**What is our research question?**

Our research question is whether or not the mean housing costs have gone up in accordance to inflation or if alternative factors have caused the market to inflate artificially. We have several ways of studying this question, but for the most part we will be studying the housing prices from 2020 to 2022 to 1. Understand accurately what was critical to the market in the past two years, and 2. We have the entirety of the data sets available rather than a partial set that is from 2023. The main concepts are to understand what causes fluctuations in the housing market, especially in respect to COVID, and to understand the inflation rate and to see if the market is solely sheer growth of the housing market or if certain extraneous factors cause the gouge.

**What is your hypothesis and why is it important?**

Our alternative hypothesis is that the mean housing cost in 2022 has increased since 2020. The null hypothesis is that there is no significant difference between the housing cost in 2020 and 2022. This is important because it seems that housing costs have grown a lot over the past few years. And we want to see if the change in housing cost has been a reasonable change, or if it has been more than we expected. There are a lot of people looking for homes. And they will be less likely to make a purchase if the average housing cost is not what they want. Furthermore, we would like to see if the mean housing cost has increased due to bigger and more expensive houses being built, or if it is actually the average family home price is increasing. And we can see this by removing outliers. For example, if a house is worth millions of dollars, then we will remove it from our dataset because that will just increase the average cost of the housing.

**How will you be able to tell/measure the success of the research question and hypothesis?**

We will be able to tell the success of the research question and hypothesis by how well we are able to answer the question and use the answer to help us confirm our hypothesis. If we are able to answer the research question through our model that will show us that it was clear and concise enough to be answered by the project. If we are able to use this answer in order to confirm either of our hypotheses, that will also show that we asked the correct research question in general for the topic at hand. The research question can also be deemed successful based on how relative our answer is to the data we choose since having a good research question should make it easier to find data. Overall we will be able to tell how successful our research question and hypothesis are by the difficulty or ease we have in figuring them out.

**Project Description**

The purpose of this project is to find out if the average housing prices in the United States have gone up since the start of the global pandemic in 2020 to 2022. Not only are we trying to discover if the average housing price has gone up but we are also trying to figure out the potential reason behind it regarding if houses have actually improved or if their prices have risen without the actual house value rising as well. We will do this by first of course researching and understanding what attributes of a house determine its value and how its value can improve. By doing this we will have a better understanding on how to view the data regarding the actual price of houses because we will be able to tell if the house is overpriced depending on its attributes and if houses in the U.S have actually improved. For example if a house has 100 acres of land and it’s expensive it wouldn’t be fair to put that piece of data in with every other house in the United States. Once we have an understanding of what makes a house valuable we will then begin making our data set by first finding data on the average housing price per year from 2020-2022 and cleaning it to remove outliers as mentioned before. The other part of our data set we will then be looking at is data regarding how much money has been put into the improvement of houses. For example the average size, age and condition of houses. Once we have collected, analyzed and cleaned the data we will then begin creating our model. Our model will be based on graphing and charting the data collected year to year based on the simultaneous creation of three graphs. The model will first analyze the data regarding housing prices over the years and create a graph showing the rise or decrease of housing prices based on the data given. The model will next create a graph based on the average attributes of each house and how much money they may cost as well as how it impacts the value of the house. Finally the model will create a graph comparing the two sets of data by comparing the house prices to the improvement of the attributes of each house. By showing these graphs the model will be able to answer 3 questions about the housing market in the U.S. The first one being answered by the first graph mentioned on if the average price of housing in the U.S rose? The next one being if the houses have actually improved in the U.S and the last one being how do those improvements(if there are any) compare to the current average price of housing. By showing these 3 graphs we will be able to show if the price of housing in the U.S has truly gone up or down since the beginning of the global pandemic.

**Schedule**

**Week 1 (02/12/2023):**

-Research on factors influencing housing prices -Sean Johnson.

-USA Housing as from 2000 data collection part 1 - Colin.

-Covid data collection part 1 - Sunny

**Week 2 (02/17/2023):**

**-**Collective discussion on the quality of the data and potential improvements

-Housing and Covid Data collection/refinement part 2 by Colin and Sunny respectively

**Week 3 (02/24/2023):**

-Data Cleansing - Jeremiah

-Exploratory Data Analysis - Sean Johnson

**Week 4 (03/03/2023):**

**-**Data modeling - Fayol

-Collective discussion on the model and potential refinements.

**Week 5 (03/10/2023):**

**-**Data modeling refinements

-Project Summary Draft 1 - Fayol

**Week 6 (03/17/2023):**

-Professor revision

-Final touches

**Week 7 (03/24/2023):**

**-Submission.**