DATS 6103 Project Proposal

Team B

Hanzhong Gao

Zhuohan Yu

Aaron Yang

**Background:**

So called second hand’s car or used car has a huge market base. As it’s feasible and cheaper, more and more people consider to buy an used car instead of buying a new one. However, the main reason for the huge used car market is that car cannot maintain its value. The price of one car will reduce because of its age, mileage, damage and so on.

**Questions:**

In order to better understand this market, our team has collected data and decided to do statistical analysis on the used car market. The following questions are the main subjects that our team focus on:

o Understand and analyze what factors have impact on the value of a used car most.

o Predict the value of each used car in our dataset.

o Give suggestions on how to buy a used car with high cost effectiveness.

**Data:**

The data in this project contains 31 variables and more than 200,000 observations. 15 variables are anonymous we will drop them later. 150,000 observations will be train set and 50,000 observation will be test test.

GitHub Link: <https://github.com/Zhuohan-Data/TeamB-FinalProject.git>

Data Resource : <https://tianchi.aliyun.com/competition/entrance/231784/information>