HOMEQUEST: A COMPREHENSIVE DASHBOARD FOR FINDING THE IDEAL STATE TO BUY A HOUSE IN THE USA

DATA-3400   DATA VISUALIZATION NARRATIVE

NAME: Satyanarayana Vinay Achanta
DATE: 08/10/2023

1. QUESTION / PURPOSE

What is the research question(s) you will be researching and answering with your data? (Be curious, what do you wonder?) In a few sentences, describe the question or problem that you are seeking to address. This should be a direct and answerable question.

Which City and State in the USA provides the most favorable conditions for buying a house?

2. AUDIENCE

Who is your audience? List the primary groups or individuals to whom you'll be communicating. If you had to narrow the audience to a single person, who would that be? What does your audience care about? What action does your audience need to take?

What are the benefits if your audience acts in the way that you want them to? What are the risks if they do not?

My audience is a real estate investor or someone looking to buy a house in the USA. If I were to narrow it down to a single person, I imagine a prospective homebuyer who is interested in finding the best state to buy a house. They care about making an informed decision and are looking for data-driven insights mainly on the average house prices, and possible things like crime rates, unemployment rates, to guide their purchase. If I can provide them with a comprehensive dashboard that presents this information in an accessible manner, it can help them make a more informed choice on where to invest in real estate based on factors that matter to them, such as affordability, safety, and employment opportunities. If the audience does not act in the desired way, they may face risks such as making uninformed decisions based on limited or unreliable information. This could lead to potential financial losses, living in an area with unfavorable conditions, or missing out on opportunities in states that provide more favorable conditions for buying a house.

3. SINGLE SENTENCE

Develop a user-friendly dashboard that presents average house prices, crime and unemployment rates for each state in the USA to assist homebuyers in finding the best state to buy a house.

4. DATA

Briefly describe your dataset(s) here. What is the source? Does it need cleaning? Include the file(s) you plan to use in your submission.

I am utilizing three datasets for this project.

The first dataset is the House Dataset, obtained from Zillow Research (https://www.zillow.com/research/data/). It contains information on states, cities, and house prices, including a time series of average house prices from 2000 to 2023. The dataset required some cleaning, such as adding state names for Tableau to read into maps.

The second dataset is the Crime Dataset, sourced from Corgis (https://corgis-edu.github.io/corgis/csv/state_crime/). It provides information on crime types and frequencies in each state from 1960 to 2019. The dataset consists of columns like State,

State_Code_Crime, Year, Data.Population, Data.Rates.Property.All, etc. Cleaning was necessary, including adding state codes to map it with the house price dataset.

The third dataset is the Unemployment Dataset, collected from the Iowa State University's Institute for Community Prosperity (https://www.icip.iastate.edu/tables/employment/unemployment-states) and the U.S. Bureau of Labor Statistics (https://www.bls.gov/lau/lastrk18.htm). It contains details of unemployment rates for each state from 1980 to 2023. The dataset required cleaning, such as adding state codes and extracting data from additional sources to include information beyond 2018.

By utilizing these datasets, the goal is to develop a user-friendly dashboard that presents comprehensive information on average house prices, crime rates, unemployment rates, and other relevant factors for each state.

House.xlsx   Crime.xlsx   Unemployment.xlsx

5. CHART TYPES + UNITS

What types of charts do you think best visualize your data? Use your best guess to predict what chart types you believe will help tell your story. Why?

Line Chart: I am considering utilizing a line chart to visualize the time series data of house prices. This chart would be interactive, allowing users to click on a specific state or city on the map. Once clicked, the line chart would dynamically display the historical trends and fluctuations in house prices for the selected location over time. This approach provides a focused and detailed view of the house price dynamics, enabling users to explore the specific changes and patterns associated with each state or city.

Bar Chart: A bar chart can be used to compare average unemployment rate across different states. It provides a clear visual comparison of prices and helps identify the states or cities with the highest or lowest values.

Multi-Line Chart: A multi-line can be used to explore the relationship between several different crimes and number of crimes. This helps identify potential correlations or patterns in the data.

6. LOCATION IS EVERYTHING

Describe how you might incorporate a spatial element in your visualization. Specifically, what data would lend to a map or space. As part of the final dashboard, you are required to have a spatial map. However, it doesn't need to be a geographical map (although many students go this route), it could be any location e.g., basketball court, grocery aisle, body, etc.

To include a spatial element in the visualization, I would use geographical maps to display the data related to different states in the USA. The datasets such as the House Dataset, Crime Dataset, and Unemployment Dataset, all have state-level information. By leveraging the geographical location of each state, I can create a map that visually represents the data and provides a spatial context.

By incorporating these spatial maps into the final dashboard, users can visually explore and compare data across states, enabling them to make more informed decisions about where to buy a house based on their preferences and priorities.

7. PROCESS

Throughout the development of the HomeQuest dashboard, I encountered various challenges and had moments of clarity that shaped the project's direction. Here's an overview of the key steps and insights gained during the process:

Data Collection and Cleaning:

I began by sourcing the required datasets from reputable sources such as Zillow Research, Corgis, Iowa State University, and the U.S. Bureau of Labor Statistics. After downloading the datasets, I had to clean and preprocess them to make them compatible with Tableau. This involved tasks such as merging data, standardizing state codes, and extracting relevant information.

Connecting Data to Tableau:

To establish a connection between the cleaned datasets and Tableau, I utilized the data connectors provided by Tableau. These connectors enabled me to import, link, and structure the data within the Tableau environment. This step was crucial in setting the foundation for building visualizations and creating a cohesive dashboard. However, I did encounter some challenges with the data interpreter in Tableau, which resulted in connection relationship issues. It remains to be determined whether these issues were specific to my Tableau setup or a broader concern. Also, I navigated challenges such as manual column data type conversion due to Tableau's inaccurate autodetection, particularly with date columns being misinterpreted as numbers and lot more.

Choosing Visualizations:

Selecting the appropriate chart types was a thoughtful process. I opted for line charts to depict the time series data of house prices, bar charts to compare unemployment rates, and multi-line charts to analyze crime patterns. These choices were influenced by the need to effectively convey insights while maintaining user-friendliness and interactivity.

Spatial Mapping:

To incorporate a spatial element, I decided to use geographical maps to represent state-level data. This choice aligned well with the geographic nature of the datasets and provided users with a visual context for their decision-making. Implementing these maps required me to explore Tableau's mapping capabilities and integrate them seamlessly into the dashboard using filters.

Dashboard Design and Interactivity:

Creating a user-friendly and informative dashboard layout was a crucial step. I aimed to ensure that users could interact with the visualizations, select states of interest, and observe the dynamic changes in charts and maps accordingly. Implementing filters, tooltips, and drill-down options enhanced the overall user experience.

Exploratory or Explanatory:

This project is primarily exploratory in nature. It aims to analyze and visualize various data sets, including house prices, crime rates, and unemployment rates, to provide insights and guidance for potential homebuyers or real estate investors. By creating interactive visualizations and incorporating geographical maps, the project enables users to explore and compare different states' conditions, facilitating informed decision-making. The project's focus on uncovering patterns, trends, and correlations within the data aligns with the exploratory approach, aiming to provide a comprehensive understanding of the factors affecting the choice of a suitable state for purchasing a house.

Future Considerations:

While the current version of the HomeQuest dashboard provides valuable insights, there are potential enhancements to explore. Incorporating more contextual information about each state, such as lifestyle factors, education quality, and healthcare facilities, could offer a more holistic perspective to users.

Conclusion:

Developing the HomeQuest dashboard has been an enlightening journey, merging data visualization techniques with real-world applications using Tableau. The process highlighted the significance of thoughtful data collection, cleaning, visualization selection using Tableau, and user-centered design.

8. SCREENSHOT