General Assembly SF 17 – Data Science Class Project

Project Milestone 1. Due Date Sept 30, 2015

Submitted by Anil G.

* The project goal is to predict the bike rental demand in Washington DC.
* The project is based on the completed Kaggle competition – “Forecast use of city bike share system”
* The competition provides training and test dataset. On initial exploration the data seems to be good quality. The dataset is quite small and can be processed easily on my laptop.
  + Training Set
    - Two years of hourly data from Jan 2011 till Dec 2012
    - Contains the predictor values for first 19 days
    - Contains the output for the first 19 days
    - 10866 rows of observations
    - No missing values
* Test Set
  + Two years of hourly data from Jan 2011 till Dec 2012
  + Contains the predictor values for last days of the month (from 20th till 28, 29, 30 or 31)
  + Predict the bike demand for the test dataset
* I chose this topic because I believe that this data set will provide me to learn and use different data science techniques.