## **Introduction**

This dataset is about the Indian Voice Call Quality Customer Experience. It has Customers Feedback Captured using TRAI MyCall App. Customers rate their perceived experience about voice call quality in real time and help TRAI gather customer experience data along with other network data.

**Released Under**: National Data Sharing and Accessibility Policy (NDSAP)

**Contributor:**

Ministry of Communications

Department of Telecommunications (DOT)

Telecom Regulatory Authority of India (TRAI)

## **Dataset**

The attached data folder contains multiple files, each file talks about data of an individual month and response from each user depicts an individual row. Each file has 8 columns as shown below:

1. Operator: Telecom operator which a particular person is using

2. Inout\_travelling: Whether the rating is for indoor or outdoor or travelling

3. Network\_type: Type of network. Eg: 4G, etc.

4. Rating: Rating provided by user

5. Calldrop\_category: Category in which an user classified a calldrop

6. Latitude: Current latitude of a user

7. Longitude: Current longitude of a user

8. State\_name: State of a user

## **Business Problems/ Questions**

1. Which operator is best rated in each state overall and in terms of indoor and outdoor

traveling?

2. Is there any correlation between network type and rating?

3. Is call drop correlated with the rating?

4. Where do we see more call drops - outside or inside? If it is outside, is it right to say that

there is a correlation? (can you do a hypothesis testing to verify this hypothesis?)

## **Methodology**

**Tools Used**

For this analysis, I utilized Python along with its data analysis and visualization libraries, including Pandas, NumPy, Matplotlib, Seaborn, and SciPy. Additionally, Power BI was used for creating interactive visualizations and dashboards.

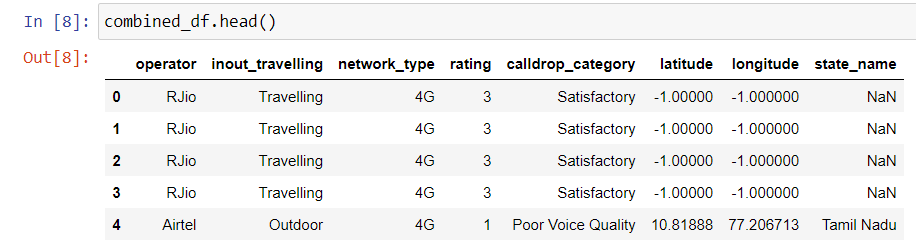
**Data Loading and Preprocessing**

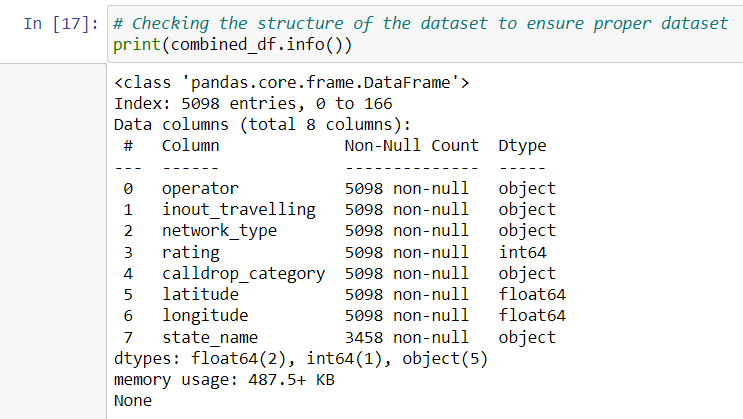
Data Loading:

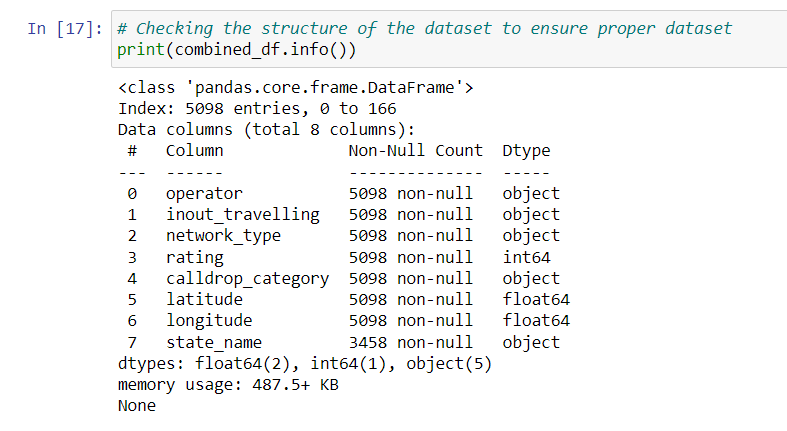
The dataset was spread across multiple CSV files, each representing data from January to October 2023. The files were named January\_MyCall\_2023.csv, February\_MyCall\_2023.csv, and so on. To combine these files into a single DataFrame for analysis, we used the Pandas library.

# **Exploratory Data Analysis (EDA)**

**Data Overview:**





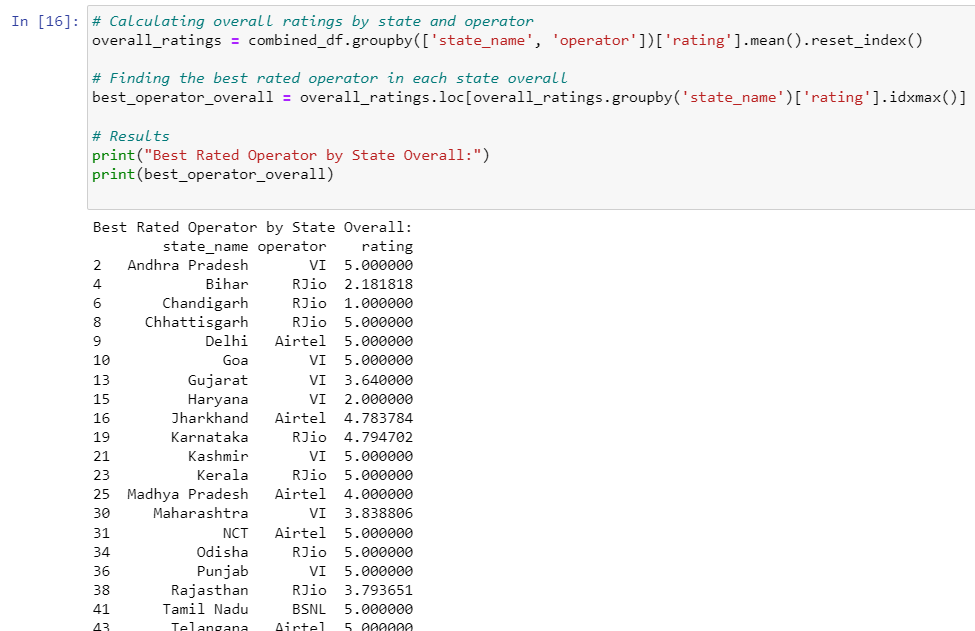


# **Analysis and Findings**

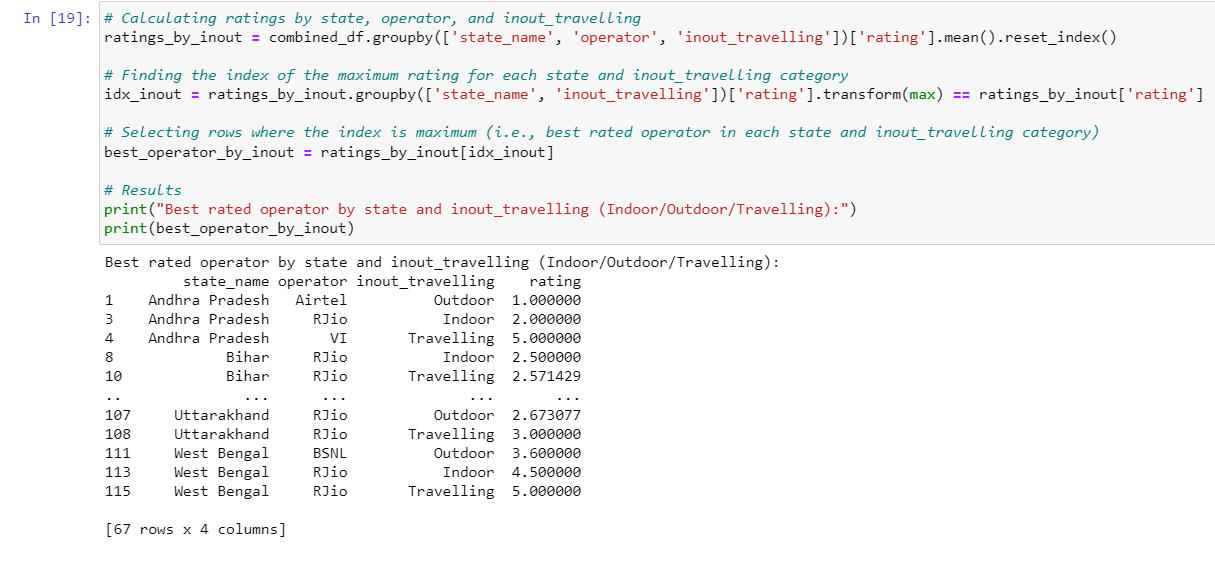
For each question, I include the following sub-sections:

* **Question**: The business problem is addressed here.
* **Methodology**: The approach taken to answer the question is mentioned here and the codes used for the analysis are shown here.
* **Output and Visualizations**: The output and visualizations are presented here.
* **Interpretation**: Possible interpretation of the results.

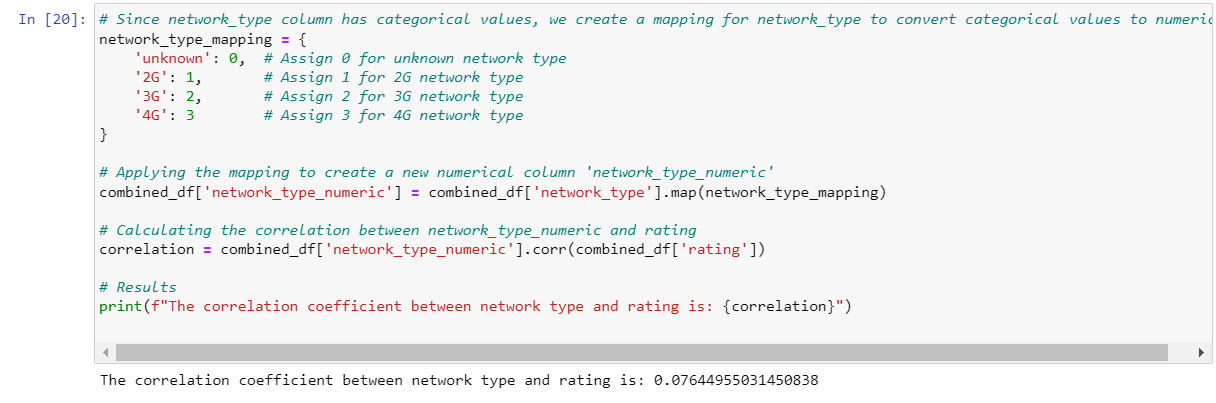
#### Question 1: Which operator is best rated in each state overall?



#### Question 2. Which operator is best rated in each state in terms of indoor, outdoor and travelling environment?

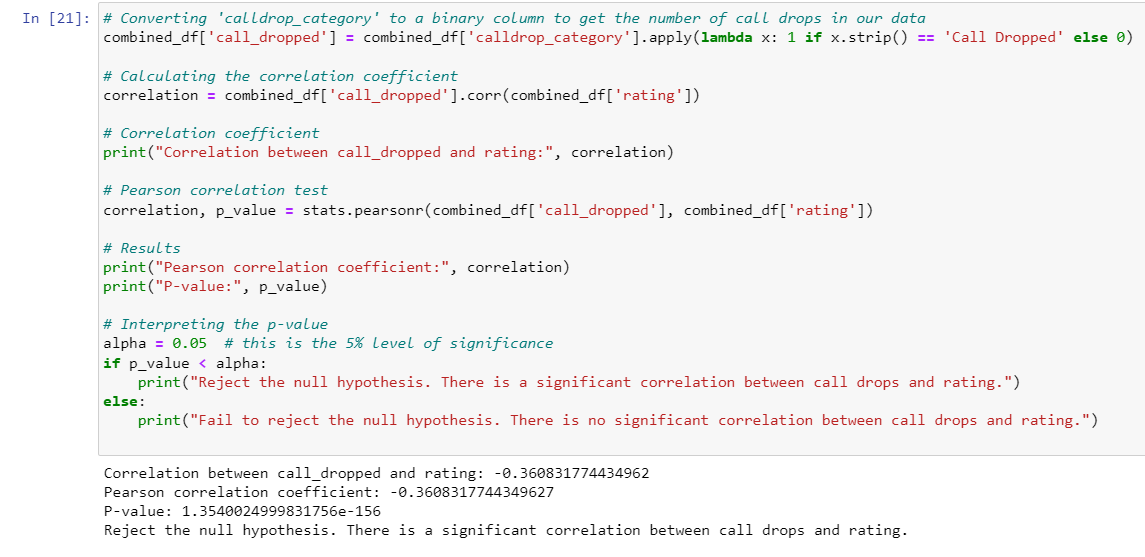


### *Question 3*. *Is there any correlation between network type and rating?*



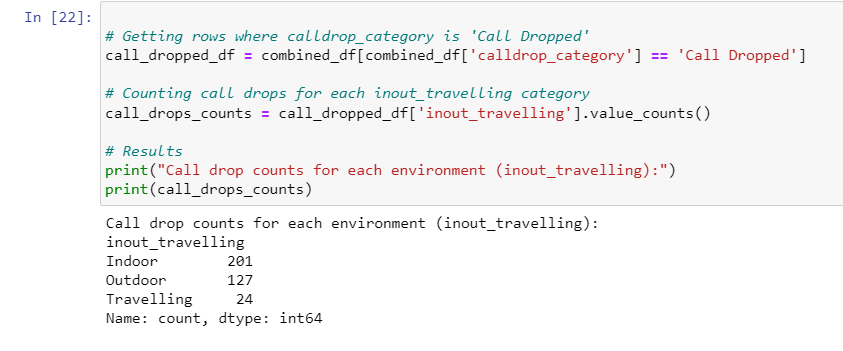
Interpretation: The correlation coefficient indicates a very weak positive correlation between network type and rating.

### *Question 4. Is call drop correlated with the rating?*



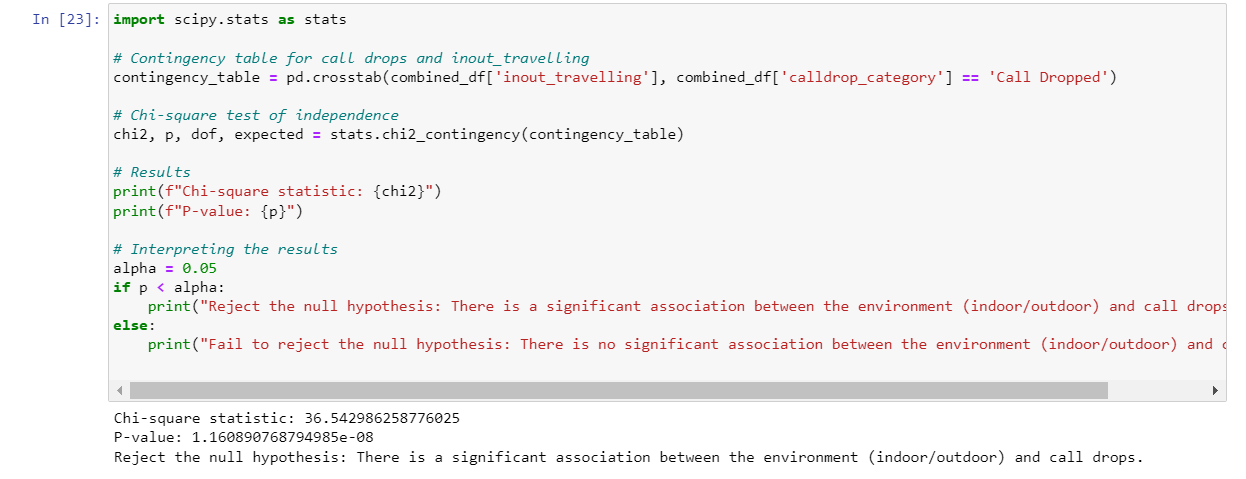
Interpretation: There is a moderate negative correlation between call drops and rating, indicating that higher call drop rates are associated with lower ratings.

#### Question 5. Where do we see more call drops - outside or inside?



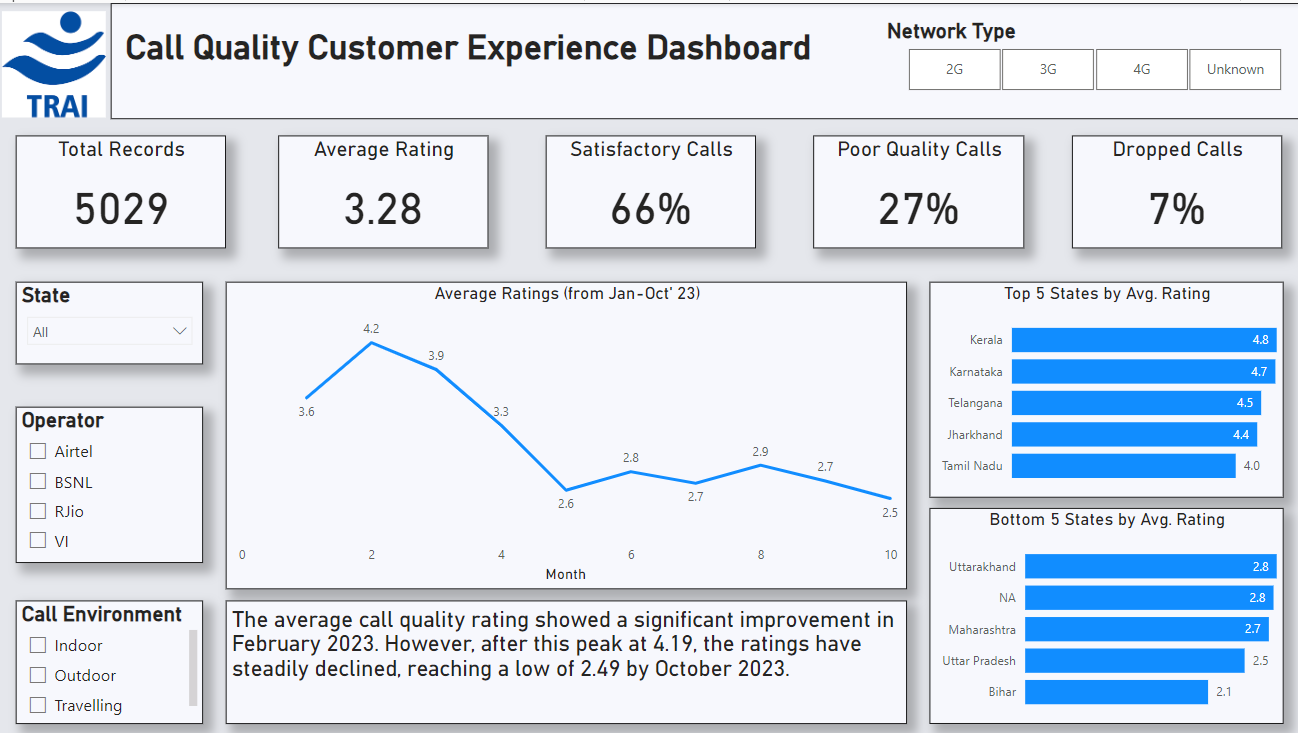
Interpretation: There are more call drops while users are indoor than outdoor.

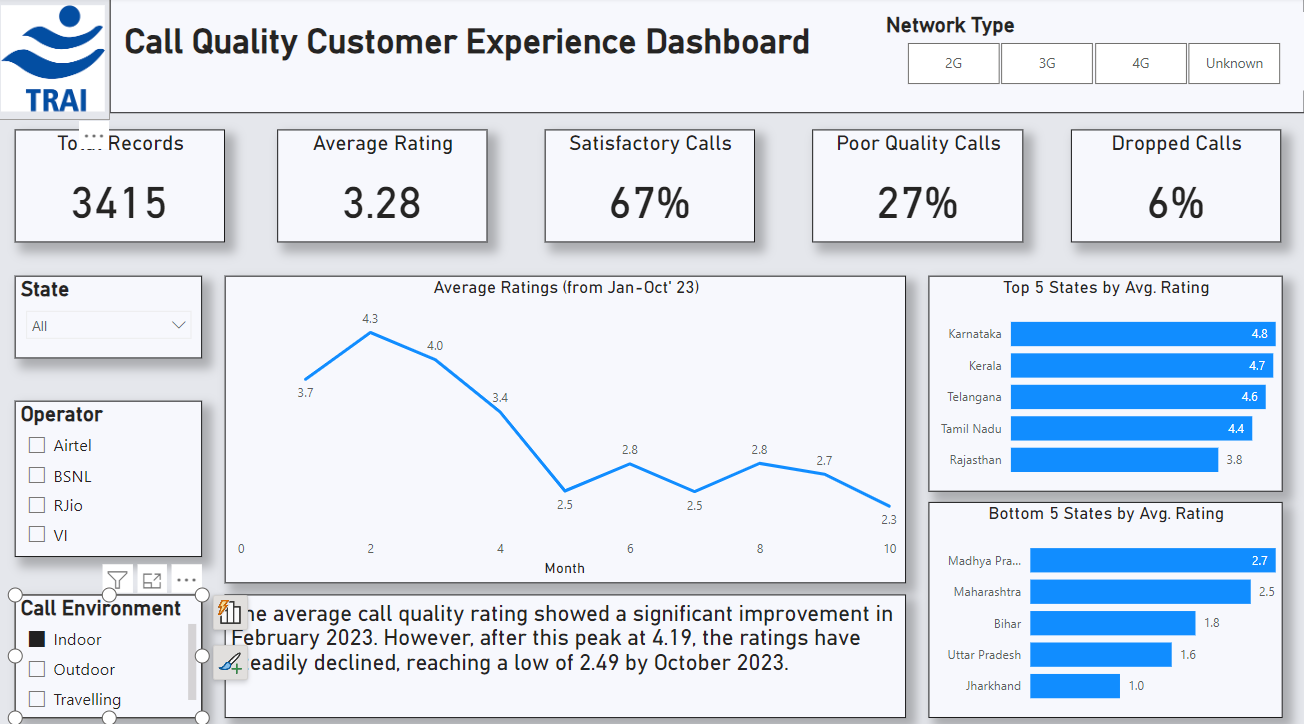
#### **Question 6: Is there any association between the environment (indoor/outdoor) and call drops?**

Interpretation: The results indicate a significant association between the environment (whether user is indoor, outdoor or travelling) and call drops.

**Visualisation (in Power BI)**

There were states with very few records, so those records have been deleted and states with more than 30 records are included only



This is the entire dashboard. I have mentioned 5 KPIs.  
  


Indoor call quality experience is better than outdoor call quality experience. There are more frequent call drops when users are outside.

