

**Lamia khalid**

Summary Report: Analysis of Salaries.csv

## **Introduction:**

This report presents the findings and insights obtained from the analysis of the dataset Salaries.csv. The dataset contains information on employee salaries, including columns such as 'Id', 'EmployeeName', 'JobTitle', 'BasePay', 'OvertimePay', 'OtherPay', 'Benefits', 'TotalPay', 'TotalPayBenefits', 'Year', and 'Agency'. And the two deleted columns The objective of the analysis was to gain insights into salary patterns, identify correlations, and uncover any notable trends within the dataset.

## **Descriptive Statistics:**

Descriptive statistics were calculated for relevant columns in the dataset. The statistics for key variables are as follows:

- 'BasePay': The average base pay across all employees was 66325.44884050643, with a minimum of -166.01 and a maximum of 319275.01.
- 'OvertimePay': The average overtime pay was 5066.0598864 with a minimum of -0.01 and a maximum of 245131.88
- 'OtherPay': The average other pay was 3648.7672968046095, with a minimum of -7058.59 and a maximum of 400184.25.
- 'Benefits': The average benefits received by employees were 18924.232838874363, with a minimum of -33.89 and a maximum of 96570.66.
- 'TotalPay': The average total pay, which includes base pay, overtime pay, and other pay, was 74768.321971703, with a minimum of -618.13 and a maximum of 567595.43.

### **Group Analysis:**

The data was grouped by 'JobTitle' to gain insights into salary variations across different job titles. The analysis revealed significant disparities in average salaries among various job titles. For example, employees in management positions tended to have higher average salaries compared to those in administrative or support roles. This finding suggests that job title plays a crucial role in determining salary levels within the organization.

### **Correlation Analysis:**

A correlation analysis was performed to examine the relationship between 'TotalPay' and BasePay. The analysis identified a positive correlation between 'TotalPay' and 'BasePay', indicating that employees with higher base pay tend to receive higher total pay. However, no significant correlations were found between 'TotalPay' and 'OvertimePay' or 'OtherPay' both of them were less than 0.5. This suggests that base pay has a stronger influence on overall compensation compared to overtime pay or other pay components.

### **Scatter Plot:**

A scatter plot was created to visualize the relationship between 'TotalPay' and 'BasePay'. The plot displayed a positive linear relationship, indicating that as base pay increases, total pay also tends to increase. This finding further supports the positive correlation observed in the correlation analysis.

### **Conclusion:**

The analysis of the Salaries.csv dataset provided valuable insights into salary patterns within the organization. It revealed variations in average

salaries across different job titles, with higher average salaries observed in management positions. The correlation analysis highlighted the strong positive relationship between base pay and total pay, indicating the significance of base pay in determining overall compensation. The scatter plot visually confirmed this relationship, showing a positive linear trend between 'TotalPay' and 'BasePay'.