Project Description

This Project involves analyzing a company's hiring process data and draw meaningful insights from it.

As a data analyst, first I will understand the project requirements and purpose and then I will understand the data which is provided in the project attachment and will perform analysis to get the meaningful insights.

Analysis: involves understanding trends such as the number of rejections, interviews, job types, and vacancies can provide valuable insights for the hiring department.
For example, the HR team might use these insights to decide when to conduct interviews and to determine the salary package to offer to the new joinees.

Approach

1. **Understanding the Data**: The first step in any data analysis is understanding the data. From the data in the given attachment, I understood that data consists of several columns including ‘application_id’, ‘Interview Taken Date’, ‘Interview Taken Time’, ‘Status’, ‘event_name’, ‘Department’, ‘Post Name’, and ‘Offered Salary’.

2. **Identifying the Analysis**: Based on your questions, I identified the types of analysis you wanted to perform, which included descriptive statistics, correlation analysis, trend analysis, comparative analysis, and hypothesis testing.

3. **Descriptive Statistics**: Calculated measures of central tendency (mean, median, mode) and measures of dispersion (range, variance, standard deviation) for the ‘Offered Salary’ column. I also Calculated the frequency and percentage of ‘Hired’ and ‘Rejected’ statuses.

4. **Correlation Analysis**: I calculated to find if there’s a correlation between the ‘Offered Salary’ and the ‘Status’ of the application.

5. **Trend Analysis**: Based on your questions analyzed trends over time, such as the number of hires/rejections per day.

6. **Comparative Analysis**: Based on your questions compared the average ‘Offered Salary’ across different departments or between different genders.

7. **Hypothesis Testing**: I suggested t-test in Excel to test if the mean salary offered to males is significantly different from that offered to females. Result shows there is no significant difference.

8. **Creating Class Intervals**: Based on your question, I created class intervals or bins for the 'Offered Salary' column to understand the distribution of salaries in the company.

9. **Visualizing Position Tiers**: Finally, created a bar chart or pie chart in Excel to represent the different position tiers within the company.

Throughout this process, my goal was to provide clear, accurate, and helpful insights and answers to questions.

Tech-Stack Used

The software and versions you used for the project:

Microsoft Excel
Pivot Tables
Data analysis tool pack

Insights:

I will provide my insights in the form of answers for the questions posed by management team

1. **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?

Output:

4		Column Labels		
5	Gender	Hired	Rejected	Grand Total
6	Male	2564	1523	4087
7	Female	1865	823	2688
8	Don't want to sa	268	125	393
9	**Grand Total**	**4697**	**2471**	**7168**

2. **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.

Output:

14	**Average of Offered Salary**
15	49984.63

Suggestion: We can perform t-test in Excel to test if the mean salary offered to males is significantly different from that offered to females

t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	49920.08074	50094.16592
Variance	817695113.5	856385562.7
Observations	4087	2688
Pooled Variance	833044476.7	
Hypothesized Mean Difference	0	
df	6773	
t Stat	-0.242879021	
P(T<=t) one-tail	0.404053236	
t Critical one-tail	1.645078635	
P(T<=t) two-tail	0.808106472	
t Critical two-tail	1.9603143	

Result is P Value (P(T<=t)) > given common significance level(0.05), which shows there is no significant difference.

3. 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.

Output:

Salary Range	Count of Offered Salary
10000-29999	1275
30000-49999	1490
50000-69999	1450
70000-89999	1445
90000-109999	659
Grand Total	**6319**

To answer this question used table from "Table exclude outliers" sheet. Salary Ranges calculated after removing the High and Low outliers from the data. Refer "Table3" sheet for outliers' calculation

Outliers calculation on Offered Salary column

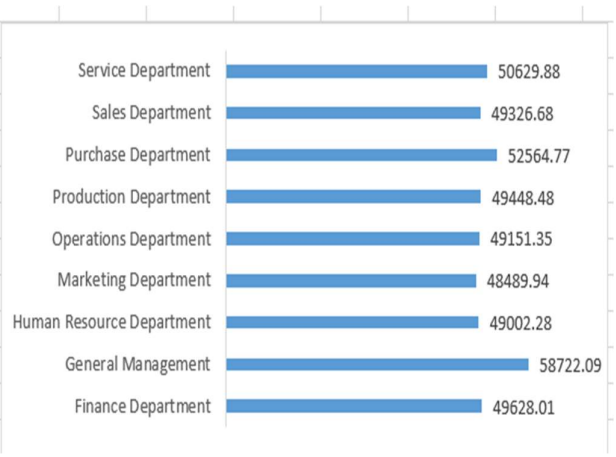
Mean	49984.634
Standard Deviation	28850.471
Lower Bound	12479.022
Higher Bound	136536.05

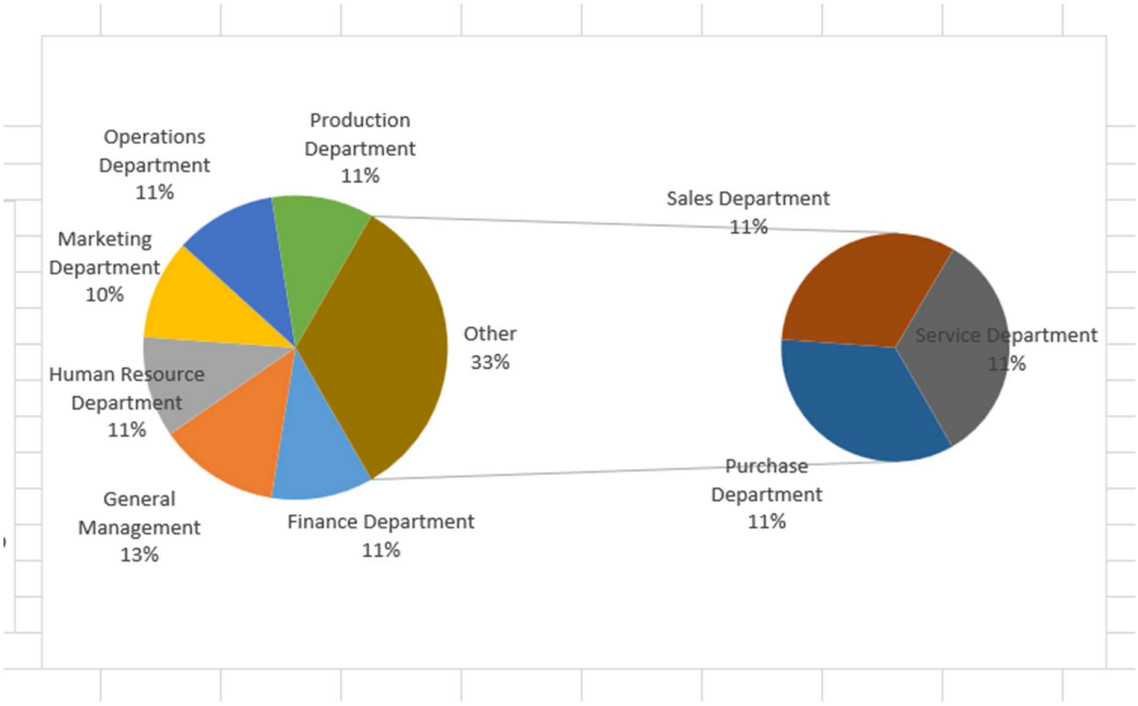
4. Departmental Analysis: Visualizing data through charts and plots is a crucial part of data analysis.

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

Output:

Department	Average of Offered Salary
Finance Department	49628.01
General Management	58722.09
Human Resource Department	49002.28
Marketing Department	48489.94
Operations Department	49151.35
Production Department	49448.48
Purchase Department	52564.77
Sales Department	49326.68
Service Department	50629.88
Grand Total	**49984.63**

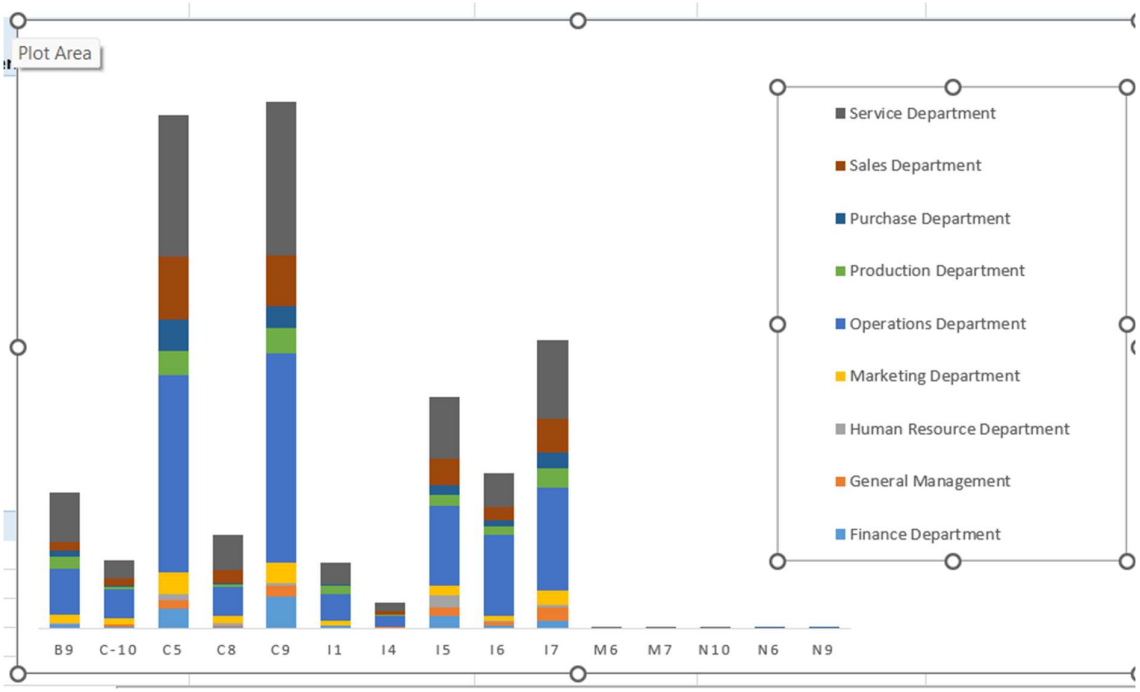




5. Position Tier Analysis: Different positions within a company often have different tiers or levels.

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.

Output: As per below chart Service and Operations Departments has more requirements under different posts or positions



Result

This project definitely given chance to play with given data and I regained my confidence on using Pivot tables and rebrushed my statistics knowledge.

When it comes to this project this analysis provides valuable insights into the company's hiring practices and salary structure. It can help the company make informed decisions about hiring, salary offers, and more. However, it's important to remember that these results are based on the provided sample data and may not represent the entire population of the company.

Drive Link

Link for the SQL query file:

https://docs.google.com/spreadsheets/d/1GeunDc6xGZfA-IWI7Lm702gR7AVDCxj_/edit?usp=sharing&ouid=106678072727235537296&rtpof=true&sd=true