HIRING PROCESS ANALYSIS

AN EDA PROJECT ON STATISTICS

BY

DEBARATI CHATTERJEE

Project Description:

- Hiring process is the fundamental and the most important function of a company. Here, the MNCs get to know about the major underlying trends about the hiring process. Trends such as- number of rejections, number of interviews, types of jobs, vacancies etc. are important for a company to analyse before hiring freshers or any other individual. Thus, making an opportunity for a Data Analyst job here too!
- Being a Data Analyst, my job is to go through these trends and draw insights out of it for hiring department to work upon.
- I am working for a MNC such as Google as a lead Data Analyst and the company has provided with the data records of their previous hirings and have asked me to answer certain questions making sense out of that data. I need provide a detailed report for the below data record mentioning the answers of the questions that follows:

Approach:

To complete the project, I have taken a certain approach.
- I spent some time familiarizing myself with the data before commencing the analysis. Examine the data structure to obtain a feel of the overall content. This allows me to identify any potential concerns or obstacles that I may encounter while I do my analysis.
- I check for any null values or missing data in the dataset, and replace with suitable option.
- Outliers have a significant impact on summary statistics and distort your analysis's results. Outliers must be identified and handled appropriately, so I analyse dataset to observe outliers and remove them to achieve proper findings.
- When I finish my exploratory data analysis , I present my insights to the audience in a clear and concise manner. I utilize visuals like charts and graphs to help me express my findings.

Teck-Stack Used:

- **MS Excel:** To clean the data, analysis and summarise given dataset and find valuable insights for the company.
- **MS Word:** To present the report in structured manner to the authority.

Removing Outliers:

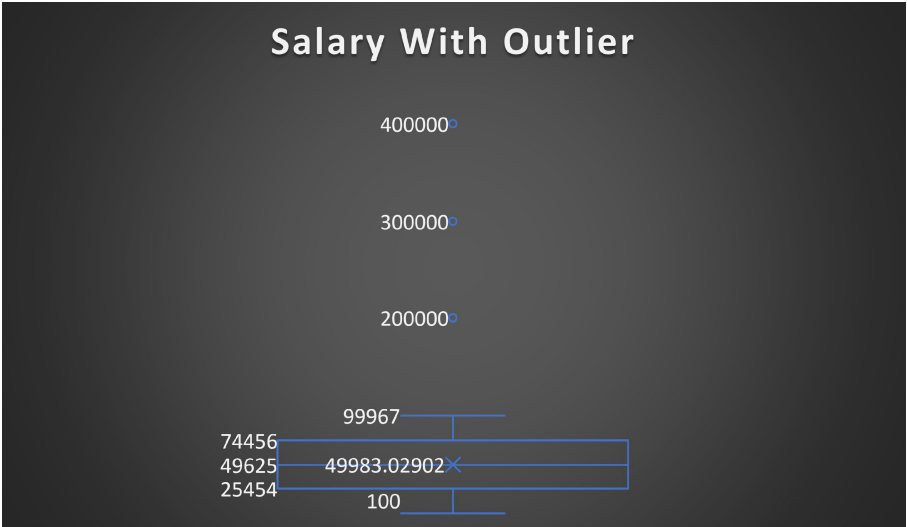In this Data set after handling null value, we observe certain outliers in Salary Column.

Fig-1 Salary with Outliers

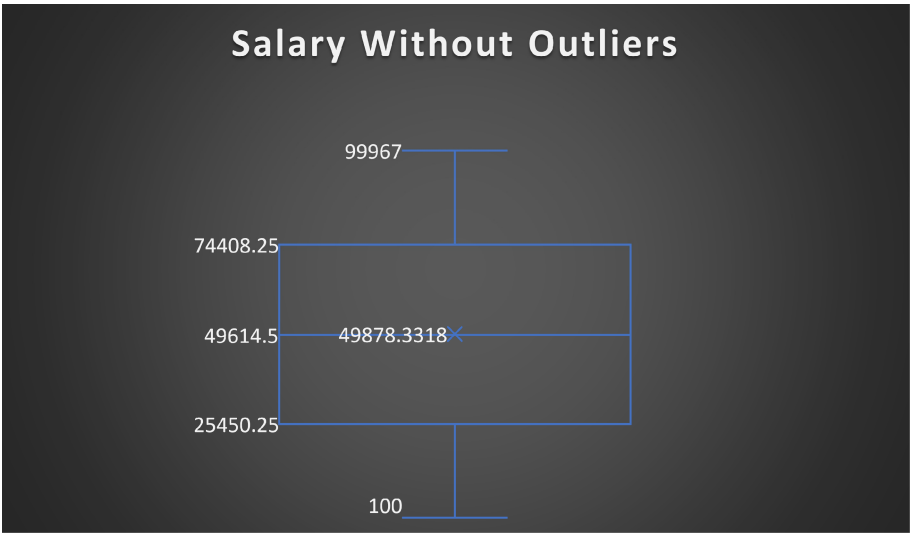After removing outliers, we are able to analysis the data properly.

Fig-2 Salary without Outliers

Final Table:

application_id	Interview Taken	Status	event_nar	Department	Post Nam	Offered Salary
863024	01-08-2014	Hired	Male	Service Department	c8	99967
913059	11-08-2014	Hired	Female	Service Department	c9	99953
124098	23-06-2014	Hired	Male	Service Department	c9	99950
690427	14-06-2014	Hired	Male	Operations Department	c5	99948
879713	23-05-2014	Rejected	Male	Production Department	i1	99939
222116	24-07-2014	Rejected	Male	Operations Department	i7	99929
137124	16-08-2014	Hired	Male	Operations Department	c9	99920
302732	07-08-2014	Hired	Male	Production Department	c10	99891
768225	07-08-2014	Hired	Male	Service Department	c9	99880
915022	04-07-2014	Hired	Female	Operations Department	c9	99852
911449	28-07-2014	Rejected	Male	Operations Department	c5	99841
932562	13-08-2014	Rejected	Male	Marketing Department	c5	99828
322454	15-08-2014	Hired	Male	Sales Department	c5	99824
841824	06-05-2014	Hired	Female	Production Department	c9	99800
819973	06-05-2014	Hired	Male	Service Department	c8	99766
496401	10-07-2014	Rejected	Female	Finance Department	i6	99762
562353	01-05-2014	Hired	Female	General Management	i7	99745
564902	24-07-2014	Rejected	Female	Sales Department	c5	99736
230925	18-07-2014	Rejected	Female	Production Department	i1	99722
234992	30-05-2014	Hired	Male	Service Department	b9	99645

Summarise Data Table:

	Column Labels									Total No of applicant	Total Salary
	⊞May		⊞Jun		⊞Jul		⊞Aug				
Department	No of applicant	Salary	No of applicant	Salary	No of applicant	Salary	No of applicant	Salary			
⊞Finance Department	73	47394.07	59	54379.73	100	48585.35	56	49395.71		288	49628.01
⊞General Management	67	56591.91	39	66665.23	45	58599.64	21	51029.24		172	58722.09
⊞Human Resource Department	12	67126.58	25	46991.60	20	53052.65	40	42796.48		97	49002.28
⊞Marketing Department	67	46863.52	87	51326.90	81	46371.62	90	48864.80		325	48489.94
⊞Operations Department	682	49553.53	689	51660.42	740	47119.96	660	48394.10		2771	49151.35
⊞Production Department	80	53143.35	82	43403.88	119	46837.86	99	54607.40		380	49448.48
⊞Purchase Department	76	53992.24	78	52335.56	104	53822.98	75	49611.95		333	52564.77
⊞Sales Department	148	51271.89	144	47409.75	210	49767.48	245	48649.51		747	49244.37
⊞Service Department	541	51586.44	437	50990.36	527	50437.75	550	49586.66		2055	50629.88
Grand Total	1746	50884.13	1640	51093.71	1946	49015.73	1836	49132.02		7168	49976.06

Insights:

A. **Hiring:** Process of intaking of people into an organization for different kinds of positions.
   Task: How many males and females are Hired?

Analysis: Among 7165 total applicant 4416 hired. It is observed that 1854 are female and 2562 are male among them.

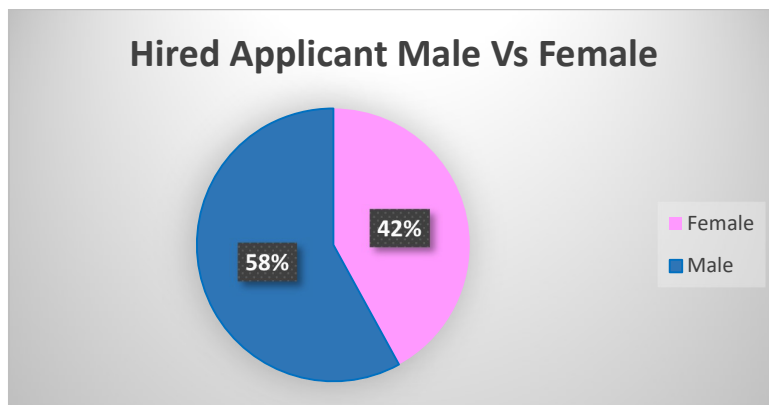| Status	Hired
Gender	**No of Applicant**
Female	1854
Male	2562
Grand Total	**4416**

Visualization:

Fig-3 Male vs Female ratio in hired applicant

B. **Average Salary:** Adding all the salaries for a select group of employees and then dividing the sum by the number of employees in the group.
   Task: What is the average salary offered in this company?

Analysis: Average salary offered by the company is 49878 where median of salary is 49614.5 (in currency).

Average	Median
49878.33	49614.5

C. **Class Intervals:** The class interval is the difference between the upper class limit and the lower class limit.

Task: Draw the class intervals for salary in the company?

Analysis: Offered salary can be divided into 10 class with 10000 class size where Min salary is 0 and max salary is 100000

Salary Class	No of Applicant
0-9999	678
10000-19999	732
20000-29999	711
30000-39999	709
40000-49999	781
50000-59999	751
60000-69999	698
70000-79999	734
80000-89999	711
90000-100000	659
Grand Total	**7164**

Visualization:

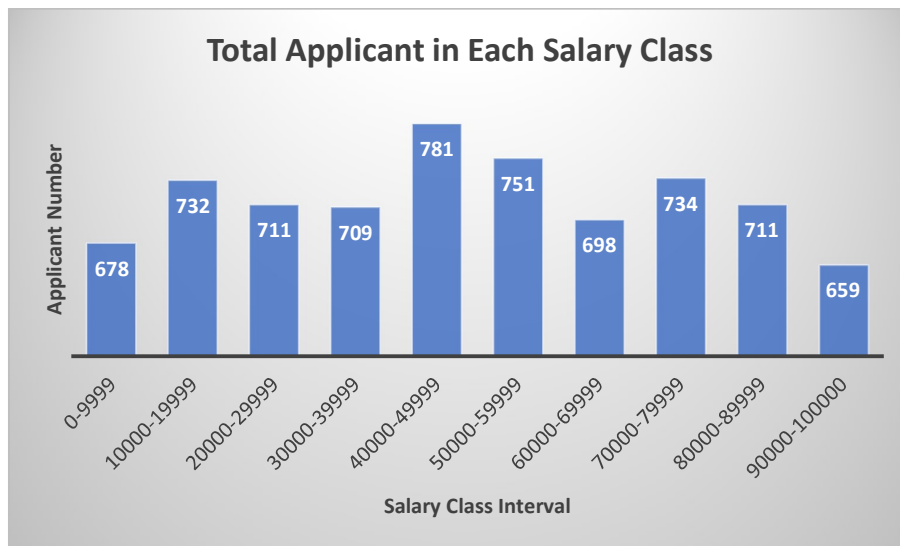Fig-4 Frequency distribution in each salary class

D. **Charts and Plots:** This is one of the most important part of analysis to visualize the data.
Task: Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working different department?

Analysis: We can observe that maximum employees hired and is working in Operations Department, which is 39% of total employees.

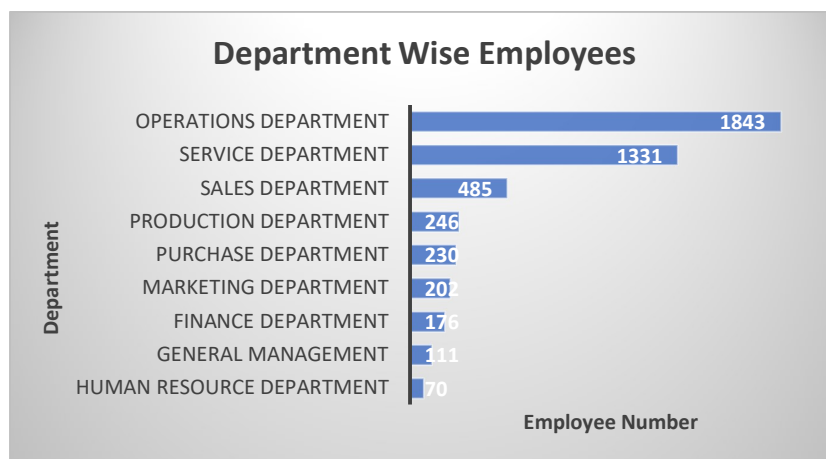| Status	Hired
Departments	**No of Employees**
Operations Department	1843
Service Department	1331
Sales Department	485
Production Department	246
Purchase Department	230
Marketing Department	202
Finance Department	176
General Management	111
Human Resource Department	70
Grand Total	**4694**

Visualization:

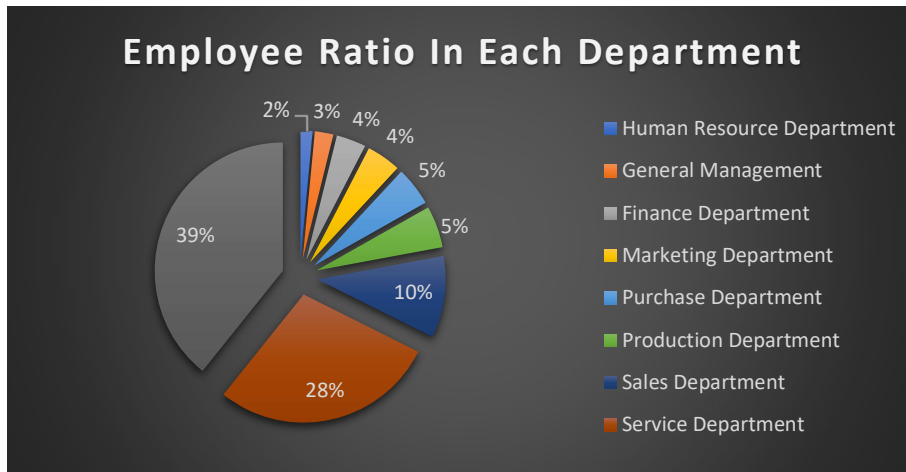Fig-5 Number of Hired Employee in each department

Fig-6 Ratio of employees in each department

E. Charts: Use different charts and graphs to perform the task representing the data.
Task: Represent different post tiers using chart/graph?

Analysis: It is founded that c9 and c5 are the most applied post from the post tier.

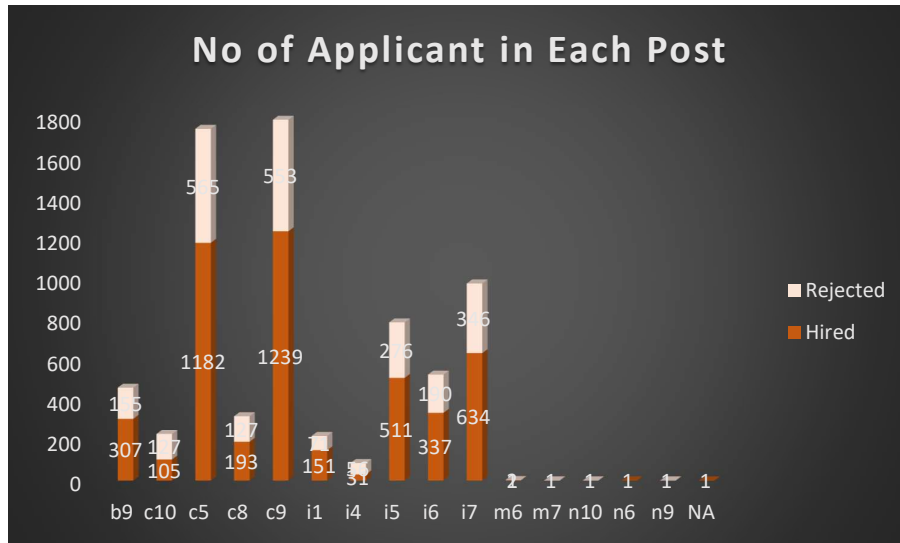| Post Name	Hired	Rejected	Grand Total
b9	307	155	462
c10	105	127	232
c5	1182	565	1747
c8	193	127	320
c9	1239	553	1792
i1	151	71	222
i4	31	56	87
i5	511	276	787
i6	337	190	527
i7	634	346	980
m6	2	1	3
m7		1	1
n10		1	1
n6	1		1
n9		1	1
NA	1		1
Grand Total	**4694**	**2470**	**7164**

Visualization:

Fig-7 Total Applicant in Each post

Result:

- Among 4416 hired applicant 42% are female and 58% are male.
- Average salary offered by the company is 49878 (in currency)
- Majority of the applicants are offered salary in the class 40k-49k
- 39% of employees are under Operation Department.
- Maximum (1792) people applied for c9 post where 1239 are hired among them.

Conclusion:

This project has helped me to learn data cleaning, data manipulation and further data visualization on real life data set. I get to experience Exploratory Data Analysis which will benefit me to create error free data model in future. I get to level up my excel skill in this process. Further I have learned about hiring process of any company.

THANK YOU