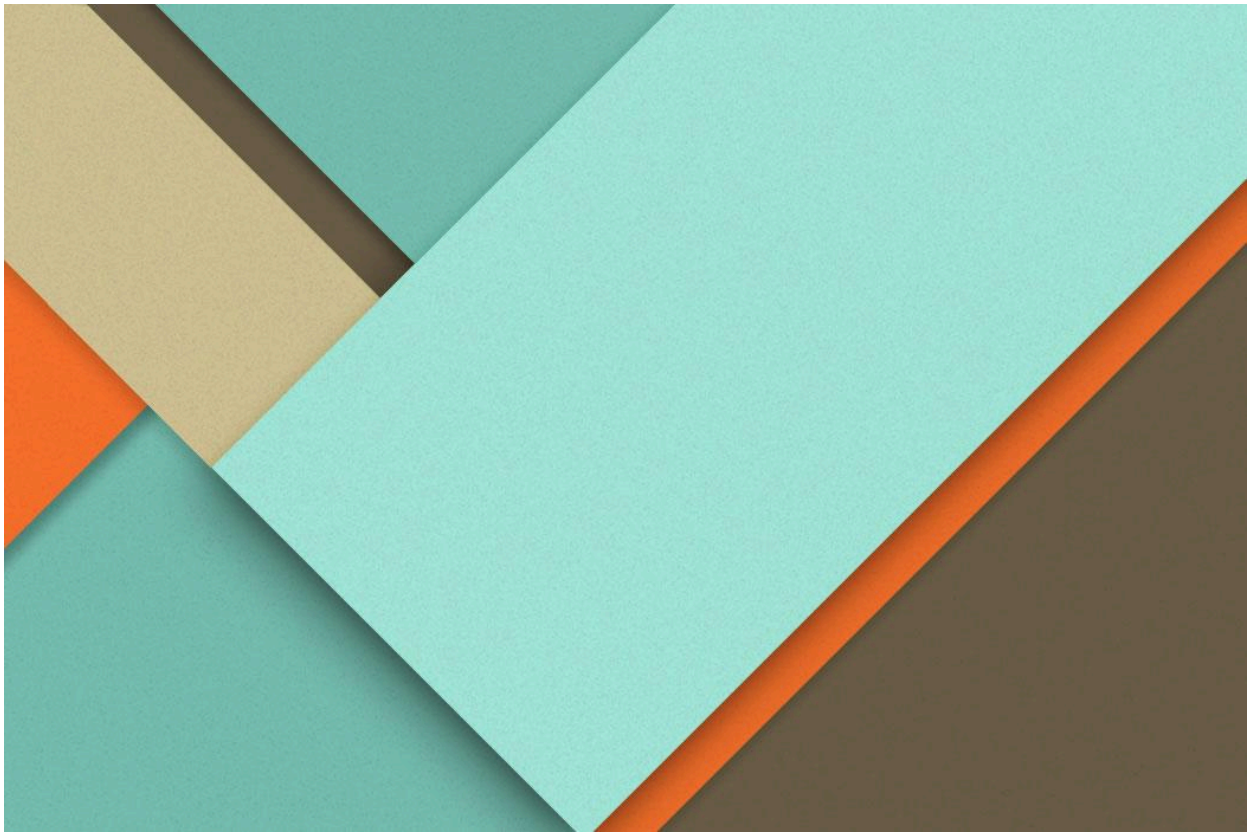


Hiring Process Analytics

November 14, 2024.

Vinay Rana

Project Summary:

This project was developed to examine data from a company's hiring process. The main purpose was to gain insights into areas such as gender diversity, salary trends, departmental structure, and job levels. By looking into these factors, the project aimed to identify patterns in hiring and provide a clearer picture of the company's organizational setup.

Methodology:

We gathered a dataset containing information on new employees, including gender, salary, department, and job level. Microsoft Excel 2022 was used for this analysis due to its strong data handling features. Techniques like pivot tables, charts, and formulas were applied to identify patterns and extract meaningful insights from the data.

Tools and Software:

- Software: Microsoft Excel 2022
- Reason: Excel's powerful tools, such as pivot tables, charts, and statistical functions, allowed us to deeply analyze, visualize, and interpret trends within the hiring data.

Key Insights:

- Gender Distribution: Provided insights into workforce diversity.
- Salary Trends: Showed patterns across various job roles.
- Department Structure: Illustrated team compositions.
- Job Levels: Revealed the distribution of roles by seniority.

These insights offered a better understanding of trends within the hiring process, shedding light on the company’s internal dynamics.

Data Cleaning

- In the 'Event_name' column, rows with a hyphen (-) as their value have been updated to 'Unknown'.

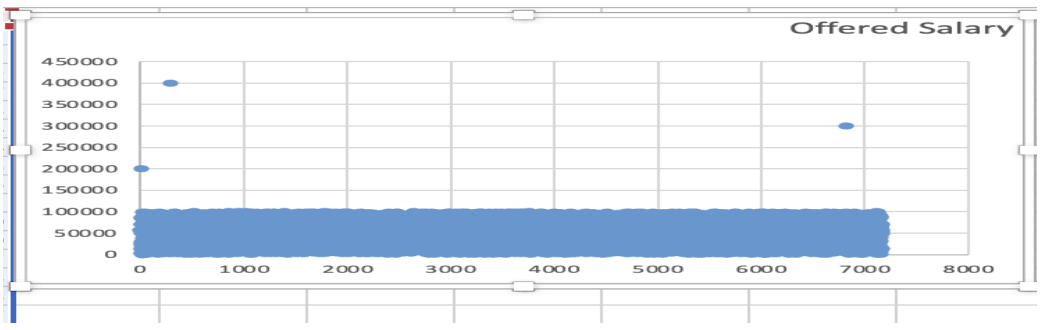
- In the "Offered Salary" column, one row had a missing value. This row belongs to the "Sales Department" with the position "i7." We filled in the missing salary with the median salary for "Sales Department" employees with the "i7" position, which is 45,400. Formula-[=MEDIAN(IF((E:E="Sales Department")*(F:F="i7"), G:G))]

- In the "Post Name" column, the category "c-10" appeared to be a typo. We corrected it to the intended category, "c10".

- The "Post Name" column contains one row with a value of "-," which can be considered a null value. In this case, the corresponding value in the "Department" column is "Sales Department," and the "Offered Salary" is "85914." To address this, we replaced the "-" value with the most frequent "Post Name" for candidates in the "Sales Department" whose "Offered Salary" falls between 85,000 and 96,000. The most common value for this group is "c9."

Handling Outliers

The scatter plot for "Offered Salary" shows outliers at 200,000, 300,000, and 400,000. The first two were replaced with the median salary for the corresponding "Department" and "Post Name," but 400,000 was left unchanged due to its unique post name.



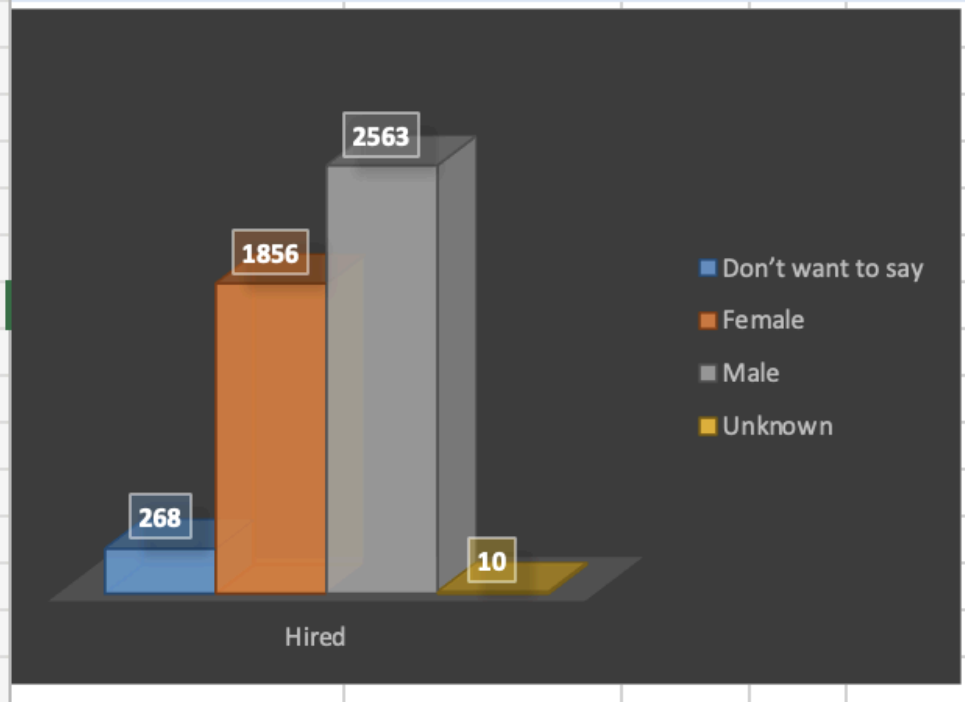
Data Analytics Tasks

A. **Hiring Analysis:** The hiring process involves bringing new individuals into the organization for various roles.

Your Task: Determine the gender distribution of hires. How many males and females have been hired by the company?

Result-

Count of event_name	Column Labels				
Row Labels	Don't want to say	Female	Male	Unknown	Grand Total
Hired	268	1856	2563	10	4697
Grand Total	**268**	**1856**	**2563**	**10**	**4697**



I utilized pivot tables to analyze the hiring data by gender, which showed that the company has hired **1,856** females, **2,563** males, **268** individuals who chose not to disclose their gender, and **10** with unknown gender information.

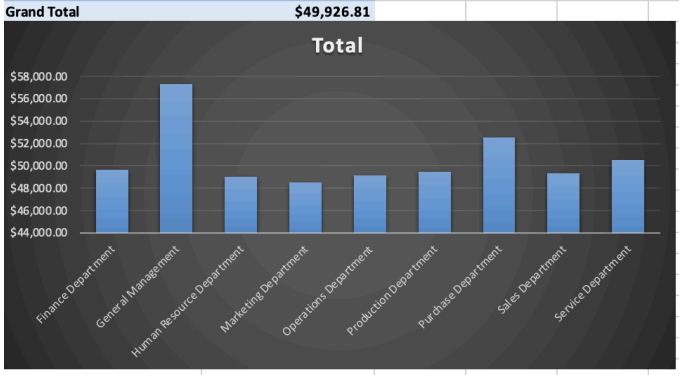
B. **Salary Analysis:** The average salary is calculated by adding up the salaries of a group of employees and then dividing the total by the number of employees.

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

Result-

Method-1

Row Labels	Average of Offered Salary
Finance Department	$49,628.01
General Management	$57,308.26
Human Resource Department	$49,002.28
Marketing Department	$48,489.94
Operations Department	$49,151.35
Production Department	$49,448.48
Purchase Department	$52,564.77
Sales Department	$49,305.15
Service Department	$50,554.37
Grand Total	**$49,926.81**



Method-2

FORMULA-=SUM(Sheet1!G:G)/COUNT(Sheet1!A:A)

49926.81473

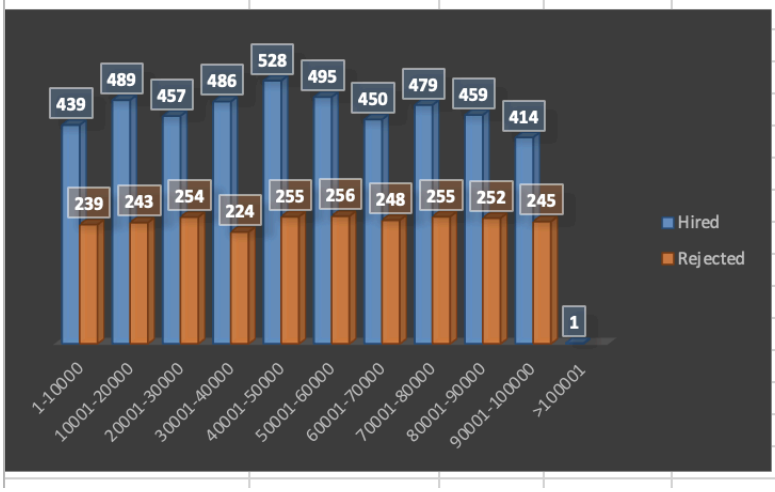
I used two methods to calculate the average employee salary across the company. The results from Method 1 showed that the "General Management" department has the highest average salary, at **$58,722.09**.

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.

Result-

Count of application_id	Column Labels		
Salary Class	Hired	Rejected	Grand Total
1-10000	439	239	678
10001-20000	489	243	732
20001-30000	457	254	711
30001-40000	486	224	710
40001-50000	528	255	783
50001-60000	495	256	751
60001-70000	450	248	698
70001-80000	479	255	734
80001-90000	459	252	711
90001-100000	414	245	659
>100001	1		1
Grand Total	**4697**	**2471**	**7168**



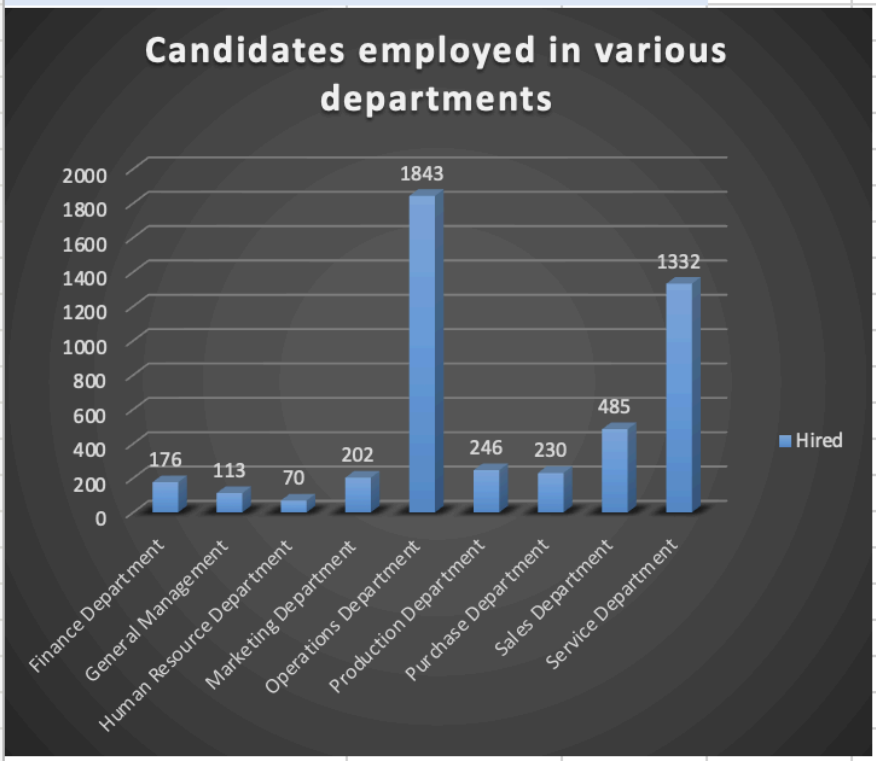
The majority of individuals hired fall within the salary range of **40,000** to **50,000**.

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.

Result-

Count of application_id	Column Labels	
Departments	Hired	Grand Total
Finance Department	176	176
General Management	113	113
Human Resource Department	70	70
Marketing Department	202	202
Operations Department	1843	1843
Production Department	246	246
Purchase Department	230	230
Sales Department	485	485
Service Department	1332	1332
Grand Total	**4697**	**4697**



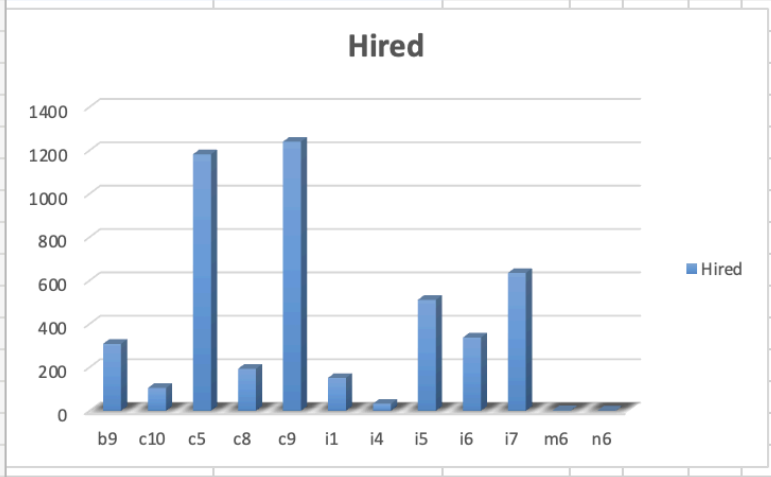
It is evident that the Operations department has the highest number of hires, with **1,843** candidates, while the HR department has the fewest, with only **70** candidates hired in the company.

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.

Result-

Count of hired as per Post name	Count of post name Hired
b9	308
c10	105
c5	1182
c8	193
c9	1240
i1	151
i4	32
i5	511
i6	337
i7	635
m6	2
n6	1
Grand Total	**4697**



It is clear that the post names "c5" and "c9" have the highest number of hires.

Insights and Key Findings:

- Salary Patterns: Salaries largely clustered between $40,000 and $50,000, while the General Management department displayed the highest average salaries, reflecting a possible emphasis on mid-level roles and departmental pay variations.

- Department Structure: The Operations department had the largest number of hires, while HR saw the fewest, highlighting a strategic focus on operational roles to support key business functions.

- Position Distribution: Positions labeled "c5" and "c9" were the most common, indicating a focus on specific job levels that might benefit from targeted talent development and succession planning.

- Data Quality and Outlier Management: Addressing salary outliers improved the reliability of our findings, emphasizing the value of data accuracy in delivering actionable insights.

Project Outcomes and Learning

This project deepened my understanding of how analytics can clarify hiring trends, including diversity, salary patterns, and departmental needs. I developed practical skills in data cleaning, visualization, and interpretation, illustrating how data-driven insights can enhance recruitment strategies, promote diversity, and align hiring practices with business goals.

Excel file link-

Excel sheet

Thankyou