

Final Summary Report: Hiring Data Analytics Project

Analyst: Jeenal Bolia

Excel Sheet: [accessible here](#)

Project Description

This project involved the assessment of a company's hiring data set for trends in recruitment of gender, salaries, department, and position levels. The goal was to create relevant business insights to provide data informed hiring patterns and optimize workforce planning. In this exploratory analysis, real-world HR data were cleaned, processed, and visualized using excel.

Approach

To begin my analyses, I preprocessed the data by cleaning missing values and addressing salary outliers with the interquartile range (IQR) method. The analyses had several sections:

- Hiring distribution by gender using PivotTables
- Average salaries calculation using conditional functions
- Classifying salaries by range using class intervals
- Department and tier hiring patterns shown using bar and pie charts

I attempted to take each task step-by-step, which was valuable for clarity and insight authenticity. The analyses involved considerable amounts of conditional formatting, calculations using AVERAGEIFS, IF, and QUARTILE.INC functions, and use of PivotTables.

Tech Stack Used

Microsoft Excel 2022

Used for data preprocessing, statistical analysis, creating PivotTables, applying conditional formatting, and charts (bar, pie, histogram, etc.) to visualize trends and distributions.

Insights

- **Gender Balance in Hiring:** The number of male candidates hired is nearly equal to the number of female candidates hired, and a very small number of hiring is either unknown or non-binary.
- **Salary Trends:** Offer amounts for the vast majority were at between ₹40K–₹80K, with most salary offers around ₹49,983.

- Outlier Detection: The salary range detection showed invalid to be above ₹147,905, which would have been considered an outlier as specified above, to maintain fairness and consistent hiring decision making.
- Department Analysis: The highest number of hires was in the Service and Operations departments, which signify possible areas of growth and strategic path direction.
- District Hiring at the Tier: There seems to be a concentration of hires at i4, i5, etc. the organization seems to be hiring mid-level positions, which also indicates a focus on operational-type roles instead of entry-level roles.

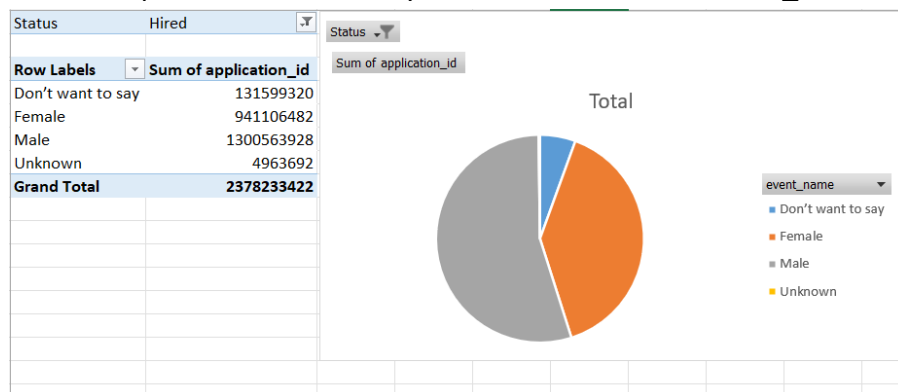
Results

1. Data Cleaning

- Handled missing values in every column, replaced '-' and blank spaces by "Unknown" using logic-based imputation.
- Checked for Outliers in Offered salary using IQR method.
 - Q1 = ₹25,460.5
 - Q2 = ₹74,438
 - IQR = ₹48,977.5
 - Upper bound = ₹1,47,905.25
- Flagged outliers using Conditional formatting in excel.

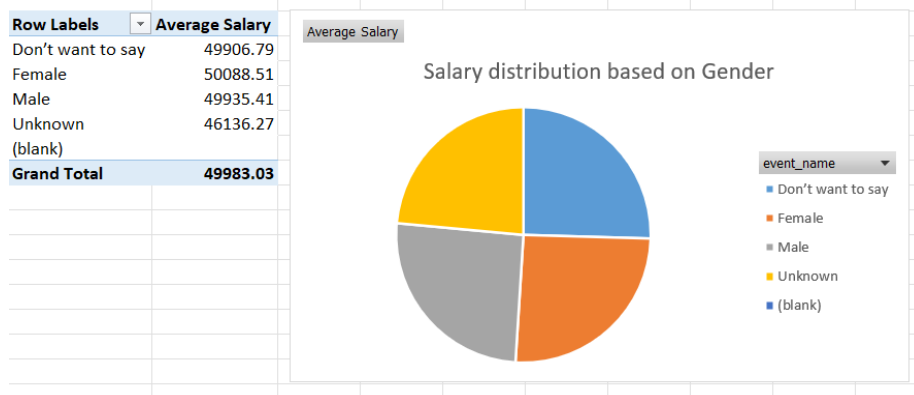
2. Gender Distribution of Hires

- Created a pivot table filtered by Status = Hired, Row = event_name



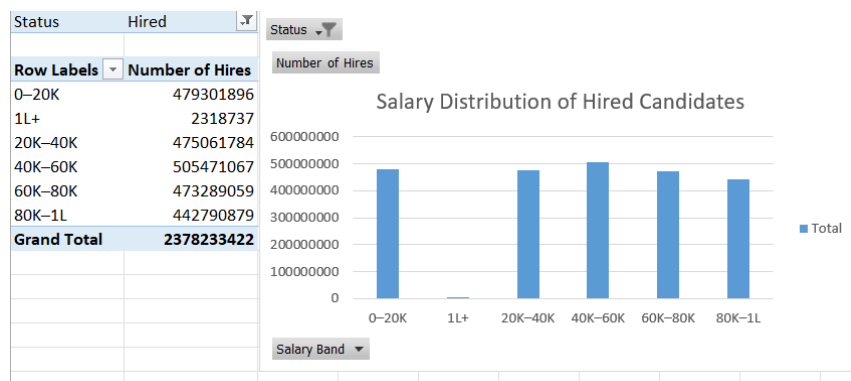
3. Salary Analysis

- Average Salary = ₹49,983.03



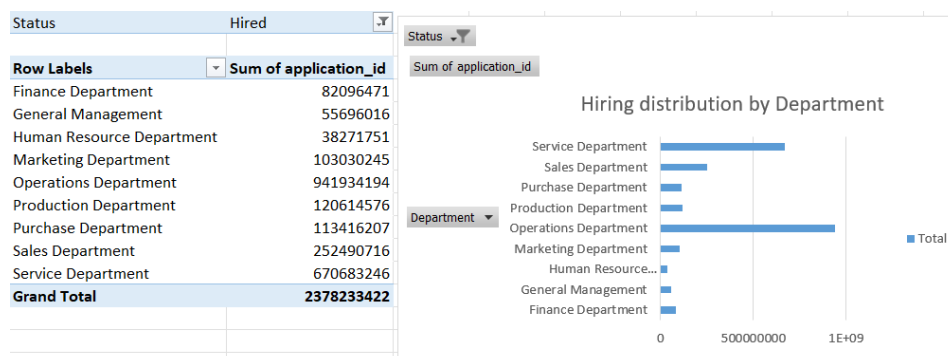
4. Salary Distribution (Histogram)

- Created a Pivot table and bar chart that shows:
 - Most salaries fall in **40K – 80K** range
 - Very few high-end outliers beyond **₹1L+**



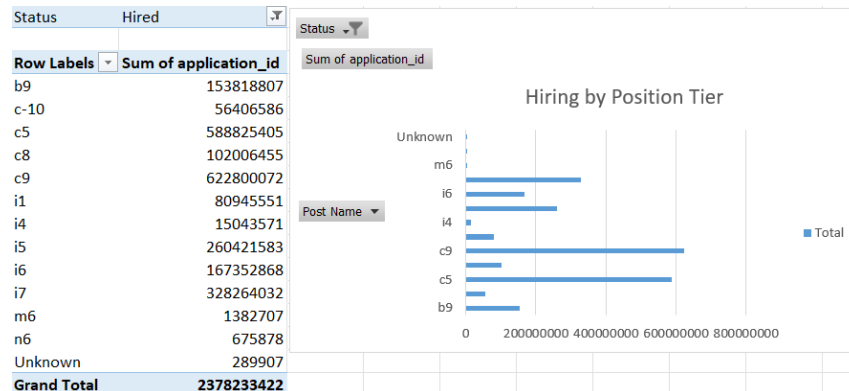
5. Departmental Hiring Analysis

- Created Pivot table with rows = Department and values = application_id
 - Service** and **Operations** Departments had the highest number of hires
 - Human Resource, General Management** had fewer hires



6. Position Tier Analysis

- Found concentrated hiring at certain tiers like – c9 and c5



Impact on business

The above information can help the hiring department:

1. Identified trends with over- and under-staffing
2. Identified the gender balance of hiring and pay
3. Identified salary outliers to have consistency
4. Adjust hiring practices based on departments and to roles

Answering to the questions –

1. Determine the gender distribution of hires. How many males and females have been hired by the company.

Status	Hired
Row Labels	Sum of application_id
Don't want to say	131599320
Female	941106482
Male	1300563928
Unknown	4963692
Grand Total	2378233422

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

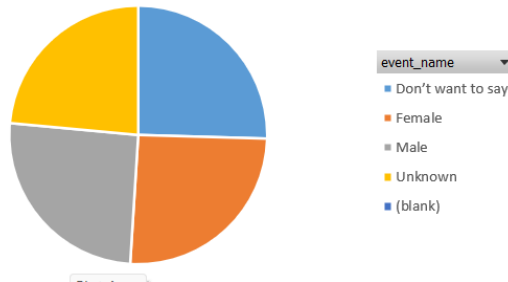
Formula used -> =AVERAGEIFS(G2:G7169, C2:C7169, "Hired")

Result = 49983.03

Row Labels	Average Salary
Don't want to say	49906.79
Female	50088.51
Male	49935.41
Unknown	46136.27
(blank)	
Grand Total	49983.03

Average Salary

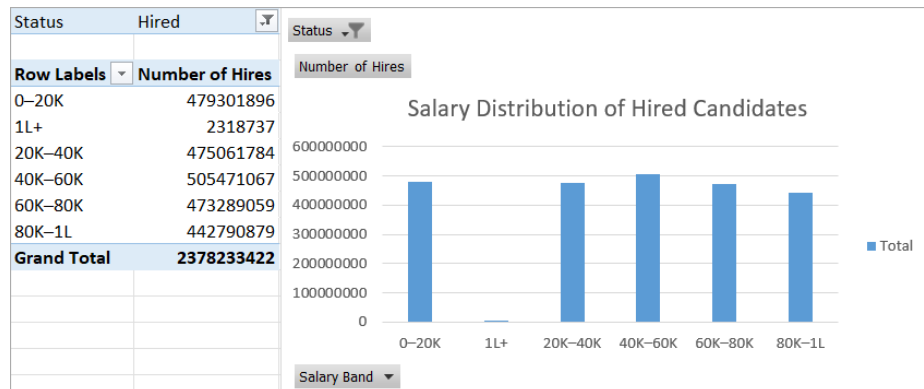
Salary distribution based on Gender



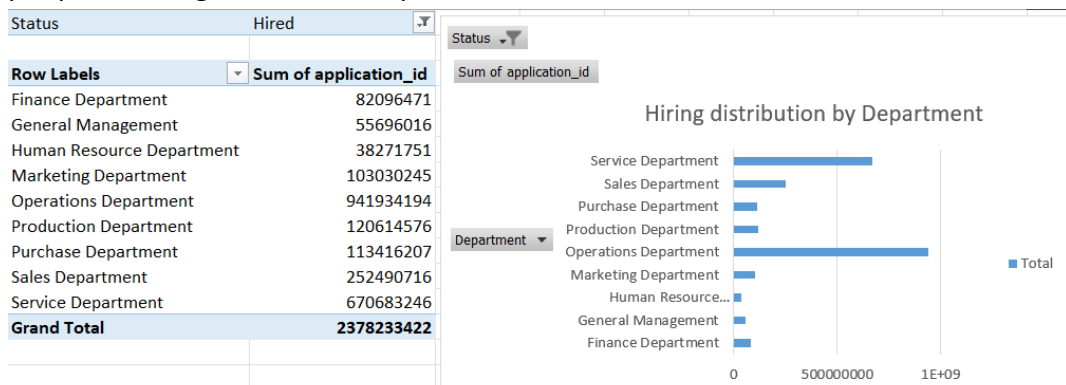
3. Create class intervals for the salaries in the company. This will help you understand the salary distribution.

- Created new column Salary Band using Formula –

=IF(G2<= 20000,"0–20K", IF(G2 <= 40000,"20K–40K", IF(G2<= 60000,"40K–60K", IF(G2<= 80000,"60K–80K", IF(G2 <= 100000,"80K–1L","1L+")))))



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



5. 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.

