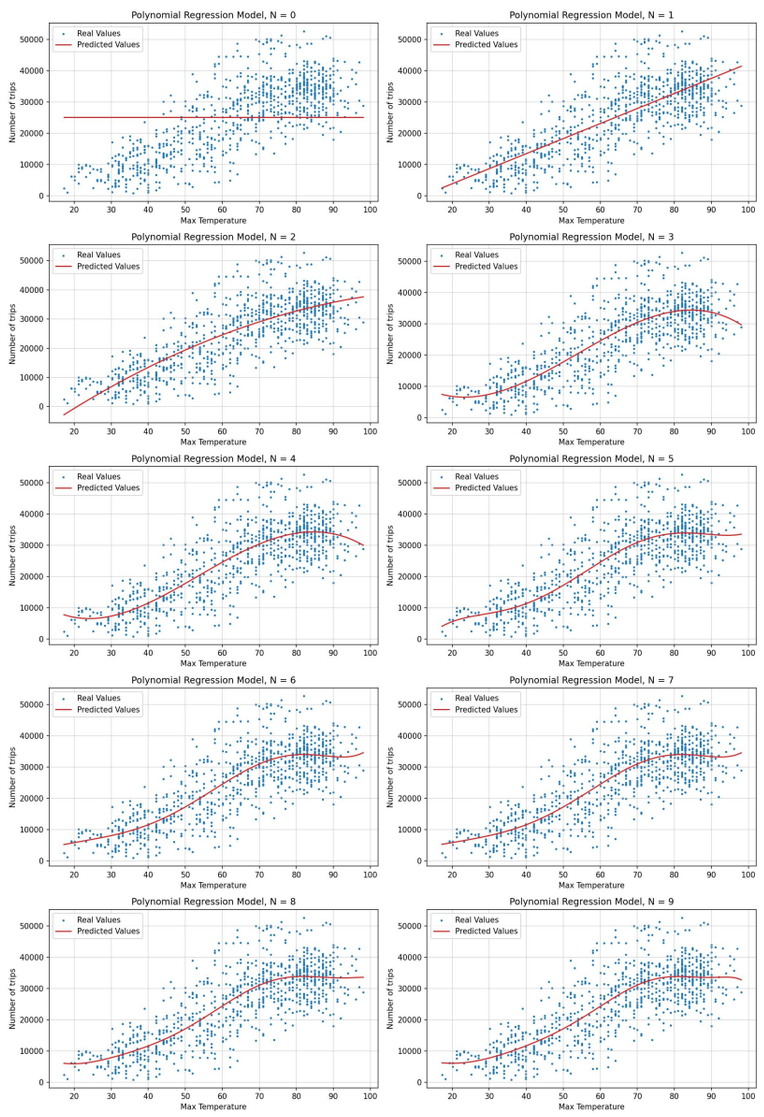
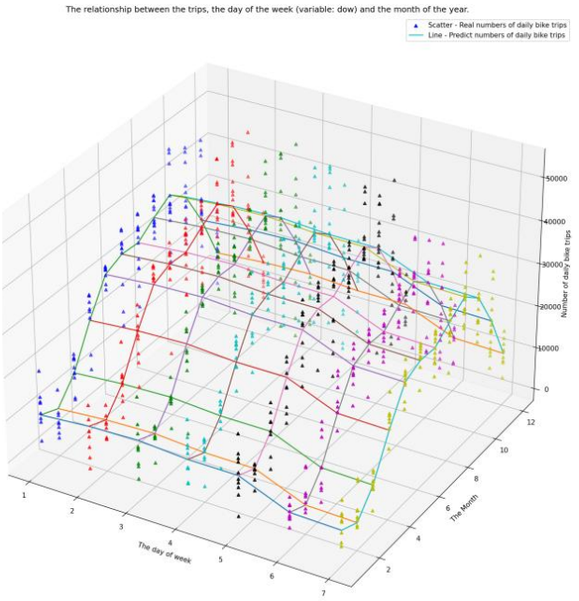
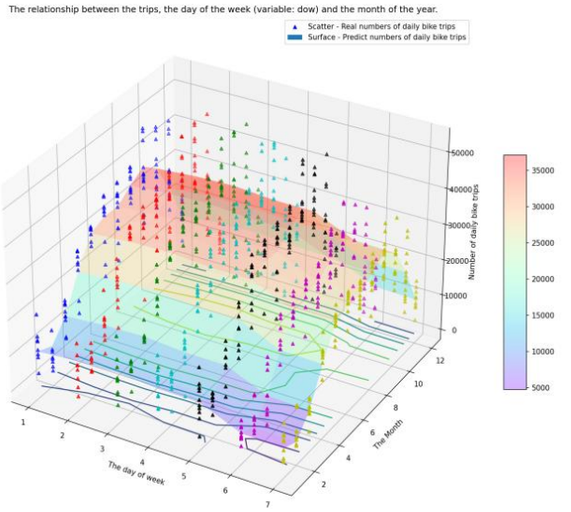
NYC’s City Bike system Analysis

Siheng Ning
sin21@pitt.edu







Legend

This project focuses on the some factor that impact the number of daily bike trips. The data includes information about the daily bike trips taken with the NYC’s City Bike system. The data frame includes also information about the weather during each day.

Figure 1

In this figure(left one) I was trying to build a model that will fit the number of daily bike trips based on the maximum temperature during the day. Groups of the polynomial regression models with different Nth degree polynomial is shown. The legends is on the left corners. The blue dots represent the daily highest temperature

correspond to the number of trips. The line represent fitted number of trips under different daily highest temperature.

Figure 2

The second plot(right two) is the number of daily bike trips correspond to the number of the day of the week (variable: dow) and the month of the year.

These two shows the real number of daily bike trips correspond to the number of the day of the week and the month of the year and them fitting value. Upper figure is the first plot which focus on showing the trend of the fitted values. Below figure more focus on showing the who is more dominate the trend in the two independent variables(day of the week(x) and the month of the year(y)).

Findings

1. After filtering I selected the model with N=4 as the final model choice. But the distribution of data points has some characteristics that the model does not describe. So I plot the second group of figures.

2. Contours at the bottom of figure at right below, mostly can be approximated as y = constants. And on the surface same month has the similar color. It all indicates that month of the year(y) is more dominate.

Github link:

https://github.com/tnbl/INFSCI2415FinalReport