Milestone 5

# Summary of Analysis

In this analysis I examined the Cost-of-Living Index data from 2024 focusing on trends and insights across various years. The dataset contains multiple economic indicators such as unemployment rates, food and nutrition costs, and medical costs for different regions. The goal of the analysis was to uncover trends in the cost of living, regional variations, and the impact of specific economic events. After cleaning and preparing the data, I performed exploratory data analysis and developed visualizations to highlight key findings.

## Findings

1. **Trends Over Time**: The data revealed significant fluctuations in cost-of-living indices over the years. For example, the UNR\_16 (Unemployment Rate) exhibited a noticeable increase in several countries during economic downturns, while food-related costs (FUNR\_16) showed more consistent growth.
2. **Regional Variations**: I observed that cost of living indices varied significantly across different regions. High-income countries generally had higher indices in medical and food costs compared to lower-income countries.
3. **Impact of Economic Events**: Certain economic events, such as global financial crises or pandemics, were correlated with sharp increases in unemployment rates and food costs.

## Assumptions

1. **Data Completeness**: It is assumed that the dataset is complete and accurately reflects the cost-of-living indices for the countries and years specified. Missing or incomplete data could affect the accuracy of the findings.
2. **Consistency Across Years**: I assumed that the method of collecting and reporting cost of living data remained consistent across the years covered in the dataset.
3. **Economic Indicators**: The analysis assumes that the indicators used (UNR\_16, FUNR\_16, MUNR\_16) are reliable and comparable across countries and years.

## Items That Still Need Clarification

1. **Data Collection Methods**: The specifics of how the data was collected and any potential biases in the data collection process need to be clarified.
2. **Missing Data**: Some data points are still missing and further clarification on their potential impact would improve the analysis. More data would also help fill in gaps using averages or other methods.
3. **Economic Context**: Additional context on the economic events and their specific impact on the cost-of-living indices would enhance the interpretation of the data.

# Direction of Story/Plan of Attack/Message You Want to Get Across

The primary message of this analysis is to highlight the significant variations in the cost-of-living indices across different countries and years, emphasizing the impact of economic events on these indices. The story will unfold by first presenting the overall trends and regional variations observed in the data, followed by a deeper dive into the impact of specific economic events. The goal is to illustrate how economic fluctuations, and regional differences influence the cost of living, providing valuable insights for policymakers, researchers, and the public.

# Target Audience

1. **Policymakers**: Individuals involved in economic policy and planning can use these insights to understand the impact of economic events on the cost-of-living and adjust policies accordingly.
2. **Researchers and Economists**: This group will benefit from the detailed analysis and trends identified in the data, which can inform further research and economic studies.
3. **General Public**: Understanding the cost-of-living trends can help the public make informed decisions about their personal finances and understand the broader economic context.

# Mediums Include

1. **Graphs and Charts (PowerPoint)**: Visualizations such as line graphs for trends over time and bar charts for regional comparisons were used to clearly present the data and findings. Visual representations of trends, such as a stacked bar chart comparing childcare costs by state, are included to clearly illustrate key insights.
2. **Written Report**: A detailed written report provides an in-depth explanation of the analysis, findings, and methodologies used. This serves as a comprehensive reference for the target audience to understand the context and implications of the data.
3. **Info Graphic (Canva)**: The infographic is designed to condense complex data into easily understandable visuals for a general audience, combining key statistics, trends, and takeaways. This approach gives a quick and easy way to absorb the data without much required from the audience.

# Desing Decisions

1. **Color**: Distinct colors were used to differentiate between indicators in graphs (e.g., unemployment rates vs. food costs), making it easier for users to distinguish between trends.
2. **Layout**: The PowerPoint and infographic use a structured layout with clear sections (e.g., findings, insights by state) to guide the audience through the information.
3. **Clarity**: Labels, legends, and simple explanations accompany each visualization to enhance understanding for a broad audience.

# Ethical Considerations

1. **Changes Made to Data**: Data was cleaned, and missing values were handled through imputation based on averages to maintain consistency across analyses. In some cases missing values were ignored or filled in with N/A.
2. **Legal or Regulatory Guidelines**: No specific legal or regulatory issues were found for using this public data. The data is from a public source and any guidelines listed were followed.
3. **Risks of Transformations**: Transforming missing values or aggregating data might obscure nuances, so the methodology for handling missing data was clearly documented to avoid misleading conclusions. Some guess work had to be done since I am not familiar with the context of the data.
4. **Assumptions in Data Presentation**: Any assumptions (e.g., filling missing values) were documented to ensure transparency in data interpretation.
5. **Data Sourcing**: The data came from credible sources. I did not find any need for additional data given the in-depth guide accompanying this data.
6. **Ethical Acquisition**: The dataset was publicly available and acquired ethically, with no privacy concerns.
7. **Mitigation of Ethical Concerns**: The use of ethical data sourcing practices and transparent documentation of any modifications made to the dataset ensures integrity in the analysis. These changes should always be made aware to the audience.

# Lessons Learned

**What Would You Do Differently?**

* In the future, I would better understand the data collection methods early on to mitigate uncertainties around potential biases or missing data. I felt learning about the topic I was working on provided me a greater understanding of the data and how best to approach designing visualization elements. I would recommend taking a couple days to research the topic and maybe spending more time understanding the reference material.

**What Did You Enjoy the Most?**

* I enjoyed creating the visualizations and seeing the data come to life, it was difficult at first not understanding what some of the columns in the data represented or understanding how those numbers impacted anything. After learning about the topic, it was fun to make predictions before making a visual and trying to guess what would happen and seeing if my guess was correct or not. Being able to shape the story how I felt best was fun.