Project Report: Exploring the Relationship between Employment Rate and Suicide Rates

I. Introduction

The project aims to investigate the correlation between employment rates and suicide rates globally, addressing the impact of economic factors on mental health using the gapminder dataset. In light of increasing automation and economic challenges, understanding the impact of employment on mental health is crucial.

II. Methodology

- **Data Cleaning and Preprocessing:**
  - Cleaned the dataset, handled missing values, and converted relevant columns to numeric formats.
  - Focused on key variables: employment rate, suicide rate per 100 people, female employment rate, and income per person.

- **Variable Grouping and Exploration:**
  - Grouped main variables into quartiles or percentiles for effective organization.
  - Utilized descriptive statistics, frequency distributions, and visualizations for data exploration.

- **Univariate and Bivariate Analysis:**
  - Employed histograms for univariate analysis and scatter/regression plots for bivariate insights.
  - Calculated the Pearson correlation coefficient, revealing a weak positive correlation.

- **Chi-Square Tests:**
  - Applied chi-square tests to explore relationships between employment rates and suicide rates in grouped categories.

- **Logistic Regression Analysis:**
  - Conducted logistic regression, considering binary outcomes based on suicide rates above or below the mean.

- **Polynomial Regression Analysis:**
  - Utilized polynomial regression to assess potential non-linear relationships between variables.

III. Results

- **Statistical Tests:**
  - Chi-square tests indicated significant associations, particularly in lower employment rate countries.
  - Logistic regression revealed odds ratios, suggesting lower suicide rates in lower employment rate countries.

- **Regression Models:**
  - Polynomial regression, specifically quadratic models, showed stronger associations compared to linear models.

IV. Model Evaluation and Additional Considerations:

- Evaluated models using residual plots and leverage plots.
- Acknowledged potential confounding variables and limitations in the analysis

V. Conclusion:

The exploration into the relationship between employment rates and suicide rates yielded nuanced findings. While a weak positive correlation was observed, lower employment rate countries tended to exhibit lower suicide rates. Limitations and potential biases were acknowledged.