Deal or No Deal?
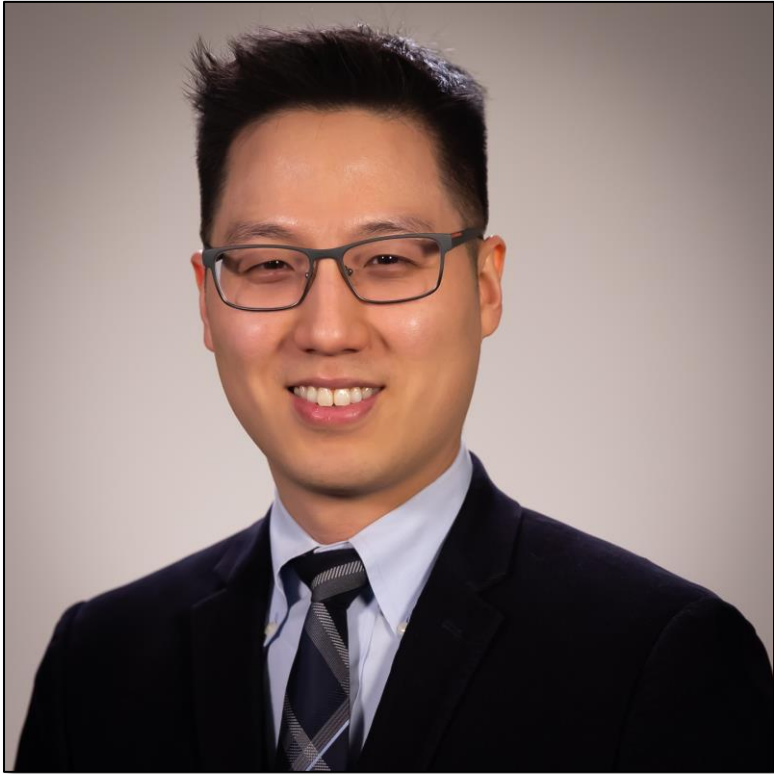
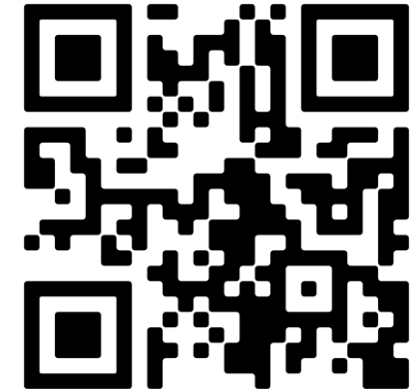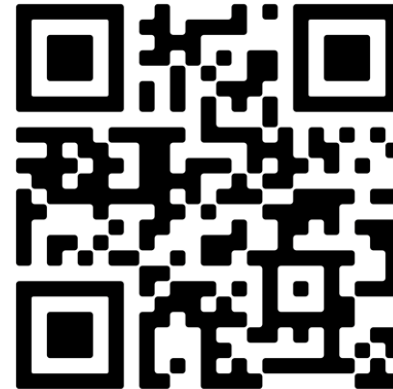A Predictive Analysis of NYC Airbnb Listings

By: Sam Choe

May 30, 2024

Meet The Presenter

Presenter: Sam Choe
Data Scientist

GitHub

LinkedIn

Agenda

Business Problem

Data Overview
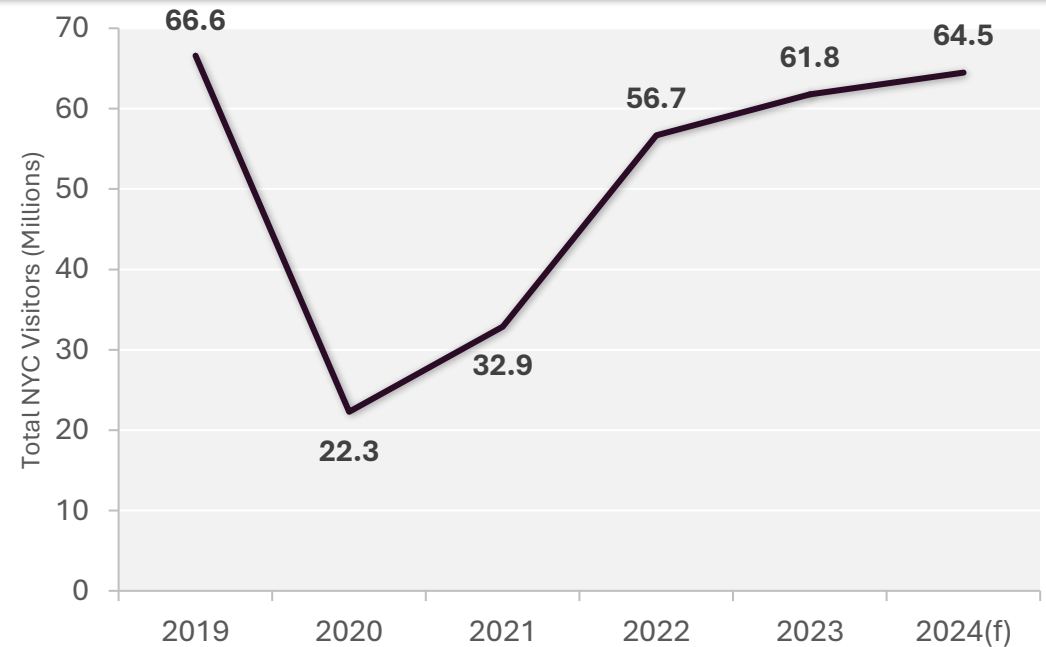
Analysis

Recommendations

Future Steps

Business Problem

Welcome to NYC!

- NYC is a popular tourist destination

- Budget conscious tourists are on the hunt for great accommodation deals

- However, this city is very expensive…

NYC Tourism On The Rise

Line chart of Total NYC Visitors (Millions): 2019: 66.6, 2020: 22.3, 2021: 32.9, 2022: 56.7, 2023: 61.8, 2024(f): 64.5

Source: NYC Tourism + Conventions (NYC Travel & Tourism Outlook)

Business Problem → Data Overview → Analysis → Recommendations → Future Steps

Data Overview

More Listings = More Options

- **Data Source**
  - o Airbnb dataset with ~67,000 unique listings
- **Exploratory Data Analysis**
  - o Data Manipulation
  - o Feature Engineering
  - o Target Variable (Airbnb Experience Quality)
- **Data Limitations**
  - o Somewhat dated
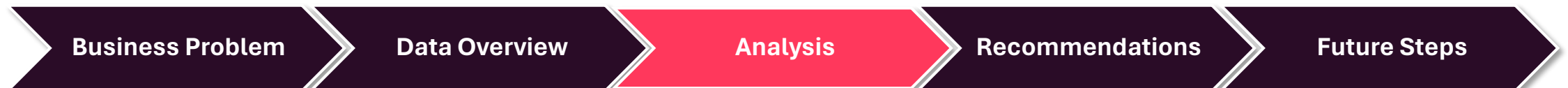 - o NYC listings mostly concentrated in Manhattan and Brooklyn

Analysis

Our Modeling Process

- **Models Explored:**
  - Logistics Regression
  - Random Forest Classifier
  - Gradient Boosting Classifier
  - XGBoost Classifier

- **Metrics:** Accuracy Scores

There Are Features That Matter

- **Model:** XGBoost Classifier (Accuracy Score: 46%)

Feature Importance Table	
Rank	**Feature**
1	**Number of Reviews**
2	**Brooklyn**
3	**Price Per Night**
4	**Bedrooms**
5	**Bathrooms**
6	**Manhattan**
7	**Amenities Count**
8	**Other Boroughs**
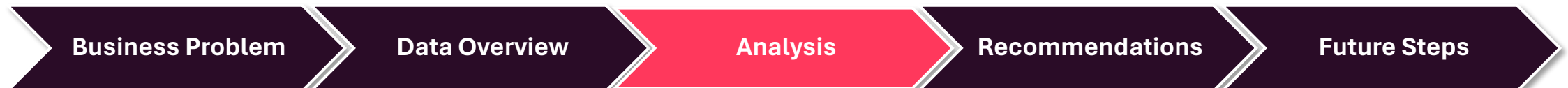
What Are The Odds?

- **Model:** Logistics Regression (Accuracy Score: 42%)
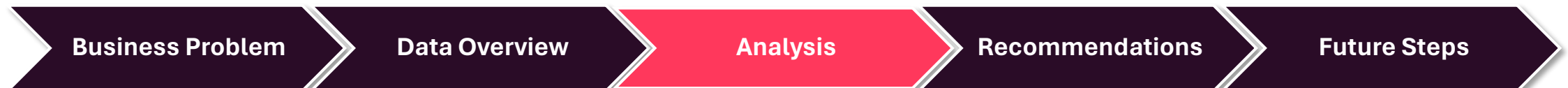
For each 1% increase in:	Increase in odds of having good quality listing
Number of reviews	**25%**
Association with Brooklyn	**16%**
Listing's Price Per Night	**6%**

Business Problem   Data Overview   **Analysis**   Recommendations   Future Steps
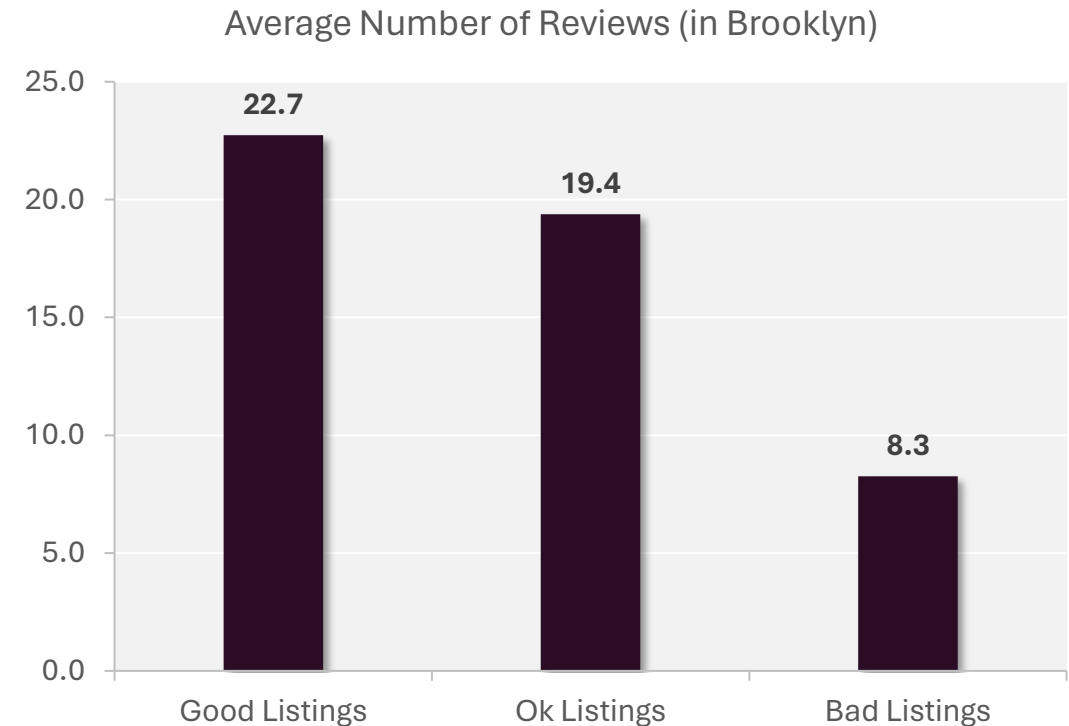
Recommendations

Deals Grow In Brooklyn

- Proportionally, Brooklyn has access to more good experience quality listings

- **Recommendation 1:** Favor Brooklyn Listings

Good Listings As % of Total Listings

Business Problem → Data Overview → Analysis → **Recommendations** → Future Steps
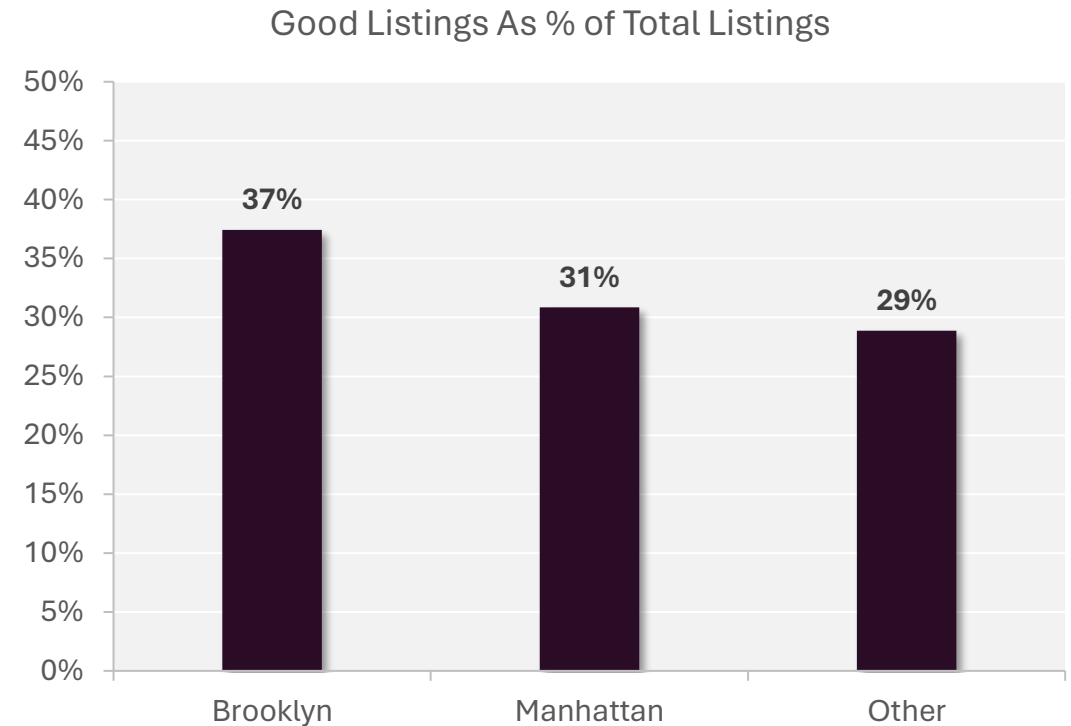
Reliable Listings Have More Reviews

- Better listings are associated with more user reviews

- **Recommendation 2:** Target listings with more than 20 reviews

Average Number of Reviews (in Brooklyn)

	Good Listings	Ok Listings	Bad Listings
	22.7	19.4	8.3

Business Problem | Data Overview | Analysis | **Recommendations** | Future Steps

Pay Up For Premium

- Good experience quality listings generally have higher price tags

- **Recommendation 3:** Pay a slight premium (~10-20%) to secure better Airbnb experiences

Average Price Per Night (in Brooklyn)

Listing	Average Price
Good Listings	$127
Ok Listings	$113
Bad Listings	$108

Business Problem → Data Overview → Analysis → **Recommendations** → Future Steps

Future Steps

We Can Do More….

- Expand our dataset to incorporate Airbnb API
- Rerun analysis to look at other geographies
- Delve deeper into quality control data around the user/host verification process
- Consider other model implementations

Questions?