

Data Analysis on Employee Attrition Dataset

Name: - Vaishnav Rajith Kalathil

Repository: - <https://github.com/zk24aao/Employee-Attrition.git>

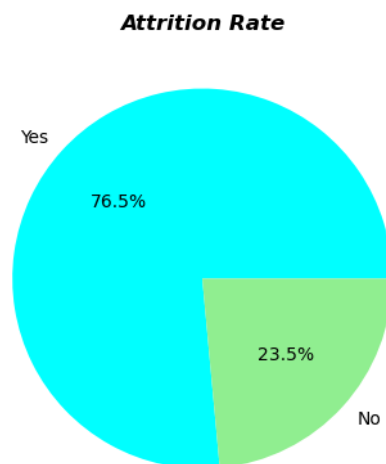
INTRODUCTION: -

Employee Attrition dataset, in this report there are few basics information of Employee Demographics, Job roles, Education field, Income etc. The report mainly focusses on Attrition Rate, Monthly Income, Age, Gender. So that we understand the reason behind increasing attrition rate. There are three types of graphs which is Relational, Categorical, Statistical Graphs. The report primarily concentrates on Pie Chart, Heatmap Correlation Matrix, Scatter Graph.

Plot 1: Pie Chart – Attrition Rate

The first plot is Pie which visualises the Attrition rate.

The Attrition rate is often used to analyse workforce reduction by showing the percentage of employees who left over a specific period.

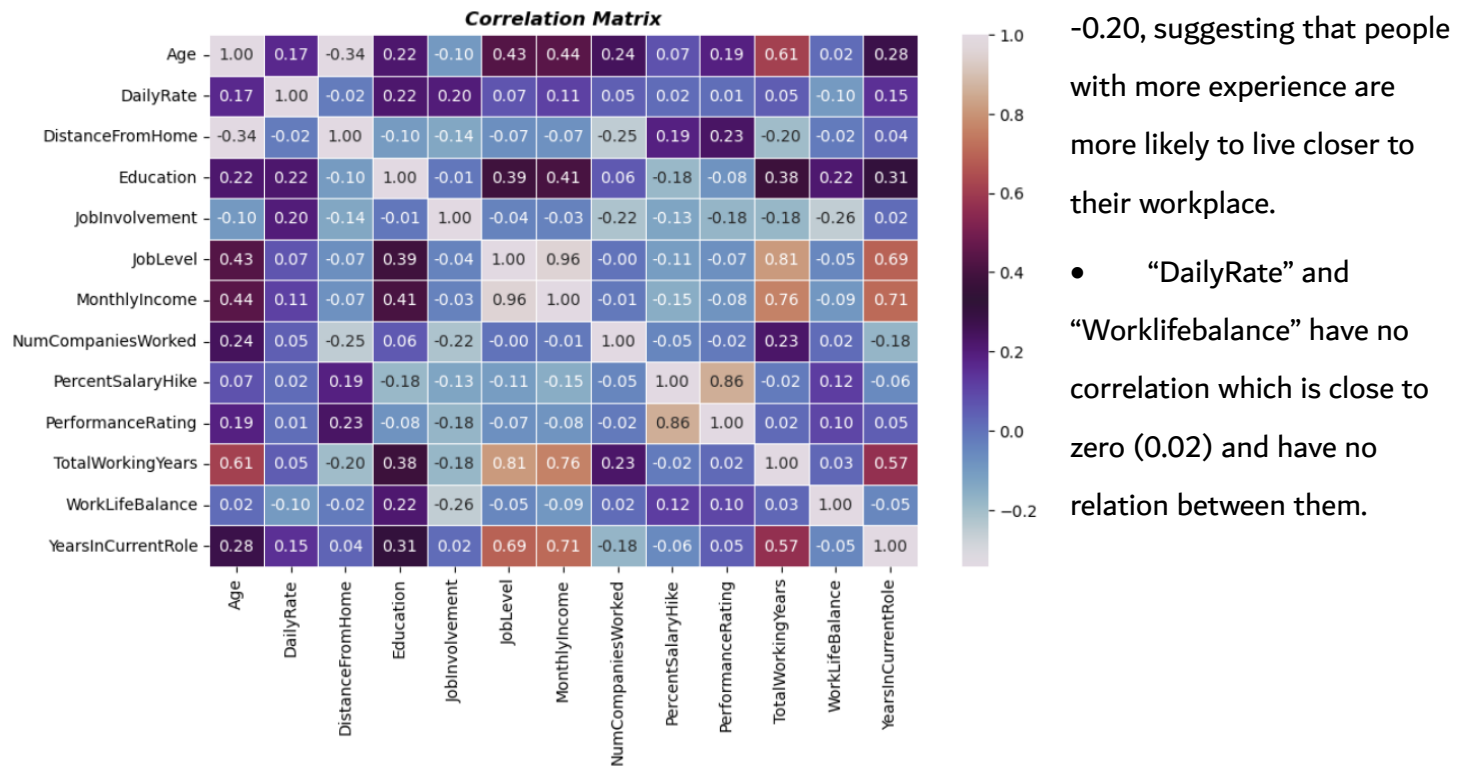


- The Chart clearly shows the majority of dataset has 'Yes' with 76.46% typically indicates that an employee has left the organization, whereas 'No' with 23.53% means the employees is still willing to work in the company.
- 'Yes' has been the dominant category.

Plot 2: Heatmap Correlation Matrix

This correlation matrix shows how different variables in the dataset are related to each other, with values ranging from -1 to +1

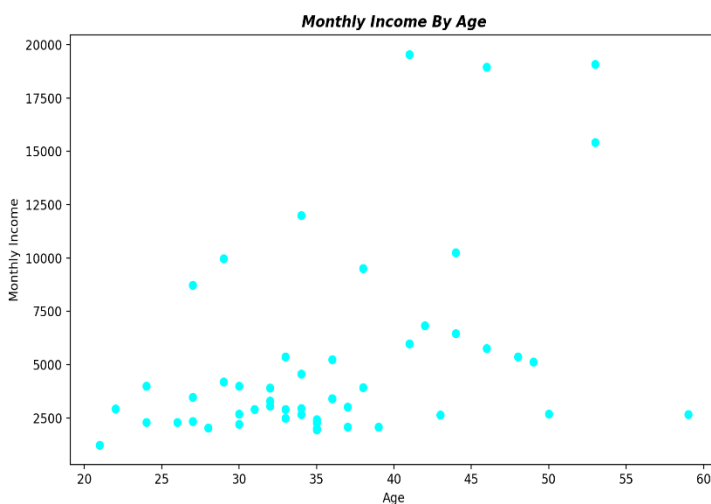
- We can see that "Totalworkingyears" and "MonthlyIncome" have strong positive relation of 0.76, means that people with more work experience have the potential to earn higher incomes.
- "DistanceFromHome" and "TotalWorkingyears" have a weak negative correlation of



Plot 3: Scatter Plot- Monthly Income by Age

The third plot is scattering plot that examines the relation between employee *Monthly Income* (Y axis) and *Age* (X axis). It helps us assess if age influences income distribution.

- From the plot, we can see a mix of income levels across various groups. It indicates that younger employees relatively have lower incomes.



- We can see a weak Positive correlation, but some employees with certain roles or seniority levels seem to stand out from the general pattern which are called as outliers.
- An intensity of value can be found in third quadrant which shows majority of people earn in range of 2000- 5000.

Conclusion:

Through these plots we can understand the reasons of increasing attrition: consistent salary, low income, more distance to work, there is weak correlation between income and age and their work satisfaction. Females have variability while male have consistent income. The income rate is not increasing with experience and age. The satisfaction rate is decreased which gave rise to quit jobs.