Economic Drivers, Immigration Dynamics, and Housing Starts: Their Collective Influence on Housing Price Index*

Navya Hooda

April 18, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

In recent years, the Canadian housing market has undergone significant changes due to various economic factors. Canadians are facing a housing crisis driven by high demand, unaffordable rates, and population growth in major cities across Canada. For example, Toronto's benchmark home price has surged by 42.8% since 2019. Back in 2019, the benchmark home price in Toronto stood at $746,500. Fast forward to 2024, and that figure has soared to $1,065,800, according to the Canadian Real Estate Association (CREA).

One major factor contributing to these changes is the notable increase in immigration over the last decade. In 2023, Canada experienced its highest influx of immigrants and students in 70 years, leading to a 2.9% surge in the population (Reuters 2023). While achieving immigration targets is a milestone for Canada, the influx of newcomers has likely exacerbated the already tight housing market.

Globally, economic shifts in recent years have had far-reaching effects. Inflation has risen not only in Canada but also in many countries worldwide due to various events. Supply chain disruptions on a global scale have driven prices higher, while geopolitical tensions, such as Russia's invasion of Ukraine, have further contributed to increased energy and food costs (Forbes Advisor Canada 2024).

In response to rising inflation, Canada saw its inflation rate peak at 6.1% in 2022, gradually easing to 3% in 2023. However, the target rate for stability is around 2% (Bank of Canada

*Code and data are available at: https://github.com/hoodanav/Housing_Market_Analysis/tree/main

2024). To curb inflation, the Bank of Canada implemented a series of interest rate hikes between March 2022 and July 2023. This aggressive monetary policy campaign saw the benchmark rate soar from 0.25% to 5%, marking one of the most substantial increases on record(The Globe and Mail 2024).

With so many factors at play, it is hard to keep track of just how the housing market is evolving. This analysis aims to explore the dynamics of the housing market and understand the impact of economic variables on the housing market, analyze if any indicators are more significant contributors to the housing market than others etc. Through this exploration we also aim to explore economic indicator changes and what significance they have to the housing market, the relationship between correlated variables, and their impact on the market.

This paper also explores the housing supply in specifically Ontario due to its increased housing activity in the past few years. We explore how the housing demand has changed over the last 20 years, and whether the current housing development initiatives are enough to contribute positively to the improvement of the housing crisis. Additionally, we touch upon government policy on housing capacity, and targets that are currently in motion.

With the complexity of the housing market influenced by numerous factors, keeping track of its evolution proves challenging. This analysis aims to understand the dynamics of the housing market, exploring the impact of economic variables and identifying significant contributors. By examining economic indicator changes, we seek to uncover their significance and understand correlations among variables and their market impact.

Focusing particularly on Ontario's housing landscape, where activity has surged in recent years, this paper investigates shifts in housing demand over the past two decades. We evaluate whether current housing development initiatives adequately address the housing crisis. Additionally, we discuss government policies regarding housing capacity and existing targets.

Through this analysis, we aim to grasp the current reality of the housing market and understand how economic conditions, including factors like immigration and housing starts, influence its trajectory. With the housing market facing constraints such as high interest rates, we seek to uncover insights regarding the future of the market.

The remainder of this paper is structured as follows.In the Data@sec-data, we note the sources of the data sets utilized in the paper, the specific techniques we used to process them, and key variables we wish to use. In the Model Section 3, we created models to assess the impact of the key variables on our outcome of housing prices. In the Results Section 5.1 we discuss the results of the model and what key correlations were found from the variables studied in the paper. In the Discussion Section 6, we explore the findings of modeling and predictions. In addition, we discuss biases and weaknesses in the data that contributed to our findings, and how we navigated those limitations.

2 Data

2.1 Data Source and Collection

The data sets used for this analysis have been picked from multiple sources, open government portals and economic institutions. By choosing a variety of data sets, many cleaning and alignment processes were applied to ensure as much consistency possible between the data sets before usage.

2.1.1 Inflation Rate

The Inflation Rate Data was found from Macrotrends (Macrotrends 2022a) which published a concise summary of the inflation rate (Consumer Price Index focused) changes from 1960-2023 sourcing its data from the World Bank. Inflation rate, from a Consumer Price Index (CPI) perspective, represents the percentage change in the average price level of a basket of goods and services typically consumed by households over a specific period. It indicates the rate at which the general level of prices for goods and services is rising, leading to a decrease in purchasing power.

2.1.2 Interest Rate

The interest rate used for our analysis is specifically the Canada Mortgage and Housing Corporation, conventional mortgage lending rate, 5-year term rate, and was sourced from Statistics Canada (Canada Mortgage and Housing Corporation 2024). The data was reported for years 2000-2024 by each month. The conventional mortgage lending rate, specifically the 5-year term, refers to the interest rate that financial institutions charge for conventional mortgages with a fixed interest rate and a term of 5 years. This rate is significant in the housing market as it influences the cost of borrowing for home buyers and impacts the overall affordability of housing. For our purposes we chose to aggregate data to use an average over all months to gain a yearly interest rate number to explore.

2.1.3 Unemployment Rate

The unemployment rate was found from Macrotrends (Macrotrends 2022b) which gave summarized year round rates, initially sourced from the World Bank. The unemployment rate generally refers to the percentage of the labor force that is unemployed and actively seeking employment. A visual of the rates discussed so far is seen in Figure 1.

2.1.4 New Housing Price Index (HPI)

The new housing price index data was found from Statistics Canada (Statistics Canada 2024b), and was reported by each month for each year from 2000-2024 specifically covering Ontario. The new housing price index is a value that is a monthly series that measures changes over time in the contractors' selling prices of new residential houses given a specific start date (typically with HPI = 100, currently this refers to December 2016). The data is reported monthly on NHPI, a survey that collects the following dwelling types: new single homes, semi-detached homes and town homes (row or garden homes). The survey also collects builders' estimates of the current value (evaluated at market price) of the land. The data reports HPI for land type, house type, and total type. For our purposes we chose the house type parameter and sought to calculate a yearly HPI given the data was sorted by months. The trend can be seen in (Figure 4)

2.1.5 Immigration Numbers

The yearly immigration numbers were obtained from Statista (Statista 2024), and was originally sourced from Statistics Canada. This data reports the number of immigrants Canada welcomes in each year from 2000-2023. The data was collected and sourced from a survey done by Statistics Canada. A visual of these numbers is seen in **?@fig-immigration-plot**.

2.1.6 Housing Starts (Ontario)

The housing starts data was found from Statistics Canada (Statistics Canada 2024a), was reported quarterly from years 2000-2024. This data refers to the number of units of housing started across Ontario of different dwelling types, in various regions. We chose to look at the housing starts of total units for all dwelling types, a yearly number by summing up each quarter's progress. For the 2024 value only quarter 1 numbers are reported hence the sharp decline for 2024 on Figure 2

2.2 Data Cleaning

Simple data cleaning operations were applied to achieve a cohesive merged data set for analysis. We used R (R Core Team 2023) for data cleaning and processing, using packages like tidyverse (Wickham et al. 2019) for data manipulation and janitor (Firke 2023) for cleaning column names. Other packages used includes `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `readr` (**read?**), `tibble` (Müller and Wickham 2023), `janitor` (Firke 2023), `knitr` (Xie 2023), `ggbeeswarm` (**gg?**), `ggrepel` (Slowikowski 2024), `kableExtra`(Zhu 2024), `readxl`(Wickham and Bryan 2023), `rstanarm`(Goodrich et al. 2022), `modelsummary`(Arel-Bundock 2022) and `here` (Müller 2020). All these different data sets were cleaned in a different way to match the final representation of the cleaned data.

2.3 Data Modifications

Overall, the values were often reported at different times, for consistency we chose to use a yearly metric. We averaged upon quarterly or monthly reported data to gain an averaged yearly value.

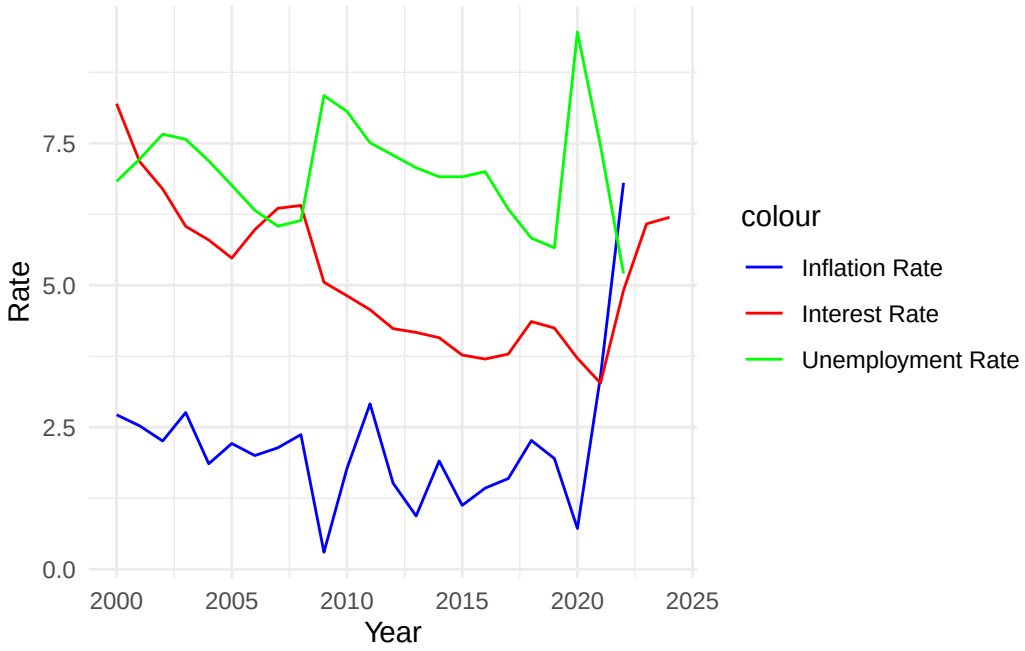
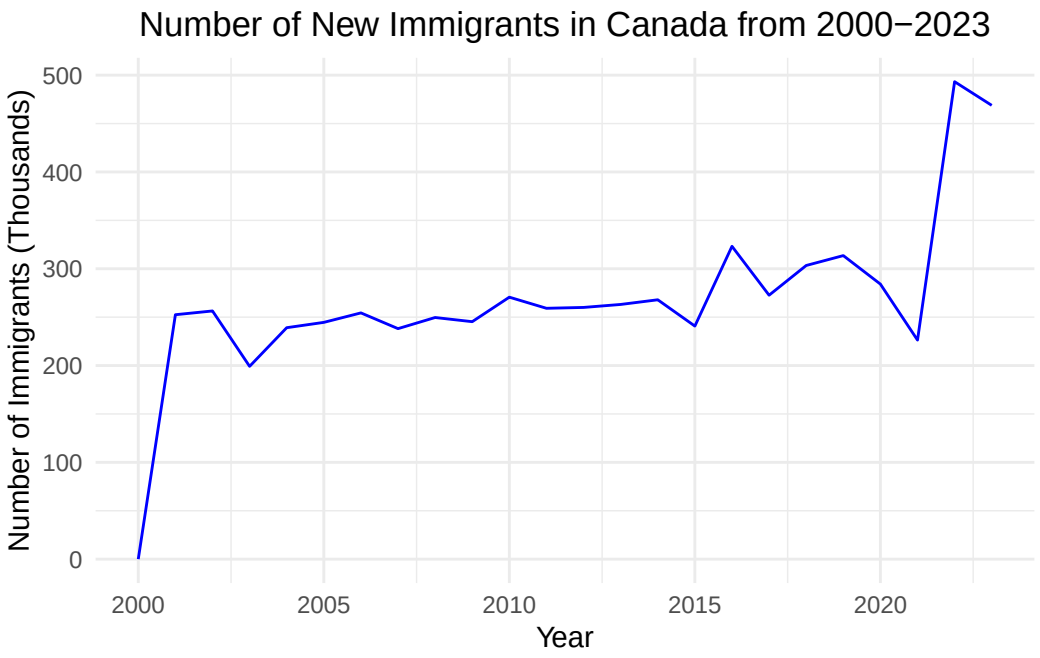


Figure 1: Inflation, Interest and Unemployment Rates 2000-2023

```
# Load the necessary packages
#| label: fig-immigration-plot
#| fig-cap: Number of immigrants in Canada from 2000 to 2023
#| echo: false
#| warning: false
#| message: false
# Filter out rows with missing values in the 'Number.of.immigrants.in.Canada.from.2000.to.
merged_data_filtered <- merged_data[!is.na(merged_data$Number.of.immigrants.in.Canada.from

# Create a line plot for the number of immigrants in Canada over time
ggplot(merged_data_filtered, aes(x = Year, y = `Number.of.immigrants.in.Canada.from.2000.t
  geom_line(color = "blue") +
  labs(title = "Number of New Immigrants in Canada from 2000-2023",
       x = "Year",
```

```
     y = "Number of Immigrants (Thousands)") +
theme_minimal() +
theme(plot.title = element_text(hjust = 0.5)) +  # Center the title
scale_y_continuous(labels = scales::comma_format(scale = 1e-3), limits = c(0, NA))  # Fo
```



3 Model

4 Calculate predicted values

predicted_values <- predict(model_summary)

5 Create a data frame with observed and predicted values

plot_data <- data.frame(Observed = model_data$House_Price_Index, Predicted = predicted_values)

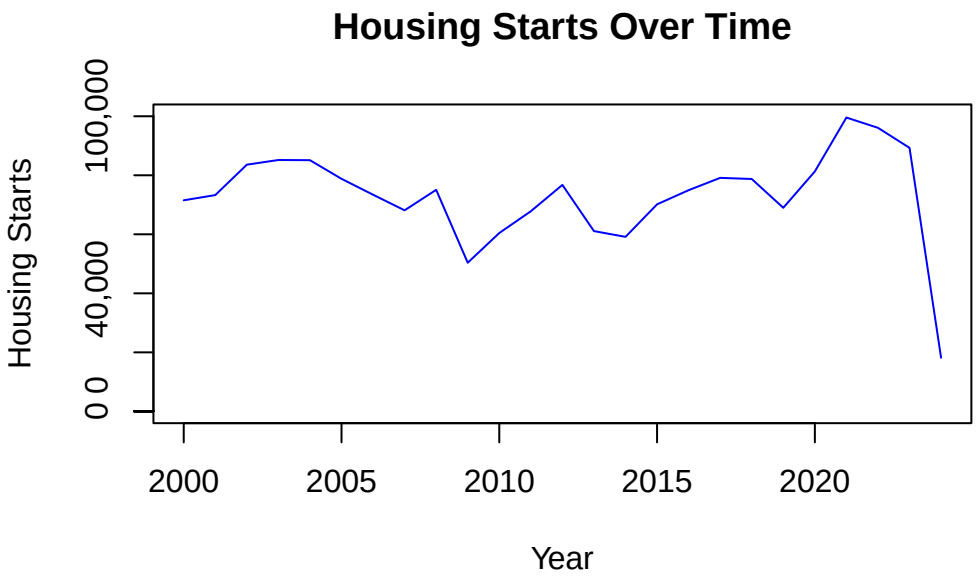


Figure 2: Housing Starts Over 2000-2023

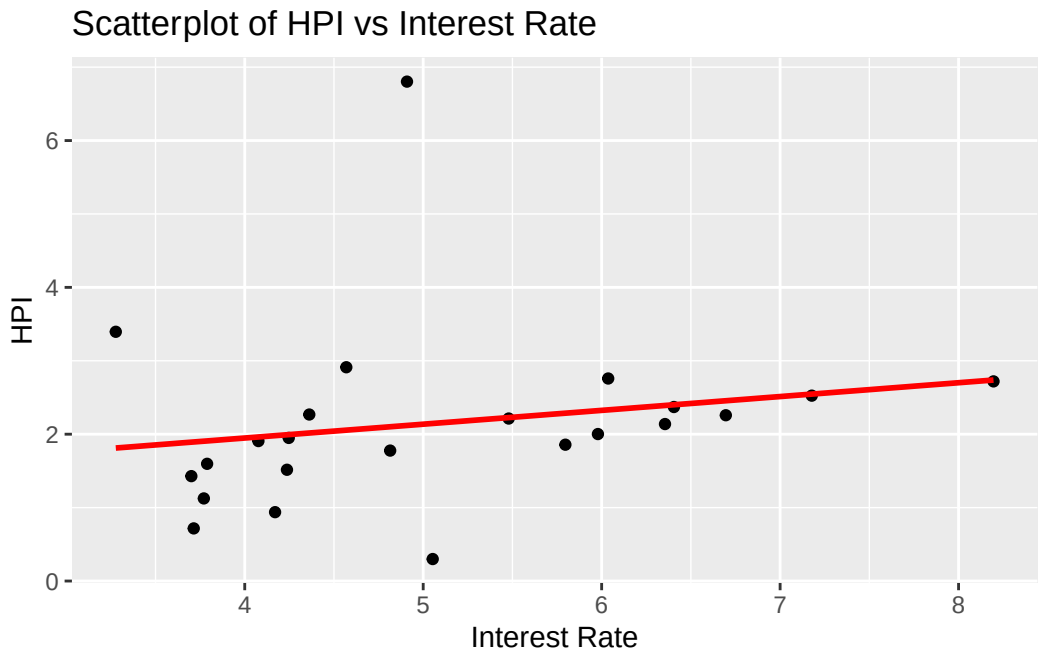


Figure 3: Number of immigrants in Canada from 2000 to 2023

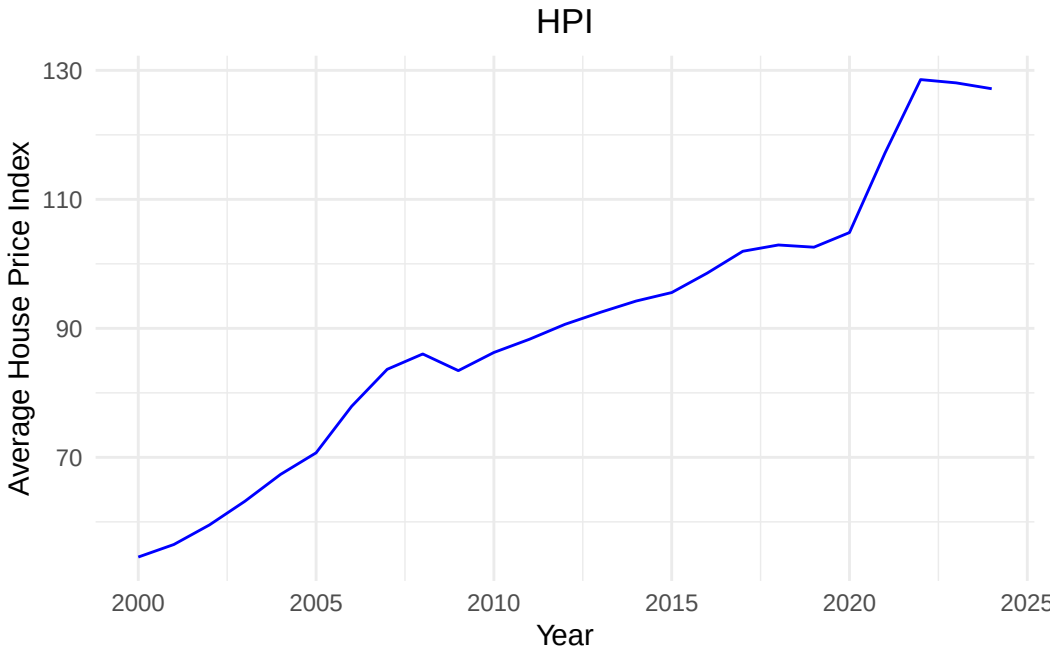


Figure 4: House Price Index 2000-2023

5.0.1 Model set-up

The goal of our modeling strategy is to analyze the relationship between the economic predictors and the housing price index (HPI), and explore how the economic variables contribute to the housing market. After exploring some of the intial data we observed a general trend between different variables, that indicated some level of relationship between them. To see the actual relationships, we implement a multiple linear regression model to take a close look at what variables seem to have an effect on the outcome variable, denoted by the house price index.

5.0.2 Multiple Linear Regression

We chose to explore a variety of economic and count variables like immigration and housing starts, to see how these impact the housing prices we use multiple linear regression model.

The basic form of the regression equation is:

The linear regression equation for predicting the House Price Index is given by:

House_Price_Index = _0 + _1 × Inflation_Rate + _2 × Interest_Rate + _3 × Unemployment_Rate + _4 × Immigration_numbers + _5 × Housing_Starts +

Where:

- House_Price_Index is the response variable (dependent variable).
- β_0 is the intercept term.
- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are the coefficients (regression coefficients) corresponding to each predictor variable.
- Inflation_Rate, Interest_Rate, Unemployment_Rate, Immigration_numbers, Housing_Starts are the predictor variables (independent variables).
- ε represents the error term.

5.0.3 Model Justifications

The above factors were chosen as they are economic indicators of the housing market, and are often the variables that have influence in the housing market given the market conditions. For inflation and interest rates, there is a cyclic relationship, as inflation goes up, the interest rates on mortgages go up as a direct consequence. This in turn usually increases monthly housing costs, detering buyers from qualifying for loan approvals. Unemployment rates are correlated to small spending. Given Canada's housing crisis, immigration numbers and housing starts seem to be changing supply and demand, as a result these indicators can be conclusive.

We expect a negative linear relationship between interest rates and housing price index, housing starts and house price index, and unemployment rate and house price index. For Immigration we expect a somewhat positive linear relationship to house price index. The increasing pace of immigration has been fueling demand and constricting the market. These relationships are estimated based upon economic patterns and past evidence.

5.1 Results

5.1.1 Model Summary

Our multiple linear regression model is summarized in **?@tbl-model-sum**. Our multiple linear regression model reveals the strength and association of each indicator to the outcome variable, while holding other variables constant. The estimated coefficients in the model indicate the magnitude and direction of each variable's association. Indicators that had a positive relationship include the inflation_rate where each unit of increase in its value correlated to a 5.398 unit (estimate coefficient) increase in house_price_index.

The variables interest rate and unemployment_rate had a negative association to the house_price_index, as expected.For each unit increase of the interest rate (percent) the house price index saw a 12.45 unit decrease. The unemployment rate had estimate coefficient -2.812

indicating that as the unemployment rate changed by a unit the house price index saw a 2.812 unit decrease. However, here the relationship is not statistically significant as $p > 0.05$.

The housing_starts suggests a small negative effect on the house price index, but is not statistically significant. Similarly, the immigration_numbers saw a small positive effect on the house price index, but it is not statistically significant. Considering the nature of these variables, they are count values, hence a change in one unit (one housing start increase or one immigrant number increase) is not expressed significantly by this model.

The model demonstrates an overall good fit with an R-squared value of 0.8522, indicating that approximately 85.229% of the variance in house price index is explained by the included predictors. The adjusted R-squared, which accounts for the number of predictors in the model, is slightly lower at 0.8088, suggesting the model is fairly robust. The residual standard error is 8.364, suggesting the amount of variability of observed responses around the fitted regression line (plot).

6 Discussion

6.1 Weaknesses and Biases

The data sets used for the analysis were combined and explored, these data sets may have added certain biases to our overall findings. There are a number of data sets that had missing values, and most of the data had to be modified to be compatible for use against the other indicators, we also note the differences in their units reported. The modelling process is prone to adding bias to our findings due to model pre-assumptions, and whether the model may have overlooked trends due to its limitations. The regression model suggested immigration numbers and housing starts as not significant indicators of the house price index. However, from a economic perspective these variables should have affected the price index as they represent supply and demand variables.

Since Our model could not grasp the nature of its data and showed a very minute relationship. To show some key insights, transforming these variables to a concise unit of change could help highlight their importance, or additionally using already aggregated data could help the regression model uncover associations more meaningfully.

6.2 Future Research

The analysis we covered only explores a handful of housing market predictors, in the future expanding the preidctors and ading a variety of variable types can help build more intuitive models for udenrstanding. To gain a new perspective on the housing market, indicatros relation t demographics, socieconmic factors, how many buyers in the market can comfortably afford mortgages and loans. This would build a foundation for a deeper analysis, and an opportunity

to learn more about the housing market and whether Canada is headed in the right direction in the coming years.

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Bank of Canada. 2024. "Inflation: Bank of Canada." https://www.bankofcanada.ca/core-functions/monetary-policy/inflation/.

Canada Mortgage and Housing Corporation. 2024. "Canada Mortgage and Housing Corporation, Conventional Mortgage Lending Rate, 5-Year Term." https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3410014501.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Forbes Advisor Canada. 2024. "Why Are Interest Rates so High?" https://www.forbes.com/advisor/ca/investing/why-are-interest-rates-so-high/#:~:text=Many%20factors%20contributed%20to%20the,inflation%20to%20spike%20to%205%25.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Macrotrends. 2022a. "Canada Inflation Rate." https://www.macrotrends.net/global-metrics/countries/CAN/canada/inflation-rate-cpi#:~:text=Canada%20inflation%20rate%20for%202022,a%200.32%25%20decline%20from%202018.

———. 2022b. "Canada Unemployment Rate." https://www.macrotrends.net/global-metrics/countries/CAN/canada/unemployment-rate#:~:text=Canada%20unemployment%20rate%20for%202022,a%200.17%25%20decline%20from%202018.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://tibble.tidyverse.org/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reuters. 2023. "Immigrants, Students Fueled Canada's Population Jump 2023: StatsCan." https://www.reuters.com/world/americas/immigrants-students-fueled-canadas-population-jump-2023-statscan-2023-09-27/.

Slowikowski, Kamil. 2024. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*. https://ggrepel.slowkow.com/.

Statista. 2024. "Number of Immigrants in Canada." https://www.statista.com/statistics/443063/number-of-immigrants-in-canada/.

Statistics Canada. 2024a. "Canada Mortgage and Housing Corporation, Housing Starts, Under Construction and Completions, All Areas, Quarterly." https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3410013501&pickMembers%5B0%5D=1.8&pickMembers%5B1%5D=4.1&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2000&cubeTimeFrame.endMonth=01&cubeTimeFrame.endYear=2024&referencePeriods=

20000101%2C20240101.

———. 2024b. "New Housing Price Index, Monthly." https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810020501.

The Globe and Mail. 2024. "Bank of Canada: The Globe and Mail." https://www.theglobeandmail.com/topics/bank-of-canada/#:~:text=The%20central%20bank%20raised%20interest,the%20past%20five%20rate%20decisions.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2023. *Readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.