

[illegible]

## ***Project Description***

The hiring process involves bringing new individuals into the organization for various roles.

This project focuses on certain topics of the hiring process such as gender distribution of the employees, salary analysis, departmental analysis and position tier analysis. Upon working on the given dataset, we can understand which posts are most demanding in the company, what are the average and the maximum salaries offered, different tiers or levels offered within different positions and lastly we can use visualization tools to understand the proportion of people working in different departments.

## ***Approach***

The dataset had the following columns - application\_id, interview\_taken\_on, status (hired/rejected), event\_name (gender), department, post name and offered\_salary.

In the Microsoft Excel, data cleaning was conducted on the downloaded dataset. I used 'find and replace' to search for all "-" (hyphens) in the event\_name column and replaced them with "Don't want to say". Secondly the "offered salary" column was sorted in the ascending order. The majority of the salary lies between the range 800 and 1,00,000. Hence, the outliers were found out using the filter function - 2,00,000, 3,00,000 and 4,00,000. Thereafter, the purpose of each of the questions was understood, and the appropriate Excel functions, formulas and visualization tools were used to conduct the analysis.

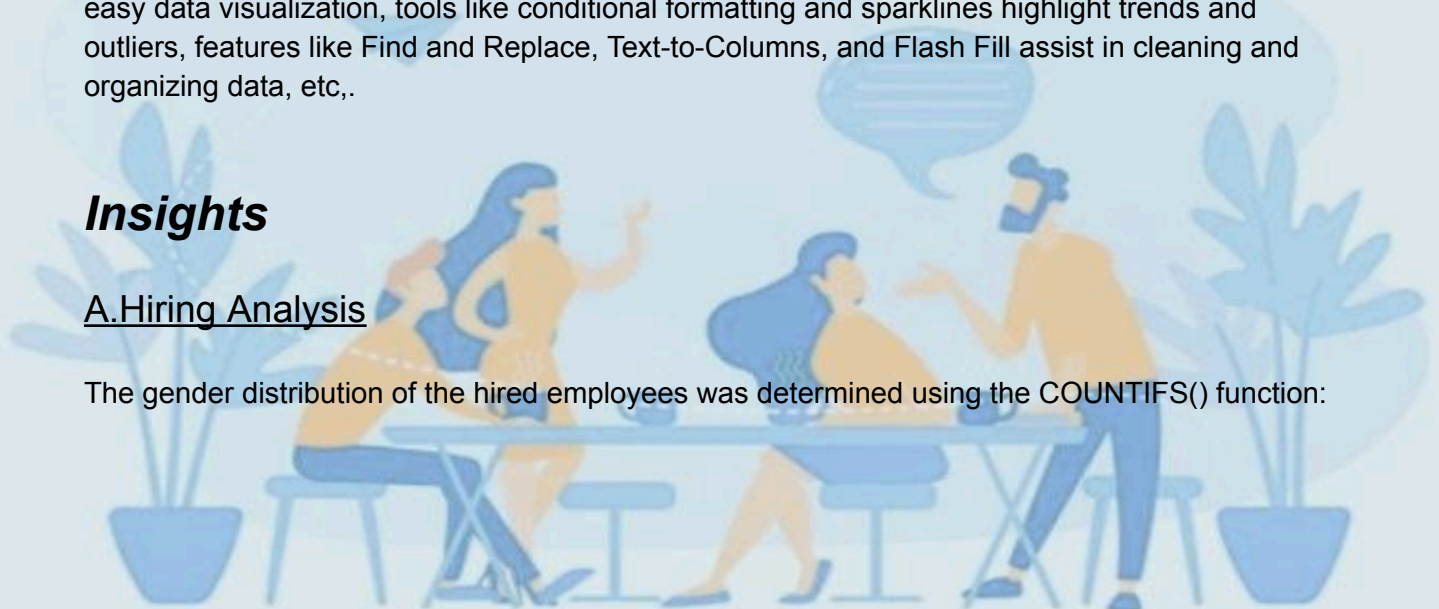
## ***Tech-Stack Used***

The software used is Microsoft Excel because of some of the advantages it provides like - availability of wide range of functions, it offers various chart types (line, bar, scatter, pie, etc.) for easy data visualization, tools like conditional formatting and sparklines highlight trends and outliers, features like Find and Replace, Text-to-Columns, and Flash Fill assist in cleaning and organizing data, etc.,.

## ***Insights***

### **A.Hiring Analysis**

The gender distribution of the hired employees was determined using the COUNTIFS() function:





For males,

```
=COUNTIFS( D:D , D4 , C:C , C3 )
```

For females :

```
=COUNTIFS( D:D , D3 , C:C , C2 )
```

The resultant is shown below.

| Hiring Analysis       |                         |
|-----------------------|-------------------------|
| Number of males hired | Number of females hired |
| 1522                  | 819                     |

A higher number of males are hired in 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. The Excel function used is as follows -

```
=AVERAGE( G2:G7169 )
```

The average salary offered is 49,983. There were outliers - 2,00,000; 3,00,000 and 4,00,000 which may have impacted the result to some extent.

| Salary Analysis        |             |
|------------------------|-------------|
| Average Salary Offered | 49983.02902 |

## 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.

The outliers in the offered salary are -



| Outliers in Offered Salary |
|----------------------------|
| 200000                     |
| 300000                     |
| 400000                     |

These were found using the FILTER() formula-

```
=FILTER( G2:G7169 , G2:G7169 >150000,"No records")
```

Before calculating the class intervals for the salaries, the following were calculated -

|                   |       |
|-------------------|-------|
| Min salary        | 100   |
| Max Salary        | 99967 |
| Range             | 99867 |
| Number of classes | 8.19  |
| Class Width       | 12194 |

- Minimum salary is calculated using the min() function.

```
=MIN( G2:G7169 )
```

- Maximum salary is calculated using the maxifs() function (outliers have been ignored).

```
=MAXIFS( G2:G7169 , G2:G7169 , "<150000")
```

- Range was calculated by subtracting the minimum salary from the maximum salary.
- Number of classes is calculated using a statistical formula-

```
=ROUND(1+3.332+LOG10(COUNT( G2:G7169 )), 2)
```

- Class width is calculated as follows-

```
=ROUND( J38 / J39 , 0)
```

Then for the class intervals, for the salaries in the company, lower and upper limits were calculated respectively as follows-

Lower limit.

= J44 +1

The cell was selected and its border was dragged down just to one cell to copy the formula in there. Then Upper Limit for the first class interval was calculated -

= \$J\$40 + J44

Similarly the cell was selected and its border was dragged down just to one cell to copy the formula there. Then the second cell from the lower limit column was selected and dragged down to one cell to copy the formula below, and so on. The frequency of each class interval is calculated using COUNTIFS() function -

=COUNTIFS( \$G\$2:\$G\$7169 , ">=" & I45 , \$G\$2:\$G\$7169 , "<=" & J45 )

The resultant is :

| Lower Limit | Upper Limit | Frequency |
|-------------|-------------|-----------|
| 1           | 12194       | 832       |
| 12195       | 24388       | 880       |
| 24389       | 36582       | 861       |
| 36583       | 48776       | 956       |
| 48777       | 60970       | 905       |
| 60971       | 73164       | 851       |
| 73165       | 85358       | 891       |
| 85359       | 97552       | 824       |
| 97553       | 109746      | 164       |
| Total =     |             | 7164      |

## D. Departmental Analysis

Visualizing the data through charts and plots is a crucial part of data analysis.

To show the proportion of people working in different departments, the unique departments were filtered with the help of UNIQUE() function in Excel.

=UNIQUE( E2:E7169 ,,0)

Then the number of hired employees from each department was calculated -

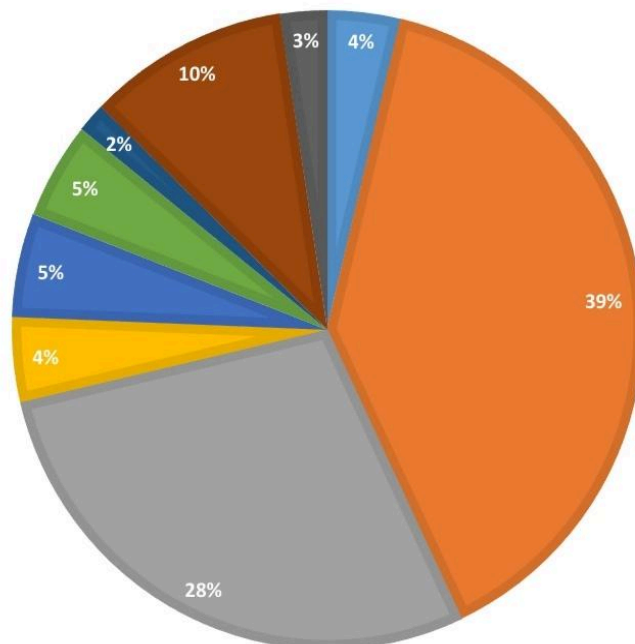
=COUNTIFS( \$E\$2:\$E\$7169 , I10 , \$C\$2:\$C\$7169 , \$J\$9 )

The resultant table so obtained is -

| DEPARTMENTS               | Hired |
|---------------------------|-------|
| Finance Department        | 176   |
| Operations Department     | 1843  |
| Service Department        | 1332  |
| Marketing Department      | 202   |
| Production Department     | 246   |
| Purchase Department       | 230   |
| Human Resource Department | 70    |
| Sales Department          | 485   |
| General Management        | 113   |
| Total                     | 4697  |

### PROPORTION OF PEOPLE WORKING IN DIFFERENT DEPARTMENTS

■ Finance Department      ■ Operations Department      ■ Service Department  
 ■ Marketing Department      ■ Production Department      ■ Purchase Department  
 ■ Human Resource Department      ■ Sales Department      ■ General Management



It can be seen that the Operations Department has the maximum employees followed by the Service Department. The remaining departments have far lesser employees.

## E. Position Tier Analysis

Different positions within a company often have different tiers or levels.

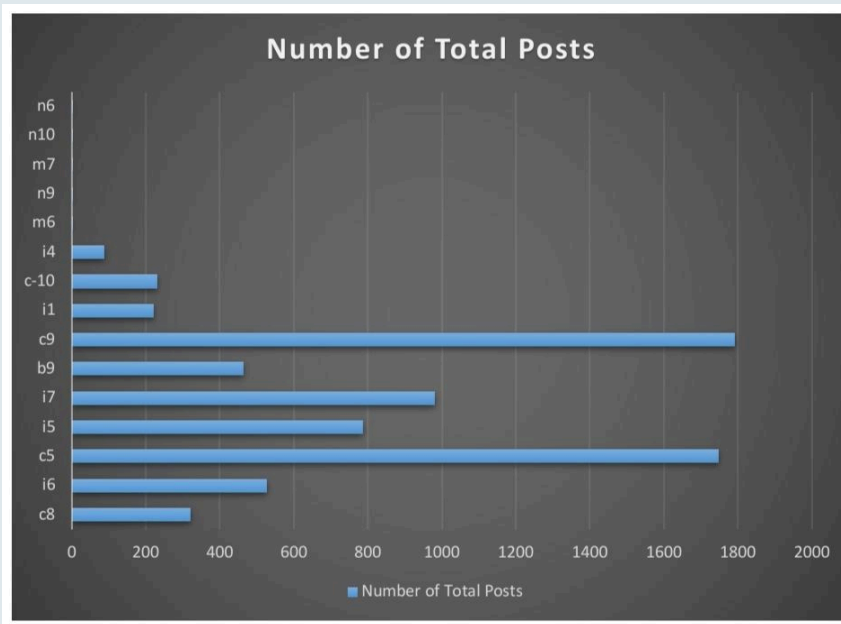
Unique post names were obtained and then their corresponding counts were calculated.

```
=UNIQUE( F2:F7169 ,0)
```

```
=COUNTIF( $F$2:$F$7169 , I60 )
```

Representing the different position tiers within the company to understand the distribution of positions a bar graph was plotted.

| <b><u>Position Tier Analysis</u></b> |                       |
|--------------------------------------|-----------------------|
| Post Name                            | Number of Total Posts |
| c8                                   | 320                   |
| i6                                   | 527                   |
| c5                                   | 1747                  |
| i5                                   | 787                   |
| i7                                   | 982                   |
| b9                                   | 463                   |
| c9                                   | 1792                  |
| i1                                   | 222                   |
| c-10                                 | 232                   |
| i4                                   | 88                    |
| m6                                   | 3                     |
| n9                                   | 1                     |
| m7                                   | 1                     |
| n10                                  | 1                     |
| n6                                   | 1                     |



Maximum employees are hired for the c9 post among various departments.

## Result

Working on the project, insights were obtained and a better understanding of the proportional distribution of departments along with the fact that the company hires more males over females, maximum of the employees are being hired in the operations department and in c9 post followed by c5 post among various departments, maximum employees are being paid salary in the bracket of 36,500 to 48,800 and lastly that some demanding posts were offered very high salaries like 2,00,000; 3,00,000 and 4,00,000.

## Drive Link

[Click here](#) to visit the Excel sheet.

By - Komal Tayade