AIRBNB V'VER

PRICE MODELING

JULY, 2023

NIMA CHINIKAR

Table Of Content

Introduction

Real Estate investment is scary

--> Data and analysis to help decision making

PRICE PREDICTION USING MACHINE LEARNING REGRESSION

Data Collection

InsightAirbnb.com

'from publicly available information from the Airbnb site. The data has been analyzed, cleansed, and aggregated to facilitate public discussion.'

Description

- **Shapes**: 5975, 75
- **Property**
  - price
  - location
  - bedrooms
- **Host**
  - total listings
  - is verified
- **Availability**
  - per month, per year
- **Review**
  - count
  - score
- **Amenities**

Transform-ations

Amenities

[TV, Microware, Wifi, ...]

Missing Columns: +20

Feature engineering, research, mean, drop

Descriptive columns

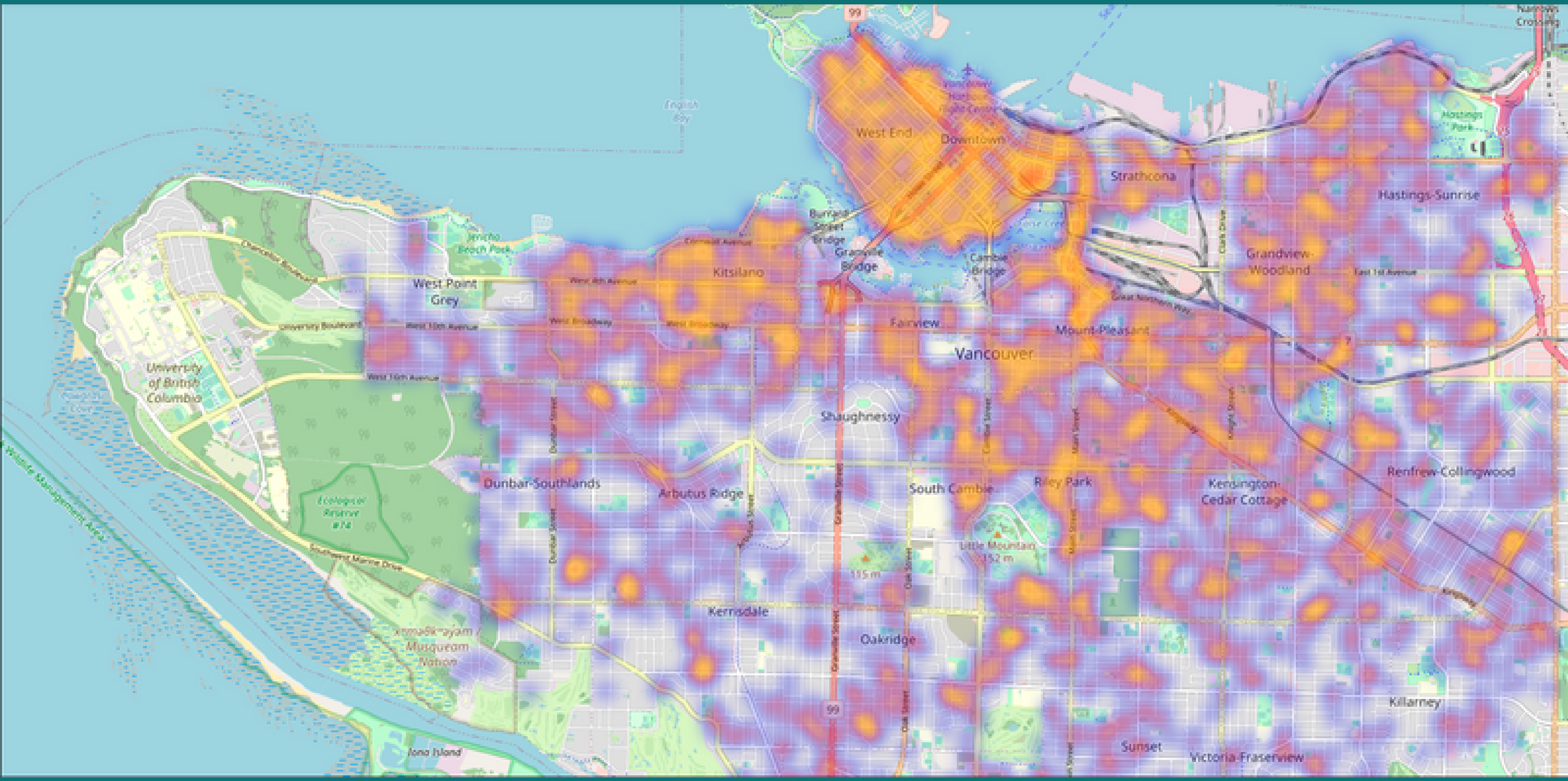
One-hot Encoding, order labeling

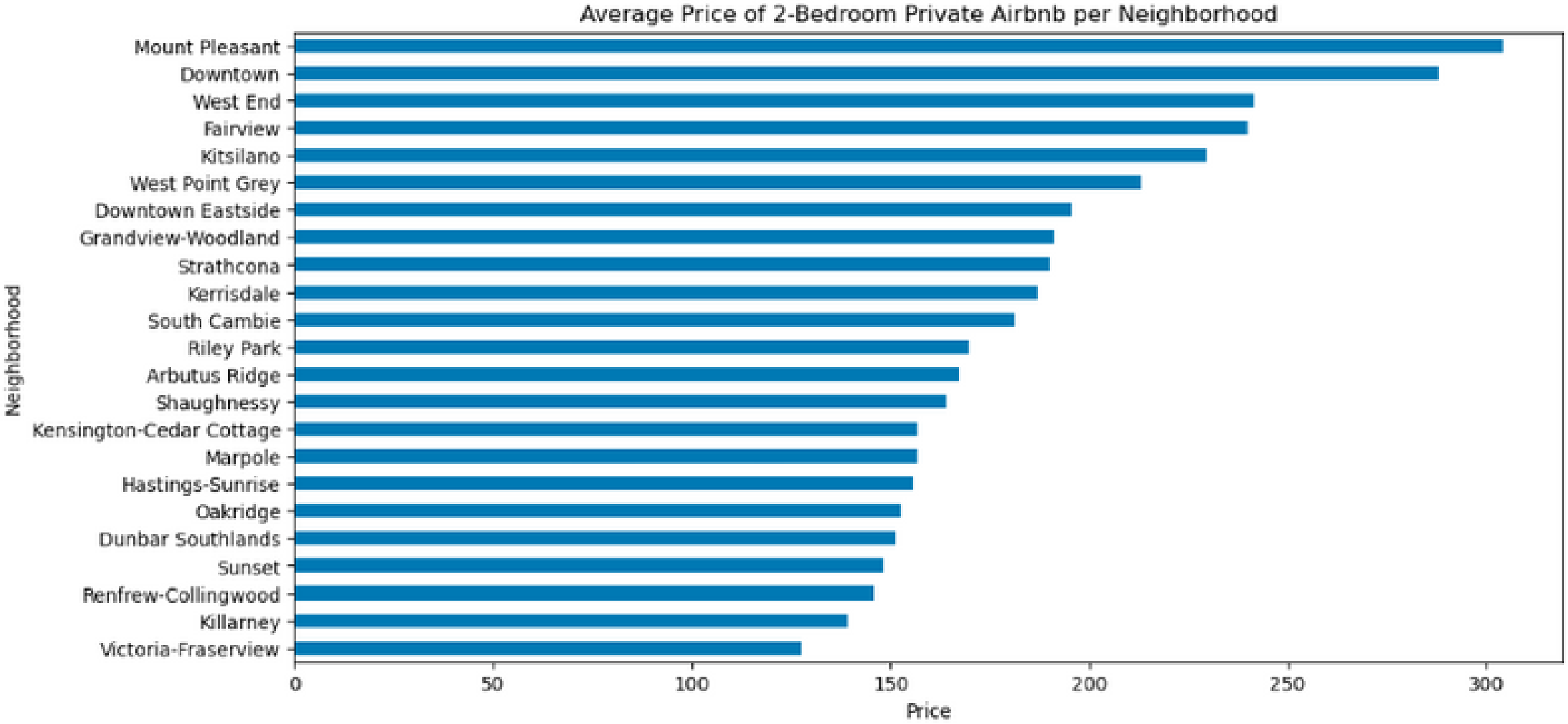
Normalized distribution

Transformation to log

EDA

- Airbnb Concentration
- Neighborhood Average Price


West End
Downtown
Strathcona
Hastings-Sunrise
Grandview-Woodland
Kitsilano
West Point Grey
University of British Columbia
Fairview
Mount Pleasant
Vancouver
Shaughnessy
Dunbar-Southlands
Arbutus Ridge
South Cambie
Riley Park
Kensington-Cedar Cottage
Renfrew-Collingwood
Kerrisdale
Oakridge
Killarney
Sunset
Victoria-Fraserview
Little Mountain
152 m
115 m
Burrard Street Bridge
Granville Bridge
Cambie Bridge
Jericho Beach Park
Hastings Park
Ecological Reserve #14
Musqueam Nation
Iona Island
English Bay
Point Grey Road
West 10th Avenue
West Broadway
Granville Street
Dunbar Street
Arbutus Street
Cambie Street
Main Street
Oak Street
Clark Drive
Knight Street
East 1st Avenue
99



The top 3 surely makes sense in Vancouver. They are all along the seawalk, which is a major touristic attraction

Modeling

- Linear Regression
- KNN
- Decision Tree
- Random Forest
- XGBoost

- Default
- Optimized

Models Comparison

	Mean Squared Error	R2 Score
Linear Regression - Default	879319141434740376076288O.00	-219956459898542354627297280.00
Ridge	0.18	0.56
Decision Tree - Default	0.29	0.28
Decision Tree - Best	0.21	0.46
Random Forest - Default	0.16	0.61
Random Forest - Best	0.16	0.61
KNN - Default	879319141434740376076288O.00	-219956459898542354627297280.00
KNN-Best	0.21	0.48
XGBoost - Default	0.15	0.62
XGBoost - Best	0.15	0.64

Interpreta-bility

Key Factors Affecting Prices

- Bathrooms
- Accomodation capacity
- Availability

Key Amenities Affecting Prices

- Dishwasher
- Shampoo
- Refrigerator
- Dedicated Workspace

Future Direction

- Expand the Dataset
- Refine the Model for Property Types
- Incorporate Neighborhood-Specific Trends
- Include External Factors

THANK YOU