**MATT GARLOCK**

**DSC 540 FINAL PROJECT MILESTONE 1**

**6-15-24**

**Topic: Housing Market Prices in Comparison to Median Income in Florida**

**Project Subject Area**

**Housing Market Prices in Florida Compared to Median Income:** This project will analyze the housing market prices in Florida and compare them to median income levels across different regions in the state. The goal is to understand the affordability of housing in relation to income.

**Data Sources**

**1. Flat File (CSV/Excel)**

**Description:** This dataset includes historical housing prices, sales data, and inventory levels by city and county in Florida. It provides data on the median sale price, the number of sales, and the available inventory.

**Link:** Florida Housing Market Data (CSV)

https://www.zillow.com/research/data/#:~:text=Zillow%20Home%20Value%20Index%20(ZHVI,and%20as%20a%20raw%20measure.

**2. API**

**Description:** This API provides data on median income levels by state and county. It includes annual income data, allowing us to analyze trends over time and compare them with housing prices.

**Link:** US Census Bureau API - Income Data

https://www.census.gov/programs-surveys/acs/data/data-via-api.html

**3. Website**

**Description:** This website contains a table with data on housing affordability indices, cost of living, and economic factors by state and county. The data includes affordability metrics that can be used to compare housing prices with median income levels.

**Link:** Missouri Economic Research and Information Center - Cost of Living Data

https://meric.mo.gov/data/cost-living-data-series#:~:text=MERIC%20derives%20the%20cost%20of,first%20quarter%202024%20was%2088.5.

**Relationships**

* **Flat File & API:**
  + **Common Variable:** City/County
  + Housing price data from the flat file can be related to median income data from the API based on the city or county.
* **Flat File & Website:**
  + **Common Variable:** City/County
  + Housing price data can be connected to affordability index and cost of living data using the city or county as a common variable.
* **API & Website:**
  + **Common Variable:** City/County
  + Median income data from the API can be related to affordability and cost of living data from the website using the city or county as a common variable.
* **Project Approach/Plan**
* **Data Collection:** The first step involves gathering data from the identified sources. The Zillow dataset will be downloaded as a CSV file, providing historical housing prices and inventory levels in Florida. The US Census Bureau API will be used to fetch median income data by state and county, ensuring the data is comprehensive and current. The Missouri Economic Research and Information Center (MERIC) website will provide cost of living and housing affordability indices.
* **Data Cleaning:** This stage involves preprocessing the data to ensure consistency, accuracy, and completeness. Data cleaning will address missing values, outliers, and inconsistencies. For example, housing price data and median income data might need to be adjusted for inflation to make them comparable over time.
* **Data Integration:** The datasets will be merged based on common variables such as city, county, and year. This integration will involve creating relationships between the datasets, such as linking housing prices with median income and cost of living indices.
* **Data Analysis:** Descriptive statistics will be performed to identify trends and patterns. Correlation analysis will be used to examine the relationship between housing prices and median income. Regression analysis will help understand the impact of economic factors on housing affordability.
* **Visualization:** Data visualization tools will be used to create line charts, bar charts, and heatmaps to present the findings clearly. These visualizations will help in identifying trends, outliers, and significant patterns in the data.
* **Reporting:** The final step involves compiling a comprehensive report summarizing the insights, conclusions, and recommendations. This report will include visualizations, statistical analyses, and a discussion of the results.
* **Concerns/Challenges**
* **Data Integration:** One of the primary challenges is ensuring the compatibility and correct merging of datasets from different sources. Variations in data formats, missing values, and different time frames can complicate the integration process.
* **Data Quality:** Ensuring data quality is crucial. Incomplete or inconsistent data can lead to inaccurate analyses. Handling missing values, outliers, and discrepancies between datasets will require careful attention.
* **API Limitations:** Using the Census Bureau API might pose challenges such as rate limits and the complexity of constructing queries to fetch the required data. Ensuring comprehensive data retrieval within these limitations is crucial.
* **Inflation Adjustment:** Adjusting historical income and housing price data for inflation is necessary to make meaningful comparisons over time. This process requires accurate inflation indices and careful calculations.
* **Ethical Implications**
* **Privacy:** While the data sources used in this project are publicly available and do not contain personally identifiable information, it's essential to ensure that data privacy is maintained throughout the project. Aggregating data at the county or city level helps protect individual privacy.
* **Bias:** Awareness of potential biases in the data is crucial. For example, economic indicators and housing market data can be influenced by socio-economic factors that may not be evenly distributed across different regions. Ensuring that analyses and conclusions consider these biases is essential for accurate and fair reporting.
* **Data Misinterpretation:** There is a risk of misinterpreting the data or drawing incorrect conclusions. To mitigate this, statistical analyses must be conducted rigorously, and assumptions should be clearly stated. Peer reviews and consultations with domain experts can help ensure the validity of the findings.
* By addressing these challenges and ethical considerations, this project aims to provide a comprehensive analysis of housing market prices in comparison to median income in Florida, contributing valuable insights into housing affordability and economic trends in the region.