HIRING PROCESS ANALYTICS

The project is carried out using Excel to analyze the hiring counts, average salary, salary distribution, distribution of employees across each dept and positions.

The dataset is cleaned before doing the analysis by removing duplicates, records with nil values, and outliers.

Cleaning the dataset

1. 27 duplicates were found based on the application ID and they were removed. 7141 unique applications are remained after cleaning.
2. One row without post name is filtered and removed.
3. 15 records without any event name are identified and deleted.
4. 2457 records are removed as they are rejected.
5. 6 outliers are identified for salary values within the department as well as position and removed.
6. After cleaning the dataset, 4662 records were remained finally.

application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary	Outlier	Q1	Q3	IQR	LOWER LIMIT	UPPER LIMIT
649039	07-05-2014 10:48	Hired	Female	Service Department	b9	200000	TRUE	229 68.5	737 27	507 58.5	-53169.3	149864.8
650516	16-07-2014 09:32	Hired	Male	Human Resource Department	c8	51638	TRUE	337 14	342 70	556	32880	35104
844493	01-07-2014 09:33	Hired	Female	Human Resource Department	i6	95194	TRUE	671 5	132 01	648 6	-30140	229302
874368	21-07-2014 15:39	Hired	Male	General Management	i7	300000	TRUE	361 31	770 31.5	409 00.5	-25219.8	138382.3
88591	17-06-2014 12:15	Hired	Male	Human Resource Department	c8	13900	TRUE	337 14	342 70	556	32880	35104
482656	29-05-2014 09:19	Hired	Female	Production Department	i7	97814	TRUE	197 67.5	506 33.5	308 66	-26531.5	96932.5

A. Hiring Analysis:

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

Gender Distribution of hires:

No. of male candidate hired =COUNTIF(,D:D,"male") = 2548

No. of female candidate hired = COUNTIF(D:D,"Female") =1847

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.

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

=AVERAGE(G:G) = 49661.54

C. Salary Distribution:

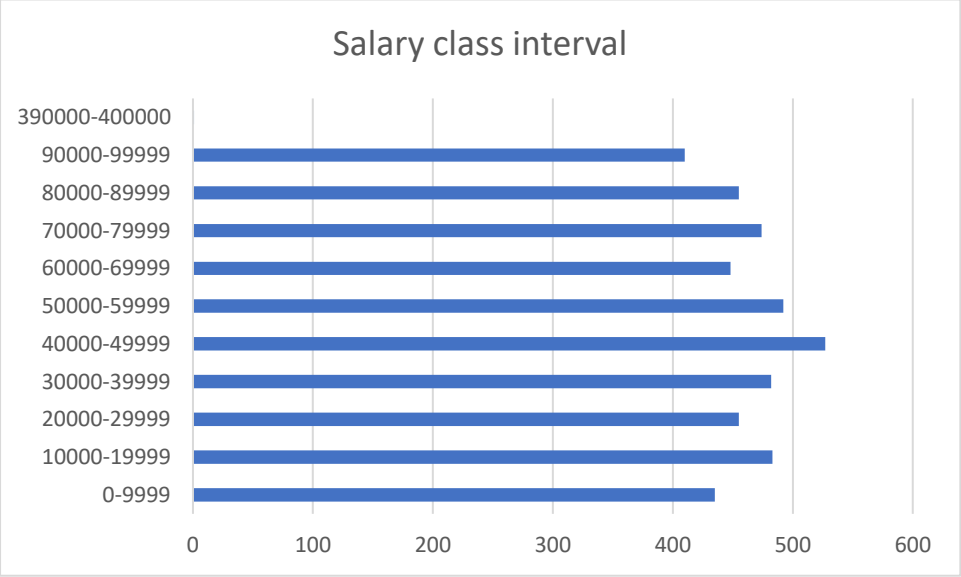
Class intervals represent ranges of values, in this case, salary ranges. The class interval is the difference between the upper and lower limits of a class.

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

A pivot table is created for the column field – offered salary. It is grouped with interval of 10000. Then count of salarties in each group is obtained from pivot table.

Class interval is created for an interval of 10000. Almost every class has employees ranging between 400 and 500. The highest class interval is noticed in the range 40000 to 49999. Lowest class interval is 390000 to 400000.

Salary Range	Frequency
0-9999	435
10000-19999	483
20000-29999	455
30000-39999	482
40000-49999	527
50000-59999	492
60000-69999	448
70000-79999	474
80000-89999	455
90000-99999	410
390000-400000	1

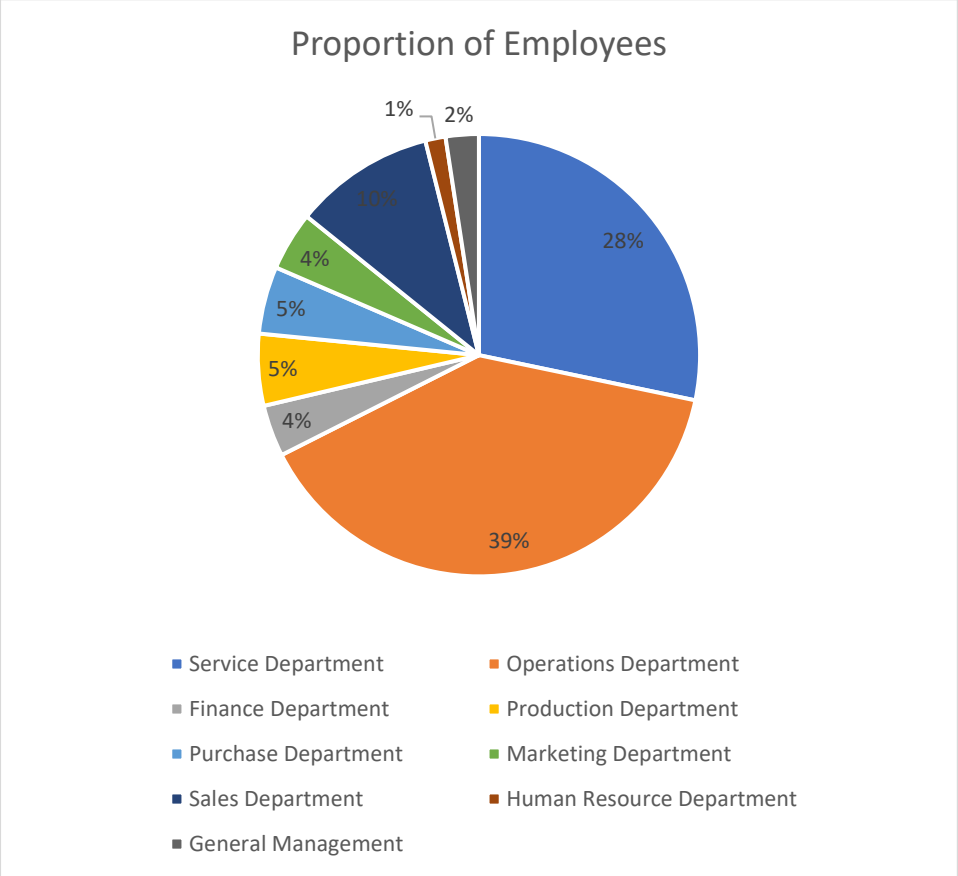


D. Departmental Analysis: Visualizing data through charts and plots is a crucial part of data analysis.

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

Operations department is the largest department holding 39% of total employees and service department is the second largest dept holding 28%. HR dept is the smallest holding just 1% of employees.

Department	No. of People	Proportion of Employees
Service Department	1319	**28%**
Operations Department	1833	**39%**
Finance Department	176	**4%**
Production Department	244	**5%**
Purchase Department	230	**5%**
Marketing Department	201	**4%**
Sales Department	481	**10%**
Human Resource Department	66	**1%**
General Management	112	**2%**
	4662	



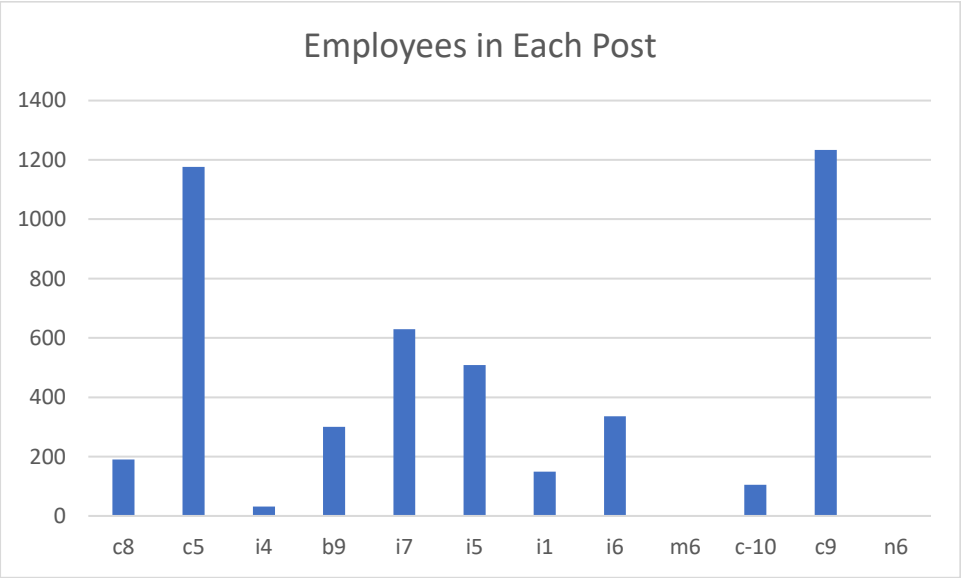
E. Position Tier Analysis: Different positions within a company often have different tiers or levels.

Your 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.

The no. of employees held in each post tiers across all departments is given below. It is calculated using countif function.

Eg: =COUNTIF(Table2[Post Name],'Post analysis'!D6)

Post	Employees
c8	190
c5	1176
i4	32
b9	300
i7	629
i5	509
i1	149
i6	336
m6	2
c-10	105
c9	1233
n6	1
TOTAL	4662



Maximum no. of employees are held in posts C9 and C5 across all departments.

The no. of employees across all tiers and department wise are given below. It is calculated using

Dept.	**Service Department**	**Operations Department**	**Finance Department**	**Production Department**	**Purchase Department**	**Marketing Department**	**Sales Department**	**Human Resource Department**	**General Management**
Post									
c8	68	51	3	3	3	19	36	5	4

c5	331	453	34	48	83	54	141	14	18
i4	10	15	0	1	1	1	3	0	1
b9	119	94	8	28	18	16	15	1	2
i7	165	233	15	39	36	31	77	8	27
i5	124	193	24	26	19	17	53	30	23
i1	50	64	4	16	0	10	2	2	1
i6	69	186	5	19	16	10	21	5	6
m6	1	1	0	0	0	0	0	0	0
c-10	30	44	2	5	2	7	8	0	7
c9	353	499	81	60	51	36	125	4	24
n6	0	0	0	0	1	0	0	0	0
TOTAL	**1320**	**1833**	**176**	**245**	**230**	**201**	**481**	**69**	**113**