Hiring Process Analytics

Project description: The hiring process refers to finding, selecting, and hiring new employees for a company. Simply, it is a procedure of placing employees according to posts available in the company in exchange for money(salary). Here, the HR leader aims to bring the best-skilled employees to the company to meet its business objectives. The hiring process is the fundamental and most important function of a company. Here, the MNCs get to know about the major underlying trends in the hiring process. Trends such as the number of rejections, number of interviews, types of jobs, vacancies, etc. are important for a company to analyze before hiring freshers or any other individual.

Here we will use EDA to find out different insights and to answer the questions asked by the company to answer certain questions making sense out of the data. **Exploratory Data Analysis (EDA)** is an approach to analyzing the data using visual techniques. It is used to discover trends, and patterns, or to check assumptions with the help of statistical summaries and graphical representations.

Steps for EDA

1. Understanding data columns and data
2. Checking for missing data
3. Clubbing columns with multiple categories
4. Checking for outliers
5. Removing outliers
6. Drawing Data Summary

From the given dataset I have answered some of the following questions using statistics and different formulas in MS Excel:

- **Hiring:** How many males and females are Hired?
- **Average Salary:** What is the average salary offered in this company?
- **Class Intervals:** Draw the class intervals for salary in the company?
- **Charts and Plots:** Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working different department?
- **Charts:** Represent different post tiers using chart/graph?

Approach:

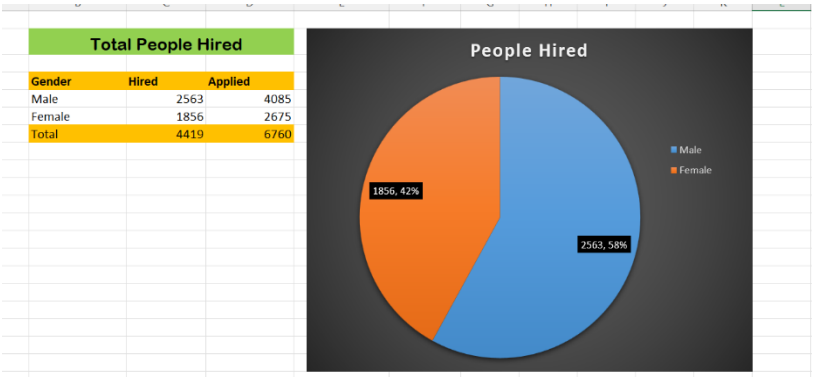
After downloading the provided dataset I used MS Excel, understood the columns and what exactly they mean, and later performed operations using statistics and different formulas to provide meaningful contexts from the dataset required by the company.

Note: The word "data" used in the formulae is the table name that is provided.

We are required to provide a detailed report for the below data record mentioning the answers of the questions that follows:

A. **Hiring:** Process of intaking of people into an organization for different kinds of positions.
Your task: How many males and females are Hired?

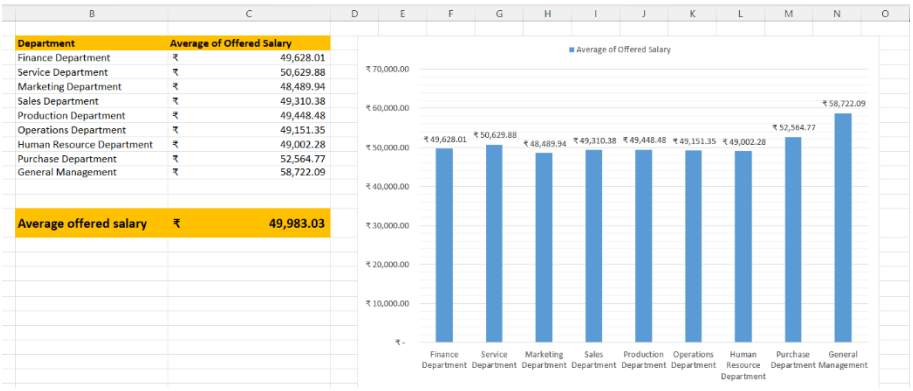
- Number of Males hired = 2563
- Number of Females hired = 1856
- Formula used (Males) = COUNTIFS(data[event_name],"Male",data[Status],"Hired")
- Formula used (Females) = COUNTIFS(data[event_name],"Female",data[Status],"Hired")



Total People Hired		
Gender	Hired	Applied
Male	2563	4085
Female	1856	2675
Total	4419	6760

B. **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.
Your task: What is the average salary offered in this company?

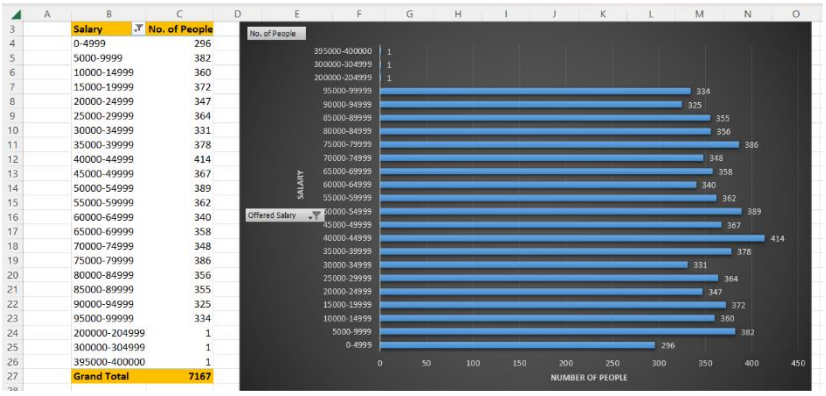
- **Average salary offered in this company = ₹ 49,983.03**
- Formula used for average salary offered = AVERAGE(data[Offered Salary])
- Also, I have calculated the average salary offered in each department where I found the highest average salary is offered in **General Management (₹ 58722.09).**



Department	Average of Offered Salary
Finance Department	₹ 49,628.01
Service Department	₹ 50,629.88
Marketing Department	₹ 48,489.94
Sales Department	₹ 49,310.38
Production Department	₹ 49,448.48
Operations Department	₹ 49,151.35
Human Resource Department	₹ 49,002.28
Purchase Department	₹ 52,564.77
General Management	₹ 58,722.09
Average offered salary	**₹ 49,983.03**

C. **Class Intervals:** The class interval is the difference between the upper class limit and the lower class limit.
Your task: Draw the class intervals for salary in the company?

- For class intervals, I have used a pivot table and taken offered salary in rows and grouped them with a difference of 5000, and taken the count of application id. as values using .
- The pivot table is visually presented using a 2D bar chart.

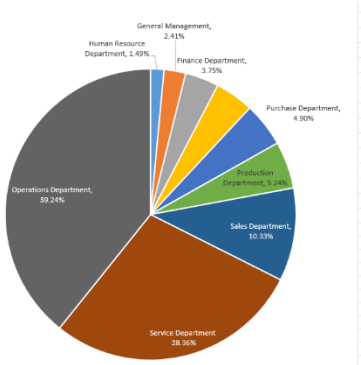


Salary	No. of People
0-4999	296
5000-9999	382
10000-14999	360
15000-19999	372
20000-24999	347
25000-29999	364
30000-34999	331
35000-39999	378
40000-44999	414
45000-49999	367
50000-54999	389
55000-59999	362
60000-64999	340
65000-69999	358
70000-74999	348
75000-79999	386
80000-84999	356
85000-89999	355
90000-94999	325
95000-99999	334
200000-204999	1
300000-304999	1
395000-400000	1
Grand Total	7167

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

- After taking all departments in a column, the number of people working in different departments is calculated using the formula =COUNTIFS(data[Department],A4,data[Status],"Hired") , where A4 is the cell location of a department name.
- And for the proportion of employees working in different departments = **(The number of people working in a department)/4697**, where 4697 is the total number of people hired in the company and then Drag and Drop.
- Then applied the “Percentage” number format which directly multiplies the selected cells by 100 and puts “%” at the end.

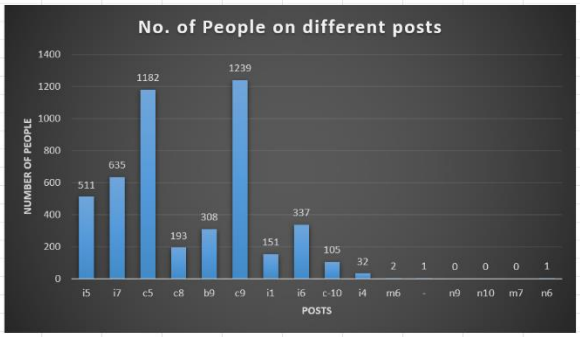
Departments	No. of people working	Percentage of People working
Human Resource Department	70	1.49%
General Management	113	2.41%
Finance Department	176	3.75%
Marketing Department	202	4.30%
Purchase Department	230	4.90%
Production Department	246	5.24%
Sales Department	485	10.33%
Service Department	1332	28.36%
Operations Department	1843	39.24%
Total	**4697**	**100.00%**



E. **Charts:** Use different charts and graphs to perform the task representing the data.
Your task: Represent different post tiers using chart/graph?

- Here, I have displayed different posts' names and the number of people on the respected post, this can be also done by using a PIVOT table in a more efficient way.
- The below image contains a graph depicting different posts along with the total number of people concerning their posts.
- Formula used (Posts Name) = UNIQUE(data[Post Name])
- Formula used (No. of People on different posts) = COUNTIFS(data[Post Name],B3,data[Status],"Hired"), where B3 is the cell location of post name i.e. i5.
- Then used Drag and Drop to apply the same formula for calculating the number of people working on each post.

Posts Name	No. of People on different posts
i5	511
i7	635
c5	1182
c8	193
b9	308
c9	1239
i1	151
i6	337
c-10	105
i4	32
m6	2
-	1
n9	0
n10	0
m7	0
n6	1
Total	**4697**



Almost most of the above questions are solved using the COUNTIFS formula meanwhile they can be answered using the PIVOT table technique which is used to summarize large amounts of data in a more effective way.

Tech-Stack used:

1. **MS Excel:** For answering all the questions I used MS Excel with statistics and different formulas for calculating the output and represented the senseful data in a graphical manner for better understanding.

2. **MS Word:** It is used for making the report to present it to the team leader.

Insight:
For solving the asked questions I have used the EDA (Exploratory Data Analysis) approach to represent the summaries of data in a graphical manner. After resolving the asked questions I discovered:

1. Males working in the company i.e. 2563 are more as compared to Females i.e. 1856 in numbers meanwhile rejection rate of males is also 30% higher than females.
2. The average salary given in the company is Rs. 49,983 (approx. 50k) whereas the General Management department has the highest average salary in the company Rs. 58,722.
3. The maximum number of people (414 people) receive salaries ranging between 40000-49999 whereas only 3 people may be managers or someone in high posts receive salaries between 2Lakhs and 4Lakhs.
4. 39.24% of employees of the company are working in the Operations department and the Human Resource department has the least employees.
5. Maximum employees are hired on post c9 following post c5 and the company has shown interest in hiring applicants on both posts and none of the employees are hired on the n9, n10, and m7 posts.
6. Operation department has the highest rejection rate at 39% and the HR department with 1% of rejection rate.

Result: In this project, I got to know how to use statistics to discover various outputs and understood the usage of charts such as 2D columns, 2D bars, Pie charts, etc. to display the summaries of data in a visual way. One of the technique PIVOT table provided by MS Excel for querying large amounts of data and subtotalling and aggregating numeric data, summarizing data by categories and subcategories, and creating custom calculations and formulas. Easily performed the hiring process and analyzed it successfully.

Excel sheet link

https://docs.google.com/spreadsheets/d/14WDssSxesYLncGLHIPX297YYS7MpmEr_/edit?usp=share_link&ouid=102746205635544467842&rtpof=true&sd=true

Please do open it in MS Excel