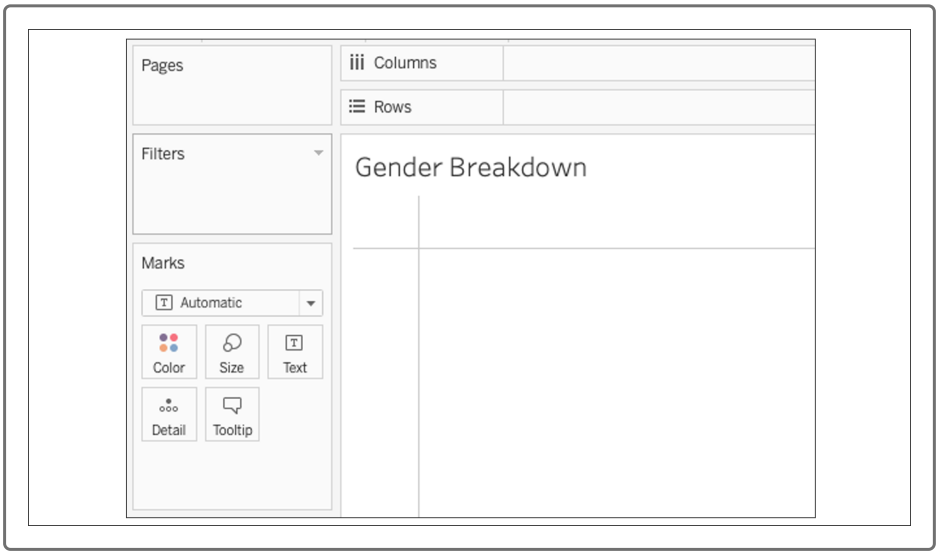
14.3.3 Find the Number of Rides by Gender

Another question we have for the data—beyond the where and when—is the who. What can the data tell us about the riders themselves? Often, the first place we start when understanding a population is gender.

We now want to know the gender breakdown of Citi Bike riders to help us learn a little more about the customers.

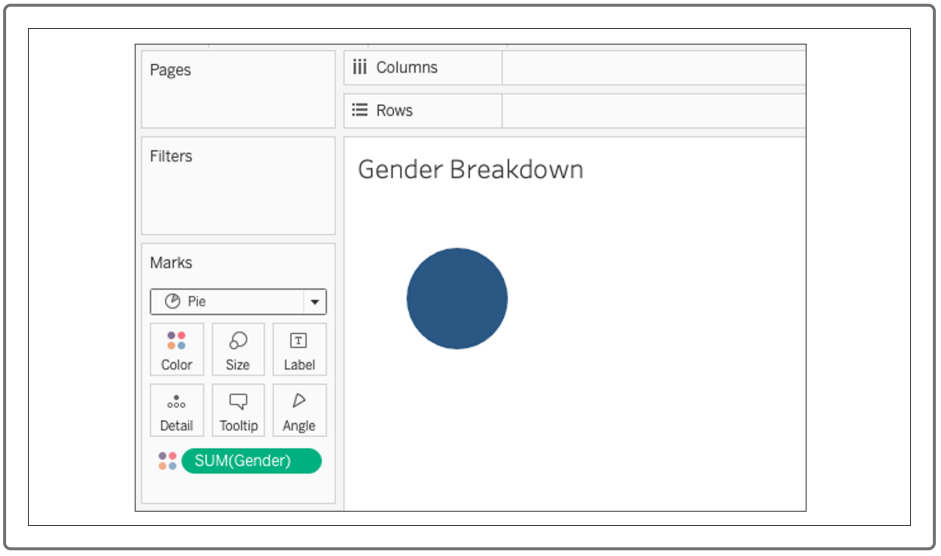
What Is the Gender Breakdown of Active Riders?

Create a new worksheet and name it "Gender Breakdown." We'll use this worksheet to create a pie chart for our data to show the gender breakdown. At this point, here's what your workspace should look like:



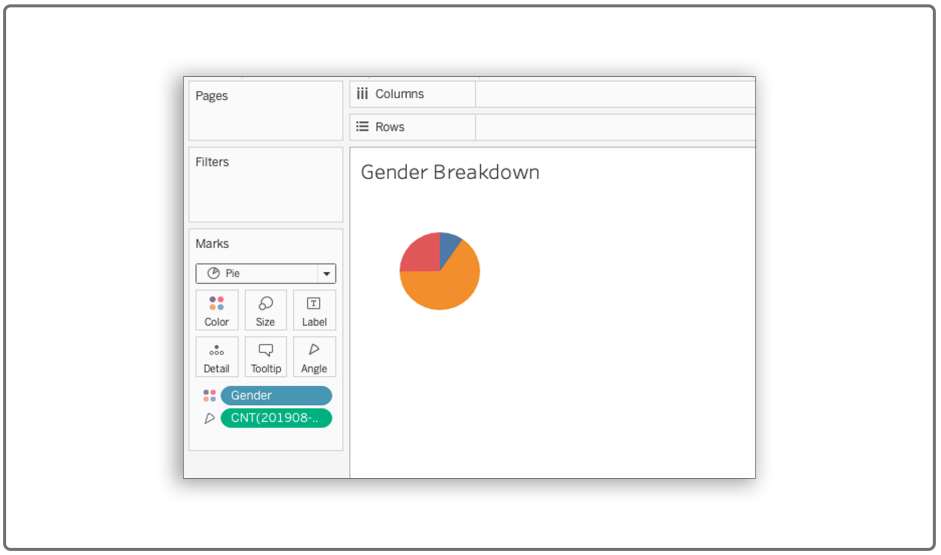
Next, identify the measures and dimensions needed for this worksheet: Gender and `201908-citibike-tripdata.csv (Count)` measures.

Drag the Gender measure to the Color mark. Your worksheet should look like the following:



In the Marks section, you'll notice Gender is within a "SUM" function. We'll want to change this to a dimension, as this will allow it to be the sum of all gender rows we have in our data. Place your cursor over "SUM(Gender)." Click the arrow and select the dimension button.

Next, drag the `201908-citibike-tripdata.csv (Count)` measure to the Angle mark. Your worksheet will be updated to look like the following:

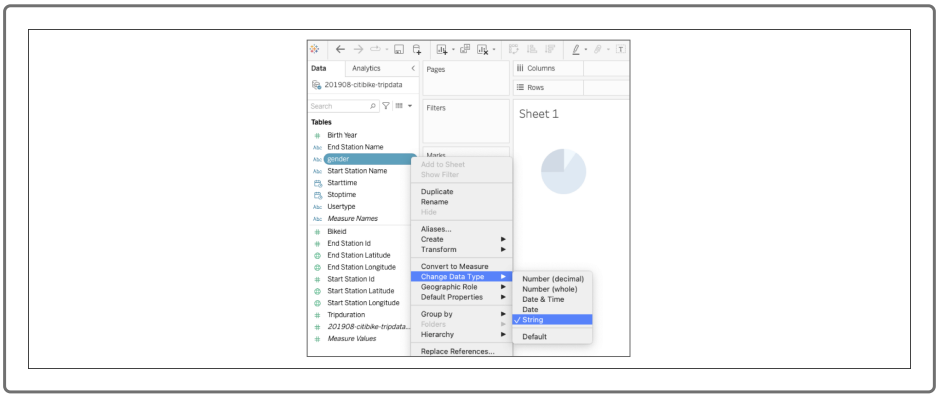


If you place your cursor over each piece of the pie, you will see 0, 1, and 2. If you go back to where we downloaded our data, Citi Bike tells us that 0 represents "Unknown," 1 represents "Male," and 2 represents "Female." But remember, even though we have this information, the audience viewing our data likely will not. Therefore, we need a calculated field.

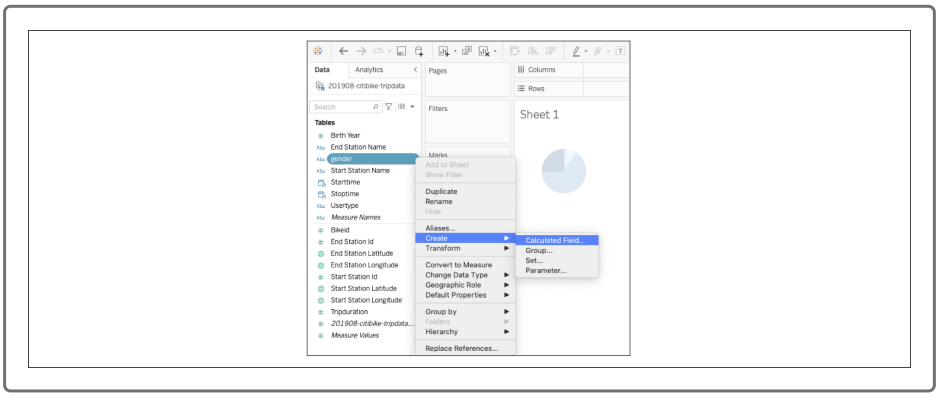
Create a Calculated Field

A calculated field essentially allows us to write code to customize our data. In this case, we'll customize so that the pie chart has a tooltip that displays Unknown, Male, and Female.

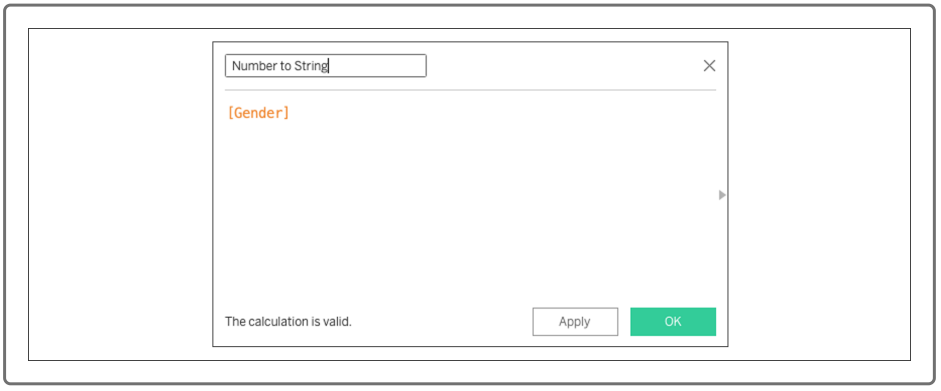
To create a calculated field, locate the Gender dimension in the Data side pane. Click the arrow in the Gender dimension, choose the "Change Data Type" option, then select the "String" option. Note that we need to complete this step to ensure the calculated field works correctly:



With the data type changed, we can start looking at the calculated field. Click the arrow in the Gender dimension again. This time choose the "Create" option, then select the "Calculated Field" option, as shown in the following image:



The image below shows the window that will pop up after choosing to create a calculated field. Change the name of the calculated field to "Number to String."



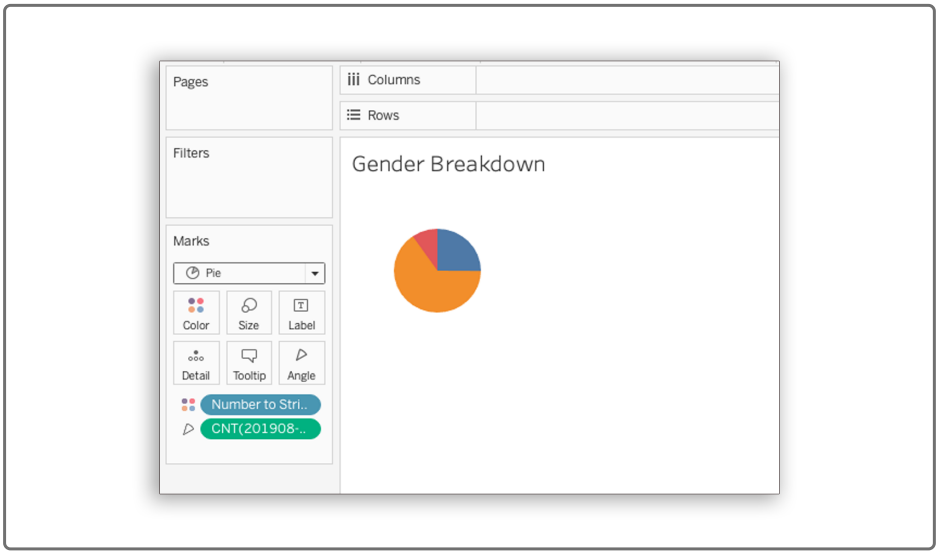
Remember that we want to convert all numbers in the Gender dimension to the string version of unknown, male, or female.

SKILL DRILL

Use your coding and Google search skills to convert the numbers in the Gender dimension to unknown, male, or female.

Once this code has been entered, click OK. Your pie chart will be updated so that you can see the number for each gender.

Now return to your worksheet and drag the "Number to String" dimension to the Color mark. This is what your worksheet will look like now.



Great work! Remember that you can place your cursor over each of the pie slices to see which gender tends to use bike sharing the most.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.