**Course Project**

**Uber Movement Analysis in Mumbai**

*Team 3: Vasanti Budihal, Sagar Borkar, Shriya Karkal, Nishit Rao*

School of Business, Stevens Institute of Technology

BIA 658-B

Bei Yan

Spring 2022

**INTRODUCTION:**

To analyze the movement of Uber in Mumbai based on rush-hour traffic to provide a solution to traffic congestion issues. Traffic congestion has become a major issue in metropolitan cities. Resolving traffic-related issues requires an in-depth analysis of travel time. Aid in the planning of urban communities and increase mobility for the residents.

**DATASET EXPLANATION:**

The dataset we have used for our project is an open-source dataset available on Uber Movements webpage. There are various parameters that can be taken into consideration for analysis of traffic in Mumbai. ‘Travel Times’ gives a measure of the time taken on a particular route in the city, ‘Speeds’ gives information about street speeds of vehicles in a given area and a new feature of ‘Mobility Heatmap (BETA)’ shows the volume of activity based on mobility of devices. We can also choose a specific location or an entire town in the origin and destination inputs. We can also choose a point of origin and keep destination open. Speeds and Mobility Heatmap is not available as of now for Mumbai. Thus, we proceed with Travel Time which is given in seconds. We wanted to compare Traffic trends from Andheri and Borivali as Origin and kept destination open. We selected year 2016 which was supposedly best year for Uber in India and compared it the latest available 2020 year. We also have a secondary dataset which consist of different wards. Mumbai is divided into 24 different wards. This secondary data set consists of Travel Time in seconds between one ward as origin and destination is open to get data of all the wards. In our dataset we set some parameters like morning peak time and evening peak time. As per uber morning peak is 7am to 10am and evening peak is 7pm to 11pm. We can also select what days data you are looking for and we select all 7 days average of mean travel times

Dataset overview:

* 2016:
  + Rows: 657
  + Columns: 11
* 2020:
  + Rows: 851
  + Columns: 11

Secondary Dataset:

* Number of Wards: 24
* Rows: 625
* Columns: 9

The data is open source and easily accessible from the Uber Movements website. Data cleaning and formatting were done based on what variables would best help us build a network of the travel time model.

Chart

Description automatically generated with medium confidence

Fig 1: Morning Peak plot of 2016

A picture containing chart

Description automatically generated

Fig 2: Evening Peak plot of 2016

Scatter chart

Description automatically generated

Fig 3: Morning Peak plot of 2020

Bubble chart

Description automatically generated

Fig 4: Morning Peak plot of 2020

The network is an undirected network

**Key Variables:**

* Origin\_Display\_Name: Origin Location
* Destination\_Display\_Name: Destination Location
* Morning\_Peak\_Mean\_TT: Travel Times in Seconds for morning period (7 am - 10pm)
* Morning\_Peak\_Lower\_Bound\_TT: Least Travel Time required in Seconds.
* Morning\_Peak\_Upper\_Bound\_TT: Maximum Travel Time required in Seconds.
* Evening\_Peak\_Mean\_TT: Mean of Lower and Upper bound Travel Time (7pm to 11pm)
* Evening\_Peak\_Lower\_Bound\_TT: Least Travel Time required in Seconds.
* Evening\_Peak\_Upper\_Bound\_TT: Maximum Travel Time required in Seconds.
* Destination\_Movement\_ID: Unique ID for destinations

**ANALYTICAL APPROACH:**

We conducted the below analysis on our dataset

* Community Detection
* Scatter plot (using ggplot)
* Centralization
* Closeness
* Highest Degree
* Graph Density
* Transitivity
* Mean Distance

Ward wise analysis:

Using a secondary dataset, we plotted the Mean Travel Times (Seconds) taken between one ward to all other wards. Mumbai city is divided into 24 different wards. This box plot shows Inter-Quartile range between Upper and Lower bound Mean Travel Times and also shows Median of Travel Time. From this analysis we found that wards R/N and R/W have the highest Mean Travel Time of all other 24 wards. These Wards belong to Andheri and Borivali respectively.

Chart, box and whisker chart

Description automatically generated

Fig :

Scatter chart

Description automatically generated

Fig : Community Detection

Chart, scatter chart

Description automatically generated

Fig : Morning Peak Travel Time (2016)

Chart, scatter chart

Description automatically generated

Fig : Morning Peak Travel Time (2020)

Chart, scatter chart

Description automatically generated

Fig : Evening Peak Travel Time (2016)

Chart, scatter chart

Description automatically generated

Fig : Evening Peak Travel Time (2020)

A picture containing graphical user interface

Description automatically generated

Fig : Centralization (2016)

A picture containing text

Description automatically generated

Fig : Centralization (2020)

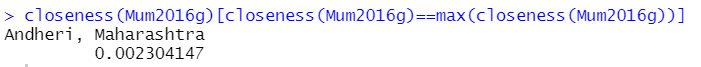


Fig : Closeness (2016)

Logo, company name

Description automatically generated

Fig : Closeness (2020)



Fig : Highest Degree (2016)



Fig : Highest Degree (2020)

Text, letter

Description automatically generated

Fig : Graph density, Transitivity, Mean distance (2016)

Text, letter

Description automatically generated

Fig : Graph density, Transitivity, Mean distance (2020)

**INTERPRETATIONS:**

Andheri is the most central node of the network. Increase in Uber clientele from 2016 to 2020. Many common Destinations between the two origin points. Changes in the longest path in the network from 2016 to 2020.

**RECOMMENDATIONS:**

Better city road planning to reduce travel time. Improve GPS for Mumbai roads to identify hidden routes which are not currently mapped. Promote the use of Uber Pool to reduce the number of vehicles on road at peak times.