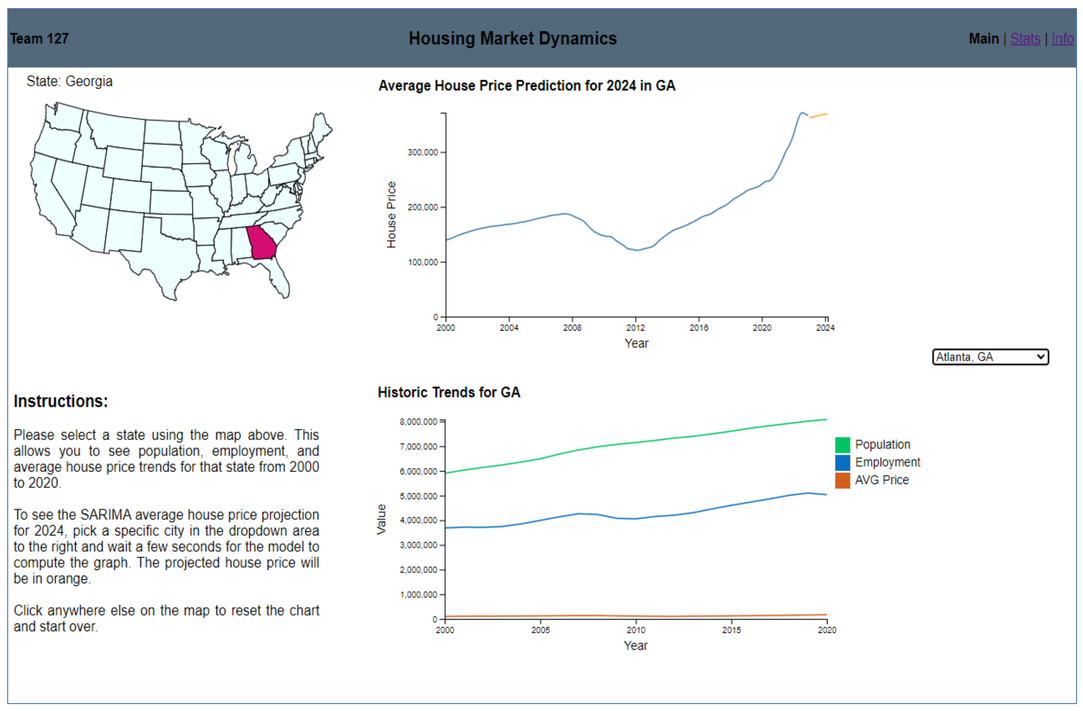
Housing Market Dynamics

Antolovich, Mira | Pierce, Kimberly | Rasouli, Houtan | Toni, Donatien

Summary

The current methods for predicting the housing market can be complicated and difficult for casual users to understand. Simpler methods may have limitations, such as relying on limited data sources that may contain errors.

To address these issues, the "Housing Market Dynamics" project aimed to create an accurate and user-friendly dashboard that explores the housing market and provides essential prediction tools for casual users, including real estate agents and potential homebuyers. The project aims to **simplify the process of navigating the housing market** and provide **valuable insights for non-experts**.



Data

We searched and found over **50 data sources** related to key performance indicators (KPIs) in the housing market. These sources covered several topics, such as population trends, median house prices, house sales count, employment, GDP, homeownership rate, and interest rates. After obtaining the data, we profiled and transformed it into time series format, enabling us to use it for predictive analysis and visualization.

Our data was obtained from various accurate online resources such as Data.gov, Statista, Zillow, and World Bank Open Data, among others. To create a housing price prediction model, we assembled a dataset with over 4 million rows representing multiple important and influential factors for more than **22,000 cities** across the United States. This data covers a 20-year period, from 2000 to 2020.

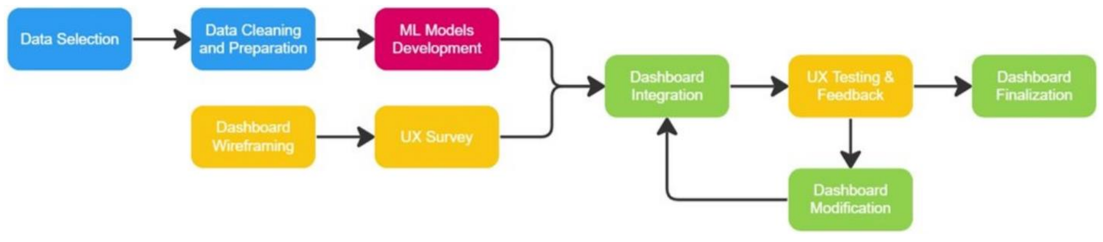
Given the size and complexity of our dataset (training parameters of SARIMA), we leveraged cloud computing services (Azure ML Studio) to run our prediction model.

Approach

Our approach was based on the following key ideas:

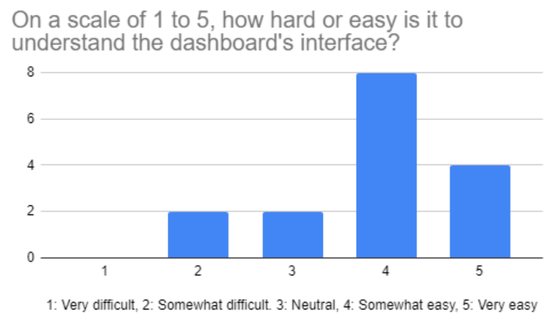
- **Integration of multiple factors:** We have included multiple factors, such as interest rates, inflation rates, and income levels, and demographic data, such as population growth, and migration, to provide a complete overview of the housing market dynamics.
- **Advanced analysis techniques:** In addition to traditional statistical methods, our project used machine learning algorithms, such as time series analysis, to provide deeper insights into the housing market.
- **Interactive data visualization:** Our goal was to deliver a data visualization dashboard that allowed users to interact with and explore the data in real-time and learn about important housing market KPIs.

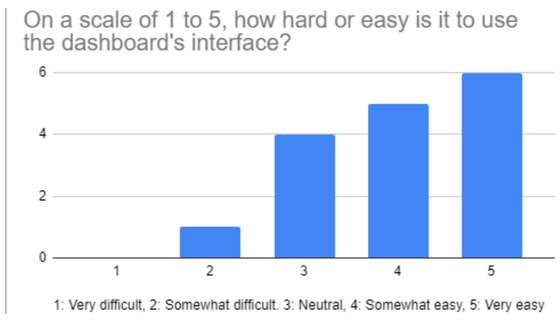
Approach Workflow Visualized



Experiments and results

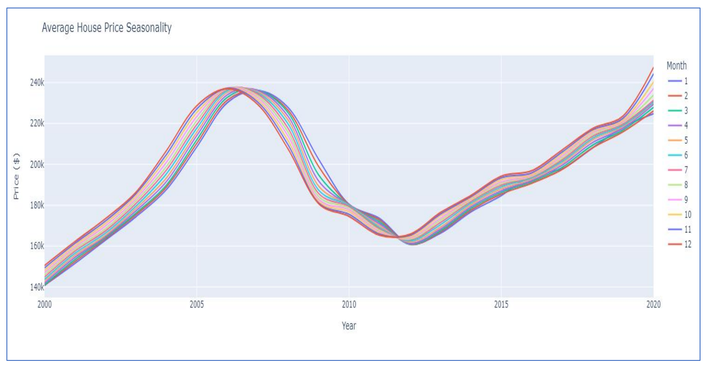
Approach evaluation: To ensure that our dashboard met the needs of casual users, we conducted **multiple user surveys** through Google surveys. By conducting this survey, we obtained valuable insights into the preferences and expectations of our target audience. This information enabled us to refine our project design and ensured that we met the needs of our users.

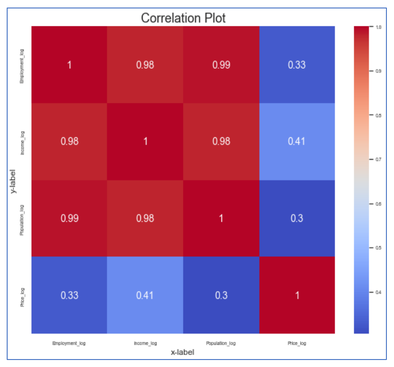




Additionally, we completed usability tests to further understand where our dashboard was successful and where it was lacking.

Modelling, analytics and results: We did a **multiple regression analysis** to identify the factors that influence house prices and then we utilized the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, which accounts for the seasonality observed in our monthly house price data.





Method comparison: Several studies have investigated the critical factors that significantly influence the housing market from various perspectives for many years. However, our approach differed from the existing methods as **we considered user experience to be just as crucial as data preparation and prediction models**. As a result, we developed a dashboard that allows users to easily explore the housing market and provides useful prediction tools. Our main objective was to ensure that our dashboard was user-friendly and accessible to everyone, providing a clear understanding of what to expect.

Next steps: In the future, we can look at our ETL process and evaluate if there are better ways to preprocess our data. Additionally, we can continue working on training and finetuning our model to allow it to be more accurate. Finally, we could continue to evaluate our dashboard for user-friendliness.

Conclusion

While the Housing Market Dynamics project has identified the primary drivers of the housing market, its most notable accomplishment is the creation of a dashboard that **empowers investors and property owners to make informed decisions**. It should be noted that this dashboard is only applicable to the American market. Nonetheless, initial surveys suggest that the project's dashboard **provides valuable insights** to casual users, especially to those who are new to the market.