Problem Statement

Housing affordability in Canada has been and continues to be an important issue for Canadians. As housing prices rise at a rate that significantly outpaces income levels, the ability for an average Canadian to purchase a home is getting further away. As the gap between income and housing prices increase, the financial strain of those wanting to buy a home, or refinance with increased value is becoming considerably greater. Research by the Bank of Canada highlights the long-term escalation in house prices relative to income, revealing a persistent imbalance in the housing market (Bank of Canada, 2015) *[1]*

Why This is Important

As students preparing to enter the workforce the hope of owning your home and not be at the mercy of a landlord is becoming less likely. Also, for those of us who have already put years into our careers this option does not seem like a possibility. If the Disparity in income to housing prices can be accurately analyzed and presented, hopefully a more informed populus will start to act against these outrageous prices.

Dataset Descriptions

Dataset authors:

- 2025.csv is from Statistics Canada. *[2]*
- income.csv is from *Canadian Real Estate Association- CREA [3]*

House Price Dataset (df_house.csv)

- This dataset is an overview of house prices from 2005-2024. It has the data separated into HPI (house price index) and benchmark. For our purposes we will be using the Combined_HPI by Province.
- samples: 14640 rows
- features: 14 columns
- key characteristics:
    - this data set has the house prices broken down by area, type and HPI vs the benchmark

Median Income Dataset(income.csv)

- This dataset contains the median and average income for various age groups of all sexes. We will be using the median income as our baseline by province.
- samples: 9072
- features: 18
- key characteristics

- this data set contains Statistics based on income type, age groups, income (median and average), sex and location

These datasets collectively allow us to evaluate how housing prices compare to income growth across regions and over time.

Key EDA Findings

- Housing Price Trends: From 2005 to 2024, the average Composite HPI has steadily increased, with significant increases starting from 2020, indicating a large growth during the pandemic.
- Income Growth Trends: Median income has not kept pace with the rise in housing prices, with minimal year-over-year increases observed in many provinces.
- Geographical Trends: Provinces like Ontario and British Columbia exhibit the highest housing price indices, while Atlantic provinces and Manitoba show relatively lower indices but still face income-price gaps.
- Data Issues: Missing values for specific demographics (e.g., age groups 65+) were identified and addressed by excluding non-relevant data.

Tools and Methods

- Data cleaning and wrangling: Python (pandas, NumPy)
- Visualizations: Seaborn, Matplotlib
- Geospatial Analysis: GeoPandas *[5]*, Folium *[6]*
- Predictive Modeling: Scikit-learn *[7]*
  - Using linear regression models, projected values for 2023–2024 show a continued divergence between median income and housing prices.
- canada_provinces.geo.json *[4]*
  - This is a tool to help visualize our data over the actual map of Canada using a heatmap distributed over the provinces

Continued Work

My proposed expansion for this project is as follows:

- Finish the comparisons YoY for Income vs House Prices
- Expand the visualization of the geo heatmaps to include greater detail, including city and income statistics
- Use financial management tools to show what percentage of an income is meant to be spent on Housing and how an individual in today's market would need to adapt their spending to achieve this
- Look deeper into difference between types of housing and overall cost in each area
- Find a scholarly paper to support my position statement

References

1. Bank of Canada. *The long-term evolution of house prices*. (2015). [Online]. Available: https://www.bankofcanada.ca/2015/08/long-term-evolution-house-prices/
2. Statistics Canada. *Table 11-10-0239-01: Income of individuals by age group, sex, and income source*. DOI: 10.25318/1110023901-eng
3. Canadian Real Estate Association. *MLS® Home Price Index (HPI) tool*. [Online]. Available:
4. canada_provinces.geo.json. [Online]. Available: *https://gist.github.com/M1r1k/d5731bf39e1dfda5b53b4e4c560d968d#file-canada_provinces-geo-json*
5. GeoPandas: GeoPandas developers. *GeoPandas: Python tools for geographic data*. (2023). [Online]. Available: https://geopandas.org/.
6. Folium: Folium developers. *Folium: Python Data, Leaflet.js Maps*. (2023). [Online]. Available: https://python-visualization.github.io/folium/.
7. Scikit-learn: Pedregosa, F., [Online]. Available: https://scikit-learn.org/.