Hiring Process Analytics

Report by – Siva Sankari H

1.Introduction –

1.1 Project Description –

Imagine you're a data analyst at a multinational company like Google. Our task is to analyse the company's hiring process data and draw meaningful insights from it. The hiring process is a crucial function of any company, and understanding trends such as the number of rejections, interviews, job types, and vacancies can provide valuable insights for the hiring department.

As a data analyst, we will be given a dataset containing records of previous hires. Our job is to analyse this data and answer certain questions that can help the company improve its hiring process. The goal of this project is to use your knowledge of statistics and Excel to draw meaningful conclusions about the company's hiring process.

2. Dataset

The dataset which we are working with is given along with the project description. The findings on the dataset are mentioned below –

1. There are no duplicates rows in the dataset. We have total 7169 records. Each record represents the applicant details in the company's database. Each row has applicant id, the date and time on which interview was taken, status of the application, event name, department name, post name and salary offered.
2. Applicant id is a unique id given to an applicant. On analysis of the dataset, it is understood that an applicant can apply to n- number of jobs he desires within the company. We have around 54 such applicants. Only application per vacancy can be found.
3. The dataset has hiring information spanning from May,2014 to August,2014. The hiring ahs been done for 9 departments namely, Finance Department, General Management, Human Resource Department, Marketing Department, Operations Department, Production Department, Purchase Department, Sales Department, and Service Department.
4. The interview is held for 16 positions of which 1 remains unknown.
5. For the ease of analysis, we have renamed the "event name" column to "Gender".

3. Approach and Analysis and Insights Derived

Data Analytics Tasks:

After downloading the dataset, use Excel to answer the below questions:

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

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

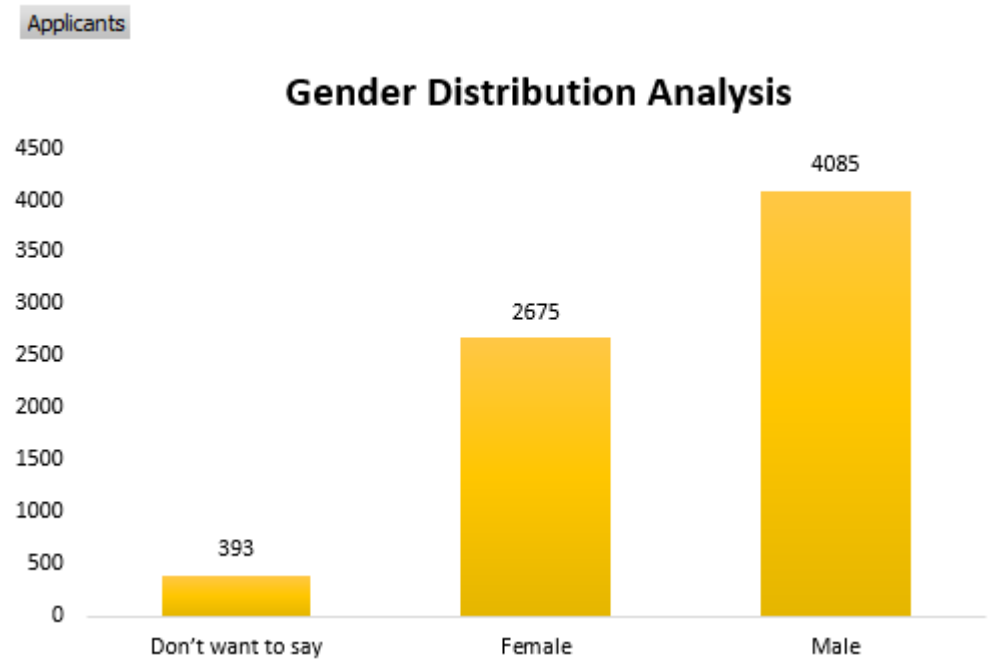
Approach –

We have used Excel's Pivot Table feature to plot the gender distribution against applications received. The findings are shown below.

Analysis –

We have got response from different gender categories. Male – 4085, Female – 2675. The company has included a "don't want to say" category for gender role – creating a more inclusive and equitable hiring process, fostering a diverse and respectful workplace. This option provides a space for individuals who may identify as non-binary, genderqueer, genderfluid, or prefer not to disclose their gender identity. We have 393 such applicants. 15 applicants have left this category blank we are ignoring them in order to get a clear gender distribution.

Gender	Applicants
Don't want to say	393
Female	2675
Male	4085
Grand Total	**7153**



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.

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

Approach –

Excel's statistical function is used here to calculate average. The pivot table is used to understand the salary distribution against various departments.

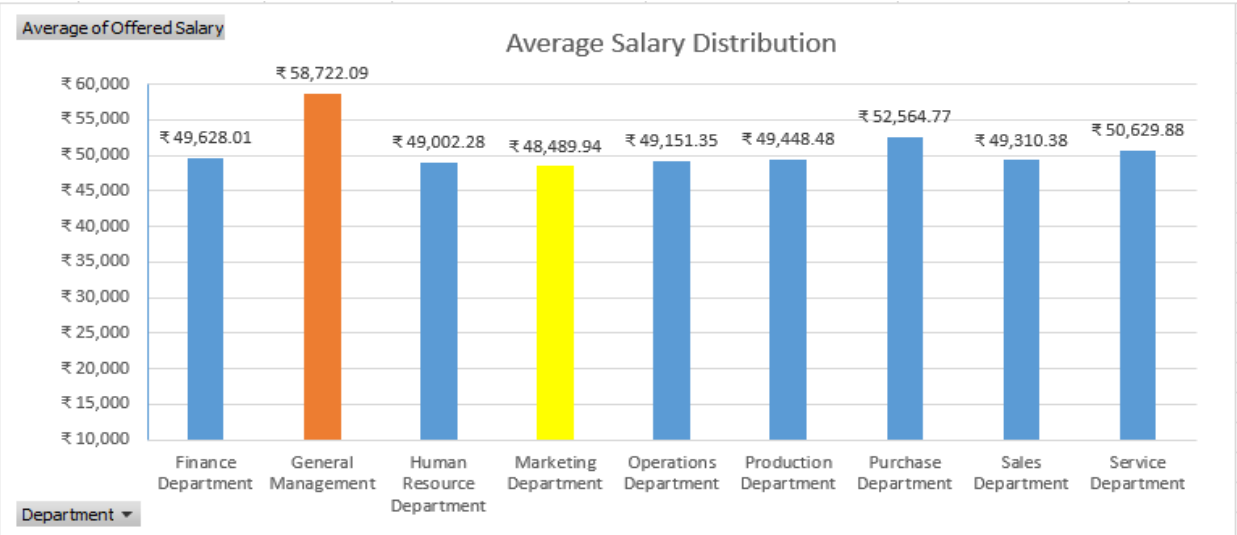
Analysis –

Average salary offered is calculated using the formula below.

$$Average = \frac{Sum\ of\ individual\ salaries}{Total\ number\ of\ applicants} \quad = ₹\ 49{,}983.$$

The figure below shows the average salary received by each department. It can be seen that "General Management" department receives maximum salary while "Marketing Department" receives the least.

Gender	Average of Offered Salary
Finance Department	₹ 49,628.01
General Management	₹ 58,722.09
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	₹ 49,310.38
Service Department	₹ 50,629.88
Grand Total	**₹ 49,983.03**
Maximum Average Salary Offered	**₹ 58,722.09**
Minimum Average Salary Offered	**₹ 48,489.94**



Gender	Average of Offered Salary	Max of Offered Salary	Min of Offered Salary
Finance Department	₹ 49,628.01	₹ 99,762.00	₹ 1,038.00
General Management	₹ 58,722.09	₹ 4,00,000.00	₹ 1,022.00
Human Resource Department	₹ 49,002.28	₹ 98,195.00	₹ 1,415.00
Marketing Department	₹ 48,489.94	₹ 99,828.00	₹ 1,007.00
Operations Department	₹ 49,151.35	₹ 99,948.00	₹ 1,027.00
Production Department	₹ 49,448.48	₹ 99,939.00	₹ 1,210.00
Purchase Department	₹ 52,564.77	₹ 99,522.00	₹ 1,258.00
Sales Department	₹ 49,310.38	₹ 99,824.00	₹ 1,487.00
Service Department	₹ 50,629.88	₹ 2,00,000.00	₹ 100.00

The figure above shows the maximum and minimum salary offered as per department. General Management department receives a highest of ₹ 4 lakh and the least being ₹ 100 offered in service department.

The salary range depends on factors like –

1. Education Qualifications and Relevant industry level certifications done by the applicant.
2. Experience and skill set, Performance and productivity in case of promotion for an already existing employee.
3. Negotiation skills.
4. The company size and industry, job role and responsibility, and financial performance of the company along with certain external factors like economic conditions and market demand also affect the salary.

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.

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

Approach –

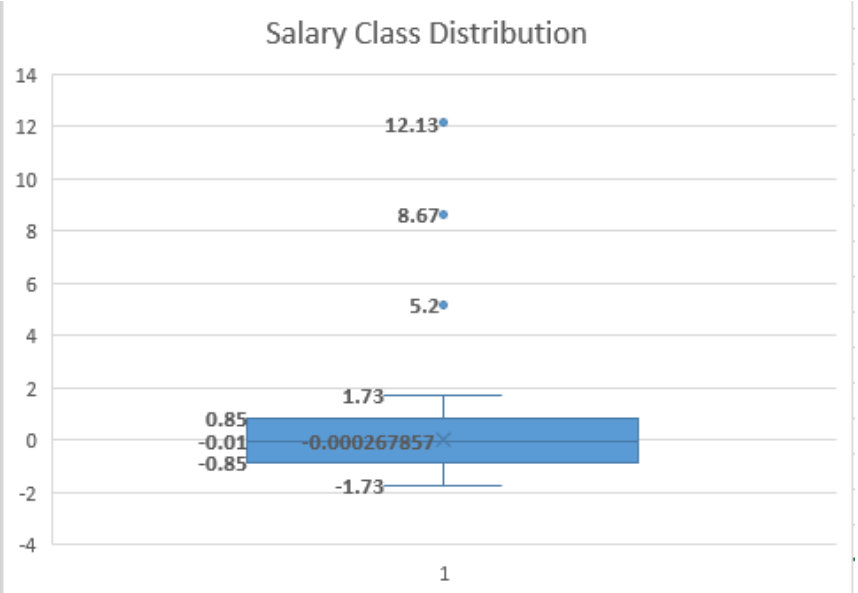
Excel's statistical function is used here to calculate average. The pivot table is used to understand the salary distribution against various departments. For the ease of understanding we have created an additional column – standardised representing standardised data.

Analysis –

From the "salary offered" field analysis it can be seen that the salaries lie in the range – ₹ 50,000, ₹ 1,00,000, ₹ 2,00,000, ₹ 3,00,000, ₹ 4,00,000. The salary lies with a minimum of ₹ 100 to a maximum of ₹ 400000.

For further understanding of the data, we have used box plot here. The box plot provides us a visual representation of the data. It provides a clear and concise summary of data,

highlighting key features like central tendency, spread, and potential outliers. The box plot representation is below –



Findings –

1. Central Tendency – the median is approximately 0, indicating half the salary lies below the median and other half above it.
2. Data Spread – IQR is +- 0.85 indicating most of the salaries are clustered around the median. The whiskers indicate the spread of data beyond the quartiles. 1.73 whiskers indicate most of the salary class lie within this range.
3. Outliers – we have 3 outliers above upper whisker. These points correspond to salary classes significantly higher than the most. A outlier with point 12.3 corresponds to a salary of 4 lakh. 8.67 point have salary range of 3 lakh and 5.2 having 2 lakh range.
4. The slight right skew suggests that there might be a few higher-paying positions or a few individuals with exceptionally high salaries.

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

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

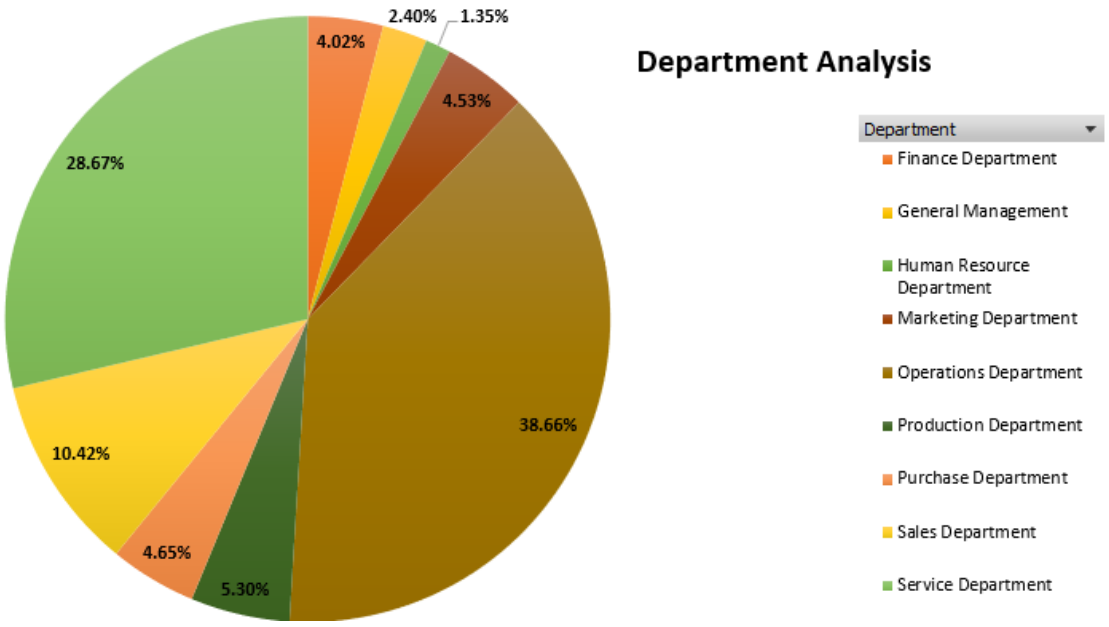
Approach –

The pivot table is used to understand the hiring among various departments.

Analysis –

Operations Department have maximum hiring with 38.66% and HR Department have least recruitment – 1.35%.

Departments	Applicants %
Finance Department	4.02%
General Management	2.40%
Human Resource Department	1.35%
Marketing Department	4.53%
Operations Department	38.66%
Production Department	5.30%
Purchase Department	4.65%
Sales Department	10.42%
Service Department	28.67%
Grand Total	**100.00%**



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

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.

Approach –

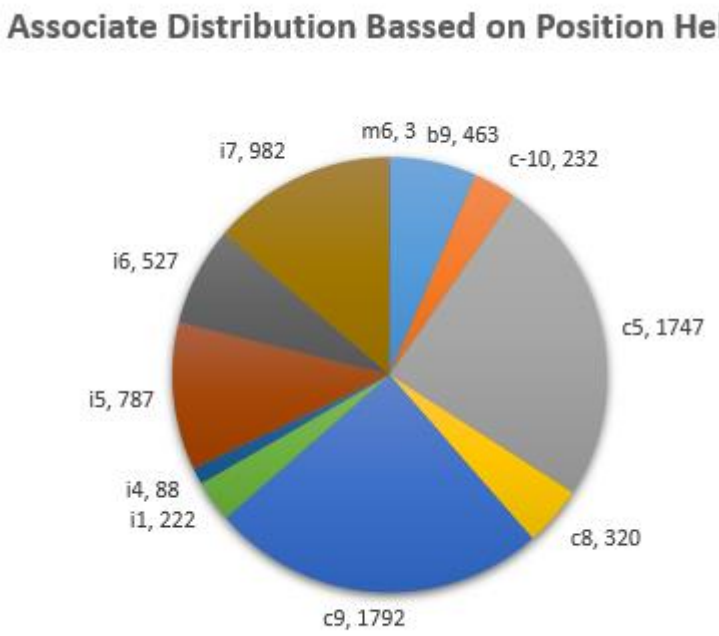
We are using Excel's Pivot chart and Line graph for analysing. We are plotting various posts open for hiring against salary class.

Analysis –

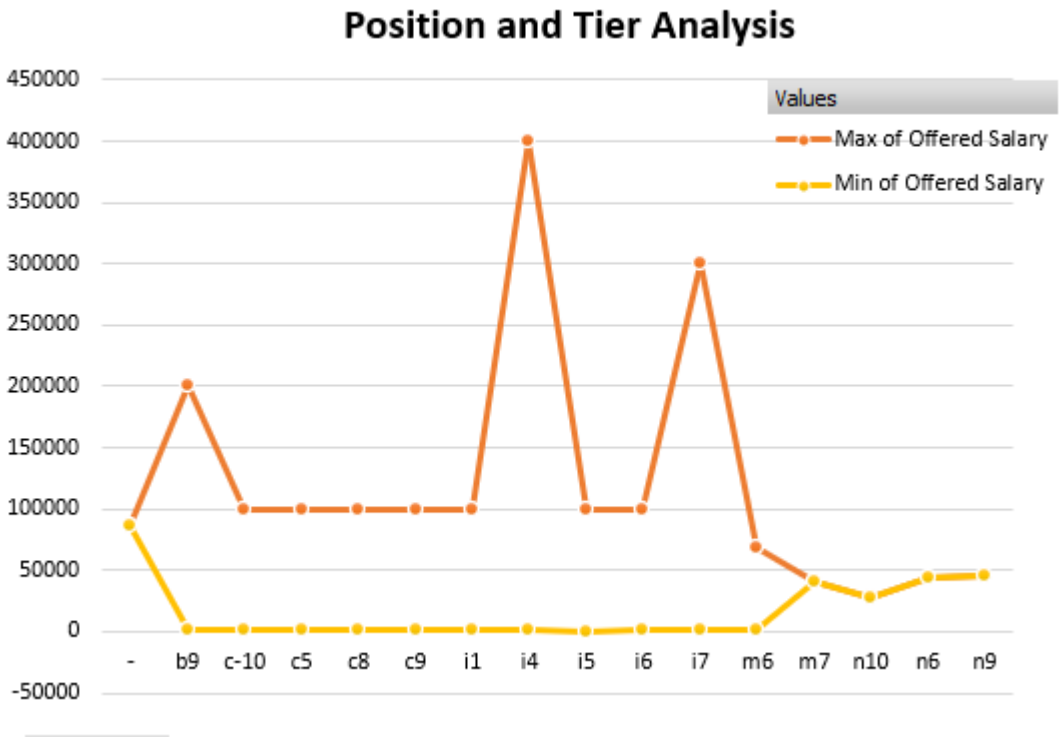
From the pie chart it can be seen that Position c9, c5, i7, i5 have most of the employee distribution.

From the line plot it can be seen that almost all of the positions offer same initial salary – maybe for new recruits with no experience. But as the experience goes on it can be seen

that the maximum salary offered varies majorly. Positions b9, i4, i7 – offers the highest salary.



Departments	Max of Offered Salary	Min of Offered Salary
-	85914	85914
b9	200000	1105
c-10	99891	1817
c5	99948	1038
c8	99967	1035
c9	99953	1007
i1	99939	1519
i4	400000	1212
i5	98926	100
i6	99762	1074
i7	300000	1022
m6	68466	800
m7	41402	41402
n10	26990	26990
n6	44700	44700
n9	46219	46219



4.Tech Stack Used –

I have used Microsoft Excel to perform this analysis. It is used to perform -

1. Collect and Organize Data: Create spreadsheets to track key metrics like time-to-hire, cost-per-hire, and source effectiveness.
2. Calculate Key Metrics: Use Excel functions to calculate metrics and identify trends.
3. Visualize Data: Create charts and graphs to understand data distribution and trends.
4. Identify Bottlenecks: Analyze data to pinpoint stages where the hiring process slows down.
5. Make Data-Driven Decisions: Use insights from the analysis to improve the hiring process.

Link to Dataset –

https://docs.google.com/spreadsheets/d/1yfwPaQA3Aa6JwOEsH6IoiWUbOZ4suszo/edit?usp=sharing&ouid=104845356380163502534&rtpof=true&sd=true