

## Udacity DAND: Communicate Data Findings Project

Data Analysis/Data Visualization Report: Bike Ride Trends and Biker Types of Ford GoBike System November, 2019.

By Rahaf Alharbi

### Dataset

This document explores the Ford GoBike's trip data for public containing bike rides from April FY2019. The attributes included the trip start/end time, as well as additional measurements such as user type, gender, and age. 16K data points were removed from the analysis due to missing values in some fields, data inconsistent, or outlier issues.

<https://s3.amazonaws.com/baywheels-data/index.html>

<https://s3.amazonaws.com/baywheels-data/201904-fordgobike-tripdata.csv.zip>

### Summary of Findings

- 1- Bike Ride Duration Time: The original duration data has right skew issue - bike durations range from less than 1 minute to 1400+ minutes with median at around 9 min and mean at around 12 min. We have to do some data transformation to make data visualization and data interpretation easier.
- 2- Regarding the bike ride trends and biker types: 2.1) Tuesday, 5:00 PM has the highest biker counts across 7 days, 24 hours. 2.2) 5:00 PM has the most male bikers compared to other hours. 8:00 AM and 5:00 PM have more female bikers compared to other hours. 2.3) 5:00 PM has the most 'Subscriber' bikers compared to other hours. It also has the most 'Customer' bikers compared to other hours. 2.4) Tuesday has the most male bikers compared to other days. It also has the most female bikers compared to other days. 2.5) Tuesday has the most 'Subscriber' bikers compared to other days. Saturday has the most 'Customer' bikers compared to other days.
- 3- Bike Ride Durations for Different Age Group Across DayofWeek and Hour: 3.1) By looking at the errorbars in both pointplots below, we can tell that there are more younger bikers (age < 40) across 7 days and 24 hours. 3.2) Bikers on Saturday and Sunday bike longer compared to bikers on other days.

### Key Insights for Presentation Below are the insights for the presentation:

- 1- The distribution of biking durations is skewed. After log transformation and outlier removing, we may visualize and interpret data better.
- 2- Peak biking day and time are shared. Biker types were analyzed.
- 3- The patterns/trends for bikers from different age group are shared.