Reviving the Taxicab Business

Presented by:

Jeevisha Anandani
Lekshmi Venugopal
Tanu Seth

Background

The Rise and Fall of Yellow Taxicab Business in the New York City

- Taxis, as we traditionally know them with taximeters, began to roll in 1907.

- For many years Yellow Taxi Cabs dominated the market in New York to provide social mobility.

- With the rise of smartphones and apps like Uber, Lyft making it easier to hail private rideshares, the yellow cab's grip on the city has dramatically waned.

- The share of passengers that chose to ride taxis dropped from 84% in April 2015 to only 65% one year later and continued to go down significantly.

- The average money brought in by yellow cab medallions dropped by nearly 30% between 2013 and 2018, from roughly $14,500 per month to $10,200.

Executive Summary

Challenge

- Decline in the yellow taxi cab business over the past decade in New York City. Close to 50% decrease in number of trips per year from 2011 to 2019.

- 90% of business concentrated in One Borough.

Our Solution

- Scope of a utility which predicts price of the trip before hand.

- Expand the business to other Areas.

Expected Impact

- Enhance Customer Experience and getting a grip back on the riders.

- 20% increase in revenue generated per mile per trip on expanding to other areas.

Visualizing The Data

The Epic rise and fall of New York's Iconic taxis
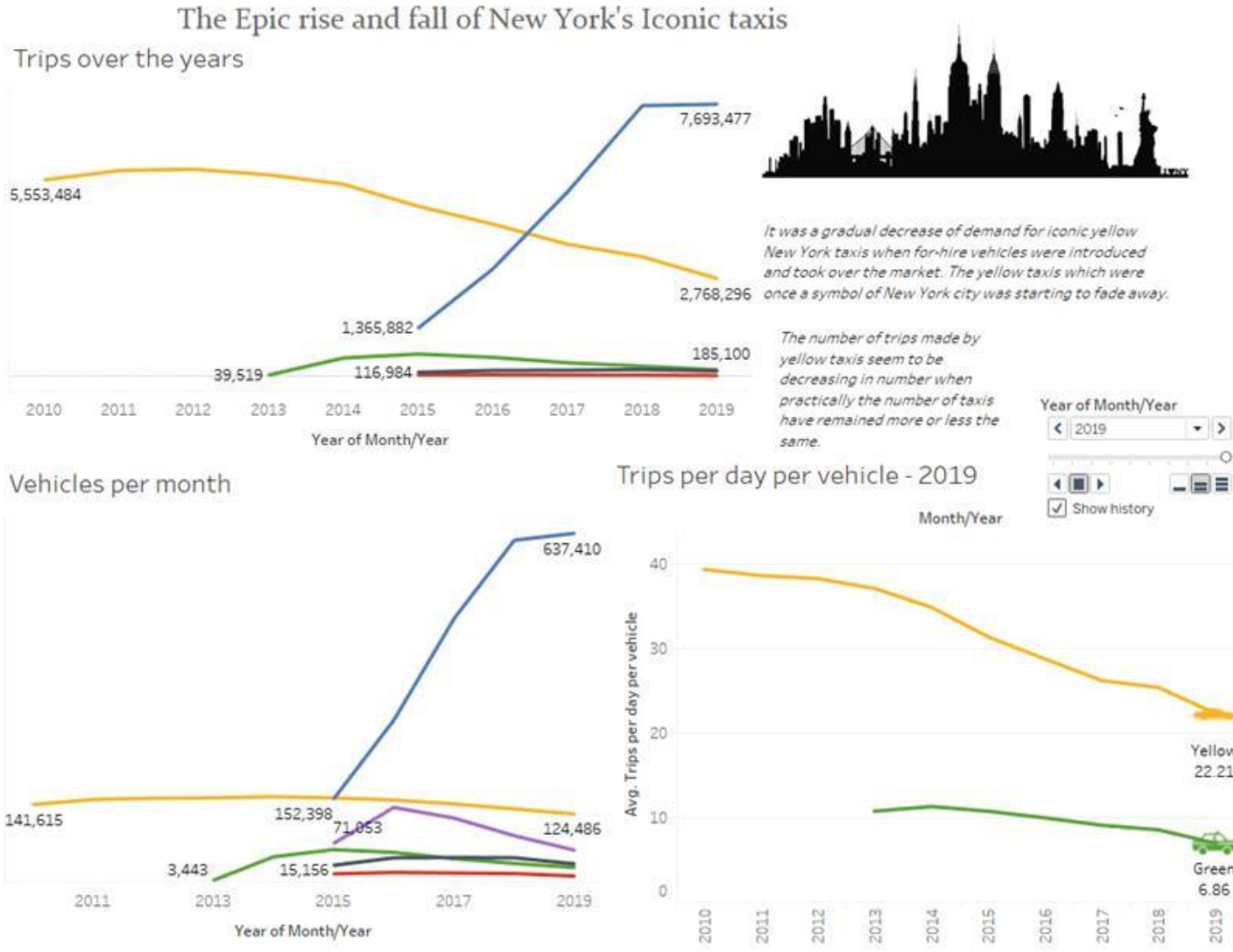
Business Problem

1 **Decline in number of trips over the years**

From 5.5 million trips per day in 2010 to 2.7 million trips per day in 2019.

2 **Decline in unique vehicles on road each month**

From 140k vehicles on road per month in 2010 to 124k vehicles on road per month in 2019.

3 **Decline in trips per day per vehicle**

From 40 trips per day per vehicle in 2010 to 22 trips per day per vehicle in 2019.

4 **Restricted business scope region wise**

Approximately 90% of the trips in 2019 January were concentrated in Manhattan.

Improving the business

Solution Overview

Improving customer experience by providing trip price before hand.

Implementation : Regression Model to predict trip fare

Scaling the business to other Boroughs

Implementation : Recognizing Boroughs to invest in for business scaling

Overview:

Price Prediction model

Aim – Predict an estimated fare amount to the passengers so that they can easily compare beforehand with other ride options

Two types of prices provided: -

- **Fare Amount –** The time and distance fare calculated by the meter

- **Total amount –** The total amount charged to passengers including tax, extra charges, improvement surcharge,etc.

Data Processing

Data Used

- January 2019 Yellow Taxi Data
- Limiting scope to Manhattan - provides the most reasonable picture for the yellow cabs.

Data Cleaning

- Tips < 50% of Fare Amount
- Fare Amount < $300
- Trip Distance < 200 miles
- Number of passengers <8

Feature Engineering

- Speed of the taxi
- Duration of the ride
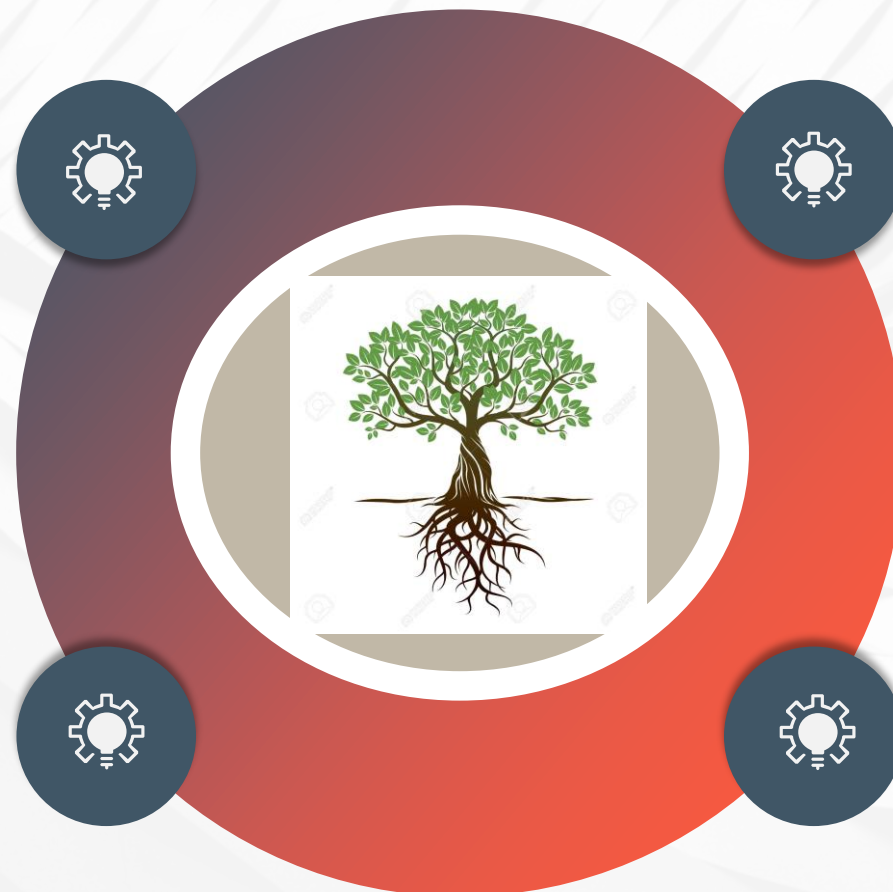- Percentage Tip
- Distance type
- Price per mile

Gradient Boosting vs Random Forest Regression

What is XGBoost?

It builds one tree at a time, where each new tree helps to correct errors made by previously trained tree.

What is Random Forest?

It is a bagging technique where trees are run in parallel without any interaction.

Top 5 features from Feature Importance

Duration of ride
Trip Distance
Rate Code ID1
Rate Code ID2
Speed

Top 5 features from Feature Importance

Duration
Trip Distance
Rate Code ID1
Price per mile
Rate Code ID 2

Model Evaluation

Summary

XGBoost

The standard deviation of the unexplained variance in the price is coming out to be 1.05

Random Forest

The standard deviation of the unexplained variance in the price is coming out to be 1.10

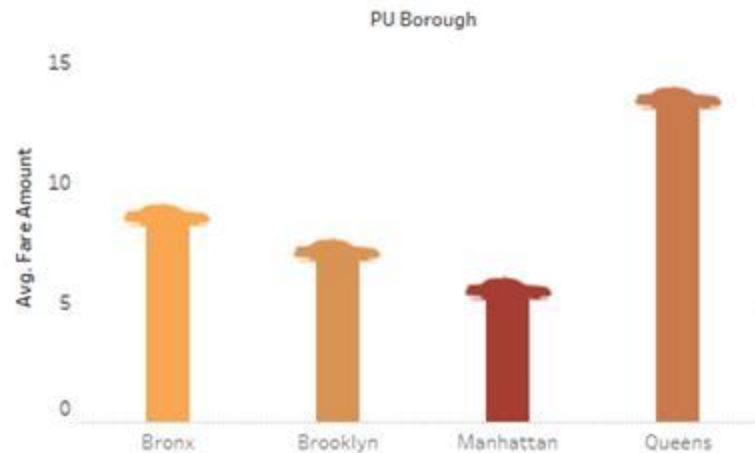# Recognizing Boroughs to invest in for business scaling

- Current Scenario: Most of the taxis are concentrated in a single Borough which is Manhattan.

- Visual Analysis of taxi ride data to know the distribution of taxis and the average fare in each Borough.

- Recommendation : To scale business in Queens and Brooklyn.
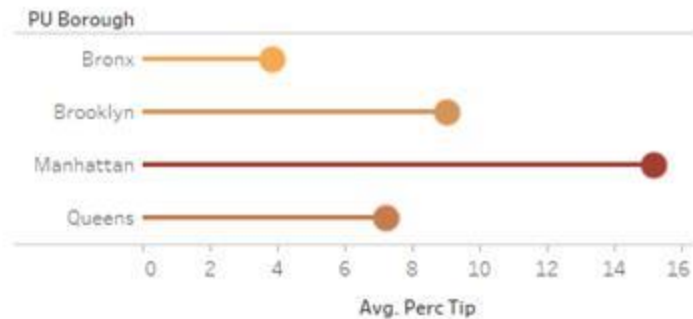
The bustling streets of Manhattan, NYC

Why did we choose Manhattan?

PU Borough
- Bronx
- Brooklyn
- EWR
- Manhattan
- Queens
- Staten Island

Manhattan is the most densely populated and thereby the busiest areas/boroughs of New York. Since the most number of people reside in Manhattan or "the city", shouldn't it be that the most number of requests for taxi need also come from the same borough? The answer given by our friend, the dataset, was a resounding YES.

From the fares, it is quite obvious that the number of trip requests that come from Manhattan is so high that the lesser fair charges can actually be balanced by the number of requests. Also, people hailing cabs from Manhattan tend to give higher percentages of the fare as tip which also can account for the highest revenue from the borough.

Average fare per trip per borough

PU Borough

Average tip percent per borough

PU Borough

12

Recommendation

Short Term

Provide better Customer Experience

Showing trip price before hand on a website by taking the pickup and dropoff locations

Optimized scheduling

Scheduling the taxi supply based on demand forecasting. Analyzing uber surge charge pattern.
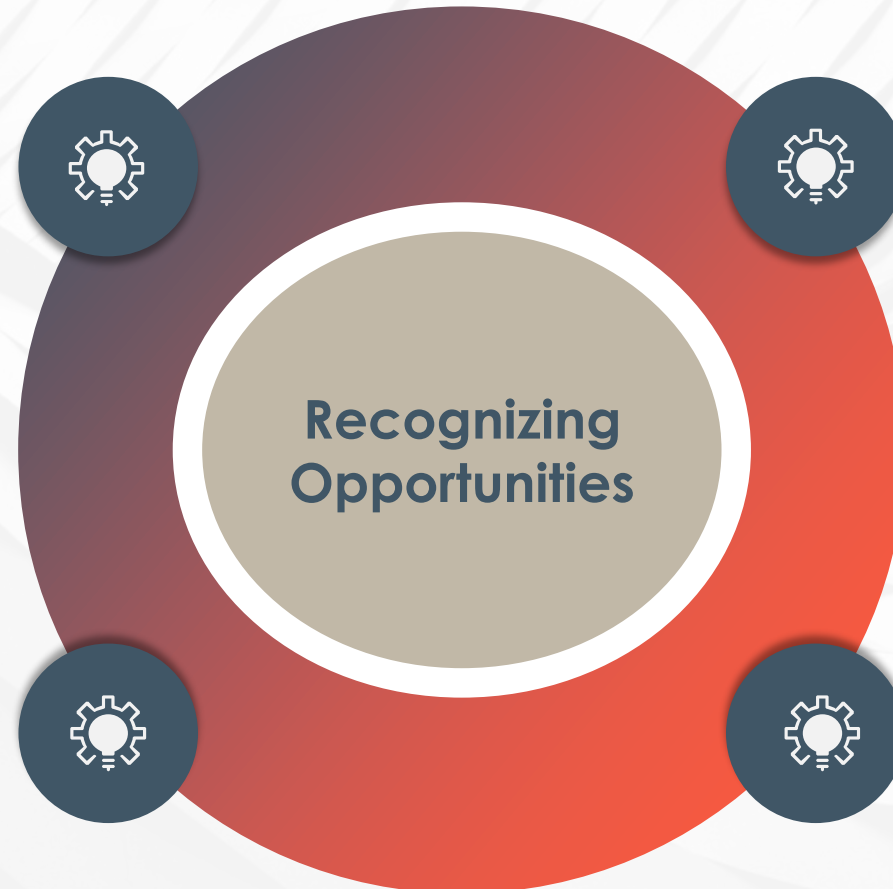
Long Term

Comfortable Rides

Upgrading the vehicles involved to offer cab services

If you can't beat the competition, Join them

Already popular in the market as street hailing cabs, providing a mobile application platform for ride booking to customers

Recognizing Opportunities

THANK
YOU

References:

- https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

- https://www1.nyc.gov/site/tlc/about/aggregated-reports.page

- https://www.nydailynews.com/new-york/ny-medallion-foreclosures-taxi-bailout-plan-uber-lyft-20200130-s2mjkhjubzgptdxasoxddwdote-story.html

- https://digg.com/2019/uber-vs-yellow-cab-in-new-york-city-visualized

- https://pubsonline.informs.org/doi/pdf/10.1287/ited.2019.0207cs