Title: COSI 116A F24 Team 1 - Jason Saez, Bryan Jiang, Alex Danilkovas, Eric Xiao
https://drive.google.com/drive/folders/1kRCfgW30uFAVctOZMuxD8tTWGC9Y7Vf9?usp=drive_link

Motivation:

Real estate trends are typically displayed as numerical values in either a table or chart. There is usually limited interactivity, much less well thought out visualizations. Specifically, there are few visualizations that show information on a county-level map and other graphical means. We decided to try to solve this problem through our project. We made a real estate visualization that utilizes a link between a map and a chart. The purpose is mainly for entertainment as well as informing as we'll be communicating how the percentage of renters to buyers has changed over time. We will also be informing users about real estate transactions as well as pricing information on a county-to-county basis. Hopefully, this data can reveal trends about where the real estate market is active. This will give the viewer the freedom to make their own assumptions and conclusions. Users will be able to look at data and compare real estate trends to gauge how the market is doing, which counties are seeing the most sales, and which are getting the most expensive sales on average.

The interactive visualization: Included as part of this page.

Demo Video: Embedded MP4 demo video using the HTML5 <video> tag if you are not doing an in person presentation.

Visualization explanation: final visualization screenshots (PNG images), design justifications, UI walkthrough, and linked presentation slides.

Data Analysis: summary of data, data types, and data preprocessing.

The primary dataset we found was on kaggle and contained very comprehensive data about real estate and can be found at the following link: https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset As you can see, the data was originally divided via state and zip code. We figured a state-level view wouldn't give as precise information as we wanted, so we attempted to show our map by zip code. We found a zip code geoJSON file that would allow us to render a map of each individual zip code. However, when we tried to render the zip code view, the size and quantity of zip codes in the U.S made the website crash whenever we tried to render it. Apparently, there are 41,683 zip codes in the U.S, far more than we thought. So, we thought that using a map of the ~3,000 counties would be much easier, but our data wasn't divided by county. We thought that maybe we should find another dataset that had county-level data, but we had already committed to our original one. Instead, we fixed and cleaned the data to make it fit our county idea. We then found a geoJSON file with all of the counties and could make our map render. We now had state, county, avg price, total price, and total transaction data.

- Data sources match the problem statement and are appropriate.

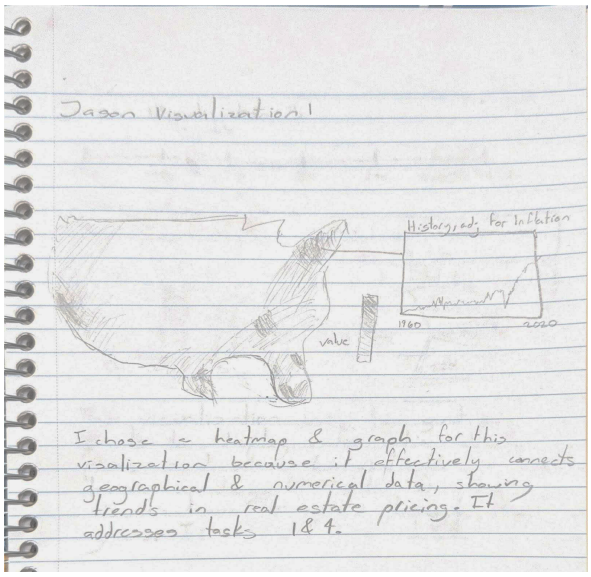
- These descriptions should be very explicit so someone could read your page and properly reproduce your results.

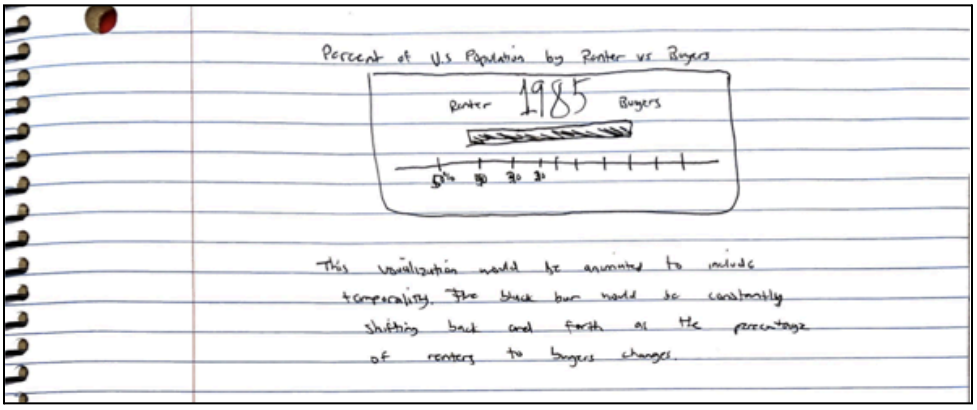
Index (ID #)	"Domain" Task	Low-Level	Mid-Level	High-Level
1.	Identifying Property Types in Different Regions	Identify	Browse	Discover
2.	Show Historical Trends	Summarize	Browse	Present
3.	Display Investment Potential	Compare	Browse	Derive
4.	Contrast Pricing of Similar Housing in Different Regions	Compare	Browse	Discover

Task Analysis: summary of task table.

The main tasks we wanted to have our visualization tackle were to identify property types in different regions, show historical trends, display investment potential, and contrast pricing of housing in different regions. For identifying property types.

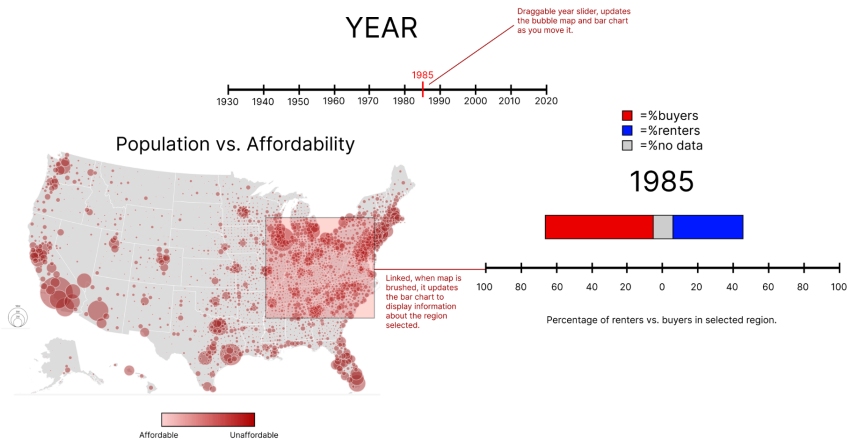
Design Process:
In the beginning, we knew our dataset consisted of data about real estate in the US, so we knew our visualization would most likely need a map of the U.S. We set out to begin drafting potential visualizations that we could use for the map as well as others that would be useful in tandem with the map.





We then voted on some of our favorite ideas (pictured above) from our initial sketches. We decided on having a heatmap/color coded map and a simpler chart view. These choices were made because we thought having an overview and a more detailed view would be a good approach to display our data. This idea came from the Schneiderman mantra that we learned in class:

> “Overview first, zoom and filter, and details on demand.” - Ben Shneiderman.



We then made this more detailed mockup of what we thought our final visualizations would look like. The temporality and single bar chart were ripped straight from our sketches, but the bubble map took some thought to come up with a way to encode population and price. Then, in our in class peer feedback, we got this comment: "I'm not sure what the size of the circle represents." We decided that, for our final visualization, we should encode data not through bubbles, but mainly color, as the former would lead to too much confusion and illegibility.

- Evidence of iterative improvement.
- Logical discussion of design choices grounded in theory from course.
- Discusses feedback from usability testing.

Conclusion: short summary of work completed and areas for improvement/future-work.

- Meaningfully wraps up project and has good future directions.

Acknowledgements: List here where any code, packages/libraries, text, images, designs, slides, etc. that you leverage come from. Include any use of LLMs and AIs.

- County geojson https://gist.github.com/sdwfrost/d1c73f91dd9d175998ed166eb216994a?short_path=bd97547
-

Additional non-required sections may be added to the website to thoroughly explain/frame the final project.

Each section must:
○ Be written in a clear manner -- your style can be less formal than academic prose but it should be serious and engage with ideas.
○ Show a high degree of attention to logic and reasoning of points.
○ Include extensive (at least 200 words), clear discussion of the section's topic.
○ Use words chosen to precisely express the intended meaning and support reader comprehension.

Overall:
○ Unity of sections clearly leads the reader to the conclusion and stirs thought regarding the topic.
○ The web page must contain no more than 3 errors in Standard written English (e.g., grammar, spelling).