RIDE HAILING ANALYSIS

Mrs.CH.SANTHI Assistant Professor, Department of CSE, NRI INSTITUTE OF TECHNOLOGY, Vijayawada, A.P., India.

SYED SALMAN[2], VALLURI SAI[3], CHENNU NARENDRA[4],

ATHI VENKATA KISHORE[5] Student, Department of CSE, NRI INSTITUTE OF TECHNOLOGY, Vijayawada, A.P., India.

ABSTRACT:

Huge amount of data is produced in Uber per every second. The project explains the working of an Uber dataset, which contains data produced by Uber for New York City. Uber is defined as a P2P platform. The platform links you to drivers who can take you to your destination. The dataset includes primary data on Uber pickups with details including the date, time of the ride as well as longitude-latitude information, Using the information, the project explains the use of the k-means clustering algorithm on the set of data and classify the various parts of New York City. Since the industry is booming and expected to grow shortly. Effective taxi dispatching will facilitate each driver and passenger to reduce the wait time to seek out one another. The model is employed to predict the demand on points of the city.Implementing of real-time data analysis on the dataset helps Uber for effective taxi dispatching which facilitate each driver and passenger to reduce the wait time to seek out one another. The

INTRODUCTION

Today transportation plays an important role in our lives. In urban area residents who are unable to drive by themselves, their transportation within the city is mainly public transportation including buses, subways and railways, taxis. Uber which is very popular among people recent year. The main difference is the way of calling and the way to pay.

So based on the need of transportation, the uber company introduced the service of ride to customers to reach their destination at a time. sometimes we didn't reach to our destiny at a time. To avoid these situations the uber company was provided cab service to the passengers.

We can maintain the data of uber trips by how many trips done per day and we can analysis the usage of service on which day

passengers need more. It is more popular in cities and the job holders more usage who are doing in shifts. We can increase their business by providing good service and behaving with customers.

By this, the customers cannot leave their service and they next time the customers prefer their service only and they prefer others also by maintaining good service to the customers.

The main motivation is for effective taxi dispatching which facilitates each driver and passenger to reduce the wait time to seek out one another. It is safest and comfortable to passengers who travelling in this cab.

EXISTING SYSTEM

Huge amount of data is produced by Uber per every second. Despite of having such an amount of data there are still some situations where the customer has to wait for a long time for the taxi. This can make the customer to use another application for booking taxi therefore the organization can lose their customers

This can affect both customer and the organization. The main reason to losing customers is delaying the customer time. And there is no analysis on customers which day they need.

The service of the uber is providing good but not in proper way. Behaving with customers also very important in business side. Sometimes due to misbehaving with customers also there is a chance to lose customers. So, they need to arrange proper drivers to behave in good manner with customers.

PROPOSED SYSTEM

Implementation of real-time data analysis on the dataset helps in extracting the useful information like: The day which is profitable day for Uber and the day on which a smaller number of people uses the services of Uber. The time which is busiest day for Uber.

The average starting rise time of Uber. The region in which there are many Uber trips and so on. This is an android application which will works like Uber. This application will provide service to the passengers and drivers.

Anyone can register on this application. Through this application the passengers can register the cabs and they can receive the cab at a time without delay. Based on the present data we can estimate which day has to get more profit.

And we can manage the daily trips of data and it is easy to check when we need that data in some situations. The customers can

reach their destiny at a time with cheapest price of service.

MODULES

File 1: Ride_Recomendation.ipynb:

Step 1**:**Import the necessary libraries

Step 2: Load and process the dataset

Step 3: Select the required features

Step 4: Implement Elbow or Knee Method to find optimal K

Step 5: Plot the Elbow Graph

Step 6: Run the K-means clustering Algorithm

Step 7: Store the cluster Centroids

Step 8: Visualize the centroids using scatter plot

Step 9: Visualize the centroids using google map

Step 10: Predict the nearest cluster for the new request

File 2: Analysis.ipynb:

Step 1**:** Import the necessary libraries

Step 2: Load and process the dataset

Step 3: Analyse the uber trips according to days of the month

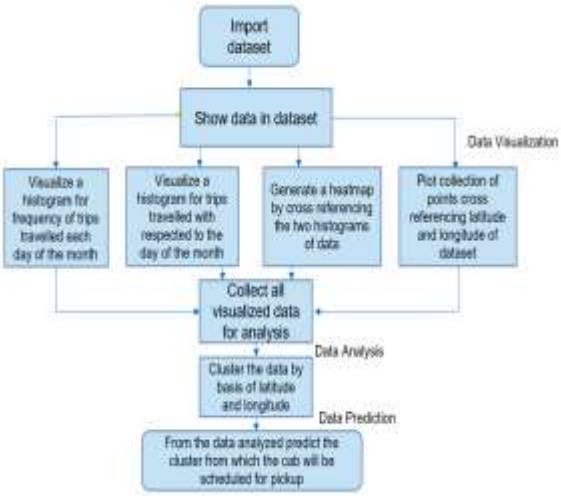
Step 4: Analyse the uber trips according to hours

Step 5: Analyse the uber trips according to Weekdays of the month

Step 6: Find corelation between hours and weekdays using a heatmap

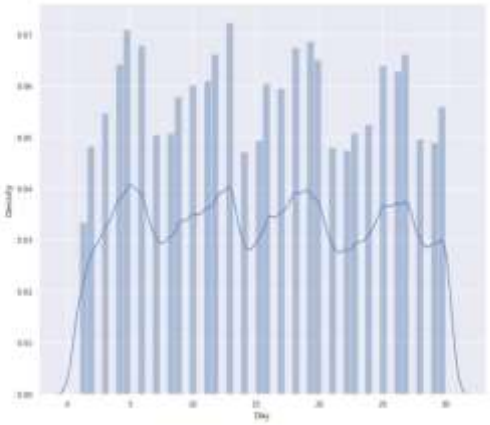
Step 7: Analyse the uber trips according to their density

SYSTEM ARCHITECTURE:



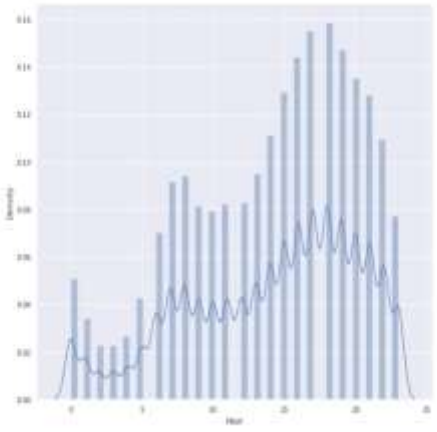
SAMPLE RESULTS

```
sns.set(rc={'figure.figsize':(12, 10)})
sns.distplot(data["Day"])
```



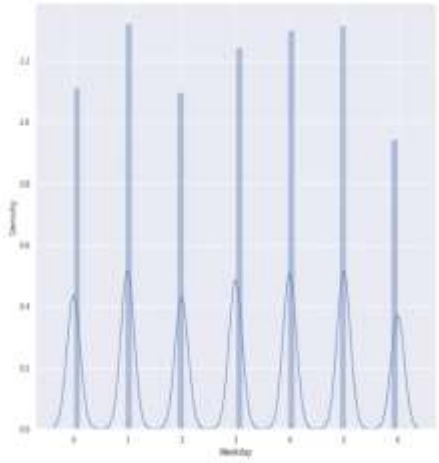
Analyze based on days

```
sns.distplot(data["Hour"])
```

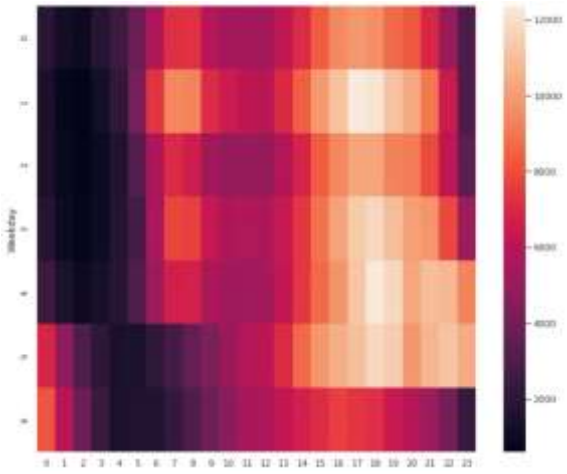


Analyze based on hours

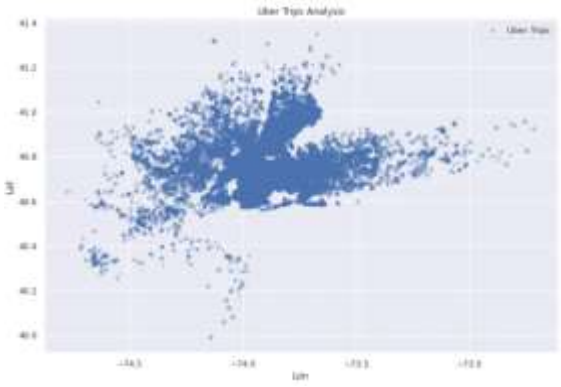
```
sns.distplot(data["Weekday"])
```



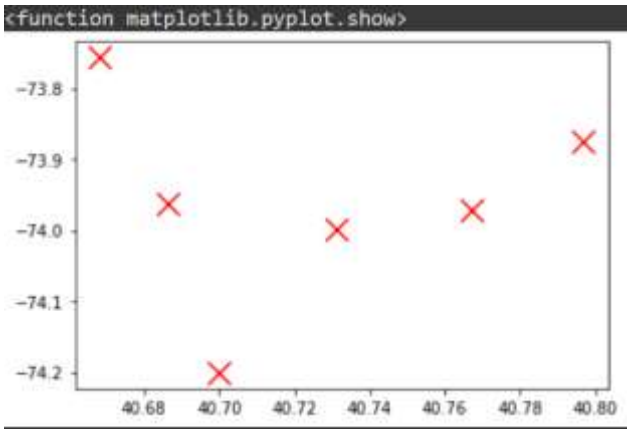
Analyse based on weekdays



Corelation between weekday and hour



Analyse based on density of trips



Visualize the cluster centroids

Visualization of Cluster centroids on Google map

CONCLUSION

The information which is extracted can be very useful for the Uber for making appropriate decisions so that they retain

their customers and as well as customers can reach their destination on time and satisfied with the service provided by the Uber.

REFERENCES:

[1] Poulsen, L.K., Dekkers, D., Wagenaar, N., Snijders, W., Lewinsky, B., Mukkamala, R.R.andVatrapu, R., 2016, June. Green Cabs vs. Uber in New York City. In 2016 IEEEInternational Congress on Big Data (BigData Congress) (pp. 222-229). IEEE.

[2] Faghih, S.S., Safikhani, A., Moghimi, B. and Kamga, C., 2017. Predicting Short-Term UberDemand Using Spatio-Temporal Modeling: A New York City Case Study. arXiv preprintarXiv:1712.02001.

[3] Guha, S. and Mishra, N., 2016. Clustering data streams. In Data stream management (pp.169-187). Springer, Berlin, Heidelberg.

[4] Ahmed, M., Johnson, E.B. and Kim, B.C., 2018. The Impact of Uber and Lyft on TaxiService Quality Evidence from New York City. Available at SSRN 3267082.

[5] Wallsten, S., 2015. The competitive effects of the sharing economy: how is Uber changingtaxis. Technology Policy Institute, 22, pp.1-21.

[6] Sotiropoulos, D.N., Pournarakis, D.E. and Giaglis, G.M., 2016, July. A genetic algorithmapproach for topic clustering: A centroidbased encoding scheme. In 2016 7th InternationalConference on Information, Intelligence, Systems & Applications (IISA) (pp. 1-8). IEEE

[7] Faghih, S.S., Safikhani, A., Moghimi, B. and Kamga, C., 2019. Predicting Short-TermUber Demand in New York City Using Spatiotemporal Modeling. Journal of Computing inCivil Engineering, 33(3), p.05019002.

[8] Shah, D., Kumaran, A., Sen, R. and Kumaraguru, P., 2019, May. Travel Time EstimationAccuracy in Developing Regions: An Empirical Case Study with Uber Data in Delhi-NCR✱.In Companion Proceedings of The 2019 World Wide Web Conference (pp. 130-136). ACM.

[9] Kumar, A., Surana, J., Kapoor, M. and Nahar, P.A., CSE 255 Assignment II PerfectingPassenger Pickups: An Uber Case Study.

[10] L.Liu, C.Andris, and C.Ratti , “Uncovering cabdrivers behaviour patterns from their digital traces”,Compu. Environ.UrbanSyst.,vol.34,no.6,pp.541-548,2010