Denver Bike Share

Overview

Denver, Boulder and Lakewood are three cities in Colorado that participate in B-cycle, a public bicycle sharing company. Data from Denver and Boulder bicycle ridership have been available since 2010 for public consumption and for use in data wrangling and munging.

Goal

The goal of this Capstone project is to explore the available data using popular data visualizations tools such as Tableau, OpenStreetMap, Mapbox, and D3js, and predict ridership using weather and time of day using a linear model such as linear regressor and non-linear classifier models such as Random Forest and Gradient Boosting.

Client

Denver B-cycle and Boulder B-cycle are the clients for this project. According to Bicycling, Denver and Boulder were among the 50 best bike cities of 2016. Naturally, both cities would like to see increased ridership year by year. Studying past trends and predicting future trends would benefit and assist these cities in planning their systems more effectively and efficiently.

Data

There are datasets for both Denver and Boulder bicycle trips that go as far back as 2010. The most recent datasets available are for 2015. It may be possible that data for 2016 might be posted on their websites in time for use in this project.

Data Acquisition

Data will be collected from the publicly available datasets and merged with geo-spatial information from Google Distance Matrix API and weather data from Dark Sky API for data wrangling and munging. The data will be explored for ridership trends using data visualization tools as described above. Data from 2015 (or 2016, if available) will be scoped for use in training, validating and testing the linear and non-linear machine learning models.

Deliverables

Python code, Report, Slide Deck to be posted on Github.