Hiring Process Analytics

PROJECT DESCRIPTION

The task requirement is to analyze the dataset given which is in Excel document. The columns are: *application_id,Interview Taken on,Status,event_name,Department,Post Name and Offered Salary.* 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. Data gives a glimpse about the past hiring activity and using this data, few questions need to be answered so as to streamline the future hiring processes.

Project file link :
https://docs.google.com/spreadsheets/d/1Cs3YETByhkPV9kEHej2OepfQkUUJaz_W/edit?usp=sharing&ouid=111360395057791388166&rtpof=true&sd=true

Approach

I will go through the dataset first and quickly see if there are any duplicates by checking the *application_id* column. I have used-*Remove duplicates* option to do the same. Then, I check for Null values using the filter option. I came across null values in the columns-*event_name(gender), Department and Offered Salary*.

I replaced the null values in gender column with *"Don't want to say",* removed the column with null value in Department as it was just one row and filled the null value row with *median* salary in the Offered Salary column so that the outlier values do not affect the filled value.

Tech-Stack Used

Microsoft Excel 2019 version

Analysis Process

Outlier Detection

I used the column- *Offered Salary* to plot a box plot to check if any outlier exists or not. Then I sorted the column in descending order.

As one can see, 3 outlier values exist. On further examination, it was found that these salaries belong to departments such as General Management. It might be possible that these salaries were offered to top level executives. Hence, I decided to retain these values.

Department	Post Name	Offered Salary
General Management	i4	400000
General Management	i7	300000
Service Department	b9	200000

Average Salary Offered: 49977.97

Median Salary Offered: 49625

Data Analytics Tasks:

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

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

Male	2562
Female	1856

=COUNTIFS($D$1:$D$7168,J2,$C$1:$C$7168,$I$2)

J2=Male, I2=Hired

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.

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

Average Salary Offered by the company: **49977.97**

=AVERAGE(G2:G7168)

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.

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

Using *Sturge's formula*, salary interval can be calculated using the below formula:

=1+LOG(7167,2) , which comes to 13 (roughly)

Where count of salary offered=7167

Salary interval width: =(400000-100)/13

(Upper Limit-Lower limit)/13

Lower Limit	Upper Limit	Count	Percentage
100.00	30777.00	2179	30.40
30778.00	61455.00	2279	31.80
61456.00	92133.00	2193	30.60
92134.00	122811.00	513	7.16
122812.00	153489.00	0	0.00
153490.00	184167.00	0	0.00
184168.00	214845.00	1	0.01
214846.00	245523.00	0	0.00
245524.00	276201.00	0	0.00
276202.00	306879.00	1	0.01
306880.00	337557.00	0	0.00
337558.00	368235.00	0	0.00
368236.00	400000.00	1	0.01

=COUNTIFS(Sheet1!$G$2:$G$7168,">="&Sheet2!N2,Sheet1!$G$2:$G$7168,"<="&Sheet2!O2)

Percentage: =(P2/$P$15)*100

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

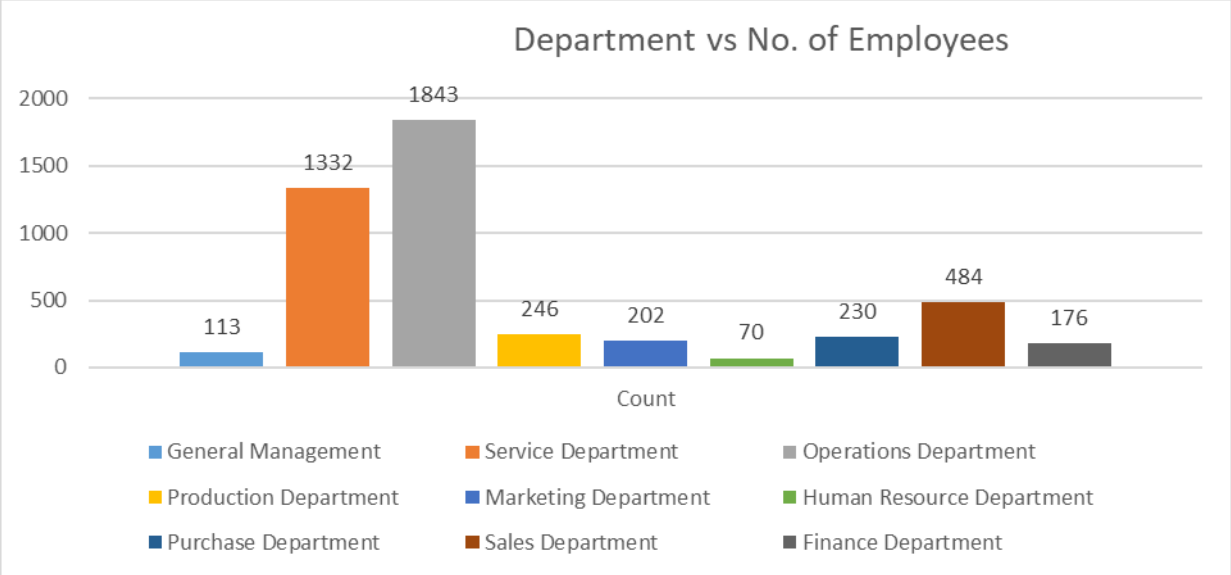
Your Task: Use a pie chart, bar graph, or any other suitable visualization to show the proportion of people working in different departments.

Department	Count
General Management	113
Service Department	1332
Operations Department	1843
Production Department	246
Marketing Department	202
Human Resource Department	70
Purchase Department	230
Sales Department	484
Finance Department	176

=COUNTIFS($E$1:$E$7168,S2,$C$1:$C$7168,$R$2)

Where, S2=Department name & $R$2=”Hired”

BAR CHART:



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

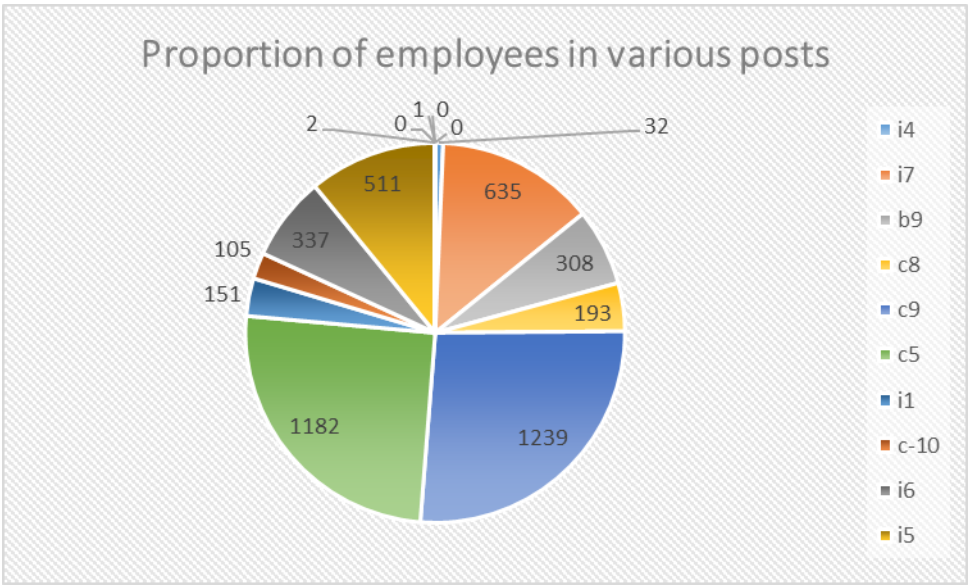
Your Task: Use a chart or graph to represent the different position tiers within the company. This will help you understand the distribution of positions across different tiers.

Post Name	Count
i4	32
i7	635
b9	308
c8	193
c9	1239
c5	1182
i1	151
c-10	105
i6	337
i5	511
m6	2
n9	0
n6	1
m7	0
n10	0

=COUNTIFS($F$1:$F$7168,Z2,$C$1:$C$7168,$Y$2)

Where, Z2=Post name & $Y$2=”Hired”

PIE CHART:



Result:

With this project, I became well-equipped with the powerful functions of MS Excel. Also, I was able to hone my statistics skills while completing the tasks in this assignment. I was able to understand

how past hires' data can be analyzed to make future hiring processes more efficient. The charts can be used to summarise mundane data in a more eye-catching way which helps various stakeholders to understand the data with just one glimpse.