The Cost-of-Living Index by Country

DSC530: Data Exploration and Analysis

Staci McDuffie

March 1, 2025

Week 12: Term Project

The primary question in this analysis was: *Does a higher Cost of Living Index correlate with a lower Local Purchasing Power Index?* The initial hypothesis was that countries with a high Cost of Living Index would have lower purchasing power, as higher costs could limit residents' ability to afford goods and services. The analysis aimed to test this assumption using multiple statistical techniques, including correlation analysis, probability distributions, and regression modeling.

The dataset contained information on 121 countries, including variables such as the Cost-of-Living Index, Rent Index, Cost of Living Plus Rent Index, Groceries Index, and Local Purchasing Power Index. The EDA revealed that while the Cost-of-Living Index and Local Purchasing Power Index have a moderate positive correlation (0.69, $p < 0.001$), the assumption that higher living costs necessarily reduce purchasing power was not always valid. Some high-cost-of-living countries, such as Switzerland, exhibited high purchasing power, whereas others, like Singapore, did not.

Histograms and descriptive statistics indicated right-skewed distributions in most indices, with significant outliers. A Probability Mass Function (PMF) comparison between high and low-cost-of-living countries showed that high-cost nations had widely varying

purchasing power levels. The Cumulative Distribution Function (CDF) highlighted that nearly 50% of countries had a purchasing power index below 50, emphasizing economic disparity.

A multiple linear regression model found that the Cost-of-Living Index and Rent Index significantly affected purchasing power, while the Groceries Index did not. The model explained 52.4% of the variance in the Local Purchasing Power Index. This suggests that rent costs contribute significantly to variations in purchasing power, while grocery costs alone do not provide a strong predictive factor.

One aspect that could have improved the analysis was incorporating income levels or GDP per capita, as these factors directly impact purchasing power. Additionally, analyzing regional trends might have provided more insights into how different continents compare in terms of cost of living and purchasing power.

Further, the dataset did not account for government subsidies, taxation policies, or cost-of-living adjustments, which may have influenced the purchasing power of residents in various countries. The impact of exchange rates and inflation was also missing, both of which could significantly alter the real purchasing power of a population.

Including tax rates, inflation levels, employment rates, and average wages would have provided a more comprehensive view of economic conditions in each country. The dataset only focused on cost indices, which do not fully capture income disparities and wealth distribution. For example, two countries with similar cost of living indices might have drastically different purchasing power due to differences in wages and disposable income levels.

The initial assumption that a high cost of living always leads to low purchasing power was proven incorrect. Some wealthy nations with high costs also have high wages, maintaining strong purchasing power. Furthermore, the assumption that grocery costs are a major determinant of purchasing power was incorrect, as evidenced by the regression analysis showing a lack of significance for the Groceries Index.

One challenge was interpreting the regression results in terms of real-world economic dynamics. While the model indicated statistical significance, external factors like government policies, labor market structures, and exchange rate fluctuations also influence purchasing power but were not accounted for. Additionally, ensuring data normalization to reduce the impact of outliers posed difficulties. Some countries, such as Switzerland, showed extremely high purchasing power, which may have skewed the regression results.

Another challenge was deciding how to handle missing contextual factors, such as income distribution, wealth inequality, and social safety nets, which could drastically influence purchasing power and cost-of-living dynamics.

In conclusion, this analysis provided valuable insights into the relationship between cost of living and purchasing power. While higher living costs do not necessarily mean lower purchasing power, rent expenses significantly impact economic well-being. Further research incorporating income data, taxation, inflation, and employment rates could refine these findings and provide a clearer picture of global economic conditions. Additionally, future studies could explore causality rather than just correlation, using time-series data to assess how changes in cost-of-living impact purchasing power over time.