# Dashboarding: Where Data Becomes Information

Hands-On Workshop

## Purpose

This activity demonstrates how to explore and visualize data using SAS Visual Analytics as a powerful dashboarding tool. Users will learn to load datasets, analyze variables, and create visualizations that transform raw data into actionable insights.

## Scenario

You've joined Altitude Express, an airline company celebrated for its innovative, data-driven decision-making. As a new data analyst, your mission is to delve into the company's flight data to uncover insights that optimize business operations and enhance customer satisfaction. Building on your predecessor's work, you'll explore the dataset and create visualizations that reveal trends, helping reduce flight no-shows, refine ticket pricing, and improve overall efficiency.

## Learning Objectives

- Analyze and prepare data for exploration.
- Create and refine interactive visual reports.
- Interpret visualizations to inform decisions.

## Software

This activity uses no-code data management and visualization tools in SAS Visual Analytics on SAS Viya 2024.09 LTS Release (September 2024).

## SAS Course Alignment

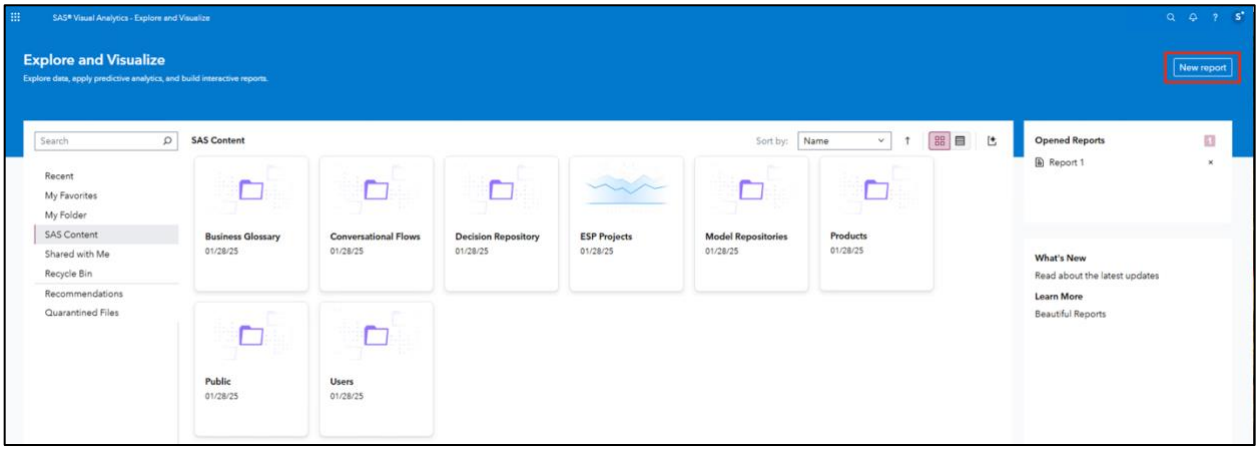
This activity compliments *SAS Visual Analytics 1 for SAS Viya: Basics* and can be used to practice skills learned in the course.

## Table of Contents

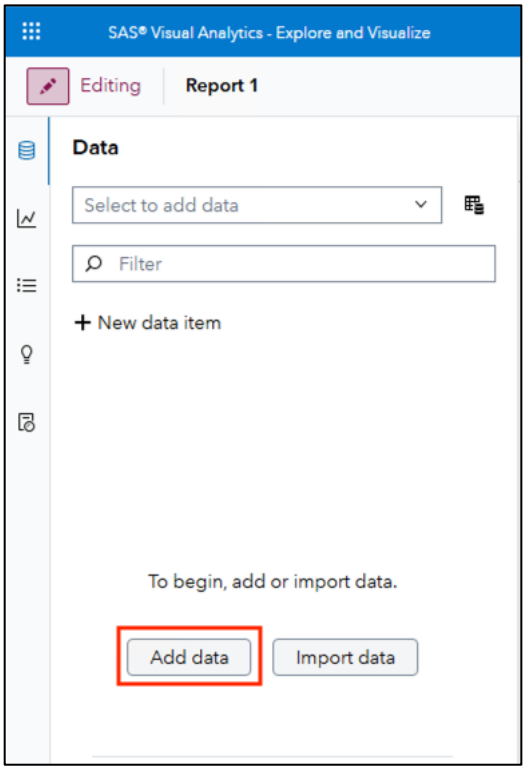
# Altitude Express: Flight Path to Smarter Data

## Part 1: Accessing Data

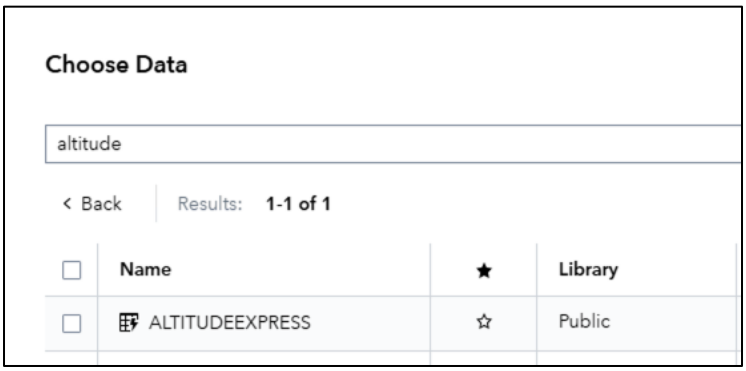
1. If you not there already, navigate to **Explore and Visualize.**
2. Click **New report.**



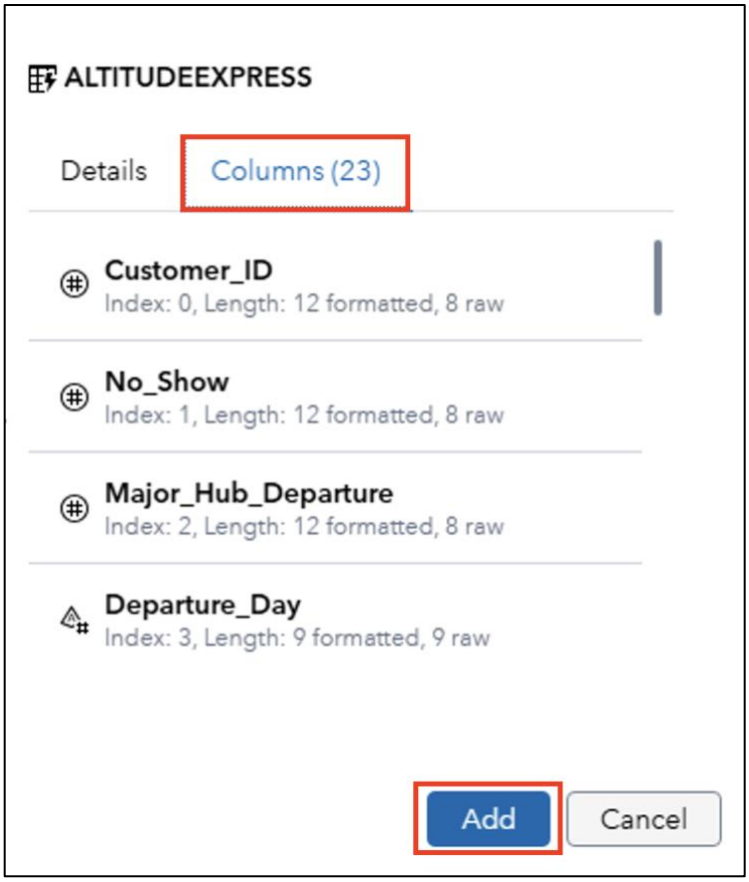

3. From the Data pane, click **Add data**.



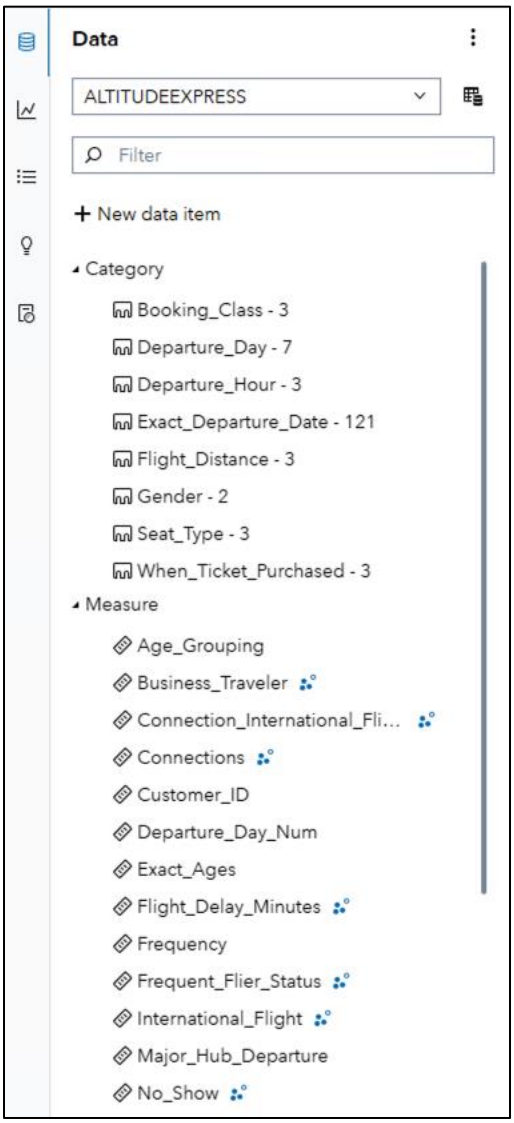

4. Using the search, find the ALTITUDEEXPRESS dataset.




5. Feel free to check out the columns in the ALTITUDEEXPRESS dataset, then **Add** it to your report.



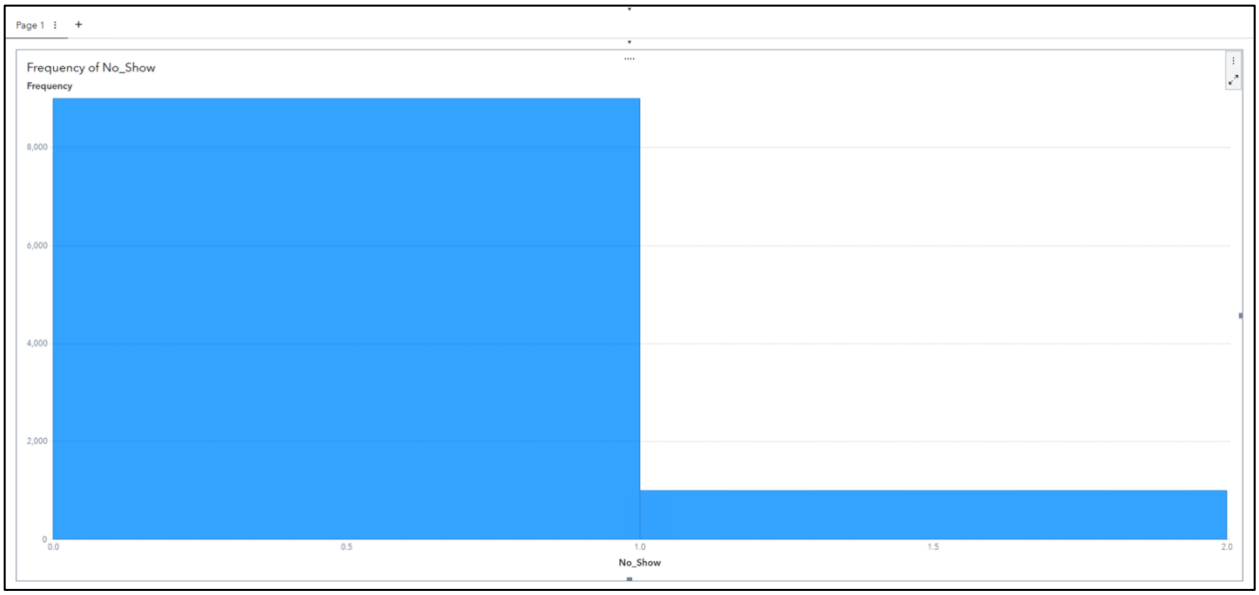

6. Now that we've loaded the dataset into our report, we see the variables in the **Data pane**.



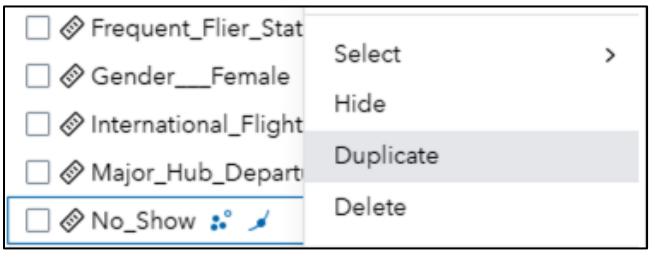

7. Notice here that next to the **Category** variables, there's a number indicating the cardinality. The icons next to **Measure** variables indicate that outliers and correlations between variables have been detected. Gotta love built-in SAS insights!

## Part 2: Exploring the Data – and Finding Stuff to Fix!

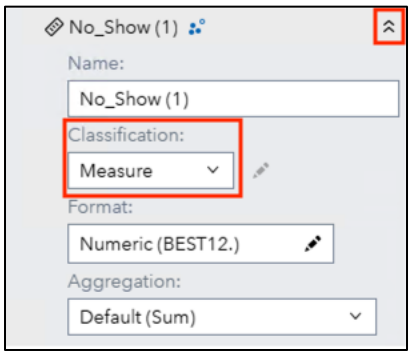
1. The *No_Show* variable is going to be of interest to us because we want to minimize the number of people who aren't appearing for flights. This is inefficient and if we could have predicted it, oversold tickets to have a full flight.
2. To investigate this variable, drag it from the **Data pane** onto the canvas. You will see a "+ Auto chart" indicator meaning that SAS is going to choose the best visualization given the variable you want to analyze. A histogram is created!



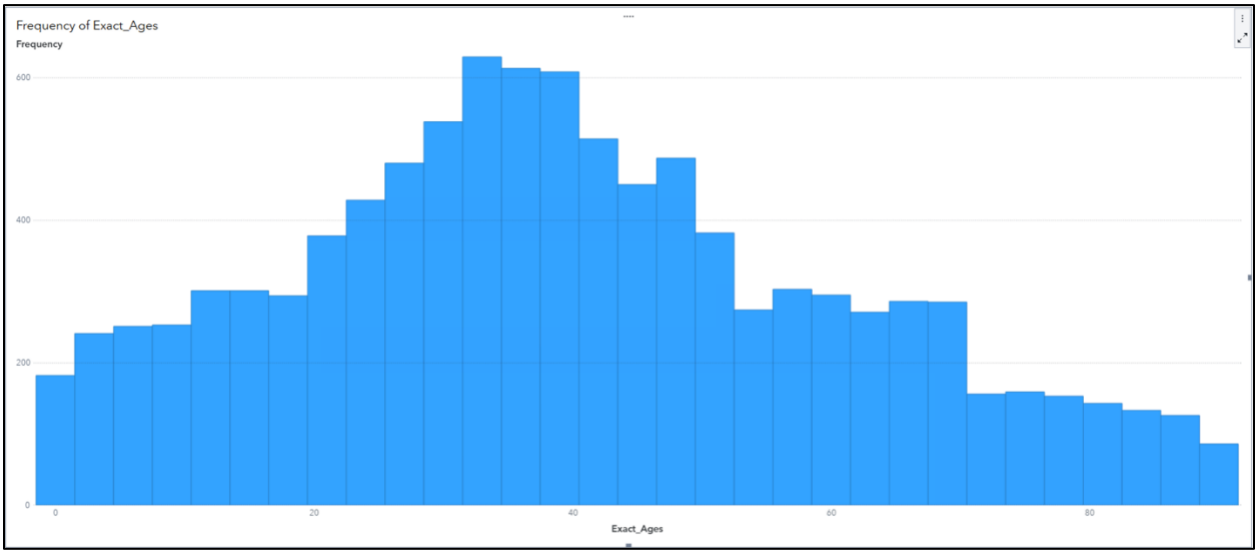

3. We can see there are significantly more *No_Show* between 0 and 1 than between 1 and 2. Well, we know that the *No_Show* variable is either equal to 0 or 1 for No or Yes. So, it would be better if we could analyze these as distinct values.
4. Duplicate the *No_Show* variable by right-clicking on the variable in the **Data pane**.



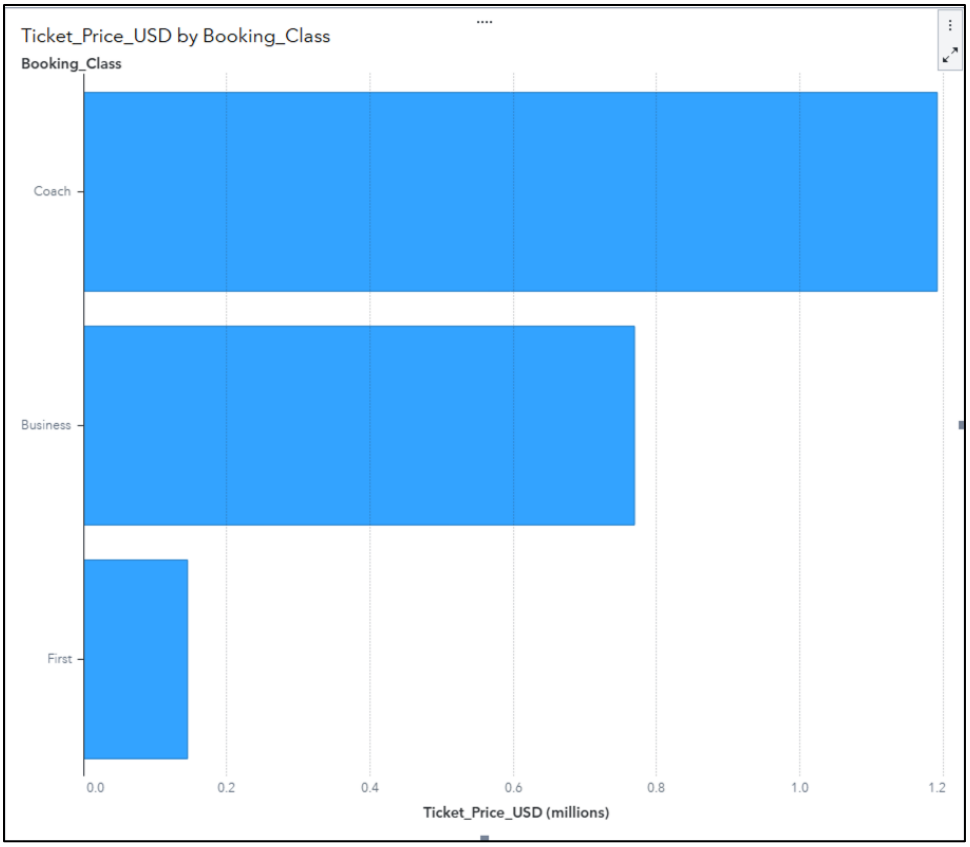

5. Using the **Edit properties** in the upper right corner of the newly created *No_Show (1)* variable, change the **Classification** from Measure to **Category**. You can also change the name if you wish.



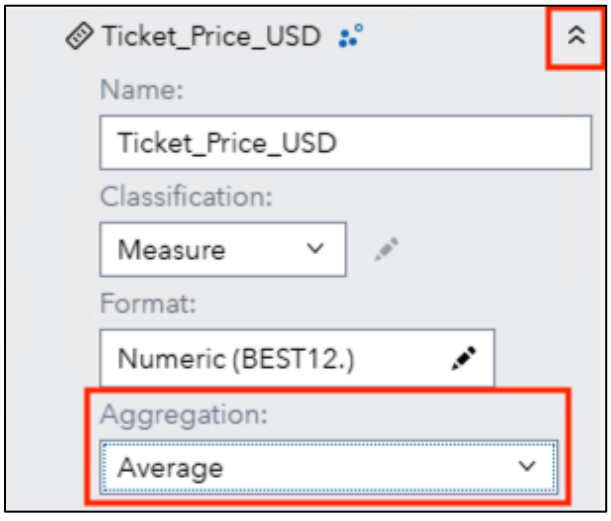

6. Now you have a *No_Show* that is a measure and a *No_Show* that is a category. Notice that the Category has 2 distinct values.
7. Feel free to explore other variables by dragging them onto the canvas. For instance, dragging *Exact_Ages* onto the canvas gives a histogram showing the frequency and variability.



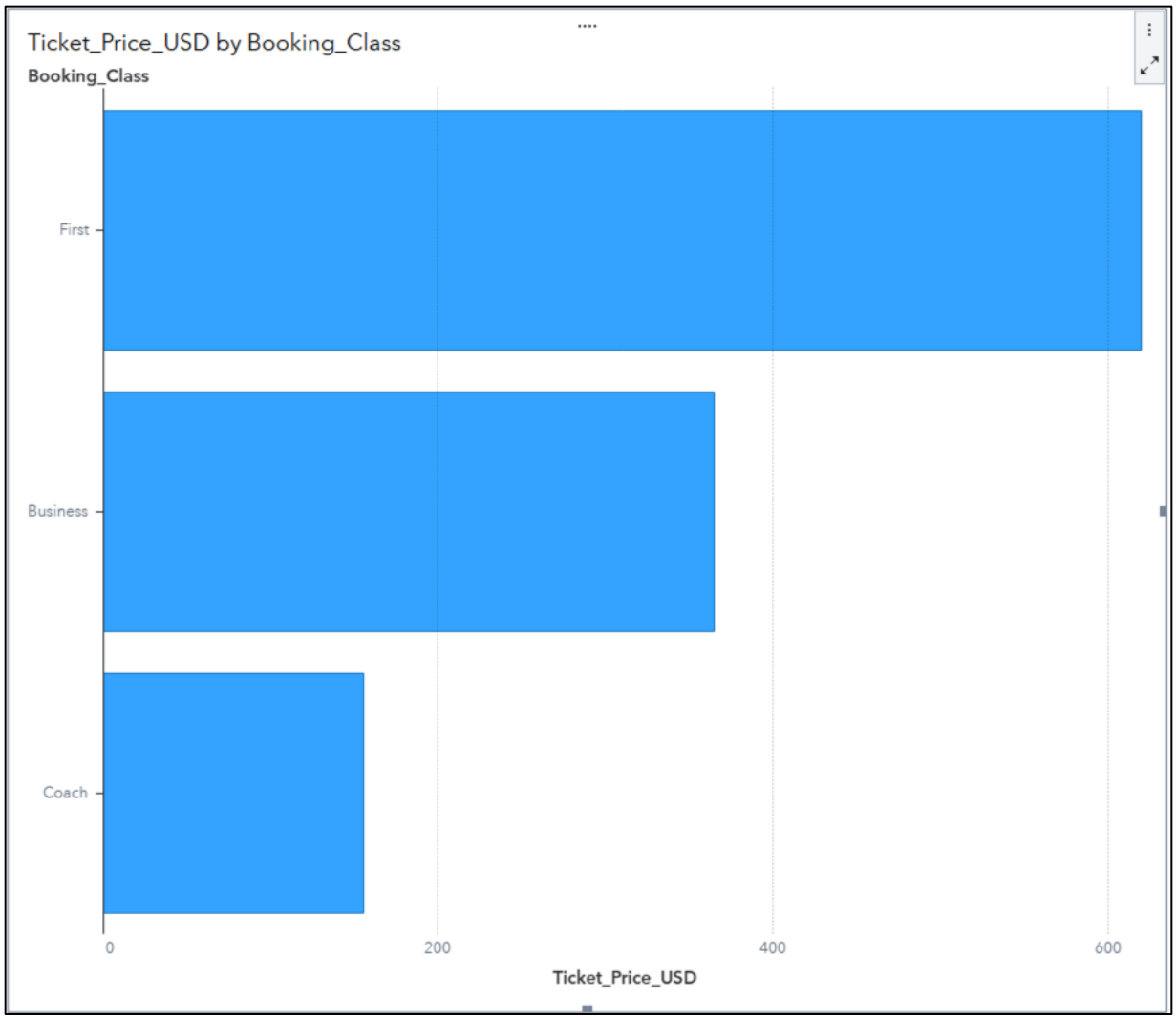

8. Wanna get fancy? Select both *Booking_Class* and *Ticket_Price_USD* then drag them onto the canvas! A bar chart is created! Notice that the scale for the *Ticket_Price_USD* is in millions – this is because the values are being aggregated, and the default is sum. This chart is likely not very representative of ticket prices since First class is more luxurious and often more expensive than Business or Coach.



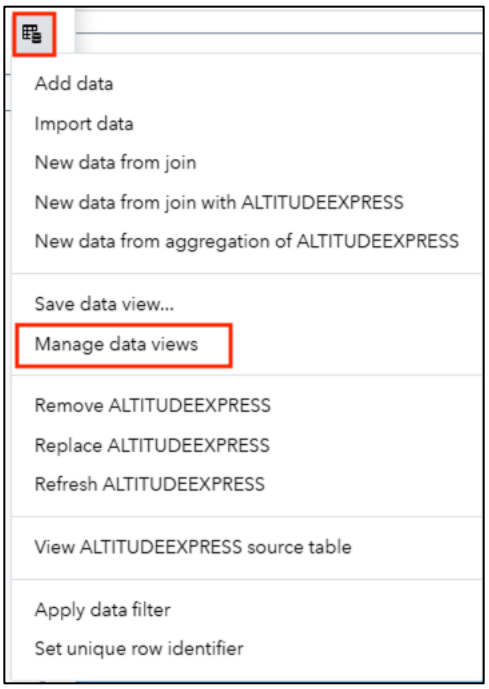

9. You can change the Aggregation of *Ticket_Price_USD* by, you guessed it, using the **Edit properties** next to the variable and changing the **Aggregation** to **Average**!




10. Now check out the bar chart!

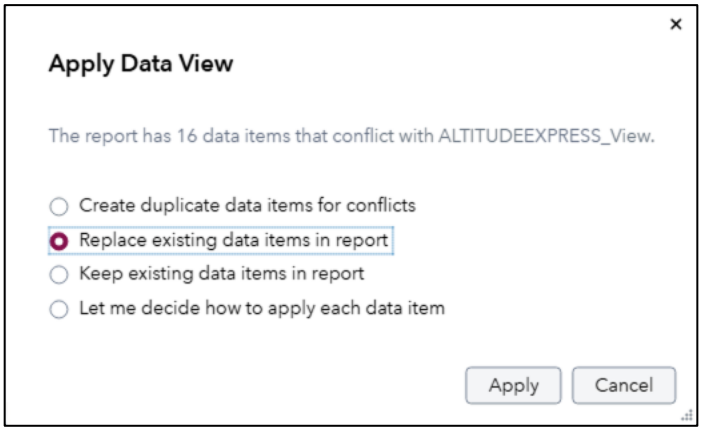



11. Whew! We could do data prep all day long – but that's not the goal. Luckily a coworker has already completed some data prep and saved their work in a "view." You can think of views as instructions that go along with a dataset to make metadata changes.
12. To load their view, go to the button next to ALTITUDEEXPRESS in the **Data pane** and select **Manage data views**.

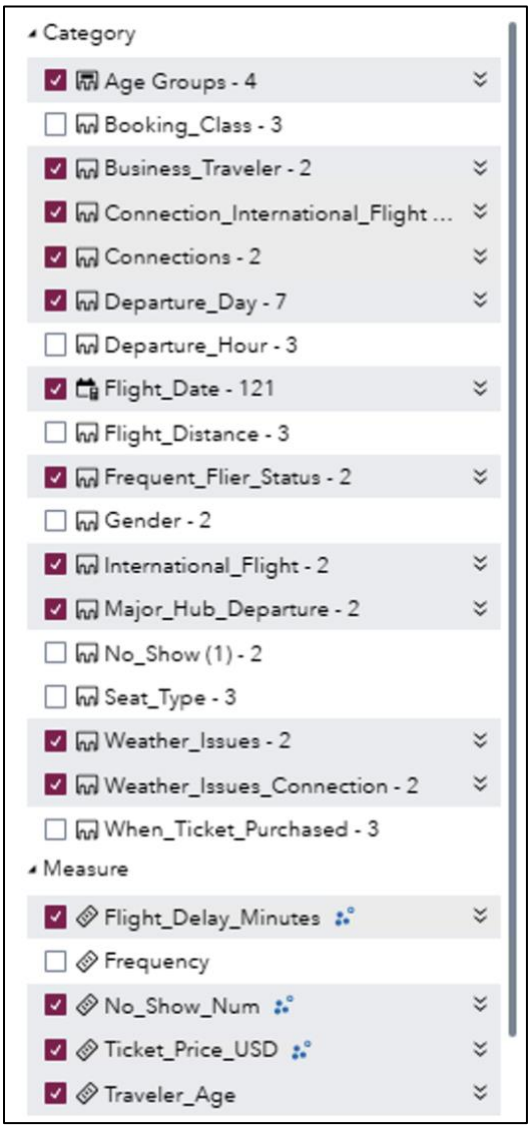



13. Select the ALTITUDEEXPRESS_View and click **Apply**.


14. When asked how to apply the data view, select the **Replace existing data items in report** option.



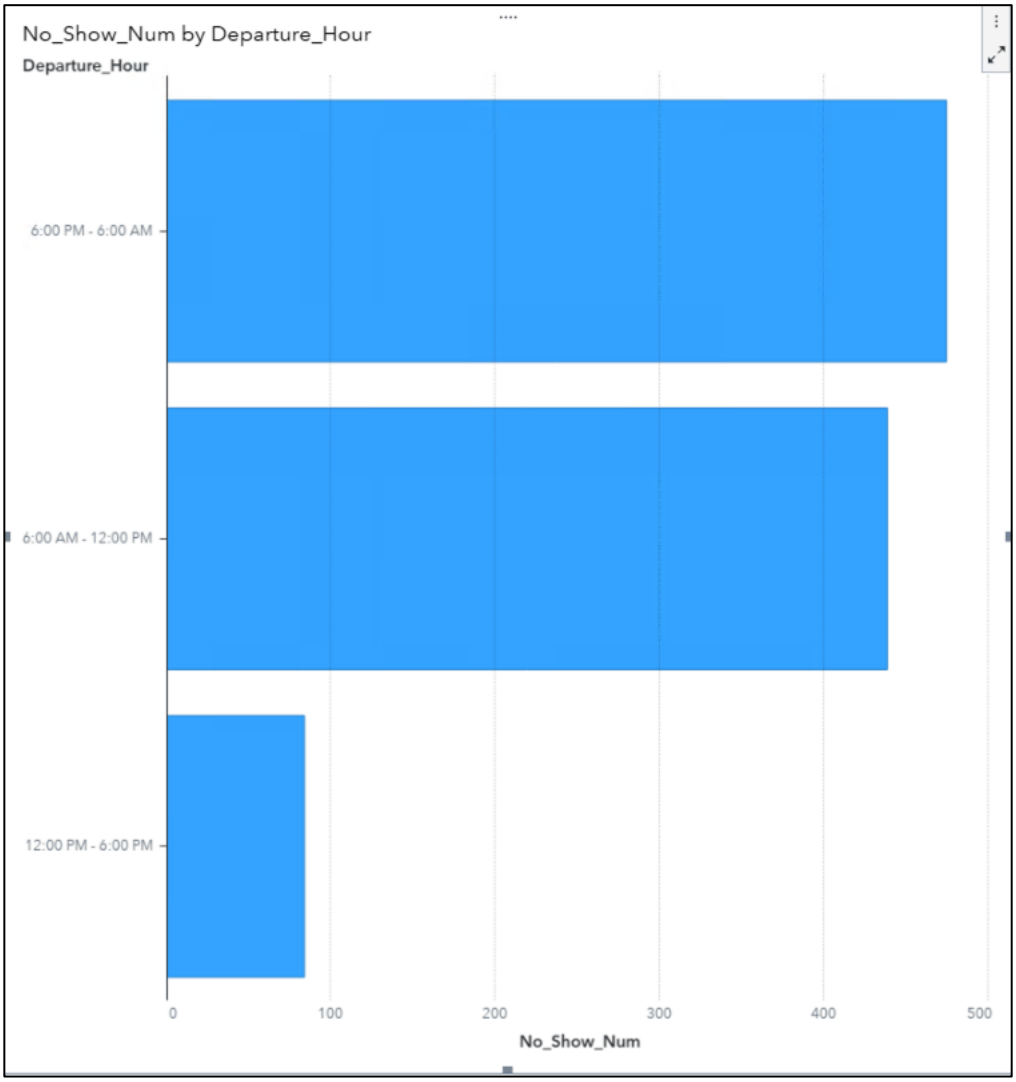
15. Your Data pane will be updated with new variables, converted variables, and updated aggregations.



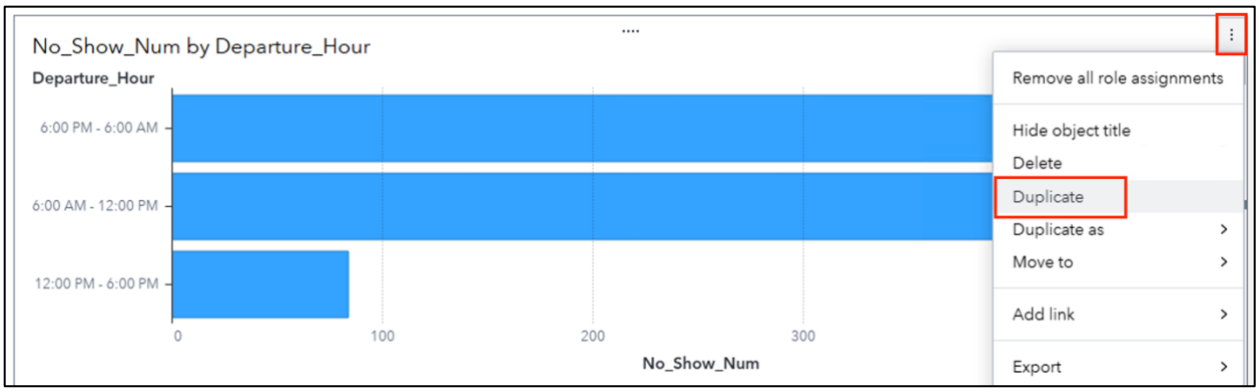

16. Click the **Clear selection option** at the bottom of the pane and get ready to really dive into the analysis!

## Part 3: Turning Data into Information

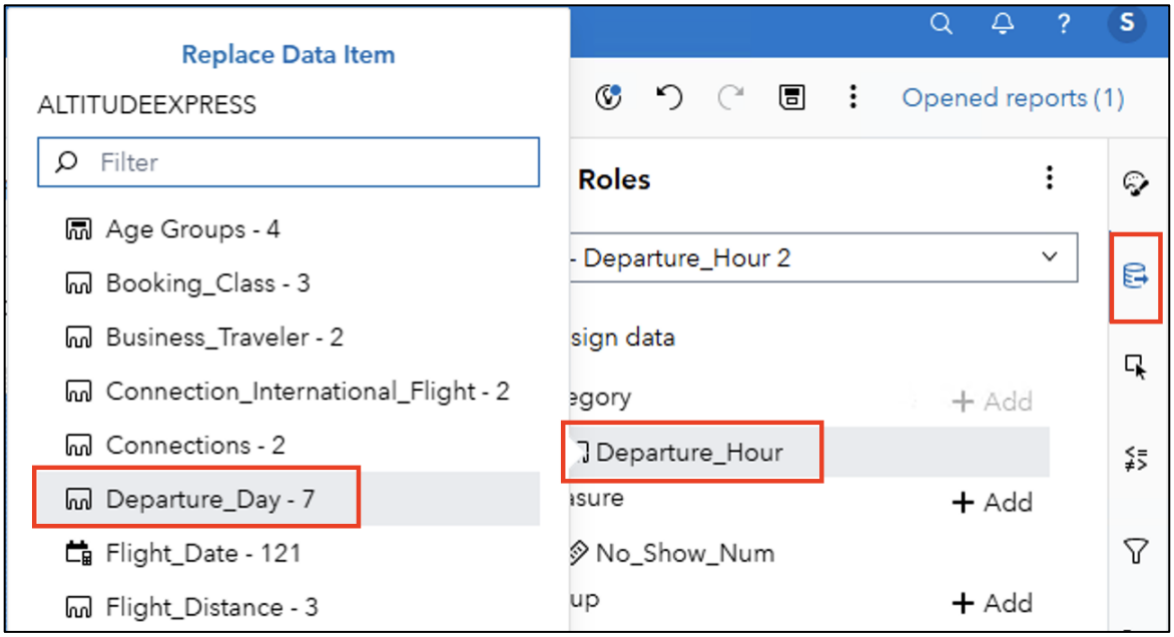
1. Add a new page to your report by clicking the **+** button next to Page 1.
2. Select *No_Show_Num* and *Departure_Hour* from the **Data pane** and drag them onto the canvas.



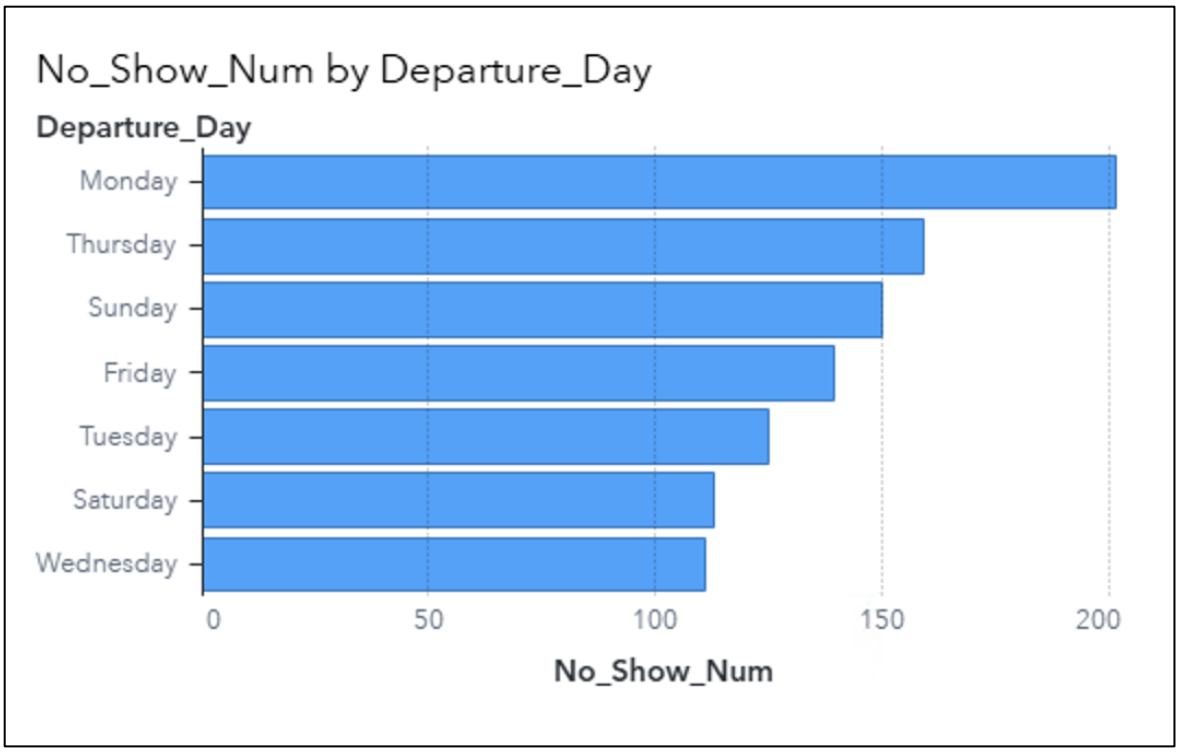

3. Looks like the greatest number of no shows happen overnight, i.e. between 6pm-6am, followed closely by the morning hours, 6am-12pm.
4. What about no shows by day of the week?
5. Duplicate the *No_Show_Num by Departure_Hour* bar chart by clicking the options button in the upper right corner and selecting **Duplicate**.



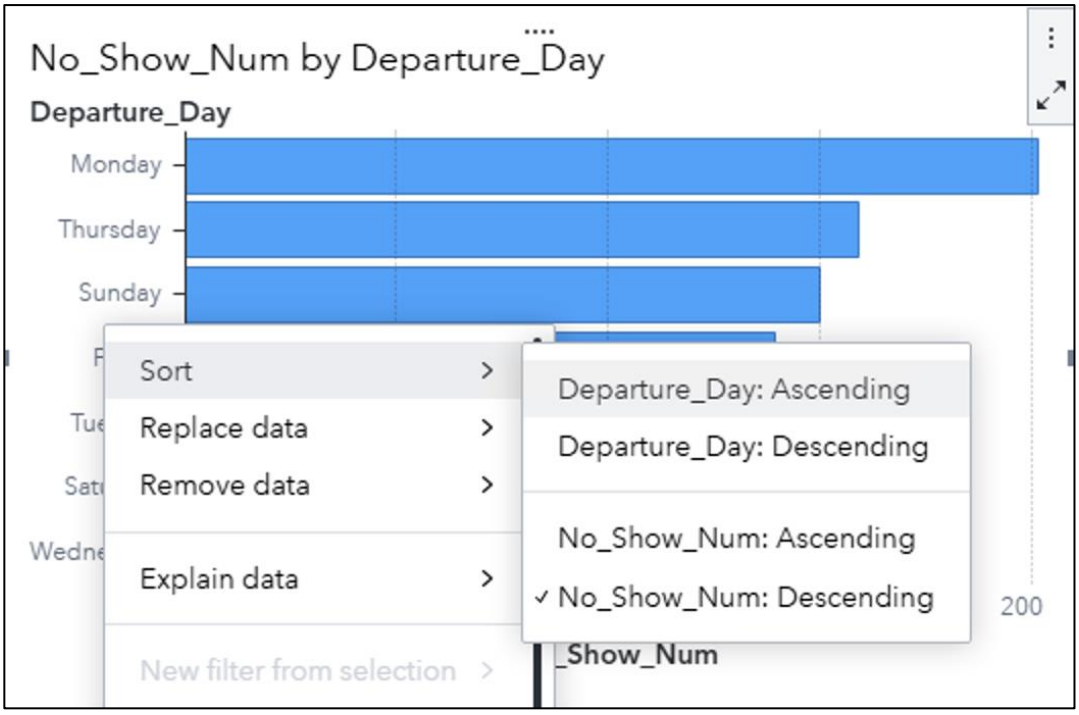

6. Seeing double? No worries - it's just a quick switch in the **Roles** pane on the left sidebar. Click *Departure_Hour* to **Replace Data Item** with *Departure_Day*.



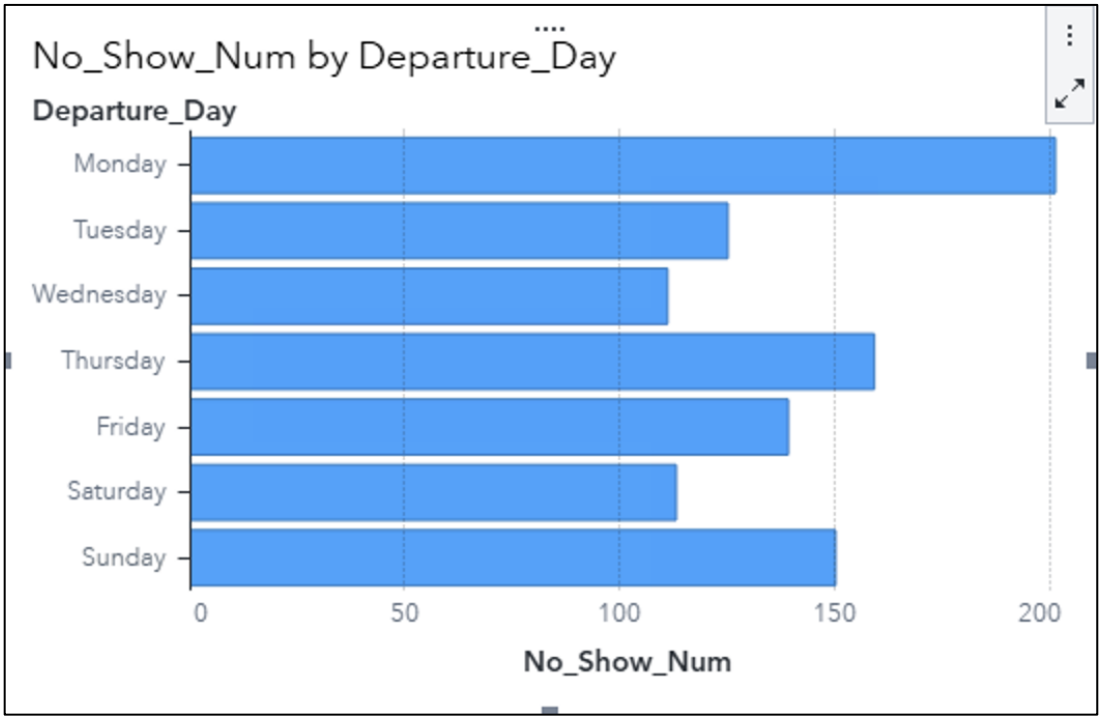

7. And voila! We can now see when no shows happen throughout the week.



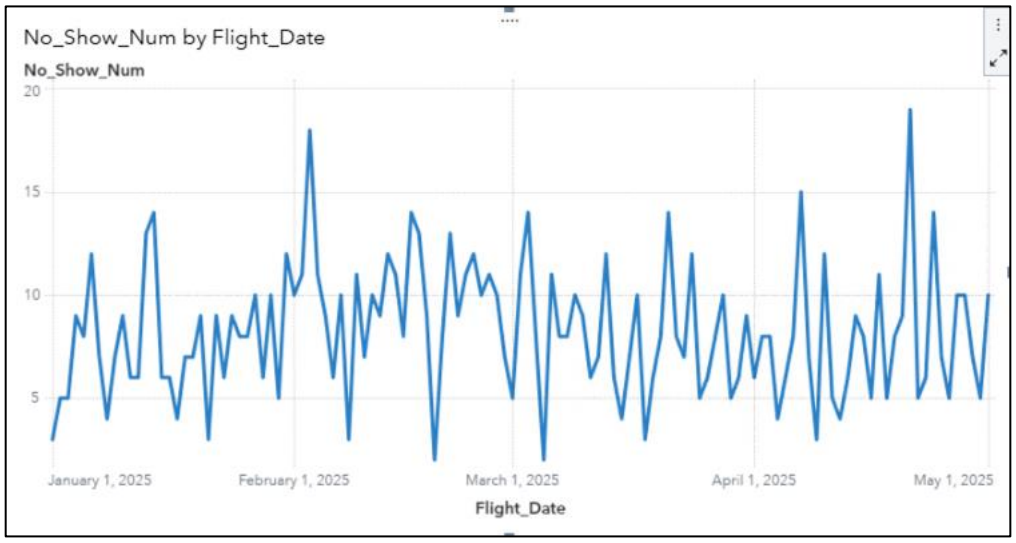

8. Looks like Mondays have the highest number of no shows, likely due to challenges associated with transitioning from weekend to weekday, such as rescheduled plans or post-weekend fatigue. You may also notice the default is to sort in descending *No_Show_Num* order. To change this, right click within the graph and select **Sort > Departure_Day: Ascending**.



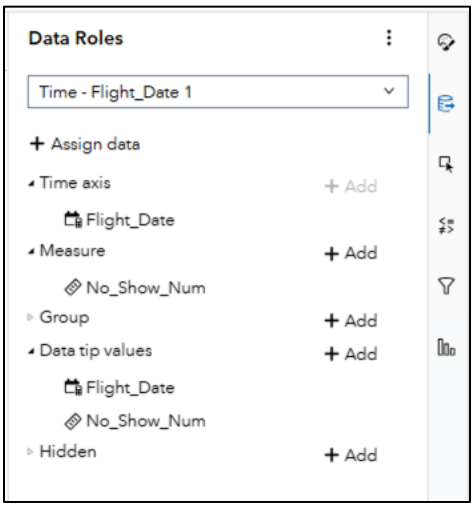
9. Much better! Our view now shows how no-shows fluctuate throughout the week.



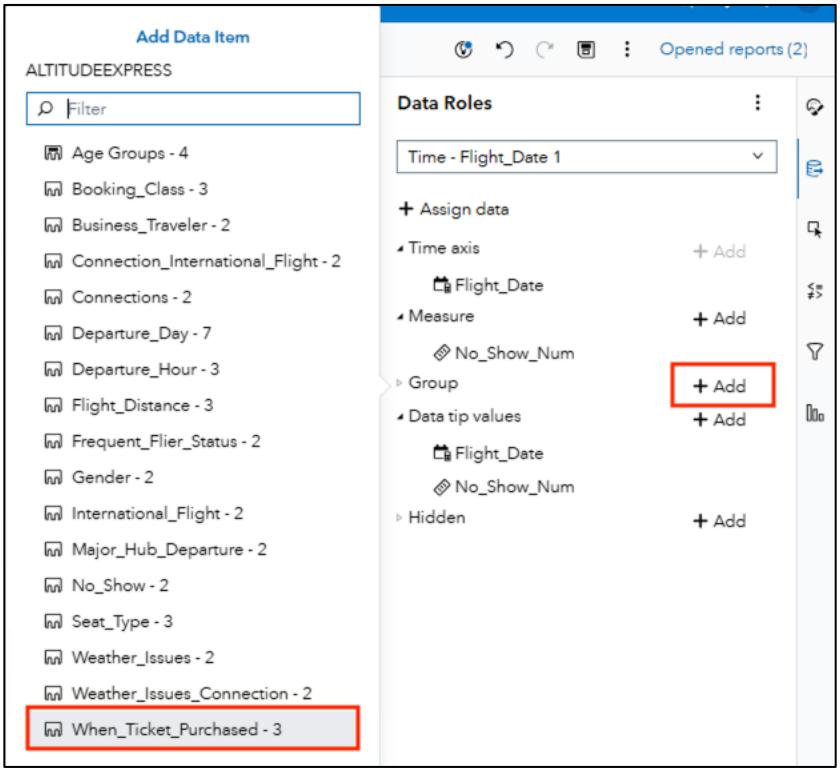

10. Let's check if there's any patterns with no shows over time. To view this, simply drag *Flight_Date* and *No_Show_Num* onto the canvas. You can place it to right, left, above, or below the bar chart. A blue rectangle will appear to indicate where the object will be dropped.



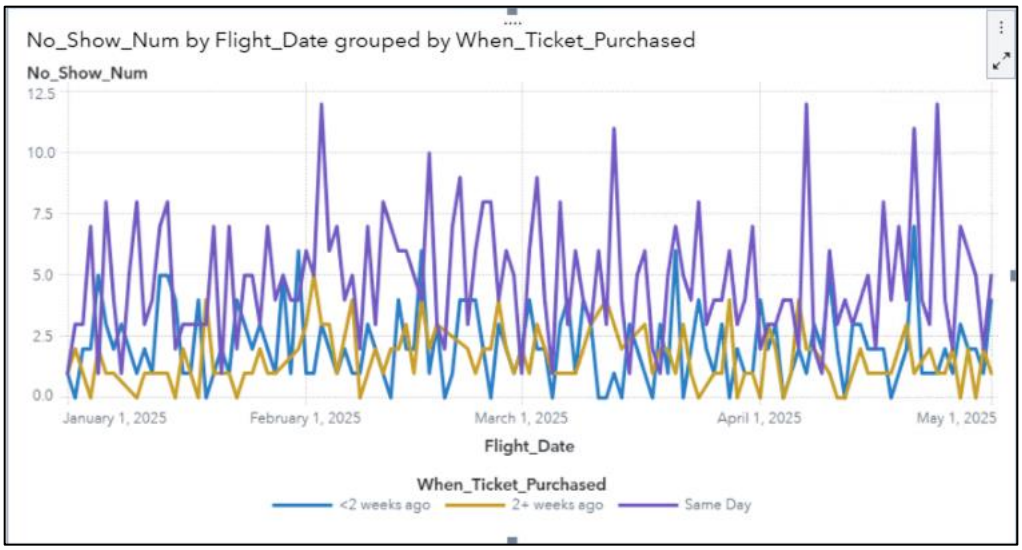

11. There are some peaks and valleys; however, nothing is jumping out. To further investigate, expand the **Data Roles** pane on the right side:



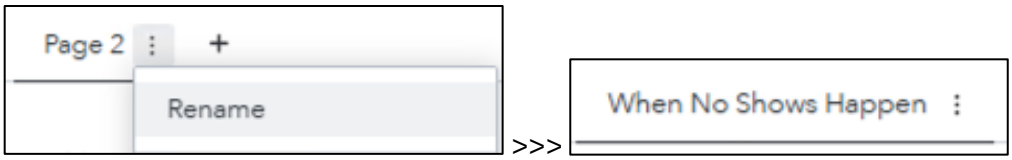

12. There are many different roles for a Time Series plot. Let's take advantage of the grouping option by selecting **+ Add** next to the **Group** role, then select *When_Ticket_Purchased*.



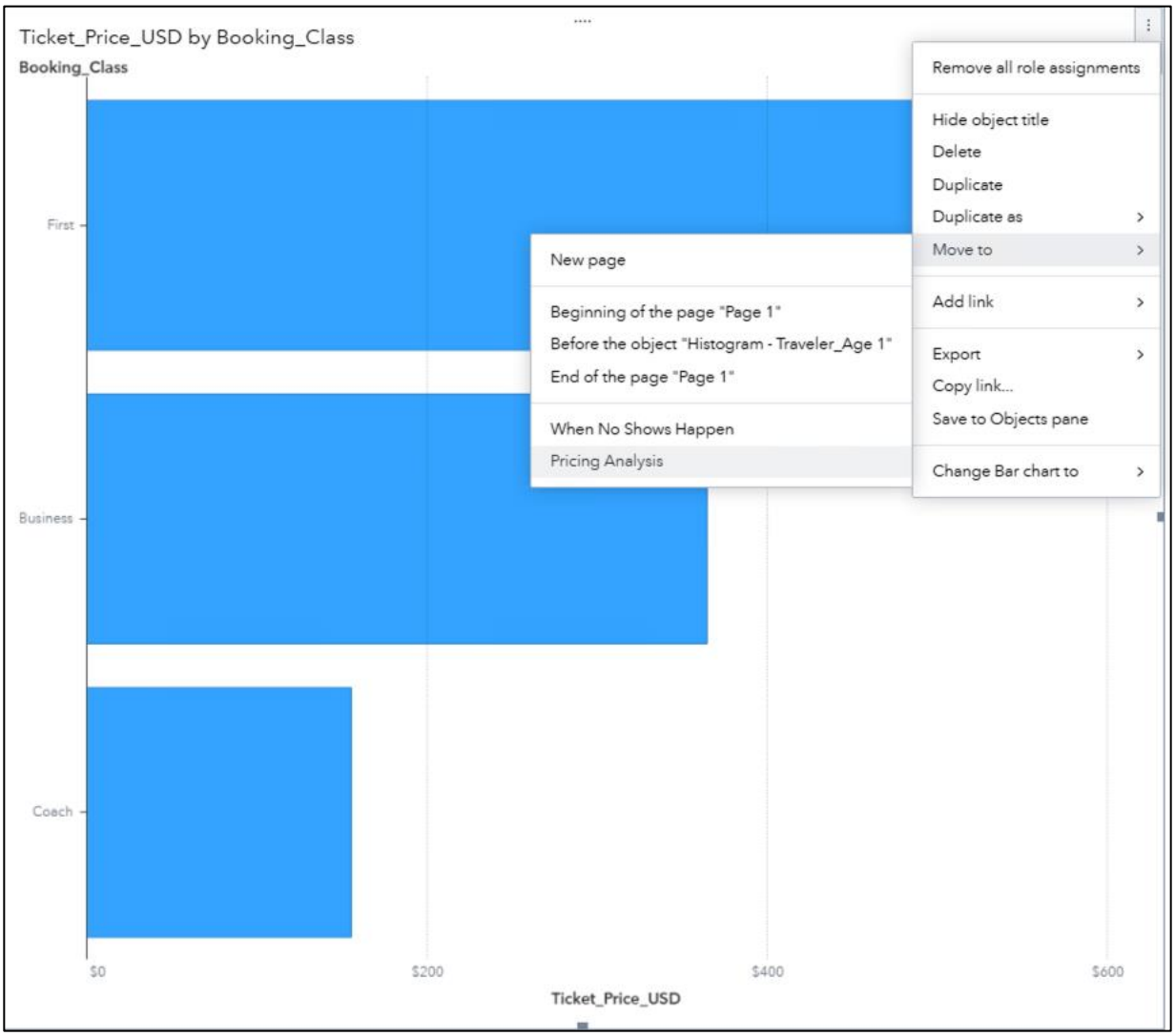

13. Your time series plot now is color coded and divided into each ticket purchase category.



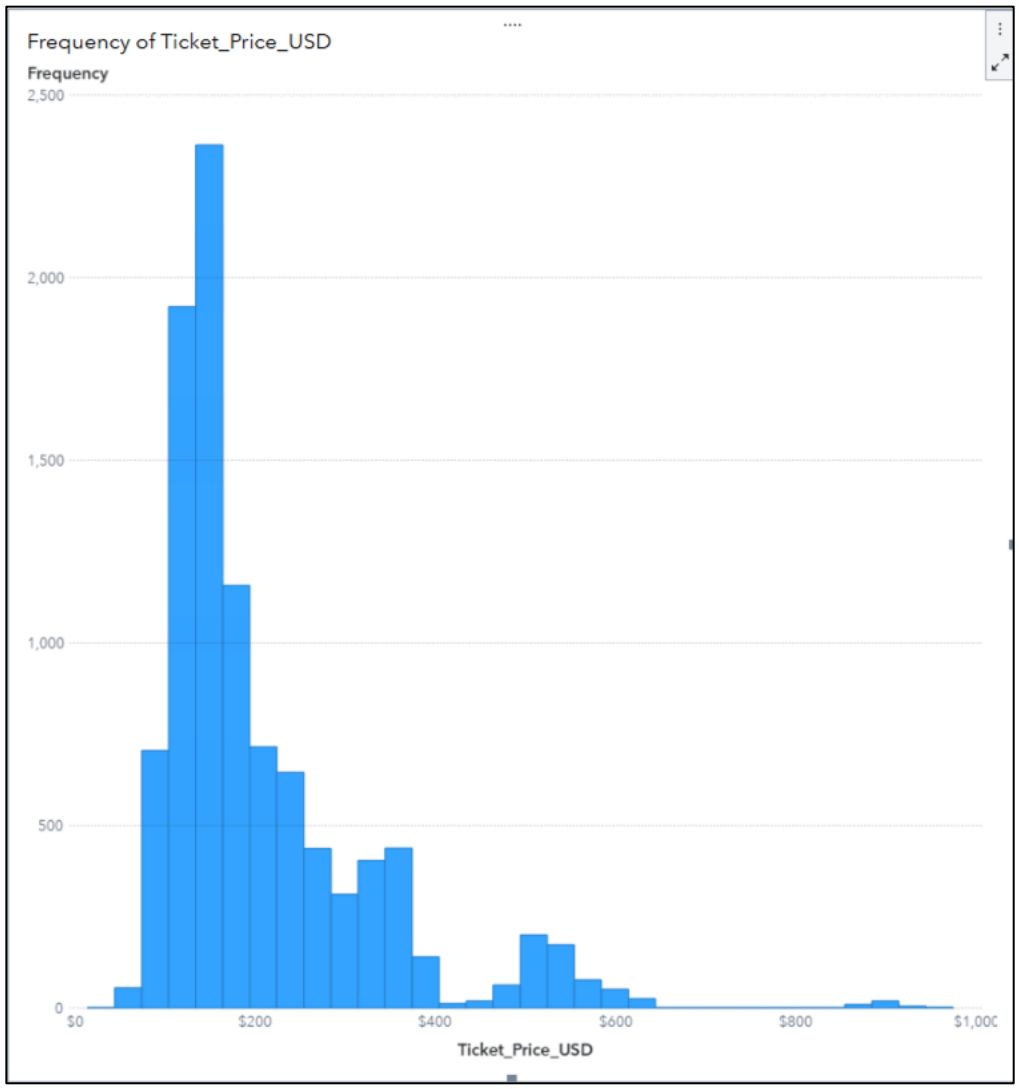

14. Wow! The most no shows happen amongst those who purchase their ticket same day!
15. This page is looking good! A more descriptive tab will help an external audience know what this page is all about before clicking on it. So, rename your current **Page 2** by clicking the **Options** menu (three dots) and selecting **Rename**. Give your page a helpful name, such as "When No Shows Happen"



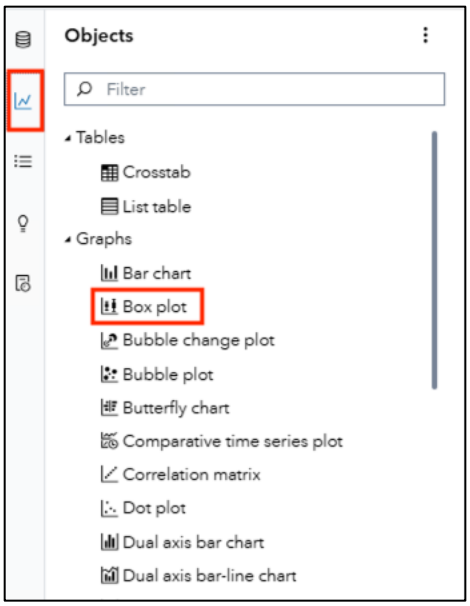

16. While you're at it, add another page and rename that to "Pricing Analysis" so we can dive into where we can make some $$$!
17. You already created a bar chart with details about ticket prices and booking class, so let's add that to your page.
18. Go to **Page 1** and click the options button in the upper right corner of the *Ticket_Price_USD by Booking_Class* chart. Hover over **Move to** and select your new **Pricing Analysis** page.



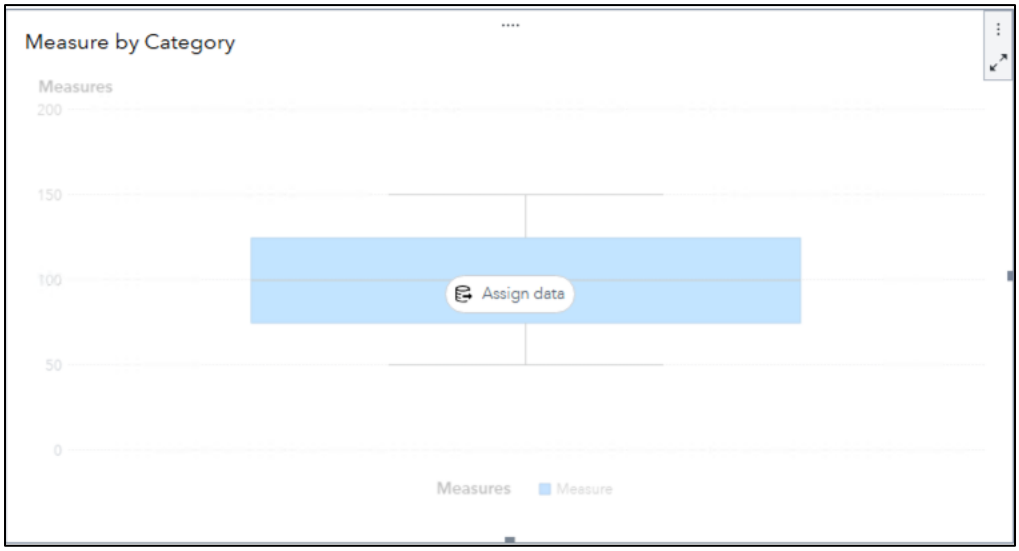

19. So, we've seen the price breakdown by booking class – but what about the overall ticket pricing? Drag *Ticket_Price_USD* onto the *Pricing Analysis* page.



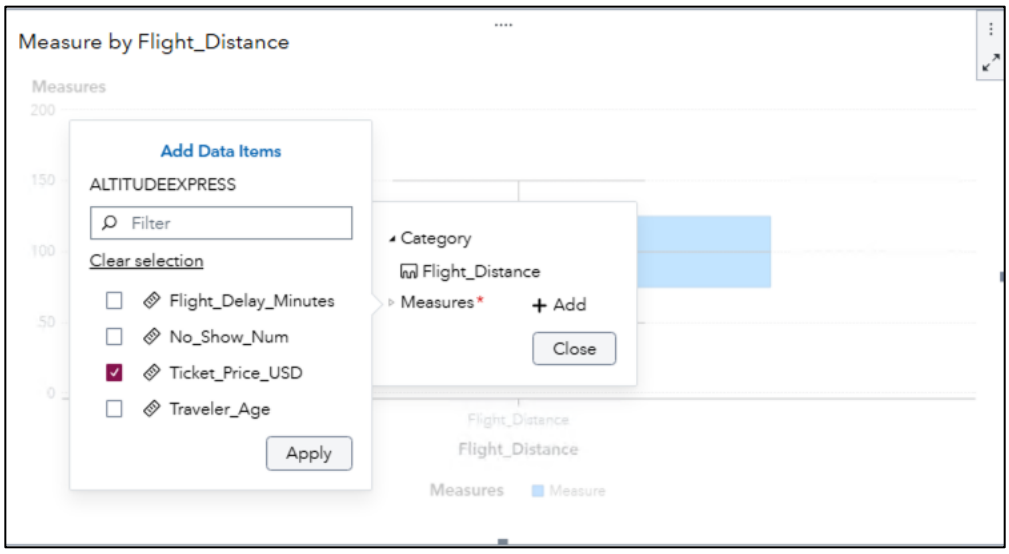

20. From this histogram, we see that ticket prices are skewed right, with the majority of tickets being sold for under $400. There is a grouping of tickets priced between $400-$600 and only a handful of tickets priced around $900.
21. Another chart that's great at showing variability is a box plot!
22. On the left side bar, expand the **Objects** pane and double click the **Box plot** object to add it to your page.



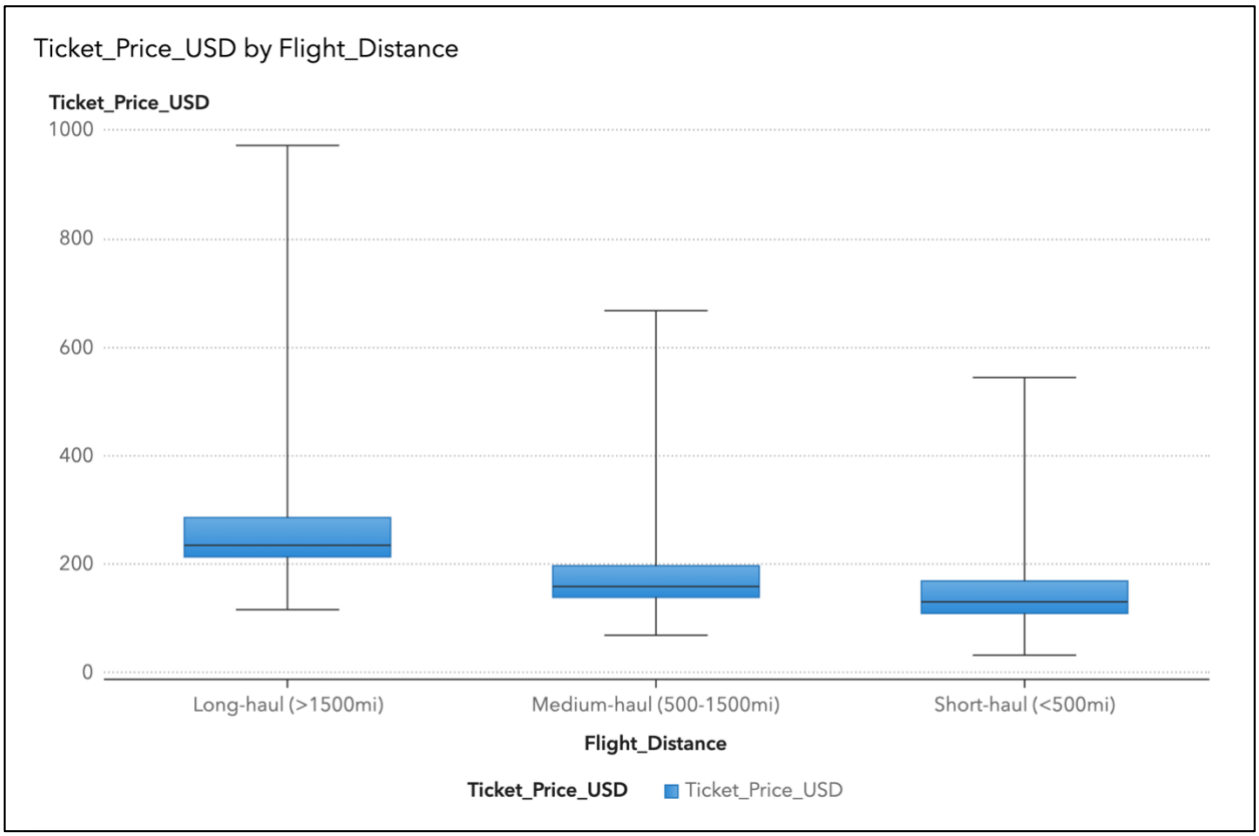

23. Yay, we have a box plot – but it doesn't have any variables associated yet.




24. Click **Assign data** and add *Flight_Distance* to the **Category** role and *Ticket_Price_USD* to the **Measures** role.

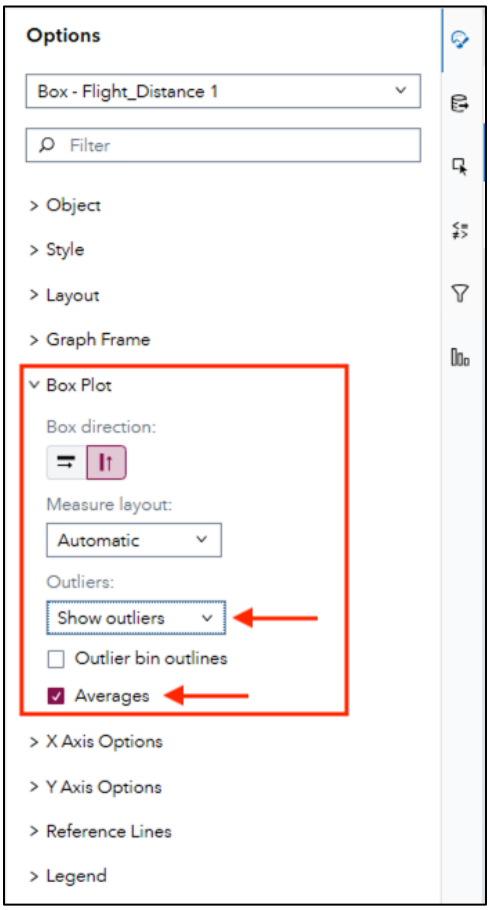



25. Our box plot shows the three categories of flight distance with the minimum, 1st quartile, median, 3rd quartile, and maximum plotted, by default. As we might expect, longer flights cost more, while shorter flights are cheaper.

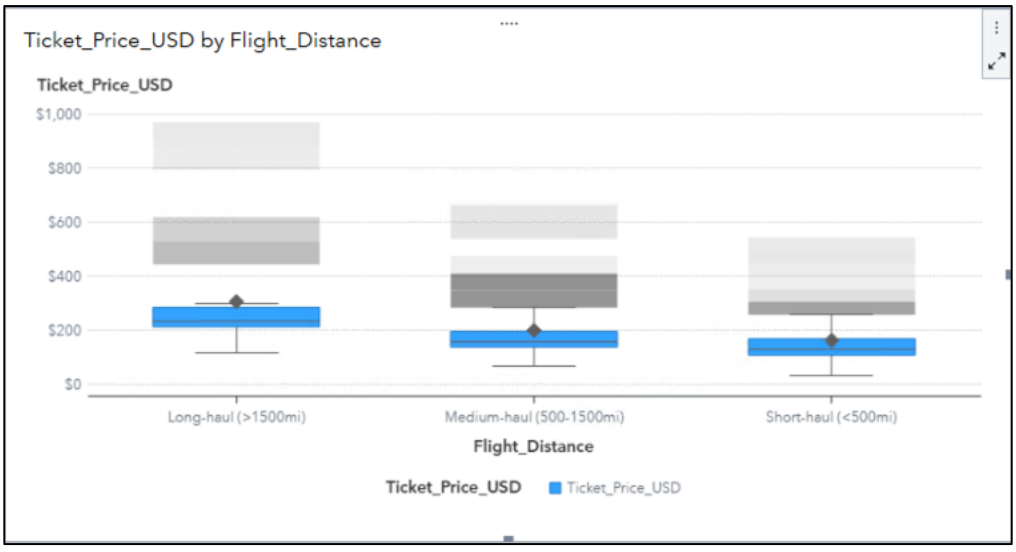



26. We could benefit from a few additional options to highlight the variability. Expand the **Options** pane on the right-side bar. Under the **Box Plot** section, make the following adjustments:

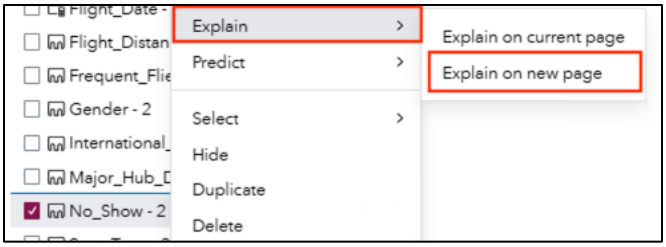
a. Outliers: Show outliers
b. Averages checked



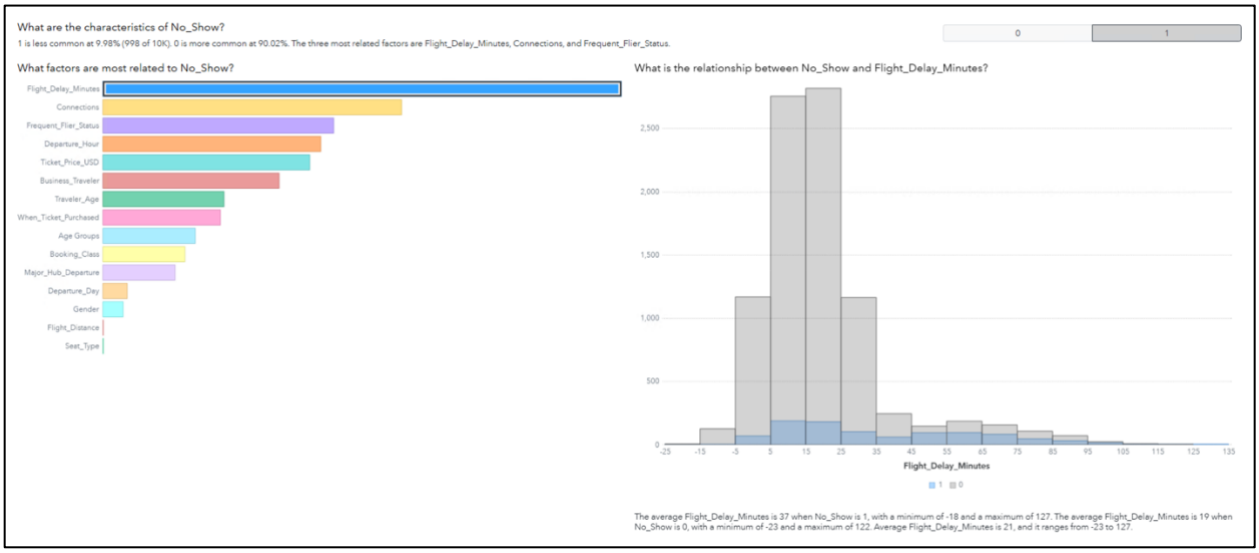

27. Averages will be denoted by a diamond and outliers will be shown using shaded bins.



28. We can now see that all averages lie above the median, sometimes even above the 3rd quartile mark, and there are several outliers for all flight distances. The box and whiskers are relatively small, indicating that without the outliers, the prices don't vary a lot within the flight distance groups.
29. This page is looking good! Let's look back at no shows – but, instead, let's focus on the category version of the variable now.
30. Right click the categorical *No_Show* variable then hover over **Explain** to select **Explain on a new page**.



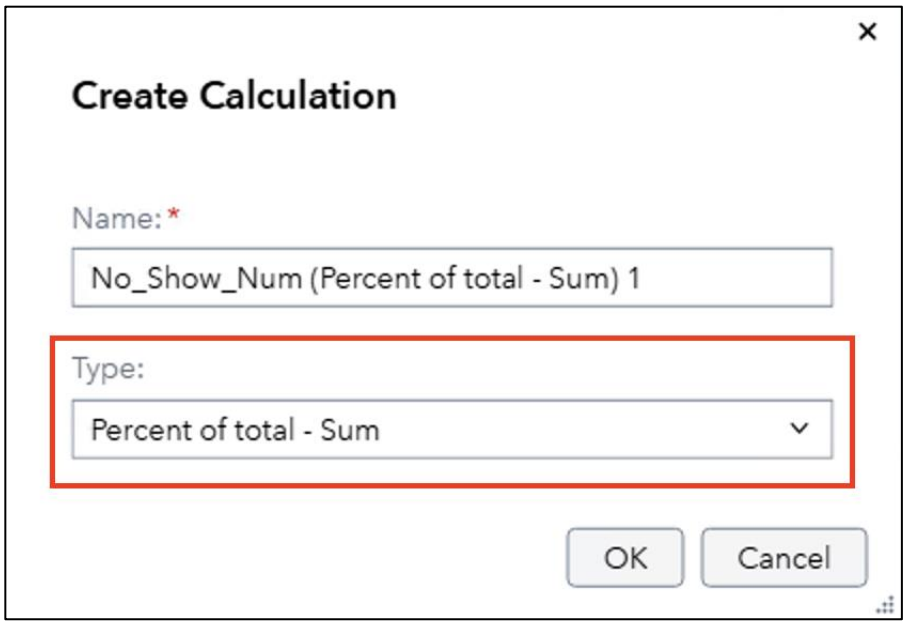

31. A new page is created with an **Automated Explanation** object, which determines the most important underlying factors for a specific response variable, *No_Show* in our case.



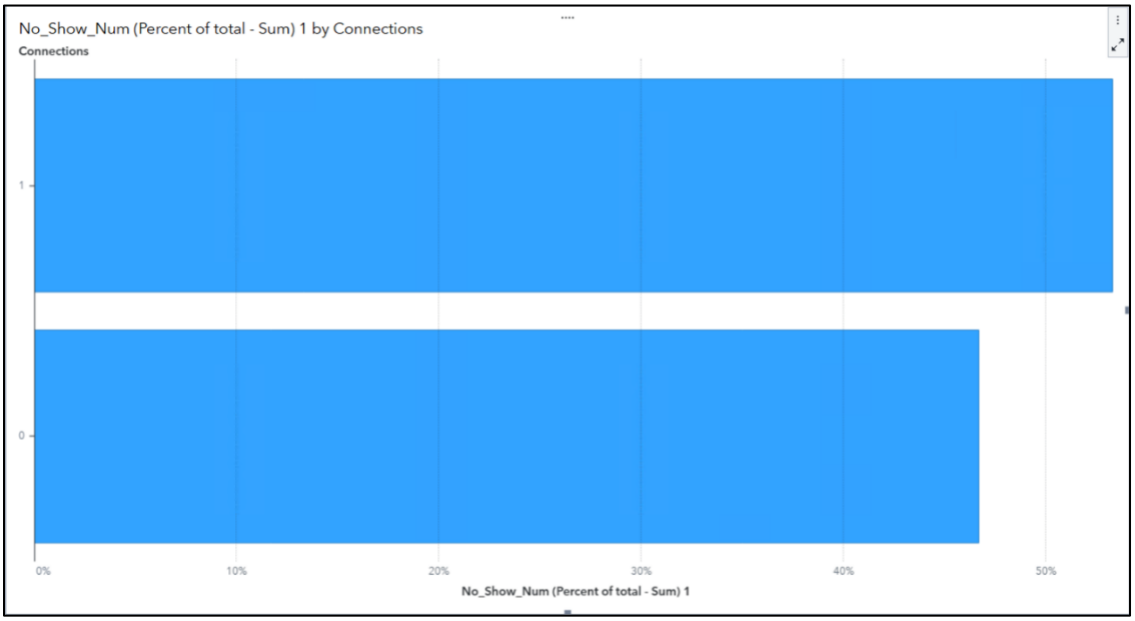

32. Rename this page and click on different factors to see the relationship with *No_Show*. Note that our Event level is 1 since we want to analyze the factors that caused the flier to "no-show."


33. For instance, clicking on the *Connections* factor shows that although there are significantly fewer cases where a flier has a connection, it appears there is an almost equal number of no-shows in the two groups… could this be right?!
34. To properly calculate the percentage of no shows in each group, we can easily create a new calculated variable!
35. Right-click *No_Show_Num* from the Data pane and select **New calculation**.

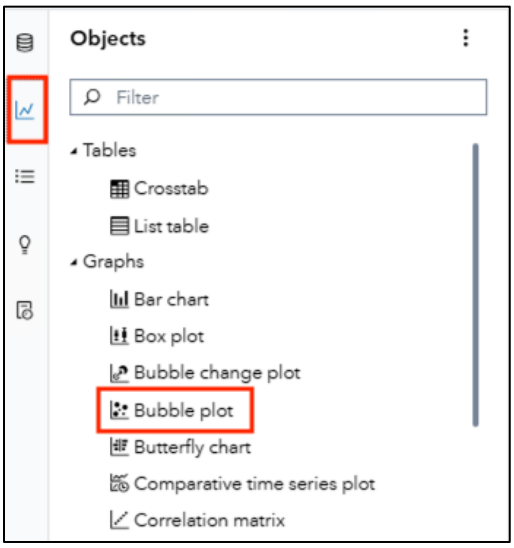

36. Use the **Type** drop down to select **Percent of total – Sum**. Feel free to rename the variable if your heart so desires!




37. Now to see the breakdown! Create a new page called “How No Show Factors Affect Profit.”
38. Drag your newly created *No_Show_Num (Percent of total – Sum)* variable – or whatever you named it – and the *Connections* variable onto the new page.

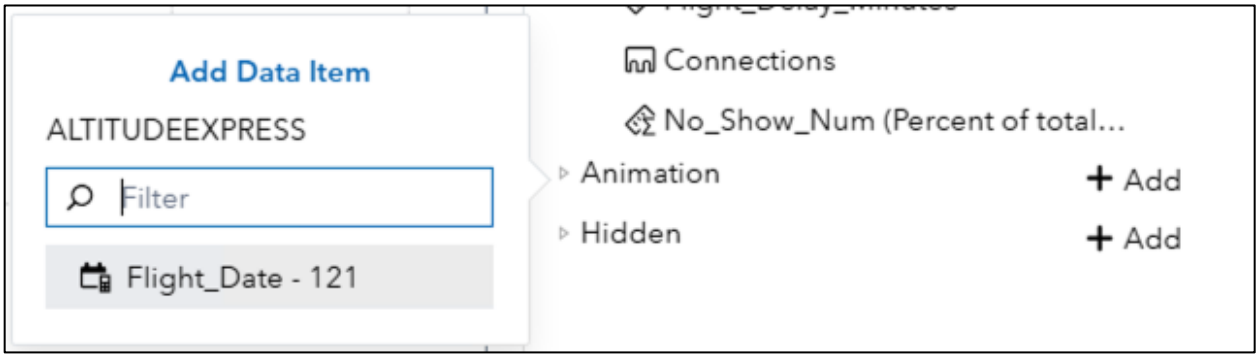



39. Woah! There’s a higher percentage of people who didn’t show up for their flight when they had a connection.
40. Let’s utilize a new type of object, the bubble plot, which is particularly effective for visualizing relationships between multiple variables by using several dimensions of data.
41. From the **Objects** pane, drag the **Bubble plot** object onto the page.

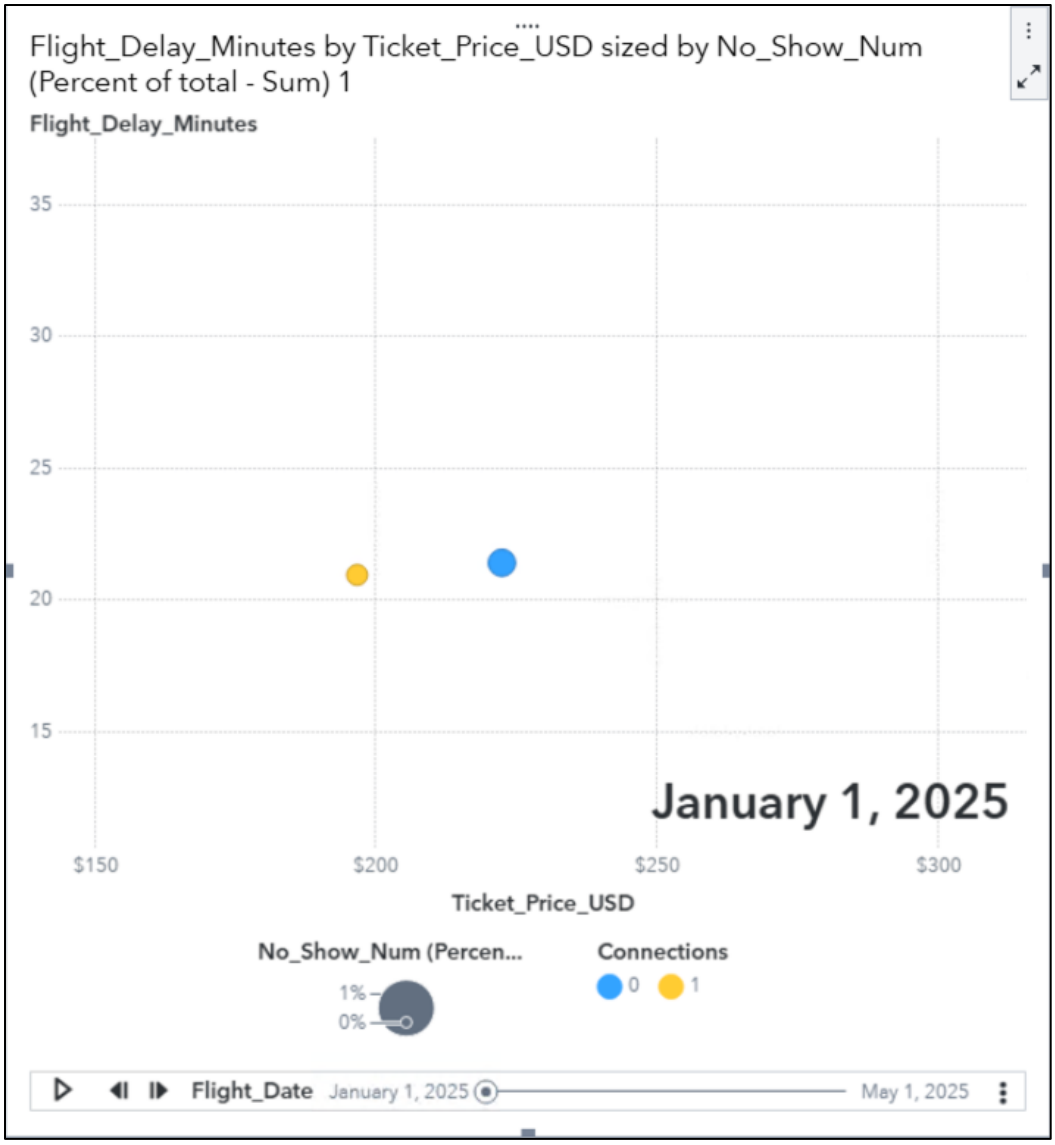



42. Populate the following variables into the roles.
    a. X axis: *Ticket_Price_USD*
    b. Y axis: *Flight_Delay_Minutes*
    c. Size: *No_Show_Num (Percent of total – Sum)*
    d. Group: *Connections*

43. As we've learned, the no show percentage is larger when a flier has a connection, but also, the flight delay is longer, and the price is higher. Pay attention to the scale, though, because these differences aren't as large as they appear.
44. One more role we might want to incorporate to our analysis is the **Animation**. We can use this to see how our bubbles move and change over time. Add *Flight_Date* to the **Animation** role.



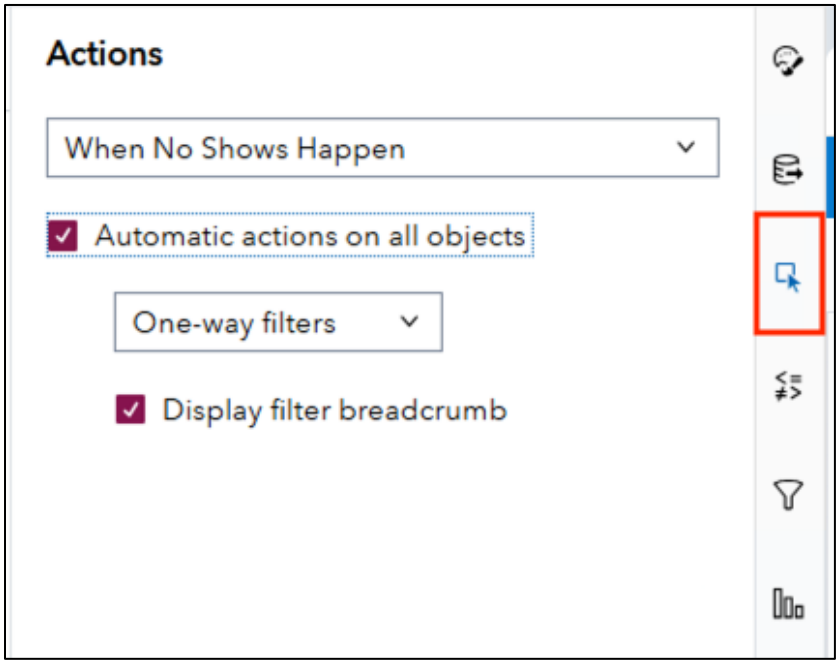

45. The bubble plot is updated to show the evolving, day-by-day relationship between our variables displayed. You can either click the play button to see the bubbles move through time, use the forward and back arrows to step through, or use the slider to navigate to a particular day.



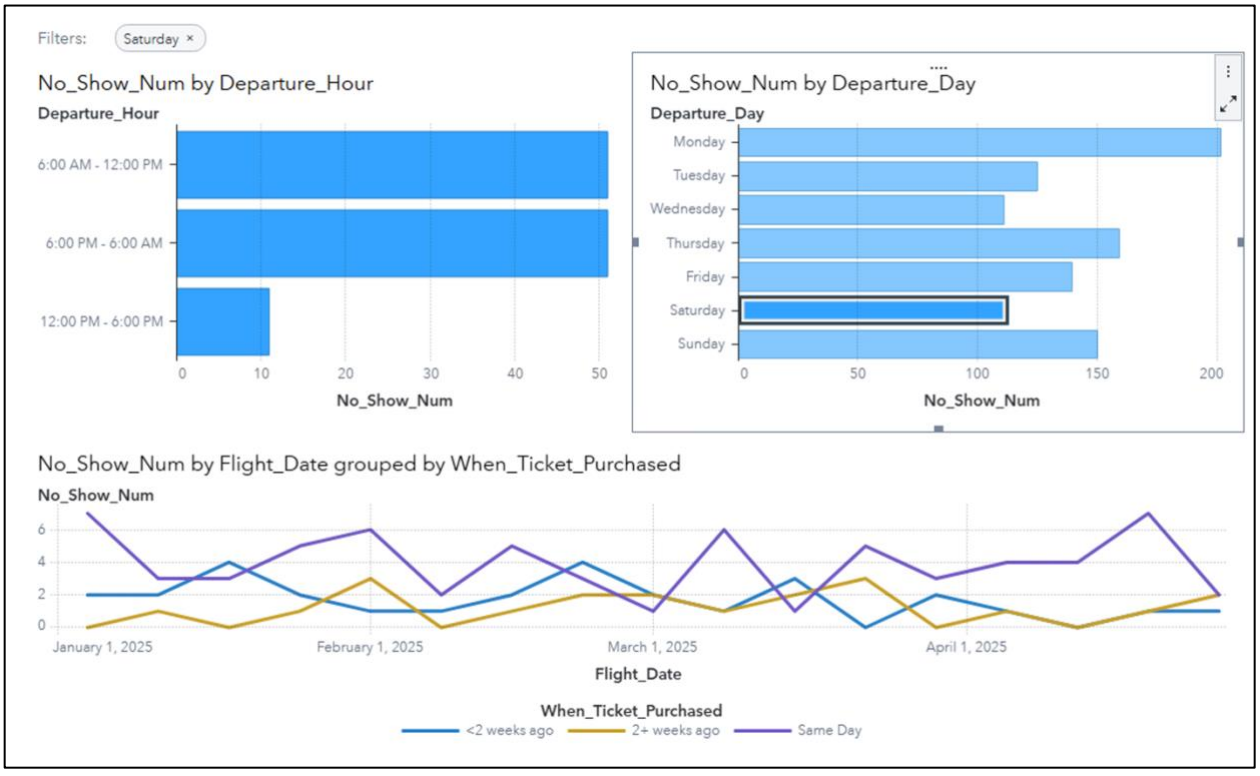

46. At this point, we've built out a lot of great visualizations! There's always more that can be done to polish your dashboard, such as standardizing colors, adding chart titles, adjusting placement within a page, etc.; however, with a few small changes, we can enhance this report for our end user!

## Part 4: Love It – But Make It Interactive

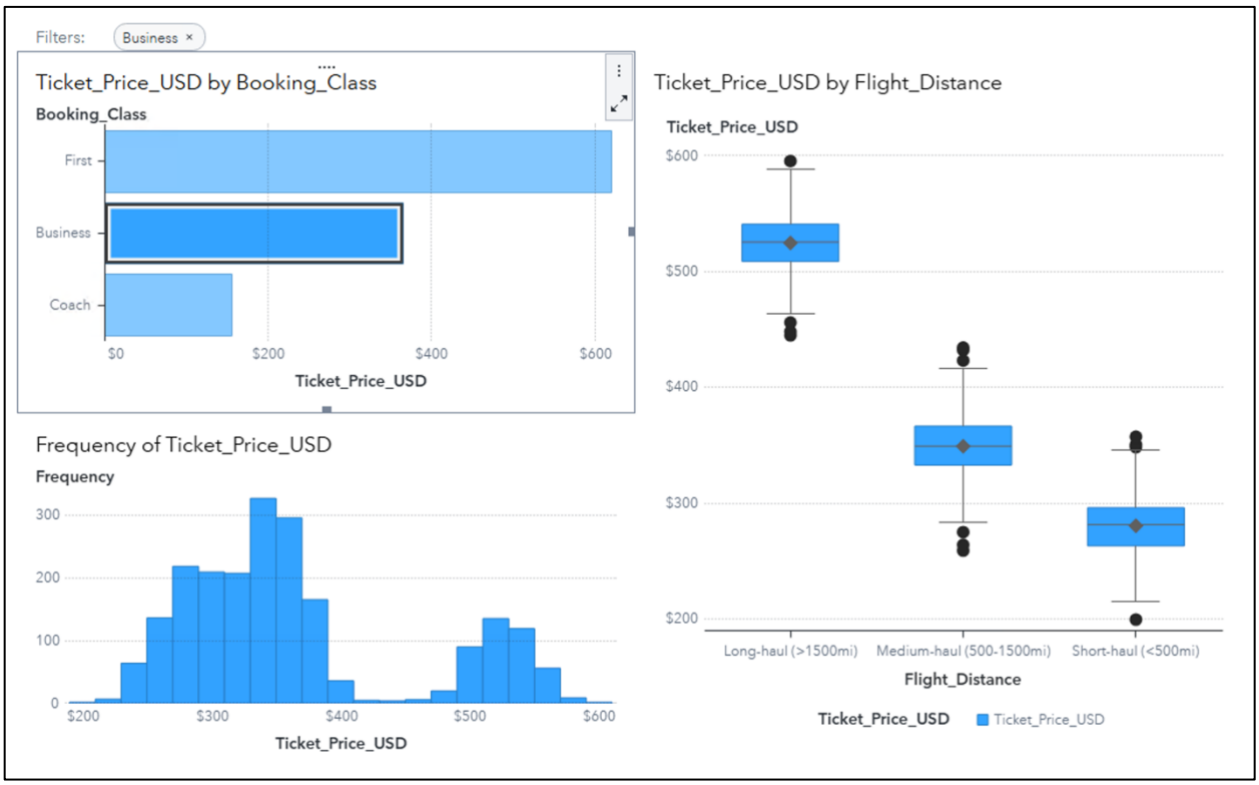
1. Head back to the "When No Shows Happen" page and expand the **Actions** pane on the right sidebar. Click the box to select **Automatic actions on all objects**.




2. With this easy click, you can now select a bar from the *No_Show_Num by Departure_Day* chart and see the other graphs on the page filter to that day of the week.



3. You could also start in the *No_Show_Num by Departure_Hour* bar chart to filter. Use this an opportunity to click around and see the interaction! Fun fact: Your end user can do the same to home in on the details they're more interested in!
4. Feelin' fancy? Make the same **Automatic actions on all objects** selection on the **Pricing Analysis** page.




5. Although your time turning data into information is coming to a close for today, there are endless possibilities to explore; the adventure of turning data into wisdom is just beginning.

# Appendix

## Appendix A: Helpful Documentation

There are numerous online resources available to deepen your understanding of creating dashboards in SAS Visual Analytics. Below are some helpful links to get you started:

- Explore and Visualize Data with SAS Visual Analytics - SAS Viya Quick Start - SAS Video Portal (14 min)
- SAS Visual Analytics Community
- Beautiful Reports

## Appendix B: Recommended Learning

The following e-learning courses are recommended to help with this activity and continue your learning journey within SAS Visual Analytics:

- SAS Visual Analytics 1 for SAS Viya: Basics
- SAS Visual Analytics 2 for SAS Viya: Advanced

You may gain access to these courses via the SAS Learning Subscription. With a free trial, learners are granted access to an extensive library of SAS eLearning courses.