**BIA 652**

**MULTIVARIATE DATA ANALYSIS**

**Spring 2018**

**Professor: M. Gandomi, PhD**

**Student:** Mukunth Rajendran & **Raphael Presberg**

**Assignment**

Submit the tentative Abstract of Project Proposal before next class:  
– Title  
– Names  
– Details about the database  
• Problem  
• Type: Classification (preferred), Regression,  
Clustering, etc.  
• Inputs and an Outpup  
• Reference: own (preferred) or borrowed

**Title:**

Green-Trip

Why?

The aim of the project is to make a taxi trip which is more effective and efficient for the customers, the drivers and the company, and reducing the fuel consumption.

We want to satisfy the customer, make the drivers happy, and improve the income of the company.

Company

Customer

Drivers

Environnement

According to Traffic Signal, Red states `do not cross`, Yellow states `continue to cross with caution`, and green states ‘safe to cross’; we are moving towards making our Yellow taxi to Green taxi, eco-friendly and safe to cross through our analytical problem solving approach and our predictions.

**Names:**

Mukunth Rajendran

Raphael Presberg

**Details about the database:**

* **Problem**

Customer:

* Reducing the waiting time for the Taxi and making it cheaper.

Driver:

* Improving the frequency of the Trips and earn more money during his/her working hours.

Company:

* Increasing their profits through satisfying the Customers, the Drivers and working towards maintaining the Eco-Friendly perspective.
* **Type**
* Classification (mainly classification, we don’t know exactly which algorithm, might be a neuronal network if we are getting better results)
* Regression
* **Inputs and Outputs**

Inputs: we have many variables that can be inputs. We will have to try different combination to discover patterns and models

Output: getting a pattern that is answering all the requirements

* **Reference**

Wharton Research Data Services (WRDS) – NYC Yellow Taxi Trips – 2015

Variables

| **Position** | **Variable Name** | **Type** | **Length** | **Description** |
| --- | --- | --- | --- | --- |
| 1 | pickup\_date | Num | 8 | Pickup Date |
| 2 | pickup\_time | Num | 8 | Pickup Time |
| 3 | dropoff\_date | Num | 8 | Dropoff Date |
| 4 | dropoff\_time | Num | 8 | Dropoff Time |
| 5 | vendor\_id | Char | 3 | Vendor ID |
| 6 | passenger\_count | Num | 8 | Passenger Count |
| 7 | trip\_distance | Num | 8 | Trip Distance (in miles) |
| 8 | pickup\_longitude | Num | 8 | Pickup Longitude |
| 9 | pickup\_latitude | Num | 8 | Pickup Latitude |
| 10 | rate\_code | Num | 8 | Rate Code |
| 11 | store\_and\_fwd\_flag | Char | 1 | Store and Forward Flag |
| 12 | dropoff\_longitude | Num | 8 | Dropoff Longitude |
| 13 | dropoff\_latitude | Num | 8 | Dropoff Latitude |
| 14 | payment\_type | Char | 3 | Payment Type |
| 15 | fare\_amount | Num | 8 | Fare Amount |
| 16 | surcharge | Num | 8 | Surcharge |
| 17 | mta\_tax | Num | 8 | MTA Tax |
| 18 | tip\_amount | Num | 8 | Tip Amount |
| 19 | tolls\_amount | Num | 8 | Tolls Amount |
| 20 | total\_amount | Num | 8 | Total Amount |

## **Public Data - NYC Yellow Taxi Trips 2015**

|  |  |
| --- | --- |
| **Library:** | public |
| **File:** | yellow\_2015 |
| **Data Range:** | 01/01/2015 - 06/30/2015 |
| **Last Updated On:** | 09/14/2015 |
| **Update Schedule:** |  |

### **Variable Reference**

| **Variable Name** | **Data Type** | **Variable Description** | **Help** |
| --- | --- | --- | --- |
| DROPOFF\_DATE | DATE | Dropoff Date |  |
| DROPOFF\_LATITUDE | NUM | Dropoff Latitude |  |
| DROPOFF\_LONGITUDE | NUM | Dropoff Longitude |  |
| DROPOFF\_TIME | NUM | Dropoff Time |  |
| FARE\_AMOUNT | NUM | Fare Amount |  |
| MTA\_TAX | NUM | MTA Tax |  |
| PASSENGER\_COUNT | NUM | Passenger Count |  |
| PAYMENT\_TYPE | CHAR | Payment Type |  |
| PICKUP\_DATE | DATE | Pickup Date |  |
| PICKUP\_LATITUDE | NUM | Pickup Latitude |  |
| PICKUP\_LONGITUDE | NUM | Pickup Longitude |  |
| PICKUP\_TIME | NUM | Pickup Time |  |
| RATE\_CODE | NUM | Rate Code |  |
| STORE\_AND\_FWD\_FLAG | CHAR | Store and Forward Flag |  |
| SURCHARGE | NUM | Surcharge |  |
| TIP\_AMOUNT | NUM | Tip Amount |  |
| TOLLS\_AMOUNT | NUM | Tolls Amount |  |
| TOTAL\_AMOUNT | NUM | Total Amount |  |
| TRIP\_DISTANCE | NUM | Trip Distance (in miles) |  |
| VENDOR\_ID | CHAR | Vendor ID |  |