Introduction & Objective

Objective:

The hiring process is a critical function of any company, impacting overall productivity, diversity, and employee retention. This project aims to analyze key hiring trends and derive insights that can enhance the recruitment strategy of a multinational company.

Aim:

- To study the gender distribution in hiring and suggest diversity improvements.
- To analyze salary distributions and detect anomalies or biases.
- To evaluate departmental hiring trends and optimize workforce allocation.
- To understand position tier distribution and improve leadership hiring.
- To leverage SQL and Excel-based analytics for actionable insights.

Dataset Overview

The dataset consists of candidate records with attributes including:

- **application_id**: Unique ID for each candidate.
- **Interview Taken on**: Date and time of the interview.
- **Status**: Indicates whether the candidate was *Hired* or *Rejected*.
- **event_name**: Represents gender (*Male* or *Female*).
- **Department**: The department to which the candidate applied.
- **Post Name**: The position applied for.
- **Offered Salary**: Salary offered to hired candidates.
- **Experience**: Years of experience before applying.
- **Education Level**: Qualification of candidates (Bachelor's, Master's, etc.).

- **Interview Score**: A score given to candidates based on their interview performance.

Tech Stack Used

Software: Microsoft Excel 2022

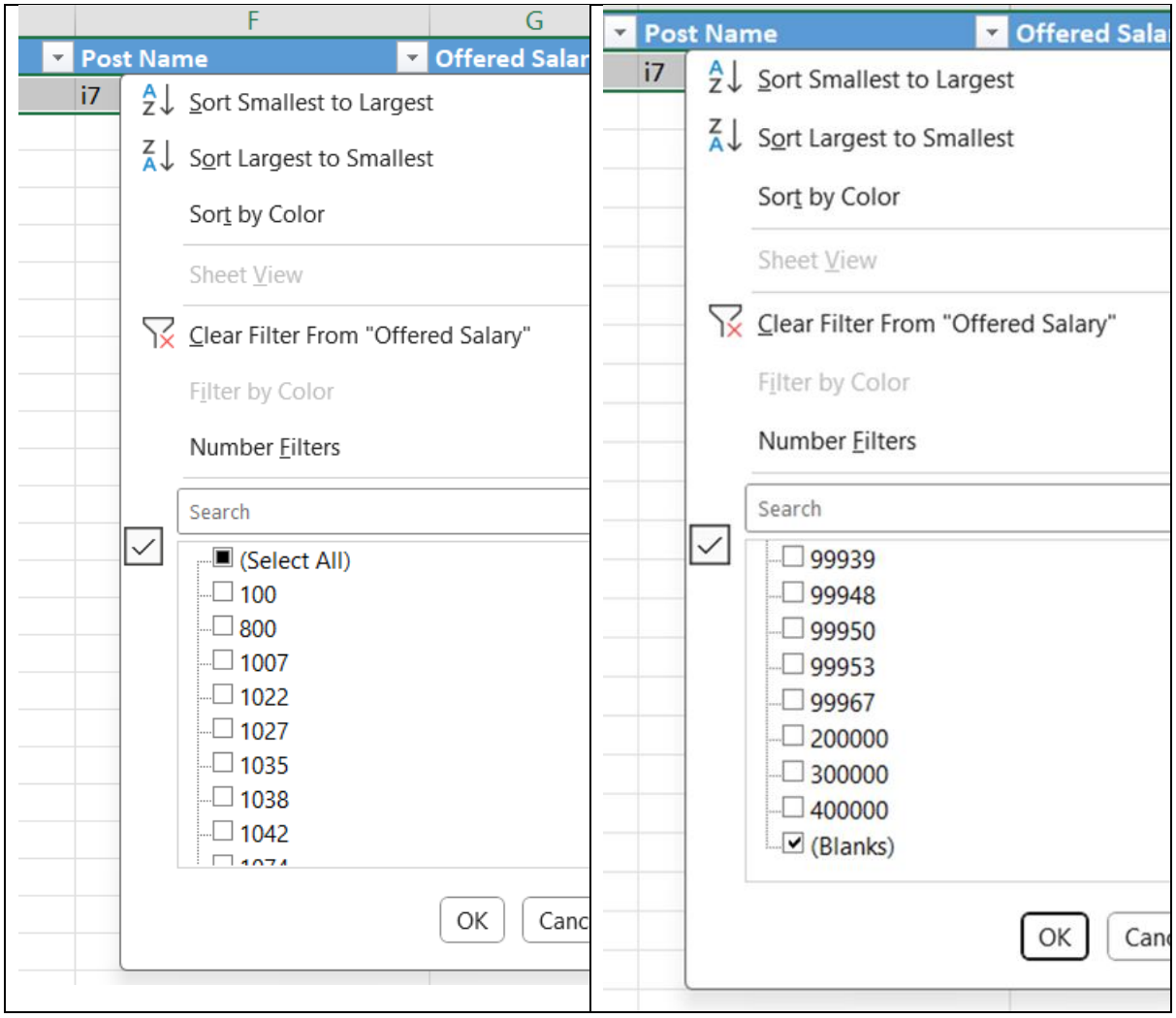
Purpose: Microsoft Excel 2022 was utilized due to its robust data analysis capabilities, including pivot tables, charts, and statistical tools. It played a crucial role in exploring and visualizing hiring data, allowing for a comprehensive examination of trends and patterns.

Data Cleaning & Preprocessing

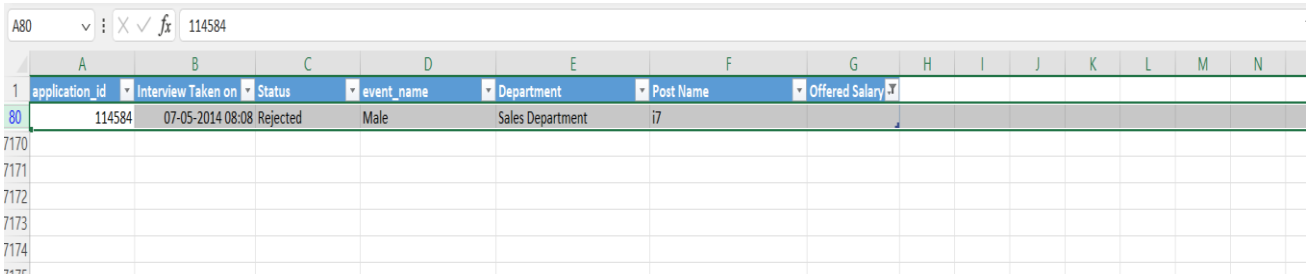
1 Handling Missing Values

- Identified missing values in the **"Offered Salary"** and **"post name"** columns.

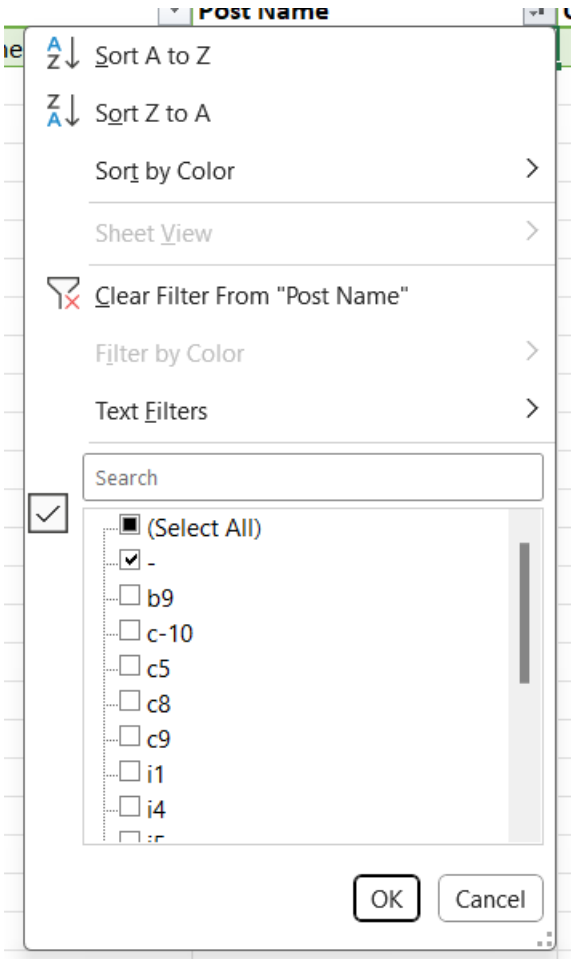
Finding missing values in **"Offered Salary"**



Only one missing value was found in offered salary

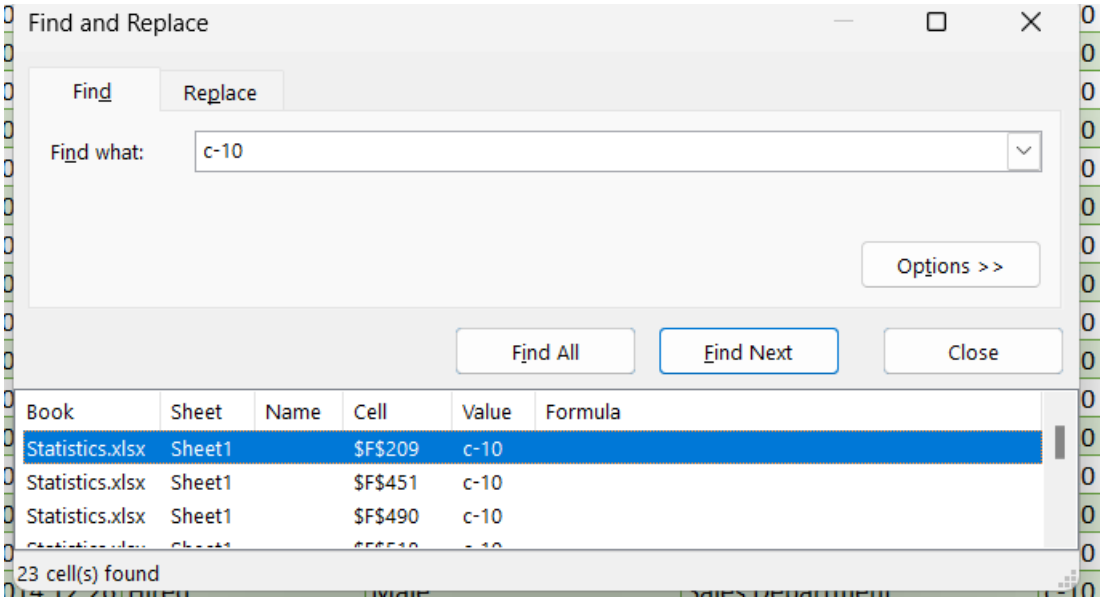


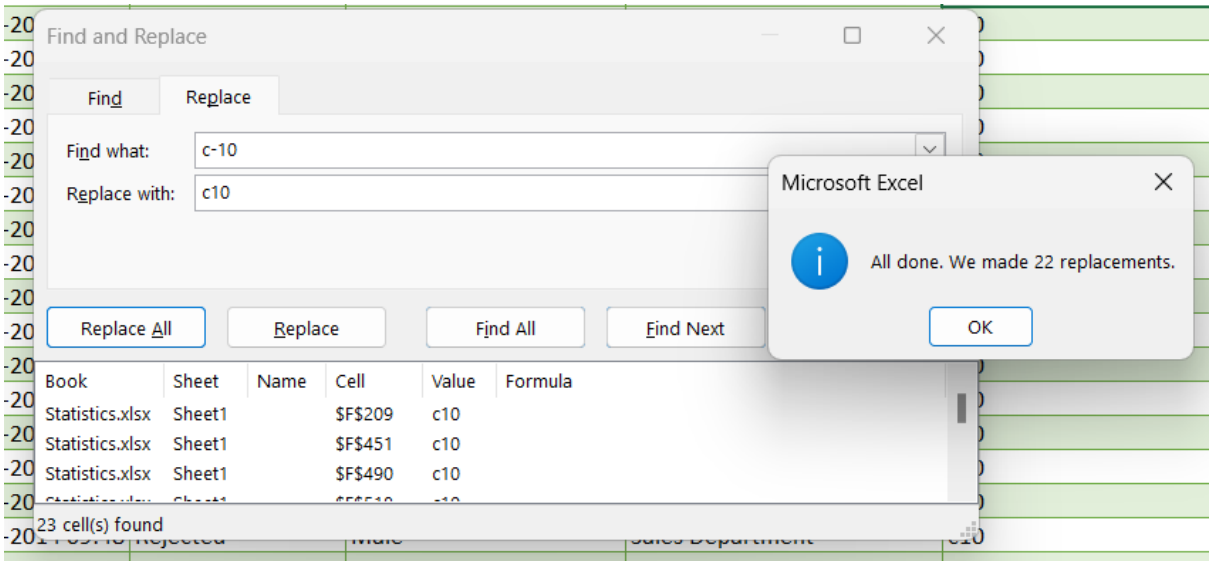
Finding missing values in **"post name"**



	A	B	C	D	E	F	G
1	application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
7	289907	01-05-2014 07:44	Hired	Male	Sales Department		85914
7170							

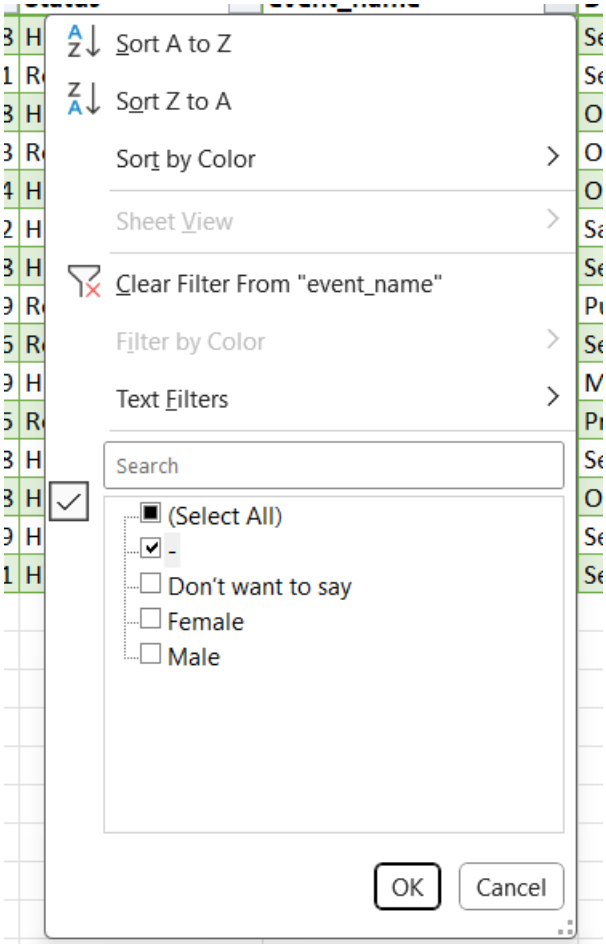
- We Will replace “c-**10**“ by “**c10**” in “**post name**” for data cleaning

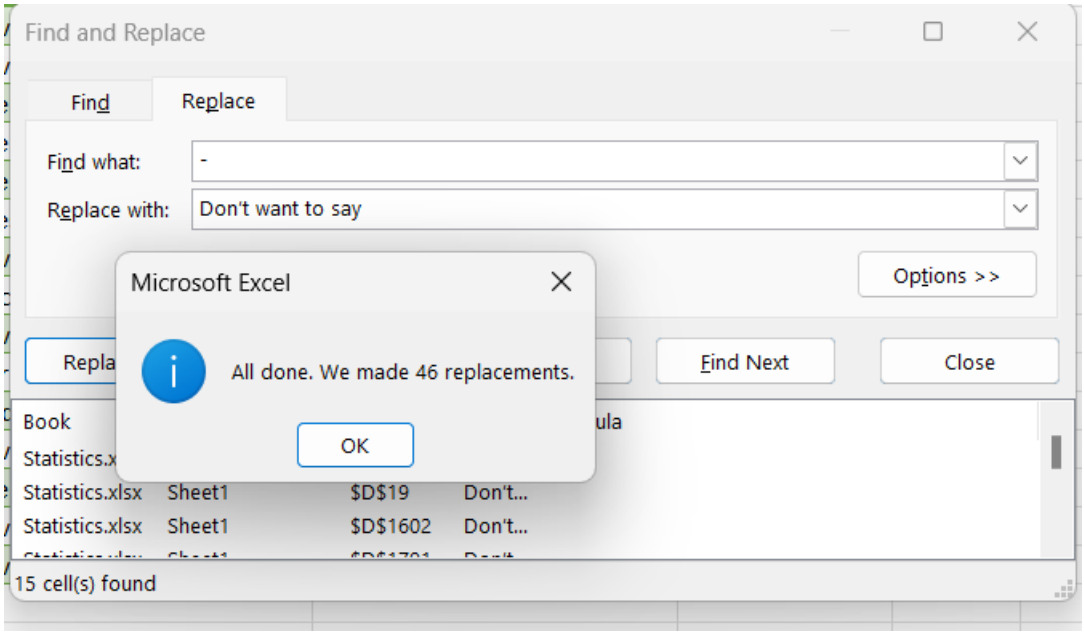




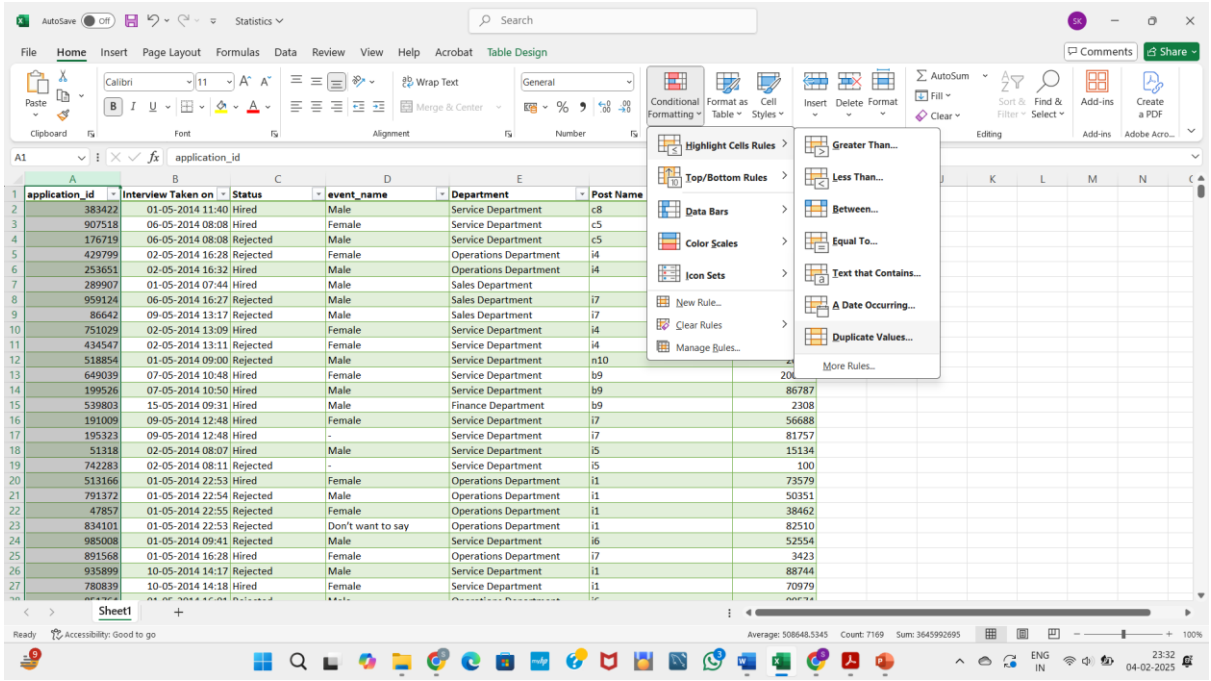
- Filtering status column by replacing" –" with" don't want to say" for data cleaning

	A	B	C	D	E	F	G
1	application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
17	195323	09-05-2014 12:48	Hired	-	Service Department	i7	81757
19	742283	02-05-2014 08:11	Rejected	-	Service Department	i5	100
602	227046	27-08-2014 18:08	Hired	-	Operations Department	b9	76730
791	711350	16-07-2014 13:33	Rejected	-	Operations Department	c-10	25785
878	835053	16-05-2014 18:34	Hired	-	Operations Department	c5	25583
259	444043	11-07-2014 14:52	Hired	-	Sales Department	c5	80262
018	352309	20-08-2014 10:38	Hired	-	Service Department	i5	4308
126	204014	09-08-2014 16:09	Rejected	-	Purchase Department	c5	96396
410	901867	18-08-2014 09:36	Rejected	-	Service Department	c5	22393
560	937905	08-08-2014 19:29	Hired	-	Marketing Department	c9	94032
607	564743	28-08-2014 10:25	Rejected	-	Production Department	c9	4076
889	245473	14-05-2014 18:48	Hired	-	Service Department	c5	66948
330	411295	22-06-2014 14:38	Hired	-	Operations Department	i1	98070
658	487617	30-05-2014 16:29	Hired	-	Service Department	c8	12470
998	827628	30-08-2014 15:51	Hired	-	Service Department	i1	3134





- Cleaning first column "**application_id"** by identifying duplicate values By conditional formatting

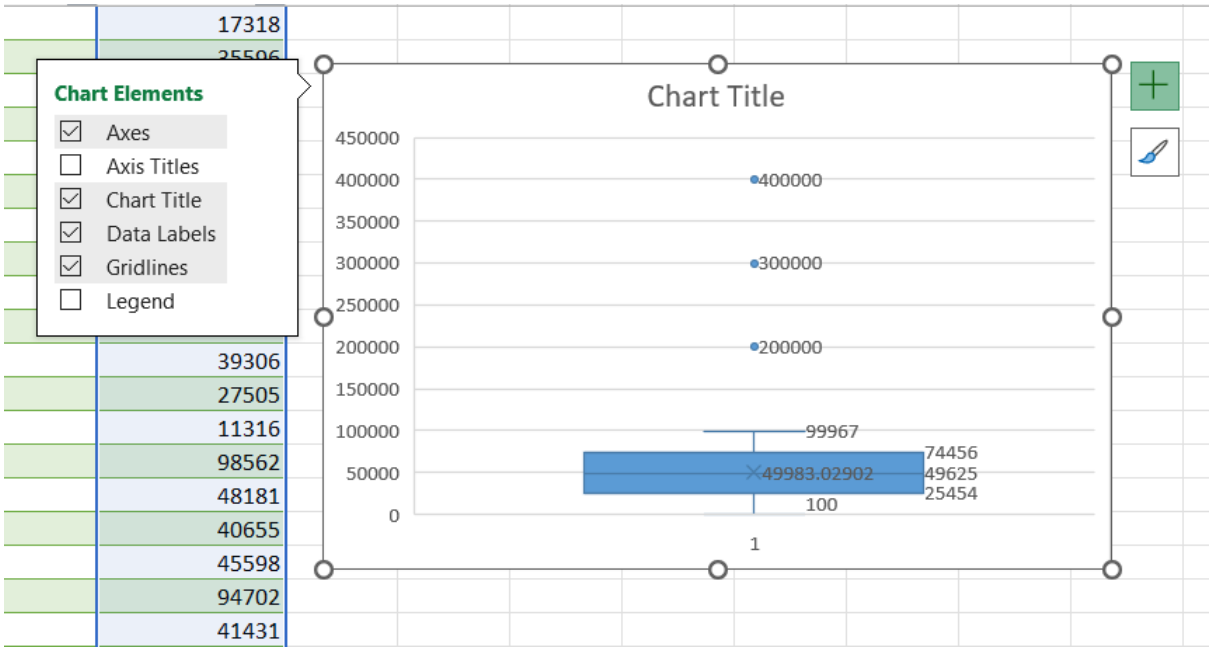


Highlighting duplicate values

2736	618051	25-06-2014 13:40	Hired	Male	Operations Department	c5	86715
2737	384224	25-06-2014 13:39	Rejected	Don't want to say	Operations Department	c5	91019
2738	487216	02-06-2014 18:05	Hired	Male	Operations Department	c5	97277
2739	471695	02-06-2014 18:10	Hired	Male	Operations Department	c5	46200
2740	708065	02-06-2014 18:10	Hired	Male	Operations Department	c5	77840
2741	500670	02-06-2014 18:11	Rejected	Male	Operations Department	c5	20716

2 Detecting & Handling Outliers

- Outlier detection in salary salary column



Removing identified outlier

	B	C	D	E	F	G
1	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
3	07-05-2014 10:48	Hired	Female	Service Department	b9	200000
86	15-06-2014 09:45	Hired	Female	General Management	i4	400000
55	01-08-2014 10:47	Hired	Male	Service Department	c8	99967
325	21-07-2014 15:39	Hired	Male	General Management	i7	300000
70						

Data Analysis & Insights

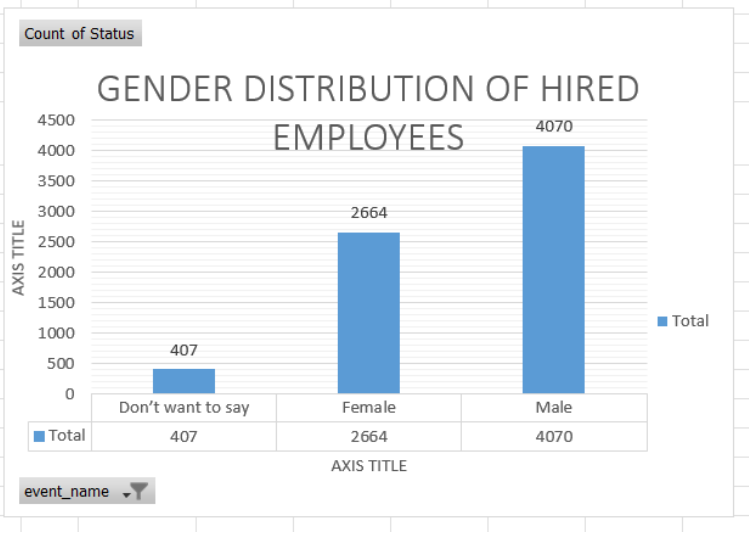
TASK:

A.Hiring Analysis: Gender Distribution

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

STATUTS	HIRED
Row Labels	Count of Status
Don't want to say	407
Female	2664
Male	4070
Grand Total	7141

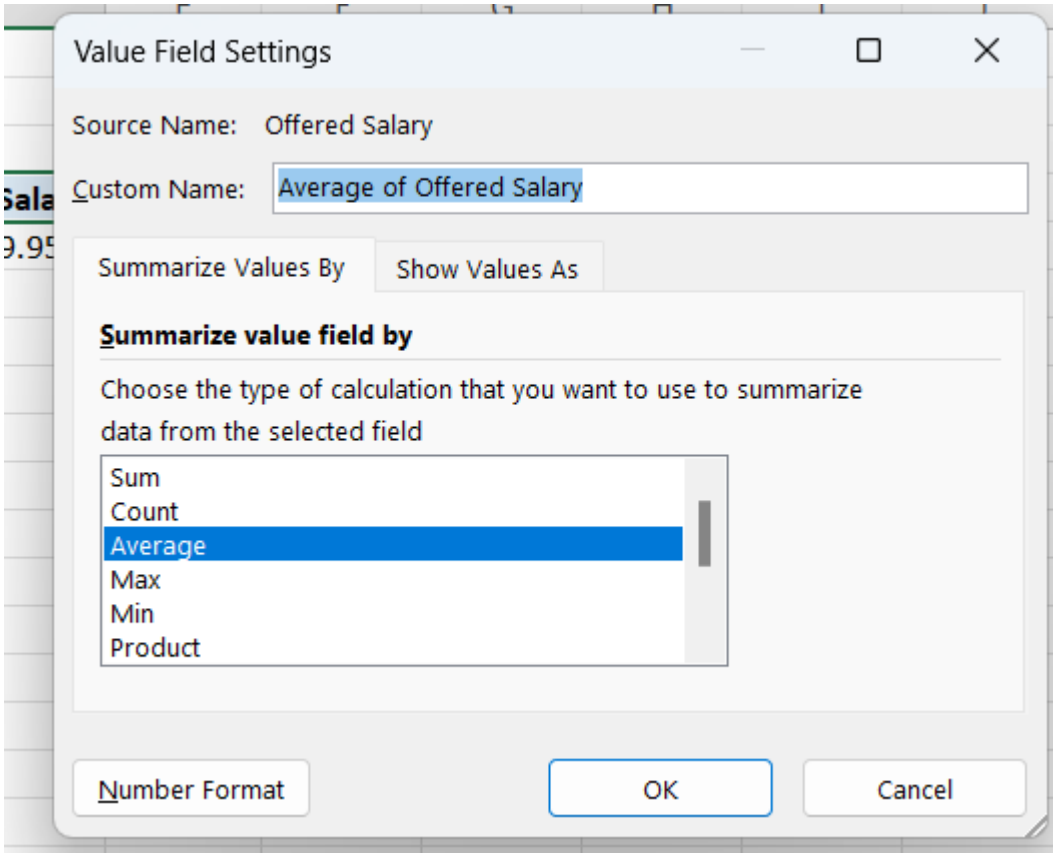
GENDER DISTRIBUTION	
Row Labels	Count of Status
Don't want to say	407
Female	2664
Male	4070
Grand Total	7141

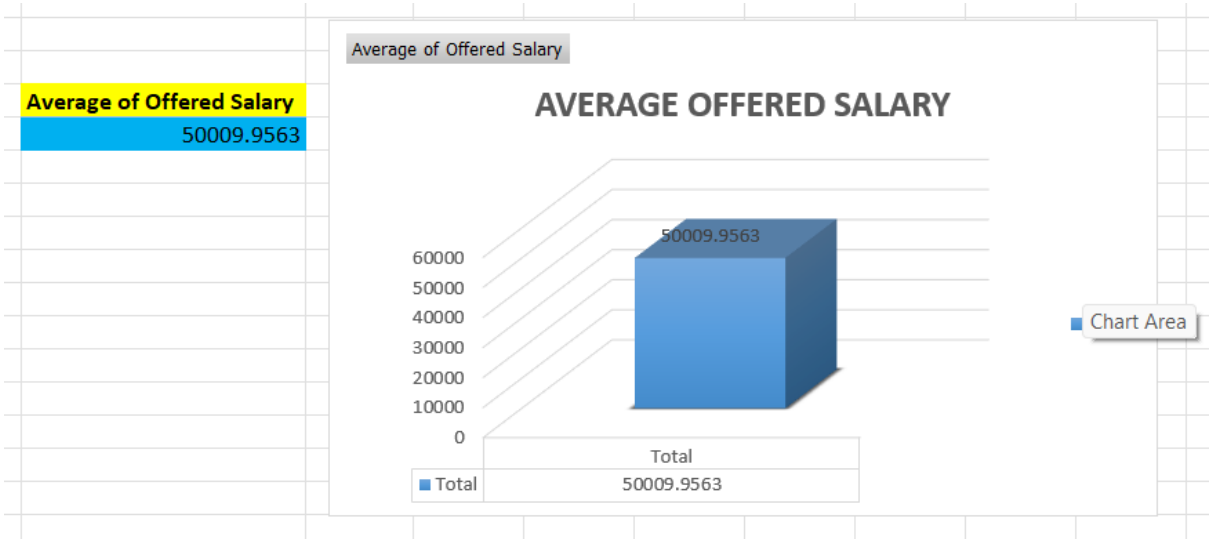


- **Insight:** There is an imbalance in gender hiring, indicating the need for improved diversity policies.

B. Salary Analysis: Average Salary Calculation

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.



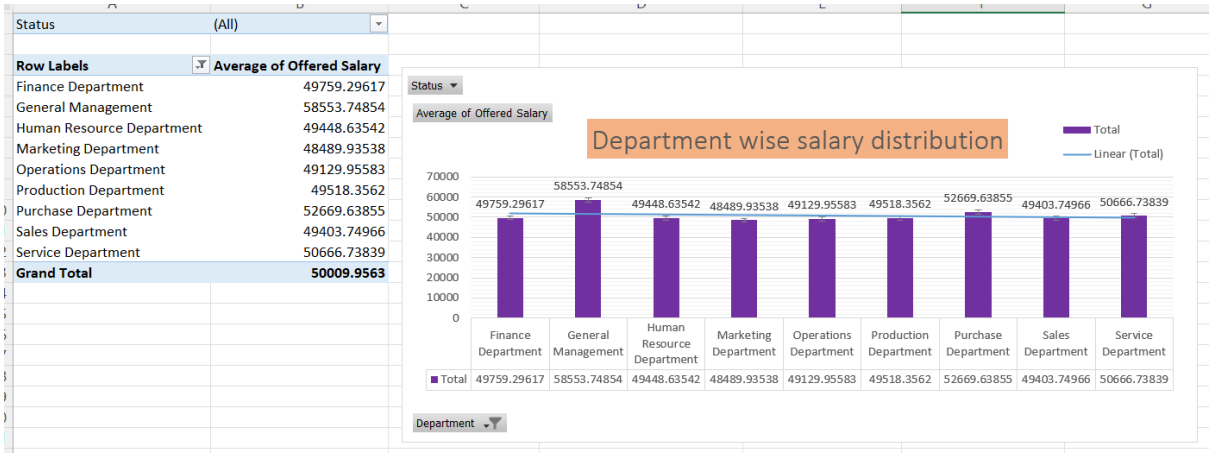


- **Average Salary Offered**: ₹ 50009.9563

- **Insight**: The average salary varies significantly across roles and department.

C. Salary Distribution: Class Intervals

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.



Status	(All)
Row Labels	**Average of Offered Salary**
Finance Department	49759.29617
General Management	58553.74854
Human Resource Department	49448.63542
Marketing Department	48489.93538
Operations Department	49129.95583
Production Department	49518.3562
Purchase Department	52669.63855
Sales Department	49403.74966
Service Department	50666.73839
Grand Total	**50009.9563**

- **Visualization**: Histogram displaying department wise salary distribution.
- **Insight**: Majority of salaries fall within ₹10,000 - ₹40,000, with fewer high-paying roles.

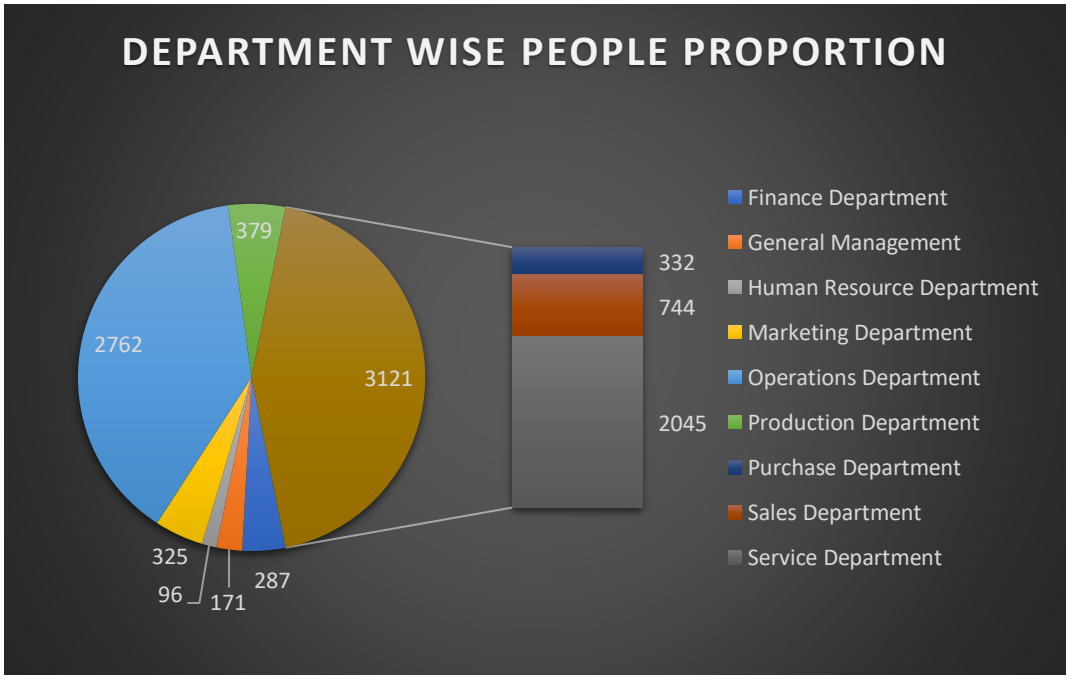
D. Departmental Analysis: Hiring Trends

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

DEPARTMENTS	NO OF PEOPLE
Row Labels	**Count of Department**
Finance Department	287
General Management	171
Human Resource Department	96
Marketing Department	325
Operations Department	2762
Production Department	379
Purchase Department	332
Sales Department	744
Service Department	2045
Grand Total	**7141**

- **Most Hired Departments**: Operation Department and Service Department
- **Least Hired Department**: Human Resource Department

- **Visualization**: Bar chart representing the number of hires per department.



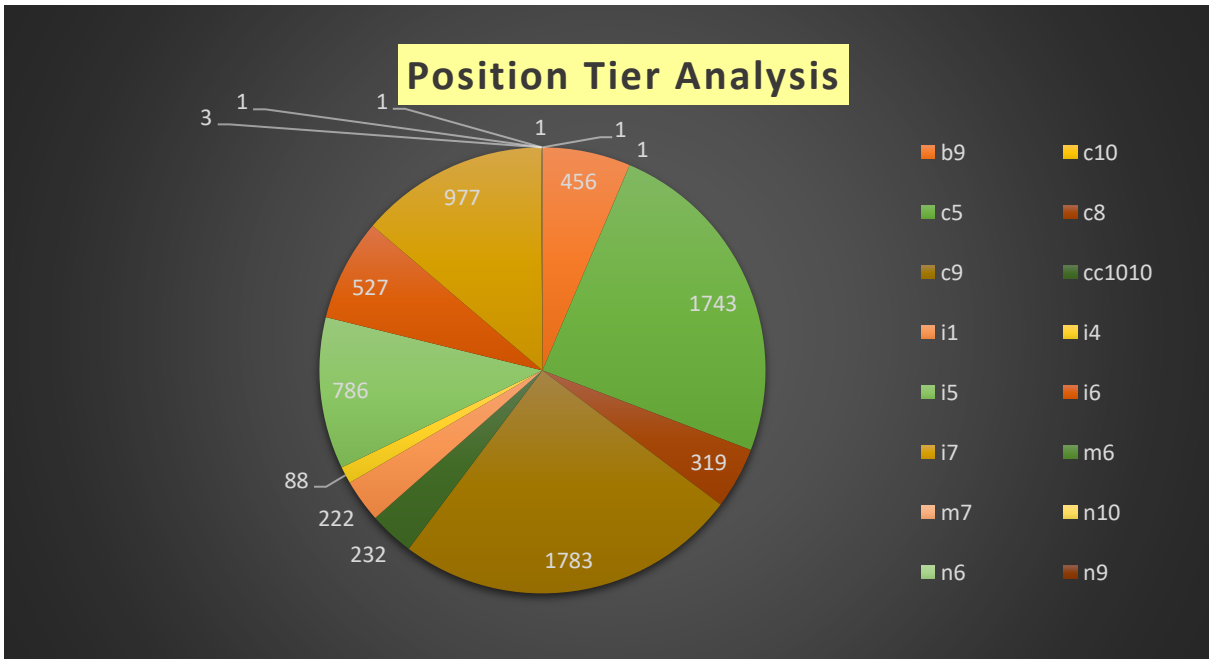
- **Insight**: High hiring in operation and service department indicates tech-driven growth, while Human Resource has limited hiring.

E.Position Tier Analysis: Job Level Distribution

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

POSITION NAME	NO OF PEOPLE
Row Labels	**Count of Post Name**
b9	456
c10	1
c5	1743
c8	319
c9	1783
cc1010	232
i1	222
i4	88
i5	786
i6	527
i7	977
m6	3
m7	1
n10	1
n6	1
n9	1
Grand Total	**7141**

Visualization: Pie chart displaying different position tiers.



- **Insight**: Entry-level roles dominate hiring, suggesting a focus on c9 talent acquisition.

summery

This report outlines the data analysis tasks conducted on the dataset. The findings provide valuable insights into various company aspects, including recruitment trends, salary distribution, department structures, and job hierarchy. These results can be further explored to aid in data-driven decision-making within the organization.

Results

The project delivered meaningful insights into recruitment analytics, offering a clearer understanding of organizational dynamics. By assessing hiring trends, salary distribution, and department composition, useful insights were obtained to refine hiring strategies and support business growth.

The final report has been saved as a PDF file and uploaded to Google Drive.

Click here to see Excel files of the data