The Canadian housing market is a complex system influenced by various factors such as location, amenities, and economic conditions in the area. The objective of this study is to explore and understand the factors influencing house prices in the top cities of Canada. By leveraging the programming language R and advanced data analysis techniques, we aim to uncover patterns, relationships, and trends within the dataset to provide valuable insights for prospective homeowners.

The primary problem addressed in this study is to analyze the Canadian housing market, specifically focusing on top cities, and identify the key factors that contribute to house price variations. Key research questions include:

1. What are the major determinants of house prices in Canadian cities?
2. How do economic indicators, such as median family income for the city, impact the housing market?
3. Are there anomalies or outliers in the dataset that require further investigation?

The dataset used for this study is sourced from Kaggle, titled Canadian house prices for top cities. (Jeremy Larcher. (2023). *Canadian house prices for top cities* Kaggle. https://doi.org/10.34740/KAGGLE/DSV/6825977)

To address the research questions and extract meaningful insights, we propose utilizing the R programming language and employing various data analysis techniques:

1. Classification: Identifying and categorizing factors that significantly contribute to variations in house prices.
2. Regression: Building predictive models to understand the relationships between independent variables and house prices.
3. Anomaly Detection: Identifying outliers or unusual patterns that may indicate unique market conditions.

By combining these techniques, we aim to provide a comprehensive analysis of the Canadian housing market, enabling homeowners to make informed decisions in the ever-evolving real estate landscape.

https://github.com/TracySchut/Final-Project-TMU-Data-Analytics