Hiring Process Analytics

Project description

The dataset contains records of candidates who were interviewed previously with information about hiring status, hiring department, salary etc.

The Dataset details are:

- Number of Data-Points: 7,168

- Number of Features: 6

- Column Details:

1. **application_id:** ID of the applicant

2. **Interview Taken on:** Date and time of the interview

3. **Status:** Hired or rejected

4. **event_name:** Gender of the applicant

5. **Department:** Name of the department for which interview was conducted

5. **Post Name:** Name of the post offered

6. **Offered Salary:** Salary offered for the job

Tools used - MS excel 2021 Because it is quite simple to use Quite effective on a small data

Data cleaning

1. Removed - and replaced it with **Don’t want to say**
2. Null Value in offered salary replaced it with median of employee in Sales Department with =MEDIAN(FILTER(G:G, (F:F="I7") * (E:E="Sales Department")))
3. Replaced c-10 with c10 in Post Name
4. For one blank cell in post name

Created a new column bin using formula =FLOOR(G2, 10000) and dragged it down till the end then used a filter to get the post name with the same department and bin.

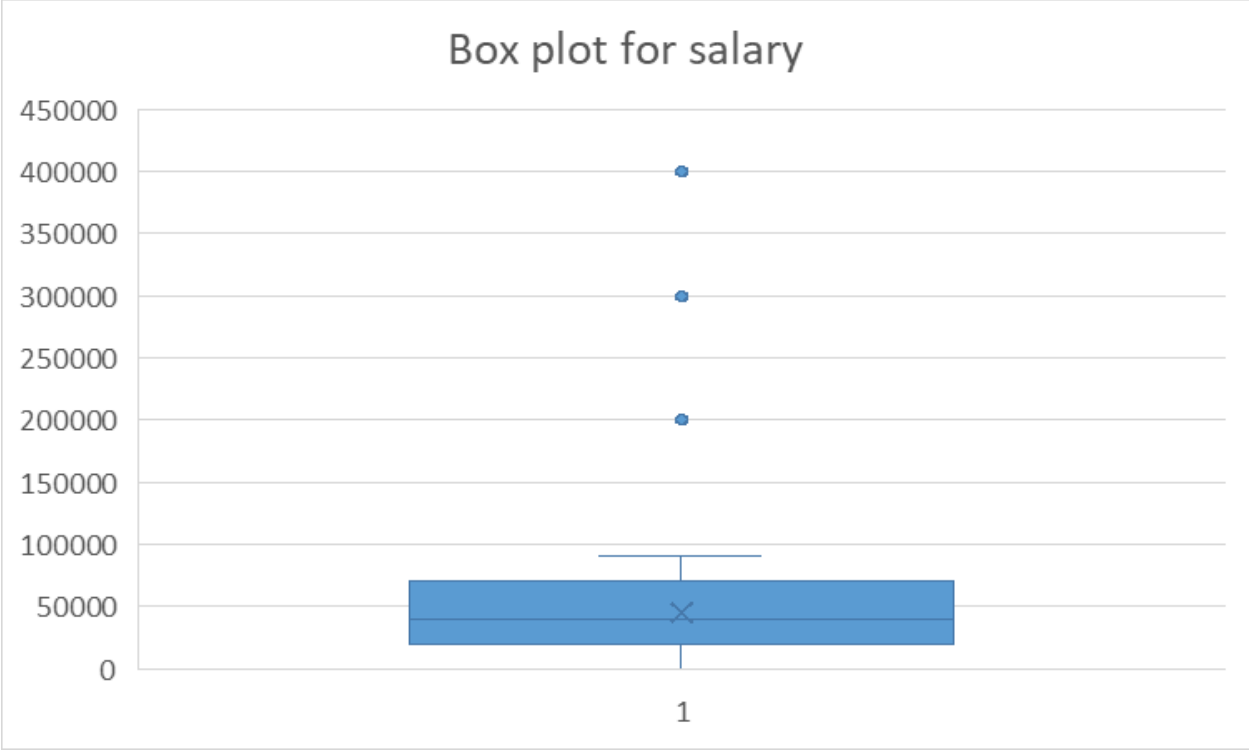
=FILTER(F:F,(H:H=H7)*(E:E=E7)) and used formula =COUNTIF(M21:M92,M21) to find it. “C5”

Handle duplicates

Deleted 54 rows used countif(a:a,a2) and filter

Handle Outliers

Used boxplot to find outliers.



Here we can see three outliers.

Replaced them with the median salary of employees in the same department and post name.

Use this to filter data ==FILTER(G:G,(F:F=F13)*(E:E=E13)) where g:g is offered salary

F:f is for comparing the post name of the outlier and e:e is to compare the department .

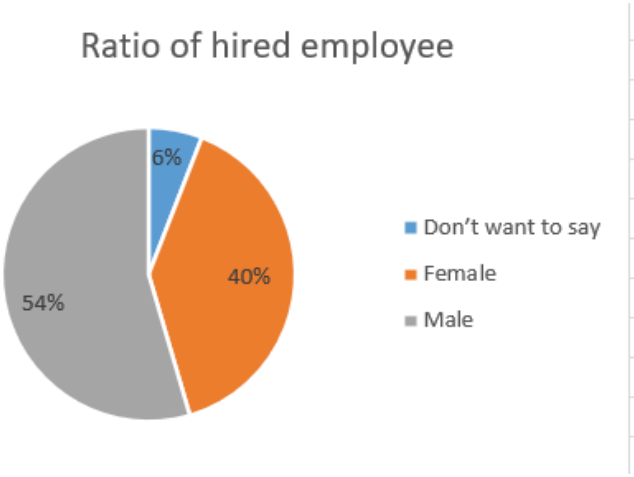
After this I used the median function on the filter output.

For 200000 and 300000 many values where available for same post and same department. 400000 was unique if both are taken in filter so i only took post name in filter .

Insights

1. Find or represent female to male ratio of hired employees.

Status	Hired
Row Labels	**Count of event_name**
Don't want to say	276
Female	1844
Male	2543
Grand Total	**4663**



Insights-

The ratio of male to female is 54 is 40% Ratio of male Is high as compared to females so it might affect company if the company is publicly traded and the hr Should look into this also the data team should look into those people who don't want to say. Overall the ratio is okay but not bad and there is quite room for improvement.

2.**What is the average salary offered by this company?**

Average salary offered	49890.01
Average salary of hired employee	49592.54

Used average and average if function to find the average of the salary offered to employees and average of the salaries of the employees who are hired.

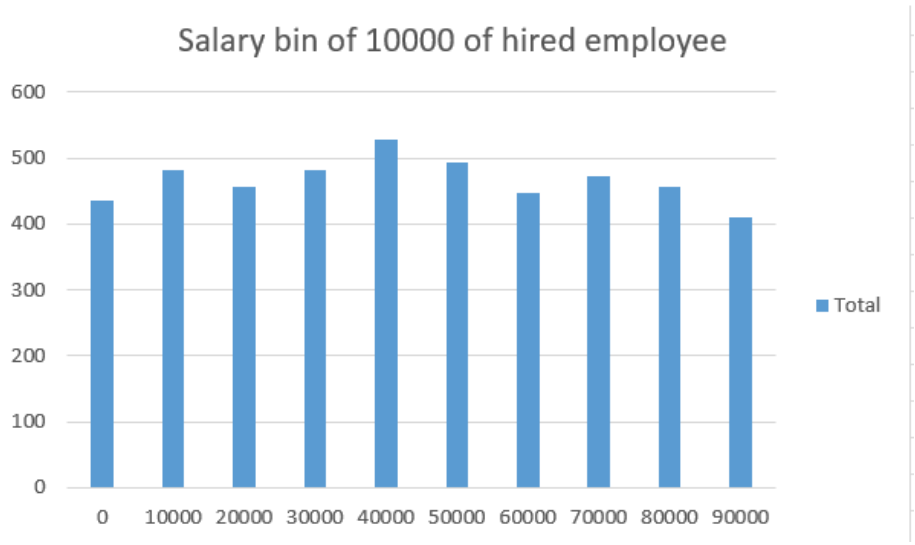
Insight-

Both the average are quite similar it means That hiring department are hiring on the basis of predetermined skill sets that are required and are also offering salary based on predetermined values that means over paying and under paying is not there on the basis of candidates they are trying to look for the right candidates and paying them accordingly

3.**Create class intervals for the salaries in the company.**

Status	Hired

Row Labels	Count of bin
0	435
10000	482
20000	456
30000	482
40000	527
50000	494
60000	448
70000	473
80000	456
90000	410
Grand Total	**4663**

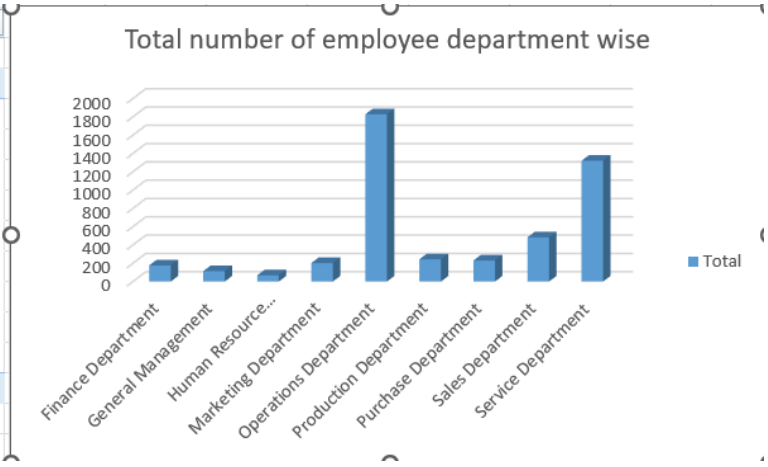


insights- So in this chart zero means zero to 10,000 similarly 10,000 means 10,000 to 20,000 so it tells me how many people are in each bracket or a bin so as we can see that it is a fairly event distribution throughout so it is a good sign the second thing we should look for is the outliers which removed in the beginning if they are true the three people are earning quite significantly more as compared to others as their salary were two lakh three lakh and 4,00,000.

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

Status	Hired

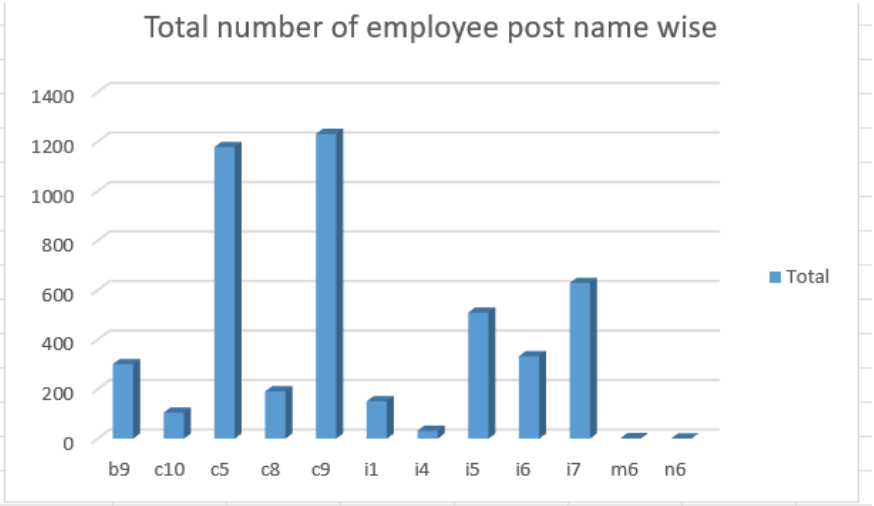
Row Labels	Count of employee
Finance Department	176
General Management	113
Human Resource Department	68
Marketing Department	202
Operations Department	1828
Production Department	243
Purchase Department	230
Sales Department	483
Service Department	1320
Grand Total	**4663**



Insights- We can see majority of the people are working in either operations or service department it tells us that this company is more focused on operational or service that is iit waste company or a service based company This also tells us that where the company want to head in the future and will hire more people from operations and service department in future as company is based on these skills only.

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**.

Status	Hired
Row Labels	**Count of employee**
b9	301
c10	105
c5	1178
c8	191
c9	1231
i1	151
i4	32
i5	509
i6	333
i7	629
m6	2
n6	1
Grand Total	**4663**



Insights-

Most number of employees are from C 5 and C9 followed by i5 and I 7 so we can know that C5 and C7 are common and as we move forward M6 and N6 are for the top management

Results and Learning

Through this project, I gained hands-on experience in applying essential data analysis techniques using Excel. I deepened my understanding of handling missing data, detecting and managing outliers, and organizing data through column grouping for clearer insights. I also strengthened my skills in using PivotTables, formulas like `AVERAGE`, `AVERAGEIF`, and `MEDIAN`, and in visualizing data through charts and graphs. Most importantly, I learned how to interpret hiring-related trends and patterns—such as gender distribution, salary ranges, departmental strengths, and job tier structures—which are critical for making data-driven decisions in recruitment and workforce planning.