Summary of Insights Report

1. Basic Data Exploration:

- The dataset contains [148654] rows and [13] columns.
- Data types for each column have been identified, and missing values have been checked.

2. Descriptive Statistics:

- Mean Salary: [74768.321971703]
- Median Salary: [71426.60999999999]
- Mode Salary: [0.0]
- Minimum Salary: [-618.13]
- Maximum Salary: [567595.43]
- Salary Range: [568213.56]
- Standard Deviation of Salary: [50517.005273949944]

3. Data Cleaning:

- Missing values in the 'Benefits' column have been handled by [imputation], ensuring the dataset is ready for analysis.

4. Basic Data Visualization:

- A histogram has been created to visualize the distribution of salaries, showing [shape/characteristics of the distribution].
- A pie chart illustrates the proportion of employees in different departments.

5. Grouped Analysis:

- The dataset has been grouped by 'JobTitle,' and summary statistics for each group have been calculated, revealing insights into the salary distribution within each department.

6. Simple Correlation Analysis:

- A scatter plot has been generated to explore the relationship between salary and another numerical column ('BasePay'). The correlation coefficient is [0.954936474296026], indicating [strong] correlation.

Conclusion:

- The analyses conducted provide a comprehensive understanding of the dataset, highlighting key patterns, trends, and relationships.

Name:Shefaa Ahmad Mestarihi

Yarmouk Yniversity