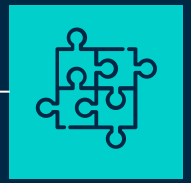
DEEPER INTO TellCo COMPANY

User Analytics in TellCo industry

BERINYUY BERTRAND MAINIMO

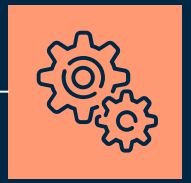
TABLE OF CONTENTS



01

PROBLEM & DATA OVERVIEW

-Business objective
-Analysis objective
-User overview analysis



02

OUR PROCESS

-User engagement analysis
-User experience analysis
-User satisfactory analysis

03

TARGET

-Results
-Recommendation
-Conclusion

UNDERSTANDING THE PROBLEM

CAN I BUY TellCo?

What is the current situation in TellCo industry?

Are there opportunities for growth of TellCo industry?

USER OVERVIEW ANALYSIS

DESCRIBING OUR DATA

TYPES OF VARIABLES

QUANTITATIVE VARIABLES

1. DISCRETE VARIABLES
-Number of xDR sessions
2.CONTINUOUS VARIABLES

-Session duration
-Total UL & DL
-Total data volume in each session

DATA TYPES

NUMERIC DATA TYPES

-Session duration
-Total ULs % DLs
-Total data volume in each session

-Number of xDR sessions

Key Insights on Univariate Analysis

1. The Data for each of the variables is symmetrical, the mean and median is similar.

2. Google has the highest mean compared to other application which shows that most of the customers spend more time and data in Google app, hence more advertisements can be done through Google app to increase the number of audience.

3. The standard deviation in Youtube data is large than the Mean which indicates that there is more variability of Customers spending data In Youtube App.

4. Youtube data and Netflix have the highest standard

deviation value which indicates greater spread in the data.

5. All the variables have a large range value which

indicates the greater dispersion in the data.

6. In most of the applications, most of the customers

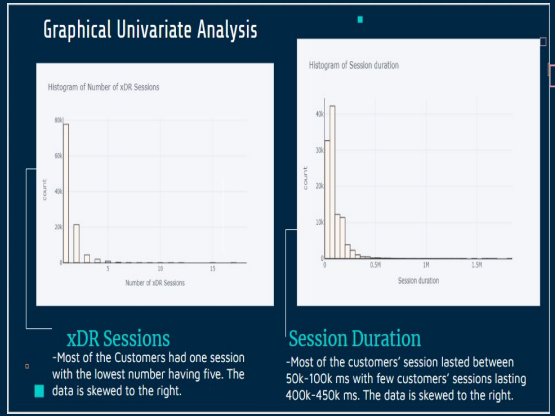
according to usage of the data, are in between the median and the upper quartile compared to between median and the lower quartile.

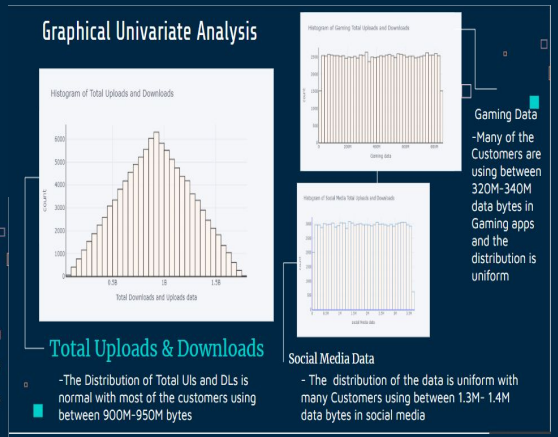
```
#Getting the statistical summary for the quantitative variables of the datamart
summary=df1.describe()
summary1=pd.DataFrame(summary)
summary1.to_csv("summary.csv")
summary1.drop(['Total Uploads','Total Downloads'],axis=1)
```

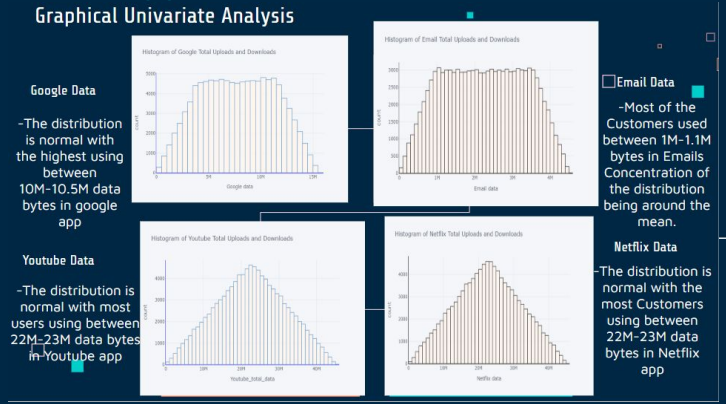
	MSISDN/Number	no.of xDR sessions	Session duration	Total UIs and DLs	Youtube_total_data	Google_total_data	Email_to
count	1.068560e+05	106856.000000	1.068560e+05	1.068560e+05	1.068560e+05	1.068560e+05	1.068
mean	4.511474e+10	1.393792	9.662390e+07	9.166177e+08	2.261432e+07	7.807186e+06	2.256
std	2.889423e+12	0.806022	8.395440e+07	3.445153e+08	9.247766e+06	3.518899e+06	1.070
min	3.360100e+10	1.000000	7.142988e+06	5.836229e+07	1.296550e+05	4.033000e+04	8.359
25%	3.365088e+10	1.000000	4.086000e+07	6.691826e+08	1.598535e+07	4.938135e+06	1.356
50%	3.366365e+10	1.000000	8.639990e+07	9.166781e+08	2.263187e+07	7.815112e+06	2.261
75%	3.368344e+10	2.000000	1.188538e+08	1.164285e+09	2.925176e+07	1.068372e+07	3.156
max	8.823971e+14	18.000000	1.823653e+09	1.780674e+09	4.519008e+07	1.552878e+07	4.518

GRAPHICAL UNIVARIATE ANALYSIS

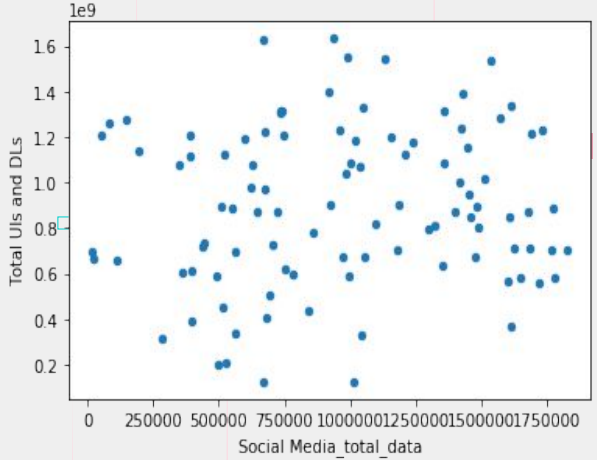
-Most of the users had one xDR session with the data distribution skewed to the right.
-Most of the users' session lasted between 50k-100k ms with few customers' sessions lasting 400k-450k ms. The data is skewed to the right.
-There is normal distribution in Total ULs % DLs with majority of users using 900-950M data bytes.
- There is normal distribution in usage of data for Google, Youtube, Email, Netflix apps and uniform distribution is exhibited in Social Media and Gaming apps.
-More users spend their data in Gaming apps, where we have a majority spending 740-760M data bytes.

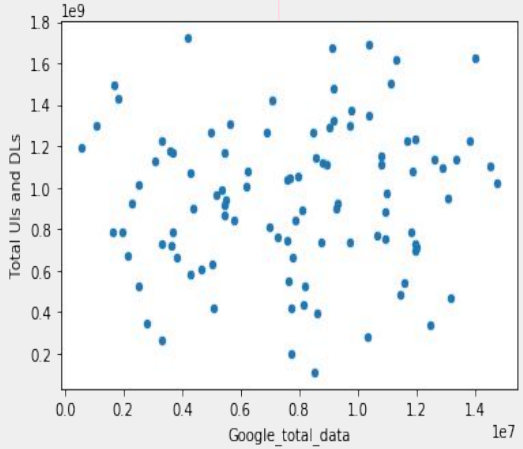


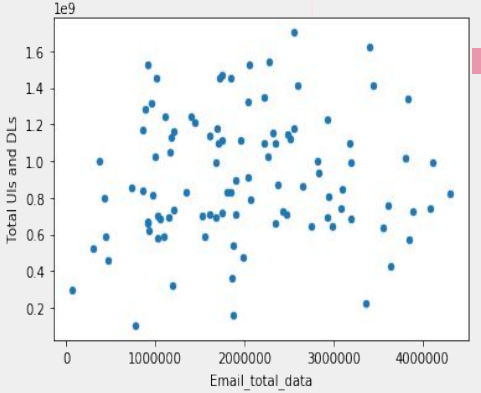


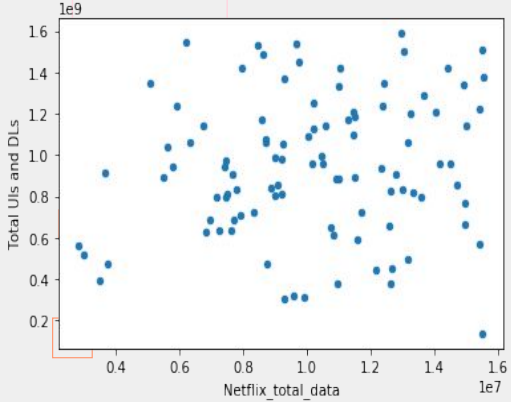


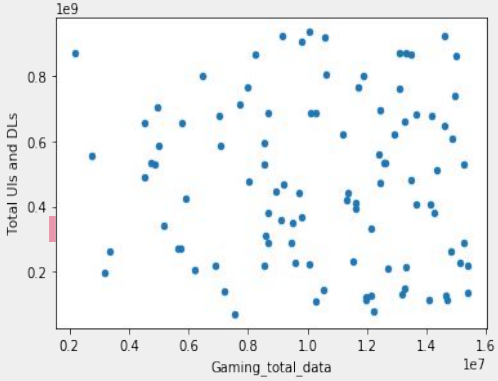
Bivariate Analysis Plots

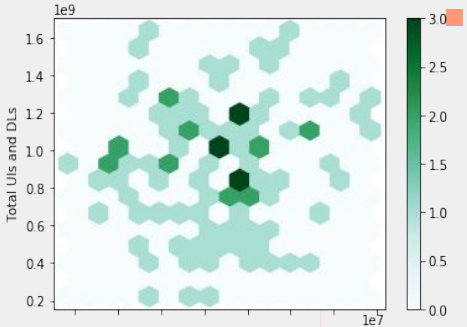










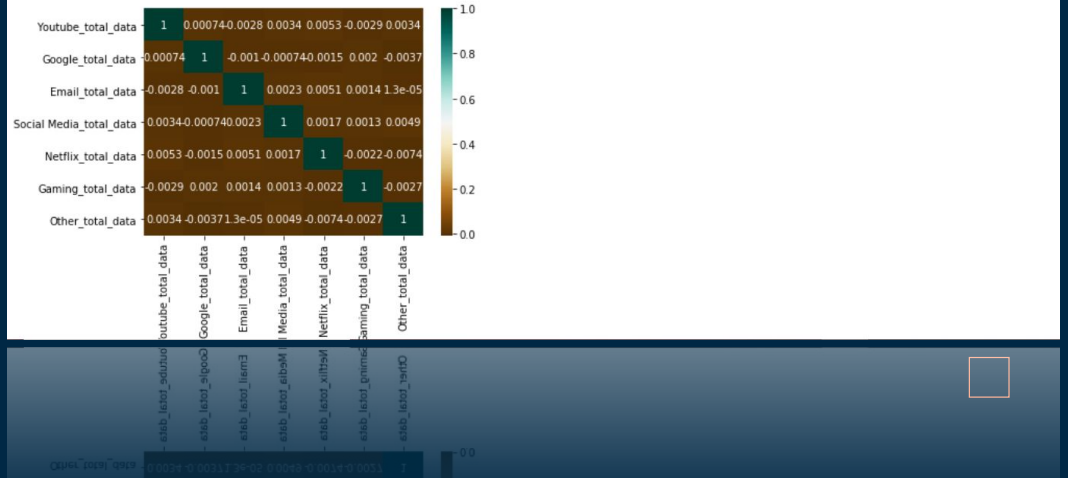


-There is Weak relationship between the data used in the application and the Total Upload,Download data.

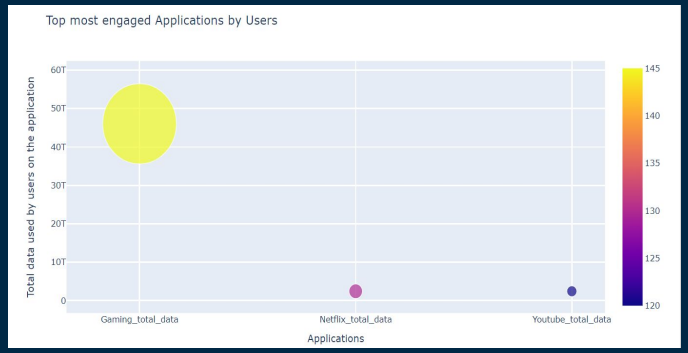
CORRELATION ANALYSIS

-The correlation between the variables is extremely weak with some of the variables exhibiting positive and other negative relationship.
- Some of the apps exhibiting relatively moderate positive relationship are, Youtube and Netflix, Netflix and Email application.

	Youtube_total_data	Google_total_data	Email_total_data	Social Media_total_data	Netflix_total_data	Gaming_total_data	Other_total_data
Youtube_total_data	1.000000	0.000743	-0.002775	0.003429	0.005332	-0.002878	0.003416
Google_total_data	0.000743	1.000000	-0.001005	-0.000738	-0.001452	0.002044	-0.003699
Email_total_data	-0.002775	-0.001005	1.000000	0.002323	0.005071	0.001405	0.000013
Social Media_total_data	0.003429	-0.000738	0.002323	1.000000	0.001700	0.001319	0.004892
Netflix_total_data	0.005332	-0.001452	0.005071	0.001700	1.000000	-0.002217	-0.007367
Gaming_total_data	-0.002878	0.002044	0.001405	0.001319	-0.002217	1.000000	-0.002729
Other_total_data	0.003416	-0.003699	0.000013	0.004892	-0.007367	-0.002729	1.000000



MOST USED APPLICATIONS



YOUTUBE APP

4.75% of the Total data is spent in Youtube app

NETFLIX APP

4.76% of the Total data is used in Netflix app

GAMING APPS

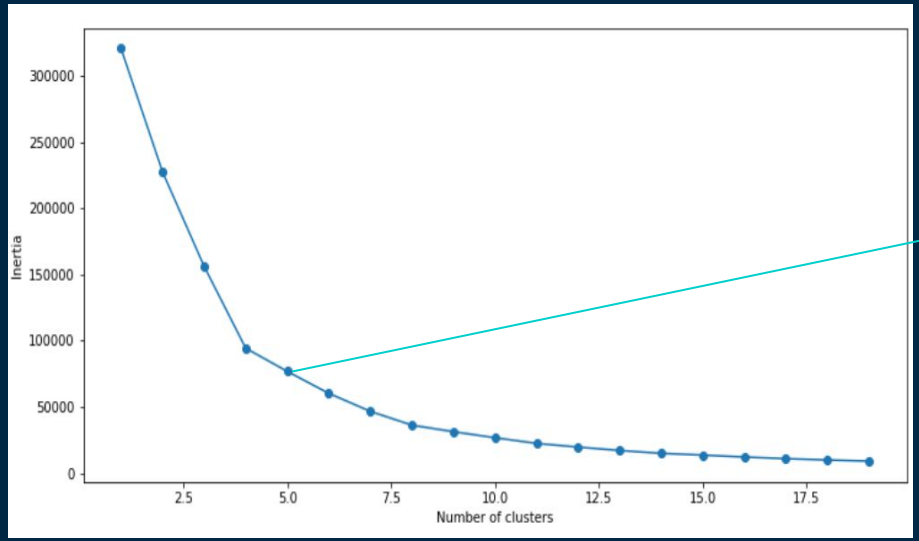
Gaming apps are the most engaged apps by users

USER ENGAGEMENT ANALYSIS

-Building & improving the QoS (Quality of Service) to leverage the mobile platforms and to get more users for the business is good but the success of the business would be determined by the user engagement and activity of the customers on available apps

- Most of the users had very few sessions but used much data bytes in the few sessions they had.

OPTIMIZED K-VALUE OF ENGAGEMENT CLUSTERS



5-8

-Using the Elbow method, we can see that the optimized k-value is between 5-8. The elbow is not clear and sharp which tells that our data was not clearly clustered.

USER EXPERIENCE ANALYTICS

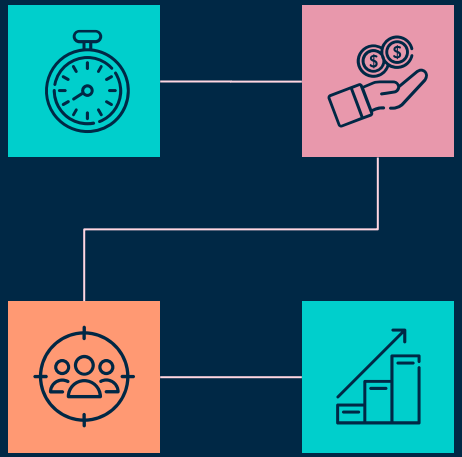
-It is so important to track and evaluate the customer's experience to optimize their products and services so that it meets the evolving user expectations, needs and acceptance.

MARS

Despite being red, Mars is a cold place

NEPTUNE

It's the farthest planet from the Sun



JUPITER

It's the biggest planet in the Solar System

SATURN

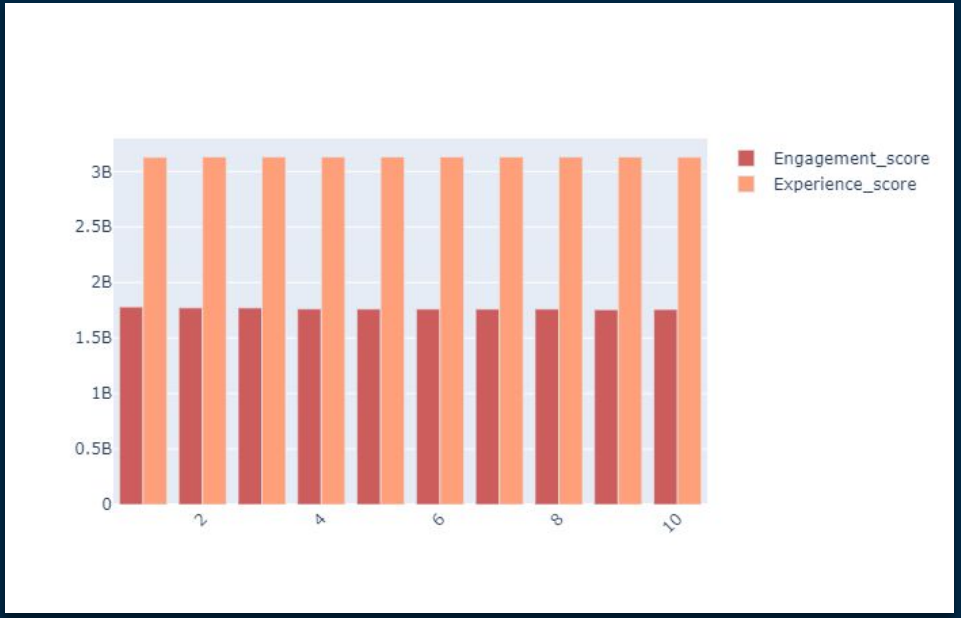
Saturn is the ringed one and a gas giant

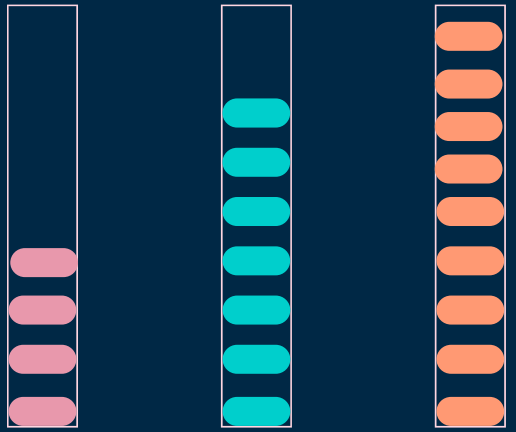
USER EXPERIENCE CLUSTERS

	COUNT	DESCRIPTION
CLUSTER 1	149595	This cluster had high numbers meaning that most of the customers were dissatisfied because of the unreliability of the network shown by high TCP retransmission.
CLUSTER 2	244	Few people showed a medium level of satisfaction from the company.
CLUSTER 3	162	Fewer people were fully satisfied according to their experience with TellCo company.

USER SATISFACTION ANALYSIS

-The Satisfaction of the customers heavily depends on the experience of the user.



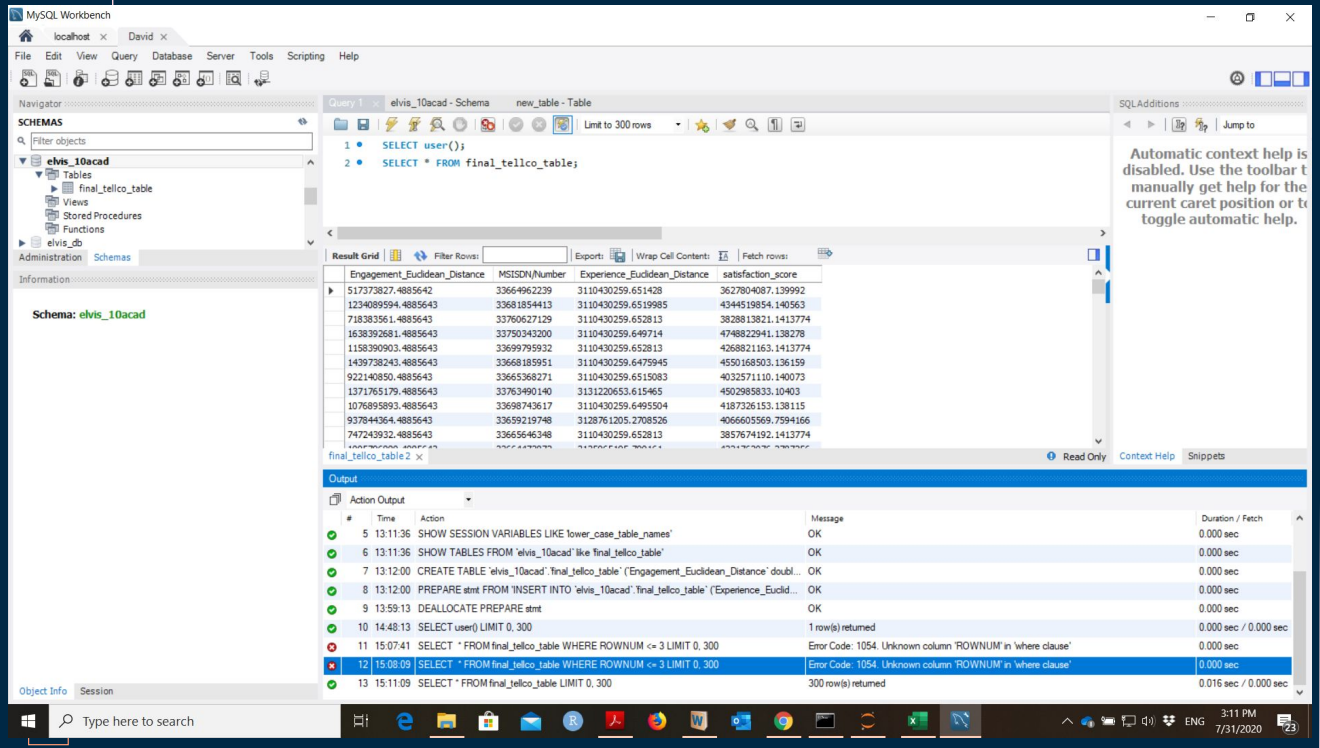


30%
Engagements score

80%
Experience score

95%
Satisfaction score

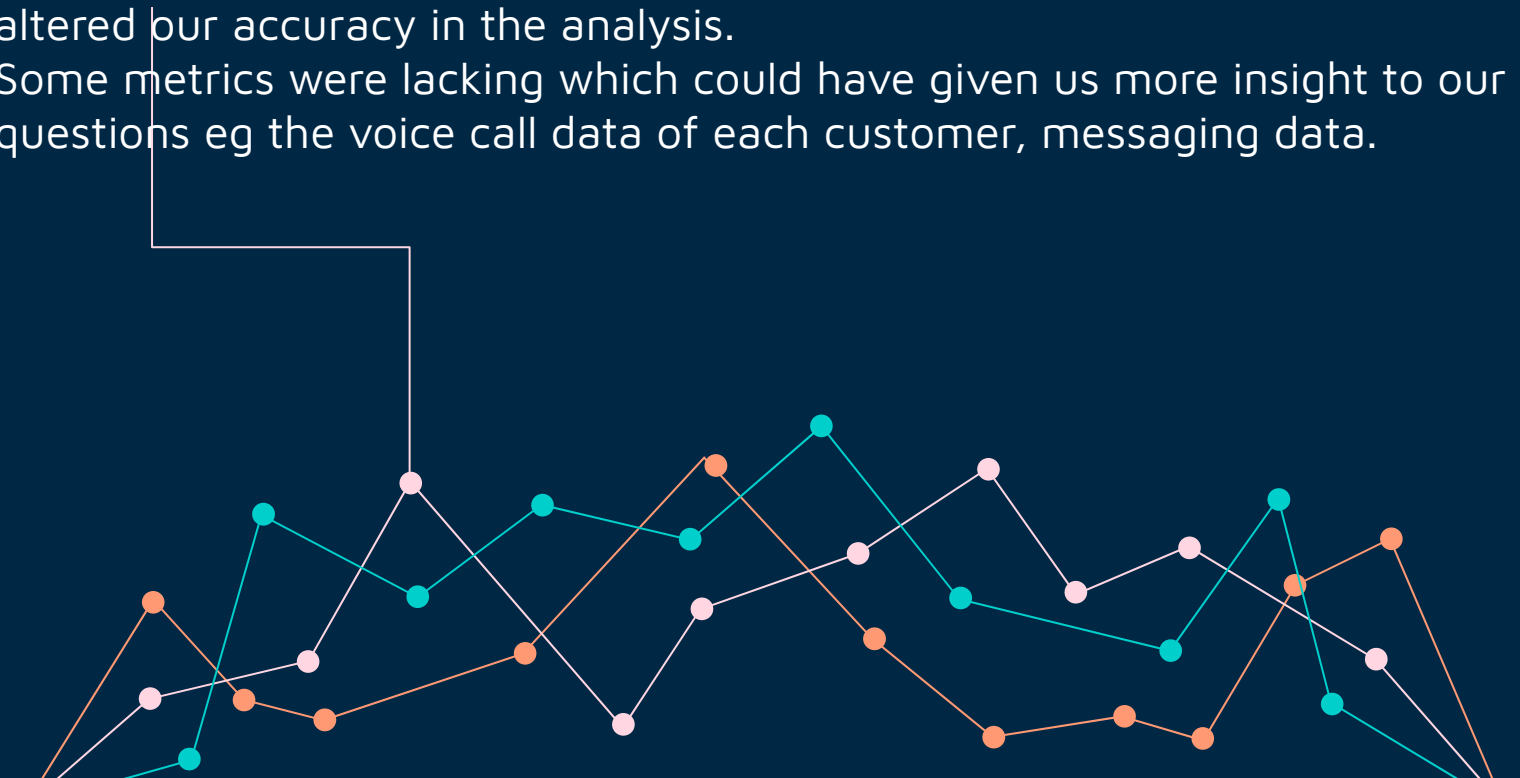
USER SATISFACTION,ENGAGEMENT AND EXPERIENCE TABLE



- Final table from the analysis showing the user's satisfaction,engagement and experience table.

LIMITATIONS TO THE ANALYSIS

- Many missing values in important columns which paralysed our analysis mostly in performing the clustering.
- There was a big number of outliers in our data which even after correcting altered our accuracy in the analysis.
- Some metrics were lacking which could have given us more insight to our questions eg the voice call data of each customer, messaging data.



RECOMMENDATIONS AND CONCLUSION

01

Increasing the experience score of the customers, by making better the network services to the customers

02

Create an engaging user experience by building and improving the Quality of Service especially in gaming applications

03

Getting closer to customers so as to be able to understand them well.

04

After Deeper dive into TellCo Company, it is clear that there is a great potential for TellCo to thrive more if the recommendations given are implemented, thus TellCo Company is a worthy business to invest in