CAPSTONE TWO: PROJECT PROPOSAL

1. UBER/LYFT DEMAND-PRICE TREND ANALYSIS:

* Problem statement formation: Analyzing demand of cab rides at a location. Also considering time of the day and weather data in determining demand and thus, see a trend throughout the day to answer a few questions:

1. What should be fare per cab ride? So as to meet the competition.
2. Which of the days business is less? and thus, employee can take leaves.
3. Which of the location have more demand at what time? And thus, transfer cabs to the location by avoiding traffic.

* Context: Uber/Lyft's ride fares varies dynamically. They are greatly affected by a lot of factors like demand of rides and supply of rides at a location, weather at a given time & time of the day. So what exactly drives this demand? Is it location specific?

How to make a balance between demand and supply?

* Criteria for success: Effectively answering questions mentioned in Context/Problem Statement Formation and/or determining other valuable insights that will help in growing business/customer base
* Scope of solution space: Determining relationship between various factors affecting cab-demand by using plots (boxplot, pairplot) and correlation matrix. We can explore other methods/models as well, during real-time execution of the project.
* Constraints: if no direct relationship exists between weather, location, cab demand, time of the day, then it may be difficult to reach any conclusion
* Stakeholders: Ankit Gupta
* Data Sources: <https://www.kaggle.com/ravi72munde/uber-lyft-cab-prices>
* Deliverables: A GitHub repo containing the work for each step of the project, including: a slide deck & a project report

1. CHURN (ATTRITION RATE) PREDICTION:

* Problem statement formation: Analyzing a rider’s activity for the cab-ride service by using last trip date, trips in January, ratings provided to/by driver etc and providing answers to a few questions:

1. How to retain riders? Predicting retention/attrition of cab-riders to determine revenue, competition and accordingly taking steps (i.e. providing discount coupons, preferred drivers) to retain/grow the riders
2. Which cities have higher retention of riders?
3. If the riders are distance specific? So may be they take a long distance ride in 2 or more months instead of taking rides every month
4. Which of the riders prefer luxury cab for their ride? So more discounts can be provided to such riders to retain them.
5. Tracking weekday trips of a rider to determine if the rider have switched job, location etc and providing the a good experience (i.e. preferred drivers) based on previous activities

* Context: A company providing cab-ride service wants to predict its rider retention. Analysis is to be performed on a group of riders who signed up for cab-ride service in January 2014 and the data was pulled on July 1, 2014. Generally a rider is considered retained if he/she is “active” if they have taken a trip in the past 30 days (since June 1, 2014).
* Criteria for success: Effectively answering questions mentioned in Context/Problem Statement Formation and/or determining other valuable insights that will help in growing business/rider base
* Scope of solution space: Determining relationship between various factors affecting rider-retention by using plots (boxplot, pairplot) and correlation matrix. We can explore other methods/models as well, during real-time execution of the project.
* Constraints: if no direct relationship exists between various rider activities such as rider/driver rating, ride distance, weekday rides etc. then it may be difficult to reach any conclusion
* Stakeholders: Ankit Gupta
* Data Sources: https://github.com/gogowenzhang/ride-sharing-churn-prediction
* Deliverables: A GitHub repo containing the work for each step of the project, including: a slide deck & a project report