**Pre-Lab Worksheet: Housing Prices**

**Learning Goals:**

Practice designing and conducting data-driven analyses to support or challenge assertions, interpreting and communicating correlation versus causation, defining orders of understanding.

**Introduction to the 2021 Housing Market Analysis:**

In this lab, you will explore the factors that influence housing prices in U.S. cities, with a specific focus on data from September 2021. Housing prices can be influenced by a variety of economic, social, and environmental factors. Your task is to analyze these factors, identify patterns, and assess their impact on housing prices. This will help you understand the interplay between economic indicators and real estate markets, and how such insights can inform policymaking.

**Key Concepts:**

**Hypothesis Formulation**: A testable prediction about the relationship between variables, such as how the number of bars in a city might affect housing prices.

**Correlation vs. Causation**: Understanding that correlation between two variables does not imply one causes the other. For example, a city with a successful sports team may have higher housing prices, but this does not necessarily mean the team's success causes the price increase.

**Statistical Significance and Effect Size**: Statistical significance indicates whether an observed effect is likely due not to random chance, often represented by a p-value.

Effect size measures the magnitude of the relationship between variables, providing insight into the practical significance of findings.

**What are some key economic indicators that could impact housing markets? Provide a brief explanation of why each might be relevant.**

|  |
| --- |
|  |

**How would you formulate a hypothesis about the relationship between an economic indicator (e.g., GDP) and housing prices? What factors would you consider in developing this hypothesis?**

|  |
| --- |
|  |

**Why is it important to distinguish between correlation and causation when analyzing data? How can you ensure that your analysis accurately reflects causal relationships?**

|  |
| --- |
|  |

**What role do visualizations play in data analysis? Describe how different types of visualizations (e.g., scatter plots, bar charts) can help you explore and interpret data.**

|  |
| --- |
|  |

**Post-Lab Worksheet: Reflecting on Housing Market Analysis**

**Learning Goals:**

Evaluate the presence of meaningful patterns in data and distinguish them from random noise. Apply Hill's Criteria to assess causality in relationships between variables. Reflect on the orders of understanding in economic data analysis and their implications for policy and research.

**Patterns vs. Random Noise:**

After analyzing the housing market data, did you observe any patterns that appear to be more than random noise? How did you differentiate between meaningful patterns and random fluctuations? Provide examples from your analysis.

|  |
| --- |
|  |

**Hill’s Criteria Application:**

Choose one variable from the Housing Prices module and apply Hill’s Criteria to assess any causal relationships you identified in your analysis.

Specifically, discuss the following criteria:

* Strength: Was the relationship between the variables strong or weak?
* Consistency: Was the relationship consistent across different datasets or analyses?
* Specificity: Was the effect specific to one variable or did it involve multiple factors?
* Temporality: Did the cause precede the effect in your analysis?
* Biological Gradient: Was there a dose-response relationship (e.g., more successful sports teams leading to higher housing prices)?
* Plausibility: Does the relationship make sense based on existing knowledge?
* Coherence: How does the relationship fit within the broader context of economic and housing market theory?
* Experiment: Were there any experimental or quasi-experimental methods used to support the causality?

|  |
| --- |
|  |

**Orders of Understanding:**

Reflect on the orders of understanding as they relate to your analysis:

* Descriptive: What descriptive statistics did you use to summarize the data? How did these statistics help you understand the general trends in housing prices?
* Associative: What associations did you find between variables? How did you determine whether these associations were likely to be meaningful or merely coincidental?
* Causal: Were you able to draw any causal inferences from your data? What evidence supports these causal conclusions, and what limitations should be acknowledged?
* Predictive: How could your findings be used to make predictions about future housing market trends? What factors would you consider in developing a predictive model?

|  |
| --- |
|  |

**Evaluating Statistical Significance:**

Discuss how you evaluated the statistical significance of your findings. What metrics to determine whether the observed relationships were likely due to chance? How did this evaluation impact your interpretation of the data?

|  |
| --- |
|  |

**Interpreting Results in Context:**

How do your findings fit into the broader economic and social context? Consider how the patterns you identified could influence housing policy and market dynamics. What additional factors might be relevant for understanding these patterns?

|  |
| --- |
|  |

**Reflecting on Methodological Choices:**

Looking back at your analytical approach, what were the strengths and weaknesses of the methods you used? How might different methods or additional data have changed your conclusions?

|  |
| --- |
|  |

**Broader Implications:**

How does your analysis contribute to the broader understanding of economic indicators and their impact on housing markets? What lessons can be applied to other areas of economic research or policy?

|  |
| --- |
|  |