Hiring Process Analytics

Google link :-

https://docs.google.com/spreadsheets/d/1rjeqCy44Avy46nj-H-2Mz764GGteQWt2/edit#gid=1522302513

Description:

The recruitment process stands as a cornerstone and paramount function within any company. It serves as a window through which multinational corporations (MNCs) can discern crucial underlying patterns. These patterns encompass various aspects such as the frequency of rejections, number of interviews conducted, job categories, and available vacancies. It's imperative for a company to meticulously scrutinize these trends before engaging in the hiring of fresh talent or any prospective candidate. Consequently, this scenario presents an enticing opportunity for a Data Analyst role.

In the capacity of a Data Analyst, the primary responsibility entails delving into these trends and extrapolating insightful conclusions. These insights are then furnished to the recruitment department, empowering them to refine and optimize their strategies.

Assuming the role of a lead Data Analyst at a prominent MNC like Google, the company has entrusted you with the task of dissecting the dataset encompassing their previous recruitment endeavours. Your objective is to elucidate meaningful patterns and draw informed conclusions about the efficacy of the company's recruitment process. This necessitates employing statistical analysis techniques, with Excel.

Process:-

The process approach utilized in analysing the given dataset entails several key steps:

1. **Data Cleaning**: The initial step involves cleaning the data to eliminate any duplicate or irrelevant entries. Ensuring proper formatting of the data is essential at this stage.
    - Column event_name has 15 rows with "-" as its values, which is replaced by "Don't want to say.
    - Row 80 of Col "Salary" had no salary mentioned. I Took the median of the salaries available and mentioned it in the missing cell. Median of Salary before adding to the blank column is 49570.
    - Column Post Name has 1 row with "-" as its value. The department is Sales department. So as per the reading the post name given to the cell is c9.

2. **Outlier Identification**: Next, the dataset is scrutinized to identify any outliers that could significantly influence the analysis. These outliers are then removed to ensure the accuracy of the results.
    - Found 3 Outliers. I have replaced them with the average value 49570

application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary	Outliers
649039	07-05-2014 10:48	Hired	Female	Service Department	b9	200000	Outlier Found
795330	15-06-2014 09:45	Hired	Female	General Management	i4	400000	Outlier Found
874368	21-07-2014 15:39	Hired	Male	General Management	i7	300000	Outlier Found

3. **Duplicate Rows**: - Application Id is unique for all the candidates. So we found 27 rows having similar Duplicate Application Ids. And removed the duplicate rows using "Remove Duplicates". 27 Duplicate values found and removed. 7141 rows remaining

4. **Descriptive Statistics**: - Relevant descriptive statistical calculations are performed to obtain a comprehensive overview of the dataset. This step aids in understanding the distribution and characteristics of the data.

5. **Insight Extraction**: - Following the descriptive analysis, the focus shifts to extracting useful insights from the data using statistical techniques. This involves identifying patterns, trends, and relationships within the dataset.

6. **Data Visualization**: - Finally, the insights obtained from the statistical analysis are visualized using data visualization methods. Visual representations such as charts, graphs, and plots are employed to present the findings in a clear and concise manner.

   By following this systematic approach, meaningful insights are derived from the dataset, facilitating informed decision-making and problem-solving.

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?

Male Hired	2563
Females Hired	1856
Don't want to say	278

Insights:-

It is observed that most of the hired candidates hired are males. It is suggested to balance to diversity as the observation can affect the organization negatively.
We should try to maintain the GENDER RATIO.

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.

Median of Salary before adding to the blank column		49625
Average of Salary of all departments together	49982.9714	
Median of Salary after filling the blank col	49625	
Minimum Salary	100	
Maximum Salary	400000	

Average Salary Department wise

	Salary
Finance Department	49628.00694
General Management	58722.09302
Human Resource Department	49002.27835
Marketing Department	48489.93538
Operations Department	49151.35438
Production Department	49448.48421
Purchase Department	52564.77477
Sales Department	49310.72825
Service Department	50629.88418

Insights:-

- Average salary offered is 49,987
- Average salary offered to the hired candidates is 49,753
- **The Average Salary of Hired Candidates is almost same as that of Offered Salary. This shows that the hiring team is recruiting candidates as per the pre-determined salary bands of the organization.**

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.

Salary Range	Salary offered
100-10099	686
10100-20099	728
20100-30099	711
30100-40099	713
40100-50099	777
50100-60099	754
60100-70099	698
70100-80099	733
80100-90099	716
90100-100099	649
190100-200099	1
290100-300099	1
390100-400000	1
Total	7168

Salary Range	Salary offered and hired
100-10099	444
10100-20099	487
20100-30099	457
30100-40099	488
40100-50099	523
50100-60099	496
60100-70099	450
70100-80099	479
80100-90099	462
90100-100099	408
190100-200099	1
290100-300099	1
390100-400000	1
Total	4697

Insights:-

We see that most of 777 number of people fall into the salary range of 40100-50099 and 523 candidates are hired having salary between 40100-50099.

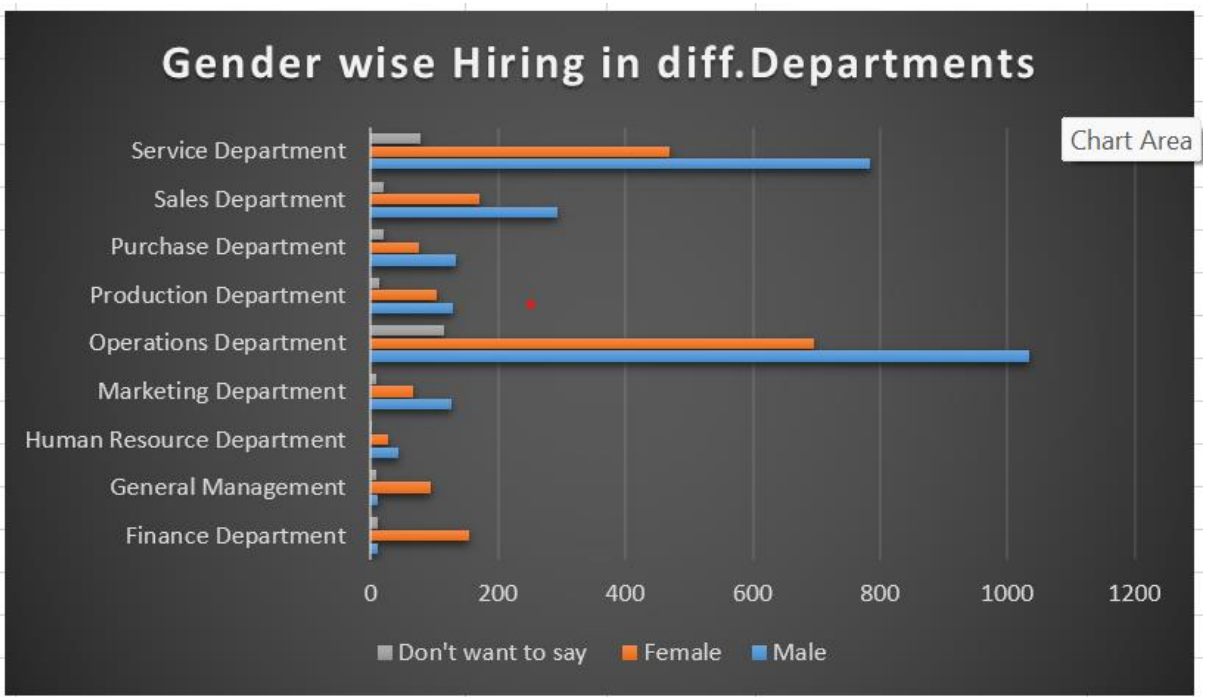
D. **Departmental Analysis:**

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.

Distribution as per the department

Department	Male	Female	Don't want to say	Total
Finance Department	10	154	12	**176**
General Management	10	95	8	**113**
Human Resource Department	43	26	1	**70**
Marketing Department	127	66	9	**202**
Operations Department	1033	695	115	**1843**
Production Department	128	104	14	**246**
Purchase Department	133	76	21	**230**
Sales Department	294	171	20	**485**
Service Department	785	469	78	**1332**
Total	**2563**	**1856**	**278**	**4697**



Insights:-

From the above pie chart, we can observe that most candidates are hired in Operations Department followed by Services Department and Sales Department and the least candidates are hired in Human Resource Department. We see that Female candidates are majorly hired in Finance and General Department.

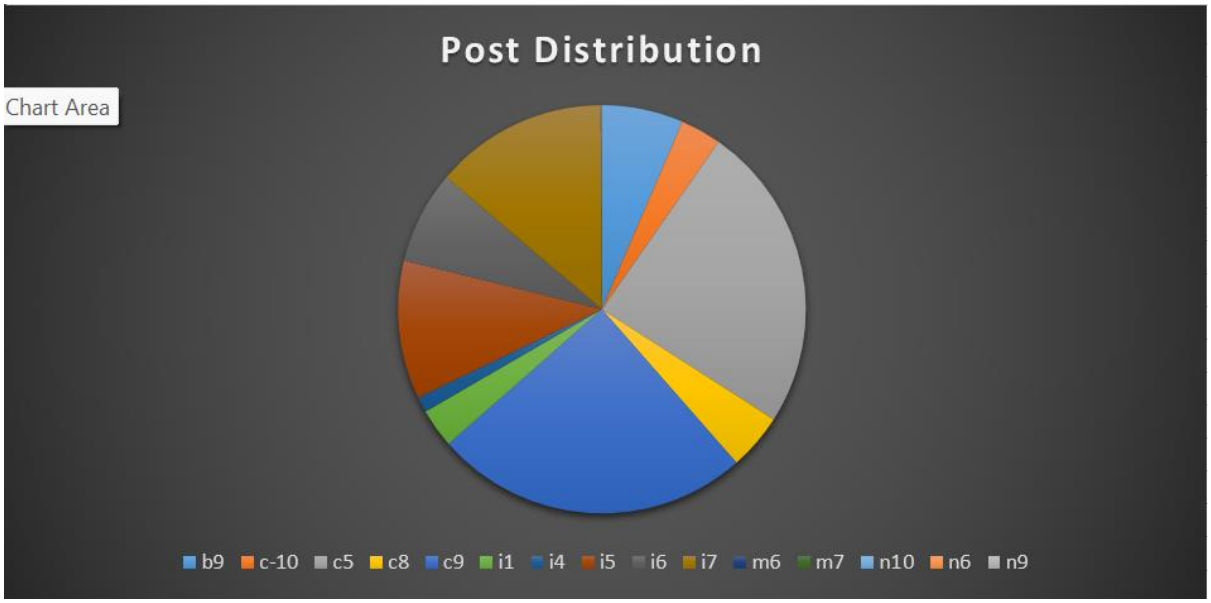
E. **Position Tier Analysis:**

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.

Department/Post Name	b9	c-10	c5	c8	c9	i1	i4	i5	i6	i7	m6	m7	n10	n6	n9
Finance Department	13	4	68	4	107	9	3	41	12	27	0	0	0	0	0
General Management	2	10	29	7	39	1	1	31	9	43	0	0	0	0	0
Human Resource Department	2	2	21	6	7	2	0	42	6	9	0	0	0	0	0
Marketing Department	28	18	74	26	70	13	1	30	15	50	0	0	0	0	0
Operations Department	158	99	671	98	711	94	38	272	278	351	1	0	0	0	0
Production Department	40	8	79	8	87	28	3	37	26	64	0	0	0	0	0
Purchase Department	22	5	107	4	74	2	3	36	23	55	0	0	0	1	1
Sales Department	28	23	216	48	176	2	10	88	43	113	0	0	0	0	0
Service Department	170	63	482	119	522	71	29	210	115	270	2	1	1	0	0
Total	**463**	**232**	**1747**	**320**	**1793**	**222**	**88**	**787**	**527**	**982**	**3**	**1**	**1**	**1**	**1**

Post Name	
b9	463
c-10	232
c5	1747
c8	320
c9	1793
i1	222
i4	88
i5	787
i6	527
i7	982
m6	3
m7	1
n10	1
n6	1
n9	1
Total	**7168**



Insight:-

Here, we can observe that the organization has hired most candidates for post tier c9 followed by c5 and then i7.

Conclusion:-

- I have completed the analysis of the provided dataset in accordance with the questions posed, offering the necessary insights and creating relevant charts and graphs as per the requirements and my interpretation. This project has been instrumental in enhancing my understanding of the Exploratory Data Analysis (EDA) process.

- Exploratory Data Analysis (EDA) involves comprehensively examining and analyzing datasets to uncover insights and facilitate informed decision-making. Through this project, I conducted EDA on a human resource dataset, which significantly enriched my grasp of the EDA process. EDA can be conducted using various tools such as Excel, Google Analytics, or WPS Spreadsheet, among others.

- In conclusion, this project has underscored the critical role of Data Analytics in the hiring process of an organization. By providing valuable insights such as the number of rejections, reasons for rejections, applicant profiles, and vacancies, Data Analytics empowers the hiring department to make data-driven decisions, ultimately enhancing the efficiency and effectiveness of the hiring process.

----------**********----------