JULY 2024

HIRING PROCESS ANALYTICS

STATISTICS

Gunisha Chopra
DATA ANALYST

PROJECT DESCRIPTION

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.

For this project, I will be the data analyst at a multinational company like Google. My task is to analyse the company's hiring process data and draw meaningful insights from it. I am given a dataset containing records of previous hires. My job is to analyse this data and answer certain questions that can help the company improve its hiring process.

Such questions might include:

A. **Hiring**: How many males and females are Hired?

B. **Average Salary**: What is the average salary offered in this company?

This analysis also includes:

Class Intervals: Drawing the class intervals for salary in the company.

Charts and Plots: Drawing Pie Chart / Bar Graph (or any other graph) to show the proportion of people working in different departments.

Charts: Representing different post tiers using chart/graph.

APPROACH

Firstly, I would understand data columns and data. I will club columns with multiple categories to make a deep analysis of trends and extract information from those. Then, I will draw a data summary.

The data set consists of the following columns:

	A	B	C	D	E	F	G
1	application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
2	383422	5/1/2014 11:40	Hired	Male	Service Department	c8	56553
3	907518	5/6/2014 8:08	Hired	Female	Service Department	c5	22075
4	176719	5/6/2014 8:08	Rejected	Male	Service Department	c5	70069
5	429799	5/2/2014 16:28	Rejected	Female	Operations Department	i4	3207
6	253651	5/2/2014 16:32	Hired	Male	Operations Department	i4	29668
7	289907	5/1/2014 7:44	Hired	Male	Sales Department	-	85914
8	959124	5/6/2014 16:27	Rejected	Male	Sales Department	i7	69904
9	86642	5/9/2014 13:17	Rejected	Male	Sales Department	i7	11758
10	751029	5/2/2014 13:09	Hired	Female	Service Department	i4	15156
11	434547	5/2/2014 13:11	Rejected	Female	Service Department	i4	49515
12	518854	5/1/2014 9:00	Rejected	Male	Service Department	n10	26990
13	649039	5/7/2014 10:48	Hired	Female	Service Department	b9	200000
14	199526	5/7/2014 10:50	Hired	Male	Service Department	b9	86787
15	539803	5/15/2014 9:31	Hired	Male	Finance Department	b9	2308
16	191009	5/9/2014 12:48	Hired	Female	Service Department	i7	56688
17	195323	5/9/2014 12:48	Hired	-	Service Department	i7	81757
18	51318	5/2/2014 8:07	Hired	Male	Service Department	i5	15134
19	742283	5/2/2014 8:11	Rejected	-	Service Department	i5	100
20	513166	5/1/2014 22:53	Hired	Female	Operations Department	i1	73579
21	791372	5/1/2014 22:54	Rejected	Male	Operations Department	i1	50351
22	47857	5/1/2014 22:55	Rejected	Female	Operations Department	i1	38462
23	834101	5/1/2014 22:53	Rejected	Don't want to say	Operations Department	i1	82510
24	985008	5/1/2014 9:41	Rejected	Male	Service Department	i6	52554

- **Application_id**: ID of applicants who appeared in the interview.
- **Interview taken on**: Date and time interview was taken.
- **event_name**: Gender of person who appeared for the interview.
- **Status**: whether a person is hired or rejected.

- **Department**: Name of the departments.
- **Post Name**: post he/she is offered.
- **Offered salary**: Salary offered to the applicant.

TECH-STACK USED:

Microsoft EXCEL

TASKS

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?

GENDER DISTRIBUTION:

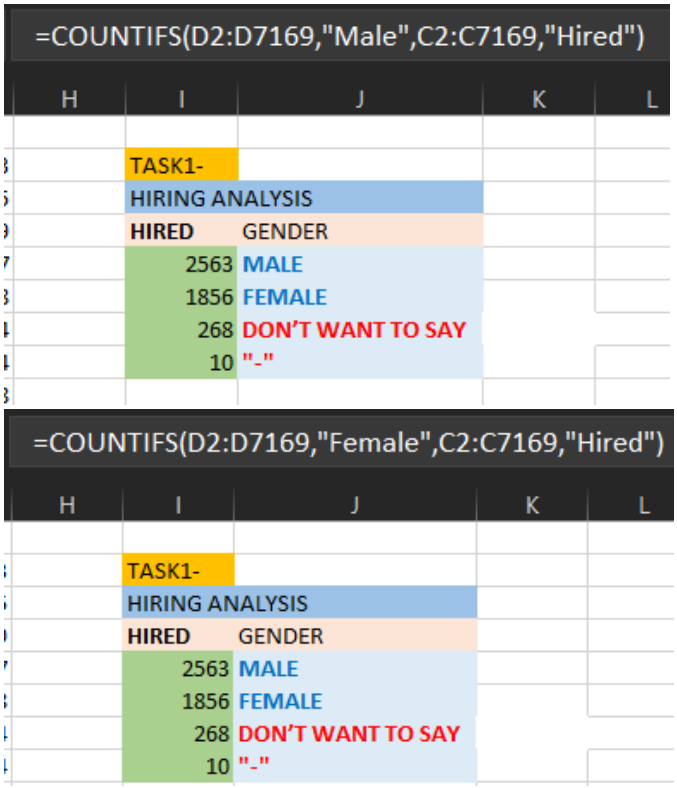
- Male
- Female
- Don’t want to say
- “ _ “

Here, the ‘**COUNTIFS’** function is used.

How many Males and Females are hired?

Formula used →

- =COUNTIFS(D2:D7169,"Male",C2:C7169,"Hired")
- =COUNTIFS(D2:D7169,"Female",C2:C7169,"Hired")



Basically here we got an analysis that the Total number of males hired is 2563 and the total number of females hired is 1856.

INSIGHT: We have more male workers than women.

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.

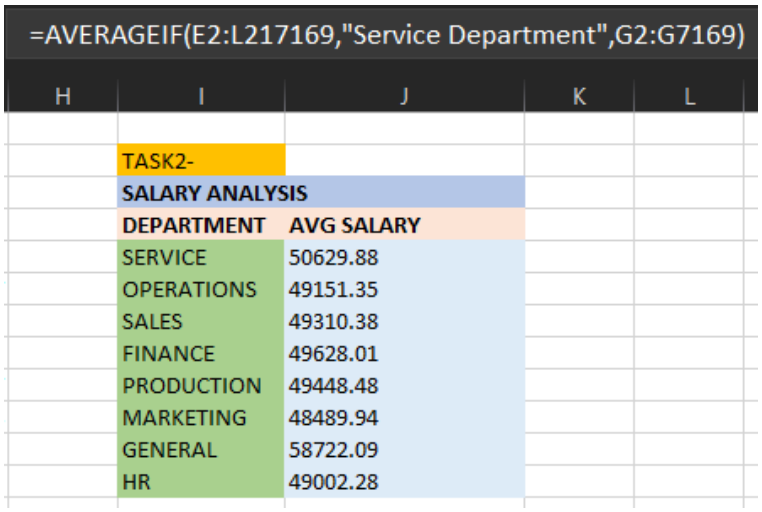
#Task: What is the average salary offered by this company? Use Excel functions to calculate this.

Here, the "**AVERAGEIF"** function is used.

Salaries offered to different departments:

formulas used →

- =AVERAGEIF(E2:E7169,"Service Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Operations Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Sales Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Finance Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Production Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Purchase Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"Marketing Department",G2:G7169)
- =AVERAGEIF(E2:E7169,"General Management",G2:G7169)
- =AVERAGEIF(E2:E7169,"Human Resource Department",G2:G7169)



	I	J	K	L
	TASK2-			
	SALARY ANALYSIS			
	DEPARTMENT	AVG SALARY		
	SERVICE	50629.88		
	OPERATIONS	49151.35		
	SALES	49310.38		
	FINANCE	49628.01		
	PRODUCTION	49448.48		
	MARKETING	48489.94		
	GENERAL	58722.09		
	HR	49002.28		

#INSIGHT: We see that the average salary of the General management department is the highest i.e. 58722.09 and the average salary of the HR department is the lowest i.e. 49002.28.

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.

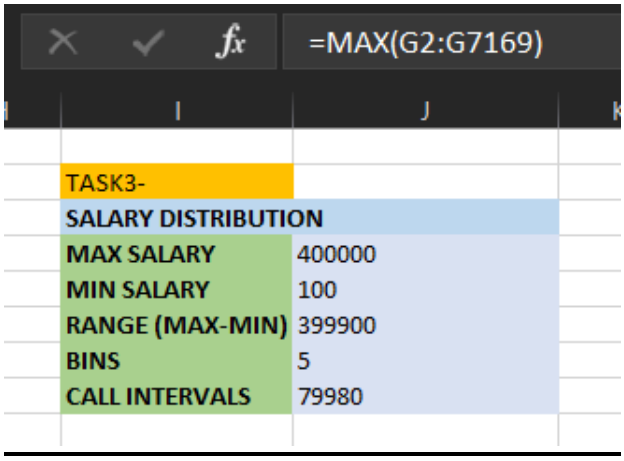
Task: Create class intervals for the salaries in the company. This will help you understand the salary distribution.

First, we find the **Maximum** and **Minimum** values from the offered salaries using the following formula:

- =MAX(G2:G7169)
- =MIN(G2:G7169)

Then, we choose the number of bins we want i.e. 5. After that, we calculated the call intervals using the formula:

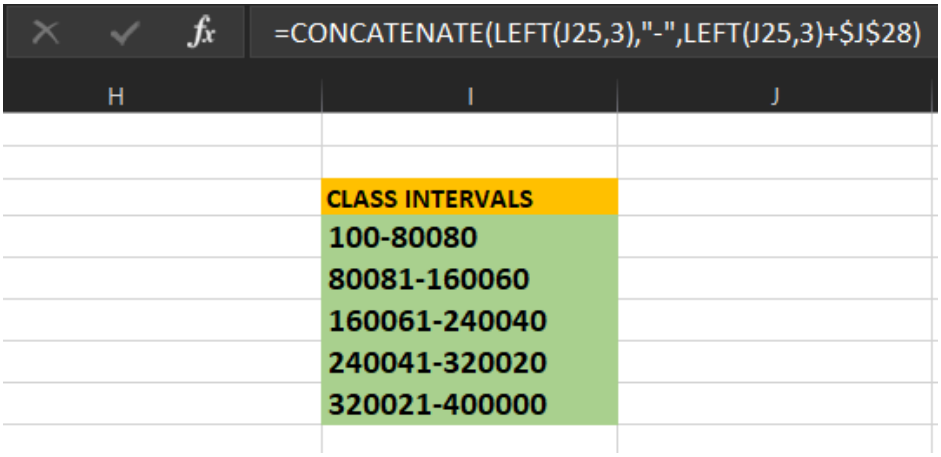
- =(Range/Bins)



Further, we create class intervals.

Formulas used→

- =CONCATENATE(LEFT(J25,3),"-",LEFT(J25,3)+$J$28)
- =CONCATENATE(RIGHT(I32,5)+1,"-",RIGHT(I32,5)+$J$28)
- =CONCATENATE(RIGHT(I33,6)+1,"-",RIGHT(I33,6)+$J$28)
- =CONCATENATE(RIGHT(I34,6)+1,"-",RIGHT(I34,6)+$J$28)
- =CONCATENATE(RIGHT(I35,6)+1,"-",RIGHT(I35,6)+$J$28)



D.) DEPARTMENTAL ANALYSIS

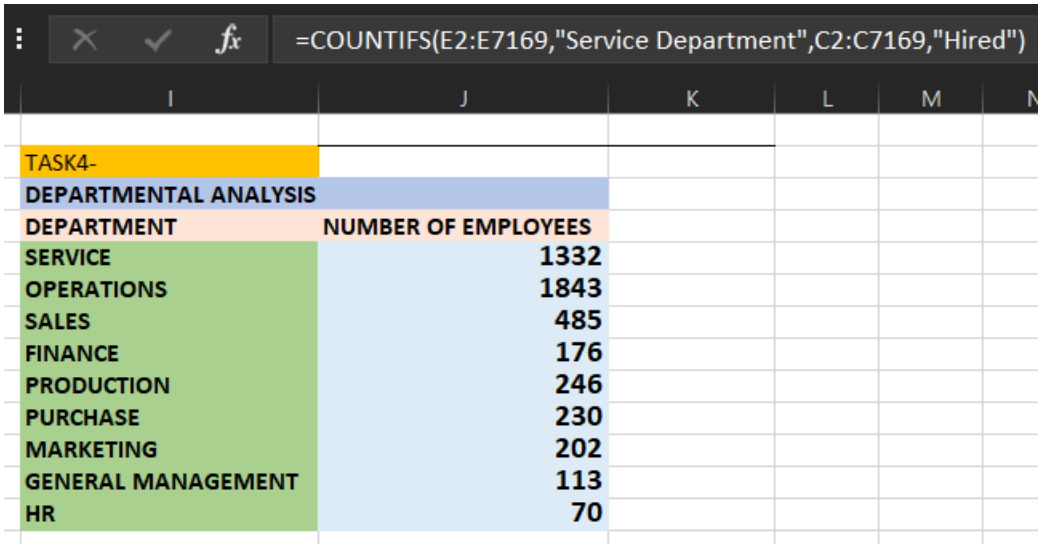
Visualizing data through charts and plots is a crucial part of data analysis.

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

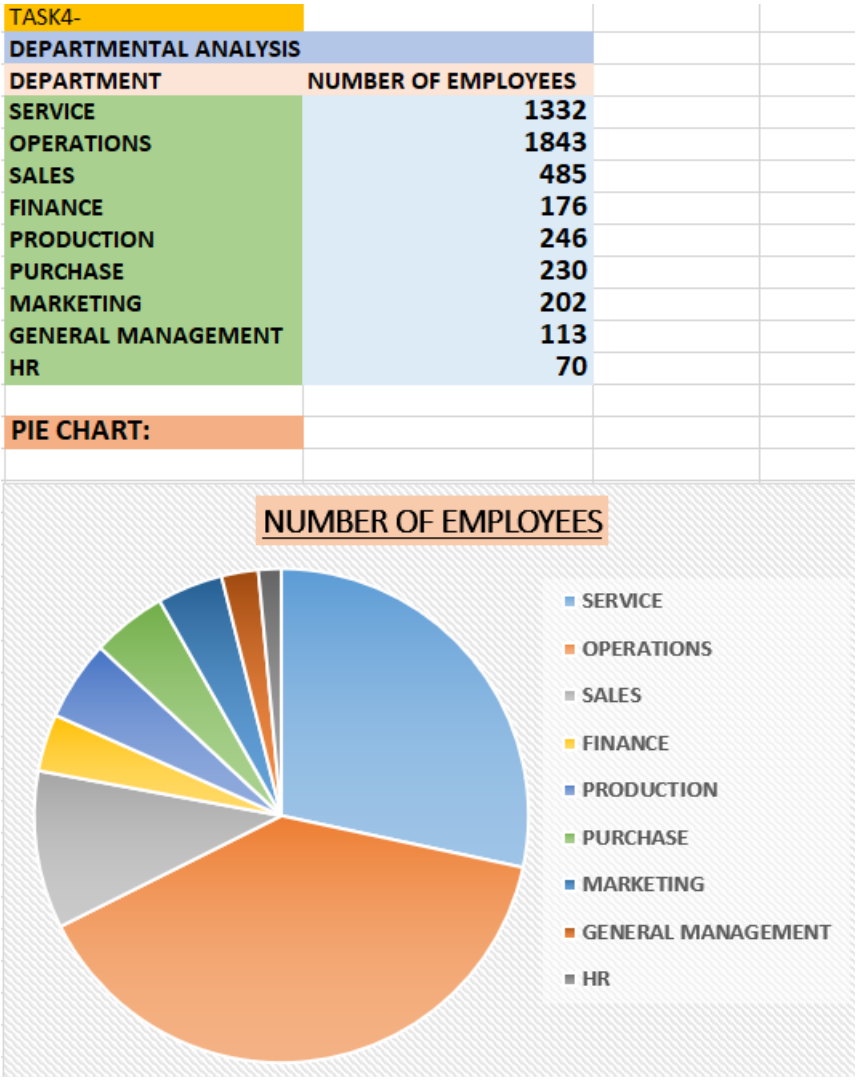
First, we find out the number of employees in each department.

Formulas used:

- =COUNTIFS(E2:E7169,”Service Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” Operations Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” Sales Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” Finance Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” Production Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,”purchase Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” marketing Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,” General Department”,C2:C7169,”Hired”)
- =COUNTIFS(E2:E7169,”Human Resource Department”,C2:C7169,”Hired”)


=COUNTIFS(E2:E7169,"Service Department",C2:C7169,"Hired")
I
J
K
L
M
TASK4-
DEPARTMENTAL ANALYSIS
DEPARTMENT
NUMBER OF EMPLOYEES
SERVICE 1332
OPERATIONS 1843
SALES 485
FINANCE 176
PRODUCTION 246
PURCHASE 230
MARKETING 202
GENERAL MANAGEMENT 113
HR 70

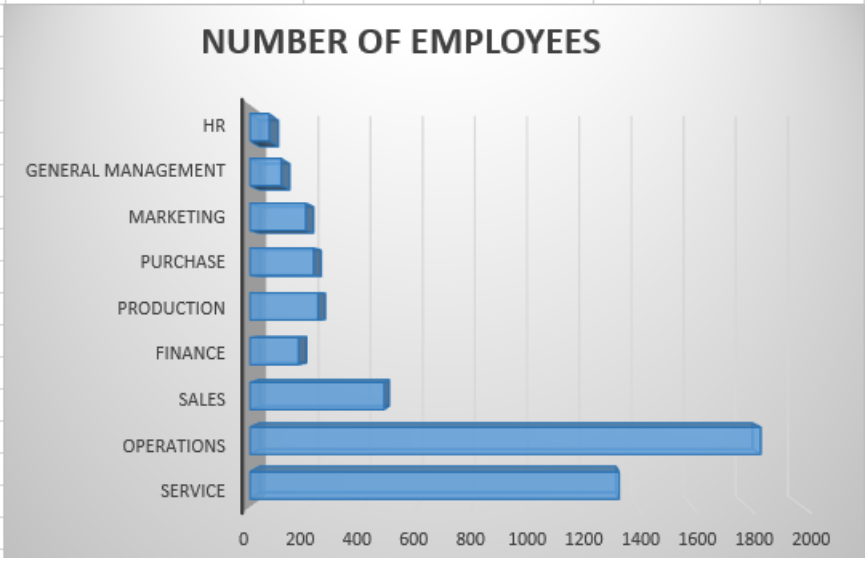
- **PIE CHART:**


TASK4-
DEPARTMENTAL ANALYSIS
DEPARTMENT
NUMBER OF EMPLOYEES
SERVICE 1332
OPERATIONS 1843
SALES 485
FINANCE 176
PRODUCTION 246
PURCHASE 230
MARKETING 202
GENERAL MANAGEMENT 113
HR 70
PIE CHART:
NUMBER OF EMPLOYEES
SERVICE
OPERATIONS
SALES
FINANCE
PRODUCTION
PURCHASE
MARKETING
GENERAL MANAGEMENT
HR

- **BAR CHART:**

TASK4-	
DEPARTMENTAL ANALYSIS	
DEPARTMENT	**NUMBER OF EMPLOYEES**
SERVICE	1332
OPERATIONS	1843
SALES	485
FINANCE	176
PRODUCTION	246
PURCHASE	230
MARKETING	202
GENERAL MANAGEMENT	113
HR	70

BAR CHART:



INSIGHT; we have most of our workers working in the operations department i.e. 1843. And the least number of workers in the human resources department i.e. 70.

E.) POSITION TIER ANALYSIS:

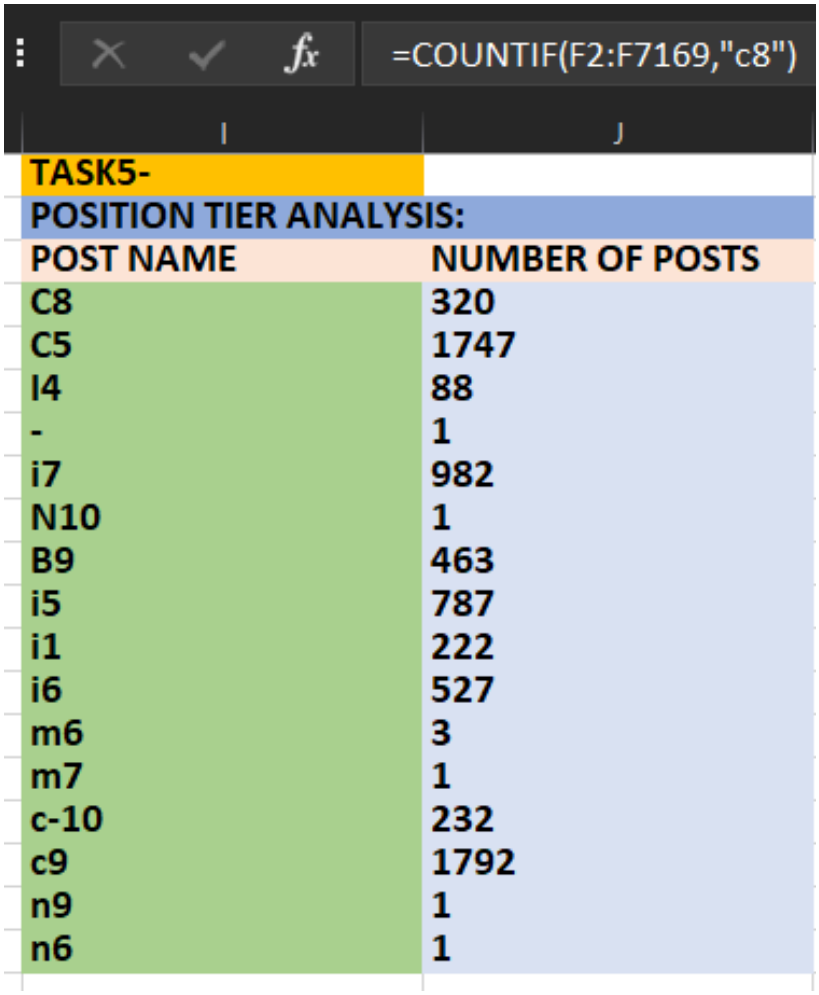
Different positions within a company often have different tiers or levels.

#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.

First, we calculate the number of all the different posts in the company**.**

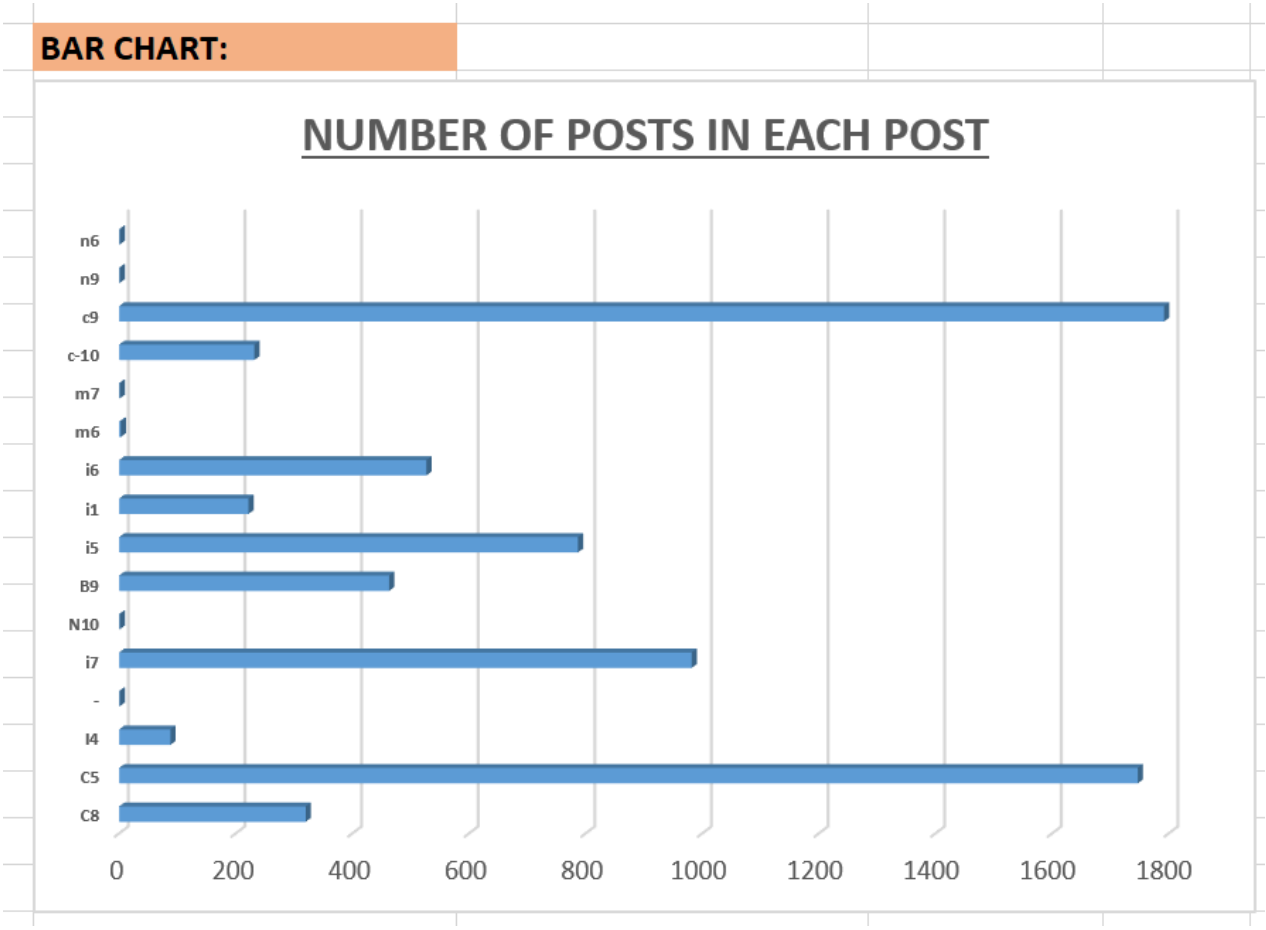
For that, we need to first use the Advanced Filter to get unique posts, and then use the following Formulas→

=COUNTIF(F2:F7169,"c8")
=COUNTIF(F2:F7169,"c5")
=COUNTIF(F2:F7169,"i4")
=COUNTIF(F2:F7169,"-")
=COUNTIF(F2:F7169,"i7")
=COUNTIF(F2:F7169,"n10")
=COUNTIF(F2:F7169,"b9")
=COUNTIF(F2:F7169,"i5")
=COUNTIF(F2:F7169,"i1")
=COUNTIF(F2:F7169,"i6")
=COUNTIF(F2:F7169,"m6")
=COUNTIF(F2:F7169,"m7")
=COUNTIF(F2:F7169,"c-10")
=COUNTIF(F2:F7169,"c9")
=COUNTIF(F2:F7169,"n9")
=COUNTIF(F2:F7169,"n6")



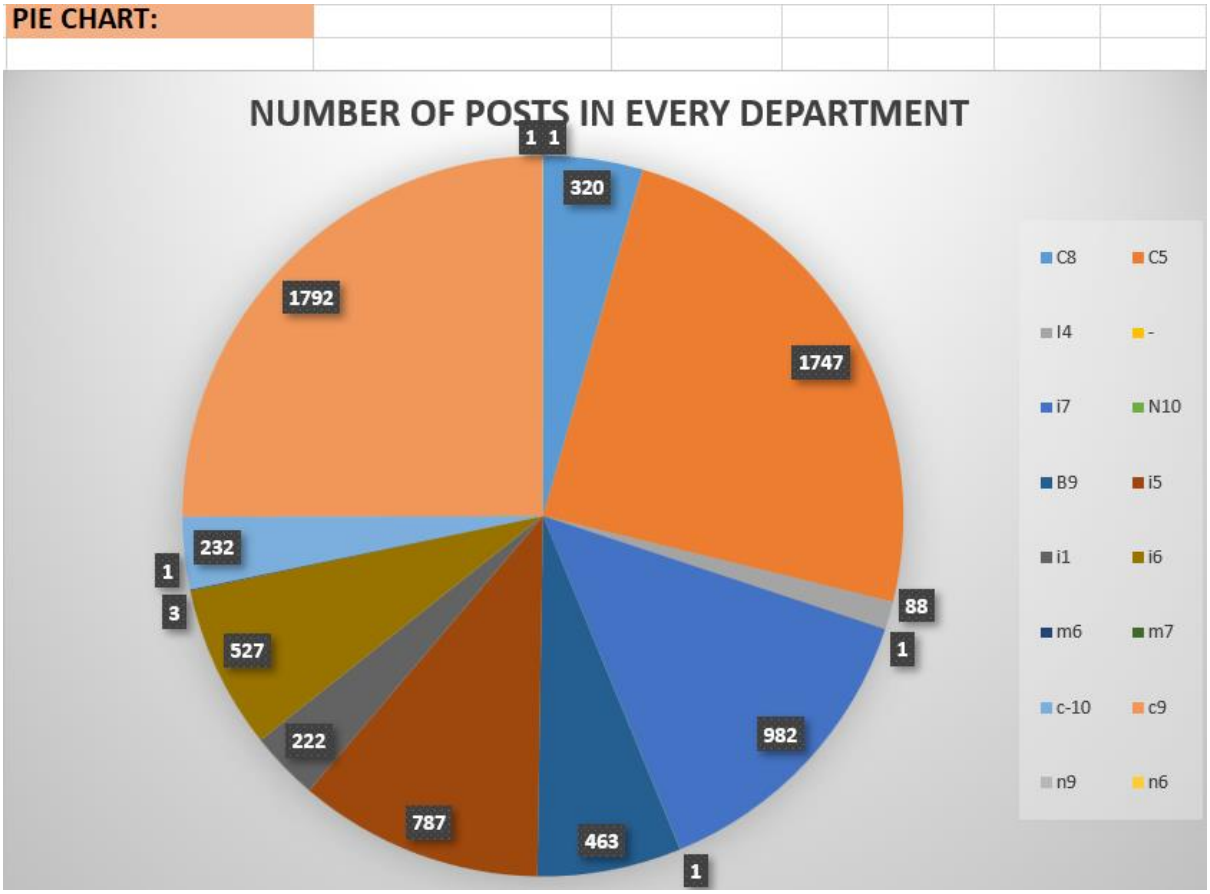
#INSIGHT: We see that the maximum number of seats is in the “c9” post. Furthermore, “-“, ”n10”, “m7”, “n9”, and “n6” are at the same level on the position tier closely followed by “m6”.

- **BAR CHART:**



#INSIGHT: After the visualization, we can see that C9 and N9 posts occupy the majority section of the pie chart

- **PIE CHART:**



#After the visualization, we can see that C9 and N9 posts occupy the majority section of the pie chart.

RESULTS AND CONCLUSIONS:

This project was very helpful for learning data analytics using Excel. I learned about pivot tables and tables. Formulas in Excel to sort data. As I created charts and graphs, it helped visualize the dataset.

A chart is a visual representation of the data. Excel provides you with many charts types and you can choose one that suits your data.

- **Averageifs ()** AVERAGEIFS, like SUMIFS, lets you take an average based on one or more parameters. SYNTAX = AVERAGEIFS (avg_rng, range1, criteria1, [range2], [criteria2], …)

- **Countsifs ()** The COUNTIFS function counts the number of values that satisfy a set of conditions. SYNTAX = COUNTIFS (range, criteria)

- **Counta ()** COUNTA determines whether a cell is empty or not. You will come across incomplete data sets daily as a data analyst. Without needing to restructure the data, COUNTA will allow you to examine any gaps in the dataset. SYNTAX = COUNTA (value1, [value2], …)

HYPERLINK TO EXCEL FILE:

https://docs.google.com/spreadsheets/d/1ZbYb4H1UwG5B4YSzCKr98IIegUTSeyZ0/edit?usp=sharing&ouid=102683227032029211056&rtpof=true&sd=true