Project Proposal

This repo is one of the T5 Bootcamp requirements.

Who are the main BAY WHEELS users?

Bay Wheels is a regional public bicycle sharing system in California's San Francisco Bay Area .

The study aims to Find out who are the most important users of this service.

Dataset

The data set contains information about bicycles rides provided by Bay Area Motivate from the Bay Wheels bicycle sharing service. The service motivate is committed to supporting bicycling as an alternative transportation option.

The dataset consisted of 15 culomns and 519700 rows.

The data are available at https://s3.amazonaws.com/fordgobike-data/index.html

The dataset contains the following features:

Trip Duration (seconds)
Start Time and Date
End Time and Date
Start Station ID
Start Station Name
Start Station Latitude
Start Station Longitude
End Station ID
End Station Name
End Station Latitude
End Station Longitude
Bike ID
User Type (Subscriber or Customer – “Subscriber” = Member or “Customer” = Casual)
User Birth Year
User Gender

The dataset is available as the .csv file. a sample of data is shown in the following table:

	duration_sec	start_time	end_time	start_station_id	start_station_name	start_station_latitude	start_station_longitude	end_station_id	end_station_r
0	80110	2017-12-31 16:57:39.6540	2018-01-01 15:12:50.2450	74	Laguna St at Hayes St	37.776435	-122.426244	43	San Francisco Public Library (Grove St at Hyde...
1	78800	2017-12-31 15:56:34.8420	2018-01-01 13:49:55.6170	284	Yerba Buena Center for the Arts (Howard St at ...	37.784872	-122.400876	96	Dolores St at 1 St
2	45768	2017-12-31 22:45:48.4110	2018-01-01 11:28:36.8830	245	Downtown Berkeley BART	37.870348	-122.267764	245	Downtown Berkeley BART
3	62172	2017-12-31 17:31:10.6360	2018-01-01 10:47:23.5310	60	8th St at Ringold St	37.774520	-122.409449	5	Powell St BAR Station (Marke at 5th St)
4	43603	2017-12-31 14:23:14.0010	2018-01-01 02:29:57.5710	239	Bancroft Way at Telegraph Ave	37.868813	-122.258764	247	Fulton St at Bancroft Way

Tools

There are tools that will be used to achieve the goal of this study, such as:

- Pandas, Numpy for discovering the data and train a model .
- Matplotlib and Seaborn for visualization .
- The work will be done through Jupyter notebook.

TO DO:

- Explore the data and come up with EDA phases then use a model to fit the data.
- **NOTE:** the used features may be increased or changed and the model as well.