

# Montreal Airbnb Price Prediction

## Introduction:

The vacation rental market in Montreal, dominated by platforms like Airbnb, presents a complex pricing challenge due to the diversity of listings and dynamic market conditions. Accurate price prediction is crucial for hosts to remain competitive and maximize their earnings while ensuring guest satisfaction.

## Project Background:

Montreal, a vibrant city with a fluctuating influx of tourists, has a varied demand for Airbnb rentals. Hosts often struggle to optimally price their listings, balancing profitability with attractiveness to potential renters. This project leverages machine learning to assist in making informed pricing decisions.

## Project Goal & Dataset Overview

### Project Goal:

The primary objective is to develop a predictive model that can accurately forecast Airbnb listing prices in Montreal based on multiple features, thus aiding hosts in pricing their properties effectively.

### Dataset Overview:

Data was sourced from Inside Airbnb, featuring detailed listings in Montreal. The dataset includes variables such as neighborhood, room type, amenities, and number of bathrooms, which are crucial for price determination.

### Data Preparation:

The data cleaning process involved:

- Removing duplicates and handling missing values.
- Transforming categorical data into machine-readable formats.
- Normalizing numerical data to ensure model accuracy.

### Exploratory Data Analysis (EDA):

Key findings from the EDA include:

- Significant price variations across different neighborhoods, indicating a strong geographical influence on pricing.
- Correlation between increased amenities and higher prices.

- Price sensitivity to room types, with entire homes/apartments commanding higher rates compared to private or shared rooms.

**Predictive Modeling:**

Models evaluated included Linear Regression, Decision Trees, and Random Forest. Each model was assessed based on its accuracy and ability to generalize from training data to unseen data.

**Results:**

The Random Forest model outperformed others, showing a high degree of predictive accuracy. Key metrics included Mean Squared Error (MSE) and  $R^2$ , with the latter indicating that the model could explain approximately 60% of the variance in rental prices.

**Future Work:**

Potential future directions include:

- Integrating real-time market data to refine predictions.
- Employing advanced neural networks to capture complex nonlinear relationships.
- Extending the model to incorporate user reviews and ratings for a more holistic approach to pricing.