Project Proposal

**Group Members:**

Gopal Menon and Kimberly Williamson

**Data Sources:**

We will be using the dataset from the Di-Tech challenge consisting of data collected around the ride-hailing company Didi Chuxing. The dataset includes data regarding the transportation behaviors Chinese citizens.

**Structure:**

We intend on applying regression techniques to the dataset. Through applying these techniques, we will be able to mine traffic patterns of Didi Chuxing customers. The traffic patterns we intend to data mine will include data regarding the time of the day, location of pick-ups and drop-offs, and other relevant information to forecast Didi Chuxing customer’s needs.

**Problem:**

With such a high reliance on non-personal transportation, it is important for the transportation industry to have a good understanding of the transportation patterns. Using this dataset, we hope to forecast the transportation needs of the customers of Didi Chuxing. By focusing better on the needs of the customers, Didi Chuxing can optimize their drivers and their revenue.

**What is New:**

We will explore a deeper understanding of regression techniques by comparing several regression techniques to determine the advantages and disadvantages to the techniques. (I need help here)