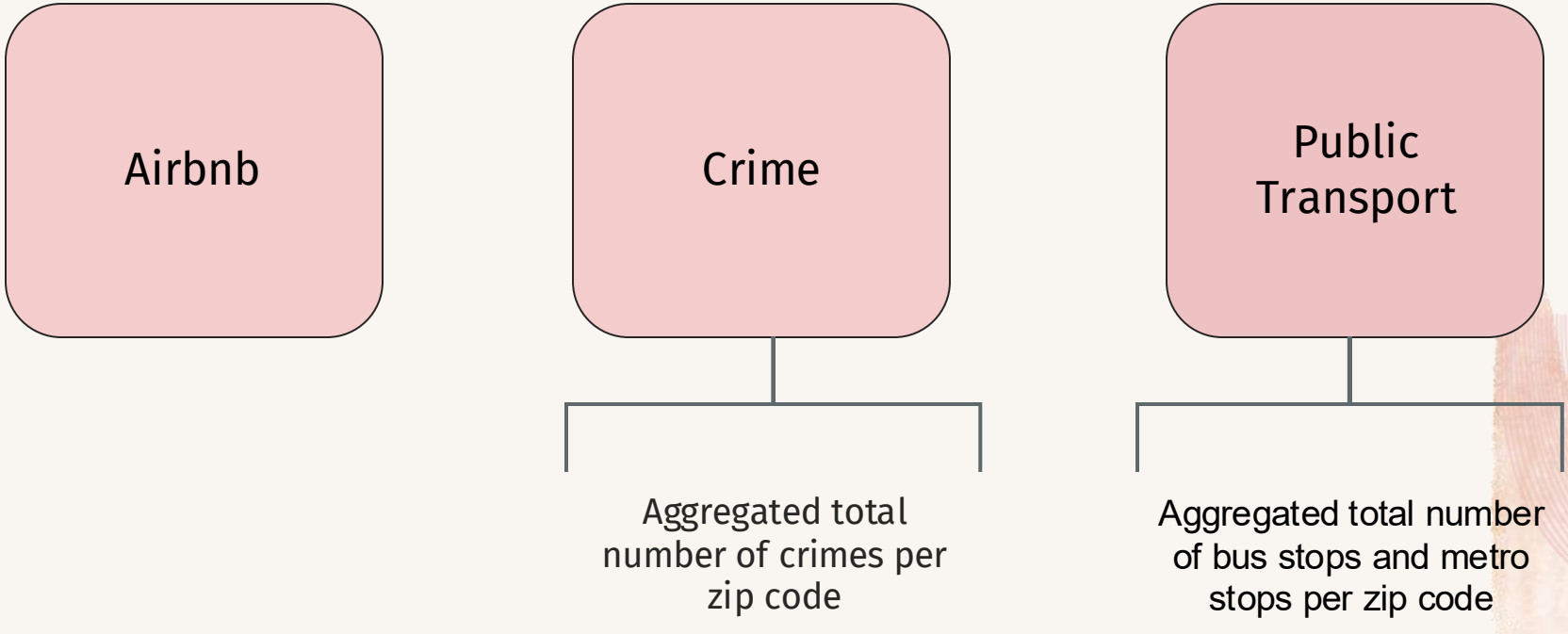
Airbnb Occupancy and Revenue Analytics

Los Angeles County

Introduction

Airbnb Prediction focuses on analyzing data related to rental prices and occupancy rates across various listings. This involves evaluating factors like local crime rates, transportation access, and property amenities, ultimately guiding hosts to optimize their listings for better revenue generation.

Data


Airbnb
Crime
Public Transport
Aggregated total number of crimes per zip code
Aggregated total number of bus stops and metro stops per zip code

Feature Engineering

- Amenities - grouped the similar amenities together, added super categories, summed up and then encoded them.
- Property type - group similar property types encoded the property types.
- Datetime columns - changed date time columns to numeric columns by counting the days since that date.
- Room Types - encoded the 4 room types.
- Bathroom count - Took shared bathrooms as 0.5 bathroom and private bathroom or nothing mentioned as 1 bathroom/
- Host response time - Used cardinality to do label encoding

- Imputed columns like days since last review where the value does not exist as -1 to mask it
- Imputed zero to columns where nulls are supposed to mean 0
- When needed to impute mean we imputed it as mean grouping listings by zip code to make sure the location is taken care of
- For price column to deal with outliers log transformed the column

Important Features

Price

- Weapons-Related Crimes
- Host ID
- Superhost Status
- Free Wi-Fi and Internet
- Number of Reviews
- Safety Features
- Swimming Amenities
- Kitchen Essentials
- Family-Friendly Amenities
- Number of Bathrooms

Occupancy

- Years as Host
- Host Acceptance Rate
- Total Listings Owned by Host
- Average Minimum Nights Required
- Host Response Time
- Time Since Last Review
- Total Amenities Offered
- Private Room Listings by Host
- Host Response Rate

R Squared Value

Linear Regression

Price: **52.97**
Occupancy: **89.99**

Random Forest

Price: **63.66**
Occupancy: **91.61**

GBT

Price: **77.07**
Occupancy: **91.5**

Current results (Price)

77.07

R Squared value

Current results (Availability)

91.5

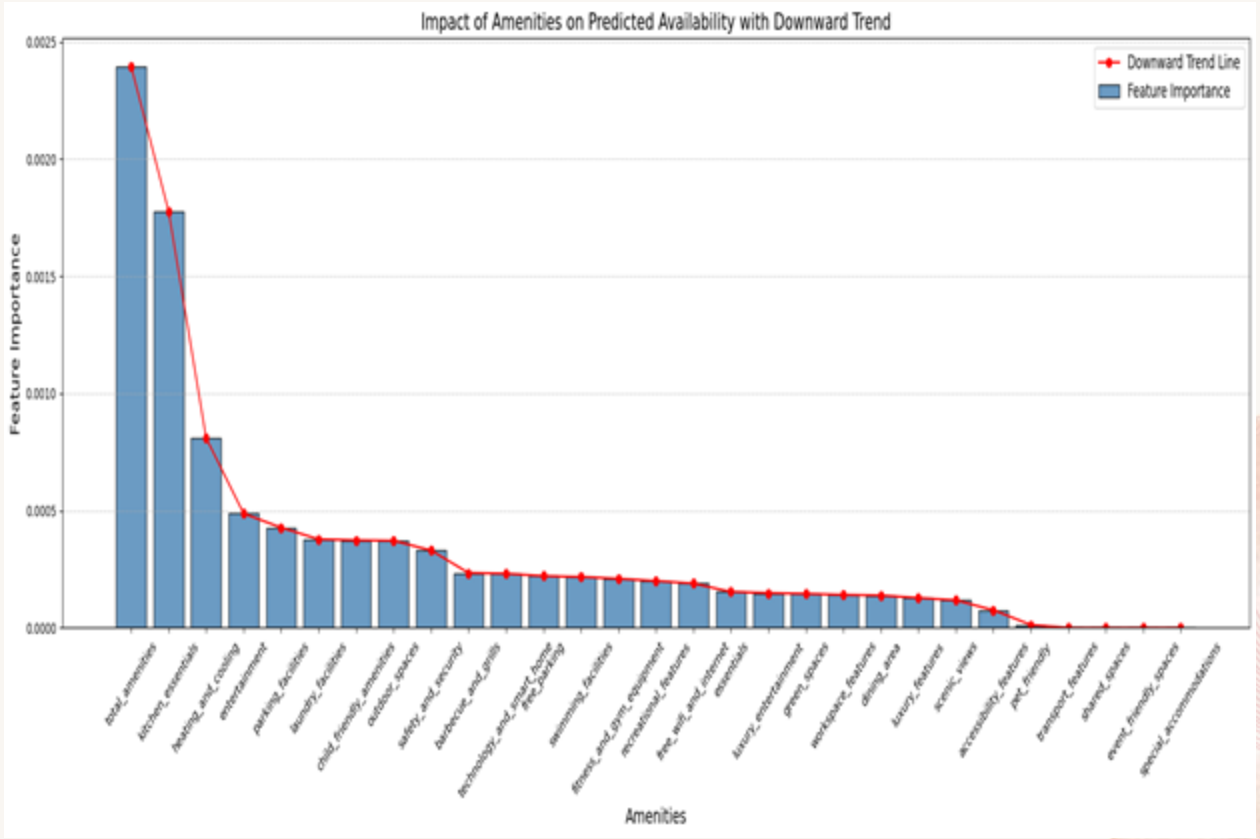
R Squared Value

Stakeholder Metrics

- Revenue prediction
- Pricing insights
- Location Recommendation
- Best Amenities
- Property type Recommendation

Stakeholder Metrics: Best amenities

To measure how different amenities impact **occupancy rates**, helping stakeholders prioritize features that drive bookings.



Stakeholder Metrics: Revenue Prediction

- Use price and occupancy predictions to forecast the host's revenue for the upcoming year
- This can enable hosts to make more informed decisions on future investments, repairs and planning
- Hosts can also use these revenue to forecasts to price their Airbnbs appropriately

Dynamic Pivot Table by ZIP Code

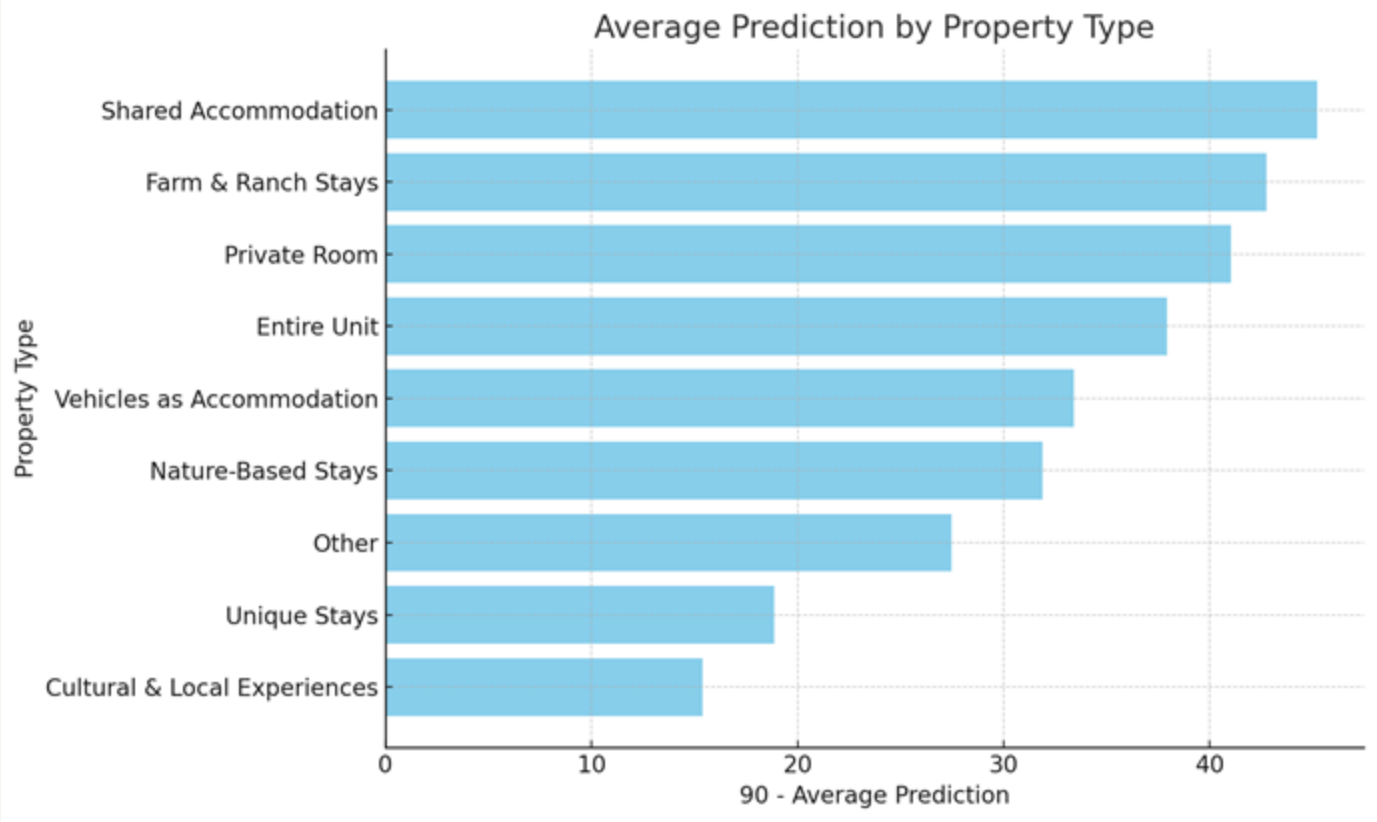
Select a ZIP Code:

90802

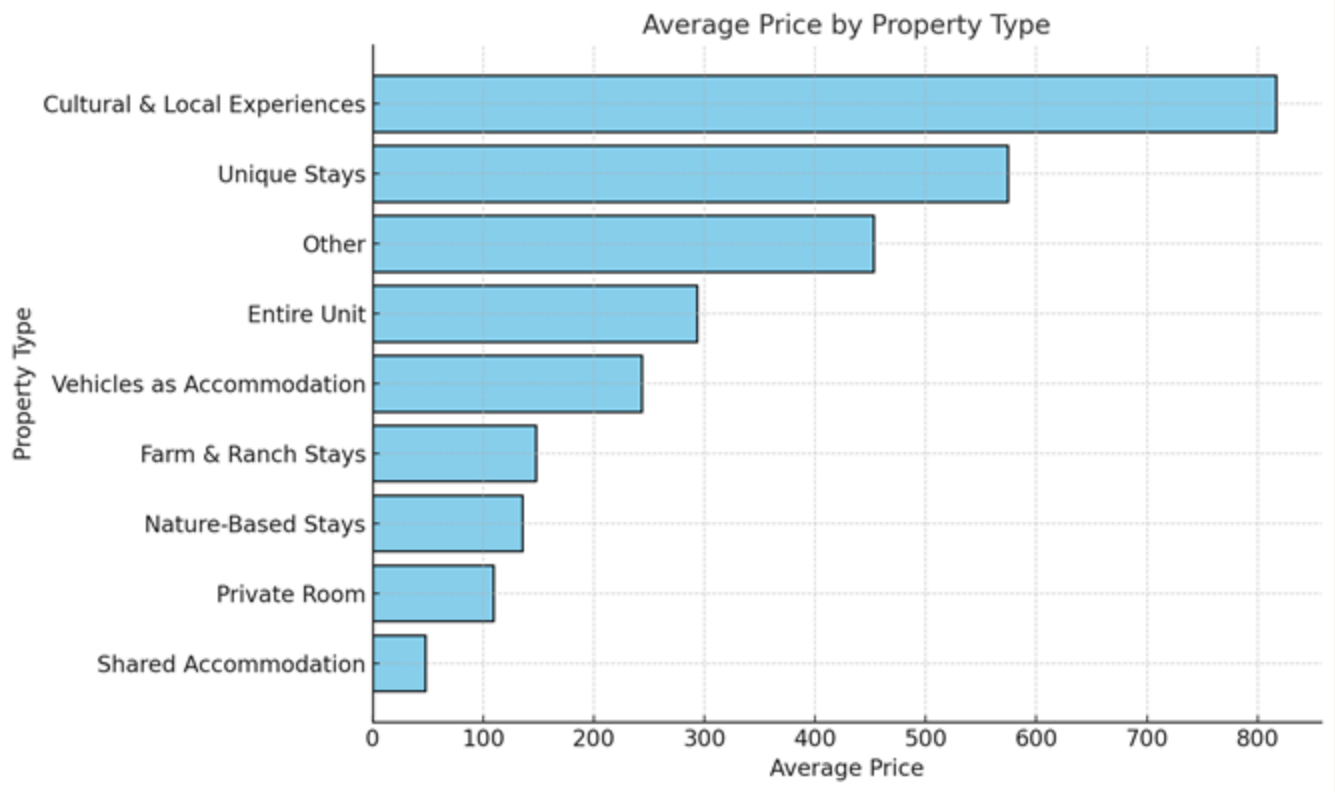
Pivot Table for ZIP Code: 90802

Property Type	Room Type	Bathroom Count	Average Revenue Generated
property_type_Entire_Unit	room_type_Entire_home/apt	0	785
property_type_Entire_Unit	room_type_Entire_home/apt	1	9156.948096885813
property_type_Entire_Unit	room_type_Entire_home/apt	2	8563.91489361702
property_type_Entire_Unit	room_type_Entire_home/apt	3	7390.25
property_type_Entire_Unit	room_type_Entire_home/apt	4	16496
property_type_Other	room_type_Entire_home/apt	1	9215
property_type_Other	room_type_Hotel_room	0	5920
property_type_Other	room_type_Hotel_room	0.5	0
property_type_Other	room_type_Hotel_room	1	5226.289473684211
property_type_Other	room_type_Private_room	0.5	123
property_type_Other	room_type_Private_room	1	49710.45
property_type_Other	room_type_Private_room	2	1320

Airbnbs with the highest occupancy



Avg price for Airbnbs with highest occupancy



Zip Codes that have high occupancy

listings_ZIPCODE	avg_prediction
93563	90.12414356290813
93543	89.93809587546657
90021	89.69905791328229
91108	87.77623024692038
90716	86.57814547436787
91020	86.36293418311
91201	85.23555708547323
90723	85.19833256311472
90002	85.17128506573172
90073	84.98617278591436

Future Scope

- Simulation Analysis with front end
- Use an Optimization algorithm to find the ideal pricing for an Airbnb to achieve maximum occupancy
- Expand our data to include other cities

Thank You!


HOW'S THE BIG DATA PROJECT COMING ALONG, HOSKINS?
© D.Fletcher for CloudTweaks.com