

Paras Patel  
Data Analytics – Udacity  
Data Visualization Project – readme

For the data visualization project, the Ford GoBike system data for all of 2018 was analyzed. The monthly system datasets are available in .csv format. To make an analysis of the 2018 data, 12 files covering months January to December were concatenated.

The main findings are outlined in the 'FGB Exploratory Report'. The main findings focused around univariate, bivariate, and multivariate relationships between key features collected by Ford GoBike. Main findings summary:

- Potential product failure for a batch of bikes.
- Most trips last 600 seconds and 750-1500 meters.
- Majority of trips made at 8:00 AM and 5:00 PM.
- Most users are in the 28-33 year range.
- Drop off of trip duration around 60 minutes due to additional service fees.
- Customer and subscriber base is primarily male.
- Longest rides occur from 12AM–3AM. The shortest rides occur in the 6AM-9PM time frames. Most rides occur in the 6AM-9AM & 3PM-6PM time frames. Longest rides tend to start from the San Francisco Ferry Building.

The results in the explanatory analysis were based on validity of the data. For instance, the new distance calculated from average velocity, in the exploratory notebook, was not included in the final version because it was not used for analysis, and was based on a feature that was not directly collected by Ford GoBike. The visualizations included in the explanatory analysis illustrated properties, trends, and relationships from features directly available from the original dataset.

Resources: Udacity Knowledge Center, Ford GoBike website, Stack Overflow, hakimel / reveal.js Github repository.