**Comparative Analysis of U.S. City Average CPI: Inflation Trends, Prices, and Compensation Dynamics**

**Purpose of the Analysis**

This comparative analysis explores the relationship between key inflation trends indicators (Consumer Price Index (CPI), Personal Consumption Expenditures (PCE), Producer Price Index (PPI), and GDP Deflator) by analyzing the U.S. City Average Consumer Price Index (CPI). The study aims to build a predictive model using linear regression to forecast CPI based on these indicators, showcasing how the rising costs of goods impact compensation trends, particularly with inflation, and whether pay adjustments are aligned with price increases.

**Key Findings:**

There is a significant correlation between CPI and the other inflation indicators. The linear regression model predicted cpi with a reasonable level of accuracy. The mean squared error (MSE) and R-squared values suggest the model has predictive potential.

**Key Indices Analyzed from Federal Reserve Economic Data (FRED)**

* **CPI: Consumer Price index**
* **PPI: Producer Price Index for All Commodities**
* **PCE: Personal Consumption Expenditures Price Index**
* **GDP: Gross Domestic Product Deflator**

**Research Questions**

**1.** Is there a correlation between the Consumer Price Index and inflation indices (PCE, PPI, AND GDP Deflator)

**2.** Can PCE, PPI, AND GDP Deflator be used to predict CPI using a linear regression model?

**3.** How well does a linear regression model predict CPI compared to actual CPI values?

**4.** What is the magnitude and direction of the relationship between CPI and the predictor variables (PCE, PPI, and GDP Deflator)?

**Hypotheses to test**

* H1: There is a significant correlation between CPI and the other inflation indices (PCE, PPI, and GDP Deflator)
* H2: PCE, PPI, and GDP Deflator can significantly predict CPI in a linear regression model.
* H3: The Linear regression analysis with statistical tests of significance for each predictor variable’s coefficient (p-values)
* H4: The coefficients of the linear regression model are non-zero, indicating that each predictor variable (PCE, PPI, GDP Deflator) has a measurable impact on CPI.

1. **Correlation between CPI and other inflation indices (PCE, PPI, and GDP Deflator)**

Using correlation analysis to determine the strength and direction of relationships between the variables we found that the null hypothesis (H0) showed no significant correlation between CPI and the other inflation indices (PCE, PPI, and GDP Deflator). Alternatively, (H1) showed a significant correlation between CPI and the other inflation indices (PCE, PPI, and GDP Deflator).

1. **PCE, PPI, and GDP Deflator can significantly predict CPI in a linear regression model.**

Using linear regression analysis with statistical tests of significance for each predictor variable’s coefficient (e.g., p-values). The test allowed us to see that the null hypothesis (H0): PCE, PPI, and GDP Deflator do not significantly predict CPI in a linear regression model (i.e., their coefficients are zero). Alternatively, PCE, PPI, and GDP Deflator significantly predict CPI in a linear regression model (i.e., at least one coefficient is non-zero).

1. **The linear regression model provides a good fit to the data, as indicated by R-square and MSE.**

Using the evolution of model performance metrics, specifically R-squared and mean squared Error (MSE) showed us that the null hypothesis (H0) linear regression model does not provide a good fit to the data (i.e., R-squared is low, and MSE is low). Alternatively, the hypothesis (H1) linear regression model provides a good fit to the data (i.e., R-squared is high, and MSE is low).

1. **The coefficients of the linear regression model are non-zero, indicating that each predictor variable (PCE, PPI, GDP Deflator) has a measurable impact on CPI.**

Examining the regression coefficients and their statistical significance (p-values). *The null hypothesis (H0) shows that the coefficients of the linear regression model are non-zero, indicating that each predictor variable (PCE, PPI, GDP Deflator) has a measurable impact on CPI.* Alternatively, hypothesis (H1) coefficients of PCE, PPI, and GDP Deflator are non-zero, indicating that these variables have a measurable effect on CPI.

**Data Visualization Used:**

1. Line Plots of the four inflation indices (CPI, PCE, PPI GDP Deflator) over time helped visualize the trends in each indicator.
2. Descriptive statistics (mean, standard deviation, min, max, etc.) were calculated for each index to summarize their distribution for CPI, PCE, PPI, and GDP Deflator.
3. Scatter Plot to show the uprising trend over the period 1982-1984 for CPI, PCE, PPI, and GDP Deflator. Also, used to plot the actual vs. predicted CPI values to visualize how well the model’s predictions align with the true values.
4. A correlation analysis matrix was computed to assess the relationships between the variables using a heatmap to show the strong correlations between CPI and other indicators, particularly PCE and PPI.
5. A Linear regression model was used to predict CPI based on PCE, PPI, and GDP Deflator. This model was chosen due to its simplicity and interpretability.

**Performance of T-Test against two groups/ One-Sample T-Test**

**Interpretation and Insights**

Our null hypothesis for both the T-Test and One-sample T-test must be rejected. We must go with the alternative hypothesis and there are some statistically significant factors that are correlated with inflation and prices and should be further investigated.

**General Insights:**

Two-Sample Tests: There are significant differences in the means of CPI compared to PCE, PPI, and GDP Deflator. This suggests that while these indices are related, they capture distinct aspects of economic activity.

One-Sample Tests—The means of CPI, PCE, PPI, and GDP Deflator are all significantly different from the hypothesized mean of 100, suggesting that none of the indices are centered around this value in the dataset.

Recommendations for interpretation—The significant difference implies that these inflation indices, while related, should not be used interchangeably without careful analysis.

These results support the need for *distract* modeling for each variable to capture its unique behavior.

**Results and interpretation**

**Prediction Performance:**

The model successfully predicted the CPI, with an R-squared value of [Value], suggesting that the independent variables (PCE, PPI, and GDP Deflator) explain a significant proportion of the variability in CPI.

A scatter plot of actual vs. predicted CPI values showed that the model’s predictions were reasonably close to the values.

**Coefficient Insights:**

The positive/negative coefficients indicate that for each unit increase in the independent variables, CPI is expected to increase/decrease by the respective coefficient value. For example, if the coefficient for PCE is [Value], then a 1 unit increase in PCE is associated with a [Value] increase in CPI.

**Conclusion**

**Summary of findings:**

In Conclusion, the analysis revealed strong correlations between CPI and the other price indices (PCE, PPI, GDP Deflator). The linear regression model was effective in predicting CPI, though the accuracy of the predictions could be improved by exploring more complex models or additional features.

**Next Steps:**

Further refinement of the model could include the use of additional economic indicators or the application of more advanced modeling techniques, such as time series analysis (e.g., ARIMA models) or machine learning models. Future work could also involve exploring seasonal or cyclical effects in inflation data.