# airbnb Toronto

Dongyan◦Fahad◦Huy◦Jeni◦Jordan◦Mariana◦Roma◦Ryan

# Agenda

---

Introduction

Data Description & Cleaning/Wrangling

Analysis Methodology

Results & Recommendations

Summary

# Executive Summary

**About us:**
Top housing platforms offering alternative and less expensive vacation rentals
Focus: Toronto, ON

**Opportunity:**
Identify which attributes most impact price and review scores

**Stakeholders:**
External: Hosts, renters
Internal: Marketing, Data engineers, Programmers
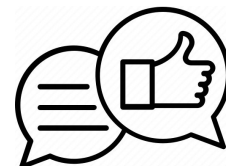
**Target:**
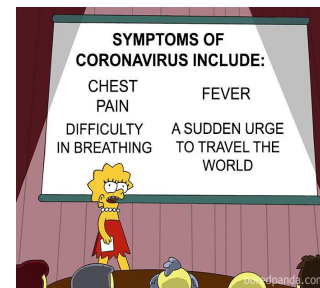
**2%** ⬆

Commissions

**Plan to increase Price:**
1. Offer review incentives
2. Recommended price adjustments

**Plan to improve Reviews:**
1. Audit listing accuracy where score is below 3

# Introduction

— — —



| 2008 | 2019 | 2020 | 2022 + |
|------|------|------|--------|
| **Established** | **Impact of Global Pandemic** | **Company goes Public** | **Opportunity to Increase Sales** |
| Online housing platform offering cheaper vacation retinal alternatives for travellers compared to more expensive hotel accommodations | Travel restrictions have little impact on airbnb sustainable business model | Business thriving, airbnb becomes a public entity | As post COVID economies in various sectors regain foothold, belief that demand for travel will flourish |

To maximize potential revenue, airbnb hosts can make use of attributes that drive booking price and improve these features for the benefit of travellers

**Goal: Increasing revenue by 2% through increasing commissions for airbnb**

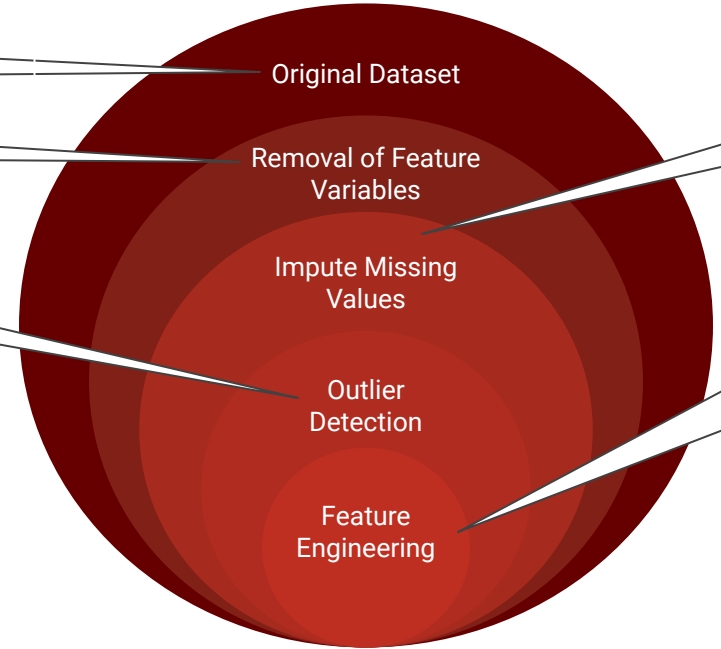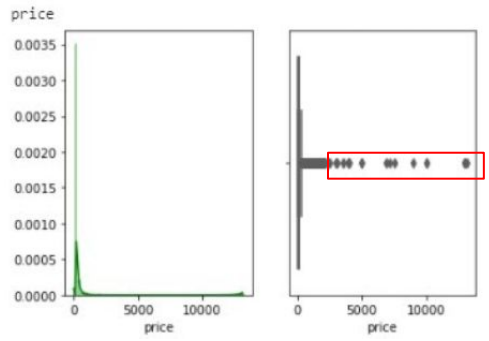Introduction → Description & Cleaning → Analysis Methodology → Recommendations → Summary

# Data Description Cleaning & Wrangling



Original Dataset

74 variables and 19,343 total records located in Toronto, ON only

SimpleImputer (Categorical) and IterativeImputer (numerical) from Sklearn Python used to impute missing values

Removal of Feature Variables

- Irrelevant data, i.e. listing_url
- Missing 40% or more of values

Impute Missing Values

- Host Since
- Host Response
- Host Acceptance
- Former City - grouped 70+ neighbourhoods into 6 cities (fig. below)

Interquartile range (IQR) method applied for detecting and removing outliers (fig. below)

Outlier Detection

Feature Engineering

# Analysis Methodology - Price Model

### 1  Feature Variables
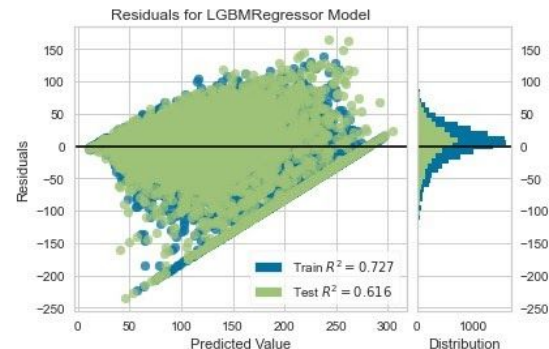


Feature Importance Plot
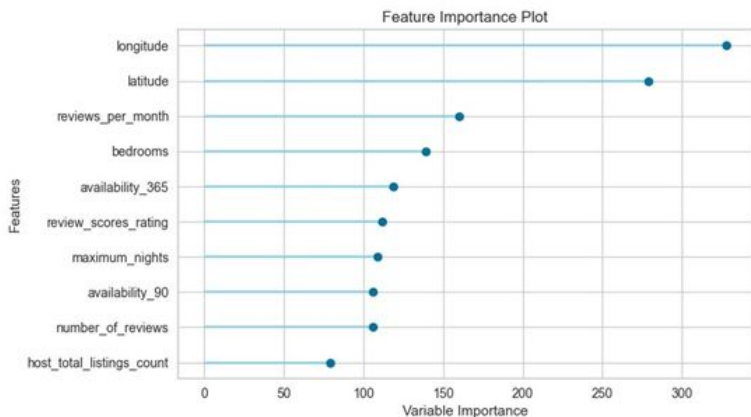
### 2  Algorithm/Model

- Test/train split (70/30)
- 10 folds for validation
- Light Gradient Boosting Machine
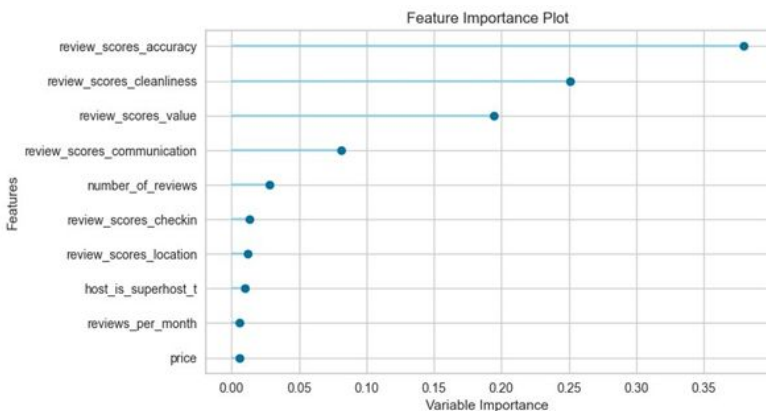- RMSE as chosen metric

### 3  Model Evaluation

- Overlap of test and training set residuals consistent
- No significant overfitting in the training set



Residuals for LGBMRegressor Model

Train $R^2 = 0.727$
Test $R^2 = 0.616$

# Analysis Methodology - Reviews Score Ratings Model

## ① Feature Variables



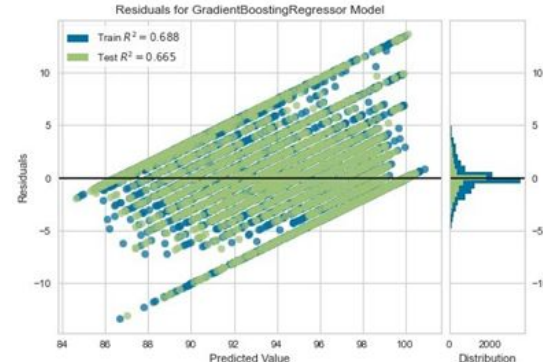Feature Importance Plot

## ② Algorithm/Model

- Test/train split (70/30)
- 10 folds for validation
- Gradient Boosting Regressor
- R-squared as chosen metric

## ③ Model Evaluation

- Overlap of test and training set residuals consistent
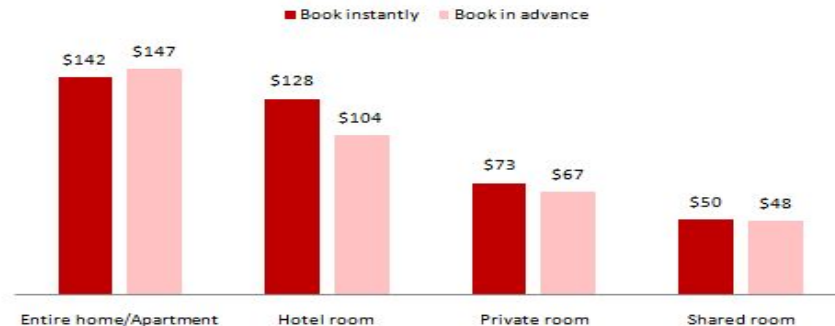- No significant overfitting in the training set



Residuals for GradientBoostingRegressor Model

Train $R^2$ = 0.688
Test $R^2$ = 0.665
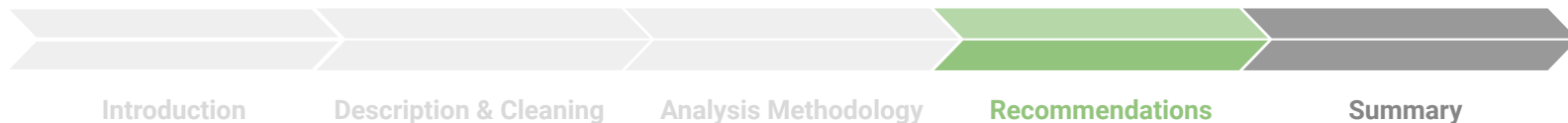
# Results & Recommendations: Price

### Average price per night vs. Number of reviews



### Ability to book instantly vs. Avg. price per night



1. Offer **incentives to renters** to leave reviews
2. Create an algorithm which provides the host with **guidance of adjusting the price based on the booking window**
3. Create an algorithm which generates guidance for hosts to identify if their set **price per night is within a standard/high/low range**

# Results & Recommendations: Review Score

— — —

Implement internal **audit/review** process when **accuracy score is low**, requiring corrective action from the host



Accuracy of listing vs. Review rating
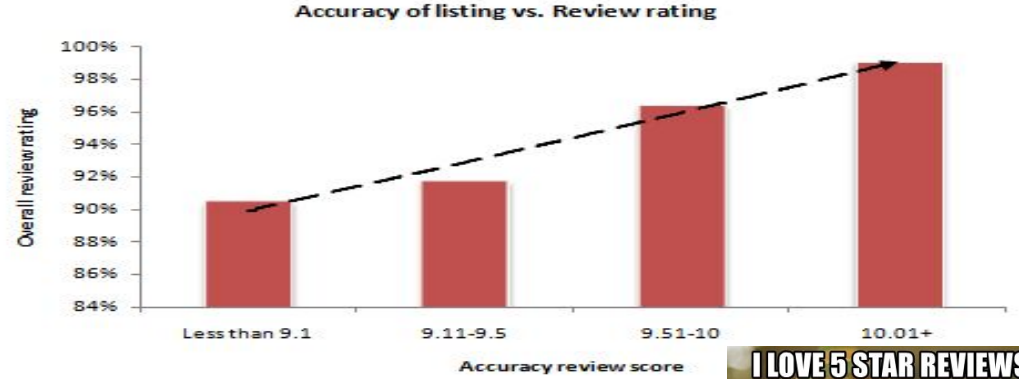


★ 5.0 · 25 reviews

| | | | |
|---|---|---|---|
| Cleanliness | 5.0 | Accuracy | 5.0 |
| Communication | 5.0 | Location | 5.0 |
| Check-in | 5.0 | Value | 5.0 |

# Summary

## Increase Price & Improved Review Scores = Added Commissions

Begin offering incentives for renters to leave reviews

Create algorithmic recommendations to hosts for price adjustments based on booking window and standard price for similar properties

Implement internal audit and review process when listing accuracy scores low

**Top Features**

Location

Number of reviews

**2%** ⬆

Commissions

**Top Features**

Instant booking

Listing accuracy

airbnb