Bust A Move!

Katy Barone (kbarone)
Shobitha Cherian (shobithac)
Norah Griffin (ncg)

Project Overview

As a consequence of the pandemic, many companies have shifted to more permanent remote-work policies. This has likely impacted migration from high density cities to smaller cities and towns, especially in the western US, that offer access to outdoor activities. Our goal was to visualize these migrations and their effects on the demographics and resources of these smaller cities and towns. We created a dash application that includes an interactive choropleth map representing the United States at the county-level, color coded by changes in housing prices as a proxy for migration to that county, overlaid with a layer of green dots representing the U.S. National Parks. The map highlights counties that have experienced dramatic changes in housing prices and that have either relatively low income or high poverty since these counties may be at risk for negative consequences of increasing housing prices and resource constraints. By clicking on a county, a user can see additional characteristics for a specific county including income, poverty rate, racial breakdown, and data on activity within the county.

Software Structure

This is a dash based application, with the main file being "app.py." The application accesses the data module and the graph_functions.py file in order to create the graphics in the application. The data module includes raw data, a clean_data.py file that includes functions to clean the data, and a folder containing the cleaned data created by the cleaning functions. The app.py file calls functions from the graph_functions.py file in order to create the data visualizations. The application also accesses the "assets" module which provides formatting for the application.

The app.py file reacts to various inputs by the user. The map has a toggle of viewing "High Risk" or "Full Map". The "High Risk" toggle highlights the counties that we consider high risk (dramatic increase in housing prices, and either low income or high poverty rate). The full map will show all counties on the same opacity level. The app also reacts to a user clicking on a county. When the user clicks on the county, the county is highlighted and the graphics on the right side change to reflect the specific county. A user can also change the types of graphics for the specific county through a dropdown, and the graphics will change based on the dropdown selection.
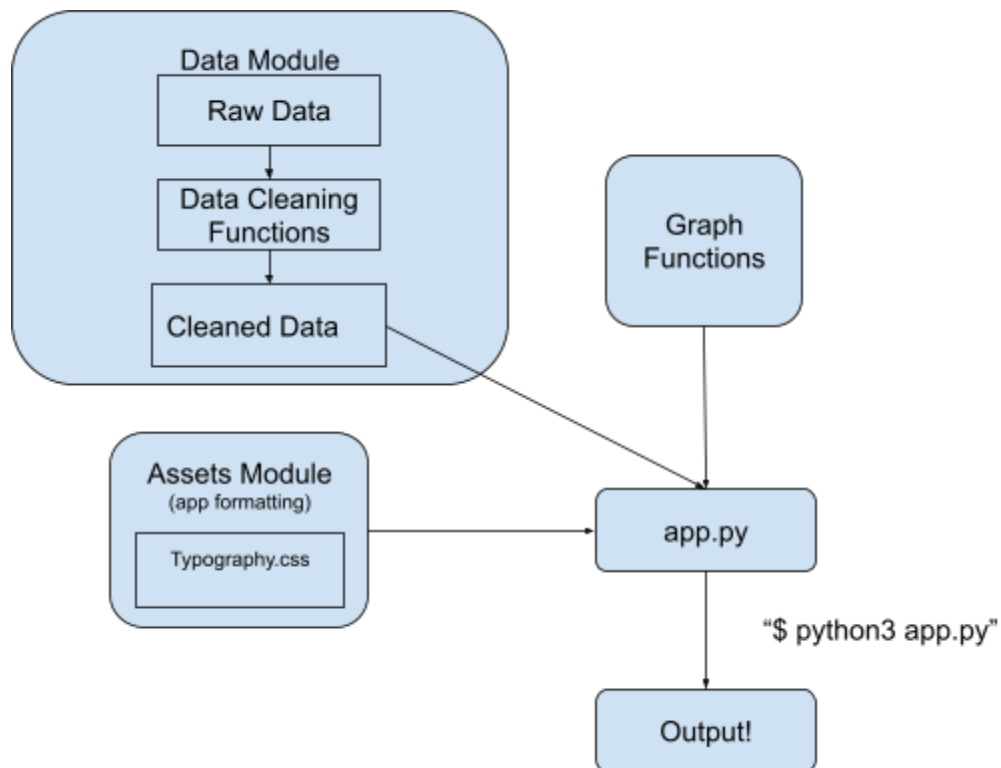
Code Responsibilities

Our project was very collaborative and everyone contributed to almost every task/ module/ file.
With that said, we have provided a rough breakdown of responsibilities below:

- Shobitha
  - Data collection (zillow)
  - Data cleaning
  - Most Dash development
  - Data visualizations (poverty/ income graphs)
- Katy
  - Data collection (census data - poverty, income, race)
  - Data cleaning
  - Data Visualizations (map, race breakdown)
  - Some Dash
- Norah
  - Data collection (zillow, google mobility)
  - Data cleaning
  - Data visualizations (map, gps activity)
  - Some Dash

How to Interact the Application:

Once within the main directory for the project, the user will access the application by entering "python3 app.py" in the command line. The user will then be directed to a web address that runs the application. There will be a description of the project, a chloropleth map color coded based on housing price changes, and a series of graphs on the right that present data for a specific county in the map. The chloropleth map can show the entire US on one color scale if the "Full Map" radio button is selected, or it will highlight counties that we've determined to be high risk if "High Risk" is selected. The map also has green points that represent national parks. By hovering over a county, the user can see general data on each county, which can help guide the user to click on counties that they find interesting. When a user clicks on a county on the map, the graphics on the right will change to reflect the characteristics of that county. The user can choose to view different characteristics by selecting different categories in the dropdown.

What the project tried to accomplish and what it actually accomplished:

Our project explored the theory that the increasing availability of remote working opportunities, partly as a consequence of the pandemic, has led to migrations from major cities to smaller cities and towns in the western US that provide access to outdoor recreational activities (close to national parks). Using historical housing price data as a proxy for recent regional mobility, we were able to create a visualization that matches our theory and confirms information in other news stories and research on this theory. Additionally, we hoped to capture the effects that these migrations may have on the local communities and resources in that county. We were unable to directly achieve this aspect of the project and instead used proxies of income and poverty level to provide some insights into how these major changes might affect the local population. We had high hopes for the google mobility data based on gps activity, but it was not as useful since it is confounded by responses to the Covid pandemic.

Data Sources:

Zillow Data:

- https://www.zillow.com/research/data/

Google Mobility Data:

- https://www.google.com/covid19/mobility/
- https://github.com/OpportunityInsights
    - We took aggregated google mobility data from Opportunity Insights

National Parks Location Data:

- https://github.com/sughodke/D3-US-Graph/blob/master/nationalparks.csv

Census/ACS data (income, poverty, race):

- https://data.census.gov/cedsci/table?q=DP05&g=0100000US%240500000&tid=ACSDP1Y2019.DP05