CSC343 Phase 1

Xinyi Song(1004152338), Min Wu(1003897839)

September 30, 2021

Domain:

Many websites state that the unemployment rate, interest rate and household debt-to-income are critical factors influencing the house price index. In this project, we will discuss whether these three factors affect the housing price index in Canada.

Dataset:

- Link to datasets:
  - Canada new housing price index:
    `https://tradingeconomics.com/canada/housing-index`
  - Canada unemployment rate:
    `https://tradingeconomics.com/canada/unemployment-rate#data`
  - Interest rate in Bank of Canada:
    `https://tradingeconomics.com/canada/interest-rate`
  - Canada household debt-to-income:
    `https://tradingeconomics.com/canada/households-debt-to-income`
- Relevant information:
  For all four tables, which are the new housing price index table, unemployment rate table, interest rate table and the household debt-to-income table, the relevant information is Country, DateTime and Value from 2001-01-01 to 2021-08-31.
- Directions for learning before interpreting data:
  In this project, we want to find whether the unemployment rate, the interest rate and the household debt-to-income have a positive or negative or no correlation with the housing market. So we realize that we need to know how to create a graph or build a data model in SQL.
- Directions for cleaning up:
  Remove the data that is before 2001-01-01. Then, remove the irrelevant column apart from the Country, DateTime and Value from the original dataset.

Investigative Questions:

1. Are housing price index and household debt-to-income in a positive correlation?
2. How does the housing price index fluctuate with the unemployment rate?
3. Among unemployment rate, interest rate, and household debt-to-income, which factor significantly impacts the housing market?

Schema:

- Relational schema:
  - HousingPI(country, <u>date</u>, index)
    A tuple in this relation represents a Canada monthly housing price index.
  - UnemploymentRt(country, <u>date</u>, rate)
    A tuple in this relation represents a Canada monthly unemployment rate.
  - InterestRt(country, <u>date</u>, rate)
    A tuple in this relation represents a daily interest rate in Canada.
  - HouseholdDTI(country, <u>date</u>, ratio)
    A tuple in this relation represents a Canada quarterly household debt to income.
- Integrity constraints:
  - UnemploymentRt[date] $\subseteq$ HousingPI[date]
  - InterestRt[date] $\subseteq$ HousingPI[date]
  - HouseholdDTI[date] $\subseteq$ HousingPI[date]
- Data dictionary:

HousingPI

Attribute	Description	Type	Required	Default
country	The country of the housing price index	TEXT	Yes	Canada
date	The released date of the housing price index data	TIMESTAMP	Yes	
index	The value of the housing price index	FLOAT	Yes	

UnemploymentRT

Attribute	Description	Type	Required	Default
country	The country of the unemployment rate	TEXT	Yes	Canada
date	The released date of the unemployment rate	TIMESTAMP	Yes	
rate	The value of the unemployment rate in Canada	FLOAT	Yes	

InterestRT

Attribute	Description	Type	Required	Default
country	The country of the interest rate	TEXT	Yes	Canada
date	The released date of the interest rate	TIMESTAMP	Yes	
rate	The value of the interest rate in Canada	FLOAT	Yes	

HouseholdDTI

Attribute	Description	Type	Required	Default
country	The country of the household debt-to-income	TEXT	Yes	Canada
date	The released date of the household debt-to-income	TIMESTAMP	Yes	
ratio	The value of the household debt-to-income	FLOAT	Yes	

- Justification of design:
  We decide to use the structure of the dataset directly without any changes. In this project, we will analyze the unemployment rate, interest rate and household debt-to-income to find how these three factors individually affect the housing market.
  Therefore, we would need a complete table for each factor. We found the date and corresponding value of these three factors, and the structures of data displayed are what we want exactly.