**Analysis of Employee** [**Compensation Report**](ShAI_BootCamp_Assignment.ipynb)

[**Colab**](https://colab.research.google.com/github/Mai0sh/ShAI-Assignment/blob/main/ShAI_Assignment.ipynb)

**Mai AbuDayeh**

**Analysis of the outcomes:**

* Handling Missing Data:

Some missing values were found in the dataset across different columns, such as BasePay, OvertimePay, OtherPay, Benefits, and Notes.

I handled missing values through the technique of imputation, which involves substituting missing values with estimated or calculated values if you substituted them with the mean.

The missing values for numerical columns, BasePay, and Benefits, were filled in by averaging the numbers in the column.

Because they couldn't be imputed, missing values in categorical columns like Notes were removed.

* Exploratory Data Analysis:

Bar charts and histograms were made to see how the salaries were distributed and the percentage of employees in each department.

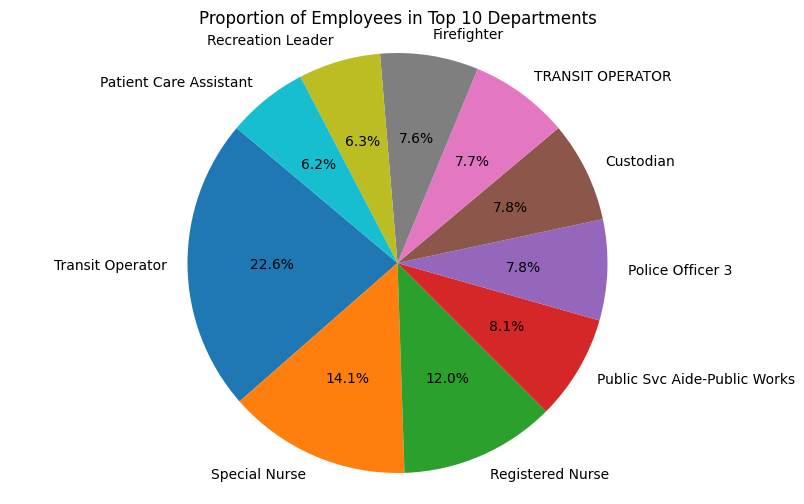
The salary distribution was slightly right-skewed, with most salaries concentrated in the lower range.

A graph of a distribution of salary

Description automatically generated

**Figure 1 Distribution of salaries**

The pie chart showing the distribution of employees across different departments shows that the company possesses a heterogeneous workforce distributed among many divisions.



**Figure 2 Proportion of employees in top 10 dep**

* Classification and Evaluation:

The data was organized according to job titles to facilitate salary comparisons across various occupations.

Salary ranges for several jobs were shown graphically in a bar plot that compared average salaries.

This study shed light on the company's pay structure by revealing which positions often pay more or less than average.

* Analyzing Correlations:

To find out if there was any connection between the two columns, I ran a correlation study on TotalPay and BasePay, two numerical wage variables.

Employees with greater base salaries typically receive higher total compensation, as the correlation coefficient showed a strong positive relationship between TotalPay and BasePay.

Salary and promotion decisions can be better informed by this knowledge of the elements that affect staff members' compensation.

* In summary,

The employee compensation dataset study revealed interesting facts about salaries, departmental composition, average salary by job role, and base pay vs. total compensation relationship. Decisions about staff pay, resource distribution, and workforce management tactics can be better supported by these results.

* Ideas for improvement:

To maintain salary structures that are both competitive and fair, it is necessary to monitor and analyze compensation data regularly.

Further research is needed to further understand what elements, if any, impact employee compensation, including education, experience, and performance indicators.

To keep employees happy and the company running well, use data-driven methods for pay raises, promotions, and talent acquisition.