HIRING PROCESS ANALYTICS

Submitted By: Rishabh Kukreja

TABLE OF CONTENTS

01

PROJECT DESCRIPTION

What we aim to do in this project

WHAT WE AIM

Our project is aimed at tracking major underlying trends about the hiring process using the data records of previous hirings to create a better strategy for future hirings

02 APPRAOCH

Step by Step framework used in the process

DATA ANALYTICS FRAMEWORK

1. ASK

Noting down key business questions asked by management team

2. PREPARE

Hiring data such as applicant_id, hiring_status,posts, salary etc is needed

3. PROCESS

Data is collected csv file, then missing and duplicate data is cleaned

4. ANALYZE

Data analysis using excel functions and pivot tables to answer the business questions

5. SHARE

Bringing data to life with visuals and sharing final report

6. ACT

Deriving meaningful insights from analysis

BUSINESS QUESTIONS

- How many males and females are Hired ?
- What is the average salary offered in this company ?
- Draw the class intervals for salary in the company ?
- Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working different department ?
- Represent different post tiers using chart/graph?

CLEANING DATA

- Column name containing gender name replaced from event_name to gender.
- We have found some duplicate application_id using conditional formatting, we will remove them as they are looking dubious (for same application_id we have both male and female applicant which might be unrealistic).
- Missing values in gender column: first we will see the missing cells belong to which department, then we will replace the blank value by most prominent gender under that department
- Missing values in post name column: first we will see the missing cells belong to which department, then we will replace the blank value by most prominent post under that department
- Missing values in salary column: first we will see the missing cells belong to which department, then we will replace the blank value by average salary value under that department
- Removing outliers from salary column: We see some very high and very low unrealistic values, we will replace the values which fall outside lower and upper limits with lower limit value (for extremely low outlier values) and upper limit value (for extremely high outlier values). We will use QUARTILE.INC() function

CLEANING DATA

- **Handling Missing values in gender column:** We will find dominant gender in each department by using pivot table, conditional formatting and if() function, then we will replace the N/A gender values with dominant gender in each department

Row Labels	N/A	Don't want to say	Female	Male	Grand Total	Dominant gender
Finance Department		16	257	14	287	Female
General Management		9	150	11	170	Female
Human Resource Department		4	34	56	94	Male
Marketing Department	1	12	101	210	324	Male
Operations Department	4	167	954	1626	2751	Male
Production Department	1	18	139	218	376	Male
Purchase Department	1	24	107	200	332	Male
Sales Department	1	31	248	463	743	Male
Service Department	7	110	661	1259	2037	Male

CLEANING DATA

- **Missing values in gender column:** We will find dominant post in each department by using pivot table and filtering then we will replace the N/A post values with dominant post in each department. We find that N/A in post value is under sales department and sales department has c5 post as most prominent one, so we will replace N/A with c5

	A	B	C	D	E	F	G
1	application_id	Interview Taken on	Status	gender	Department	Post Name	Offered Salary
7	289907	5/1/14 7:44	Hired	Male	Sales Department	N/A	85914

Sales Department	500
c5	216
c9	173
i7	111

CLEANING DATA

- **Missing values in salary column:** We will find avg salary in each department by using pivot table and filtering then we will replace the blank values with avg salary of that post in each department. We find that blank value in salary value is under sales department + i7 post and sales department and i7 post has avg salary of 48405.7636. So we will replace blank value with 48406

application_id	Interview Taken on	Status	gender	Department	Post Name	Offered Salary
114584	5/7/14 8:08	Rejected	Male	Sales Department	i7	

Sales Department	49420.7682
b9	59006.2857
c5	50043.1567
c8	45004.6667
c9	50215.1098
i1	41585
i4	60945.3
i5	48091.0341
i6	50223.2558
i7	48405.7636
c10	39233.1739

CLEANING DATA

- **Removing outliers from salary column:** We will calculate the lower and upper limit using quartile 1, quartile 3 and IQR values obtained by using quartile.inc() function. We found lower limit as -47641.75 and upper limit is 186020.625. Now we find the data in salary column that is out of the lower and upper limit range using OR() function and conditional formatting. After finding outliers we will replace it with lower limit for lower values and upper limit for higher values. We will replace these values by 186021

application_id	Interview Taken on	Status	gender	Department	Post Name	Offered Salary	Outlier status
649039	5/7/14 10:48	Hired	Female	Service Department	b9	200000	TRUE
795330	6/15/14 9:45	Hired	Female	General Management	i4	400000	TRUE
874368	7/21/14 15:39	Hired	Male	General Management	i7	300000	TRUE

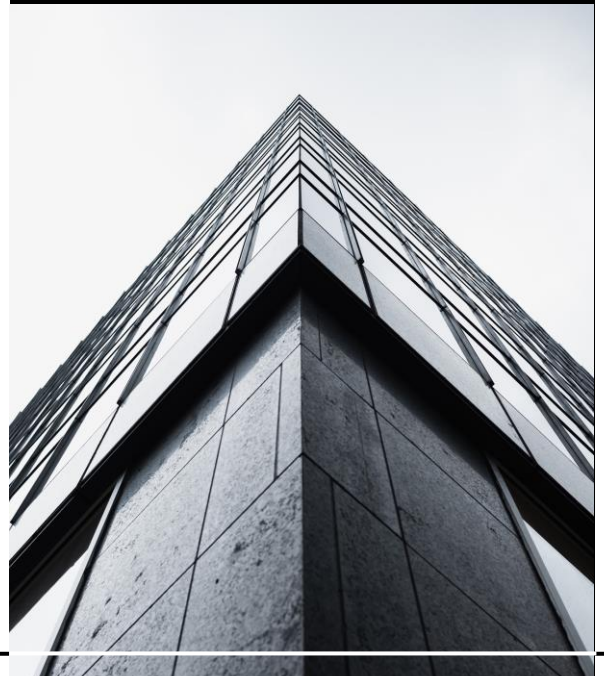
QTL 1	25588.25
QTL 3	74408.25
IQR	48820
Lower limit	-47641.75
Upper limit	186020.625

03

TECH-STACK USED

Armoury of tools used during the project

TOOL USED



MS-EXCEL

MS-POWERPOINT

04 RESULTS

Making sense of why we did the project

ANSWERS TO BUSINESS QUESTIONS

A) How many males and females are Hired ?

Males hired	4072
Females hired	2651

B) What is the average salary offered in this company ?

49948

Average salary offered to hired employees:

49681

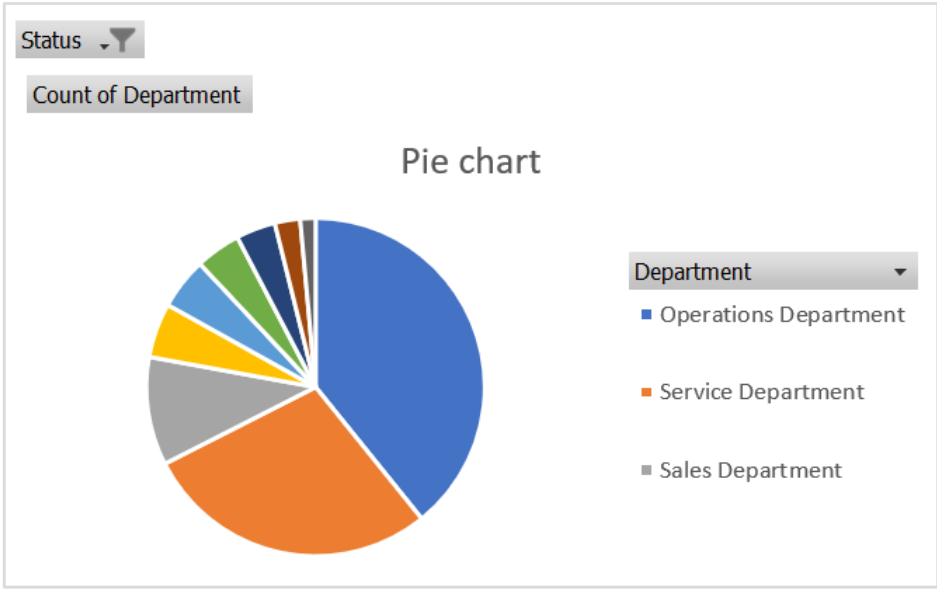
C) Draw the class intervals for salary in the company ?

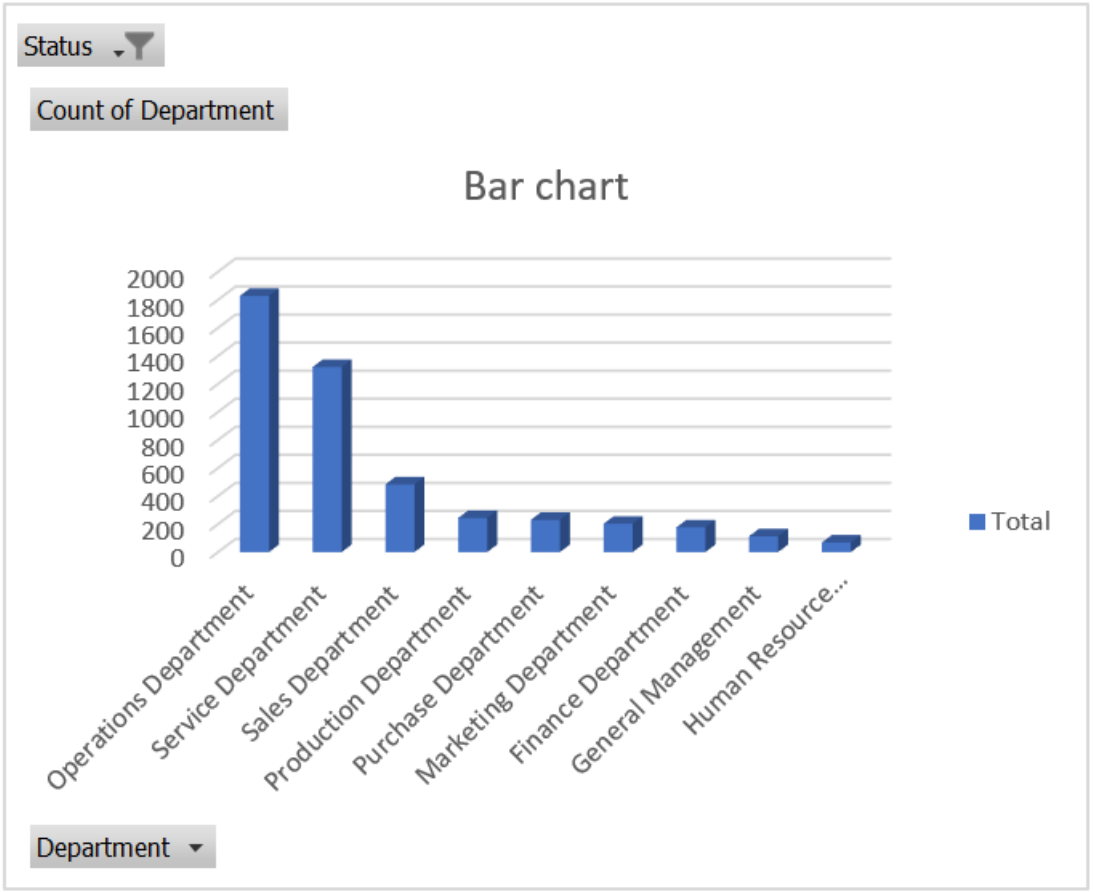
Salary class interval	Count of Offered Salary
100-20099	1397
20100-40099	1416
40100-60099	1524
60100-80099	1420
80100-100099	1354
180100-200099	3
Grand Total	**7114**

D) Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working in different department ?

Status	Hired

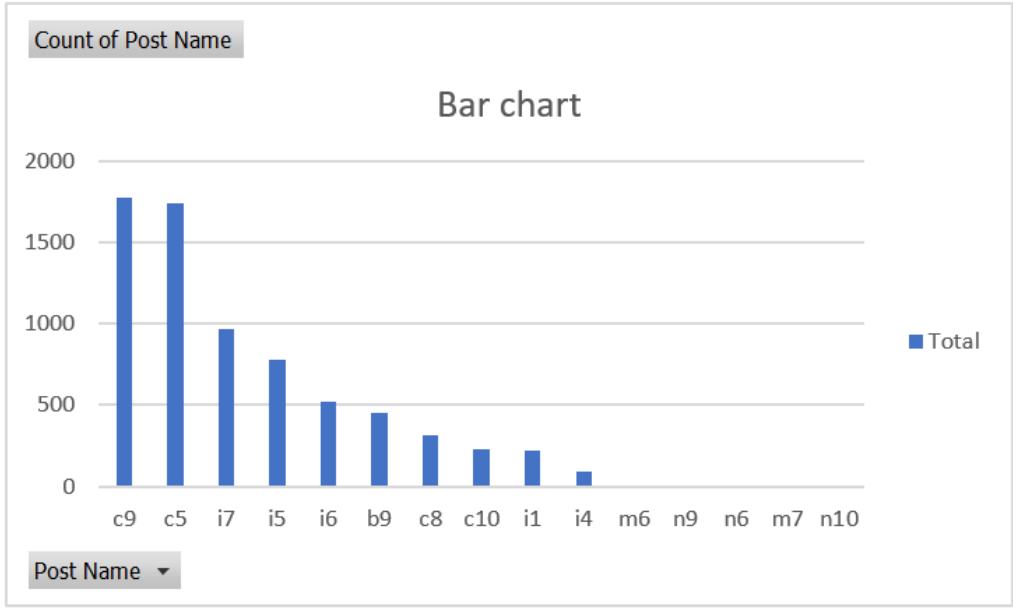
Departments	Count of Department
Operations Department	1828
Service Department	1320
Sales Department	483
Production Department	243
Purchase Department	230
Marketing Department	202
Finance Department	176
General Management	113
Human Resource Department	68
Grand Total	**4663**



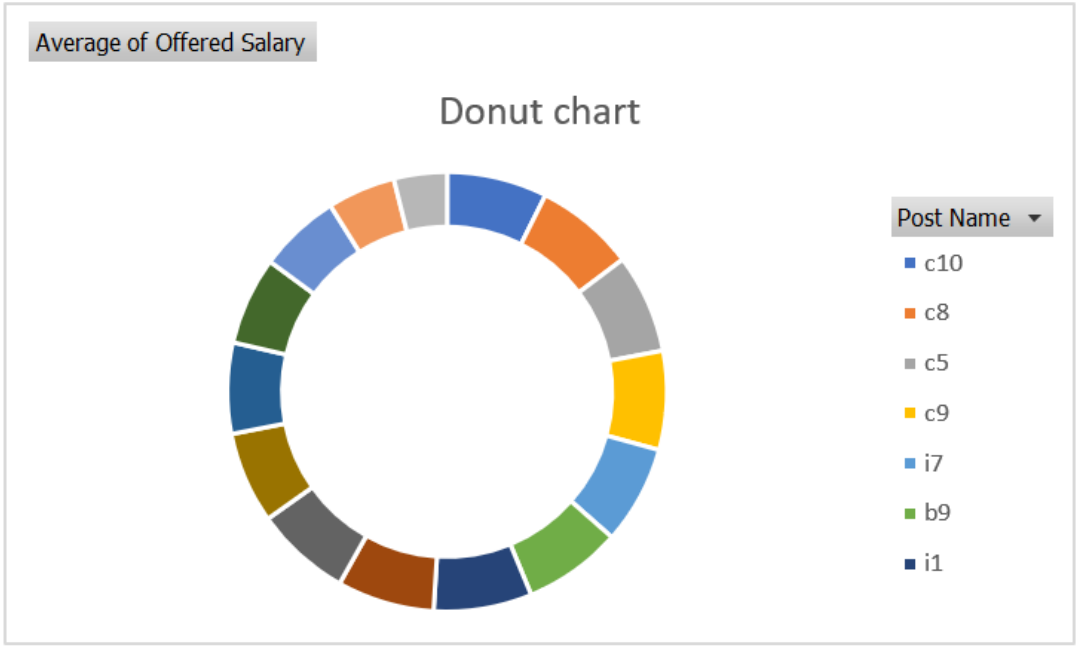

Status
Count of Department
Bar chart
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Operations Department
Service Department
Sales Department
Production Department
Purchase Department
Marketing Department
Finance Department
General Management
Human Resource...
Total
Department

E) Represent different post tiers using chart/graph?

Post names	Count of Post Name
c9	1775
c5	1741
i7	973
i5	783
i6	523
b9	455
c8	316
c10	231
i1	222
i4	88
m6	3
n9	1
n6	1
m7	1
n10	1
Grand Total	**7114**



Post name	Average of Offered Salary
c10	50957.18182
c8	50623.17405
c5	50245.47674
c9	50211.18085
i7	50012.40082
b9	49967.98681
i1	49943.93694
i5	49374.79821
i6	48678.11855
i4	46446.26136
n9	46219
n6	44700
m7	41402
m6	34521.33333
n10	26990
Grand Total	**49947.22744**



05 INSIGHTS

Insights from the analysis

INSIGHTS

- 40,000 – 60,000 is the most sought after salary bracket for jobs which is inline with the average salary of 49948 offered by the company.
- Operations department and Service department are the most active departments, General management and Human resource departments are least active departments in hiring.
- c9 and c5 are the most sought after jobs among all the posts and m6, n9, n6, m7 and n10 are the least sought after jobs.
- c10 and c8 are the departments offering top salaries and m6 and n10 are the departments offering least salaries.

THANKS