spotify streaming data to identify streaming behavior trends across multiple variables, including time and location. This end-to-end guide is a reproducible resource and should provide the framework for additional analysis.

In this hands-on-lab, you will learn to:

- Collect a large-scale data set using Kaggle
- Create a data set and data model in SAP
- Construct data stories and visualizations in SAP
- Build a predictive model in SAP

#### **Step 1: Collect the data**

In this step, you will collect a data set from an external data science website.

1. Download the data set - *Spotify's Worldwide Daily Song Ranking*. Remember your file name and location.

NOTE: It may take a few minutes to download the file as it is quite large (45 MB).

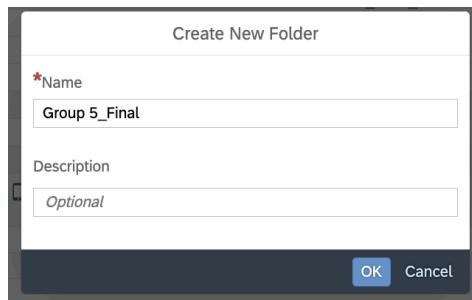
## **Step 2: Create a data set in SAP**

In this step, you will import the downloaded file into your SAP instance. You will transform the data set to improve its usability.

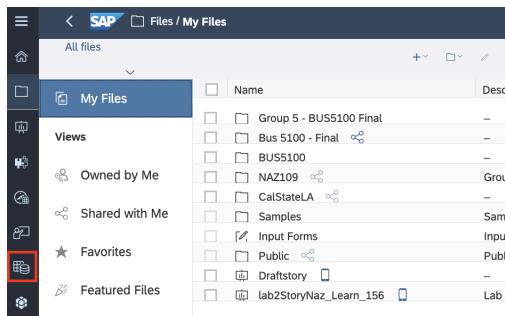
1. Sign into your SAP account
2. Select the **Folder** icon on the side navigation panel. Then, select the **Folder** icon on your action ribbon.



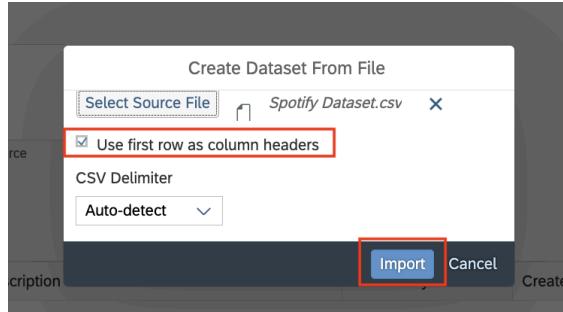
3. Select **New Folder** then update the **Name** field.



4. Select **OK**
5. Select the **Dataset** icon located on the side navigation panel.



6. Select **From a CSV or Excel File** then **Select Source File**.
7. Find the file you saved in section 1. Select **open**.
8. Your source file field should be updated. Verify the field **Use first row as column headers** is enabled. Then, select **Import**.



9. A “Save Dataset” modal will display. Name your file (i.e. “Spotify Dataset 2017”) and store the file in the Folder you created in section 1. For example, we have the following structure **My Files/Group5 - BUS5100 Final**. Then, click **OK**. NOTE: It may take a few minutes to import the file due to its size. Please keep your session active during this period.
  10. You will be redirected to your data table. Select the “URL” column. Then, select the **More icon (...)** → **Delete Column**.

Create Transform						
	Position	Track Name	Artist	Streams	URL	Date
1	1	Reggaetón Lento	CNCO	19272	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01
2	2	Chantaje	Shakira	19270	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01
3	3	Otra Vez (Feat. J. J.)	Zion & Lennox	15761	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01
4	4	Vente Pa' Ca	Ricky Martin	14954	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01
5	5	Safari	J Balvin	14269	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01
6	6	La Bicicleta	Carlos Vives	12843	https://open.spotify.com/track/1DwvLJXWzZGQHgkPjCqfA	2017-01-01

11. Select the “Region” column. Then, select the **Transform** icon (HINT: above the ...) → **Change Case** → **Change to UPPERCASE**



12. Select the “Region” column. On your action ribbon, select the **Geocoordinates** icon then **Area Name**.

The screenshot shows the SAP Datasets interface with a modal window titled "Geo Enrich By: Coordinates". The modal contains a table with two columns: "Area Name" and "Region". The "Region" column has three entries: "Area Name" (highlighted with a red box), "EC", and "EC". A red box also highlights the "Region" button in the modal's header.

13. A modal will display on the right side of your screen: “Geo Enrich by Area”. For Country/Region, select **Region** then **OK**.

The screenshot shows the "Geo Enrich by Area" configuration dialog. It has sections for "Geo Hierarchy", "Country/Region", "Region", and "Sub-Region". The "Country/Region" section is highlighted with a red box. It contains a radio button for "Select a column in your dataset" (selected) and a dropdown menu showing "Region" (also highlighted with a red box). Below it is another radio button for "Specify from a list of countries/regions" and a dropdown menu "Select from a list of countries...". The "Region" and "Sub-Region" sections have dropdown menus labeled "Select a column as Region (e.g. ...)" and "Select a column as Sub-Region (...)" respectively.

14. Select **SAVE** on your action ribbon (HINT: under FILE).

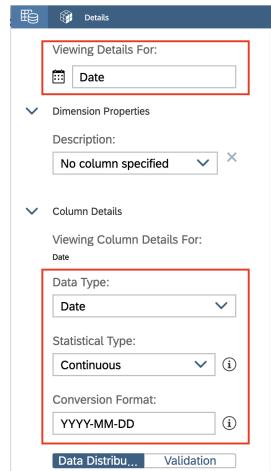
15. Select the **Level-Based Hierarchy** icon located in your Action ribbon.

- Name the Hierarchy “ArtistHierarchy”
- Add **Artist** as a dimension
- Then, add **Track Name** as a dimension
- Then, add **Region** as a dimension
- Click **OK**

The screenshot shows the "Level-Based Hierarchy" configuration dialog. It has a table of data on the left and a "Dimensions" panel on the right. The "Dimensions" panel shows a tree structure with levels: "Artist" (CNCO, Shakira, Zion & Lennox, Ricky Martin, J. L), "Track Name" (Reggaeton Lento, Chantaje, Otra Vez (feat. J. L), Vente Pa' Ca), and "Region" (EC). A red box highlights the "Artist" dimension. At the bottom, there is a "Select Dimension" dropdown and a "Cancel" button.

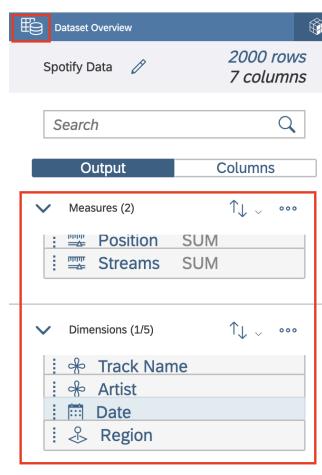
16. Select each column to view its settings. On your builder, verify the following the **column details** are as follows:

- a. **Track Name:** String, Nominal
- b. **Artist:** String, Nominal
- c. **Region:** String, Nominal
- d. **Position:** Integer, Continuous
- e. **Streams:** Integer, Continuous
- f. **Date:** Date, Continuous, YYYY-MM-DD



17. Select the **Data Set Overview** icon on your ribbon. Verify the measure and dimensions are configured as follows:

- a. **Measures:** Position, Streams
- b. **Dimensions:** Track Name, Artist, Date, Region

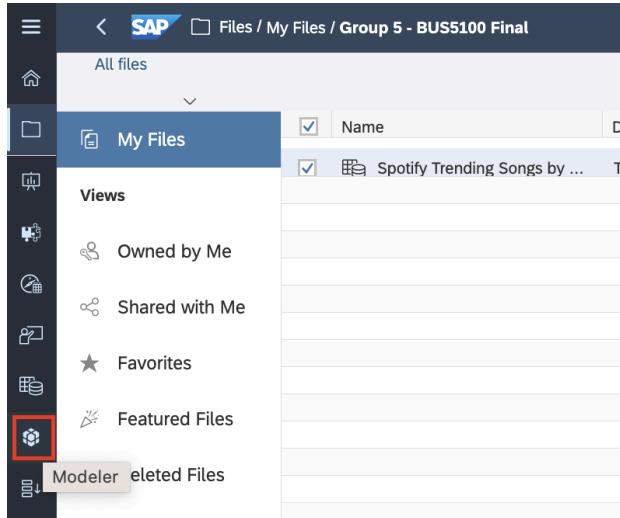


18. Select **SAVE** on your action ribbon (HINT: under FILE).

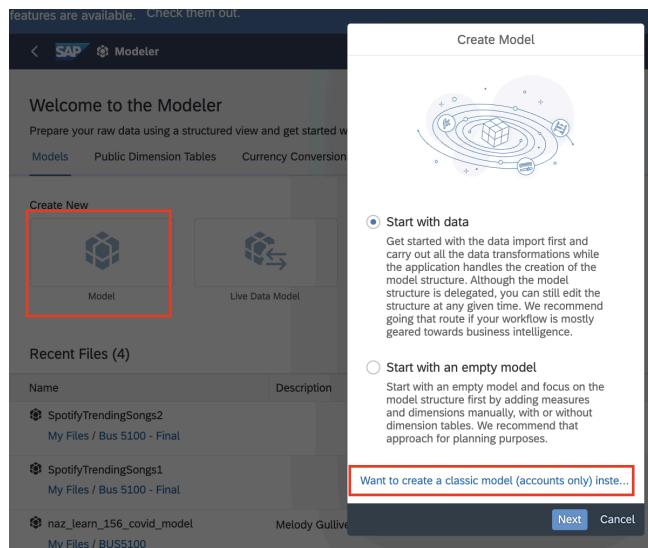
### **Step 3: Create a Model**

In this step, you will use the data set you created in step 2 to build a data model. The data model will allow you to define and format data fields, such as date and region.

1. Select the **Modeler** icon on the side navigation panel.



2. Select the **Model** icon then **Want to create a classic model (accounts only) instead?**



3. Select **Next**
4. Select **Start with data** then **Next**
5. Select **Dataset** from the list of options
6. Find the dataset you created in step 2. Select the file name. For example, we have the following structure **My Files/Group5 - BUS5100 Final**.  
NOTE: It may take a moment to complete this action.
7. Select **SAVE** on your action ribbon (HINT: under FILE).

8. Select the “Date” column then the **More (...)** icon → **Duplicate column**

A screenshot of a data editor interface. A table is displayed with columns: Date, Date\_2, Track Name, and Artist. The 'Date' column is currently selected, indicated by a blue header bar. A context menu is open over the second row of the 'Date' column, with the 'Duplicate column' option highlighted and surrounded by a red box.

9. Select the “Date\_2” column then the **Transform** icon. Select **Split on ‘-’**

A screenshot of a data editor interface. A table is displayed with columns: Date\_2, Track Name, Artist, and Region\_A. The 'Date\_2' column is currently selected, indicated by a blue header bar. A context menu is open over the second row of the 'Date\_2' column, with the 'Split on '-'' option highlighted and surrounded by a red box.

10. Select the “Date\_2” column. On your builder, verify the type is **Generic**.

A screenshot of the Modeling builder. Under the 'Type' section, a dropdown menu is open, showing 'Generic' as the selected option. The dropdown has a red border around it.

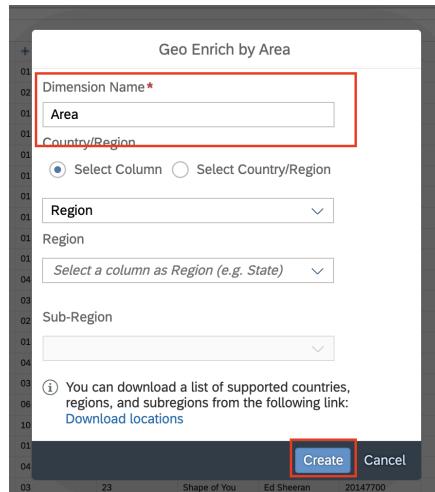
11. **Split** and **Label** the “Date\_2” data into Year, Month, and Day categories

- a. Double-click the “Date\_2” column. Rename it “Year.”
- b. Select the “Date\_2\_2” column then the **Transform** icon. Select **Split on ‘-’**
- c. Double-click the “Date\_2\_2” column title. Change it to “Month”. Select **Enter** on your keyboard to keep the new column name.
- d. Double-click the “Date\_2\_2\_2” column title. Change it to “Day”. Select **Enter** on your keyboard to keep the new column name.

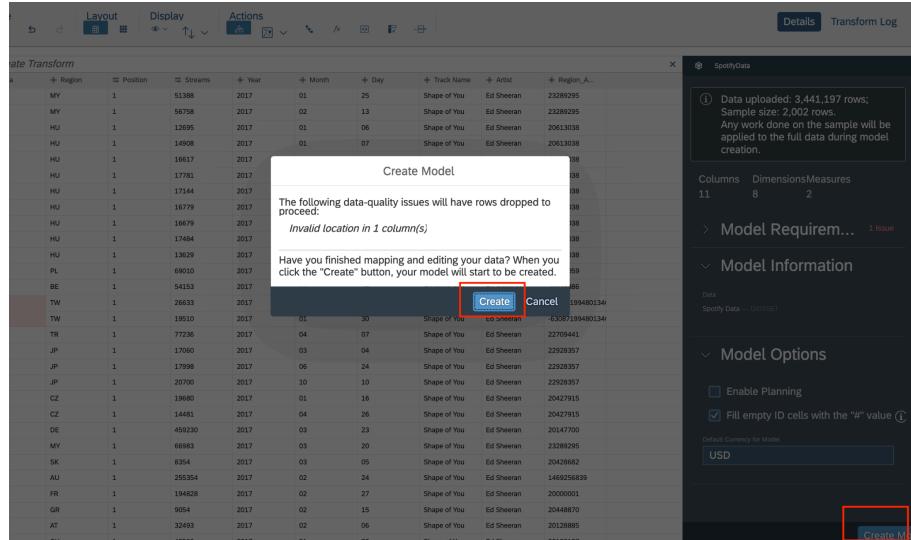
	Year	Month	Day	Track N
	2017	01	25	
	2017	02	13	
	2017	01	06	

12. Select the “Region” column. On your action ribbon, select the **Geocoordinates** icon then **Area Name**.

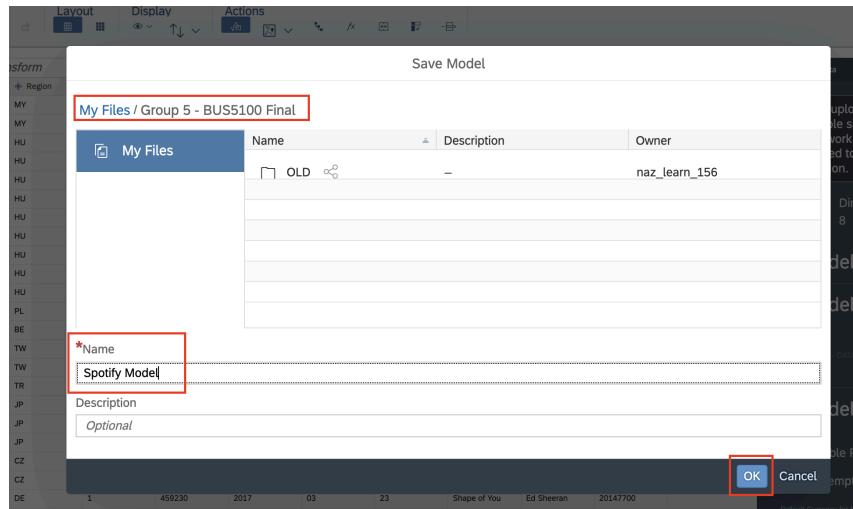
13. A modal will display on your screen: Geo Enrich by Area. For Country/Region, select **Region** then **Create**.



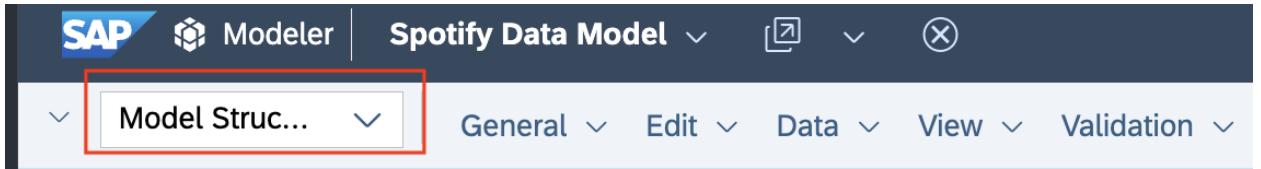
14. Select **SAVE** on your action ribbon (HINT: under FILE).  
 15. Select **Create Model** on your builder. It may take a moment to complete this action. If an alert modal displays, select **Create** again.



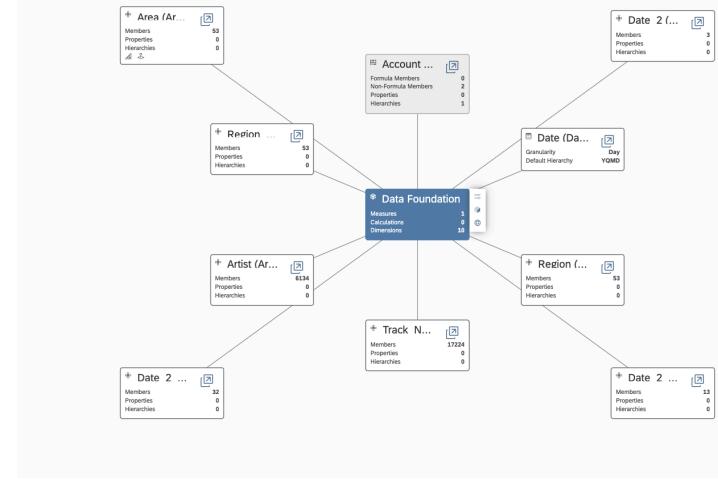
16. Save your model in the folder you created in section 2. For example, we have the following structure **My Files/Group5 - BUS5100 Final**. Update the Name field then click **OK**.



17. It may take a few minutes to create your model.  
 18. Your screen will redirect to your model. If you do not see your model, find the model in your saved files. For example, we have the following structure **My Files/Group5 - BUS5100 Final/Spotify Data Model**.  
 19. Verify you are on the **Model Structure** tab.



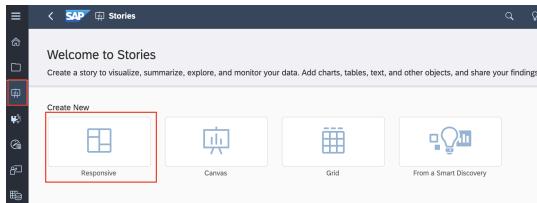
20. View your model.



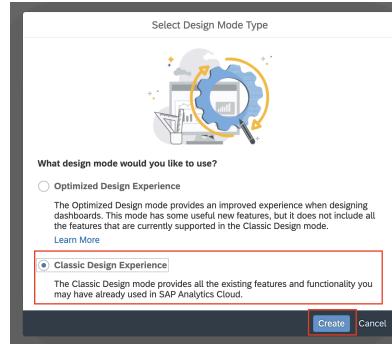
## Step 4: Create a Story

In this step, you will create a story using the data model from step 3. The story will contain diverse visualizations, including a geo-map and line/bar charts. These visualizations will help assess streaming trends in your data set.

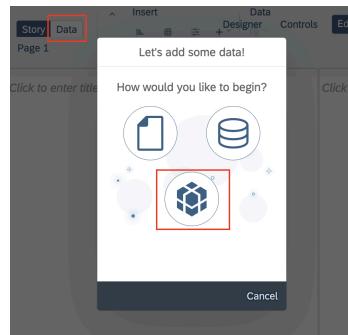
1. Select the **Story** icon on the side navigation panel. Then, select **Create New Responsive** page



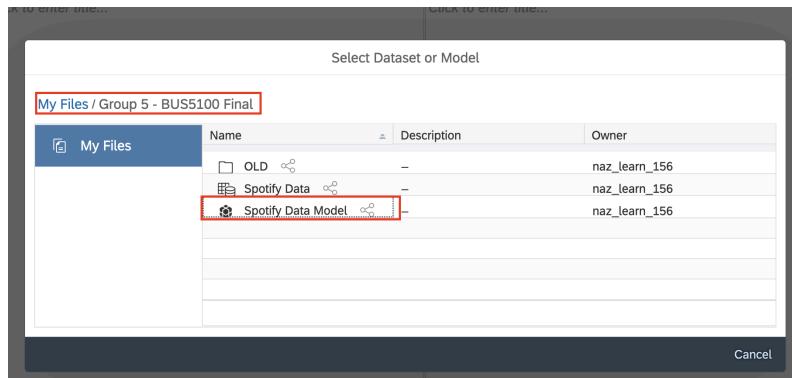
2. A Select Design Mode Type modal will display. Select **Classic Design Experience** then **Create**.



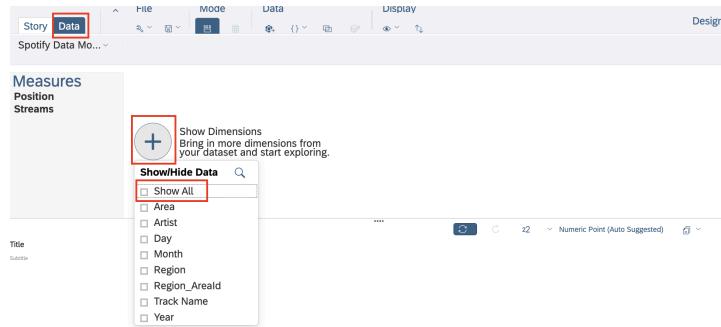
3. Select the **Data** tab. A modal will display. Select the **Dataset** icon.



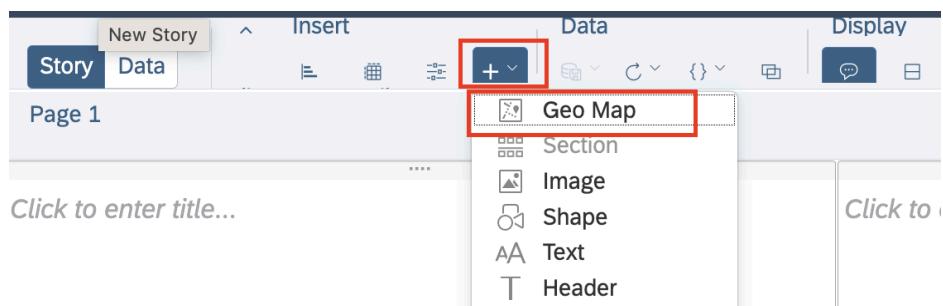
4. Find and select the data model you created in section 3. For example, we have the following structure **My Files/Group5 - BUS5100 Final/Spotify Data Model**



5. Select the **Data** tab. On the Show Dimensions area, select the + icon then **Show All**.

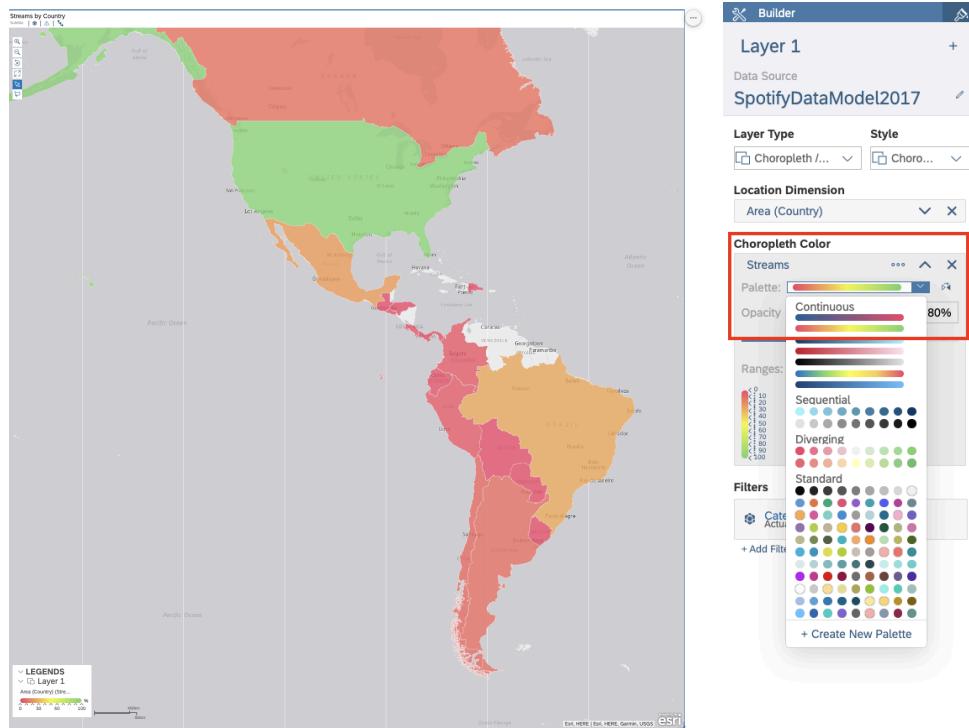


6. Return to the **Story** tab. Select the **Insert** + icon on the action ribbon. Then, the **Geo Map** button.



7. Modify your table through the builder.
- Under the Content layer, select **Add A Layer**.
  - For Layer Type, select **Choropleth / Drill Layer**.
  - For Style, select **Choropleth**.
  - For **Location Dimension**, select **Area (Country)**.
  - For **Choropleth Color**, select **Streams**.
  - Under **Choropleth Color**, select the arrow. Change the range to 10%.
  - Under **Choropleth Color**, select your color palette. For instance, we selected the second continuous value.
  - Select **OK**.
  - Double-click the Content Layers value. Renaming it streaming percentage.
  - For the **Base Layer**, select any base map option. For instance, **Light Gray**.

### Streams by Country in 2017



- Double-click the Geo-map's title. Change it to "Streams by Country."



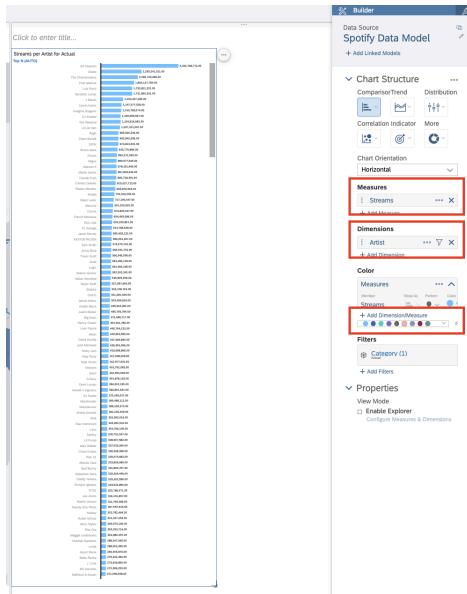
- Select **Click to enter title**. Add the title, "Streams by Country in 2017." You may use the Styling tool to customize the title's font size, color, etc.



- Select the second panel on your Story grid. Click the **Insert Chart** icon. Use the builder tool to modify your chart.

- For **Chart Structure**, select the Bar.
- For **Chart Orientation**, select Horizontal.

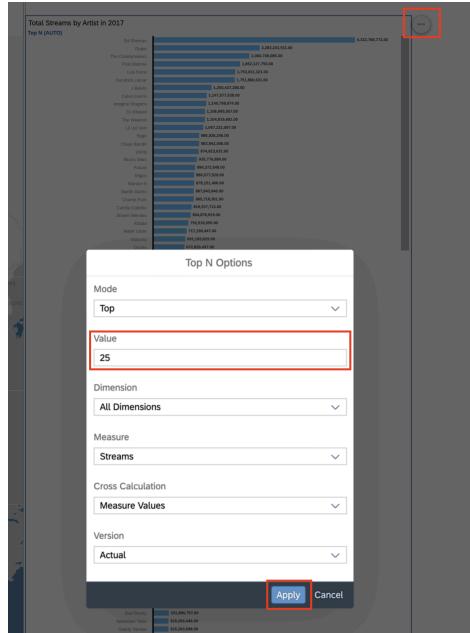
- c. Under the **Measures** section, **Add Stream**
- d. Under the **Dimensions** section, **Add Artist**. Change the color palette if desired.



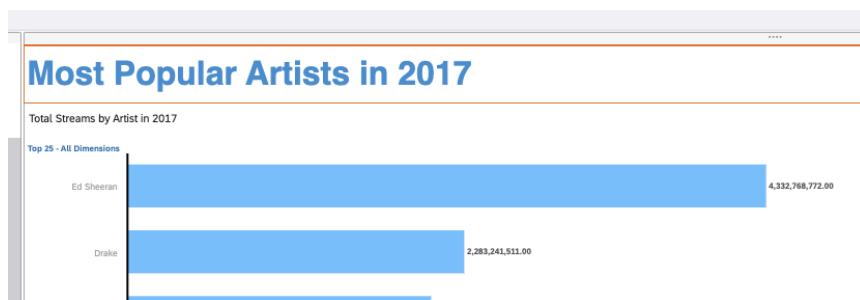
- 11. Double-click the bar chart's title. Change it to “Total Streams by Artist in 2017.”



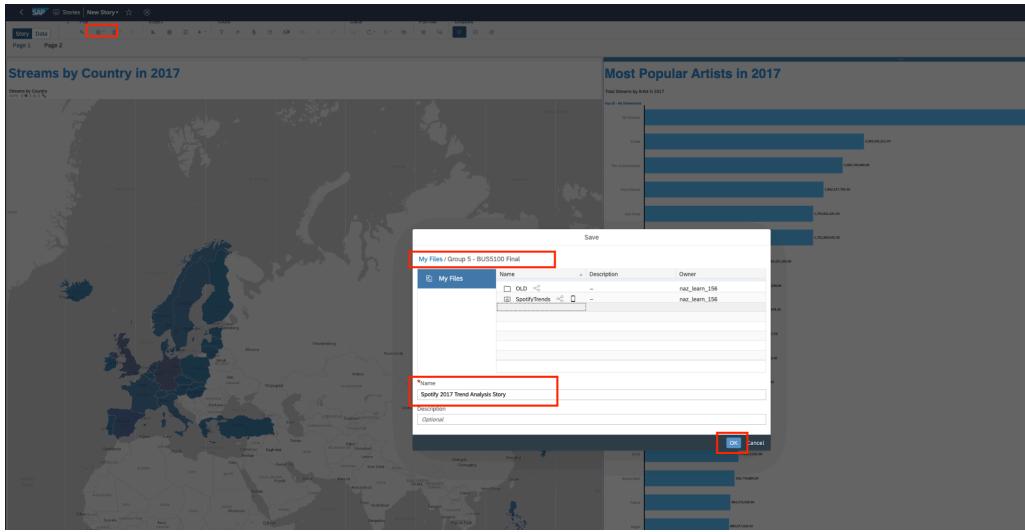
- 12. Select the **More (...)** icon then **Rank → Top N Options**. For value, enter **25**. Click **Apply**.



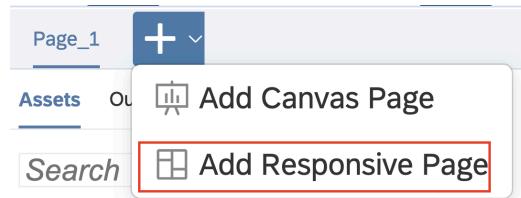
13. Select **Click to enter title**. Add the title, “Most Popular Artists in 2017.” You can modify the title through the styling tool.



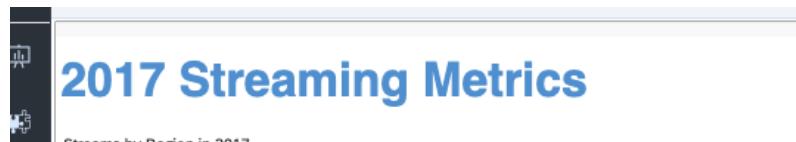
14. Save your story in the folder you created in section 2. For example, we have the following structure **My Files/Group5 - BUS5100 Final**. Name the model (i.e., “Spotify 2017 Trend Analysis Story”) then click **OK**.



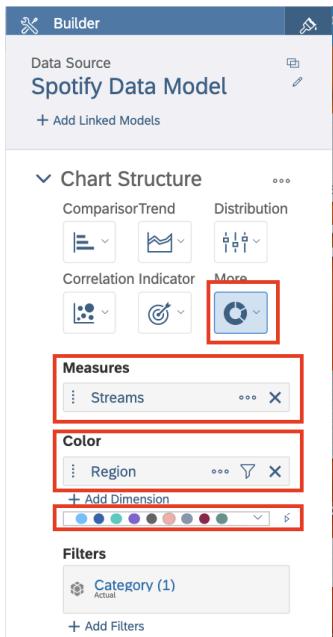
15. Navigate to your action ribbon to add a new page → select **Responsive**. This will be the 2nd page of your story.



16. Select Page 2, Panel 1. Select **Click to enter title**. Add the title, “2017 Streaming Metrics.” You may use the Styling tool to customize the title’s font size, color, etc.



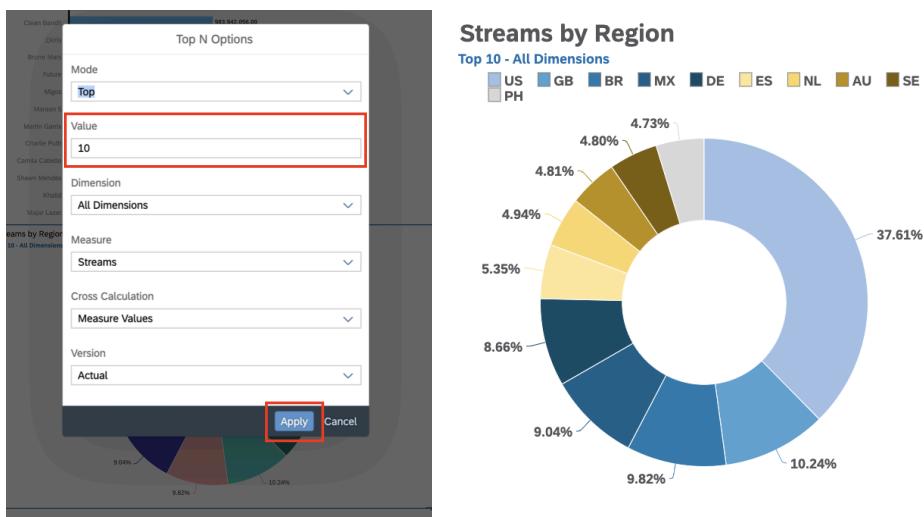
17. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
- For **Chart Structure**, select the **Donut**.
  - Under the **Measures** section, **Add Stream**
  - Under the **Dimensions** section, **Add Region**. Change the color palette if desired.



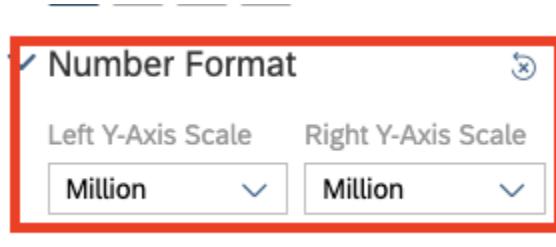
18. You may use the Styling tool to customize the chart's display font size, color, etc.
19. Double-click the bar chart's title. Change it to "Streams by Region in 2017."



20. Select the **More (...)** icon then **Rank → Top N Options**. For value, enter **10**. Click **Apply**.

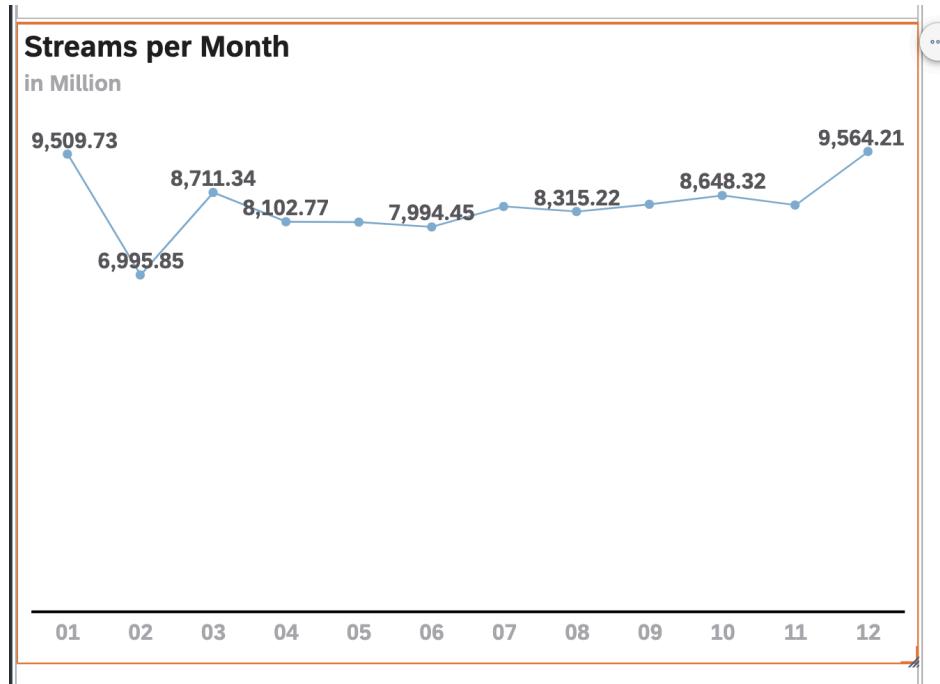


21. Remain on Page 2, Panel 1. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
- For **Chart Structure**, select the **Line** (HINT: Under Comparison Trend)
  - Under the **Left Y-Axis Measures** section, **Add Stream**
  - Under the **Dimensions** section, **Add Month**.
22. You may use the Styling tool to customize the chart's number formatting, display font size, color, etc. For instance, we changed the Number format to **Million** for both the left and right y-axes.



Scale Format

23. Double-click the line chart's title. Change it to “Streams Per Month”

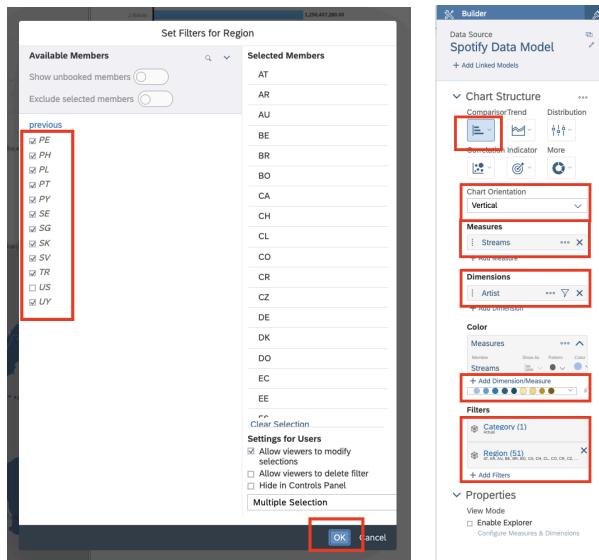


24. Select **SAVE** on your action ribbon (HINT: under FILE).
25. Select Page 2, Panel 2. Select **Click to enter title**. Add the title, “Most Popular Domestic vs. Foreign Markets”

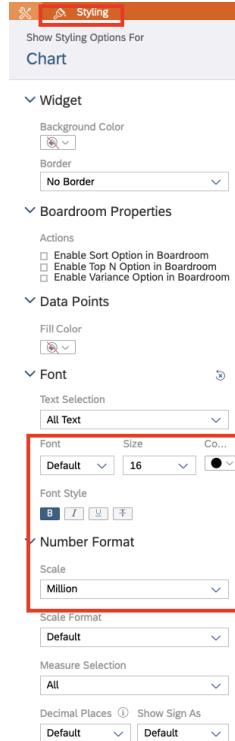
## Most Popular Domestic vs. Foreign Markets

26. Remain on Page 2, Panel 2. Click the **Insert Chart** icon. Use the builder tool to modify your chart.

- a. For **Chart Structure**, select the **Bar**.
- b. For **Chart Orientation**, select **Vertical**.
- c. Under the **Measures** section, **Add Stream**
- d. Under the **Dimensions** section, **Add Artist**. Change the color palette if desired.
- e. Under **Filters**, add **Region**. Select **all values except US**.



27. Select the paintbrush to navigate to the Styling tab. You may modify the display as desired—for instance, Size 16, Color Black, and Scale Million.



28. Double-click the bar chart's title. Change it to “Most Popular Artist (Foreign)”



29. Select **SAVE** on your action ribbon (HINT: under FILE).

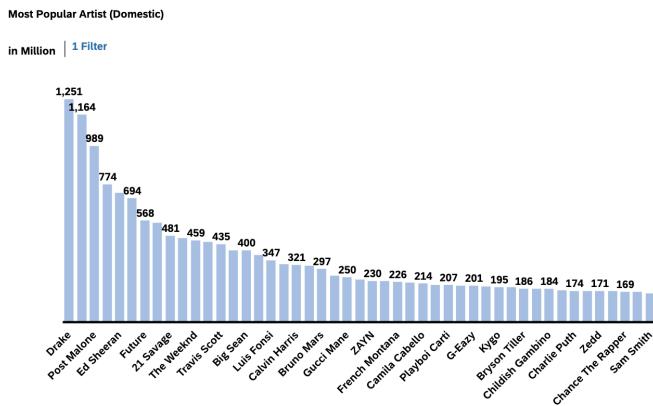
30. Remain on Page 2, Panel 2. Click the **Insert Chart** icon. Use the builder tool to modify your chart.

- For **Chart Structure**, select the **Bar**.
- For **Chart Orientation**, select **Vertical**.
- Under the **Measures** section, **Add Stream**
- Under the **Dimensions** section, **Add Artist**. Change the color palette if desired.
- Under **Filters**, add **Region**. Select **US** only.

31. Select the paintbrush to navigate to the Styling tab. You may modify the display as desired—for instance, Size 16, Color Black, and Scale Million.

32. Select the **More (...)** icon, then **Streams → Highest to Lowest**

33. Double-click the bar chart's title. Change it to “Most Popular Artist (Domestic)”



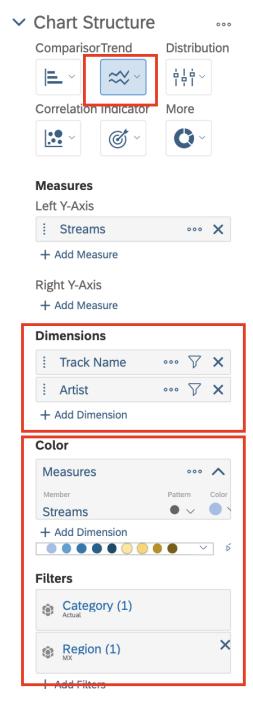
34. Remain on Page 2, Panel 2. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
- For **Chart Structure**, select the **Line** (HINT: Under Comparison Trend)
  - Under the **Left Y-Axis Measures** section, **Add Stream**
  - Under the **Dimensions** section, **Add Track Name**.
  - Under the **Dimensions** section, **Add Artist**.

35. Select the paintbrush to navigate to the Styling tab. You may modify the display as desired. For instance, Size 16, Color Black, and Decimal Places = 0.

36. Select the chart, then select the **More (...)** icon. From the list of options, select **Rank → All Dimensions → Top 5**.
37. Double-click the line chart's title. Change it to “Most Popular Songs Worldwide”



38. Remain on Page 2, Panel 2. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
- For **Chart Structure**, select the **Line**.
  - Under the **Left Y-Axis Measures** section, **Add Stream**
  - Under the **Dimensions** section, **Add Track Name**.
  - Under the **Dimensions** section, **Add Artist**.
  - Under the **Filter** section, **Add Region**. Select any region for comparison. We chose **MX**. (HINT: It is the second page.)

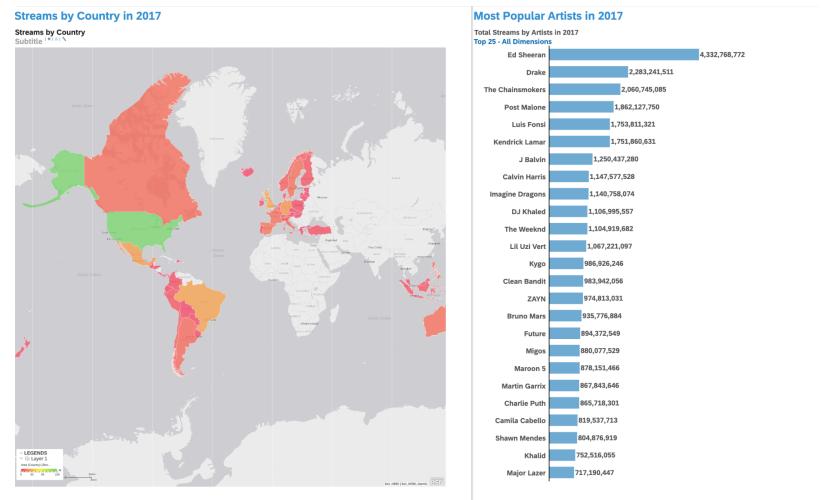


39. Select the chart, then select the **More (...)** icon. From the list of options, select **Rank → All Dimensions → Top 5**.
40. Select the paintbrush to navigate to the Styling tab. You may modify the display as desired. For instance, Size 16, Color Black, and Decimal Places = 0.

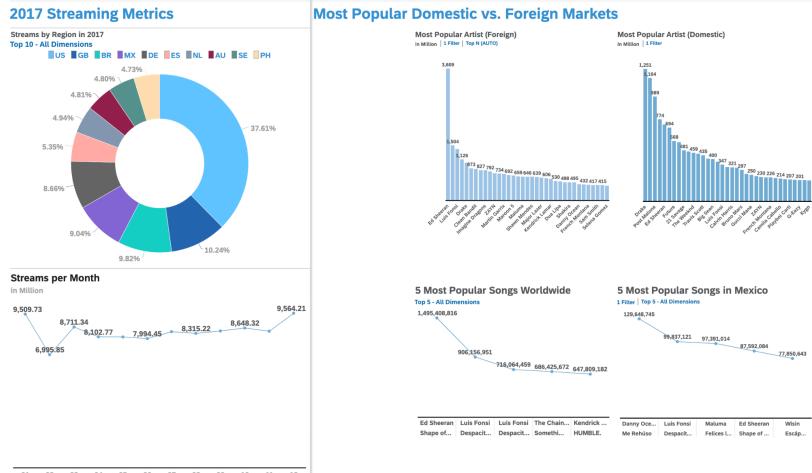
41. Double-click the line chart's title. Change it to “Most Popular Songs in [The Region You Selected]”



42. Your story should look similar to the images below:



Page 1



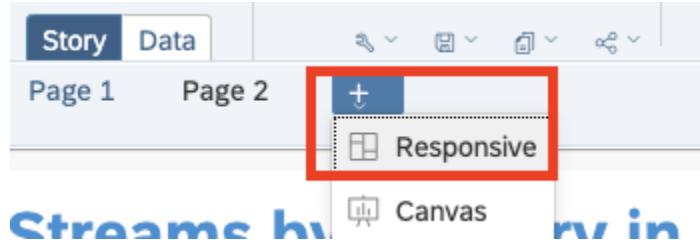
Page 2

43. Select **SAVE** on your action ribbon (HINT: under FILE).

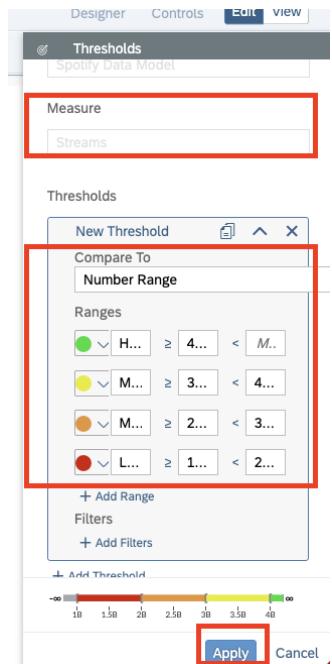
## **Step 5: Add Conditional Formatting to Your Story**

In this step, you will add a conditional formatting graph to your story. Conditional formatting will highlight changes in domestic streaming behaviors on a month-to-month basis.

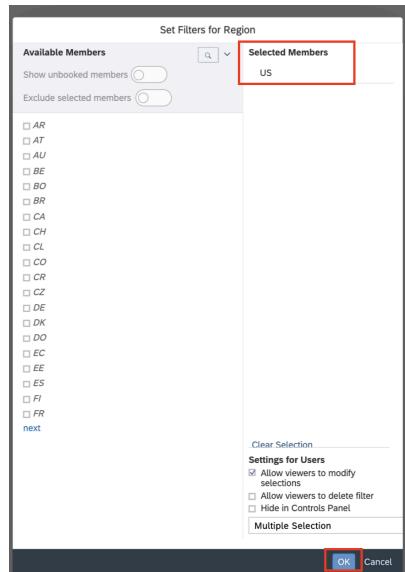
1. Return to your story. For example, we have the following structure **My Files/Group5 - BUS5100 Final/Spotify 2017 Trend Analysis Story**
2. Add a new page → select **Responsive**. This will be the 3rd page of your story.



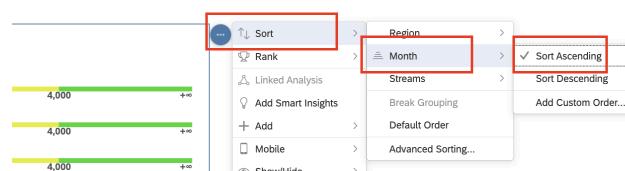
3. Follow the below steps to create a bullet chart
  - a. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
  - b. For Chart Structure, select the **Bullet Indicator** icon → **Bullet**
  - c. For Measures, add **Stream**
  - d. For Dimensions, add **Month**
  - e. For Color, add **Threshold** → select **Create Threshold**. A threshold panel will display.
    - i. For Measure, select **Stream**
    - ii. For Range, set the values as follows:
      1. High - Green - 4000000000 (Min)
      2. Mid-High - Yellow - 3000000000 (Min)
      3. Mid-Low - Orange - 2000000000 (Min)
      4. Low - Red - 1000000000
    - iii. Select **Apply**



4. Return to your builder. Select **Add Filters**, then select **Region**. From the list of options, select **US** then **OK**.



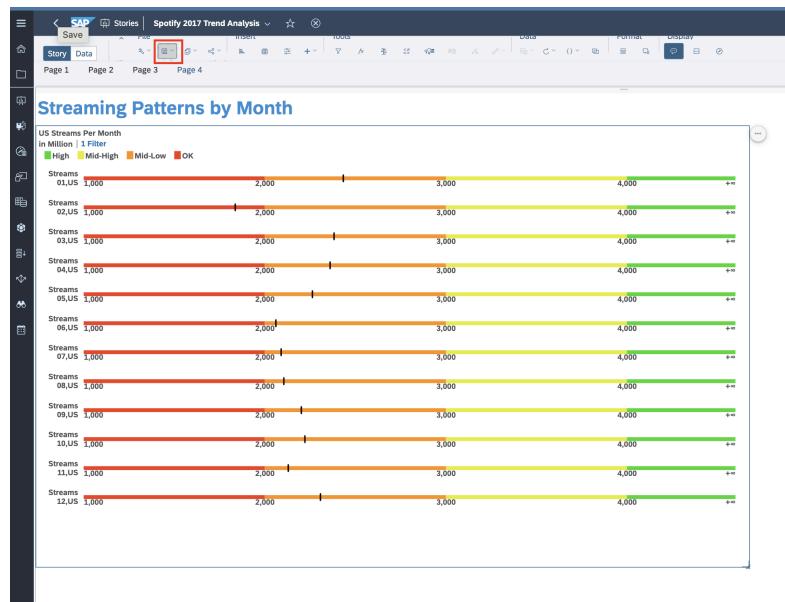
5. Select the paintbrush to navigate to the Styling tab. You may modify the display as desired—for instance, Size 20, Color Black, Bold, Scale Million.
6. Select the chart, then select the **More (...)** icon. From the list of options, select **Sort → Months → Sort Ascending**.



7. Double-click the chart's title. Change it to "US Streams per Month"
8. Select **Click to enter title**. Add the title, "Streaming Patterns by Month". Select the paintbrush to navigate to the Styling tab. You may modify the display as desired.



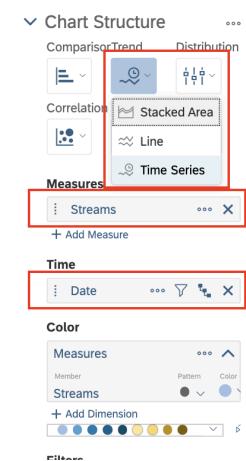
9. Select **SAVE** on your action ribbon (HINT: under FILE).



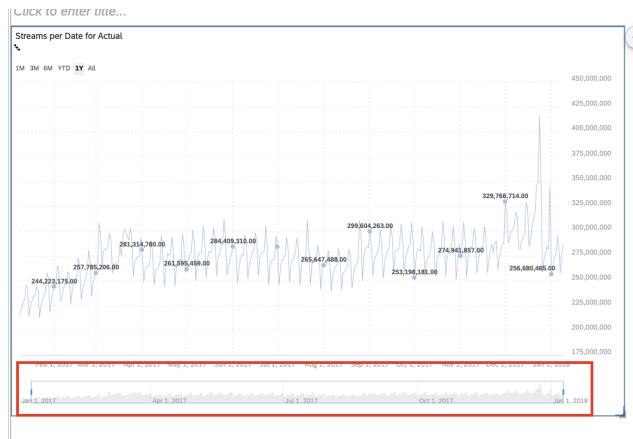
## Step 6: Create a Time Series Analysis

In this step, you will implement a predictive analysis using an SAP machine learning service. You will create a Time Series chart to visualize historical streaming behavior. Then, you will enhance the graph by applying an automatic forecast. This will allow the graph to measure future metrics.

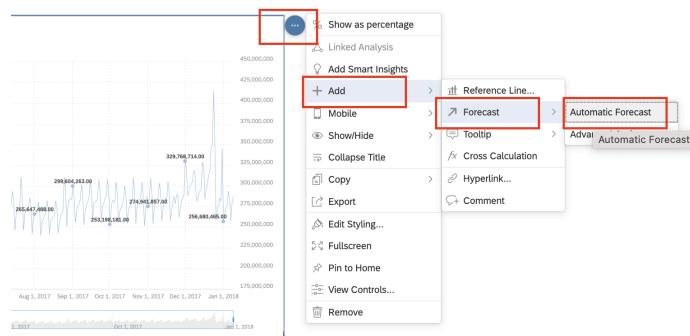
1. Add a new page → select **Responsive**. This will be the 4th page of your story.
2. Click the **Insert Chart** icon. Use the builder tool to modify your chart.
  - a. For Chart Structure, select **Time Series** (HINT: Under Comparison Trend)
  - b. For Measure, **add Stream**
  - c. For Dimension, **add Date**



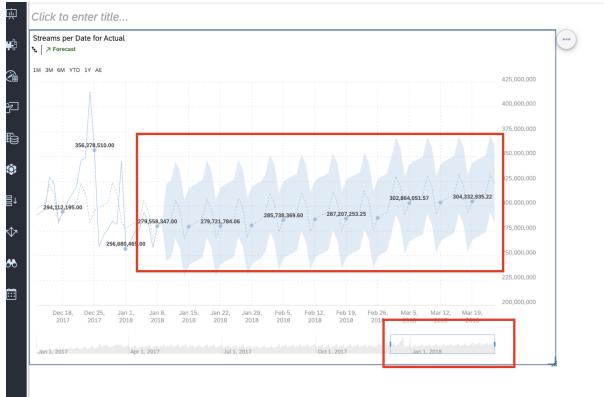
- The chart is defaulted to 1Y. You can adjust the time frame by adjusting the blue-lined date section underneath the graph.



- Select the **More (...)** icon → **Add** → **Forecast** → **Automatic Forecast**. It may take a few moments to complete this action.



- Adjust the slider to examine the forecast prediction.



- Select **Click to enter title**. Add the title, “Streams Per Date Forecast”. You may also use the styling tool to customize the text displays.



- Double-click the chart’s title. Change it to “Streams per date”

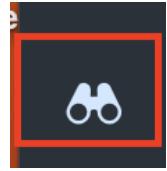


- Select **SAVE** on your action ribbon (HINT: under FILE).

## Step 7: Create a Predictive Model

In this step, you will implement a predictive analysis using an SAP machine learning service. You will create a Predictive Model to conduct a regression analysis. This will require you to customize and train your data model. When the model is complete, you will assess its quality and examine the results.

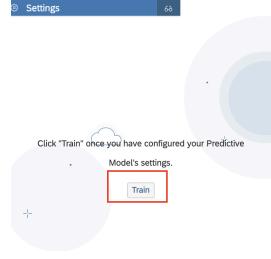
- Select the **Predictive Scenarios** icon located on the left-side navigation menu



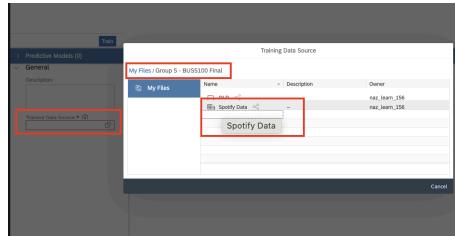
2. Select **Create New Regression** Predictive Scenario.



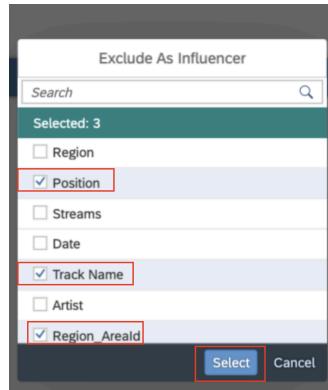
3. Name the scenario (i.e., “SpotifyTrendPredictiveScenario”). Save the file in the folder you created in section 2 then select **OK**. For example, we have the following structure **My Files/Group5 - BUS5100 Final**
4. Select the **Train** button. You will be redirected to a Predictive Model section.



5. Navigate to the Training Data Source field then select the file data set you created in Step 2. For example, we have the following structure **My Files/Group5 - BUS5100 Final/Spotify Data**



6. Modify your model settings.
  - a. For **Target**, select **Streams**
  - b. For **Influencers**, exclude the **Position**, **Track Name**, **RegionArea\_ID** then click **select**



7. Verify your predictive model settings.

A screenshot of the "Predictive Models" configuration screen. It shows a list of models with their details. One model is selected, showing its configuration: "Training Data Source" set to "Spotify Data", "Target" set to "Streams", and "Exclude As influencer" set to "Position X 2 More". A red box highlights the "Training Data Source" field. Other sections like "Predictive Goal" and "Influencers" are also visible.

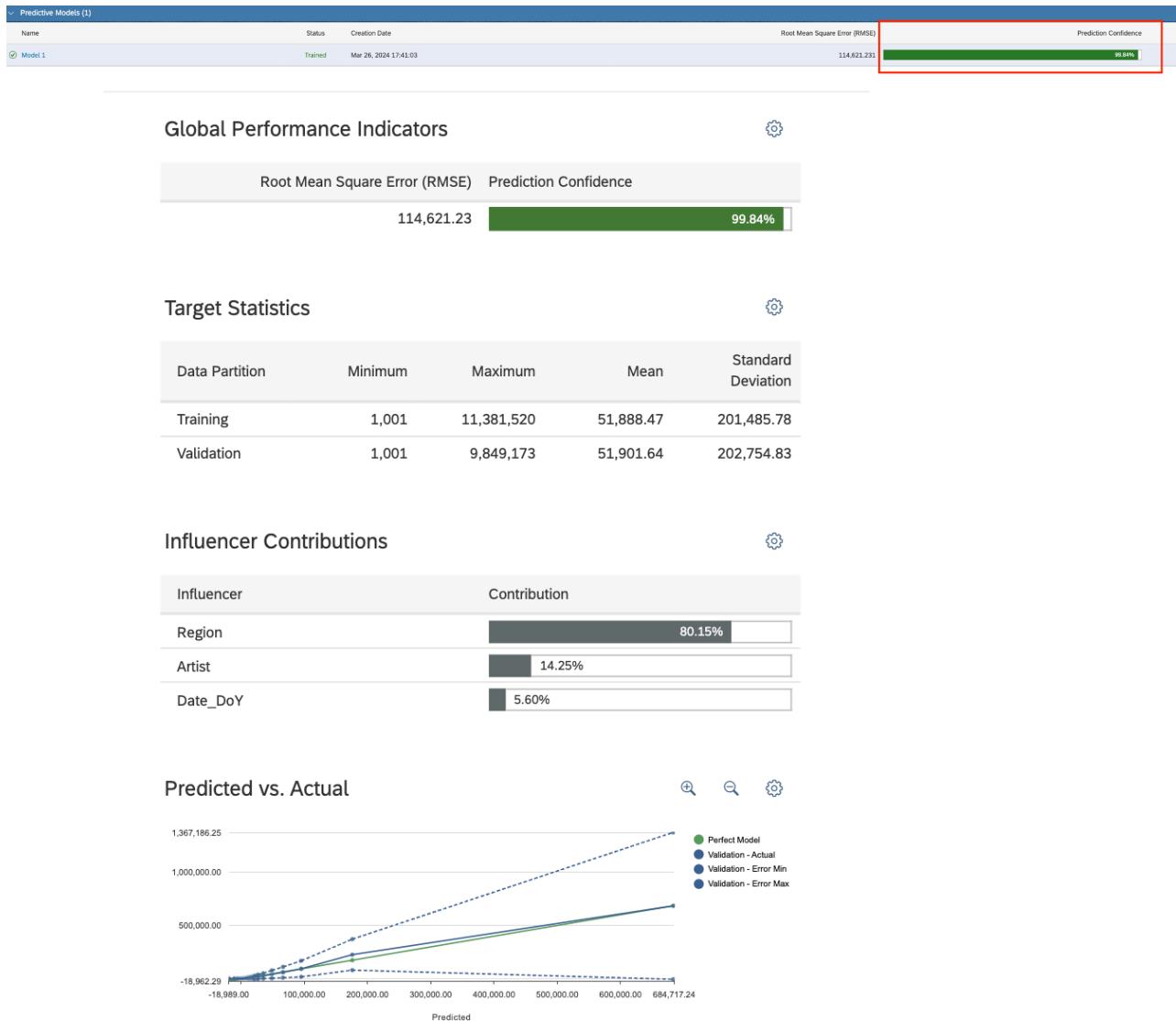
8. Select Train.

Train

9. Your model will show a status of “Training”. NOTE: It may take over 5 minutes to complete this action.

Predictive Models (1)						
Name	Status	Creation Date	Predictive Power	Prediction Confidence	Influencer Count	Record Count
*** Model 1	Training	-	-	-	-	-

10. After the training is completed, view your predictive model. Notice the Prediction Confidence is over 99%.



## References

---

1. Data set URL:  
<https://www.kaggle.com/datasets/edumucelli/spotifys-worldwide-daily-song-ranking>
2. Github URL: <https://github.com/BUS5100group5/2017spotifytrends>