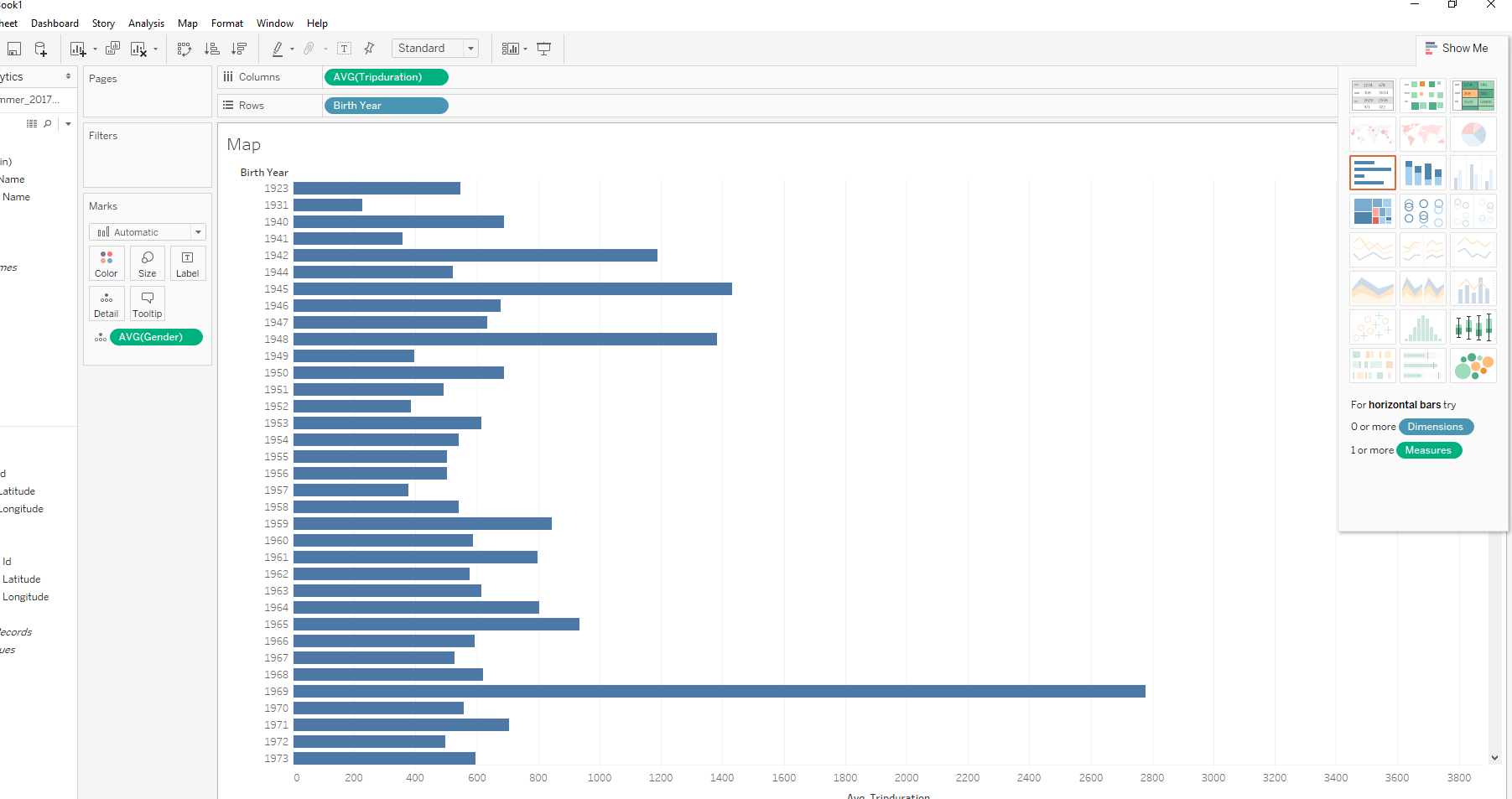
Interesting factors to analyze…

Age, gender, location…

**Main Data Points of Study… Growth opportunities... Money Makers…**

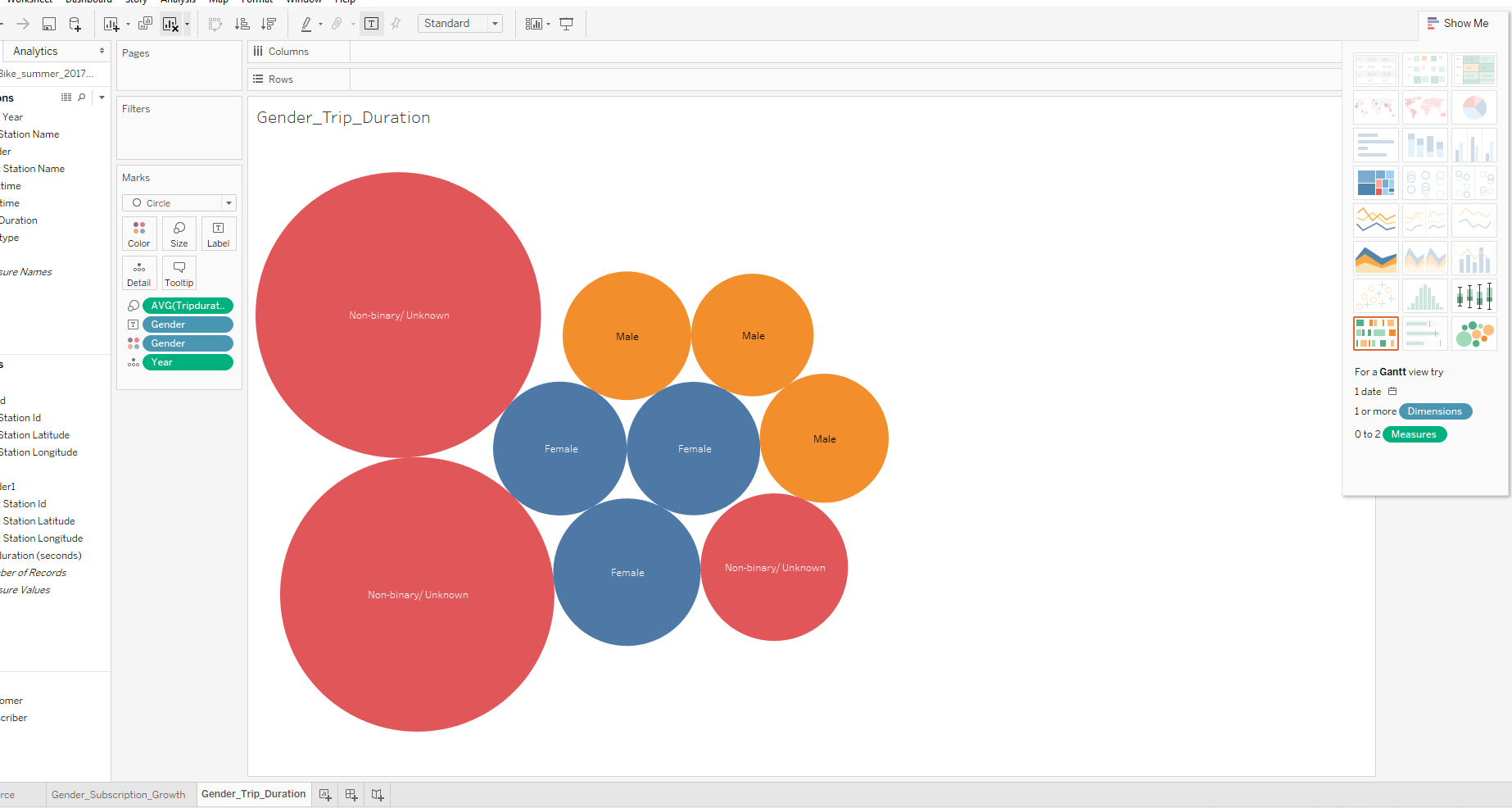
**Age**- What is the heaviest age concentration of riders and what is the lowest age concentration? People between 20 and 40 do short runs and people under 20 or so do shorter runs.

-Average trip duration by age

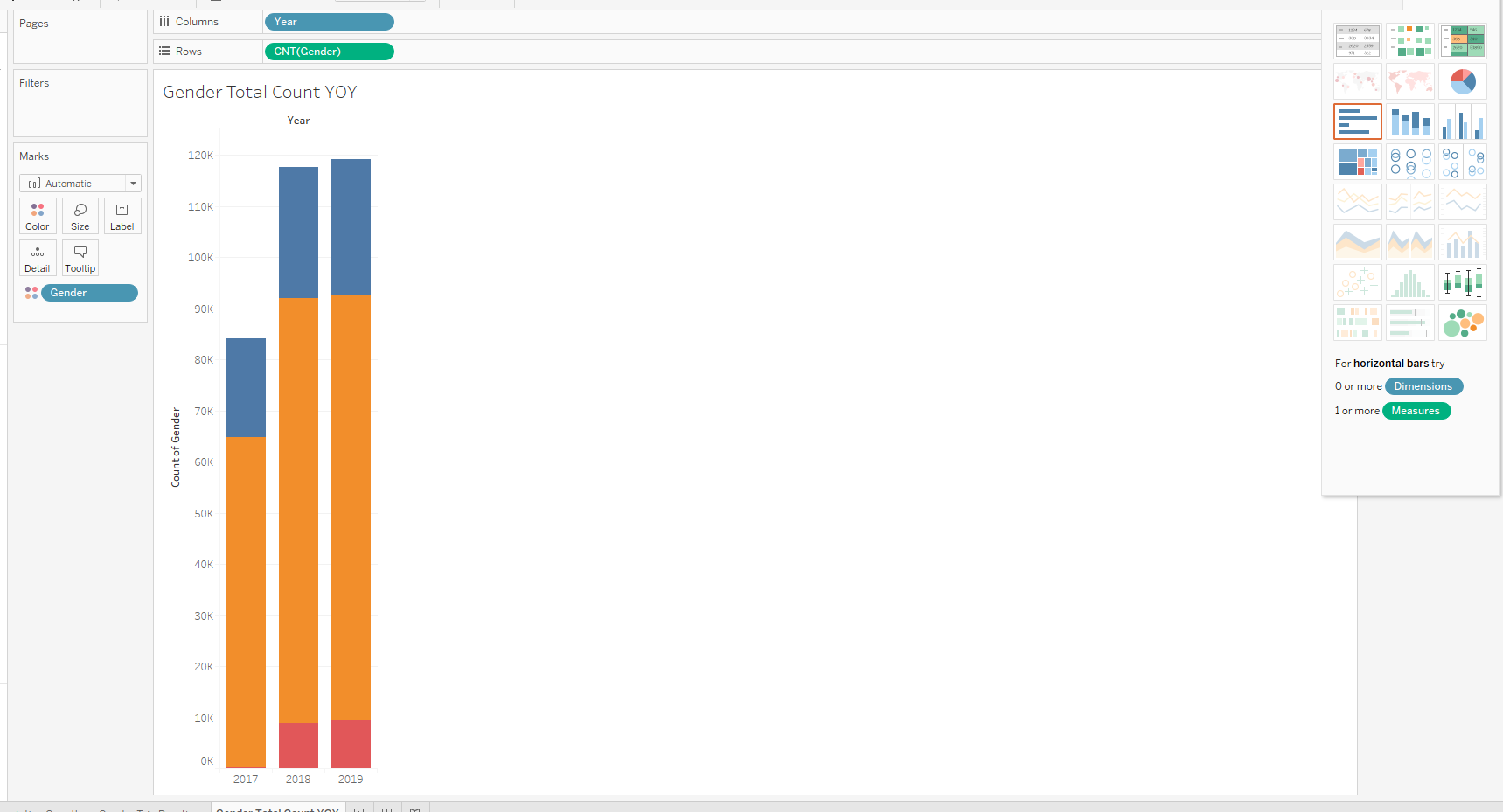


**Gender**- What is the gender break down for usage? Is there any growth in any particular area (i.e. 0 gender…)?

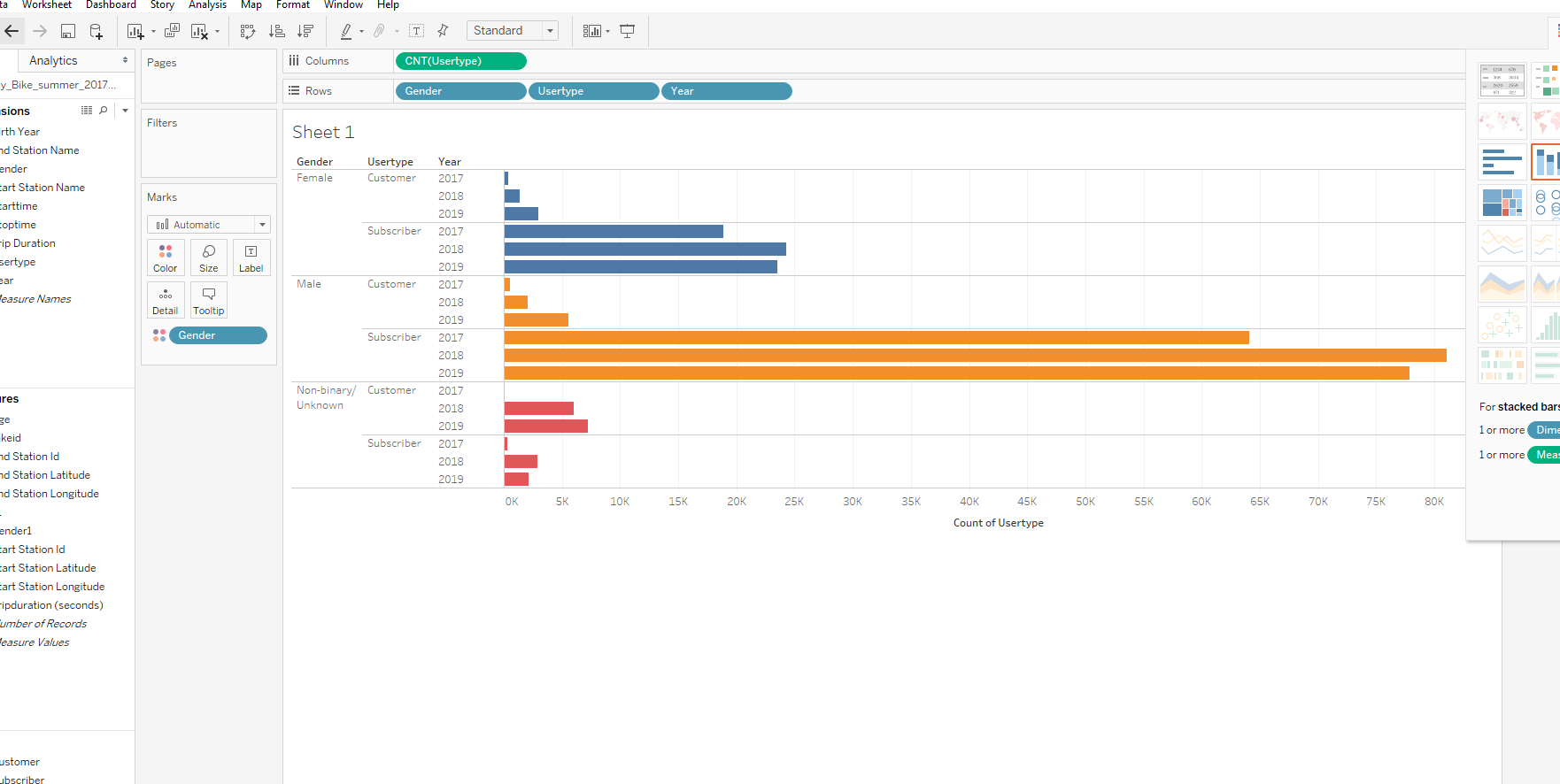
-Average trip duration by gender



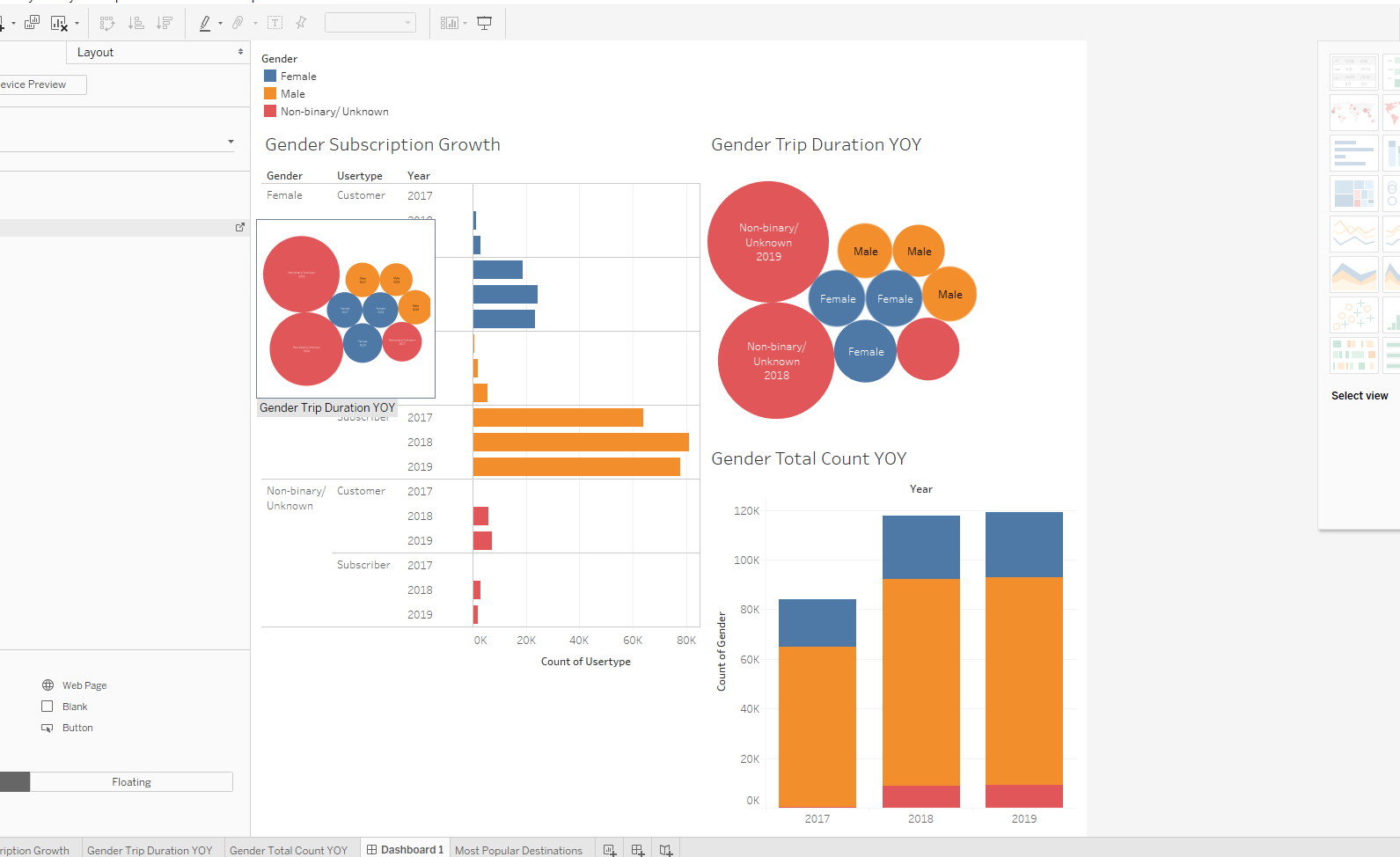
-Distribution of genders by year (YOY Gender Growth)



-Break down of genders by subscriber growth

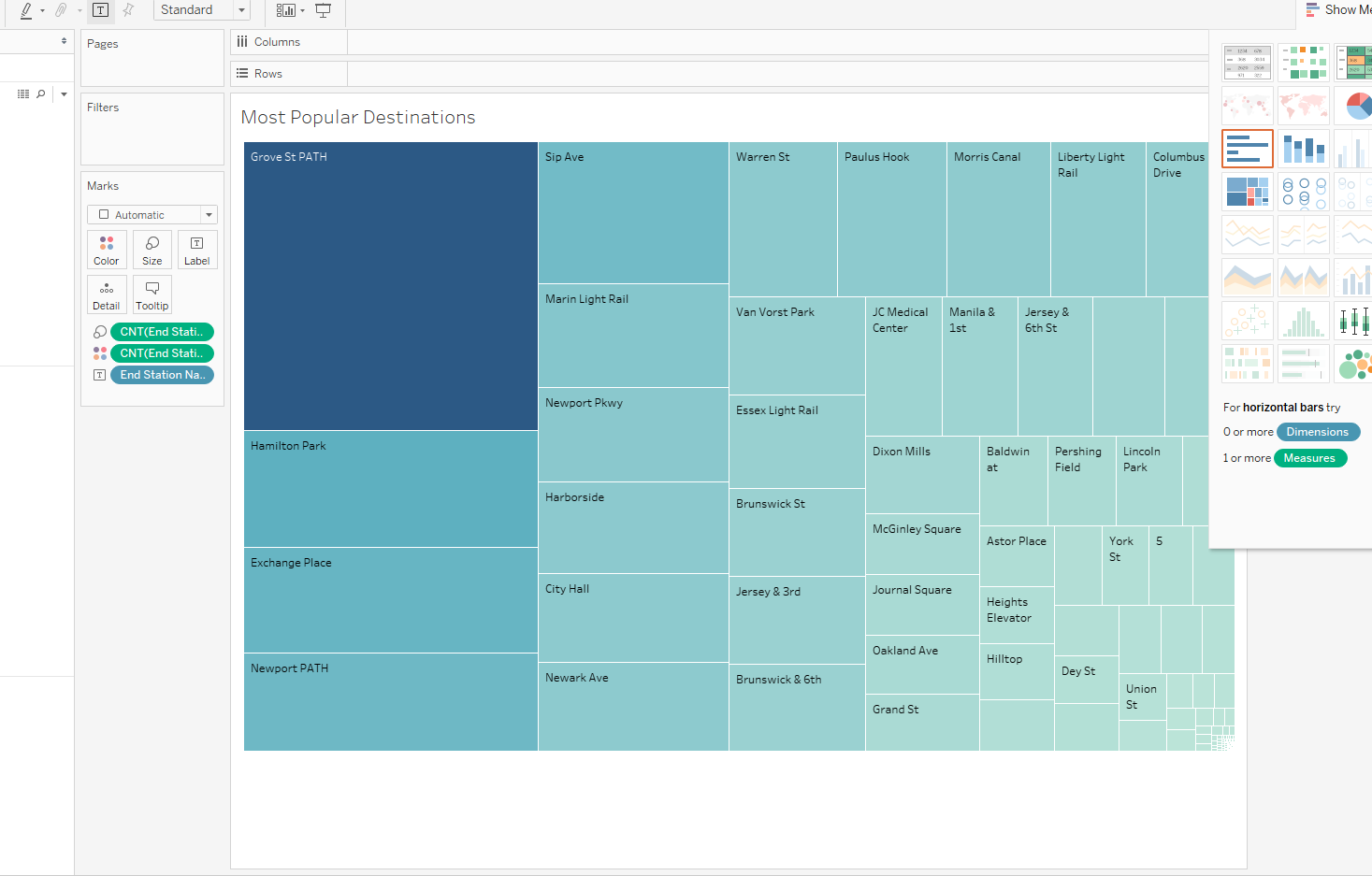


-Gender Dashboard

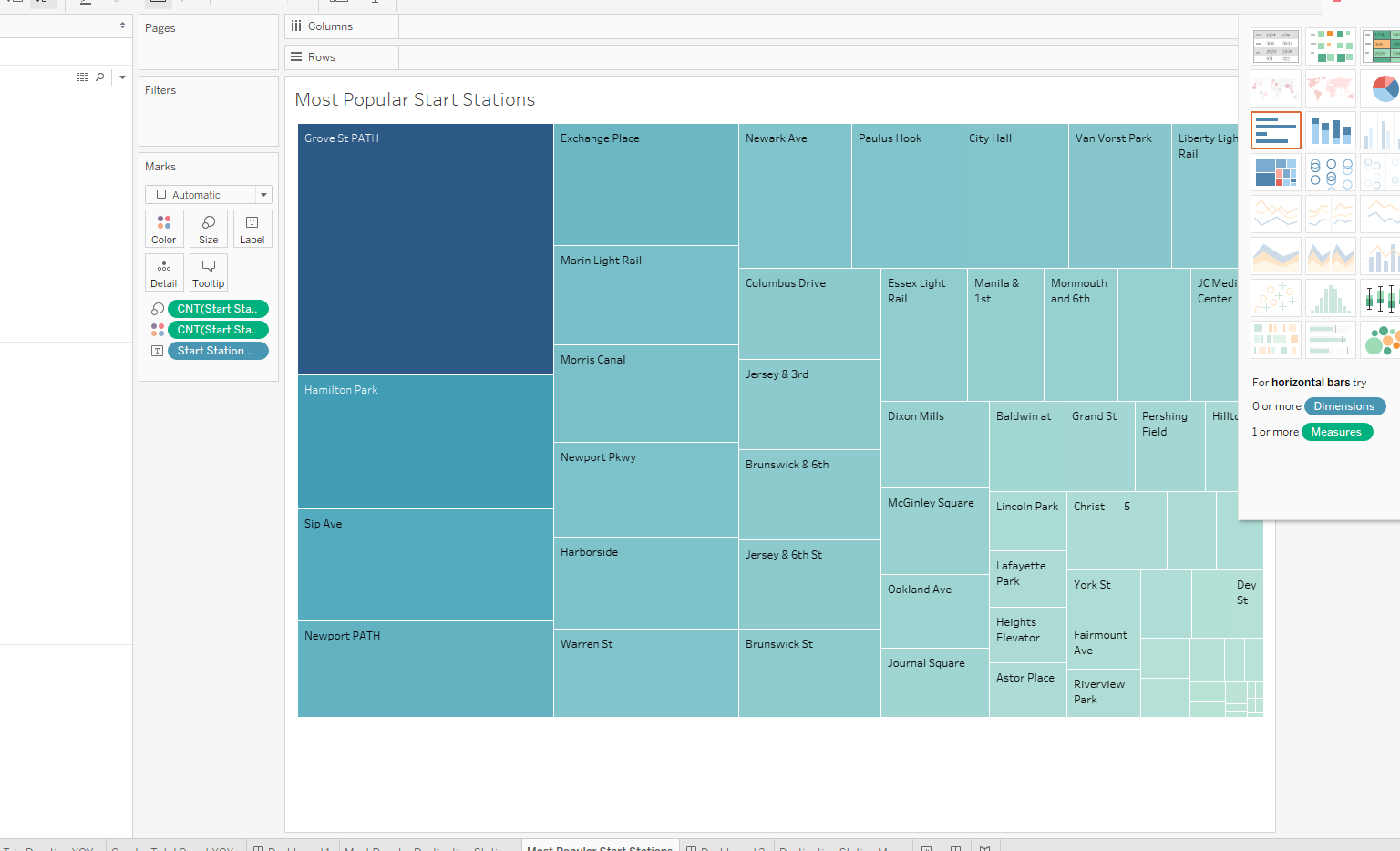


**Location**- What locations look to be the most popular? Are there any opportunities to do cross campaigns with other companies in certain **destination** areas? Think bundle campaigns that can be mutually beneficial for all parties involved….

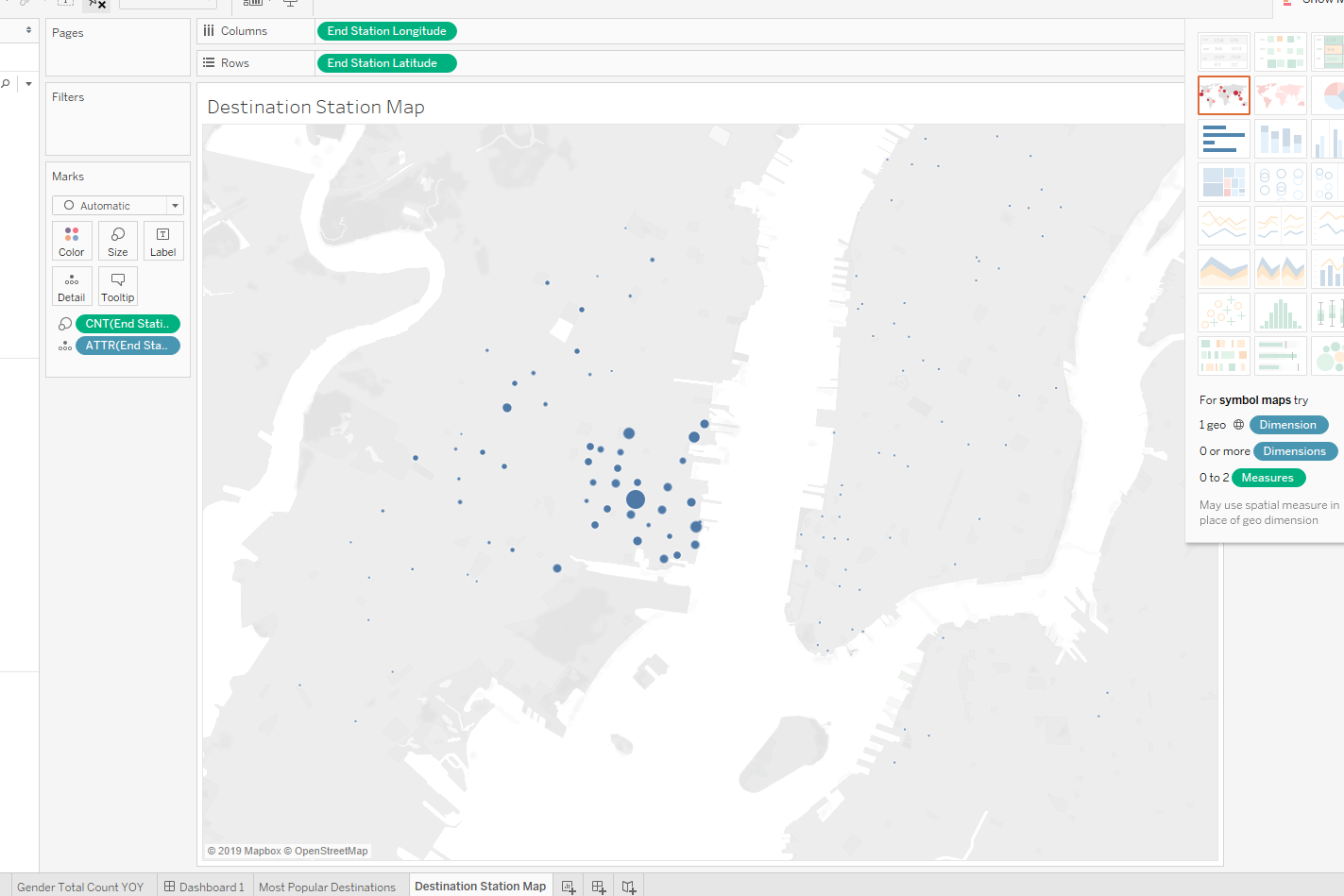
-Most popular destinations



-Most Popular Start Stations



-Destination Station Map



-Least popular destinations

Ways to Position the Data

-Try studying a time of the year where you know you’ll have the densest data to pull from.

-Three years’ worth of **summer** data should be good for the study. Will consider adding a fourth year as well depending on how good the visualizations look with the 3 points first.

-The goal here is to make sure we have our most abundant yet manageable distribution of data, not to com pare seasons etc… Thus, we can just keep it to the summer months.

**-May-July** since we are still in summer and there is no guarantee that they will release data in time

**\*\*\*\*\*Typing.com\*\*\*\*\***

**\*\*\*\*\*\*Typing.io\*\*\*\*\* >>> concentrates more on code typing (since code uses a bunch of weird characters)>>> You can also use different languages on this resource…**

**Getting Started**

Start simple with just three single months of data going back three years. i.e. May 2019, May 2018, and May 2017…

To Do…

-Questions… Ask about the validity of my strategic overview and some of the finer points of the advanced dynamic mapping.

-Start the project…

-Anything past 1920 or past 90 years of age is much less feasible, thus it is an outlier that can be deleted.

-People between 20 and 40 do short runs and people under 20 or so do shorter runs

**Machine Learning and BIG Data Notes…**

AWS can offer very powerful virtual machines…

-Anything over 10mb starts approaching the land of BIG Data.

-Zeplin Notebooks Big data (kind of like multi-threading…)

-Make sure to get very familiar with all of the activities in the machine learning portion… BUT, I should get really familiar and intimate with the

To do…

* ~~Make a discrete gender column in plain English (i.e. male, female, non-binary)~~
* ~~Make an age column~~
* ~~Add a data year column(i.e. summer 2017 or just 2017 or both etc…)~~
* ~~Convert trip duration into something that is a readable time~~
* ~~Rename columns where necessary (i.e. trip duration (in minutes))~~