**Hiring Process Analytics**

**Project Description**

**Taking the role of a Data analyst at a multinational company, my task is to analyse the company’s hiring process data and draw meaningful insights from it. The hiring process is a crucial function of any company, and understanding trends such as the number of rejections, interviews, job types, and vacancies can provide valuable insights for the hiring department.**

**Approach**

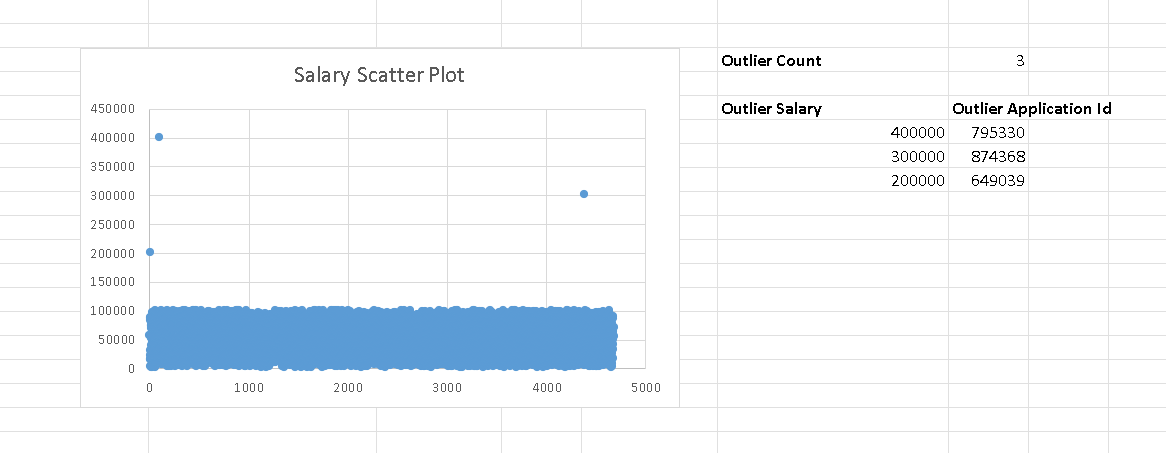
1. **Handling Missing Data:**

**Missing values are left as such as the population size for the data analysis is big enough that we can leave them as such. Infact, they are removed in the next step involuntarily.**

1. **Clubbing Columns:**

**All the columns from the original data is remade except this time only the data with the status as “Hired” is considered so as to make the data more relevant to the company’s current situation.**

1. **Outlier Detection:**



**Upon Scatter Plotting the salary, we can find the 3 extreme Salary, we use an Index Match to find the Application Id from the 3 Salary we obtained from the plotting.**

1. **Removing Outliers:**

**The outliers are kept as such instead of disregarding.**

1. **Data Summary:**

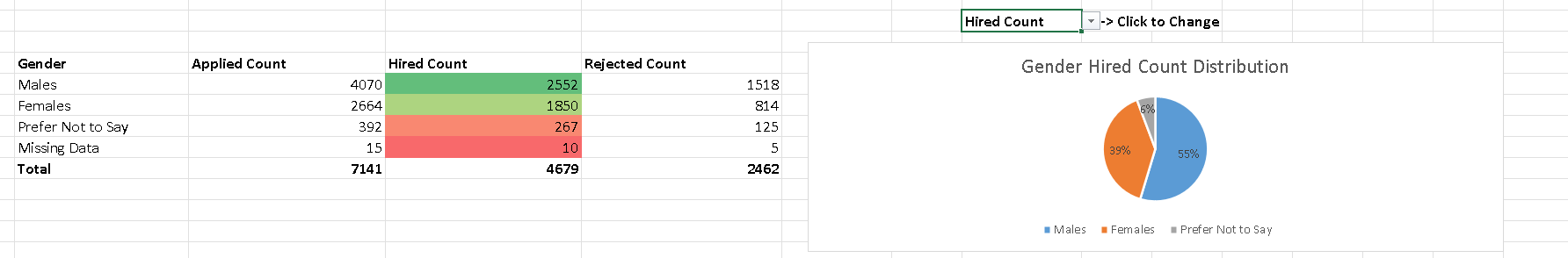
**The new data contains the knowledge of outliers as well as a cleaned data table of the hired candidates to perform data analysis on.**

**Tech-Stack Used:**

**Microsoft Excel**

**Result:**

**The Gender Ratio/Distribution is obtained as a Pie Chart from the Data.**

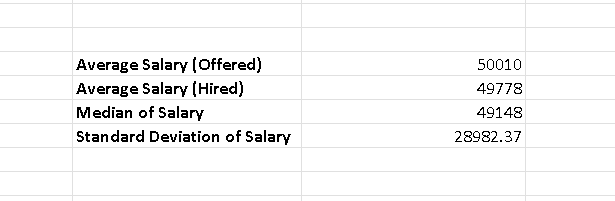


**Insights:**

1. **The Number of Males Applied is significantly more than the Number of Females Applied.**
2. **The Number of Females Hired from them is significantly more than the Number of Males Hired from the applications.**
3. **This indicates that the Females applied are more capable, or that simply females perform better in the job selection process.**

**Result:**

**The Mean, Median and the Standard Deviation of the Salary is found.**

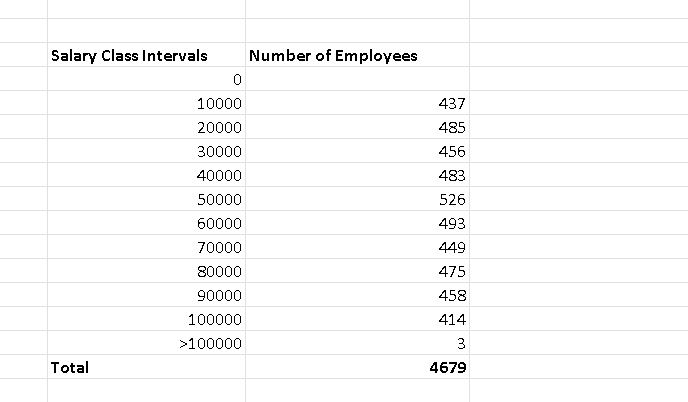


**Insights:**

1. **The mean and median being close to each other suggest that the distribution might be approximately symmetric.**
2. **The standard deviation being relatively large compared to the mean indicates a significant amount of variability or dispersion in the dataset. This is due to the presence of 3 outliers found in the start of the project and can be mitigated by removing or replacing them.**

**Result:**

**The Number of Hired Employees in each Class Intervals of Salary is obtained.**

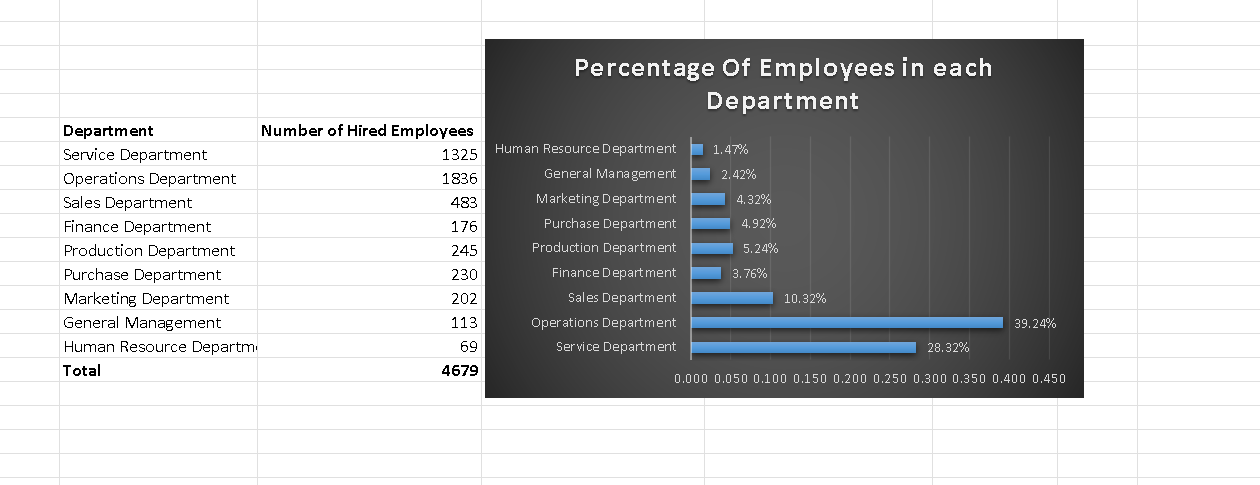


**Insights:**

1. **The Distribution of Salary is almost equal in all the intervals, but the highest interval of Salary is 40000-50000.**
2. **There are 3 salary that are above 100000, these are the 3 outliers found in the start of the project.**

**Result:**

**The distribution of employees in each department is obtained.**



**Insights:**

1. **The Number of employees is highest in the Operations Department.**
2. **The Number of employees is lowest in the Human Resources Department.**

**Result:**

**The Number of Employees in each Position is obtained.**



**Insights:**

1. **The most populated position in the company is the c9 position and then the c5 with more than 1000 employees each.**
2. **There are 2 positions with only 1 member which seems to executive positions, i.e. n9 and n6.**
3. **There are 3 positions with no members.**

**Drive Link:**

[**https://docs.google.com/spreadsheets/d/1RjaeKsmBjwYnIQRIkH7lHKqt4xc7oCbP/edit?usp=sharing&ouid=106630094537104676920&rtpof=true&sd=true**](https://docs.google.com/spreadsheets/d/1RjaeKsmBjwYnIQRIkH7lHKqt4xc7oCbP/edit?usp=sharing&ouid=106630094537104676920&rtpof=true&sd=true)

**!!Analysis at the Bottom of the Data Table!!**