

USER ANALYTICS IN THE TELECOMMUNICATION INDUSTRY



AGENDA

- User Overview analysis
- User Engagement analysis
- User Experience analysis
- User Satisfaction Analysis

INTRODUCTION

In telecommunication, CDR or Call Detail Record is the voice channel and XDR is the data channel equivalent. So here, considering xDR as data sessions Detail Record. In xDR, user behavior can be tracked through the following applications: Social Media, Google, Email, YouTube, Netflix, Gaming, and Others.

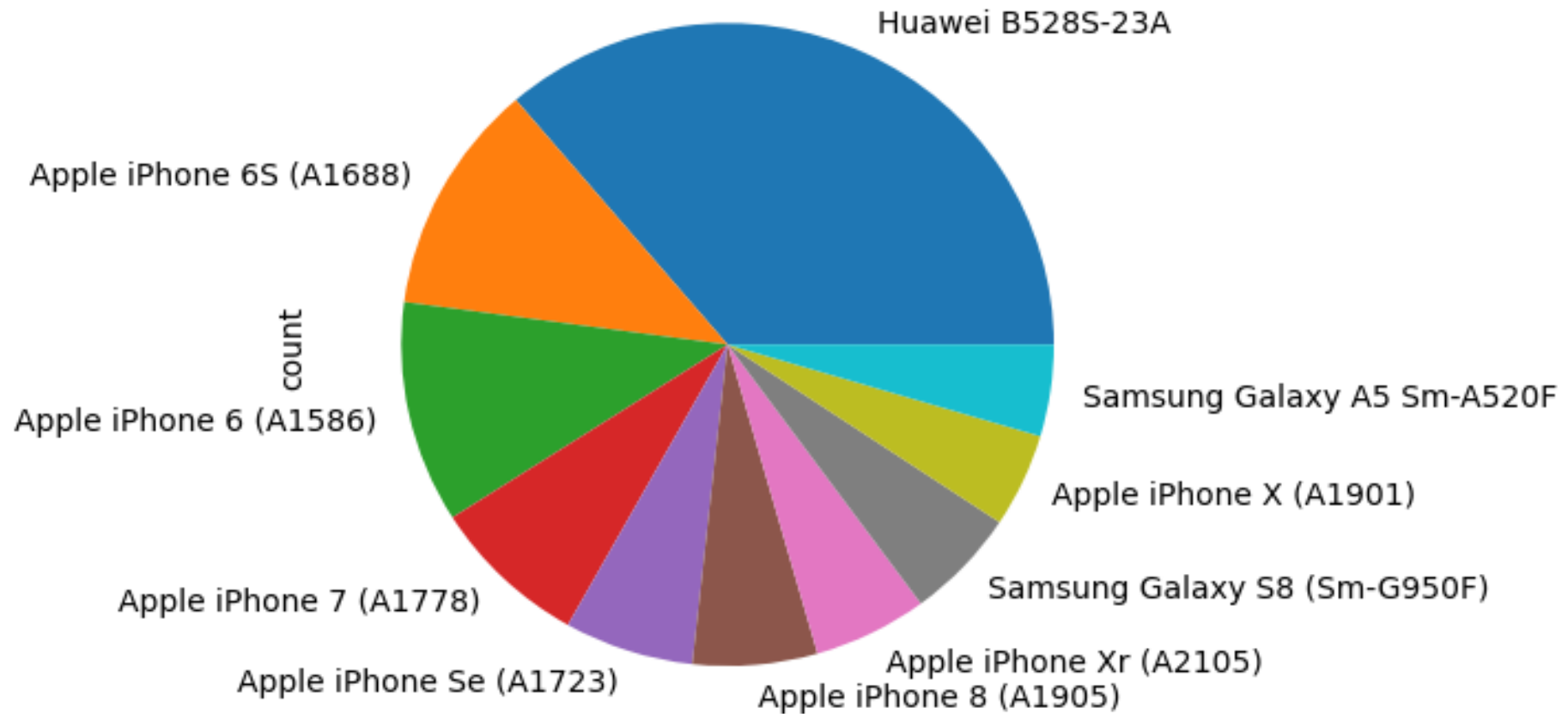


PRIMARY GOALS

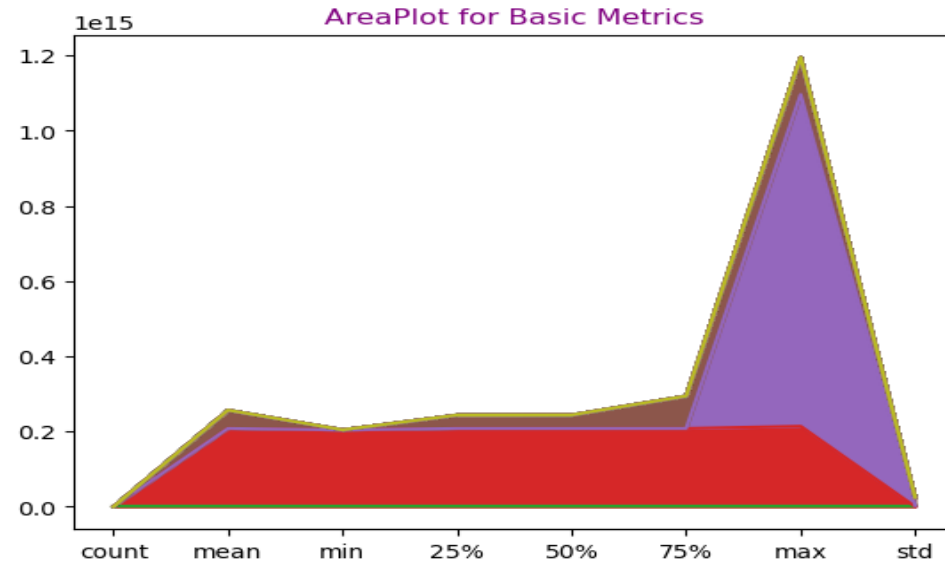
Growth potential of the company (positive or negative) based on the data

USER OVERVIEW ANALYSIS

PieChart for Top 10 Handsets used by the customers

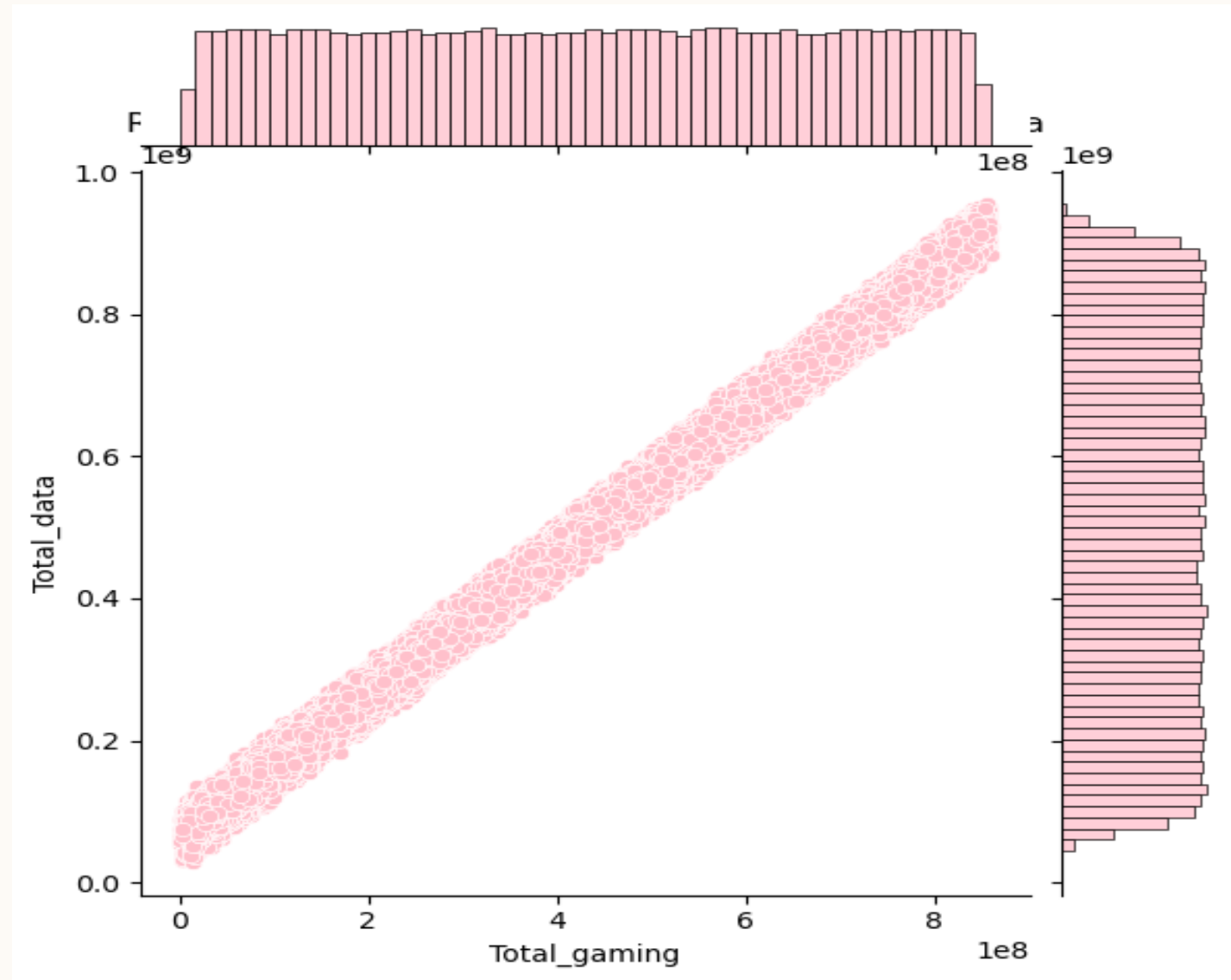


Basic Metrics are
Count, Mean,
Minimum, Median,
Maximum and
Standard Deviation.
Basic metrics of all
the columns are
shown here in this
chart.



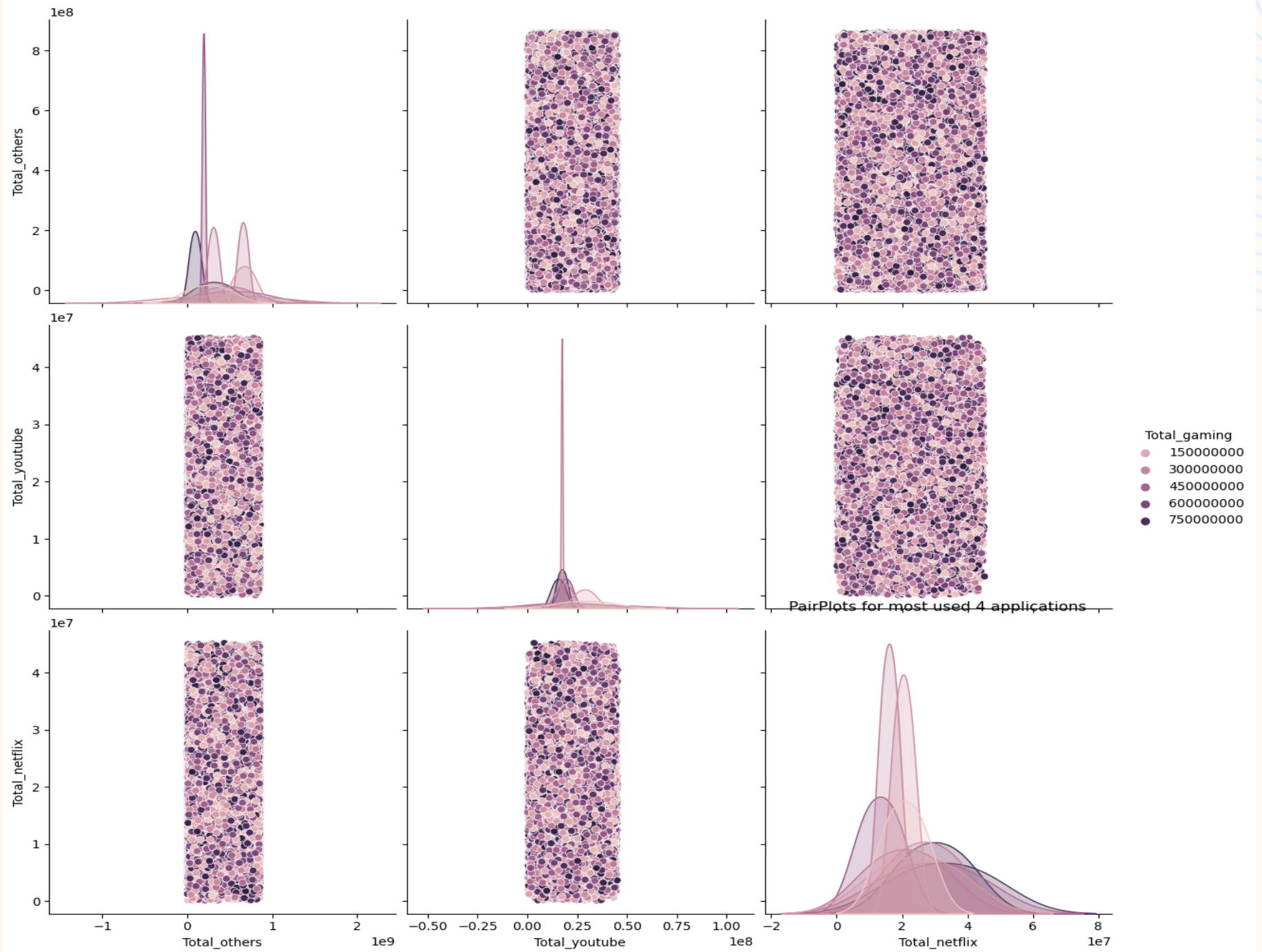
- Start ms
- End ms
- Dur. (s)
- IMSI
- MSISDN/Number
- IMEI
- Avg RTT DL (ms)
- Avg RTT UL (ms)
- Avg Bearer TP DL (kbps)
- Avg Bearer TP UL (kbps)
- TCP DL Retrans. Vol (Bytes)
- TCP UL Retrans. Vol (Bytes)
- DL TP < 50 Kbps (%)
- 50 Kbps < DL TP < 250 Kbps (%)
- 250 Kbps < DL TP < 1 Mbps (%)
- DL TP > 1 Mbps (%)
- UL TP < 10 Kbps (%)
- 10 Kbps < UL TP < 50 Kbps (%)
- 50 Kbps < UL TP < 300 Kbps (%)
- UL TP > 300 Kbps (%)
- HTTP DL (Bytes)
- HTTP UL (Bytes)
- Activity Duration DL (ms)
- Activity Duration UL (ms)
- Dur. (ms)
- Nb of sec with 125000B < Vol DL
- Nb of sec with 1250B < Vol UL < 6250B
- Nb of sec with 31250B < Vol DL < 125000B
- Nb of sec with 6250B < Vol DL < 31250B
- Nb of sec with Vol DL < 6250B
- Nb of sec with Vol UL < 1250B
- Social Media DL (Bytes)
- Social Media UL (Bytes)
- Google DL (Bytes)
- Google UL (Bytes)
- Email DL (Bytes)
- Email UL (Bytes)
- Youtube DL (Bytes)
- Youtube UL (Bytes)
- Netflix DL (Bytes)
- Netflix UL (Bytes)
- Gaming DL (Bytes)
- Gaming UL (Bytes)
- Other DL (Bytes)
- Other UL (Bytes)
- Total UL (Bytes)
- Total DL (Bytes)
- Actual Duration
- xDR sessions

Gaming is the maximum used application out of all other applications. Its relationship with Total Data is 86%
Here is the JointPlot of Gaming and Total Data

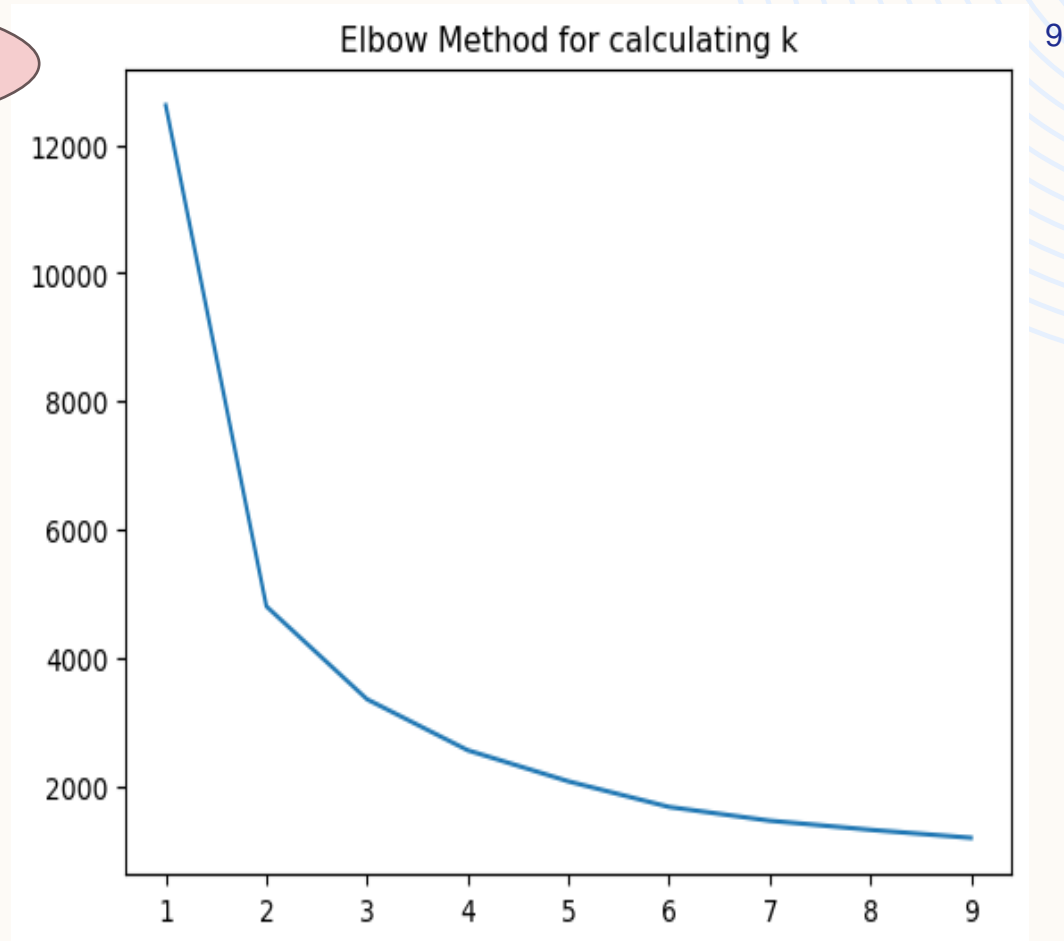
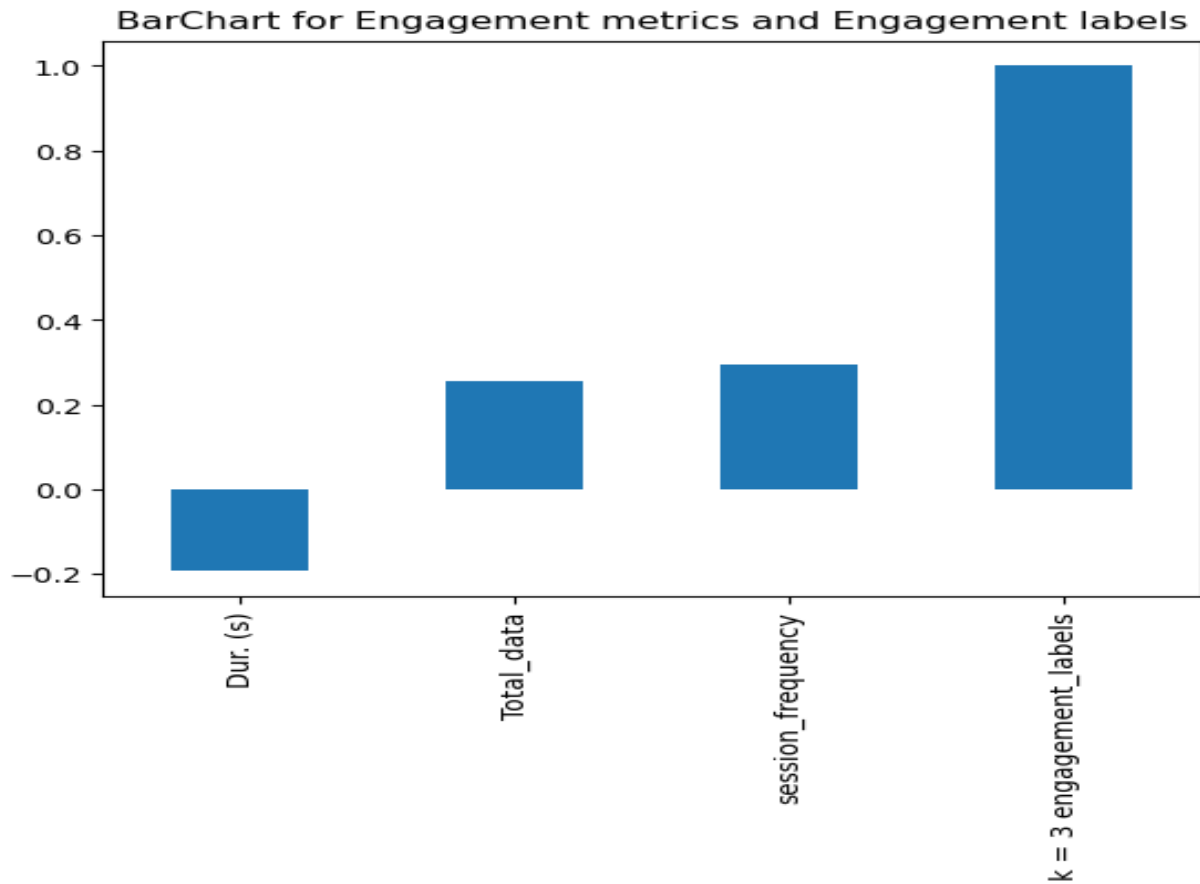


This is the PairPlot for 4 main applications which are mostly used. The list of those applications are:

Gaming
Others
Youtube
Netflix



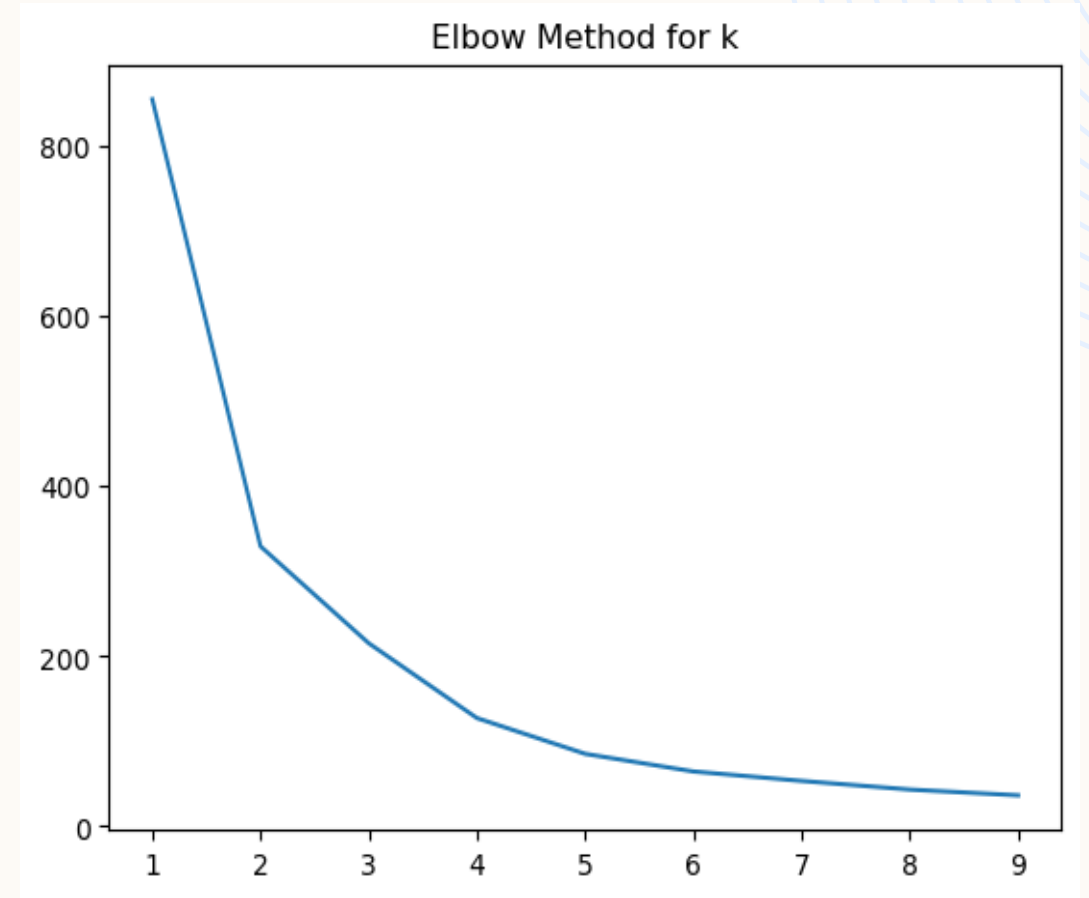
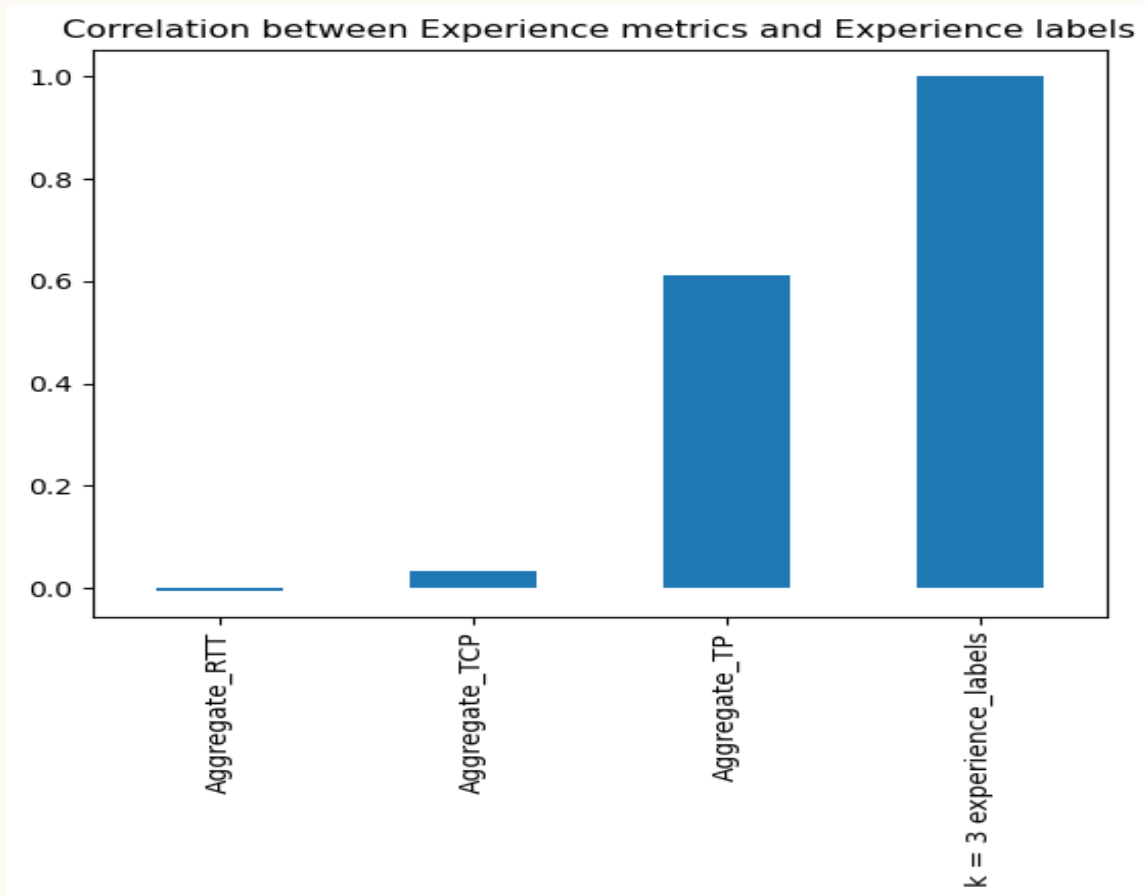
USER ENGAGEMENT ANALYSIS



Value of k should be 2

This is the correlation between Engagement Metrics and Engagement Labels
Engagement Metrics are Duration, Total data and Session frequency

USER EXPERIENCE ANALYSIS



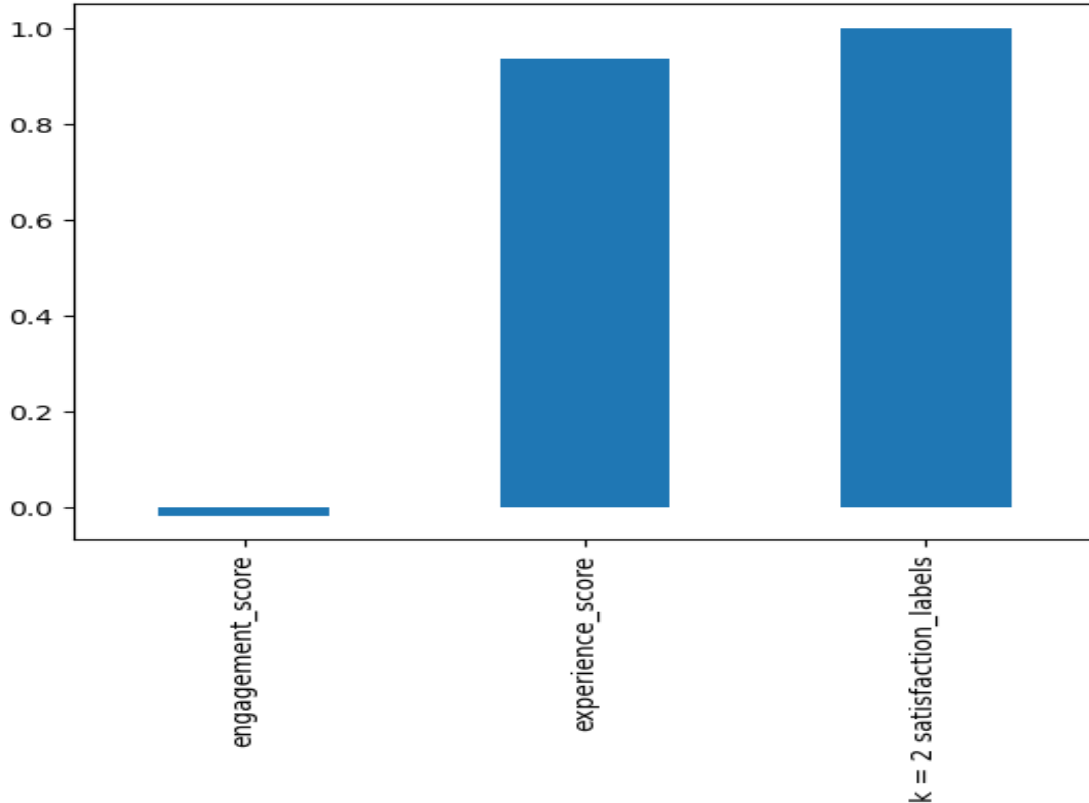
10

Value of k should be 2

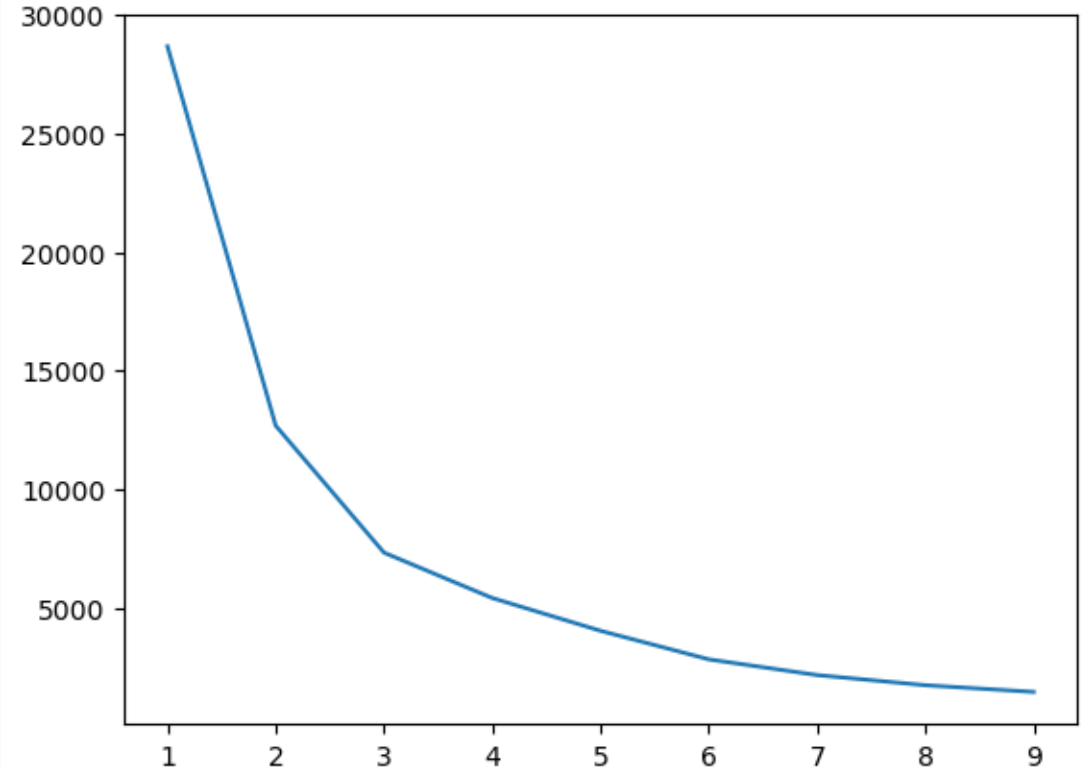
This is the correlation between Experience Metrics and Experience Labels
Experience Metrics are RTT, TCP and Throughput

USER SATISFACTION ANALYSIS

Correlation between Satisfaction metrics and Satisfaction labels

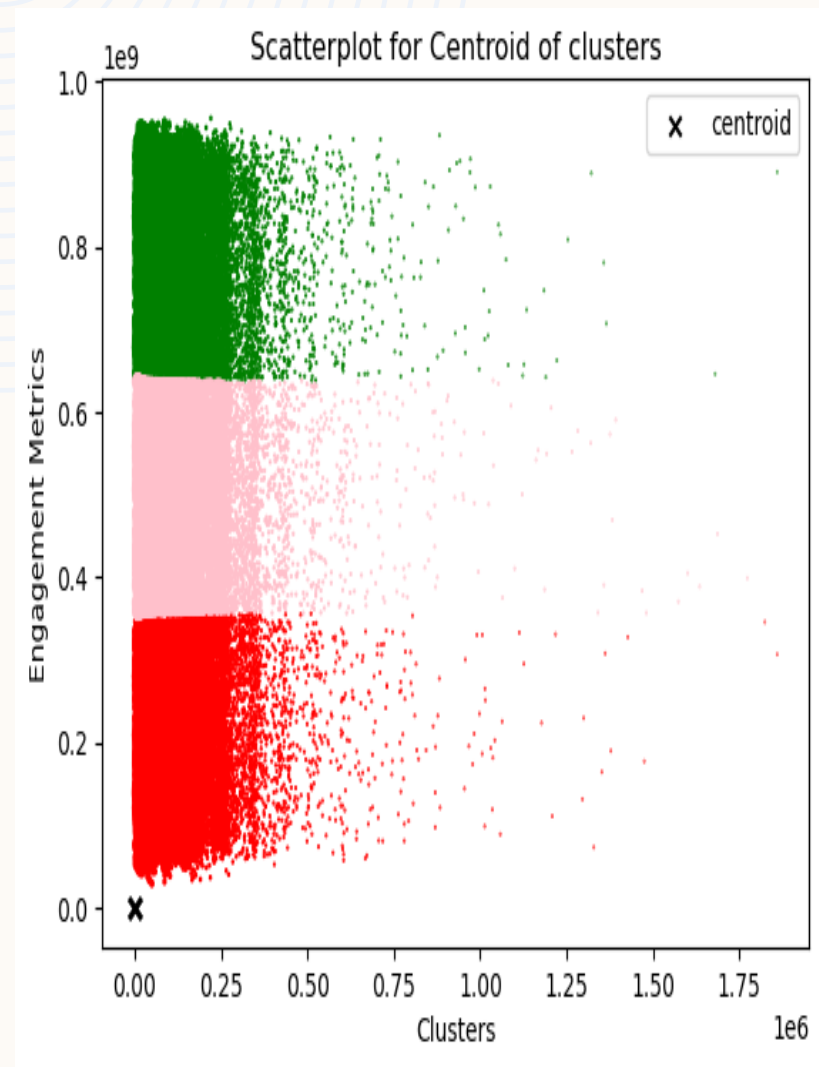


Elbow method for k

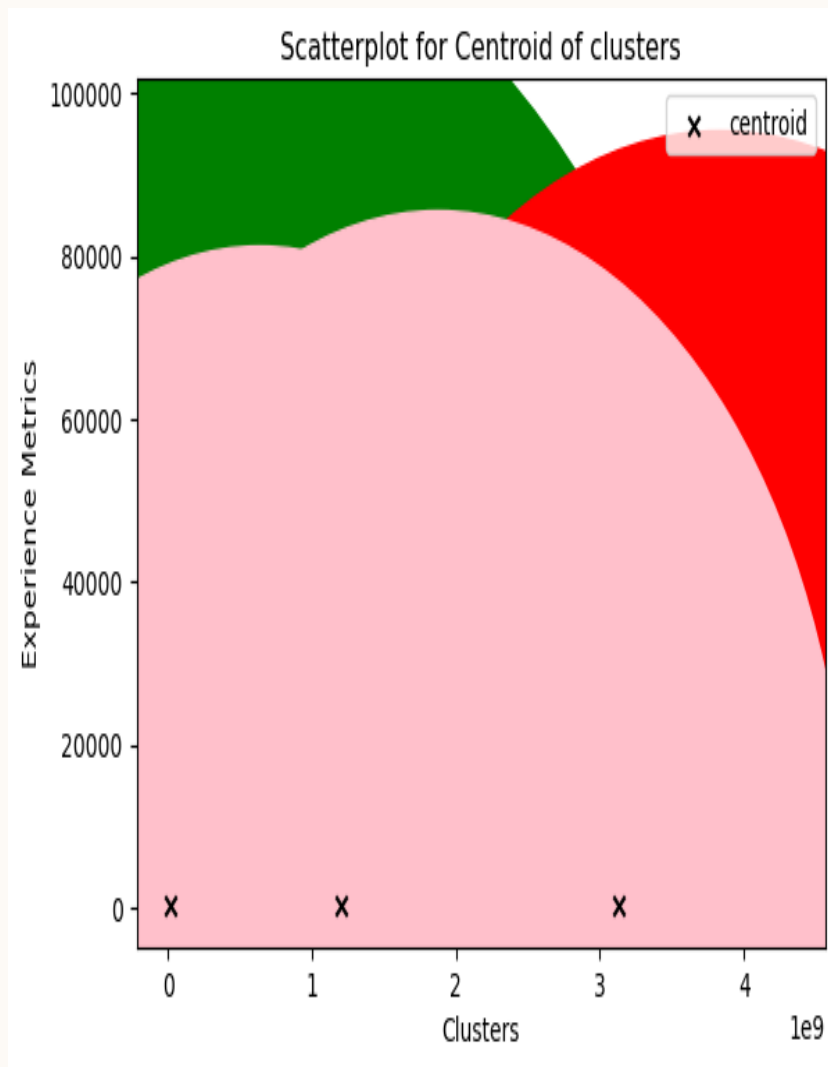


Value of k can be 2 or 3

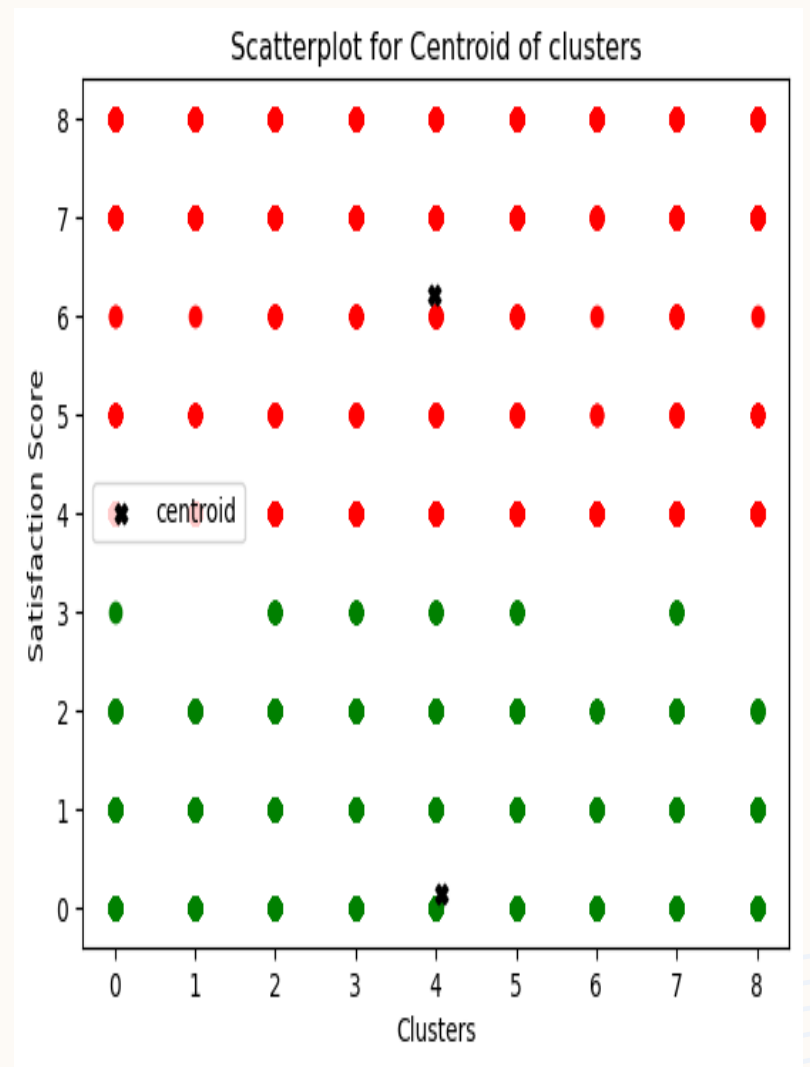
This is the correlation between Satisfaction Metrics and Satisfaction Labels
Satisfaction Metrics are Engagement Score and Experience Score



Engagement Metrics Clusters



Experience Metrics Clusters



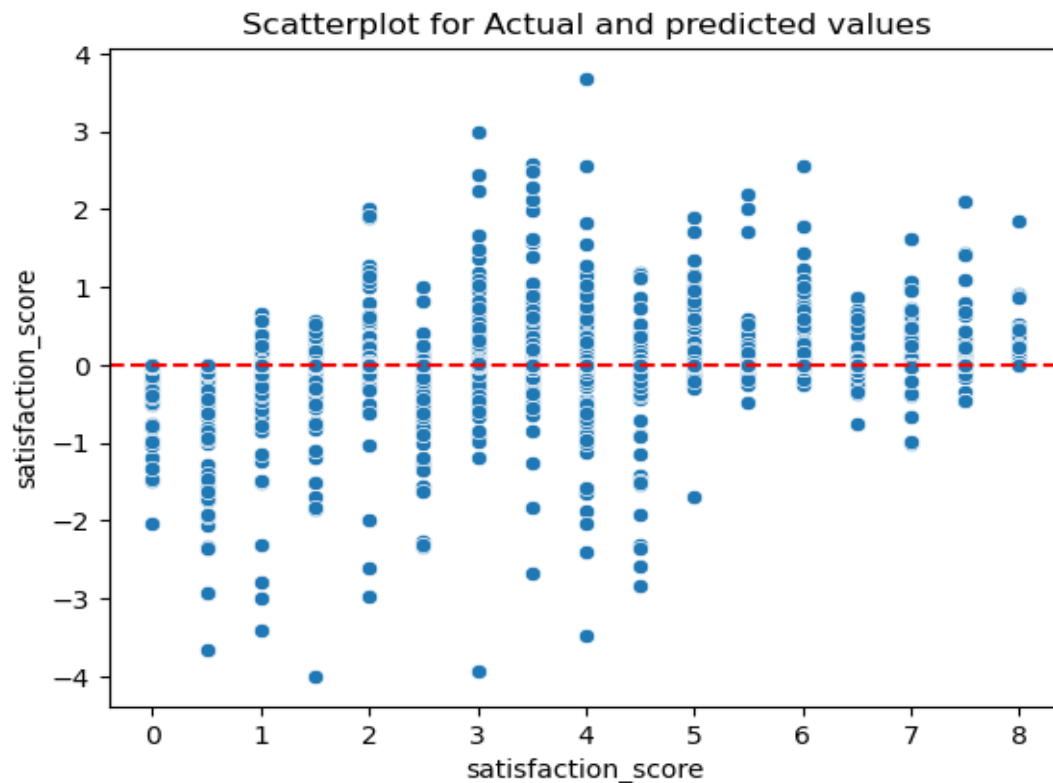
Satisfaction Metrics Clusters

Random Forest Regression Model

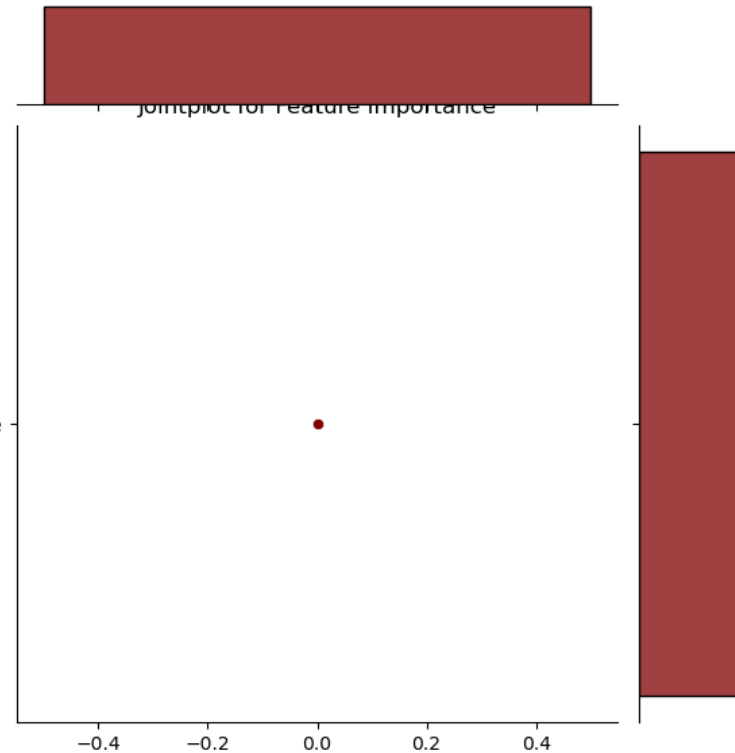
Accuracy = 99%

MSISDN/Number	session_frequency	Dur. (s)	Total_data	Aggregate_TCP	Aggregate_RTT	Aggregate_TP
3.360100e+10	1.285135	116720.0	878690574.0	2.156957e+07	46.000000	76.0
3.360100e+10	0.827683	181230.0	156859643.0	2.156957e+07	31.000000	99.0
3.360100e+10	1.111374	134969.0	595966483.0	2.156957e+07	127.458589	97.0
3.360101e+10	3.007358	49878.0	422320698.0	7.607247e+05	84.000000	248.0
3.360101e+10	8.086743	18549.0	585578364.0	2.156957e+07	28.000000	63.0

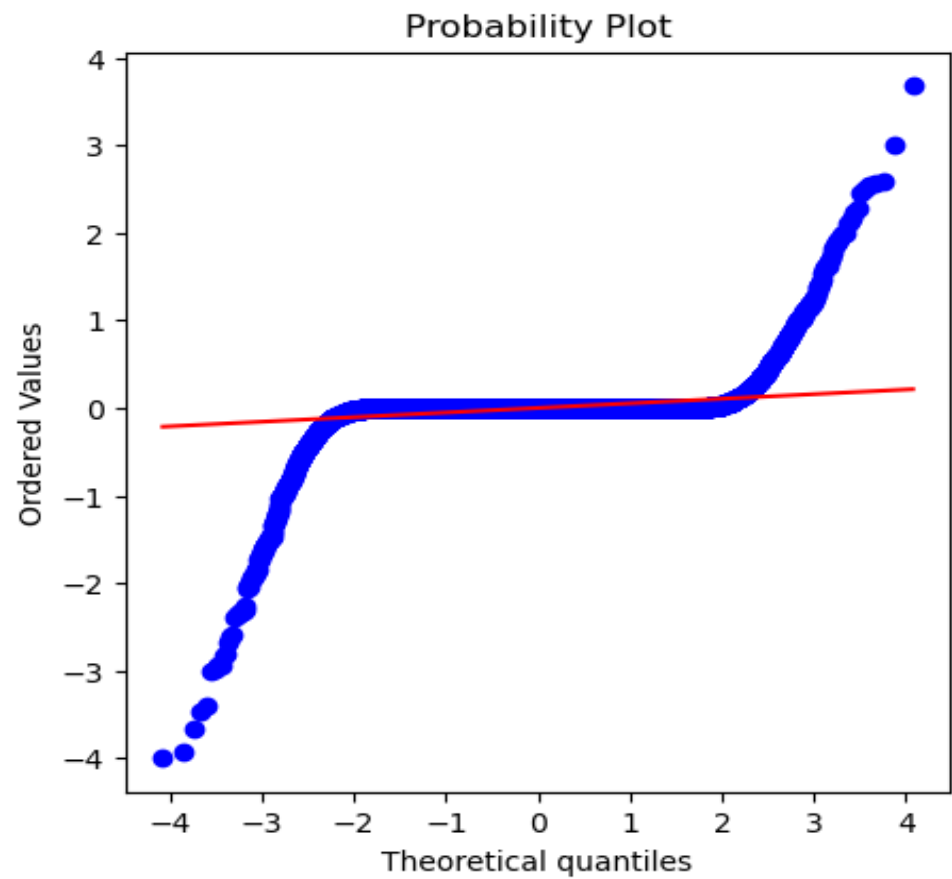
MSISDN number as Customer Id, Session frequency, Duration, Total Data, TCP, RTT and Throughput have been taken as features and Satisfaction Score as Target



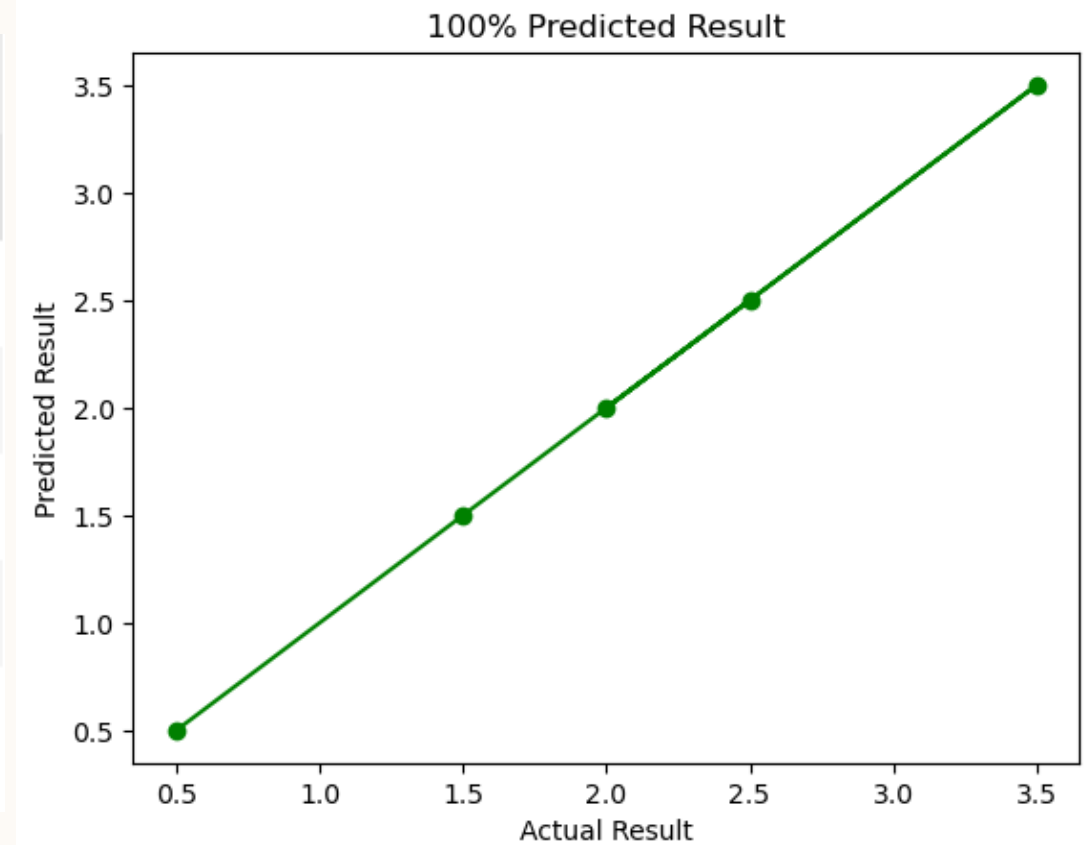
feature_importance



JointPlot for Feature Importance Analysis



final_prediction	
satisfaction_score	
2.5	2.5
2.0	2.0
3.5	3.5
1.5	1.5
0.5	0.5



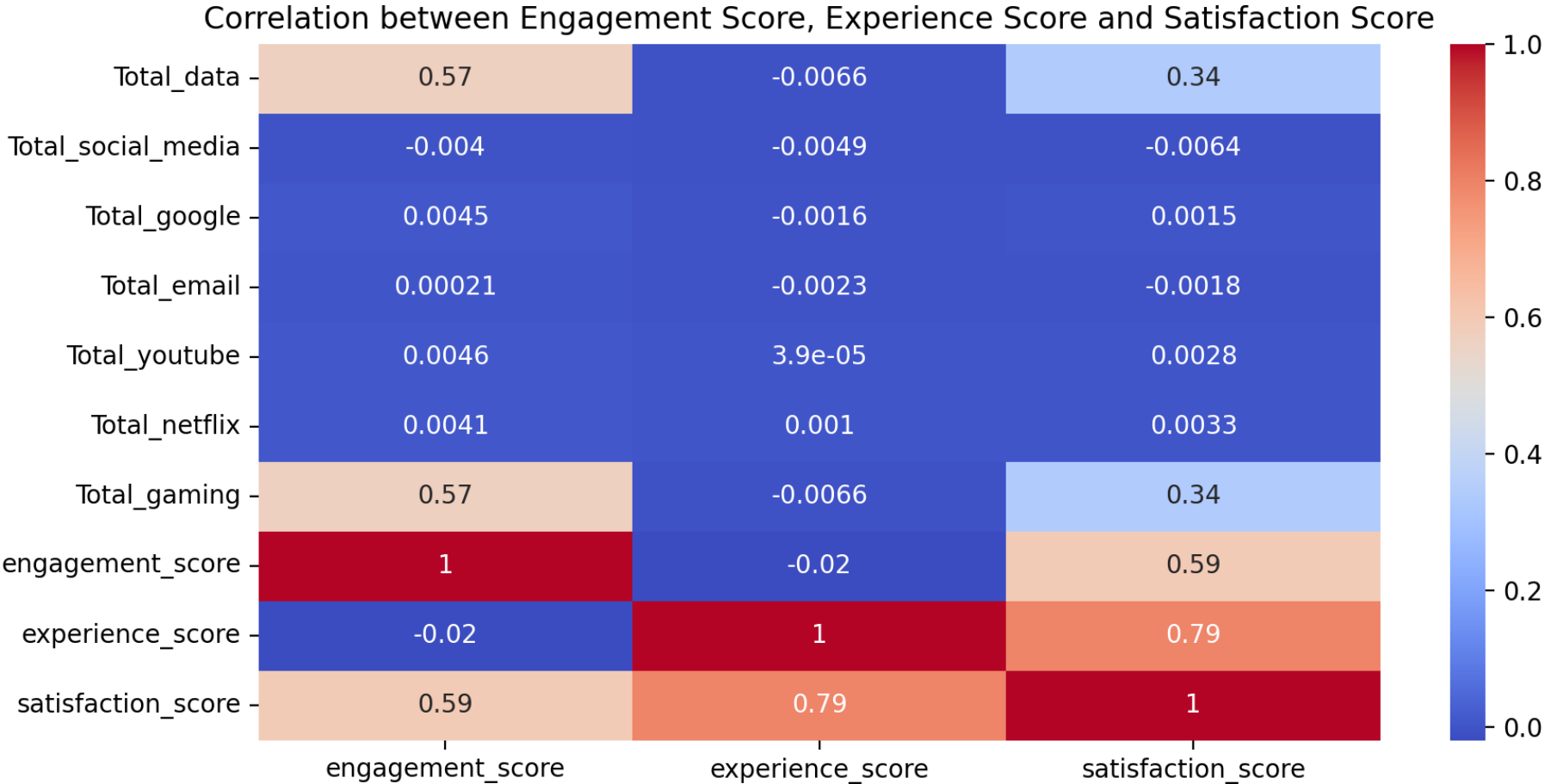
This is the probability plot of Actual and predicted and it shows constant result

Actual values of Satisfaction Score and predicted values of satisfaction score are exact same

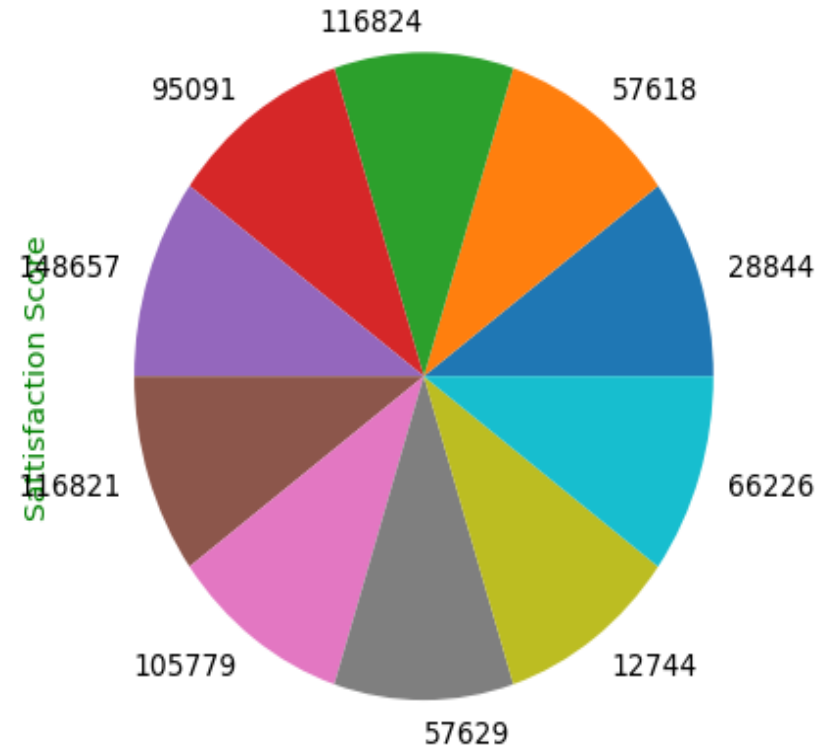
AREAS OF GROWTH

	MSISDN/Number	engagement_score	experience_score	satisfaction_score
0	3.360100e+10	5	0	2.5
1	3.360100e+10	4	0	2.0
2	3.360100e+10	7	0	3.5
3	3.360101e+10	3	0	1.5
4	3.360101e+10	1	0	0.5

Positive Correlation
Analysis between
Engagement Score,
Experience Score and
Satisfaction Score



PieChart for Top 10 Satisfied Customers



Top 10 satisfied customers are calculated on the basis of Satisfaction Score

PLAN FOR TELECOM COMPANY



PLANNING

User Overview
Analysis



MARKETING

Engagement and
Experience
metrics



DESIGN

Coordinate
business applications
like Gaming and
others



STRATEGY

Build and Improve
Quality of Service to
leverage the mobile
platforms



LAUNCH

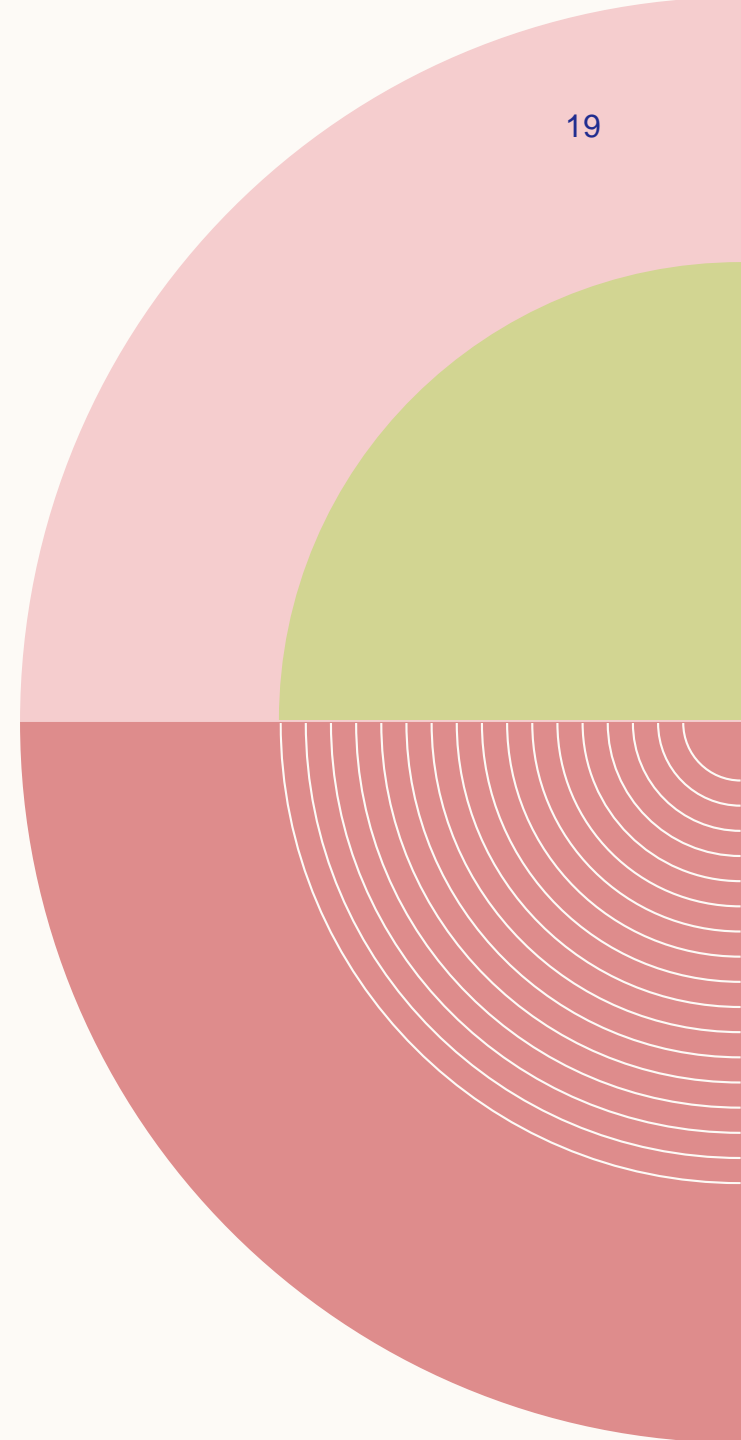
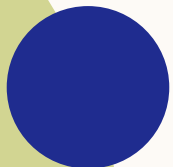
Deploy strategic
networks with
different clusters
based on metrics
score

AREAS OF FOCUS

- Social Media and Google are the two applications which are not being used by the customers according to this data or we can say that these 2 applications are rarely used because the relationship between these 2 applications and total DL & UL data is almost 0.

SUMMARY

Growth potential of the company is positive (based on this data).
As the model is also predicting positive results and the correlation between all 3 analysis is also positive, so the employer can buy this company.





THANK YOU

NITIMA SAIGAL