NATIONAL INSTITUTE OF TECHNOLOGY CALICUT

FOUNDATIONS OF DATA ANALYTICS

EC6435D

COURSE END PROJECT REPORT

ANALYSIS OF VOICE CALL QUALITY CUSTOMER EXPERIENCE

Submitted By -

Ashwini Khandave

M220464ME

INDEX

SR. NO.	PARTICULARS	PAGE NO.
1.	Objective	3
2.	Introduction	3
3.	Data Sources	3
4.	Data Wrangling	4
5.	Exploratory Data Analysis 1. Call Quality Rating, 2. Call Quality Experience, 3. Number of Customer Feedback %, 4. Comparison Of Voice Call Quality Across TSP's, 5. Voice Call Quality Rating Throughout the Year 2020 and, 6. Call Quality Rating State Wise.	5 5 6 7 8 8 9
6.	Conclusion	11
7.	References	11

ANALYSIS OF VOICE CALL QUALITY CUSTOMER EXPERIENCE

Objective

Analysis of Voice Call Quality Customer Experience across TSP's - IDEA, AIRTEL, and RJIO.

Introduction

Voice calling is the ability to contact and converse with people in real-time with a telephone/ mobile. Voice interactions give more insight into conversations than any other communication channels. With voice communications, a person's tone, inflection, and emotions come through, providing greater context, and meaning.

To get the idea of which Telecom Service Provider in India provided best service in 2020 the analysis of Voice Call Quality from customer feedbacks for the year 2020 is done. The data is taken from Open Government Data Platform India, a Gov. of India's official website. The analysis is done for Idea, Airtel, and RJio TSP.

Python, a high-level programming language and SQL, Structured Query Language are used to deal with large dataset, analyse it and communicate the analysis via Exploratory Data Analysis.

Data Sources

The required data was taken from the following sources:

1. https://data.gov.in/catalog/voice-call-quality-customer-experience?page=1# - Voice call quality customer experience data of 2020.
2. https://geographicalanalysis.com/download-free-india-shapefile-including-kashmir-and-ladakh/ - Indian state shape file.

Data Wrangling

Voice Call Quality:

1. Original Data
    - The data about Voice Call Quality for each month of 2020 is in separate .csv files. Each row is representing customer feedback.
    - Size of the data frame = 168481 rows × 9 columns
    - The data includes the following columns:
    - data.columns = ['Operator', 'In Out Travelling', 'Network Type', 'Rating', 'Call Drop Category', 'Latitude', 'Longitude', 'State Name']
2. Derived Data

    From the original data about Voice Call Quality, data calculated is what follows:

    - Mean rating of respective mode across all TSP's and for individual TSP.
    - Count and percentage of each Call_Drop_Category for all TSP's and for individual TSP.
    - Count and percentage of no. of feedbacks corresponding to the TSP.
    - Mean of rating with respect to respective TSP throughout the year.
    - Mean of rating with respect to respective TSP month wise.
    - Mean of rating with respect to State Names throughout the year.

Data Loading

The data is transformed from the csv file to a data frame, one additional column is created and corresponding month is the value assigned. It is then appended in one list. New data frame is created by concatenating the list at axis=0.

Data cleaning

- Columns renamed to replace “space” in names with "_" .
- Filtered data by removing rows containing “NAN” and “unknown” values.
- Analysis is of only 3 TSP's, deleted the rows corresponding to remaining operators.
- Reset Index
- Renamed the index column to "Customer_ID".

Filtered data size = 75228 rows × 10 columns

	Customer_ID	Operator	In_Out_Travelling	Network_Type	Rating	Call_Drop_Category	Latitude	Longitude	State_Name	Month
0	1	RJio	Outdoor	4G	1	Call Dropped	28.237968	76.832227	Rajasthan	January
1	2	RJio	Indoor	4G	4	Satisfactory	22.673484	88.490379	West Bengal	January
2	3	Airtel	Indoor	4G	5	Satisfactory	12.879084	77.724270	Karnataka	January
3	4	RJio	Indoor	4G	5	Satisfactory	25.281607	83.123418	Uttar Pradesh	January
4	5	RJio	Indoor	4G	5	Satisfactory	25.284936	83.127254	Uttar Pradesh	January

Table 1. Cleaned and Filtered Data

Exploratory Data Analysis

Total 6 plots are plotted to analyse the data of Voice Call Quality:

1. Call Quality Rating,
2. Call Quality Experience,
3. Number of Customer Feedback %,
4. Comparison Of Voice Call Quality Across TSP's,
5. Voice Call Quality Rating Throughout the Year 2020 and,
6. Call Quality Rating State Wise.

For each plot:

1. the required data is calculated and
2. then exploratory data visualisation is done.

1. Call Quality Rating

Voice call quality rating mean is taken for each mode corresponding to all TSP's and individual TSP.

Library used -

- Matplotlib
- Palettable - python color library used for plots plotted with Matplotlib.

Visualization - Donut plot (Pie plot superimposed with a white circle.)

Call Quality Rating Data:

All Operators

	Indoor	Travelling	Outdoor
Rating	3.7	3.8	3.4

Idea

	Indoor	Travelling	Outdoor
Rating	4.5	4.6	4.3

Airtel

	Indoor	Travelling	Outdoor
Rating	3.6	3.5	2.8

RJio

	Indoor	Travelling	Outdoor
Rating	3.4	3.6	3.4

Table 2. Call Quality Rating Data

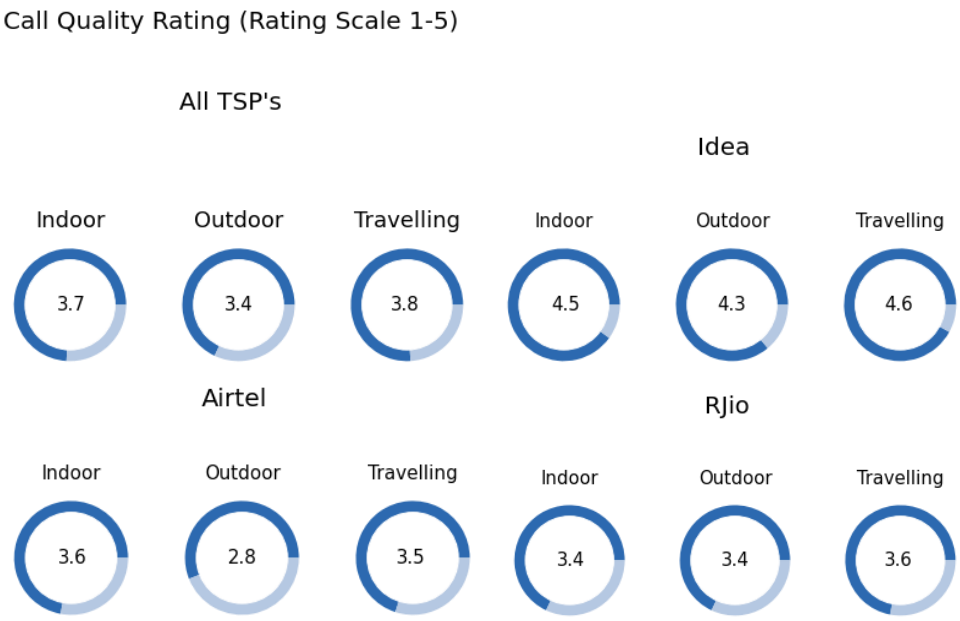


Fig 1. Call Quality Rating with respect to modes and TSP's

Call quality rating is high in travelling mode amongst all modes for all TSP's as well as individual TSP's. Idea is having good rating for all the modes among all the TSP's.

2. Call Quality Experience

Call quality experience of customers is visualised by taking the % of entries of call drop category across all TSP's and individual ones.

Library Used - Matplotlib, Plalettable

Visualization by - Pie plot

Call Quality Experience Data:

	Call_Drop_Category	All_TSP's
0	Satisfactory	52211
1	Poor Voice Quality	15618
2	Call Dropped	7399

	Call_Drop_Category	Idea
0	Satisfactory	14179
1	Poor Voice Quality	1248
2	Call Dropped	686

	Call_Drop_Category	Airtel
0	Satisfactory	15718
1	Poor Voice Quality	6397
2	Call Dropped	3563

	Call_Drop_Category	RJio
0	Satisfactory	22314
1	Poor Voice Quality	7973
2	Call Dropped	3150

Table 3. Call Quality Experience

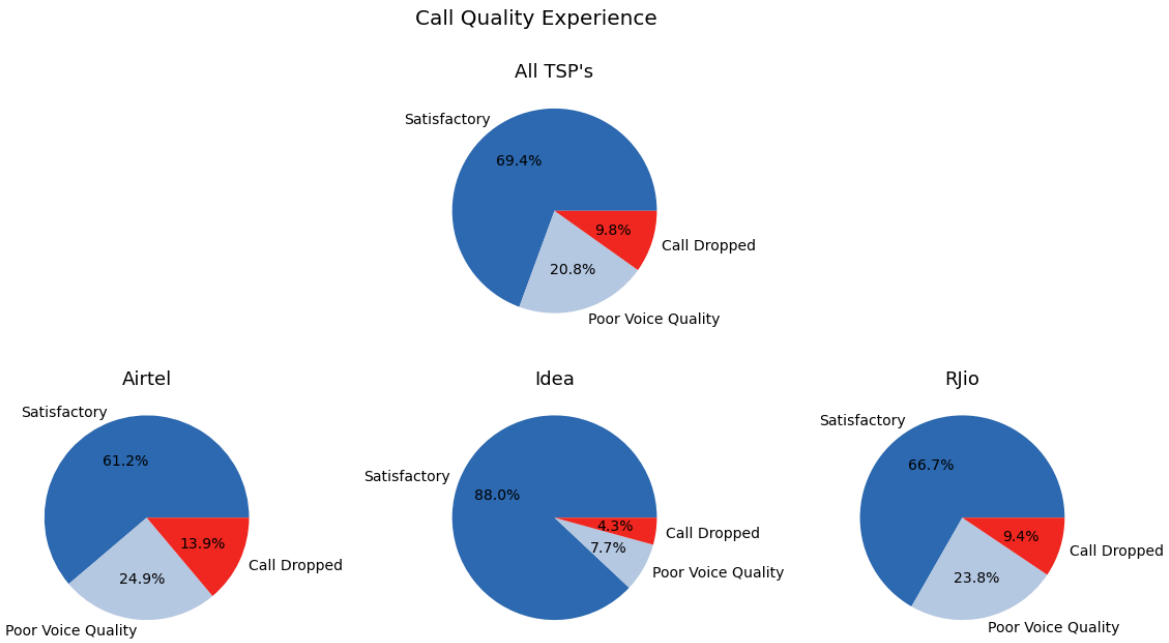


Fig 2. Call Quality Experience for all TSP's and respective TSP's.

Customers are satisfied with voice call quality provided by Idea operator. Airtel is having highest % of call dropped.

3. Number of Customer Feedback

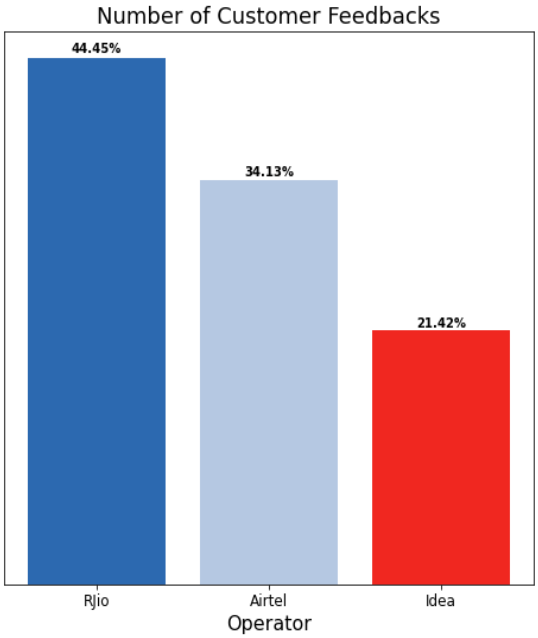
Library Used - Matplotlib, Plalettable

Visualization by - Bar plot

Percentage of value count is calculated of all operators.

	No_Of_Feedbacks	Percentage
RJio	33437	44.45
Airtel	25678	34.13
Idea	16113	21.42

Table 4. No. of Customer Feedbacks w.r.t TSP's.



The data collected from customer is dominated by RJio in numbers.

Fig 3. No. of Customer Feedbacks w.r.t TSP's.

4. Comparison of Voice Call Quality Across TSP's

Average rating of each operator over the year data is visualised.

Voice call quality of Idea service provider is best with respect to others.

Operator	Avg_Rating
Idea	4.43
Airtel	3.40
RJio	3.38

Table 5. Average rating with respect to operators.

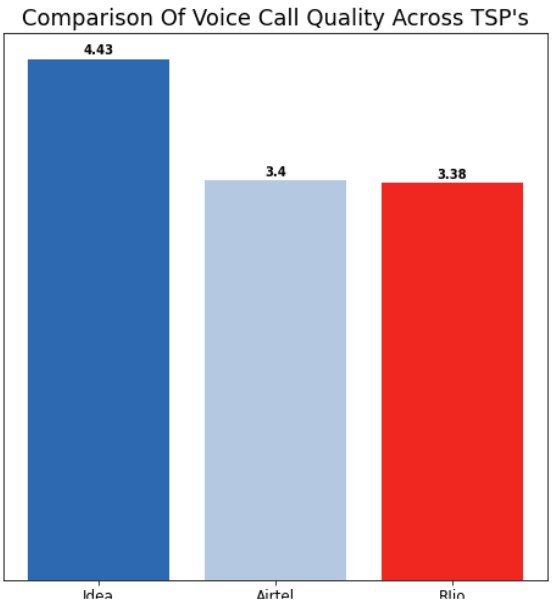


Fig 4. Average rating with respect to operators.

5. Voice Call Quality Rating Throughout Year 2020

The average rating of every month for each operator is visualised. The visualisation shows the performance of the operator over the months.

Library – Plotly.express

Visualisation - Interactive line plot.

	Month	Airtel	Idea	RJio
0	January	3.49	2.72	3.31
1	February	3.42	3.13	3.29
2	March	3.64	2.40	3.24
3	April	3.60	2.99	3.09
4	May	3.39	2.77	3.53

Table 6. Voice Call Quality Rating Throughout the Year 2020

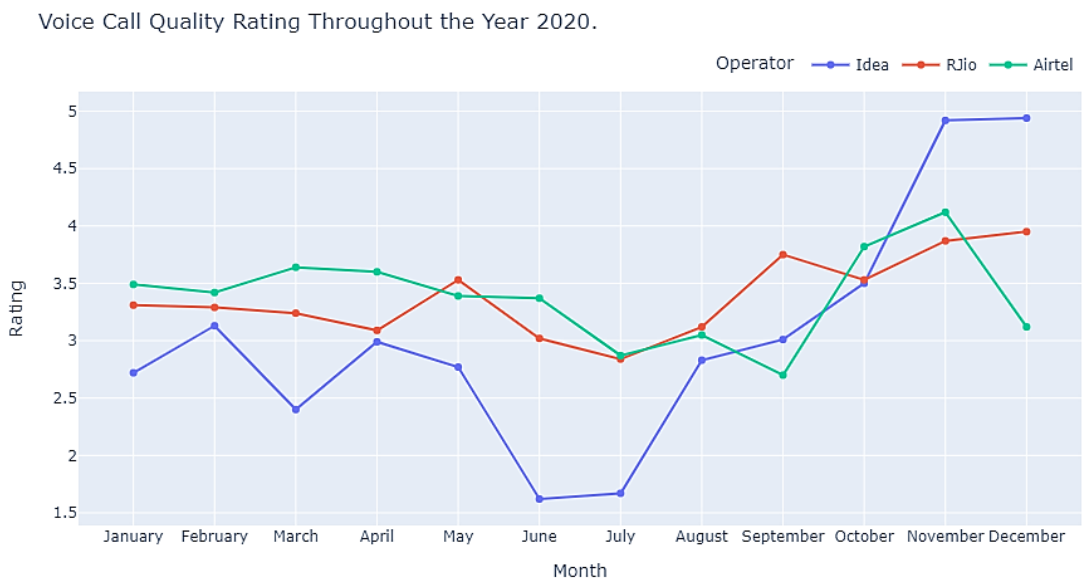


Fig 5. Interactive line plot – Voice call quality rating over the year 2020.

Operator Idea can be seen to have improved its service as per the customer feedbacks. Rating of RJio is consistent throughout the year.

6. Call Quality Rating State Wise

Average rating with respect to state is visualised.

The data is retrieved with sqlite3. For this the main data frame is converted into sql database ('Voice_Call_Quality.db'). New table is created with "CREATE TABLE" query, "Customer_ID" is set as the primary key. The data is then transferred to the new table created earlier and is stored in 'Voice_Call_Quality.db' file. The table created is in 3NF.

Library - Geopandas (reads the shape file)

Visualisation - Choropleth map (A thematic map that visualizes geographical areas or regions clearly divided with colors, shades, or patterns in relation to a given variable.)

A separate table for each state is created with 'Customer_ID' as the primary key and required data is transferred. The table is in 3rd NF.

```
#  data from Maharashtra table
c.execute("select * from Maharashtra")
r = c.fetchall()[:5]
for i in r:
  print(i)
conn.commit()
```

```
(6, 'RJio', 5, 18.66099833, 73.895535)
(7, 'RJio', 5, 18.66099833, 73.895535)
(9, 'Airtel', 2, 19.26543333, 73.12978167)
(10, 'RJio', 5, 18.66089833, 73.895555)
(16, 'RJio', 3, 18.45305378, 73.88656172)
```

Fig 6. Sql query for retriviewing the data from “Maharashtra” table.

Refer to the Voice_Call_Quality_Analysis.ipynb for sql query for obtaining avg. rating for each state.

Avg. rating is calculated with respect to state and state names are replaced as per the Indian State Shape File and the data is merged into new data frame.

	Name	Type	geometry	State_Name	Avg_Rating
0	West Bengal	State	MULTIPOLYGON (((88.01861 21.57278, 88.01889 21...	West Bengal	3.800971
1	Andaman & Nicobar	Union Territory	MULTIPOLYGON (((92.90124 12.91071, 92.90157 12...	Andaman & Nicobar	2.375000
2	Chandigarh	Union Territory	POLYGON ((76.77232 30.79420, 76.77286 30.79343...	Chandigarh	1.533333
3	Daman and Diu and Dadra and Nagar Haveli	Union Territory	MULTIPOLYGON (((72.96339 20.33227, 72.96390 20...	Daman and Diu and Dadra and Nagar Haveli	2.500000
4	Delhi	Union Territory	POLYGON ((77.10591 28.87005, 77.10668 28.86989...	Delhi	2.162712

Table 7. Merged Indian State Shape geometry data with avg. rating of states.

The data of some states is not available, those states are filled with light grey colour.

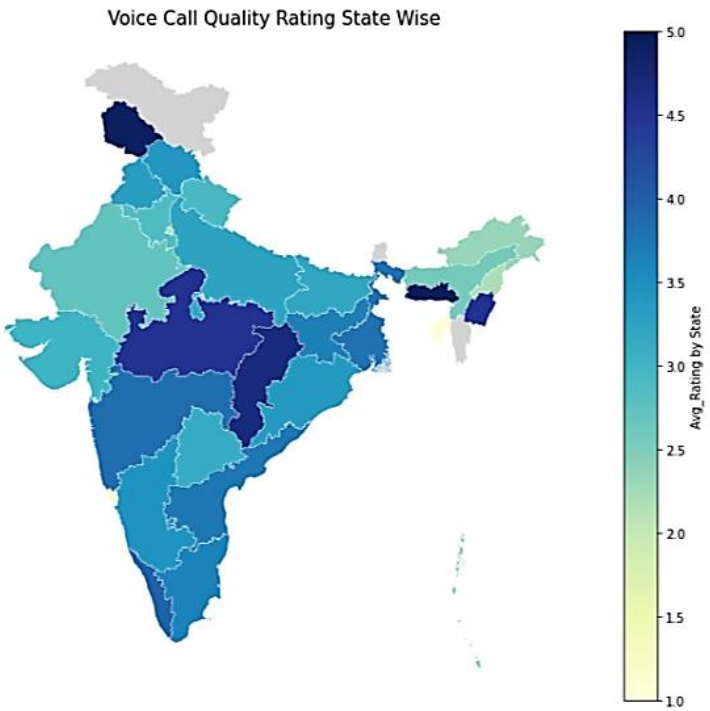


Fig 7. Choropleth map of India with states filled with colours corresponding the avg. rating.

The choropleth map visualisation inference:

- The voice call quality rating is low in East States of India.
- Middle states have the best ratings.
- South rating is moderate.
- Kashmir is having the best rating.

Conclusion

With visuals obtained we can clearly state TSP - Idea was able to satisfy customers with its good voice call quality. Airtel needs to improve its service to stay in league. RJio is consistent with its voice quality service. Since, the mean of rating was taken into consideration the analysis stays true to its nature.

References

1. https://data.gov.in/catalog/voice-call-quality-customer-experience?page=1#
2. https://trai.gov.in/
3. https://jiffyclub.github.io/palettable/
4. https://geographicalanalysis.com/download-free-india-shapefile-including-kashmir-and-ladakh/