HIRING PROCESS ANALYTICS

Project Description:

The goal of this project is to examine the underlying trends and patterns in the hiring process, providing a comprehensive understanding of key metrics such as the number of rejections, number of interviews conducted, types of jobs, and overall job vacancies.

By analysing these trends, the aim is to derive actionable insights that can assist the hiring department in improving their strategies and making informed decisions when hiring freshers or individuals for various roles within the organization.

Areas of analysis in this project included-

1. Number of male and female candidates hired
2. Average salary offered
3. Class intervals of salary offered
4. Department-wise employee distribution
5. Different tiers of post in the company

Approach:

For this project I downloaded the dataset and worked on Microsoft Excel to analyse the data and visualize the outcomes of the analysis.

The steps included-

- Understanding the columns and the data in them
- Identifying null and duplicate values
- Grouping columns with multiple entries in different categories using pivot tables
- Identifying outliers and working around them to analyse the data using excel built-in functions and pivot tables
- Using charts and graphs to visualise trends and draw conclusions

Tech-Stack Used:

Application used for this project was Microsoft Excel which is an excellent tool to perform a range of statistical analysis of data and visualise the outcomes.

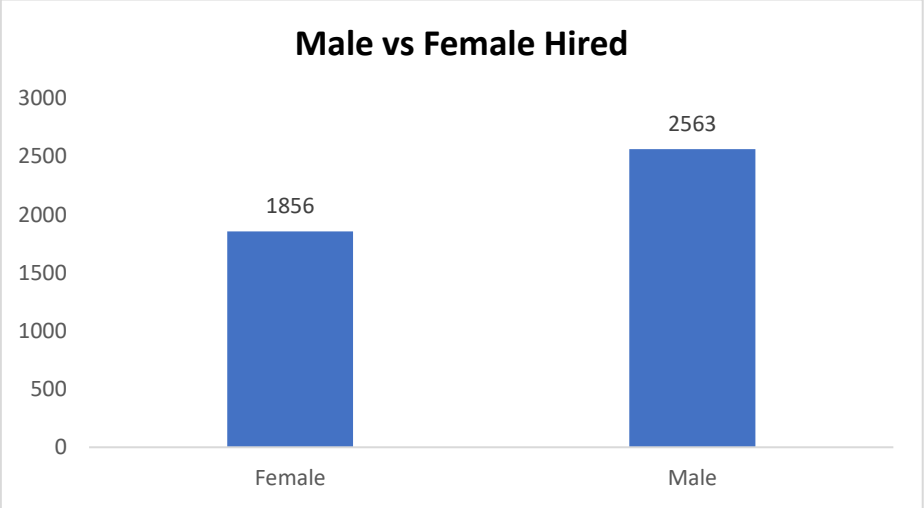
Insights:

Exploratory data analysis of the given dataset provided the following insights:

1. **Hiring:** A grand total of 7168 interviews were conducted out of which there were 4697 hirings. Among the hired candidates the distribution of male and female individuals was found to be-

- Male: 1856
- Female: 2563

Gender	Hired	Rejected	Grand Total
-	10	5	15
Don’t want to say	268	125	393
Female	1856	819	2675
Male	2563	1522	4085
Grand Total	**4697**	**2471**	**7168**



The number of males hired was observed to be more than the female hires.

2. **Average Salary:** The average salary offered by the company was found using the built-in 'AVERAGE' function in Excel and the result is as follows:

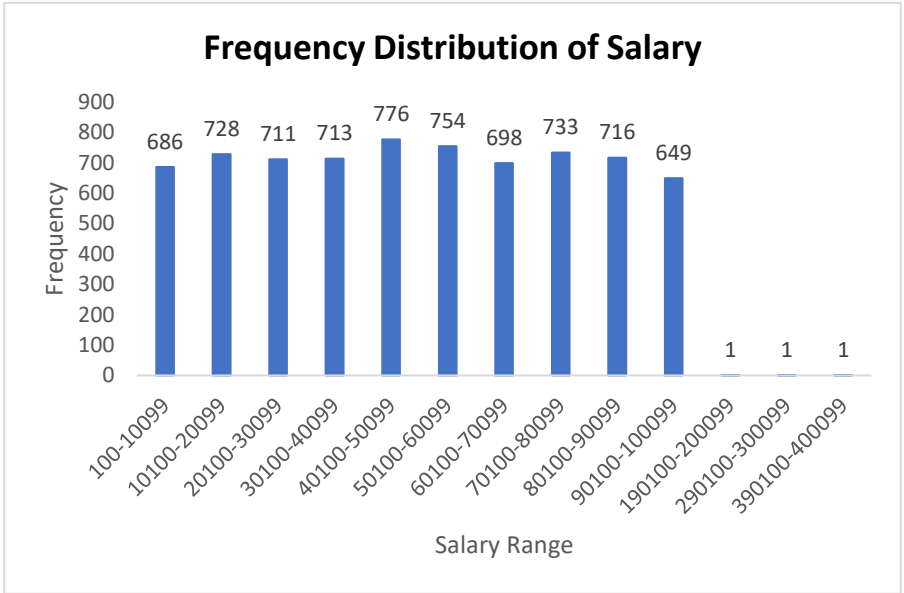
Average Salary	49983.02902	AVERAGE(G2:G7169)

3. **Class Interval:** To determine the class interval of the salary offered in the company, the difference between the maximum and minimum salary offered was calculated as shown below:

Salary max	400000	MAX(G2:G7169)
Salary min	100	MIN(G2:G7169)
Class interval	399900	MAX(G2:G7169)-MIN(G2:G7169)

To find out the frequency distribution of different salary ranges a histogram was plotted using the grouping function of the pivot table to club multiple entries for different salary ranges as shown below:

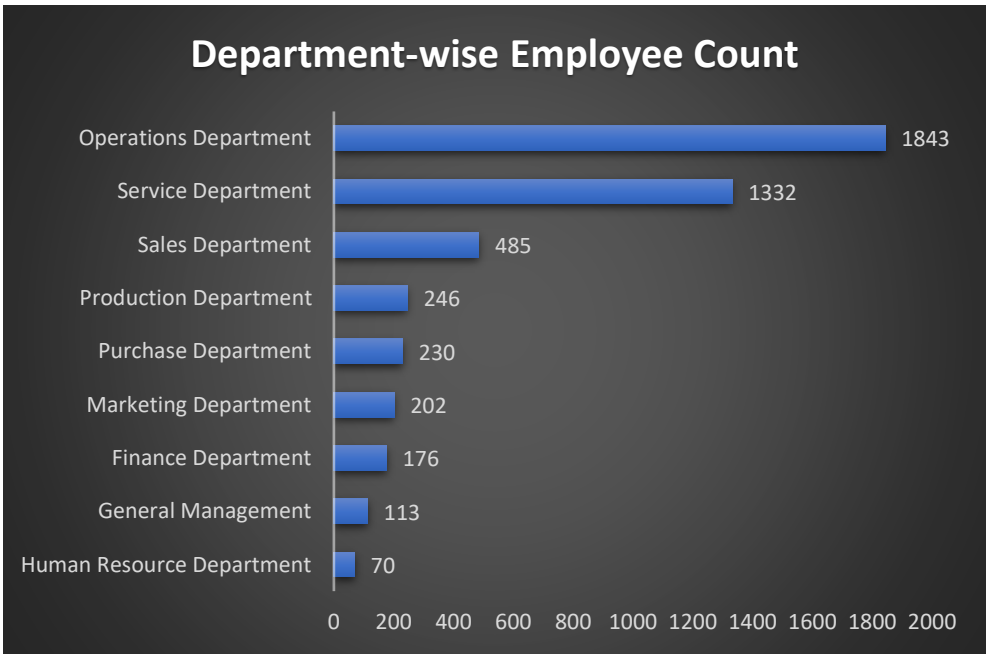
Salary Range	Frequency
100-10099	686
10100-20099	728
20100-30099	711
30100-40099	713
40100-50099	776
50100-60099	754
60100-70099	698
70100-80099	733
80100-90099	716
90100-100099	649
190100-200099	1
290100-300099	1
390100-400099	1
Grand Total	**7167**

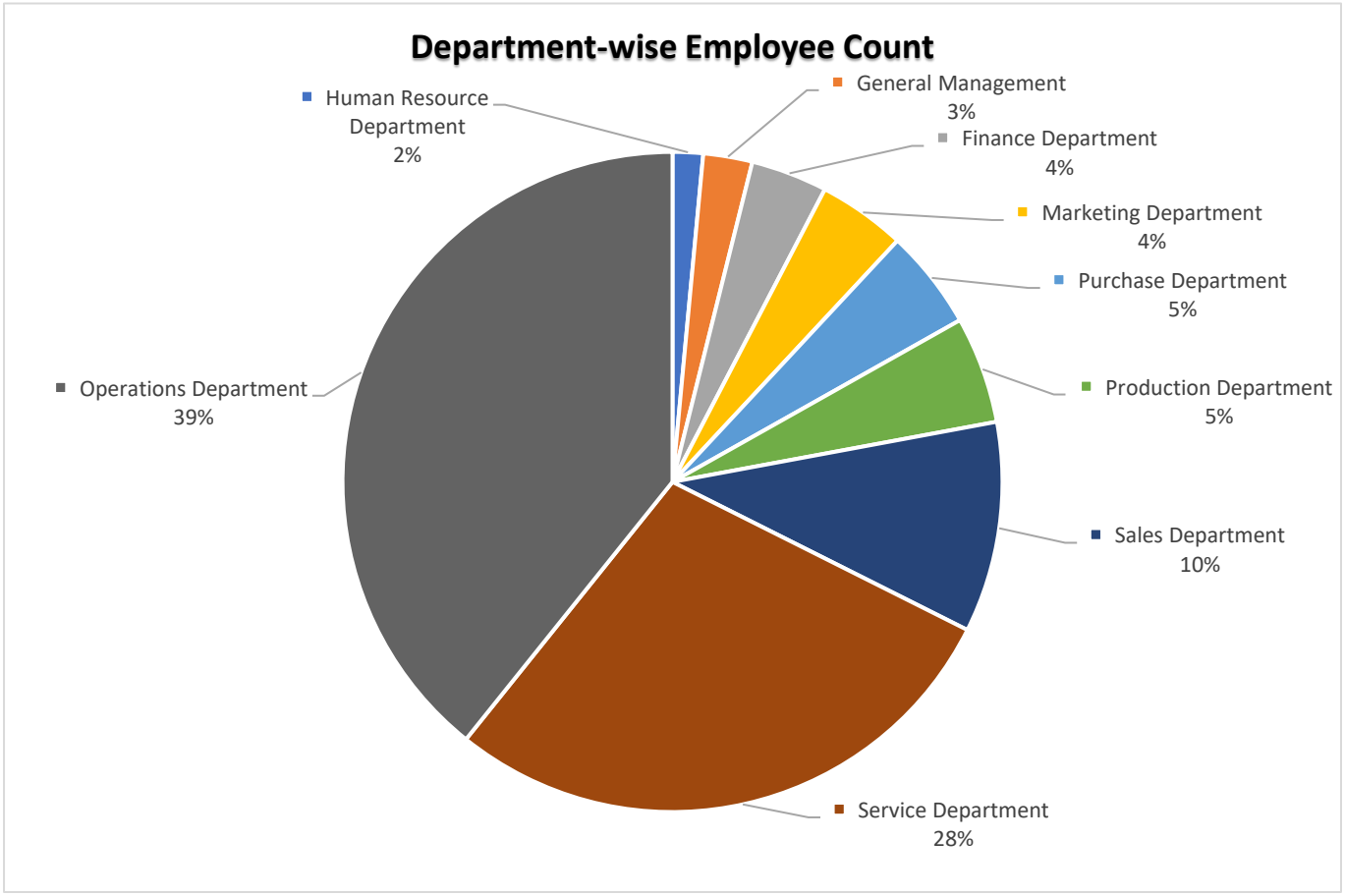


Most of the salary offered to employees belonged to the range of 40100-50099 and very few employees had salary offered in the upper bracket above 100099.

4. **Department-wise Employee Distribution:** To determine the employee count of different departments, the number of employees working (hired) in different departments was determined using pivot table and the result is as follows:

Departments	Hired
Human Resource Department	70
General Management	113
Finance Department	176
Marketing Department	202
Purchase Department	230
Production Department	246
Sales Department	485
Service Department	1332
Operations Department	1843

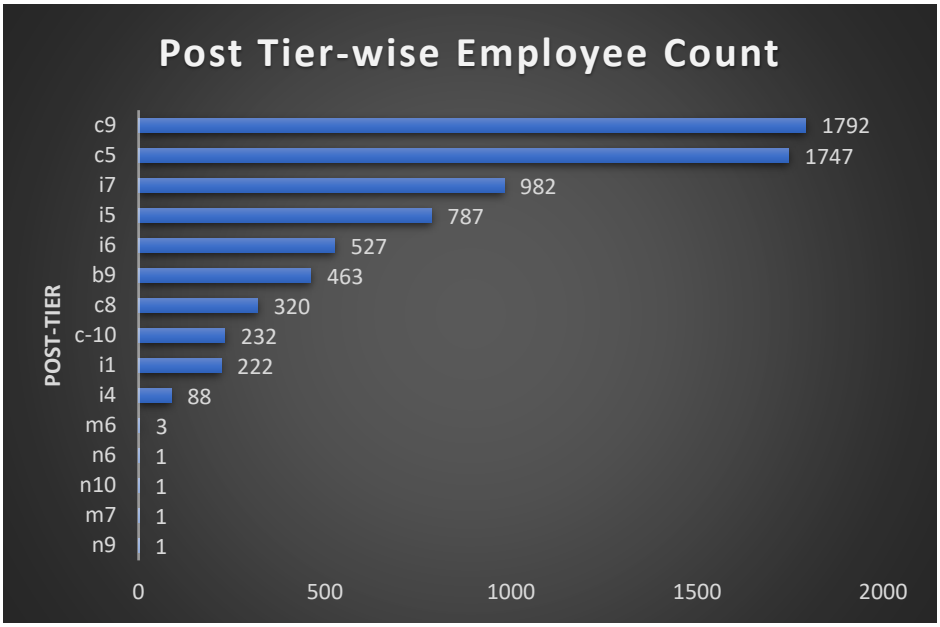




The graphs show that the Operations Department has the greatest number of working employees followed by the Service Department whereas the HR Department has lowest employee count.

5. **Post Tier-wise Employee Distribution:** The different tiers of posts and their respective number of positions offered by the company was as follows:

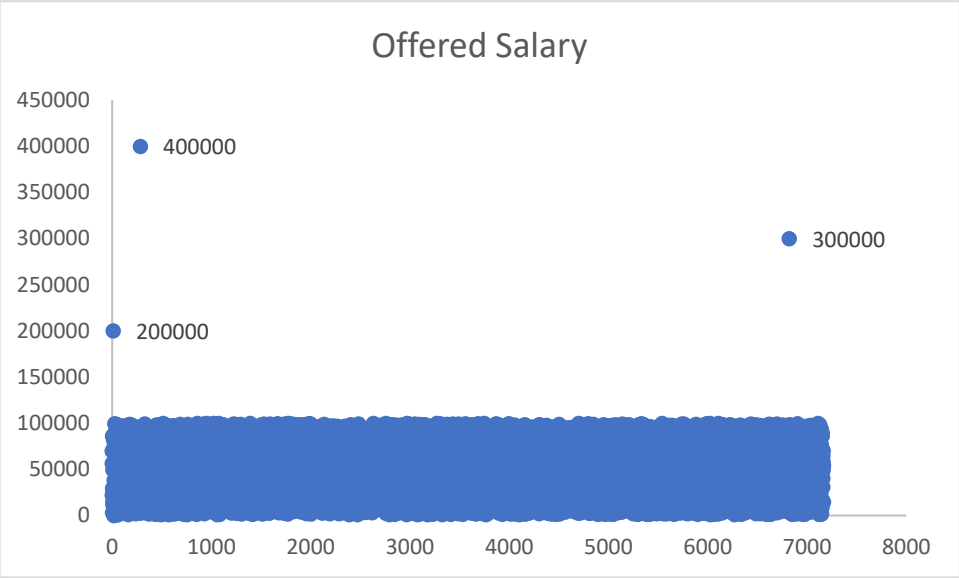
Post_tier	Count of Post Name
n9	1
m7	1
n10	1
n6	1
m6	3
i4	88
i1	222
c-10	232
c8	320
b9	463
i6	527
i5	787
i7	982
c5	1747
c9	1792



Out of the 15 tiers in the company, c9 was found to have the most employee count followed closely by c5, whereas tiers n6, n10, m7 and n9 had only 1 position.

Outliers:

The outliers are extreme values in the dataset that can affect the central tendency. They can be unusually high or low values compared to the overall dataset.
The outliers of the salary offered were detected using a scatter plot:



The graph depicts that most of the salaries lie within 0 -100000 range, however, three values are found to be unusually higher and therefore are considered as outliers.
The average salary with and without outliers were:

Average Salary with outliers	49983.02902
Average without outliers	49878.3318

The average salaries with and without outliers were found to be different suggesting that outliers can effect the central tendency of the dataset and hence needs to identified and dealt with.

Results:

Doing this project helped me understand -

- How useful Excel can be for preparation of data by cleaning and manipulation such as finding null values and removing duplicates, to avoid errors and biases during analysis.
- Using pivot tables helped grouping columns for easy visualization and filtering data.
- The built-in function of Excel provides an easy method for basic to advance statistical analysis of a dataset like finding average, class-intervals etc.
- Making charts and graphs for data visualization and drawing insights from them.
- Scatter plots can be used to detect anomalies in the dataset such as outliers easily.