Hiring Process Analytics

Authored by: Sushant Karmakar

Index

1. <u>Project Description</u>

Hiring process is the fundamental and the most important function of a company. Here, the MNCs get to know about the major underlying trends about the hiring process. Trends such as- number of rejections, number of interviews, types of jobs, vacancies etc. are important for a company to analyse before hiring freshers or any other individual.

Here, Data analysts are responsible for analyzing data using statistical techniques, implementing and maintaining databases, gathering data from primary and secondary sources, identifying, analyzing and interpreting trends from the data.

Being a Data Analyst, the job is to go through these trends and draw insights out of it for hiring department to work upon.

Here I have been told that I'm working for an MNC such as Google as a lead Data Analyst and the company has provided with the data records of their previous hirings and have asked me to answer certain questions making sense out of that data. Also, the dataset is ROCCC compliant as it has been provided directly the company, hence making it a good data.

2. Problems & Planning

I have been told to provide a detailed report for the below data record mentioning the answers of the questions that follows:

- Hiring: Process of intaking of people into an organization for different kinds of positions.
  Task: How many males and females are Hired?

- Average Salary: Adding all the salaries for a select group of employees and then dividing the sum by the number of employees in the group.
  Task: What is the average salary offered in this company?

- Class Intervals: The class interval is the difference between the upper-class limit and the lower-class limit.
  Task: Draw the class intervals for salary in the company?

- Charts and Plots: This is one of the most important part of analysis to visualize the data.
  Task: Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working different department.

- Charts: Use different charts and graphs to perform the task representing the data.
  Task: Represent different post tiers using chart/graph.

3. Preparations and Approach

I have been provided a dataset of a company where the details about people who registered for a particular post in a department of this company. It is required to use the knowledge in statistics and use different formulas in excel and draw necessary conclusions about the company.

Before diving into the analysis phase, I did the following steps:

- To prevent any data loss, I made a copy of the raw data so that any changes or analysis made to the copied data won't harm the original data.

- I checked the rows for any blank spaces or NULL values.

- Later, I checked for outliers and replaced them with the median of that particular column.

- If there happens to be any blank or NULL values, I replaced them with the value of the highest cell of that column.

- Later, I checked the rows for any duplicates present and then removed them.

- I checked for any unnecessary or irrelevant columns from the dataset that might not be required for the analysis.

- I made some changes in the name of some data for the convineince of data handling. For eg., in Post Name column, I changed c-10 to c10 using find and replace

4. Process & Analysis

The dataset has already been provided by the company; hence it makes easier to carry out the analysis.

Task 1: Calculating the number of males and females hired.

Total Number of Males and Females Hired
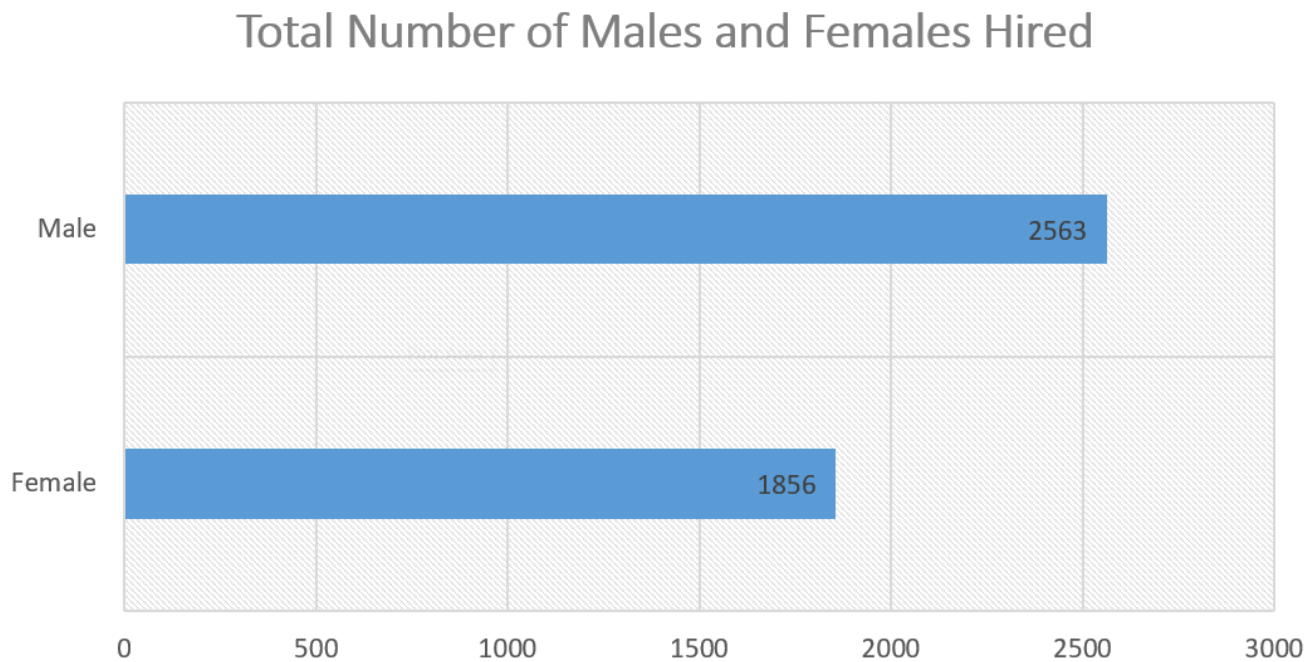
STATUS : HIRED

EVENTS	No. of hiring
MALE	2563
FEMALE	1856

Hence, from the above chart and table it can be concluded that 2563 Males has been hired whereas 1856 Females has been hired.

Task 2: Calculating the average salary offered in this company.

To calculate the average salary, I'll be removing the unnecessary outliers from the "Offered Salary" column and create a pivot table under average values.

Departments	Average of Offered Salary
Finance Department	49628.00694
General Management	58722.09302
Human Resource Department	49002.27835
Marketing Department	48489.93538
Operations Department	49151.35438
Production Department	49448.48421
Purchase Department	52564.77477
Sales Department	49310.3807
Service Department	50629.88418
Grand Total	**49983.02902**

Therefore, the average salary offered 49983.03

Task 3: Create a graphical representation of class interval for salary of the company.

Salary Range	Frequency of Salary
1007-6106	391
6107-11206	371
11207-16306	381
16307-21406	341
21407-26506	393
26507-31606	360
31607-36706	341
36707-41806	403
41807-46906	406
46907-52006	365
52007-57106	392
57107-62206	370
62207-67306	338
67307-72406	366
72407-77506	382
77507-82606	385
82607-87706	355
87707-92806	351
92807-97906	327
97907-103006	125
>103007	22
Grand Total	**7165**

I'll be using the pivot table to create a visualization for the class interval. Then using the grouping process under Pivot Table Analysis, I'll be creating the interval and frequency. Then I'll convert the data into a suitable pivot chart representing Class interval V/s Frequency.

The Chart below depicts the range of the salary clubbed with the number of the applicants that falls under the range. And from the chart it can be concluded that the highest number of applicants is 406 who has been offered salary between 41807-46906.
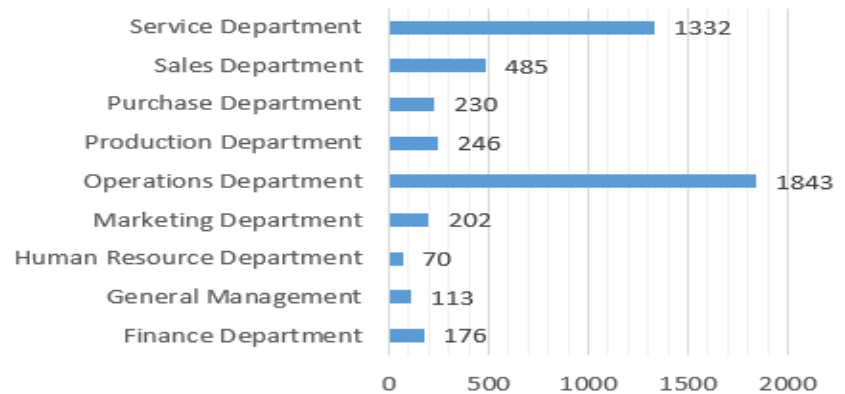
**CLASS INTERVAL OF SALARY IN THE COMPANY**
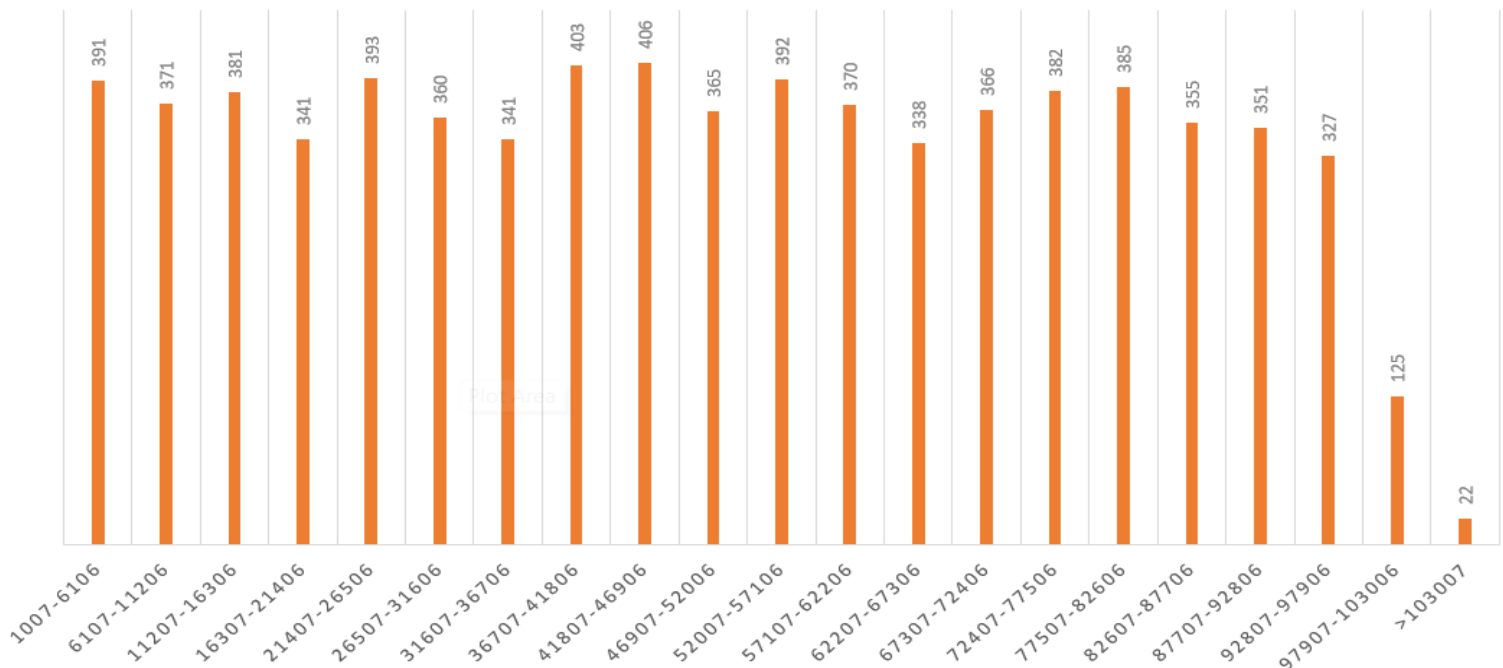
Task 4: Creating a chart/graph to show the proportion of people working in different department.

Departments	Number of People Working
Finance Department	176
General Management	113
Human Resource Department	70
Marketing Department	202
Operations Department	1843
Production Department	246
Purchase Department	230
Sales Department	485
Service Department	1332
Grand Total	**4697**

Total Number of People Working in Different Departments
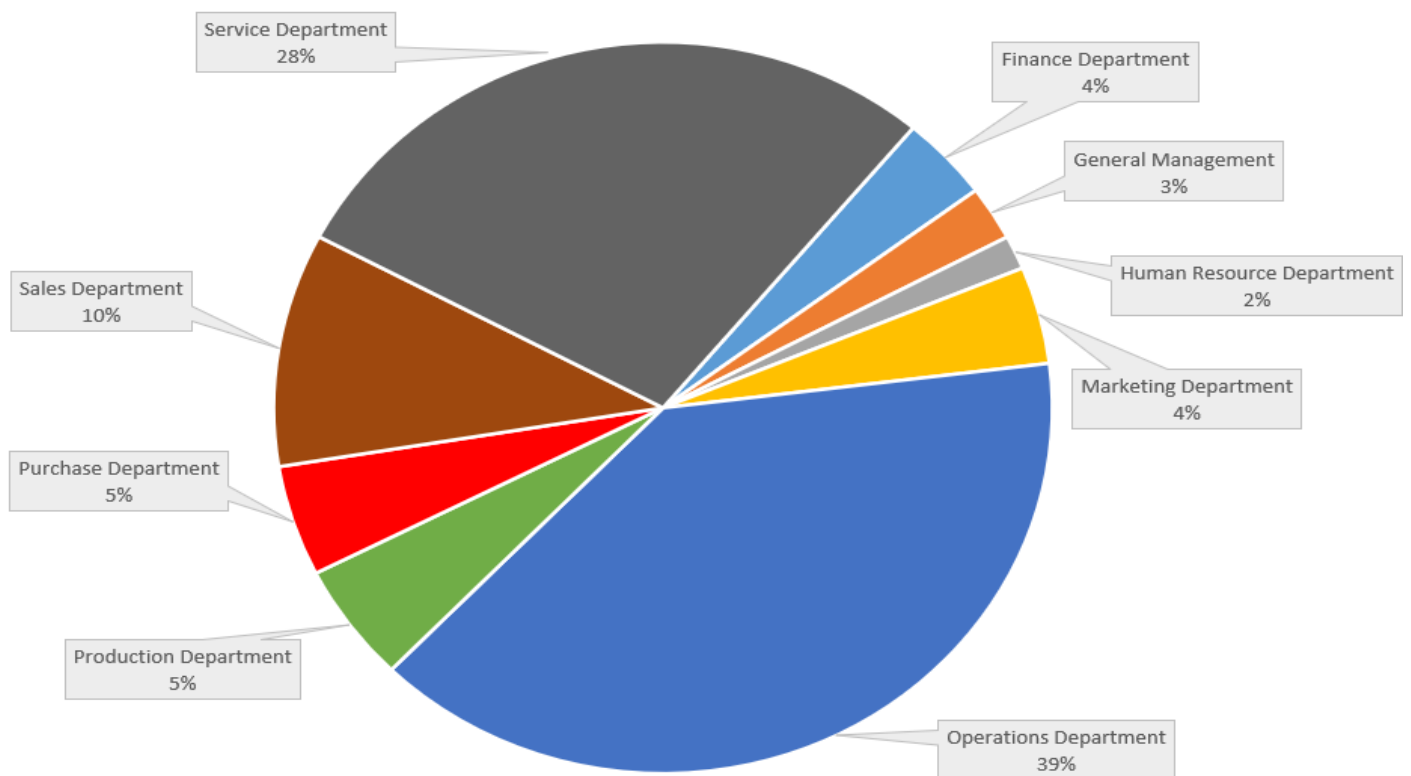
From the above chart and bar plot, it can be concluded that maximum no. of people is working in Operations Department i.e., 1843 with 39% capacity in the company.

Task 5: Representation of different post tiers using chart/graph.

Post Name 🔽	Number of Applicants
b9	463
c10	232
c5	1747
c8	320
c9	1792
i1	223
i4	88
i5	787
i6	527
i7	982
m6	3
m7	1
n10	1
n6	1
n9	1
Grand Total	**7168**

Using the pivot table, I created a pivot bar plot where the post tiers is shown graphically. So, from the plot below, it can be concluded that the c9 post tier has the greatest number of applicants present i.e., 1792.

Post Tiers
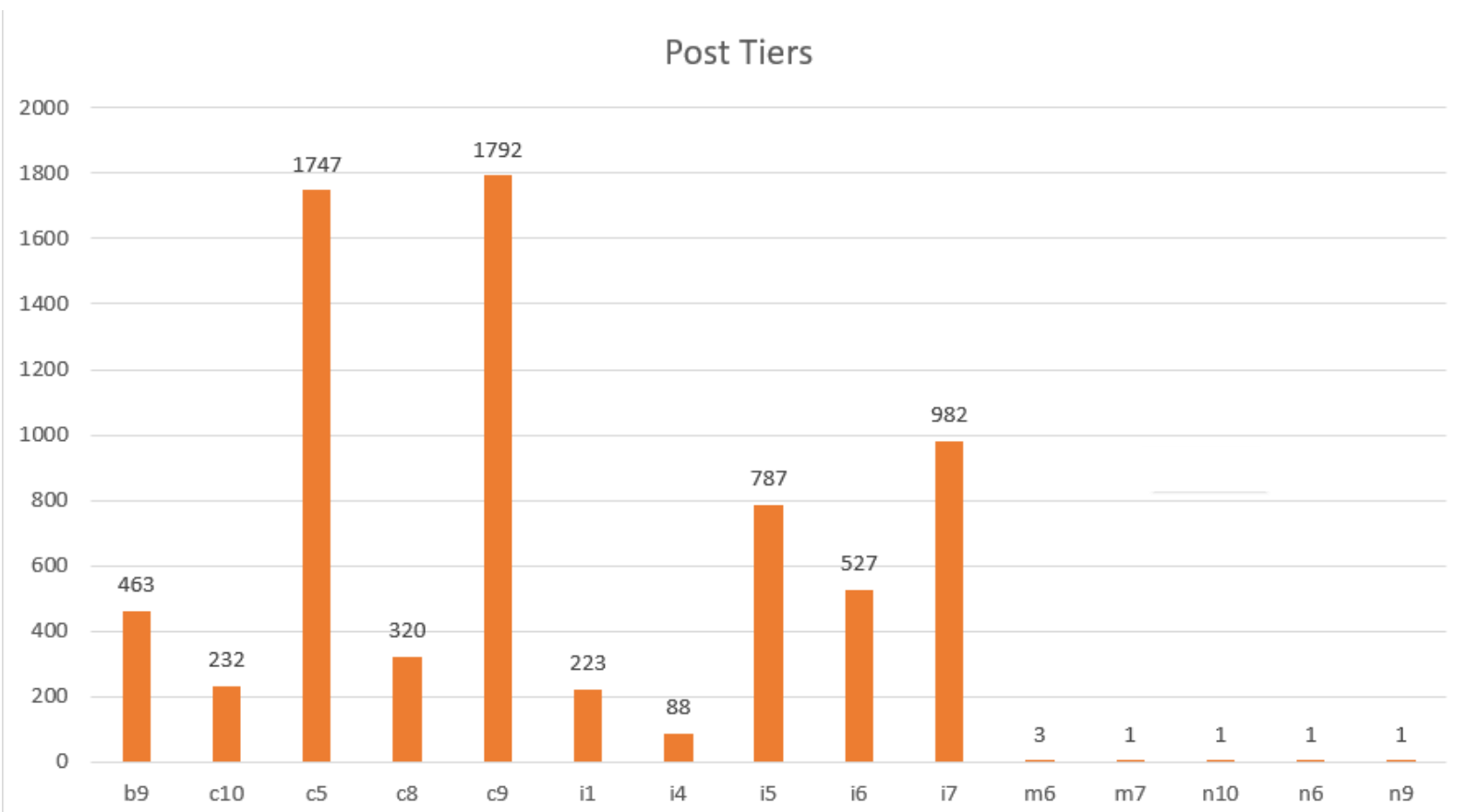
5. Tech Stack Used

For the convenience, I download the dataset from the google sheet and opened it in Microsoft excel 2019.
At Excel 2019, Microsoft has significantly improved many data transformation features. Excel is known for its flexible and powerful analysis experiences, through the familiar PivotTable authoring environment.

Also, Microsoft Word 2019 has been used for making this project.

6. Insights

After completion of this project my experience of handling data on excel increased eventually. I successfully understood the basics of Exploratory Data Analysis known as EDA. This project helped me understand the statistics in a simple yet challenging way. Before, I had no idea how to implement statistical formulas and approach in excel for a dataset but now after completing this project the doubt has been cleared. This project also made me dive into the process of data cleaning, data manipulation and take some important data driven decisions that were necessary for the project. It also gave me a handy experience on pivot table which is the most important asset for data analysis in Excel.

While working with the charts and graphs, I came to know the importance and details of each and every chart and graph. Although it was a bit challenging for me to select the appropriate chart for the report but slowly, I understood the necessity of charts and graphs in data analysis and finally I was able to answer each and every question asked.

The learning experience was excellent. The following websites helped me understand the basics and advanced forms of excel in a simple way possible

- https://trainity.link/data/learning
- https://towardsdatascience.com/an-extensive-guide-to-exploratory-data-analysis-ddd99a03199e
- https://www.excel-easy.com/data-analysis/pivot-tables.html

The time given for this project was enough to learn and work with new and advanced forms of Excel and Spreadsheets simultaneously.

7. Conclusions and Results

The project on hiring process analytics definitely gave a brief idea on the process of hiring and managing candidates on HR level.

This project also gave us an insight on why there are more male candidates working in the company. The gender imbalance in work, where there are more males than females, can be influenced by a variety of factors. It's important to note that gender disparities in the workplace are complex and multifaceted, and there is no one-size-fits-all answer. The issues like occupational segregation, work-life challenges, cultural and organizational norms and lack of representation and role models play a vital role in generation of this imbalance. Addressing the gender imbalance in a company requires systemic changes at multiple levels, including changes in societal norms, organizational policies, and individual attitudes and behaviors.

Salary differences within a company can arise due to a variety of factors, some of which may be intentional or unintentional like job responsibilities, skills and qualifications, performance, seniority and experience and individual bargaining power. It's important to note that salary differences should be based on legitimate and non-discriminatory factors, such as job responsibilities, skills, qualifications, and performance, and should comply with applicable laws and regulations related to pay equity and anti-discrimination. Employers should strive to ensure that compensation practices are fair, transparent, and based on objective criteria to promote equity and fairness in the workplace.

The distribution of employees across different departments in a company can be influenced by a variety of factors and may vary depending on the specific context and business requirements of the company. Proper workforce planning, organizational design, and strategic alignment are crucial in determining the optimal distribution of employees across different departments, including the operations department, to ensure efficient and effective operations within the company.

8. Drive Link

https://drive.google.com/drive/folders/16o-EQhXx3339ZOF6zLno4vRRmOz4vCWM

Name of the project: project_4(hiring process analytics)