Data Scientist

Case Study

EVERY CONNECTION COUNTS

NYC Green Taxi Case Study

- Executive Overview of Analysis like Trip Summary, Trip Fare, Yearly Highest and Lowest trips, Rate per trip, Weekdays vs weekend, Airport vs. Non Airport etc. using visual analytics techniques
- Descriptive Statistics
- Treatment of data (outlier, missing etc.)
- Demand of each pickup and drop points and its comparison based on time
- Best and worst route based on revenue and trip rate
- Predict the future demand for a given location and timeslot using AI/ML Algorithms
- Future Scope of Analysis
- Optimization and Suggestions

Dataset link: https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page (Green taxi, 2018)

Data Dictionary: https://www1.nyc.gov/assets/tlc/downloads/pdf/data_dictionary_trip_records_green.pdf

Tools: **Python** or R. Tableau, Power BI can be used for support in visualization

Output: Final presentation in MS Power point/PDF Format along with R/Python working file

Note: Plagiarism must be avoided