PROJECT ON HIRING PROCESS ANALYTICS

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.
Here comes the role of a Data Analyst to analyze the company's hiring process data and draw meaningful insights from it.

A dataset containing records of previous hires is provided to the Data Analyst, and their job is to analyze this data and answer specific questions that can help the company improve its hiring process.

This project aims to use our knowledge of statistics and Excel to draw meaningful conclusions about the company's hiring process and get insights that could help the company improve its hiring process and make better hiring decisions.

Here, we will be doing the job of a Data Analyst to understand how the hiring process analytics is done at big Multi National Companies like Google.

Our primary work is to clean data by following steps like:

1. **Handling Missing Data:** Check for any missing values in the dataset. If there are, decide on the best strategy to manage them.
2. **Clubbing Columns:** If columns with multiple categories can be combined, do so to simplify your analysis.
3. **Outlier Detection:** Check for outliers in the dataset that may skew your analysis.
4. **Removing Outliers:** Decide on the best strategy to handle outliers. This could be removing them, replacing them, or leaving them as is, depending on the situation.
5. **Data Summary:** After cleaning and preparing your data, summarize the findings. This could involve calculating averages, medians, or other statistical measures. It could also include creating visualizations to understand the data better.

Data Analytics Tasks:

From the given dataset we will find answers to the questions below using Excel:

A. Hiring Analysis: The hiring process involves bringing new individuals into the organization for various roles.

Task: Determine the gender distribution of hires. How many males and females have been hired by the company?

B. Salary Analysis: The average salary is calculated by adding up the salaries of a group of employees and then dividing the total by the number of employees.

Task: What is the average salary offered by this company? Use Excel functions to calculate this.

C. Salary Distribution: Class intervals represent ranges of values, in this case, salary ranges. The class interval is the difference between the upper and lower limits of a class.

Task: Create class intervals for the salaries in the company. This will help you understand the salary distribution.

D. Departmental Analysis: Visualizing data through charts and plots is a crucial part of data analysis.

Task: Use a pie chart, bar graph, or any other suitable visualization to show the proportion of people working in different departments.

E. Position Tier Analysis: Different positions within a company often have different tiers or levels.

Task: Use a chart or graph to represent the different position tiers within the company. This will help you understand the distribution of positions across different tiers.

My Approach:

- At first, I made a copy of raw data and, converted it into a Table and named it as **Stats.**

  The link to access data and pivot table is https://docs.google.com/spreadsheets/d/1GOnzbdTmtwizb2tSEPyfGwA3jKoE1VEU/edit?usp=drive_link&ouid=117198255086466412144&rtpof=true&sd=true

- Then, I detected outliers in the dataset by calculating the 1st Quartile, 3rd Quartile, Inter-Quartile Range (IQR), Upper bound and Lower bound.
  Where
  IQR=3rd Quartile-1st Quartile
  Upper bound=3rd Quartile+(1.5*IQR)
  And
  Lower bound=1st Quartile-(1.5*IQR).

1st QUARTILE	25460.5
3RD QUARTILE	74438
IQR	48977.5
UPPER BOUND	147904.25
LOWER BOUND	-48005.75

- After that, I deleted the corresponding rows of outliers as shown below.

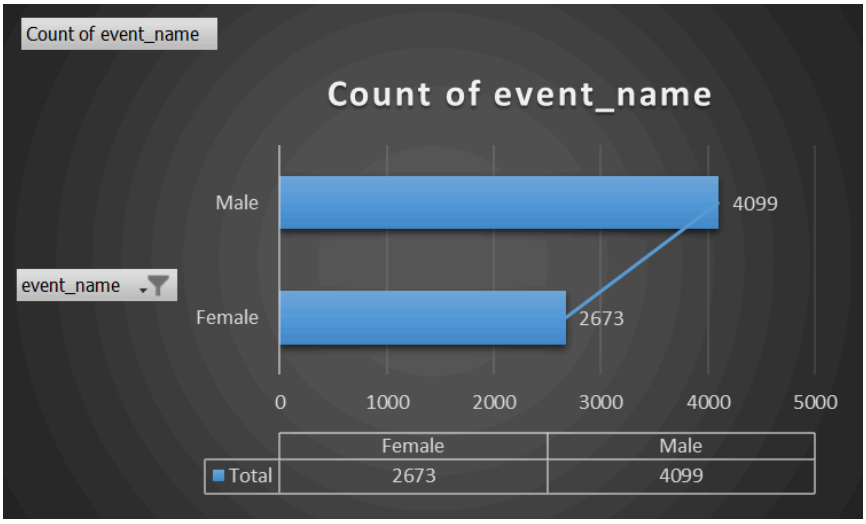
	A	B	C	D	E	F	G
1	application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
13	649039	07-05-2014 10:48	Hired	Female	Service Department	b9	200000
286	795330	15-06-2014 09:45	Hired	Female	General Management	i4	400000
6825	874368	21-07-2014 15:39	Hired	Male	General Management	i7	300000
7170							
7171							

- The next step is to handle the missing values.
- For numeric value columns like Offered Salary we replaced the missing value of Post Name i7 of Sales Department with the median of corresponding Post Name of Sales Department.

- For the categorical variables, we replaced the missing values with the mode of the corresponding columns.
- Thereafter, we started analyzing our cleaned dataset.

Findings

1. Gender Distribution of Hires:



From above, it can be inferred that approximately 65% of hired employees are male in the organization.

2. Salary Analysis:

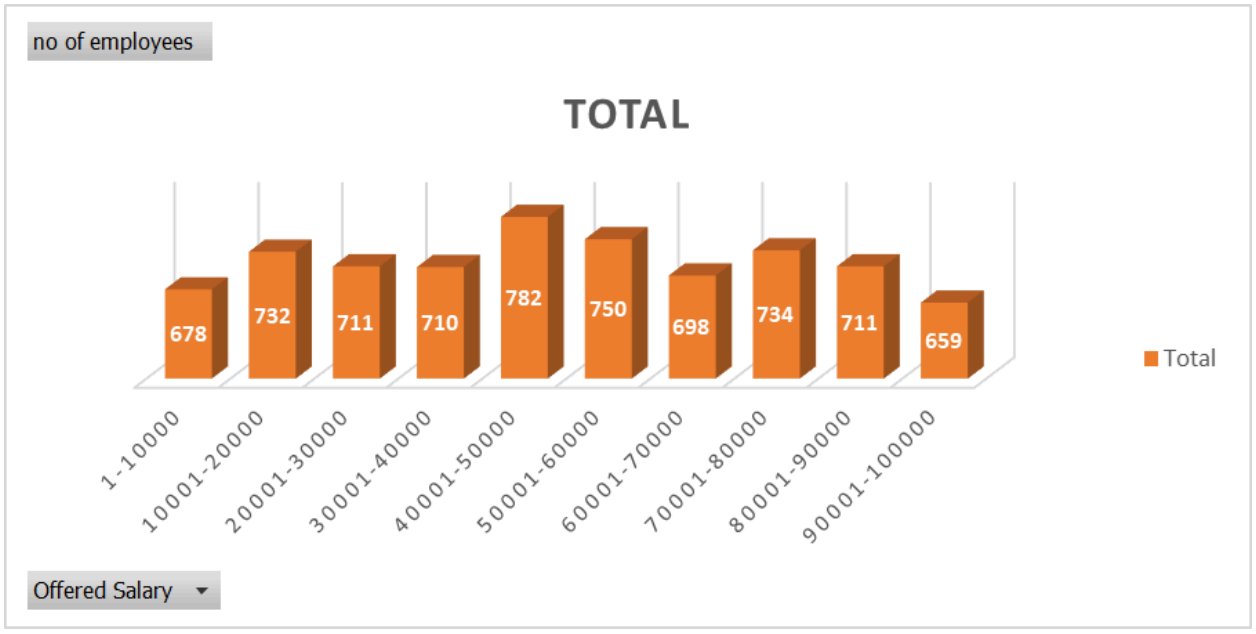
Calculated it using the average function present in Excel.

Average of all employee=AVERAGE(Stats[Offered Salary])

And it resulted in **49878.29**

3. Salary Distribution by Class Interval:

Using the Pivot Table, we created a class interval for the salary offered column with an interval of 10,000 and then represented it by a column chart as below.

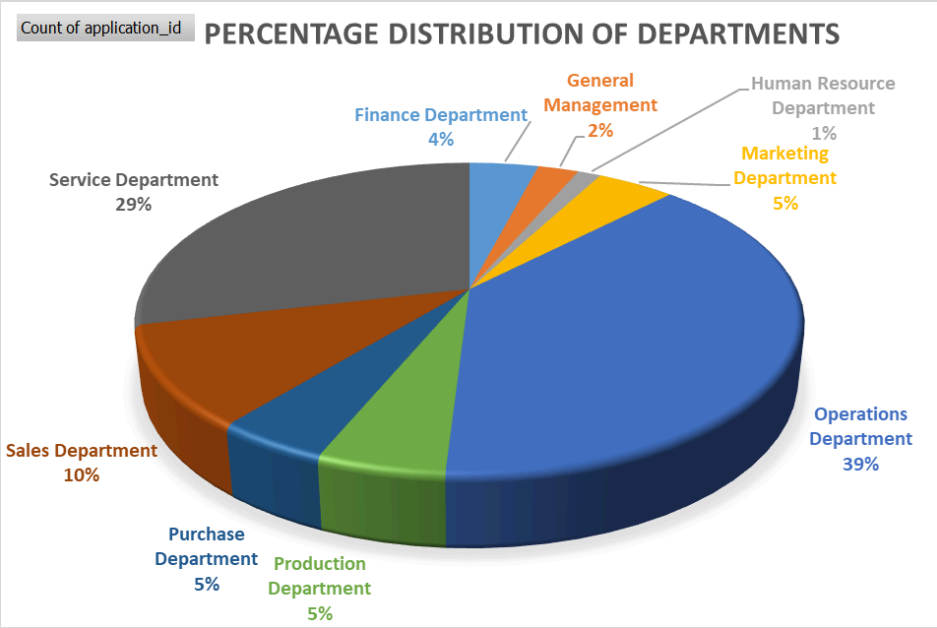


From this chart, we can conclude that 782 employees (Hired+Rejected) get salaries in the range between 40,001-50,000.

4. Departmental Analysis

We found the distribution of employees over the departments as below and represented their percentage distribution in pie-chart.

Department	No of employees
Finance Department	288
General Management	170
Human Resource Department	97
Marketing Department	325
Operations Department	2771
Production Department	380
Purchase Department	333
Sales Department	747
Service Department	2054
Grand Total	**7165**

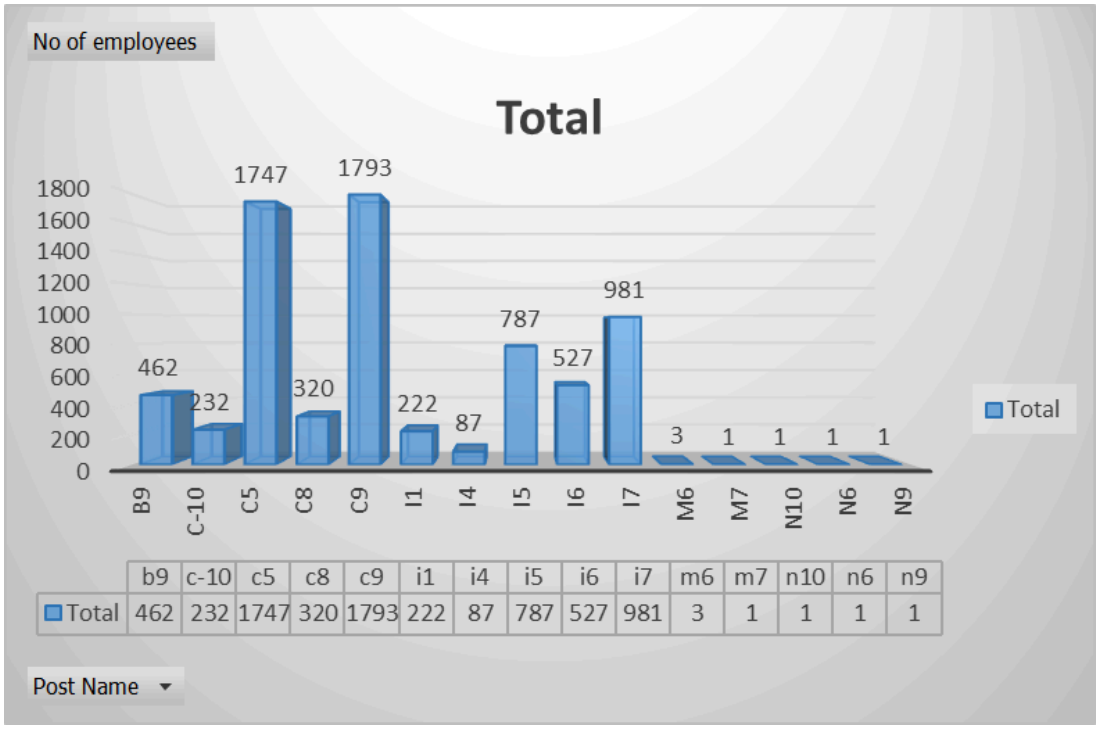


It shows that 39% of employees work in the Operations Department, which is the highest among all, and the least employees work in the Human Resources Department.

5. Position tier Analysis:

It helps us understand how the employees are distributed to different positions in the company.

Post Name	No of employees
b9	462
c-10	232
c5	1747
c8	320
c9	1793
i1	222
i4	87
i5	787
i6	527
i7	981
m6	3
m7	1
n10	1
n6	1
n9	1
Grand Total	**7165**



From the above column chart, it is clear that most of the employees are working for post name c9 which indicates this post has maximum job openings in the company.

CONCLUSION

We can conclude that Hiring Process Analytics is crucial for an organization to understand the hiring trend.
As we can see there is a disparity in the number of male and female employees of the organization and this could be for so many reasons. So, analyzing this and working on it can help the organization have gender equality in the workplace.
Also, it gives a clear understanding of which department and post have more job openings.