10 minute presentation, not including the follow up Q&A. Expect at least one question, if not more. Why this data source? What would you like to do next?

Strict 10 minute limit

Practice transitioning speakers

Keep on person in control of the screen

No code on the slides, unless it’s really cool and prudent to explaining the project

Push everything to the main, anything left off will not be graded

Clean up the repo and keep it nice

Why you chose your project. High level descriptive statistics. Visualizations

DON’T MIX UP MEAN AND MEDIAN DUMMY

Project 1: Housing Affordability, The Real Estate Market, & The Economy

* Introduction, summary
* Questions to be asked
  + Data source/clean up
    - US census API
    - Zillow API
    - Freddie Mac (interest rate)
  + How these questions relate to the industry
    - Real estate market is crucial to the health of the economy as it makes up a large portion of US GDP. Affordability plays a large factor in the real estate market.
* Question 1: What does affordability look like in the US? Are there differences among regions?
  + HAI
  + Data
    - Zillow
    - Freddie Mac
    - US Census
  + Stat analysis
    - Heat map of the US
    - Line plot of NYC, LA, Houston, Miami affordability vs time
  + Conclusion
    - Affordability and homeownership
      * Homeownership and economic activity
    - Government policy and how it effects affordability
      * Affordability and housing bubbles
* Question 2: How significantly does the price of a home change from list price to sale price?
  + Data
    - Zillow
  + Stat analysis
    - Scatter plots of list price vs sales price
  + Conclusion
    - Supply and demand overview of housing market
      * If the final sales price is consistently close to or above the listing price, it may indicate a seller's market where demand is strong, and buyers are willing to pay the asking price or more.
      * Conversely, if the final sales price tends to be below the listing price, it may suggest a buyer's market where there is more negotiation room, and sellers may need to be flexible on pricing.
* Question 3: What relationships exist between markers of economic performance and number of homes sold?
  + Economic factors:
    - Interest Rate
      * Freddie Mac and how the interest rate is calculated
    - Unemployment
    - Income
      * Median household income
  + Data
    - Zillow
    - Freddie Mac
    - US Census
  + Stat analysis
    - Scatter plots of home sales vs:
      * Interest rate
      * Unemployment
      * Income
  + Conclusion
    - What is the relationship between the housing market and the economy as a whole? Do they undulate in a commensurate manner?
    - Can economic factors predict the housing market and vice versa?
* Conclusion
  + Any issues encountered and solutions
    - Missing data for some years
    - Merging data sets
    - Using the Zillow index vs raw data
    - Selecting types of visualizations
  + Potential next steps
    - Predicting market activity based on HAI or vice versa
    - Predicting the differences between sales price and list price using economic factors/HAI
    - How housing affordability effects rental market

https://apps.bea.gov/iTable/?reqid=150&step=2&isuri=1&categories=gdpxind&\_gl=1\*16i9y2a\*\_ga\*MzA1NzYzNDc0LjE3MDI1MjI2Mjk.\*\_ga\_J4698JNNFT\*MTcwMjUyMjYyOS4xLjEuMTcwMjUyMjcxNy40Mi4wLjA.#eyJhcHBpZCI6MTUwLCJzdGVwcyI6WzEsMiwzXSwiZGF0YSI6W1siY2F0ZWdvcmllcyIsIkdkcHhJbmQiXSxbIlRhYmxlX0xpc3QiLCI1Il1dfQ==

Missing data filled with median values for that city. Data imputation. Makes it possible to perform calculations.

Median income data missing for 2007, 2008, 2023

The goal of this project is to explore the local housing markets of metropolitan areas in the United States through the lens of housing affordability and economic measures. Metrics to represent the state of the housing market are median house prices and number of homes sold. Economic factors include median household income, unemployment, and the national mortgage interest rate set by Freddie Mac. The data will be narrowed further to four major metropolitan areas—New York City, Los Angeles, Houston, and Miami to compare housing affordability in different geographic regions of the US. The datasets will also be confined to the period between 2009 and 2022. The specific questions to be addressed are the following:

1. **What does affordability look like in the US? Are there differences among regions?**
2. **How significantly does the price of a home change from list price to sale price?**
3. **What relationships exist between markers of economic performance and number of homes sold?**

The housing market is closely linked to the economy as it has contributed roughly 12% of the GDP for the last five years. This makes it the current largest contributor to the economy, almost tied with manufacturing at 11%. Housing affordability and the size of the difference between the listing price and final sales price can both illustrate the volatility of the housing market. A more stable housing market is a factor in overall economic growth as it encourages more people to become homeowners. Of course, there are more variables that influence a person’s ability to make such a large investment, such as levels of income, unemployment, and interest rates. **Transition sentence here.**

Data for the median income and unemployment levels was sourced from the US Census API while information regarding housing was pulled from the Zillow API. A challenge that presented itself early on was merging these different data sources. The data set contains many variables represented throughout time on a yearly or sometimes monthly basis. Additionally, many metropolitan areas did not have information for all variables at all times. **Go over the data cleaning process that allowed for merging on the city name.** The mortgage rates used in this project are calculated monthly by the Federal Home Loan Mortgage Corporation, or Freddie Mac.

**End Intro**

The first area to be explored is the state of affordable housing today across the US. The Housing Affordability Index (HAI) arose as shorthand for how likely a typical family would be able to afford the average mortgage in their area. HAI is calculated by dividing the Mean Family Income (MFI) by the Qualifying Income (QI) and multiplying that by 100. The QI is a tool used to evaluate mortgage loan applications and estimates how much income is required to afford a standard mortgage. If the QI and the MFI are equal, that suggests the typical family can afford to pay one complete mortgage (an HAI index of 100). As the MFI decreases relative to the QI, a family would only be able to afford to partially pay the mortgage. The lower this number is, the less affordable a market is considered (an HAI of < 100). Conversely, an MFI greater than QI suggests that a single family could afford to pay for their mortgage and then some, implying a healthy economy (an HAI > 100).

The following chart (Figure 1) shows the HAI in New York, Los Angeles, Houston, Chicago, and Phoenix from 2011 through 2022. The graph shows there was a general upwards trend in HAI throughout the country until the COVID-19 pandemic of 2020, where rates plummeted. The HAI bounced back quickly, but the figure indicates a volatile market as the government implemented various economic programs. This graph also illustrates that the HAI of New York and Los Angeles remains low while Phoenix and Houston(?) tend to be relatively high, indicating that a high demand for housing may inversely correlate with the HAI (is this supported by our data?)

For the next few charts, Figures 2 through 3, population size, number of homes sold, and national interest rate are compared to HAI across all cities (?) and times in the data set. Correlation between the variables was represented by scatter plots. In Figure 2, there appears to be a slight negative correlation between population and HAI, but there is not a significant correlation. Similarly, Figure 3 shows that HAI and number of homes sold are slightly positively, though not significantly, correlated. In Figure 4, (is this a line graph??) there is evidence of a government policy, the Federal Reserve increasing the interest rate, negatively impacting HAI. Conversely, Figure 2 shows that as HAI decreased from 2020 to 2022, the number of sales rose. This further solidifies the weak relationship between HAI and the housing market, suggesting that the HAI is not an accurate predictor of market activity. **Something about housing bubbles?**

**End Question 1**

When a house is put on the market, the price it is initially listed as and its final sales price are most often different. When there is a trend of homes selling below asking price, it can be assumed that they are in a “Buyer’s Market,” meaning the supply of homes outweighs the demand. A “Seller’s Market” is the reverse, where there are not enough homes to meet demand and buyers must compete over the same properties. The following charts (Figures 5 through 7?) show this relationship:

**Input Ishicka’s part here**

**End Question 2**

The answer to Question 1 appears to be that there is only a weak correlation between HAI and measures of economic growth. Now those factors will be compared to amount of homes sold to explore any potential relationship there. Figures 8, 9, and 10 are scatter plots of number of sales (x-axis?) by unemployment, mortgage interest rate, and median household income. Once again, there was no significant correlation between these factors. Since these factors seemingly cannot predict one another, it can be concluded that the housing market is complex with many variables contributing to its fluctuations. The housing market is resistant to (or maybe just slow to?) the factors that put pressure on regular goods and services. **Add the charts, do some more descriptions. Discuss how the interest rate is used to represent government policy?**

**End Question 3**

This project explored the housing market and its relationship to economic performance through the lens of housing affordability, unemployment, median family income, and interest rates. **Concluding thoughts, etc**