Dataset Overview: Response variable is Attrition with 35 features of a combination of numerical and categorical. There are 1470 data entry with clean and no missing data however there is only 237 – 16% "No" responses for Attrition which is rather low to the typical attrition rate.

**Demographics Analysis**

Deep diving demographic categories such as gender, age, marital status, and education. Through this analysis, we identified the following regarding attrition:

* **A graph of different age groups

  Description automatically generated with medium confidenceGender:** While the company has an imbalanced male-to-female ratio, the percentage of attrition is marginal between both genders.
* **Age:** Grouping ages into different age groups, we observed a trend of younger individuals 18-25 followed by those aged 26-35 having a higher attrition rate compared to other age groups.
* **Marital Status:** Singles have a higher attrition rate compared to divorced and married individuals. This could stem from singles having fewer responsibilities and commitments, allowing for more change.
* **Education:** No distinct visual trends were detected, with a slightly lower attrition rate for individuals with a “Doctor”.

**Job Motivation Analysis**

A graph of different sizes and numbers

Description automatically generated with medium confidenceIn our next phase, we delved into core and internal job-related metrics to uncover visual relationships with attrition. Through this analysis, we noticed (Note: For scoring categories 1=weak and 4=Best) :

* **Work-Life Balance:** A notable trend emerged, showing higher attrition among individuals with lower scores for work-life balance.
* **Number of Companies Worked For:** We observed a higher attrition rate among individuals at their second job, followed by a plateau before seeing higher attrition among "serial job movers."
* **Job Satisfaction:** There was a clear negative correlation between attrition and higher job satisfaction.
* **Environment Satisfaction:** We found that the lowest score in this category was associated with a significant attrition rate, while other scores did not indicate clear trends.

**Performance Indication Analysis**

Finally, we delved into the monetary motivations of employees to uncover insights regarding attrition and their compensation packages:

* A graph of different sizes and numbers

  Description automatically generated with medium confidenceMonthly Income Group: We observed a substantially higher attrition rate among lower income earners in the group. This trend may indicate a search for better opportunities.
* Percentage Hike: There were no clear indications of trends related to attrition and its impact.
* Performance Rating: Similarly, there were no clear indications of trends related to attrition.
* Stock Option Level: both the highest and lowest levels were associated with the highest attrition rates, possibly due to disappointment or high performers seeking better opportunities externally.

**Correlation Matrix across the variables**

A white square with blue squares and red text

Description automatically generated

While establishing a correlation matrix for the non-categorical/text data, we identified variables with high correlation. These variables are primarily related to years within the company, manager, current role, and total working experience. Therefore, we should select certain variables and avoid selecting duplicates, as this may result in the independent variables incorrectly influencing the response.

Based on our initial high-level analysis, we've found that performance indicators may not independently drive attrition significantly. However, demographics and job motivation appear to have a strong relationship with attrition. This discovery provides a solid foundation for more advanced analysis.