

**CAPSTONE PROJECT**

Assignment 1

Analyzing historical rent prices in Canada

Professor: Samer Al-Obaidi

Submitted on 2nd June 2024

**Group: Ottawa**

**Team Members**

|  |  |
| --- | --- |
| **NAME** | **Student ID** |
| Tanya Nimesh | N01579781 |
| Yshika Shrikhande | N01545656 |
| Apeksha Hipparagi | N01579432 |
| Shrutika Desai | N01580471 |
| Gurleen Banga | N01579137 |

**Introduction**

The aim of this project is to analyze historical rent prices across different regions in Canada with population over 10,000, forecast future rent prices, considering several types of housing structures and units. Based on the identified rent prices dataset, the project aims at looking for uncovered trends, patterns, and insights which help real estate stakeholders, policymakers, and business owners.

The analysis is focused on the region-based rent prices and type of housing such as Apartments, Houses, Condos across the country. We will initiate the project by acquiring the data and move forward with its exploratory data analysis so that we can develop a perfect errorless database to build predictive models on. We will end the process by shaping our data into visualizations to make it easier for the audience to interact, understand and draw actionable insights from it.

To be more concrete, the project will highlight changes that have taken place over time, providing a precise understanding of market trends. This project analyzes future rent prices in various geographic areas to assist policymakers and stakeholders in understanding and putting the necessary initiatives into business. It will be easy to determine the causes of high-price demand by looking at factors that impact rental prices, such as population, economic conditions, and climatic factors. For instance, rent is less expensive in areas like Alberta and Saskatchewan than it is in British Columbia's Ontario.

### **PROJECT IMPORTANCE**

**Real Estate owners:** This project is especially important for real estate investors, developers, and property managers to make informed decisions regarding rent prices. This analysis gives information about the market environment to aid the investors, developers, and managers in making essential decisions pertaining to properties.

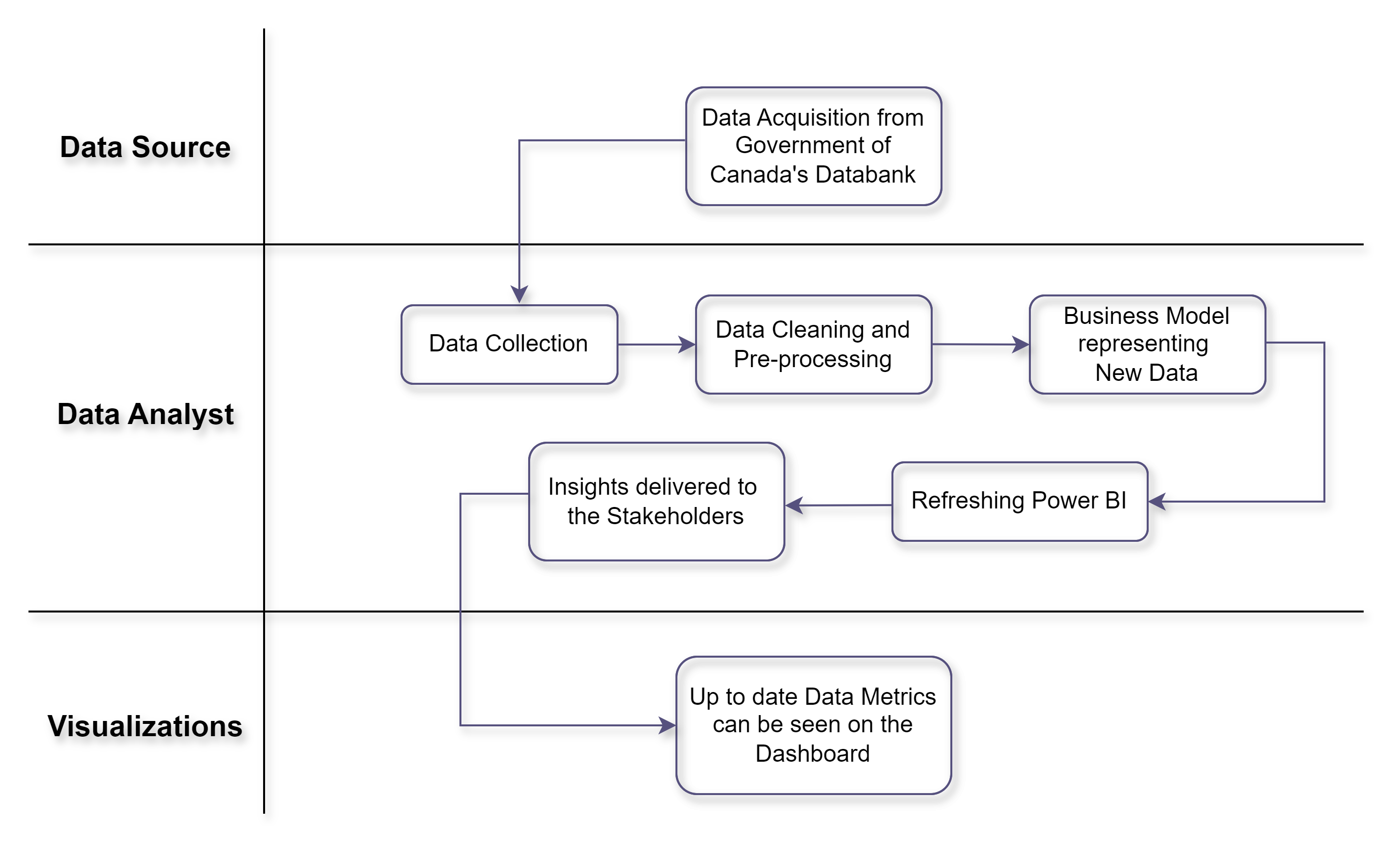
**Policymakers:** Policymakers can use the findings in formulating policies that can help work out strategies that address issues touching on housing affordability and set out strategies to develop sustainable urban centers. This analysis of the data will help identify the locations experiencing a high rise in rent price, to apply specific measures to prevent housing from becoming a preserve of a certain class.

**Business Insights and Analytics:** This project is a testimony of how information can be loaded with potential, translated from formats of unorganized data into solutions. It elucidates how firms can utilize big data to manage the performance of their operations competitively and strategically.

**Forecasting and Planning:** Predicting rent prices is useful in managing financial planning and risks. Businesses can enhance their ability to predict market shifts, enabling them to adopt proactive solutions instead of reactive ones. This proficiency is crucial for sustaining stability and fostering growth amid market volatility.

**Technology Integration:** The project highlights the importance of integrating advanced analytics and visualization tools such as Power Bi in business operations. It highlights how crucial it is to implement technologies like interactive dashboards and predictive modelling, which enhance data interpretation to stakeholders.

### **Process Flow Diagram**

**Involved Systems:**

1. **Data Sources:** CSV files and external databases have information on rent price.
2. **Data Processing System:** Tools for cleaning and preparing data by extracting, transforming, and loading (ETL) it.
3. **Analysis Tools:** R and Python notebooks are examples of tools for data analysis and predictive modelling with statistics and machine learning.
4. **Visualization Tools:** Tools for producing in-depth reports include Jupyter notebooks and Power BI.
5. **Reporting System:** Tools for producing comprehensive reports, including Word documents and PDFs.

**Process Description**

*The process is initiated by Collecting the data and preparing it through the following steps:*

**Step 1: Data Acquisition:** The data analyst oversees the project's first stage, which is called data collection. At this point, data is collected from several sources, most notably CSV files. This stage culminates in the compilation of the raw datasets. The process concludes after all pertinent data has been successfully collected.

**Step 2: Data Cleaning:** The Data Analyst leads the Data Cleaning phase. During this stage, errors in the datasets are fixed, redundant information is removed, and missing values are managed. The outcome is a set of datasets that have been cleaned. This process concludes when the data is clean and consistent.

**Step 3: Data Transformation:** The Data Transformation step is managed by the Data Analyst. This stage involves creating variables, structuring data, and aggregating data as appropriate. The end product is a set of altered datasets ready for examination. This process concludes when the data is prepared for additional analysis and in the right format.

*Further the process of Exploratory Data Analysis (EDA) is initiated which includes the following steps:*

**Step 4: Trend Analysis:** Both the Year-over-Year Analysis and the Regional Comparison phases are managed by the Data Analyst. In the Year-to-Year Analysis stage, rent rates are compared across many years to identify trends. The process moves on to the Regional Comparison phase, when trends in rent pricing are examined across several places, if these patterns have been clearly identified. After these variances and patterns are fully understood, the outcome is a comprehensive compilation of insights into regional variations in rent pricing and trends from year to year.

**Step 5: Statistical Analysis:** The step of hypothesis testing is managed by the data analyst. Experiments are carried out at this stage to find significant differences in rental prices. The result is a compilation of results indicating statistical significance. The procedure concludes with the hypotheses being tested and conclusions being drawn.

**Step 6: Forecasting:** The Data Scientist oversees a one-time frame that is dedicated to both model development and evaluation. To predict future rent prices, predictive models are constructed, and their performance is then assessed using relevant criteria. The process involves the development of prediction models and model performance measures, culminating in the selection of the optimal model through assessment outcomes.

**Step 7: Reporting and Visualization:**

Report Generation and Dashboard Creation are done in a single process directed by the Data Analyst. Comprehensive reports are prepared to provide an overview of the research and insights, and interactive dashboards are developed to graphically depict the project's findings. The result, which culminates when both deliverables are completed and sent to stakeholders, consists of both interactive dashboards and comprehensive reports.

**Movement between the steps**

* **Sequential Activities:** The process of acquiring data involves several steps, including cleaning, transformation, EDA, trend analysis, statistical analysis, forecasting, and reporting
* **Unordered Activities**: Descriptive statistics is an EDA task that can happen anytime through the process.
* **Concurrent Activities**: Statistics analysis and trend analysis might take place at the same time.
* **Conditional Activities**: The performance of the produced models determines how effectively the models are evaluated; if the models perform poorly, the process may revert to model development.

**Business Problem**

**1. Rent Affordability Analysis:** Both politicians and citizens have serious concerns about how affordable rental housing is. Rent increases can put households under more financial stress, which might affect their standard of living and financial stability. Analyzing rental housing affordability across various geographies and pinpointing places where rent costs are excessively high in relation to typical salaries is the business challenge at hand.

Distinguishing Factors:

* Affordability Metrics: Assess rent prices in relation to average household incomes in various regions.
* Regional Comparison: Identify regions where rent prices are significantly higher than the national average.
* Impact Assessment: Evaluate the impact of rising rent prices on different demographic groups.

2. **Market Demand and Supply Dynamics**: There are vital aspects that seek to determine the demand and supply of this housing to assist developers, investors and policymakers within the government. The problem specifically in concern with renting business is about getting the heat and the feel of present market, all about knowing the current state of the market, about the domains in which there is low demand and supply.

Distinguishing Factors:

* Vacancy Rates: Assess the implications of the available jobs in the mentioned areas as to limited experience and this can help in identifying the need.
* New Constructions: Based on the explanation, the following critical question can be developed: How many new rental housing units are being produced at present, and what is the consequence for the supply?
* Price Elasticity: To analyze charges of rent you should consider factors of the demand and supply determinants of price in response to the above question.

To curb the above stated problems, we can solve it by the following:

**Predicting Future Rent Prices:** Precise projection of rent rates in the future is essential for investment and strategic planning. Creating prediction models that project future rent rates based on previous data and pertinent economic variables is the commercial challenge. Following could help us drill down the solution:

* Historical Trends: Utilize historical rent price data to identify trends and patterns.
* Economic Indicators: Incorporate economic indicators (e.g., employment rates, inflation) into predictive models.
* Forecast Accuracy: Evaluate the accuracy of predictive models and refine them for better performance.

**Stakeholders**

As a team tackling a business issue utilizing this data, we have an interest in making sure the data utilized and the conclusions drawn from it are valuable. Among our stakeholder roles and contributions could be: Among our stakeholder roles and contributions could be:

1. Data Analysis: For strategic purposes, the team members should be able to compute, comprehend, identify, and comprehend the rent price data they have collected.

1. Problem Identification and Solution Development: From our understanding of the current situation presented by the analysis, the members of the team can state potential business issues or issues of concern from the rental market and come up with potential solutions or recommendation. Possible business issues or challenges related to the rental market could be determined by the team members using the analysis, and they are also able to generate solutions or strategies.

1. Reporting and Presentation: For better understanding by the audience, the team will also prepare analytical reports and recommendations alongside the analysis.

1. Decision Impact: Fiscal policies, legislative measures, and managerial strategies regarding the rental market and the residential rental industry will also be informed by the work of the team with the help of which several governmental agencies, real estate companies, and housing interest groups will decide on the given issues.

1. Project Management: Another factor that will affect the overall efficiency of the project will be the distribution of tasks among the members of the team, where individuals will be assigned to manage the project and perform data cleaning, modelling, and reporting.

The team members have an interest in guaranteeing the precise interpretation and efficient application of the historical data for rent prices since they are actively involved in the analysis and problem-solving process. Their research and suggestions will have a direct bearing on future rent price projections as well as several other housing-related issues, housing regulations, and business decisions made by other stakeholders.

|  |  |  |  |
| --- | --- | --- | --- |
| **Group contribution** | | | |
| Page no | Points | Content | Stakeholder |
| 1. | 25 | Project discussion and importance | Apeksha  Tanya |
| 2. | 15 | Process flow diagram | Shrutika |
| 3. | 25 | Business problem and requirements definition | Apeksha  Yshika  Gurleen |
| 4. | 10 | Stakeholders | Yshika |
| 5. | 15 | Analytics questions | Tanya  Shrutika |
| 6. | 10 | Scope statement | Gurleen |

**Analytics Questions**

To analyze the data better and extract the best possible results we would be focusing on a few analytic questions that would enable us to streamline the entire process and help us focus on the most important and valuable prospects of the whole project. We would be discussing Descriptive Questions (This will focus on the analyzation of the existing historical data), Diagnostic (This will focus on diagnosing the reason behind the increase and decrease of the rent rates based on the historical data), Predictive (We can predict the future rent prices and based on the historical data by building a machine learning algorithm) and Prescriptive (This will enable us to provide vital steps to be taken, based on analyzing the historical data, diagnostic insights, and predictive models.

1. **Descriptive:** Are there any clear patterns or trends in rent costs across geographies or residential types?
2. **Predictive:** How are rent prices expected to fluctuate over the next 5-10 years for various geographies and housing types, and what factors are most relevant in forecasting future rent price changes?
3. **Prescriptive:** How can the real estate developers and property managers enhance their pricing strategies based on expected rent price trends?

### **Scope of the project**

This Capstone Project is designed to analyze historical rent prices across various Canadian regions (with population over 10,000) to predict future trends, focusing on different housing types and units. Utilizing a comprehensive dataset, the project aims to uncover trends, patterns, and insights that will inform stakeholders in real estate, policymaking, and business sectors. The project will progress through several stages: data collection and cleaning, exploratory data analysis (EDA), statistical analysis, and predictive modeling. The final deliverables will include:

1. **Cleaned Dataset:** A well-prepared, clean dataset ready for analysis, free of missing values, duplicates, and errors.
2. **Exploratory Data Analysis:** Initial insights and visual representations of rent price trends and patterns.
3. **Statistical Analysis Results:** Detailed findings on the factors influencing rent prices, including hypothesis tests and regression models.
4. **Predictive Models:** Validated models capable of forecasting future rent prices for various regions and housing types.
5. **Interactive Dashboards:** Visual tools (e.g., Tableau, Power BI) to display trends, forecasts, and key insights interactively.

The predictive models assume that there will be no drastic, unforeseen changes in the economic conditions that could significantly alter rent price trends.

**Project Limitations:**

* The project will not include the integration of real-time data feeds or continuous data updates.
* Long-term implementation, ongoing maintenance and updates of the predictive models and dashboards beyond the project completion are not covered.
* While the project will provide recommendations, it will not involve the actual implementation of policy changes or interventions.

### **Risks:**

* **Model Accuracy:** The risk that predictive models may not accurately forecast future rent prices due to unforeseen market dynamics or data limitations.
* **Stakeholder Buy-In:** Risk of insufficient engagement or feedback from stakeholders, impacting the relevance and applicability of the findings.

**References**

<https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3410013301> (March 17, 2023)