HIRING PROCESS ANALYTICS
-YAADHAV R

PROJECT DESCRIPTION :

THE OBJECTIVE OF THIS PROJECT IS TO TAKE THE ROLE AS A DATA ANALYST. AS AN ANALYST, MY JOB IS TO ANALYZE THE DATASET OF PREVIOUS HIRES OF A COMPANY THAT CONTAINS VARIOUS INFORMATIONS LIKE THE NUMBER OF REJECTIONS, INTERVIEWS, JOB TYPES, & VACANCIES AND DERIVE VALUABLE INSIGHTS FROM IT.ALSO TO ANSWER THE VARIOUS QUESTION POSTED BY THE COMPANY AND IMPROVE ITS HIRING PROCESS.

TECH-STACK USED :

MICROSOFT EXCEL 2023 - VERSION 16.80 IS USED IN THIS PROJECT AS IT IS :

- IT A SIMPLE AND EASY TO USE SOFTWARE.
- ALL THE TOOLS FOR DATA ANALYSIS IS AVAILABLE.
- EASY FOR DATA VISUALIZATION WITH VARIOUS CHARTS.

PREPROCESSING

A) HANDLING MISSING DATA:

ONLY THE LAST ROW HAD THE MISSNG DATA AND THAT ROW HAS BEEN REMOVED AS IT WAS THE KEY FEATURE 'OFFERED SALARY'.

application_i	Interview Taken or	Status	event_name	Department	Post Name	Offered Salary
114584	07/05/14 8:08	Rejected	Male	Sales Department	i7	

B) CLUBBING COLUMNS:

NO COLUMNS WERE CLUBBED AS THERE ARE ONLY 7 COLUMNS AND ALL THE 7 COLUMNS CONTAIN DIFFERENT TYPES OF DATA WHICH CANNOT BE CLUBBED WITH ONE ANOTHER.

C) OUTLIER DETECTION:

THE FOLLOWING STEPS ARE CARRIED OUT IN ORDER TO FIND THE OUTLIERS:

- Q1 & Q3 IS FOUND.
- INTER QUARTILE RANGE(IQR) = Q3-Q1.
- LOWER BOUND(LB) = Q1- (1.5*IQR).
- UB BOUND(UB) = Q3+ (1.5*IQR).

Q1	25449
Q3	74414
Q3-Q1	73447.5
LB	-47998.5
UB	147861.5

D) REMOVING OUTLIERS:

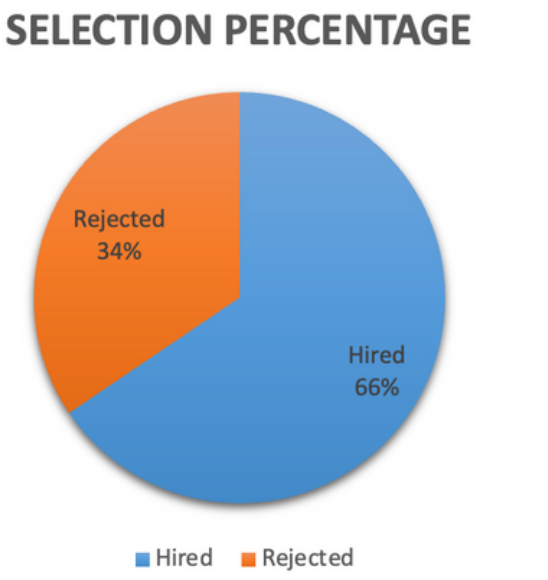
THE LOWER BOUND(LB) AND UPPER BOUND(UB) HAS BEEN CALCULATED USING THE INTERQUARTILE RANGE(Q3-Q1) ON THE 'OFFERED SALARY' COLUMN AND THE 3 ROWS WITH 'OFFERED SALARY'>UB HAS BEEN REMOVED AS IT SKEWS THE DATA. THOSE ROWS ARE SHOWN BELOW.

application_id	Interview Taken on	Status	event_name	Department	Post Name	Offered Salary
795330	15/06/14 9:45	Hired	Female	General Management	i4	400000
874368	21/07/14 15:39	Hired	Male	General Management	i7	300000
649039	07/05/14 10:48	Hired	Female	Service Department	b9	200000

E) DATA SUMMARY:

AFTER CLEANING AND PREPARING THE DATA, HERE ARE SOME STATISTICS MEASURES AND DATA VISUALISATION.

Mean(total)	49878.33
Mean(hired)	49592.96
Median	49614.5



TASKS

A) HIRING ANALYSIS:

Females Hired	1854
Males Hired	2562

APPROACH:

FORMULA USED FOR FEMALES :

=COUNTIFS(D2:D7165,"FEMALE", C2:C7165,"HIRED")

FORMULA USED FOR MALES :

=COUNTIFS(D2:D7165,"MALE", C2:C7165,"HIRED")

INSIGHTS:

- THE NUMBER OF MALES HIRED ARE MORE THAN THE NUMBER OF FEMALES HIRED.
- HOWEVER THE PERCENTAGE CONVERSION OF JOB IS HIGHER FOR FEMALES (69.36%) COMPARED TO MALES(62.75%).

B) SALARY ANALYSIS:

Avg Offered Salary	49878.33

APPROACH:

FORMULAE USED:

=AVERAGE(G2:G7185)

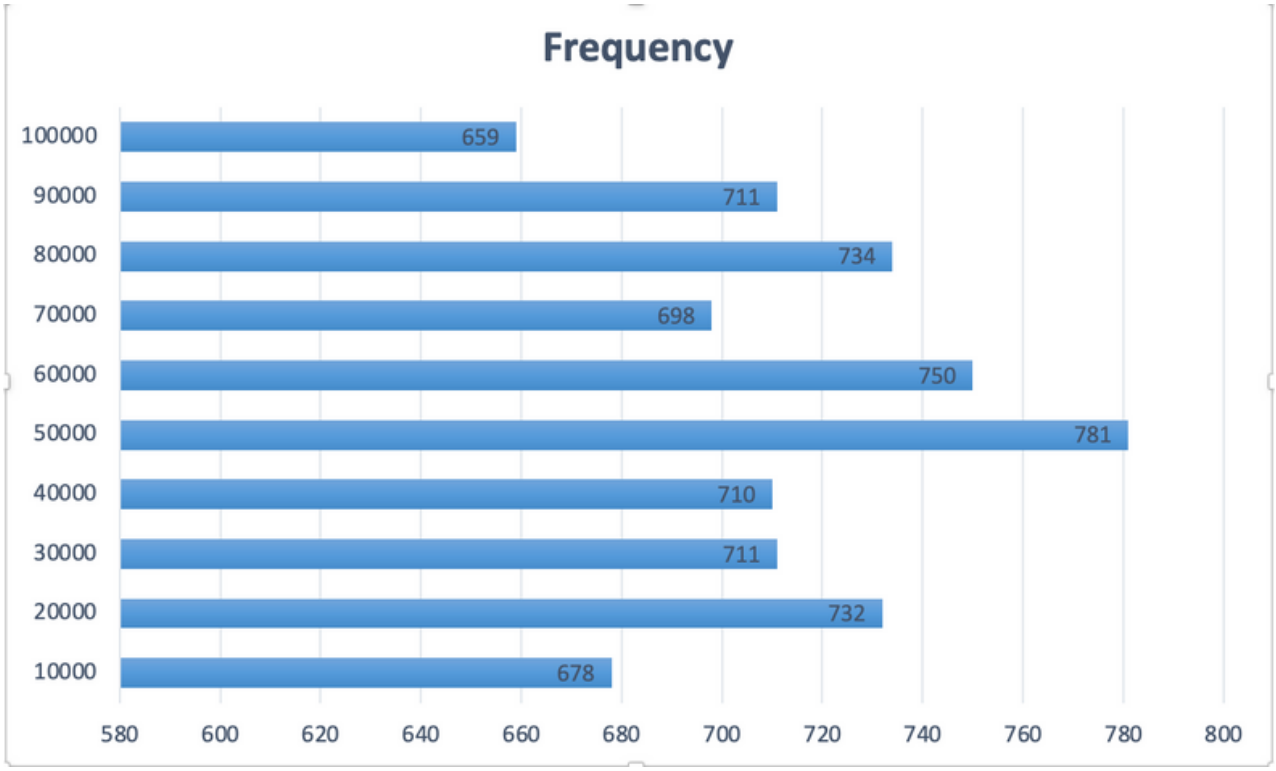
INSIGHTS:

- THE AVERAGE SALARY OFFERED BY THE COMPANY IS 49878.33.
- THE MEAN OF OFFERED SALARY IS SLIGHTLY MORE THAN THAT OF THE MEDIAN.

C) SALARY DISTRIBUTION:

APPROACH:

- FIRST A COLUMN(L3:L12) OF UPPER LIMITS OF A CLASS INTERVAL IS CREATED BASED ON THE RANGE OF OFFERED SALARY.
- THEN ANOTHER COLUMN OF FREQUENCIES IS CREATED USING =FREQUENCY(G2:G7165,L3:L12)
- NOW THESE 2 COLUMNS ARE SELECTED AND A BAR GRAPH IS INSERTED.



INSIGHTS:

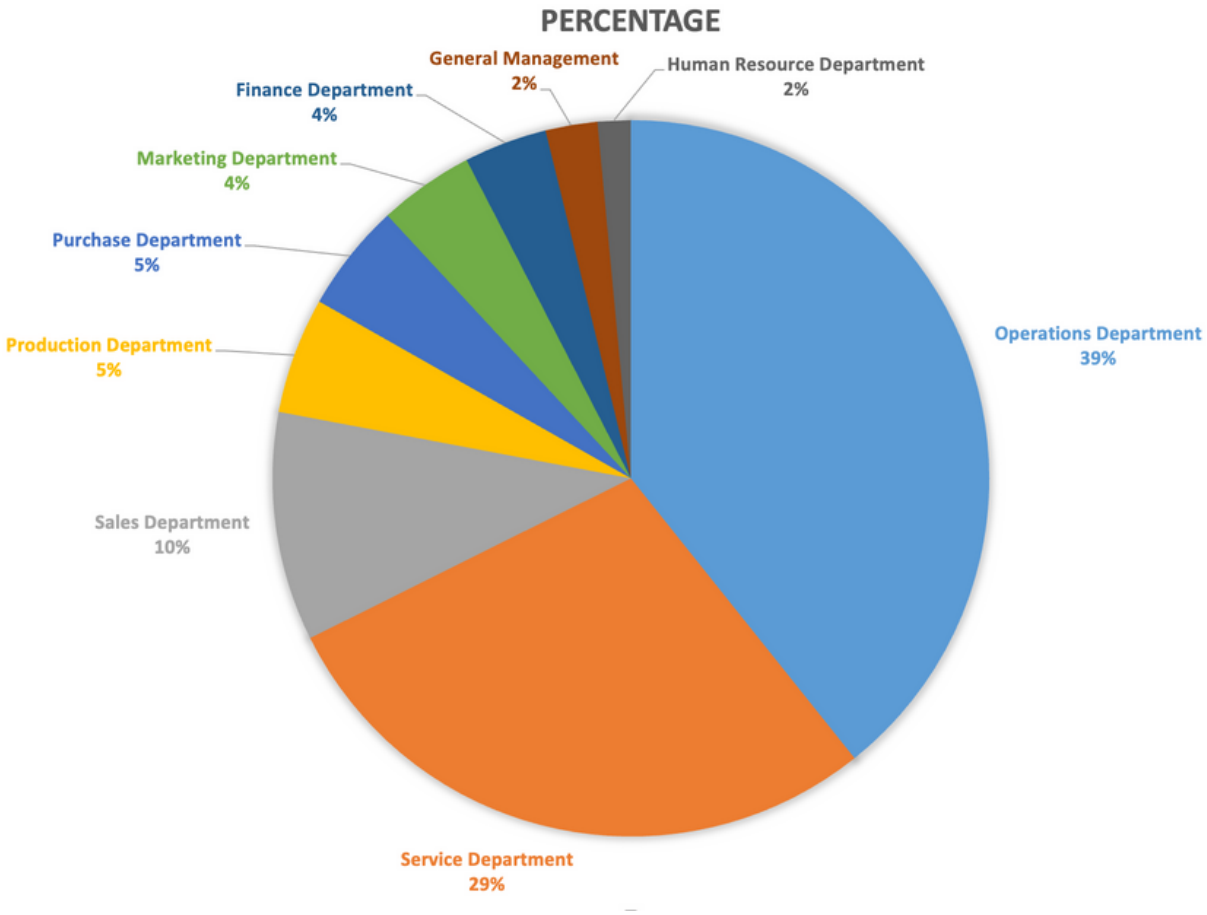
- MOST OF THE HIRES WERE OFFERED BETWEEN 40K TO 50K SALARY.
- THERE IS NO SIGNIFICANCE DIFFERENCE IN THE DISTRIBUTION AS THE RANGE OF IT IS AROUND 120 ONLY.
- THIS DISTRIBUTION WOULB HAVE BEEN CLOSE TO NORMAL DISTRIBUTION, IF 2ND AND 7TH CLASSES ARE NOT CONSIDERED.

D) DEPARTMENTAL ANALYSIS:

APPROACH:

- FIRST A COPY OF DEPARTMENT IS CREATED ON THE ‘O’ COLUMN.
- THEN DUPLICATES ARE REMOVED FROM THIS COLUMN BY SELECTING THIS COLUMN, THEN DATA > REMOVE DUPLICATES.
- NOW THE COUNT OF CANDIDATES WHO ARE HIRED IN EACH DEPARTMENTS IS CALCULATED BY
  =COUNTIFS(E2:E7165,K20,C2:C7165,"HIRED")
- THIS COUNT IS NOW DIVIDED BY THE TOTAL AND MULTIPLIED BY 100 WHICH GIVES THE PERCENTAGE OF EACH DEPARTMENT.
- THE DEPARTMENT AND PERCENTAGE IS SORTED BY DECREASING VALUES OF PERCENTAGE.
- FINALLY AFTER THE SELECTING THESE 2 COLUMNS, A PIE CHART IS INSERTED.

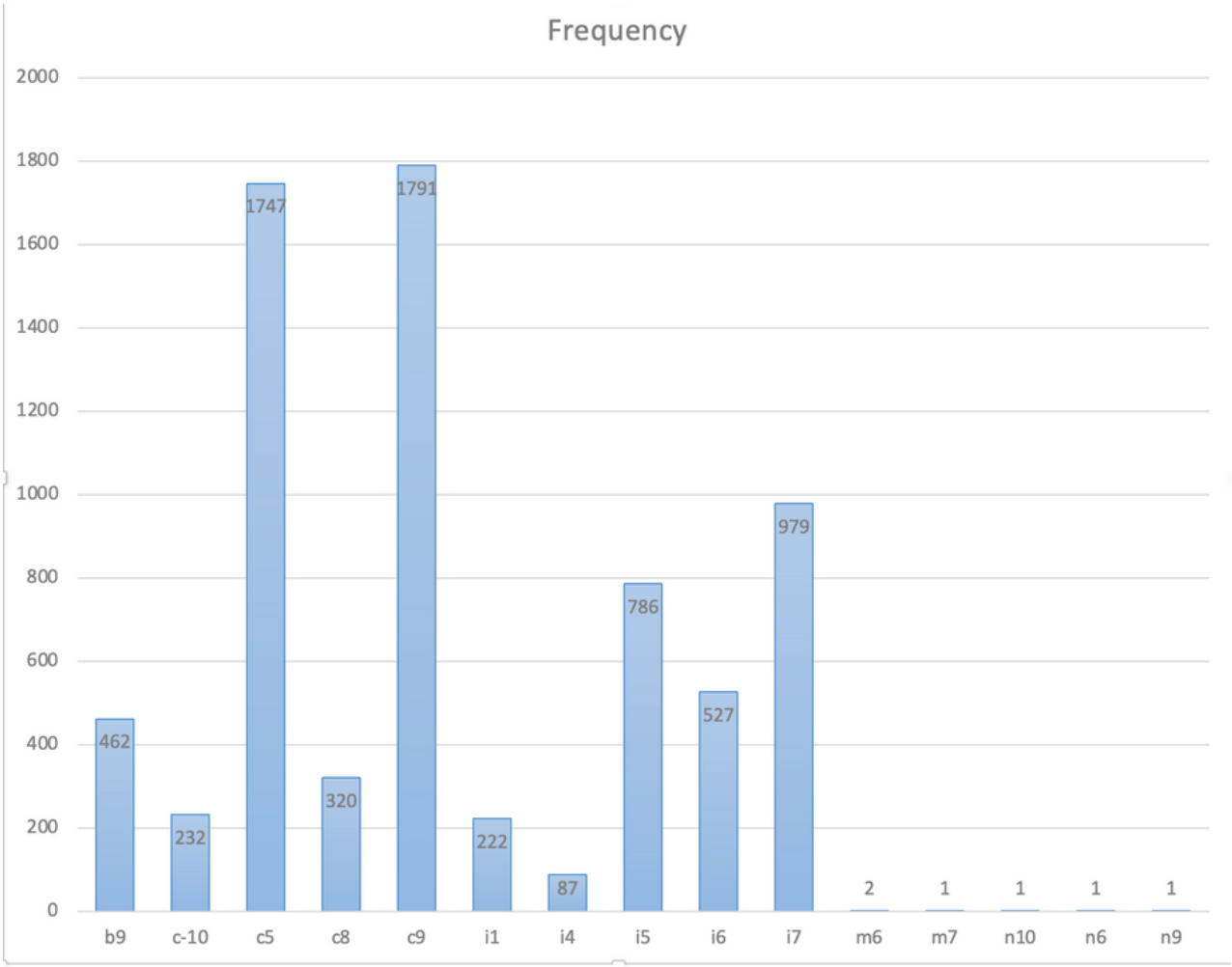
Department	Percentage
Operations Department	39.26
Service Department	28.36
Sales Department	10.33
Production Department	5.24
Purchase Department	4.9
Marketing Department	4.3
Finance Department	3.75
General Management	2.34
Human Resource Department	1.49



INSIGHTS:

- OPERATIONS AND SERVICE DEPARTMENT TOGETHER HAS 68% OF HIRES WHILE THE REST COMPRISED ONLY 32%.
- HUMAN RESOURCE AND GENERAL MANAGEMENT HAD THE LEAST SHARE WITH 2% EACH.

E) POST-TIER ANALYSIS:



APPROACH:

- FIRST A COPY OF POST NAME IS CREATED ON THE 'Q' COLUMN.
- THEN DUPLICATES ARE REMOVED FROM THIS COLUMN BY SELECTING THIS COLUMN, THEN DATA > REMOVE DUPLICATES.
- NOW THE FREQUENCY OF EACH POST IS CALCULATED BY =COUNTIF(F2:F7165,N2).
- FINALLY AFTER SELECTING THESE 2 COLUMNS, A BAR GRAPH IS INSERTED.

INSIGHTS:

- C5 & C9 ARE THE POSTS THAT 50% OF THE HIRES APPLIED TO.
- THE POSTS M6, M7, N10, N6, N9 HAD ONLY 5 APPLICATIONS IN TOTAL, WHICH IS ONLY 0.6% OF TOTAL APPLICATIONS.

RESULT:

- FIRST OF ALL THIS PROJECT HELPED TO UNDERSTAND THE CONCEPTS I LEARNED IN BETTER AND INTRESTING WAY.
- SO NOW FEEL CONFIDENT IN APPLYING EXCEL AND STATISTICS SKILLS AS I WAS ABLE TO COMPLETE ALL THE GIVEN TASKS.
- IT WAS SO EXICTING TO ANALYZE THE OUTPUTS AND DERIVE INSIGHTS FROM IT.
- OVERALL IT WAS GREAT TO EXPERIENCE TO APPLY THE SKILLS AND LEARN ALONG THE WAY.
- LOOKING FORWARD TO FACE THE UPCOMING CHALLENGES WITH CONFIDENCE AND EXICTMENT.