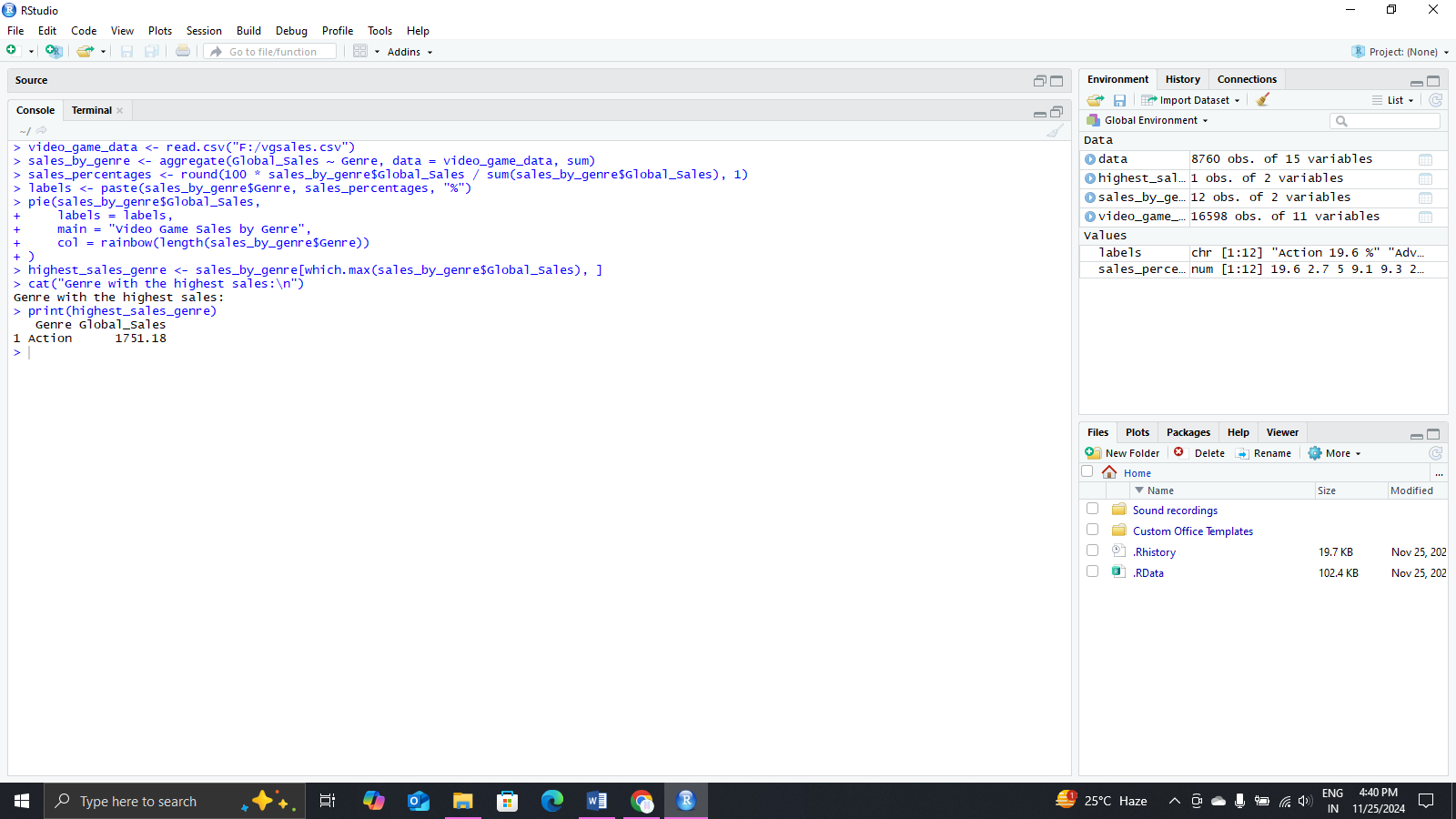
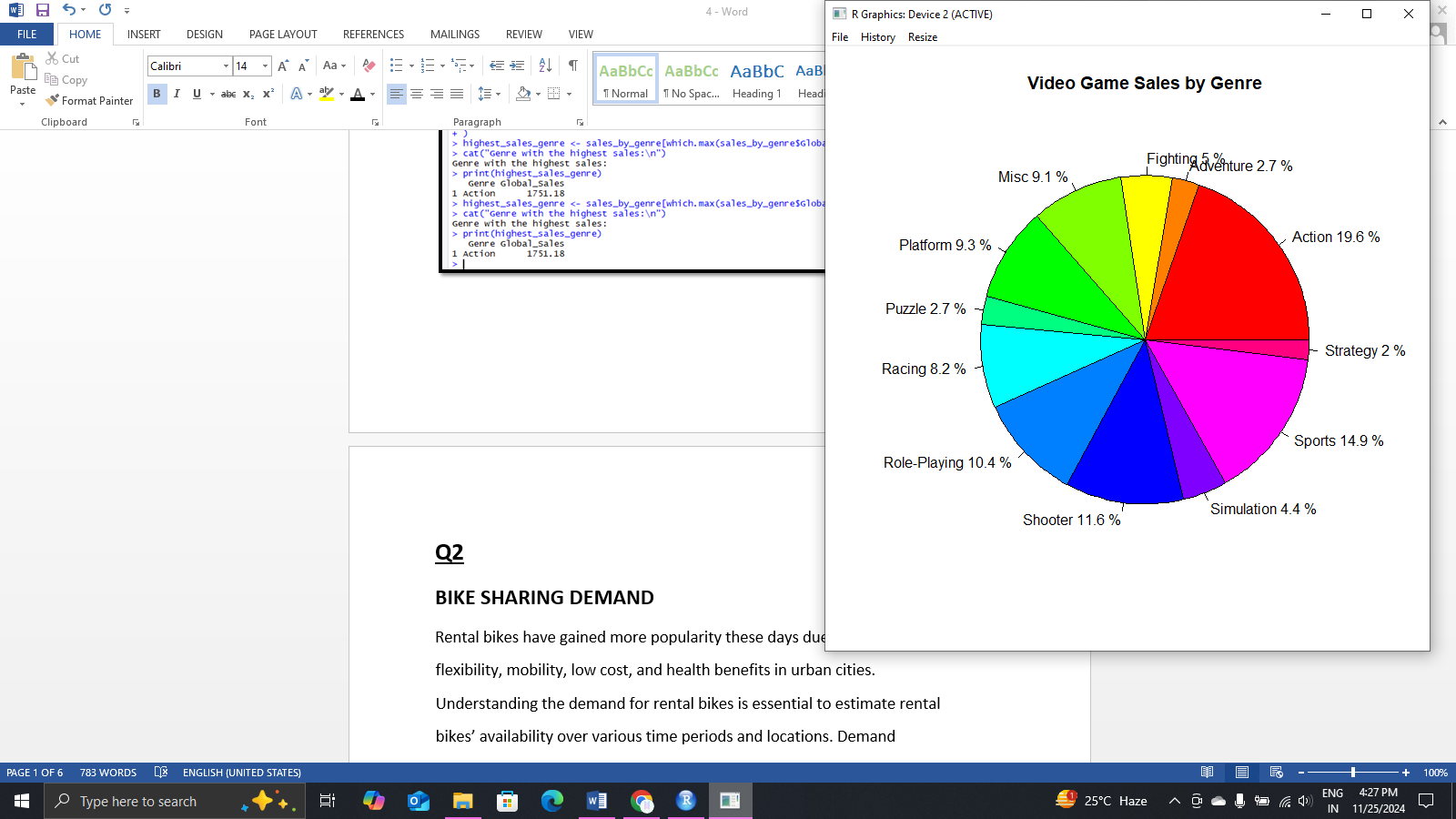
Big Data Analysis Using R

Q1

Video game sales data, consisting of rank, genre, publisher, and global sales amount (in millions) worldwide. Using R create a Pie chart and find out which genre accounts for a high portion of global sales in a video game. This would help understand the potential needs of the video game that will be published by the company.

data set ::vgsales.csv





Q2

BIKE SHARING DEMAND

Rental bikes have gained more popularity these days due to their flexibility, mobility, low cost, and health benefits in urban cities. Understanding the demand for rental bikes is essential to estimate rental bikes’ availability over various time periods and locations. Demand information can also help to allocate the optimal amount of supply of rental bikes. These factors could be critically related to customer satisfaction and traffic.

data set ::seoulbikedata.csv

Here are the explanations of each variable of Dataset

Date: Year – Month – Day

Rented Bike Count: number of bokes rented per hour

Temperature: Temperature in Celsius

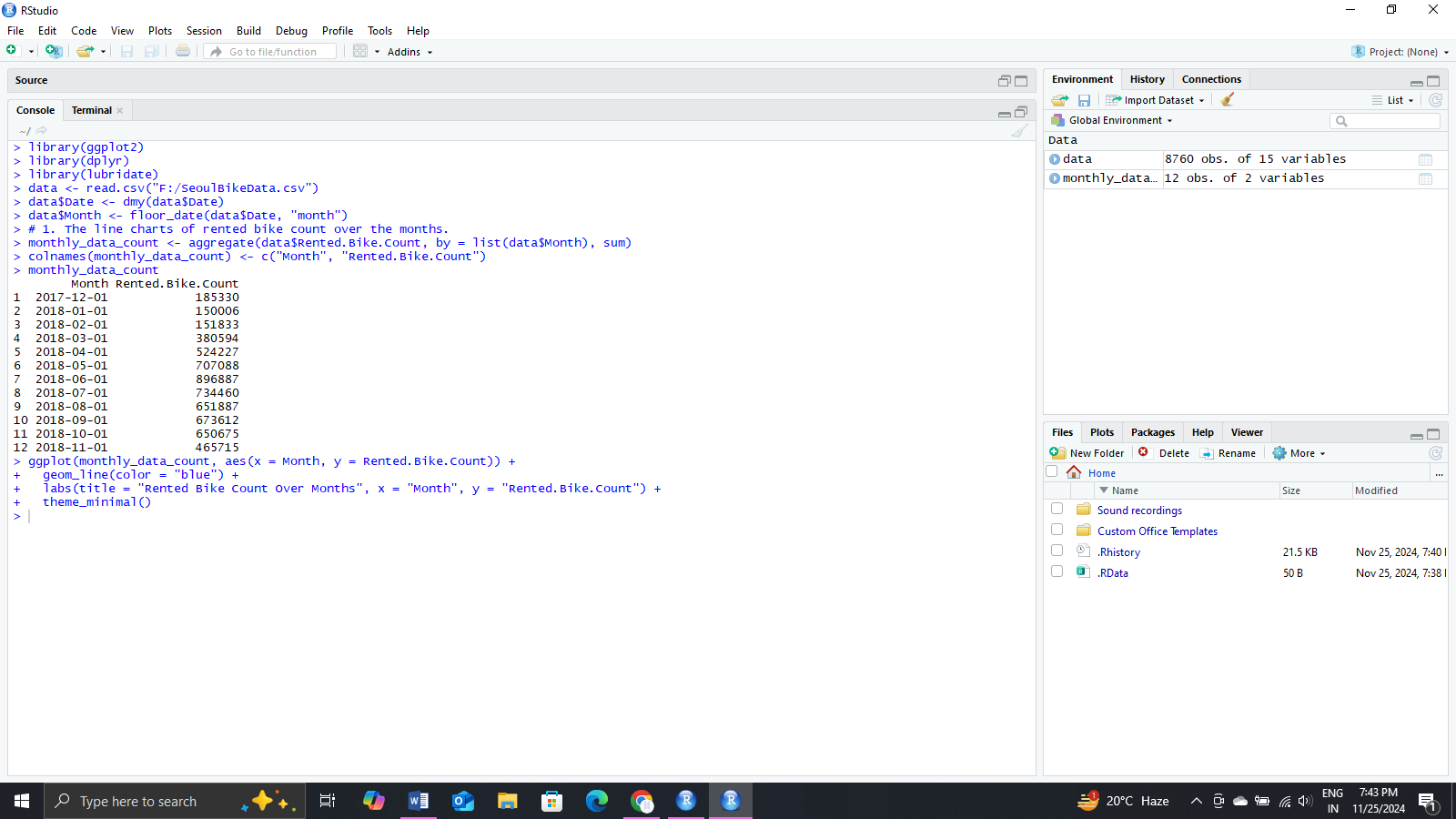
Rainfall: Rainfall in mm

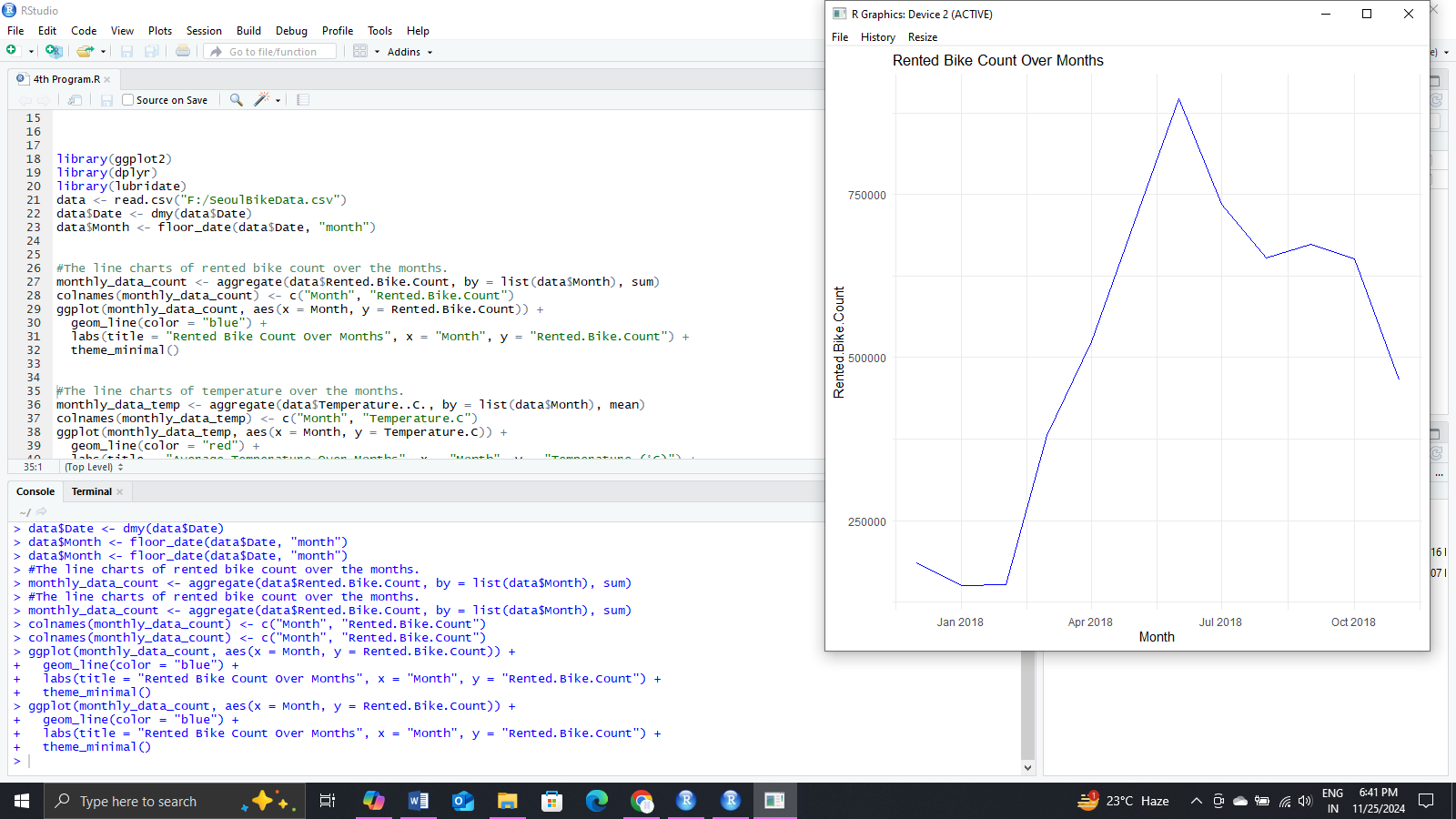
Seasons: Winter, Spring, Summer, and Autumn

Holiday: Holiday and no holiday

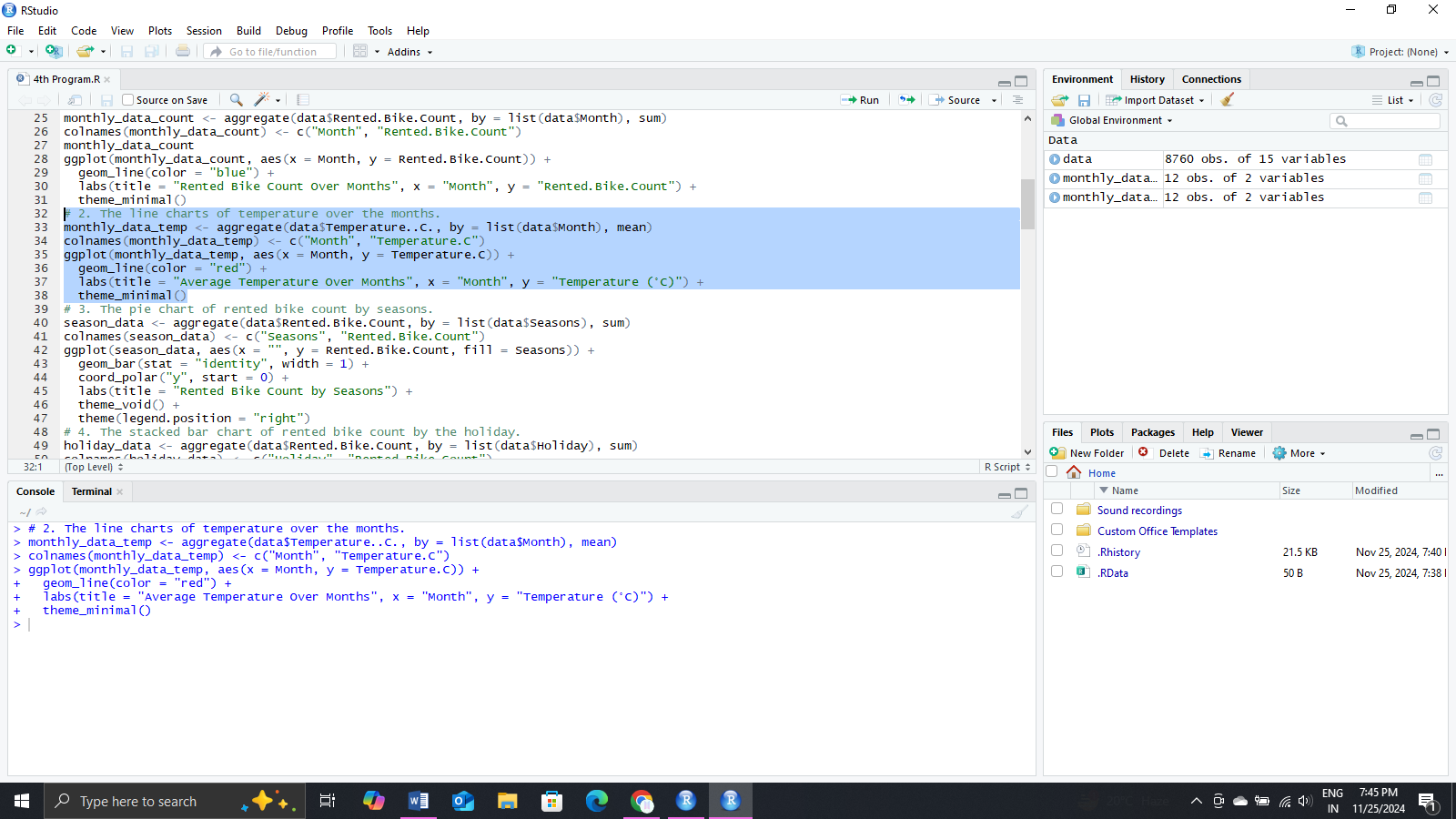
Create a Visualization using R

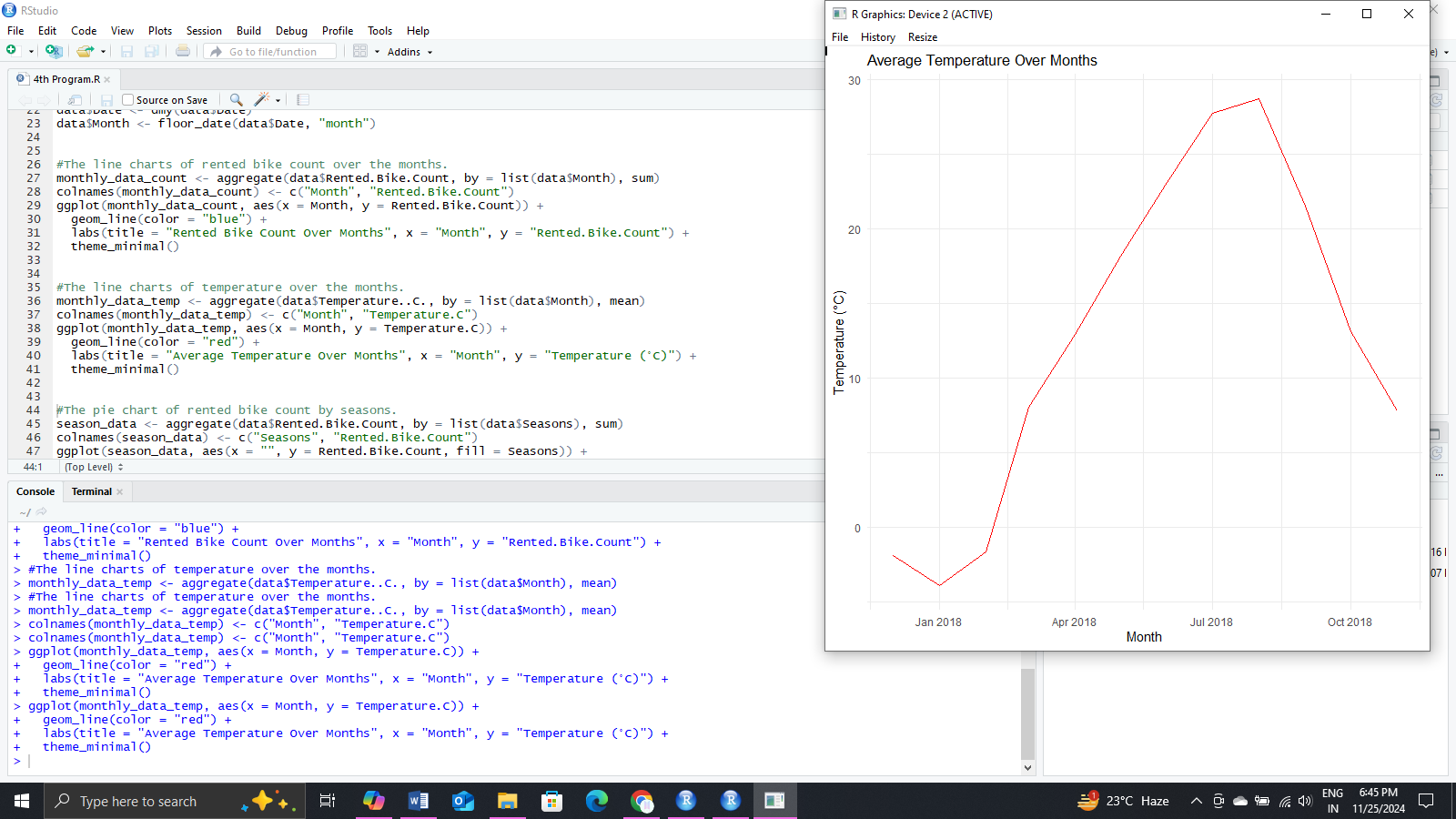
* The line charts of rented bike count over the months.



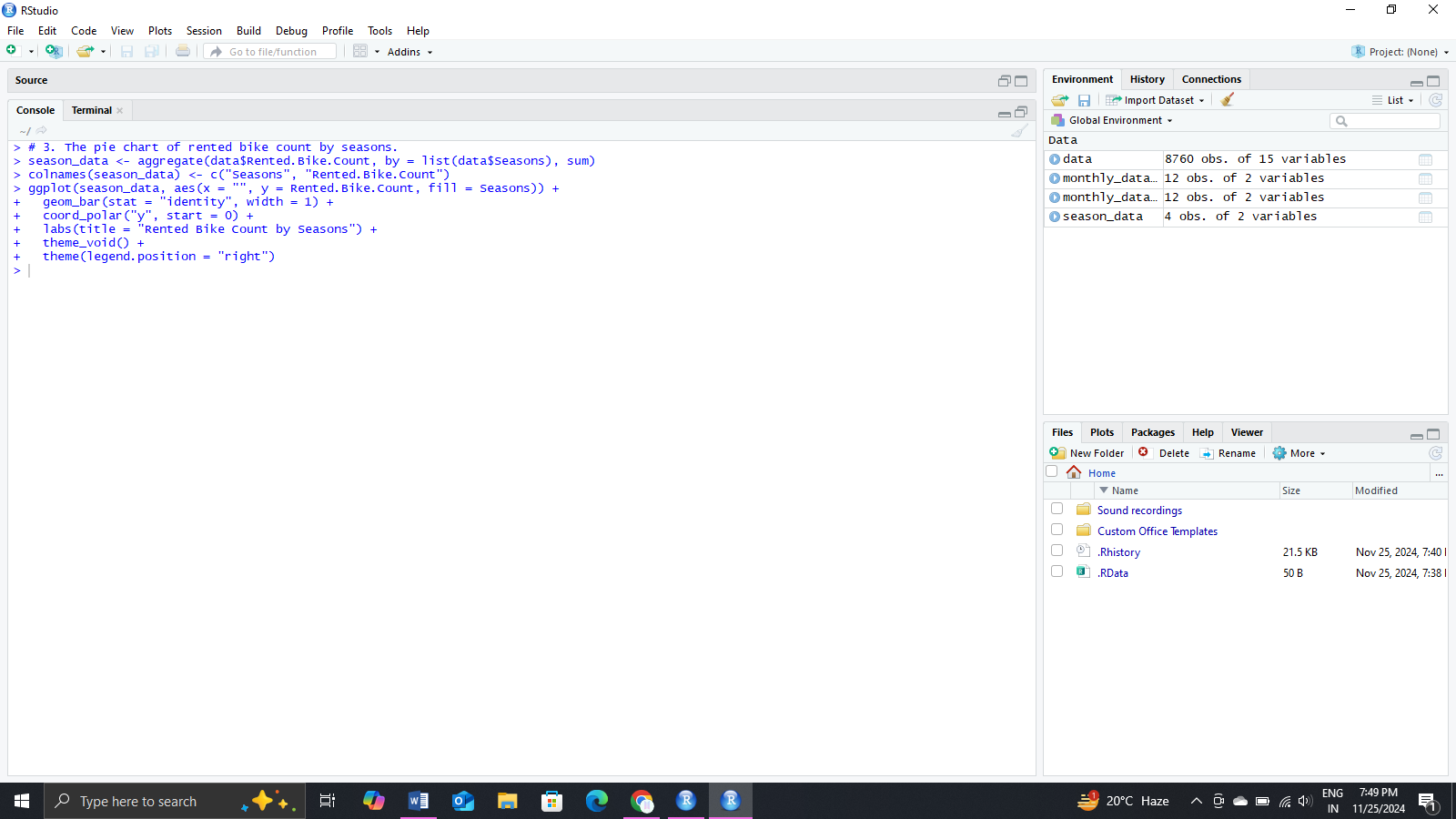


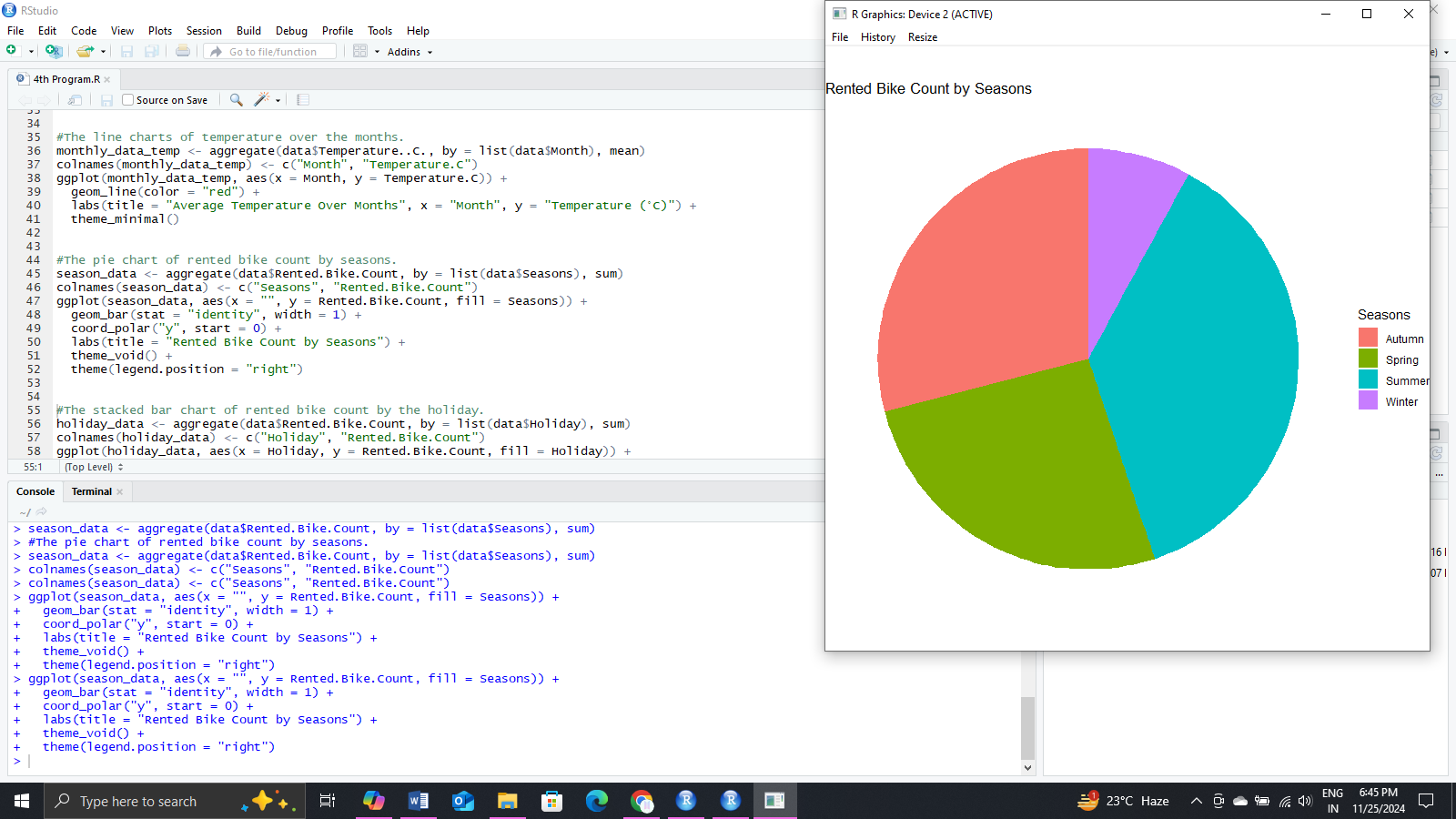
* The line charts of temperature over the months.



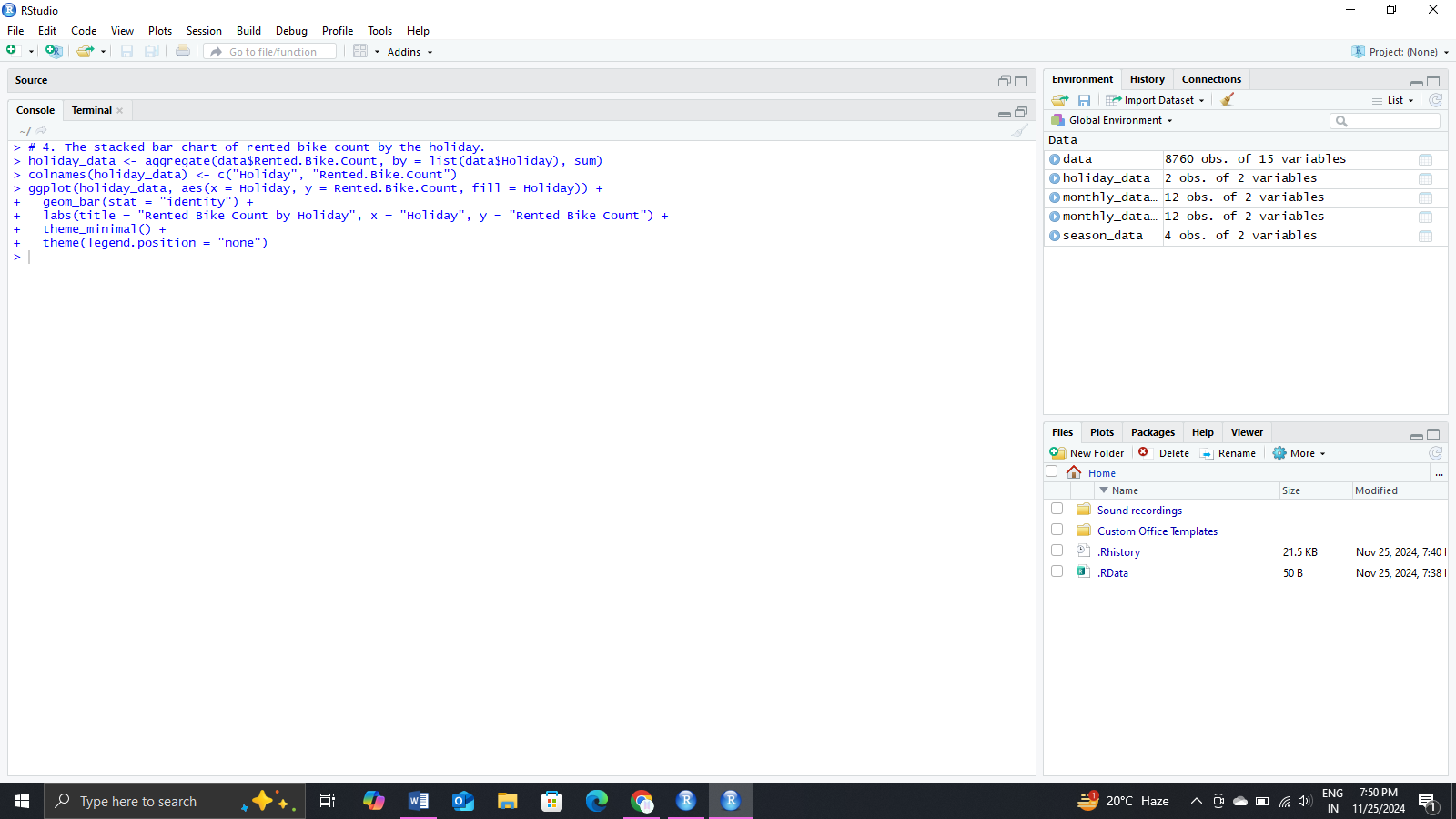


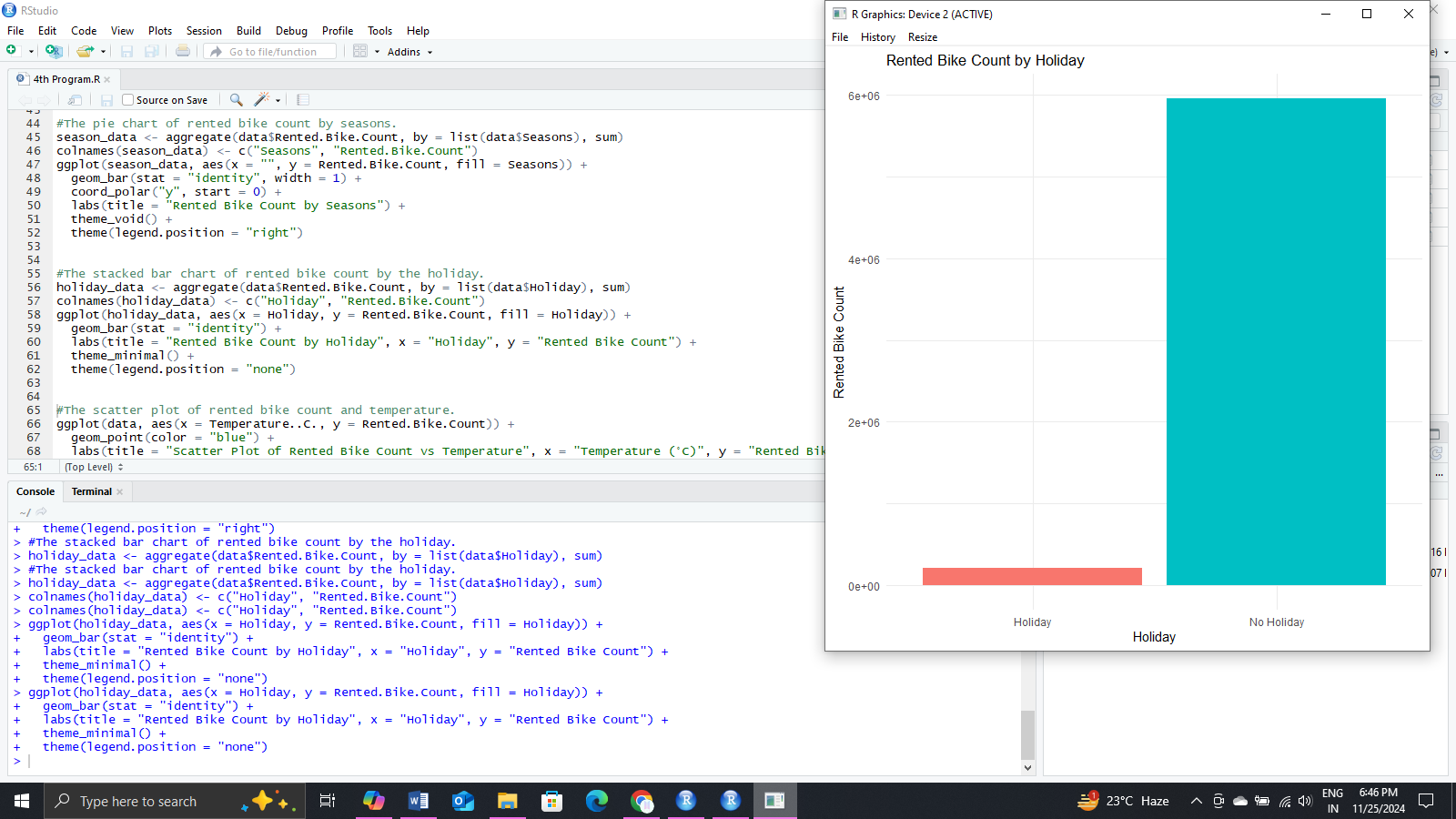
* The pie chart of rented bike count by seasons.



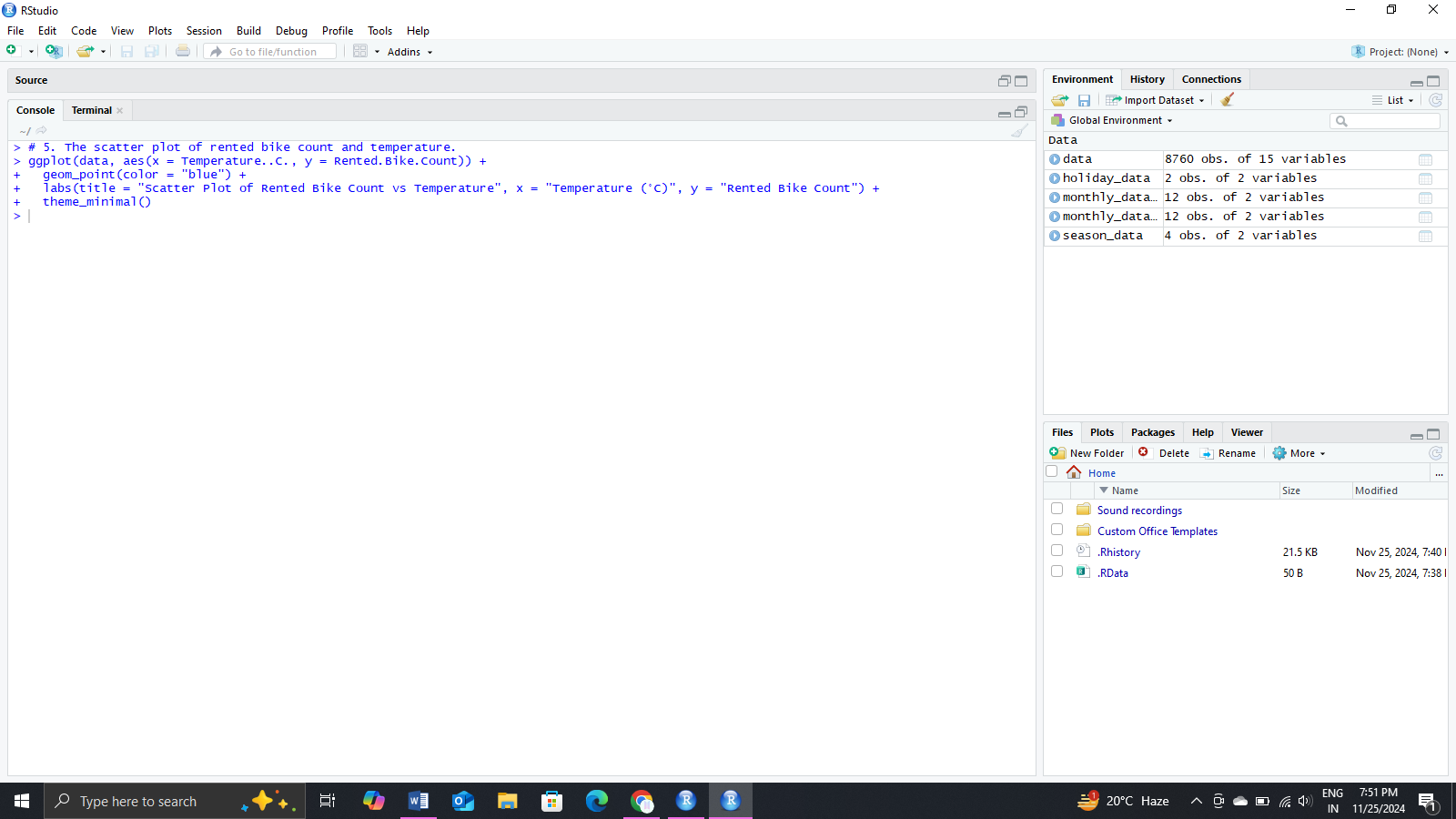


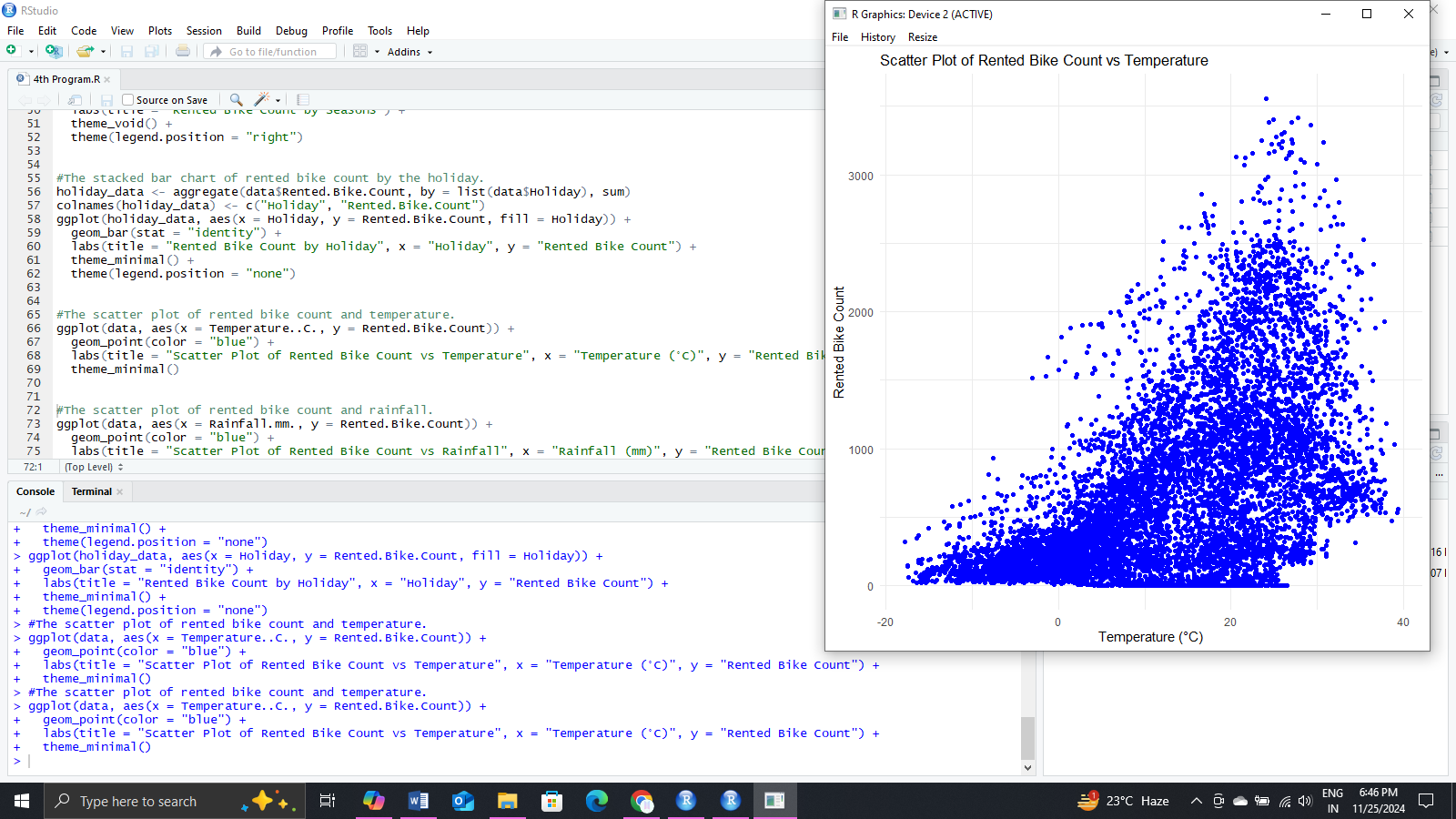
* The stacked bar chart of rented bike count by the holiday.



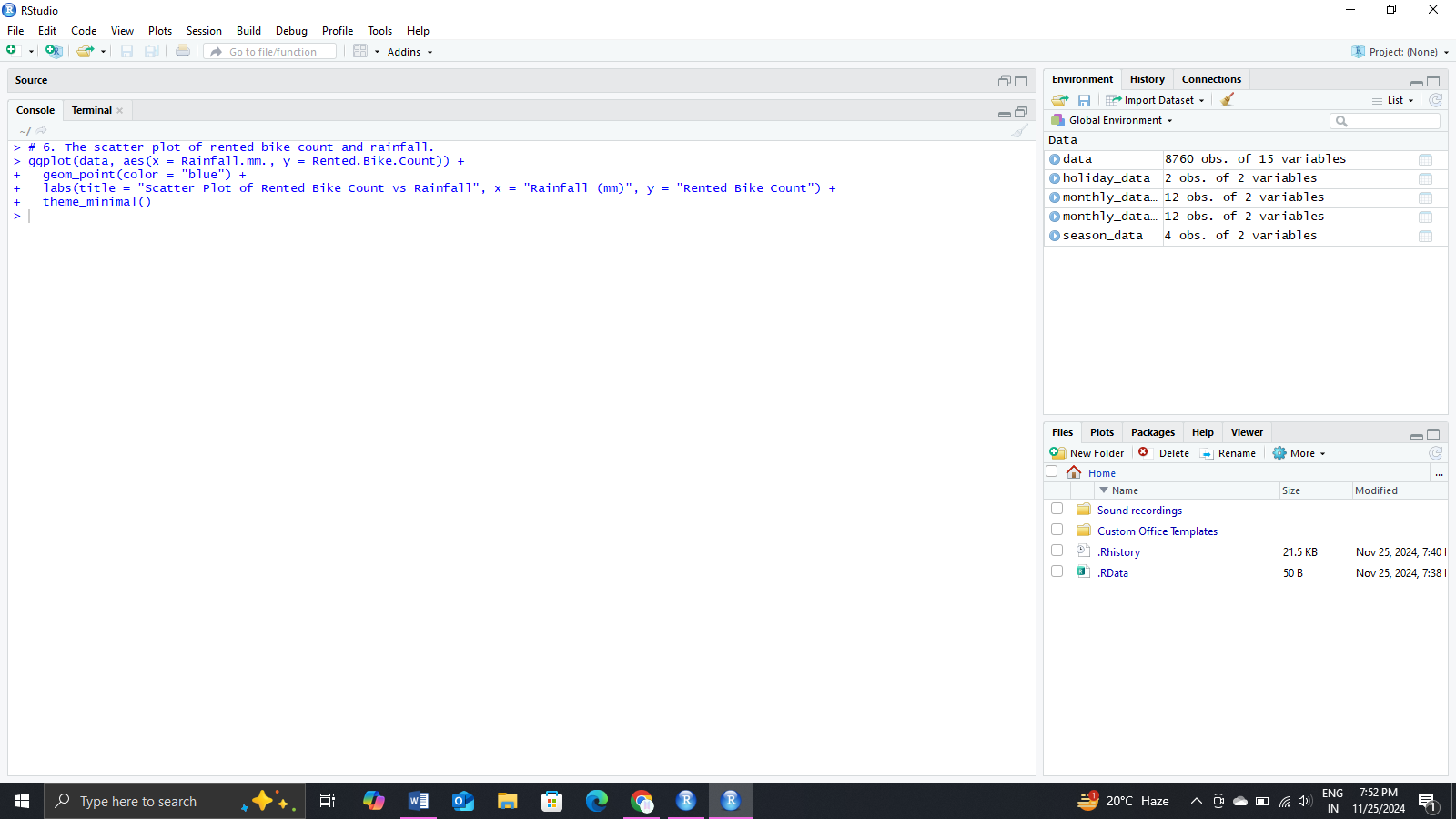


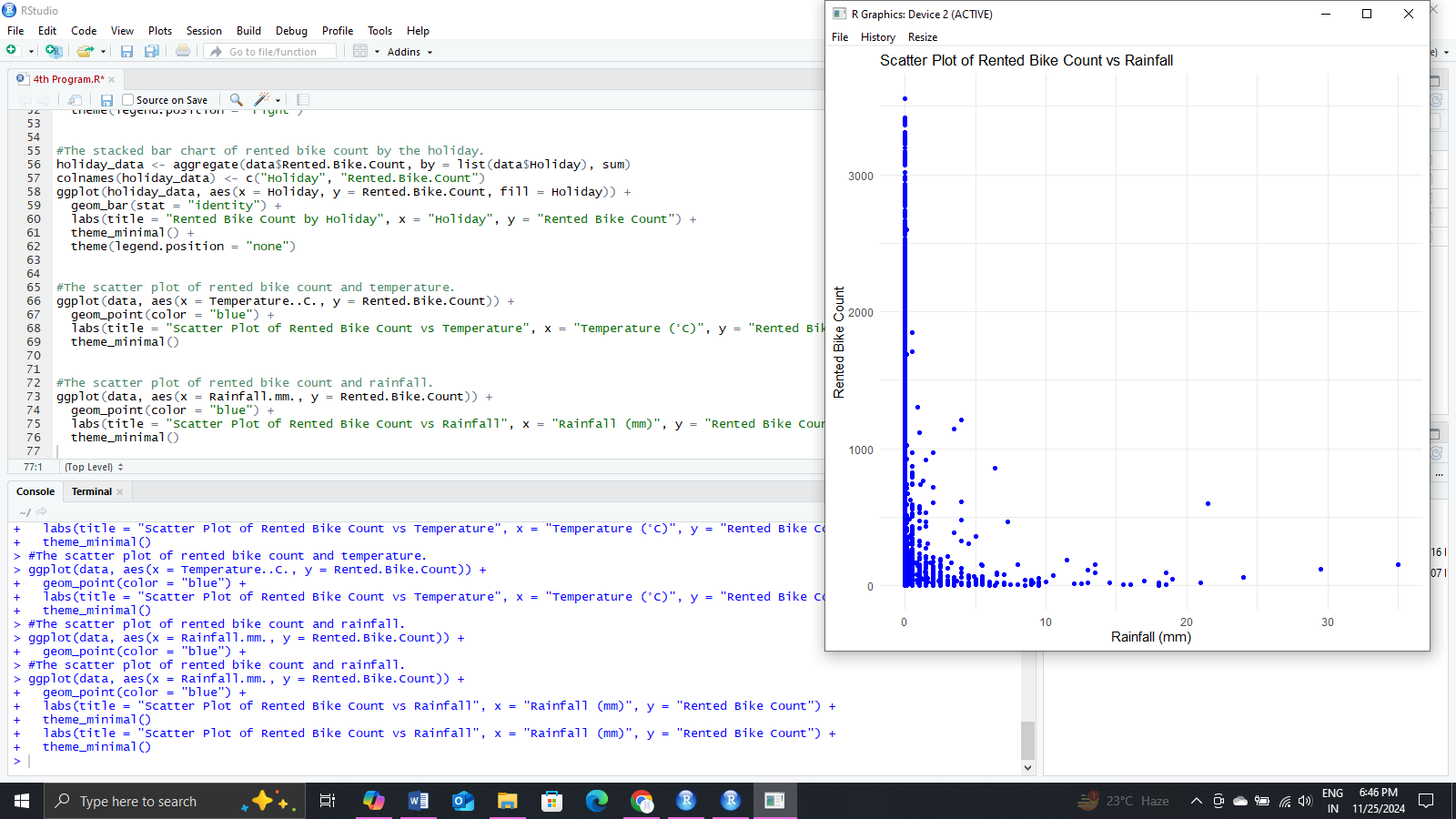
* The scatter plot of rented bike count and temperature.





* The scatter plot of rented bike count and rainfall.





Answer following Questions:

1. What is the trend of rented bike count over the months?

After declining in trend of rented bike from December it changed From February, as the trend of rented bike increased rapidly till June then there are several ups and downs and it decreased at slow rate.

1. Which seasons show greater demand for the rented bike than other seasons?

Summer seasons show greater demand for the rented bike than other seasons

1. Is the demand for rented bikes affected by the holiday season?

Yes, demand for rented bikes affected by the holiday season. As Rented bike count on Holiday and No Holiday are 215895 and 595641 respectively.

1. Is there a relationship between the rented bike demand and temperature?

Yes, with increase in temperature there is increase in the rented bike demand. As at the time of less temperature there is less demand of rented bikes but it increased as the temperature increased.

1. Is there a relationship between the rented bike demand and the amount of rainfall?

Yes, with increase in rainfall there is decrease in the rented bike demand. As the demand was at its peak when there was no rainfall but it decrease readily when the rainfall increase.

Bicycle rental companies can monitor this dashboard to predict the rented bikes’ demand and find appropriate counter measures.

Through the dashboard analysis, stakeholders can think of the following questions and find solutions to them.

* Does the demand for bicycle rental vary by season?

Yes, demand for bicycle rental vary by season.

Demand: Winter < Spring < Autumn < Summer

* What is the difference between the monthly demand for bike rental?

There is the maximum difference between the monthly demand for bike rental in January (150006) and June(896887) and that is 746881

* Is the demand for bike rental affected by holidays?

Yes, demand for rented bikes affected by the holiday season. As Rented bike count on Holiday and No Holiday are 215895 and 595641 respectively.

Q3

REAL TIME VOICE CALL QUALITY DATA FROM CUSTOMERS

data set : :call voice quality.rar

Since 2000, mobile phones have spread rapidly, and since 2010, many people use data communication through smartphones. Voice communication is the most basic service in the mobile communication business and understanding the quality and the performance of voice calls is critical to ensuring great customer experiences. Bad call experiences lead to frustrated customers, lost customer relationships, and have a real financial impact on businesses. However, measuring call quality was not an easy part for mobile carriers since users’ subjective factors are reflected a lot. Thus, mobile communication companies have used customer survey techniques to check call quality, continue to monitor, and trace call performance to improve service quality based on the survey data. The data set in this case captures the Customers Feedback using the MyCAll App developed by TRAI (Telecom regulatory authority of India) which is a statutory body set up by the Government of India The data is captured for various service providers in India, at multiple locations, network types 2G, 3G, 4G, ratings, coordinates, etc. Customers rate their experience with voice call quality in real-time and help TRAI gather customer experience data along with Network data. The Variable and Description of the Voice Call Quality

Data

Operator : Telecom service provider

Indoor\_Outdoor\_Travelling : The location status information when users rate service quality

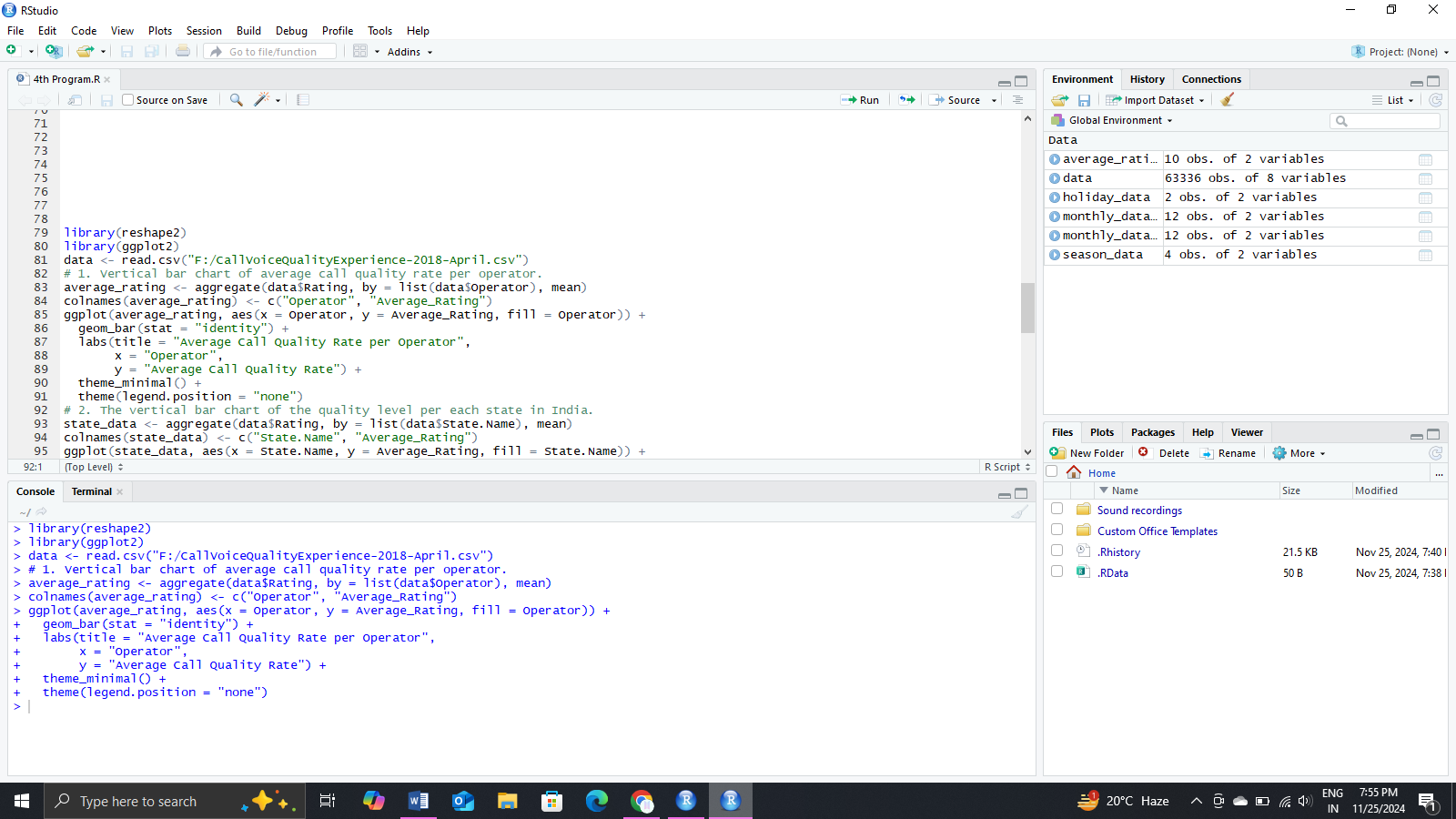
Network Type : The type of Network and standard for telecommunication

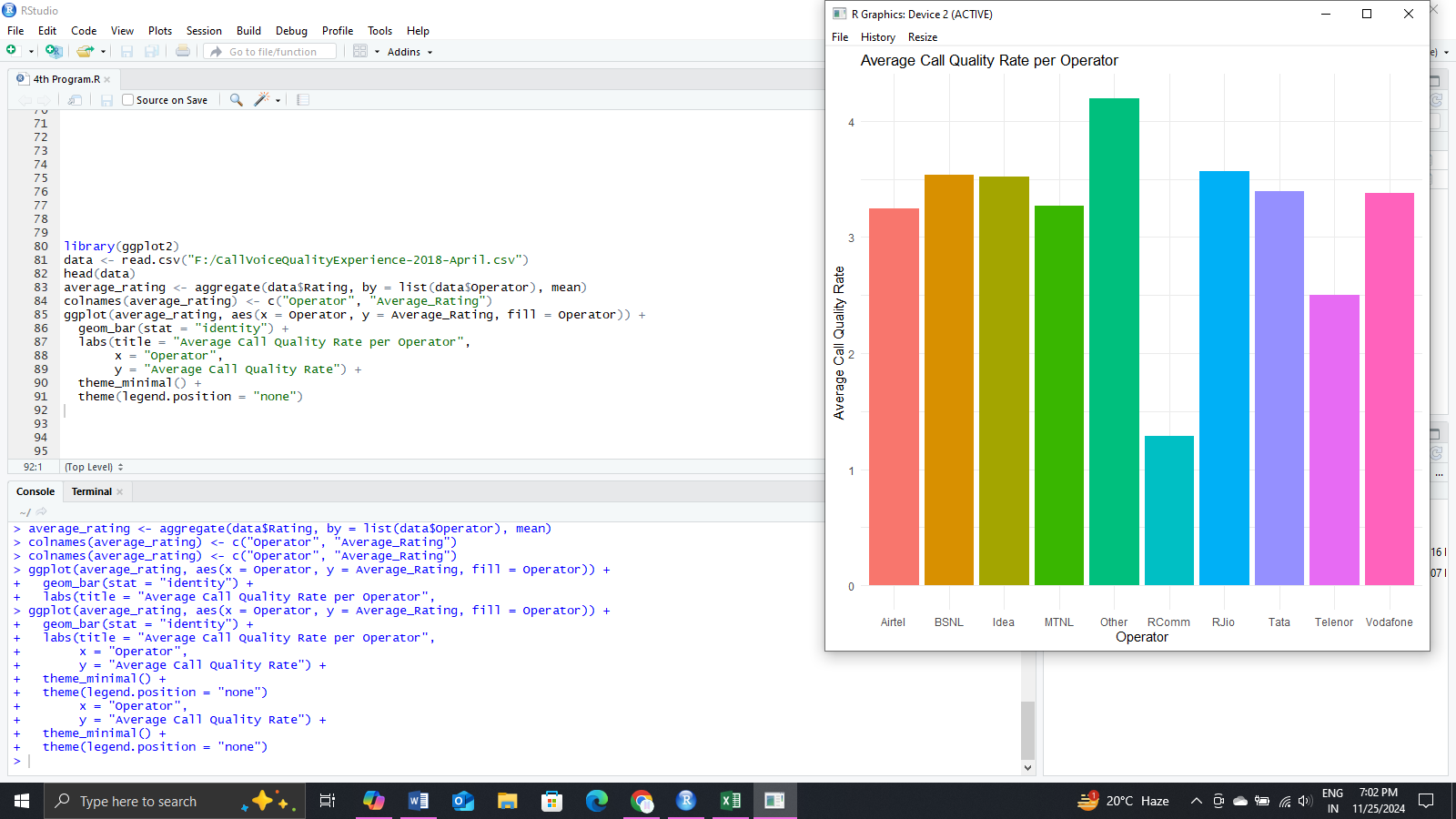
Rating : Quality rated by customers (0: lowest quality; 5: highest quality)

Call Drop Category : Reported data If a call is dropped (call drop issue) Latitude : Geographic coordinate located when the quality was rated Longitude : Geographic coordinate located when the quality was rated State : Name Location information

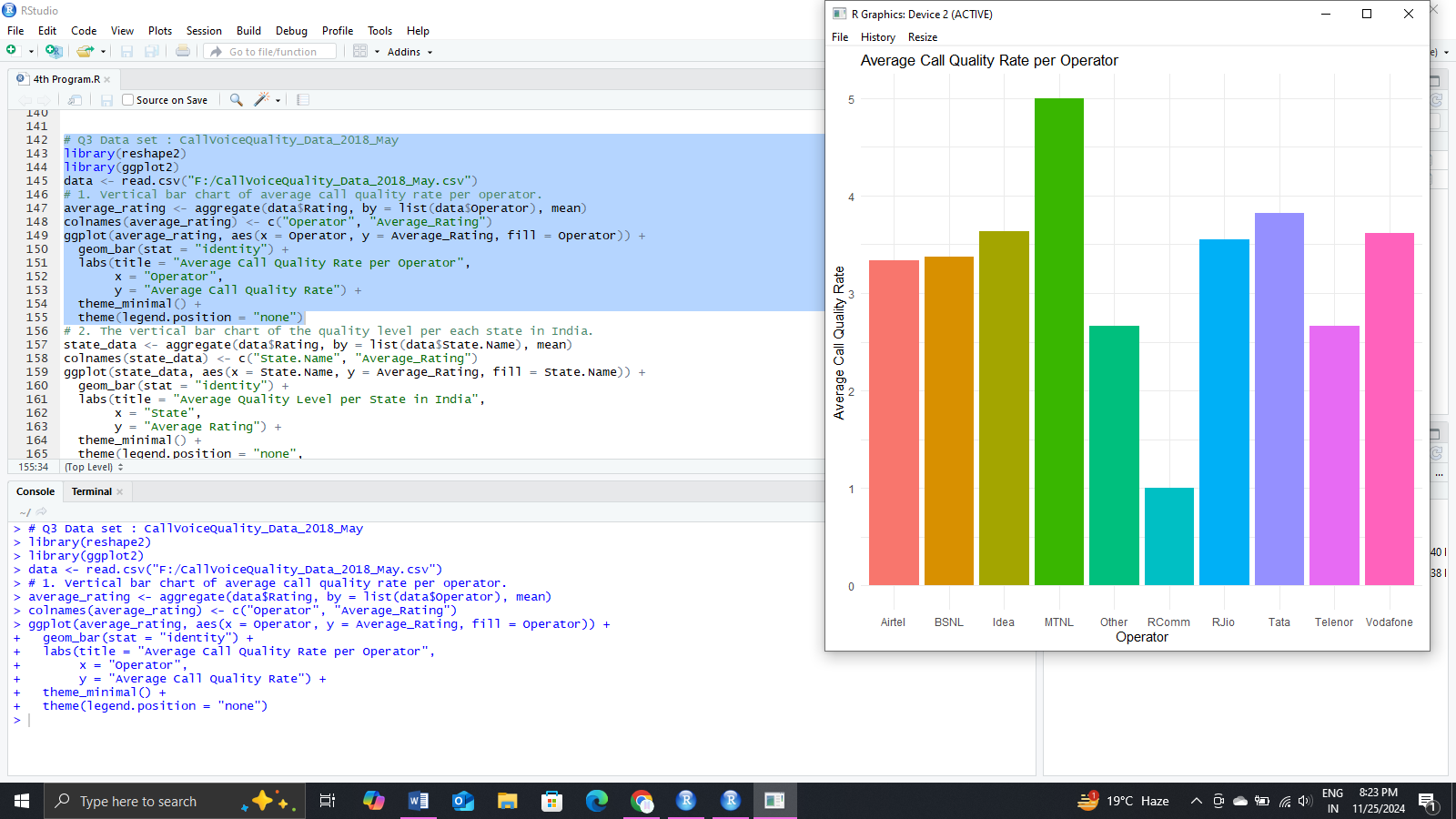
Create Data visualization using R for:

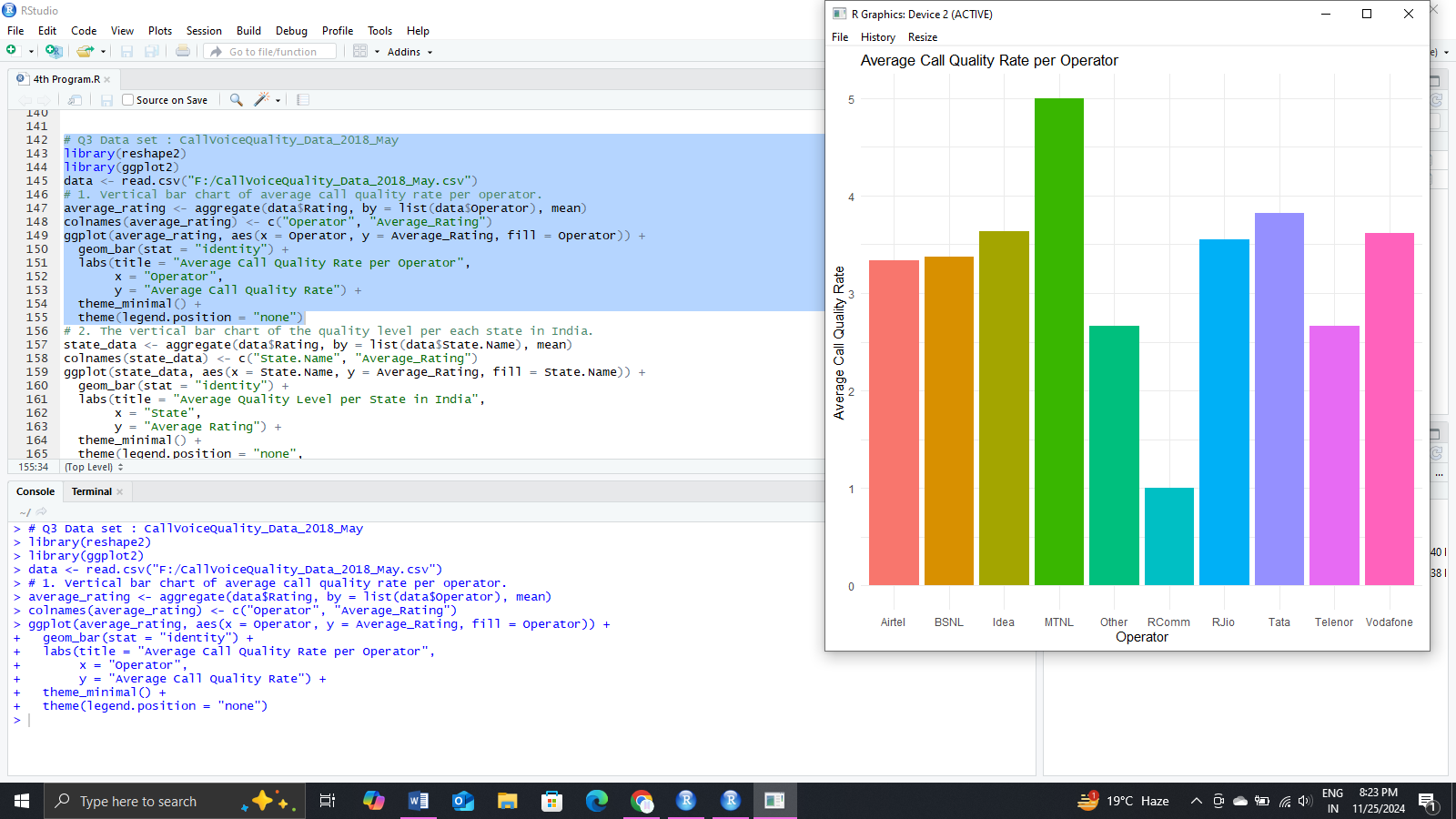
* Vertical bar chart of average call quality rate per operator.
* Data set : CallVoiceQualityExperience-2018-April



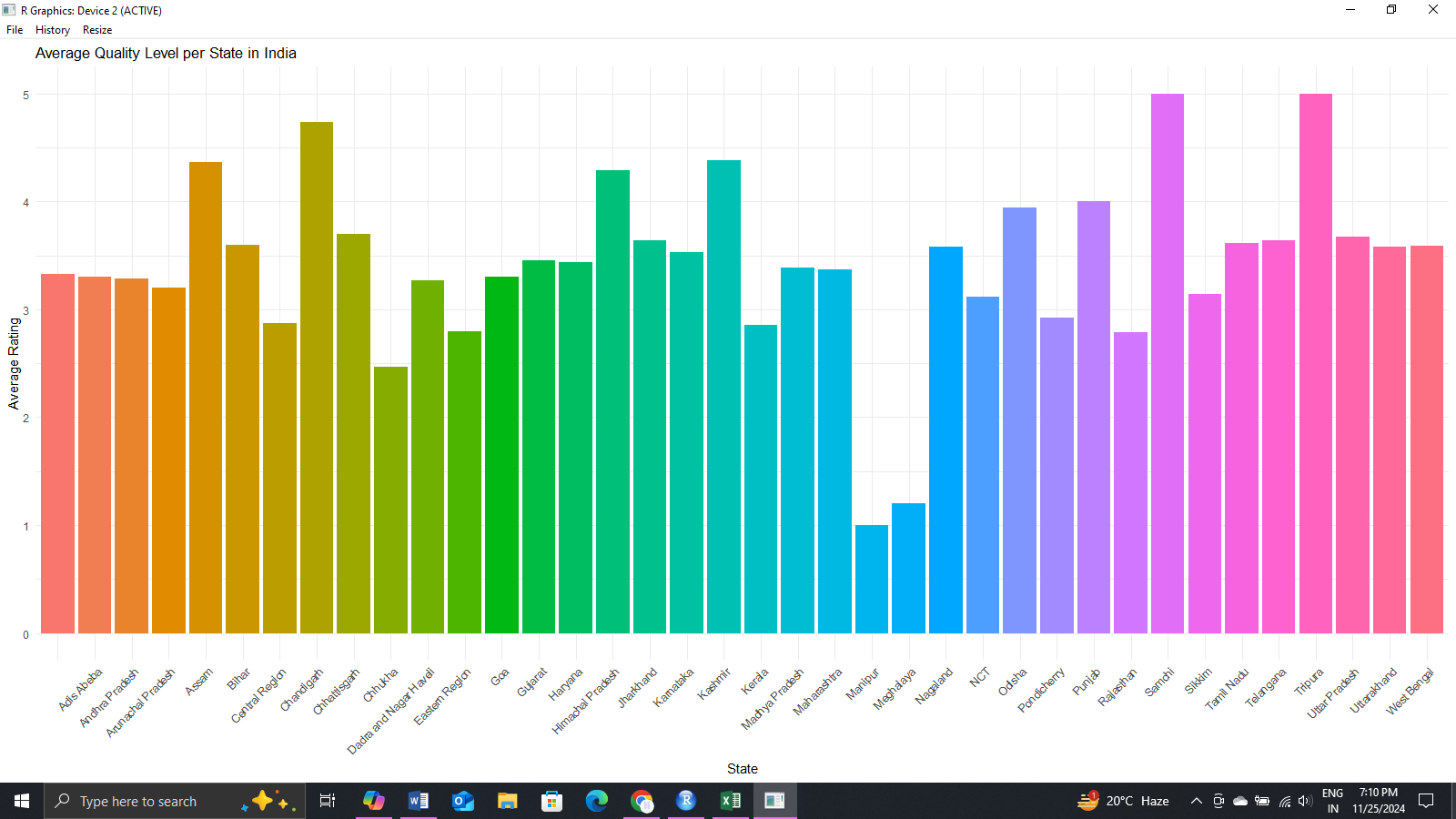
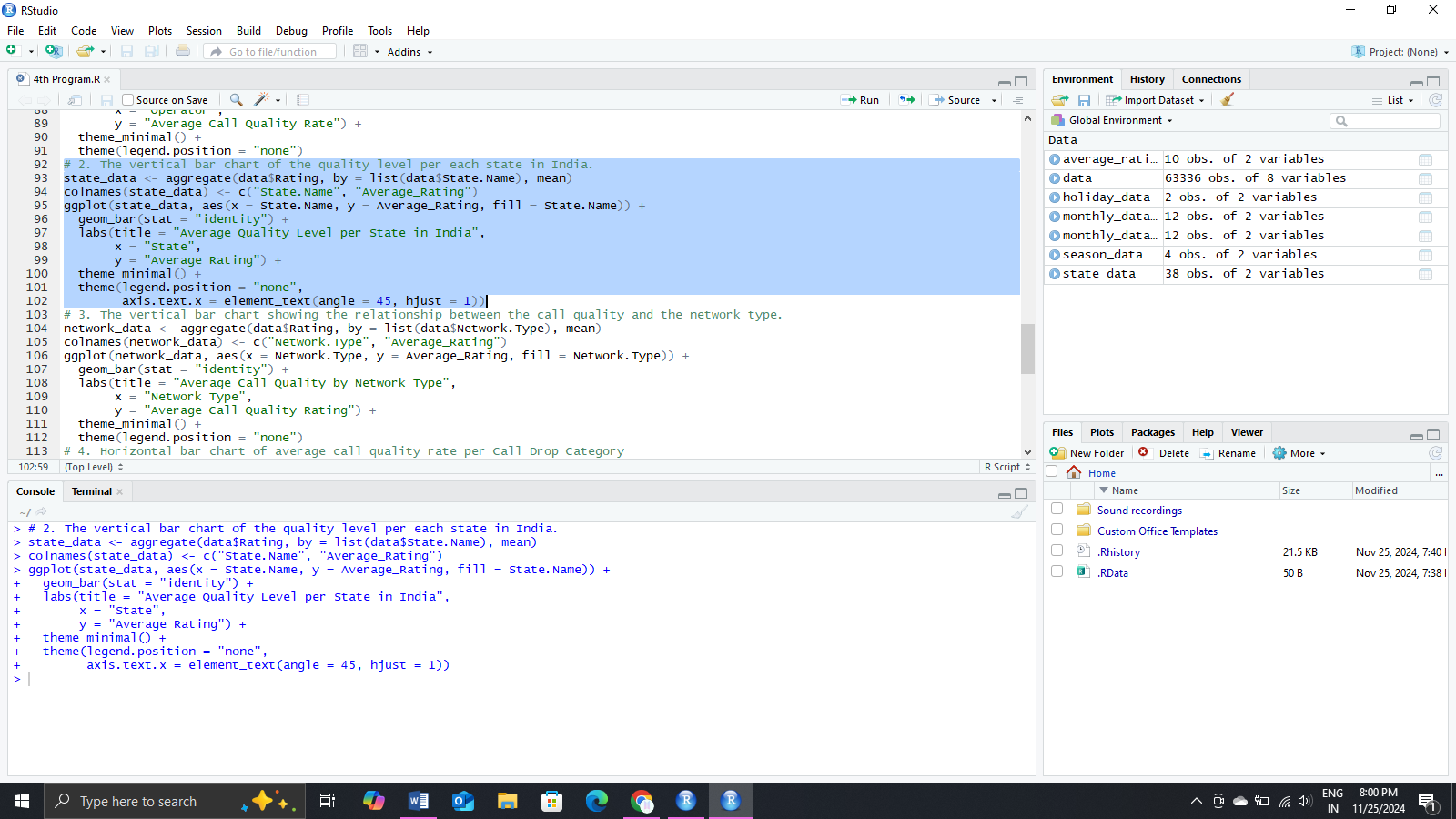


* Data set : CallVoiceQuality\_Data\_2018\_May

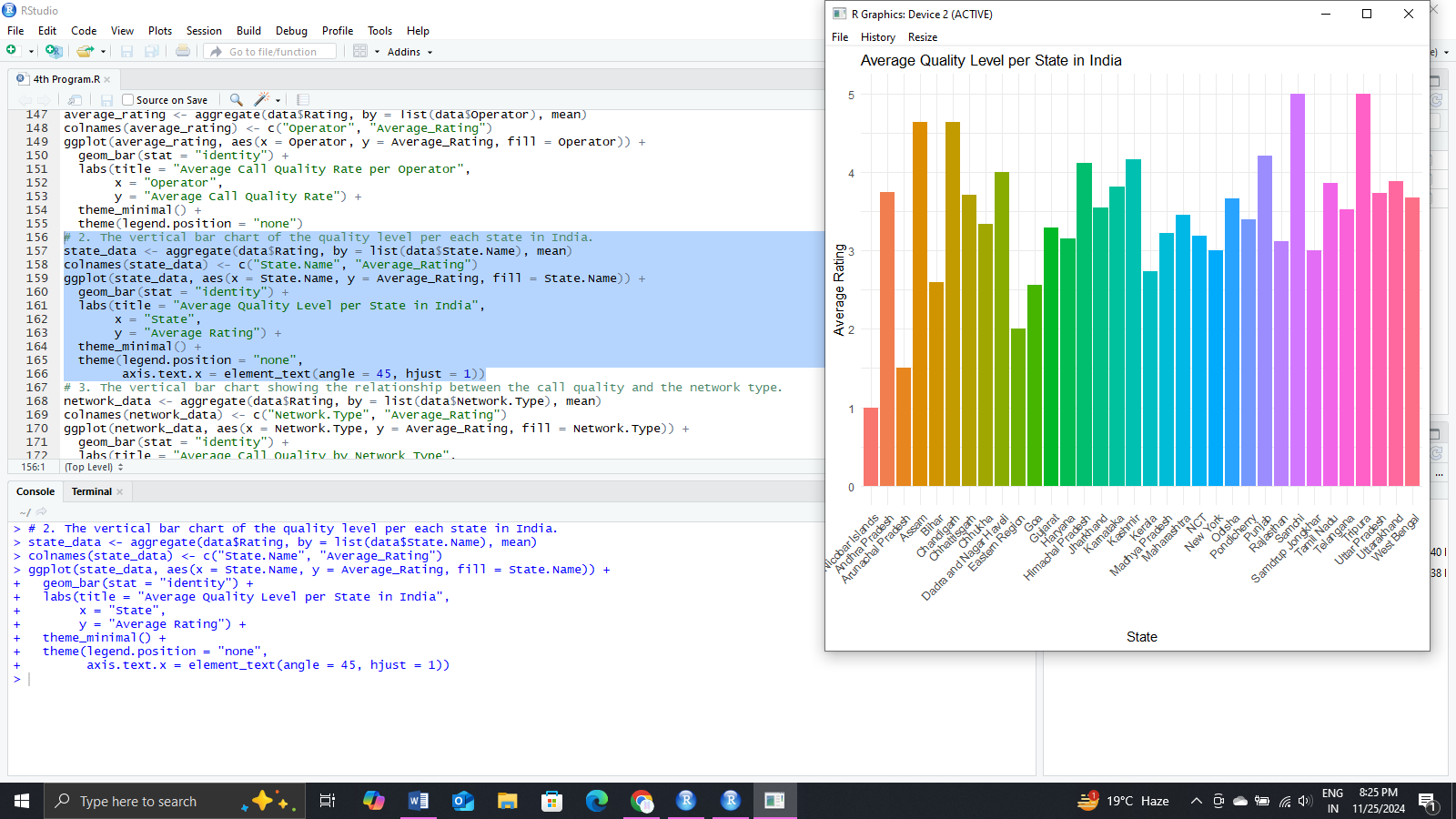


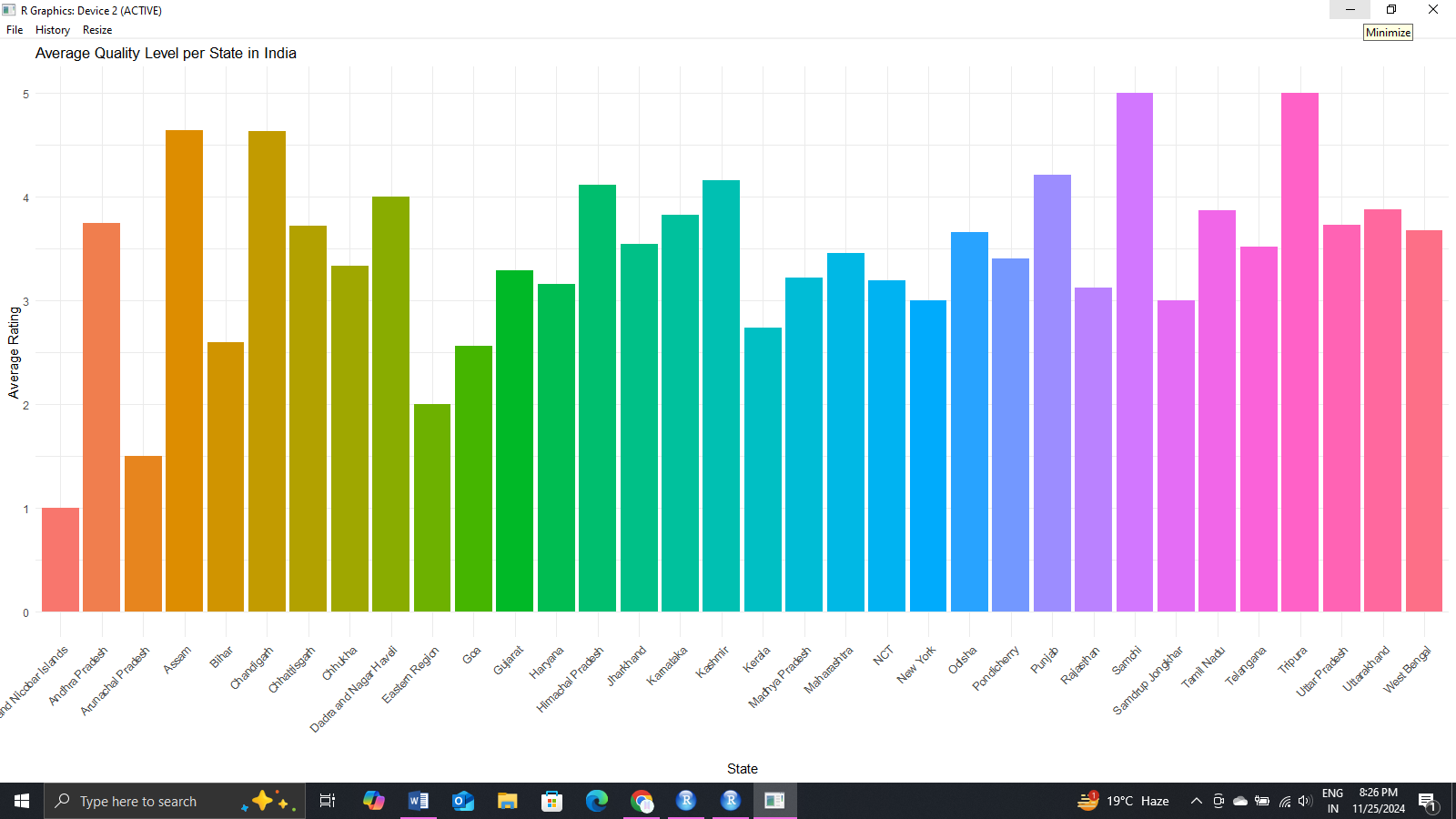


* The vertical bar chart of the quality level per each state in India.
* Data set : CallVoiceQualityExperience-2018-April

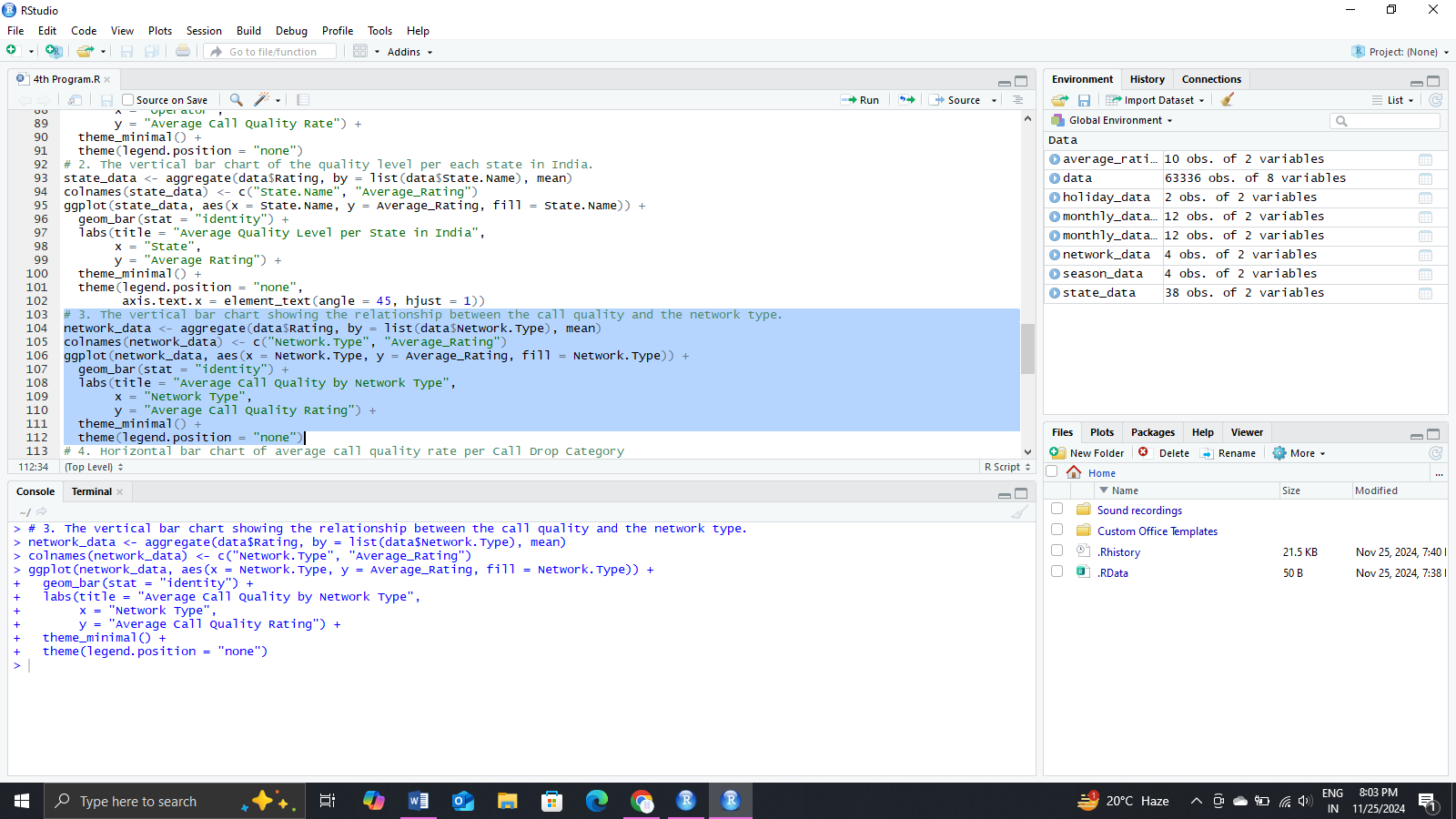


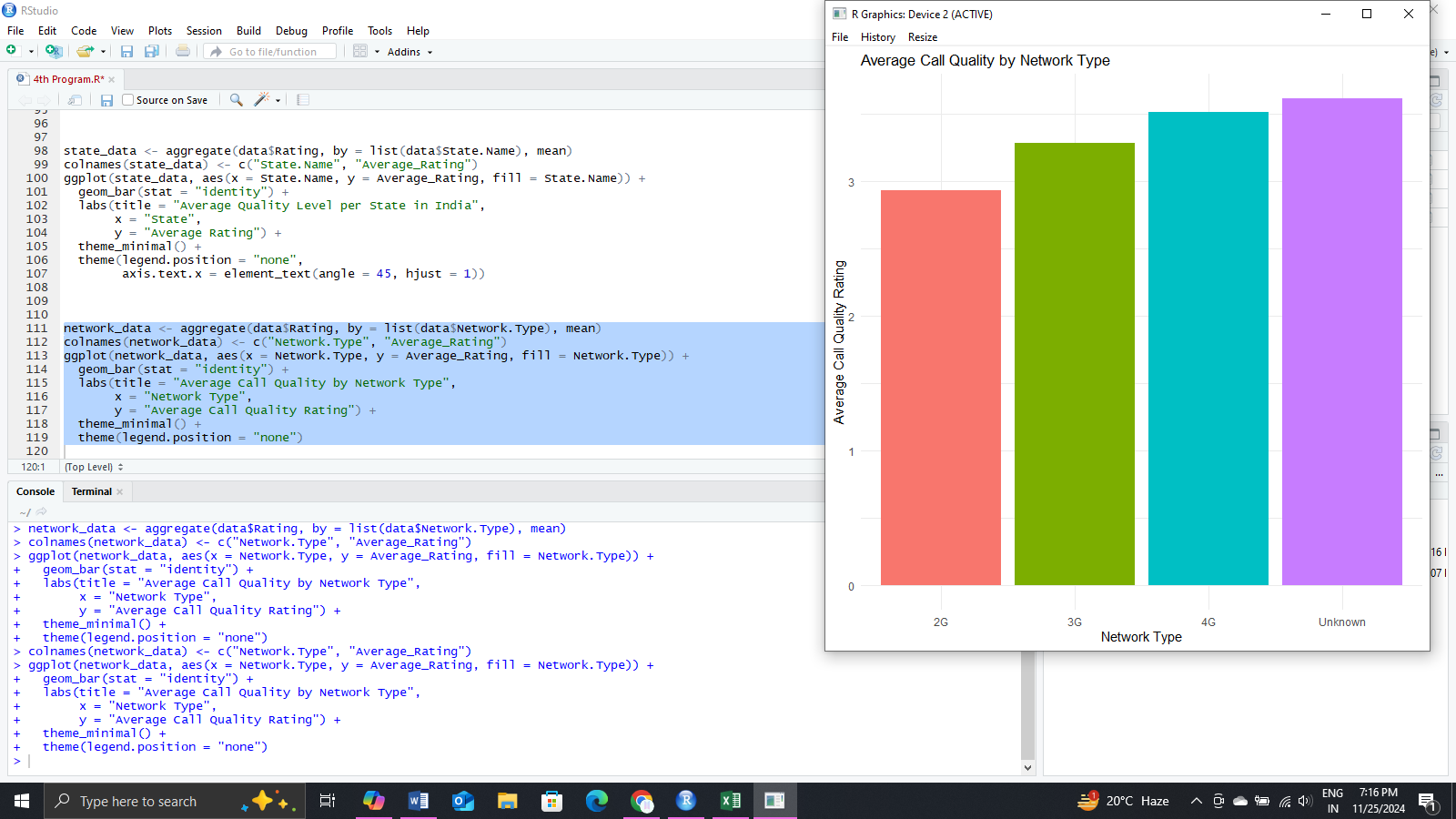
* Data set : CallVoiceQuality\_Data\_2018\_May



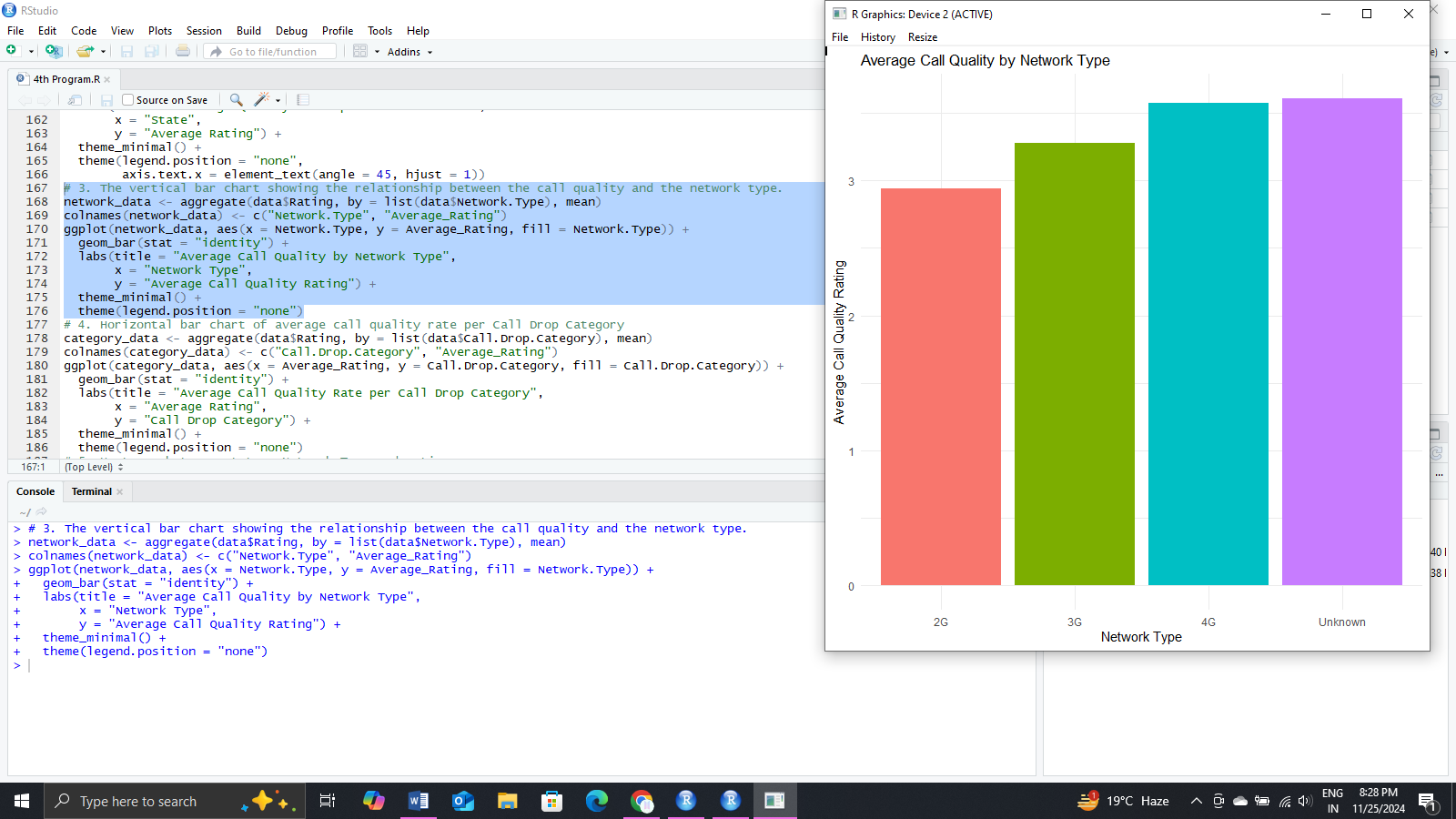


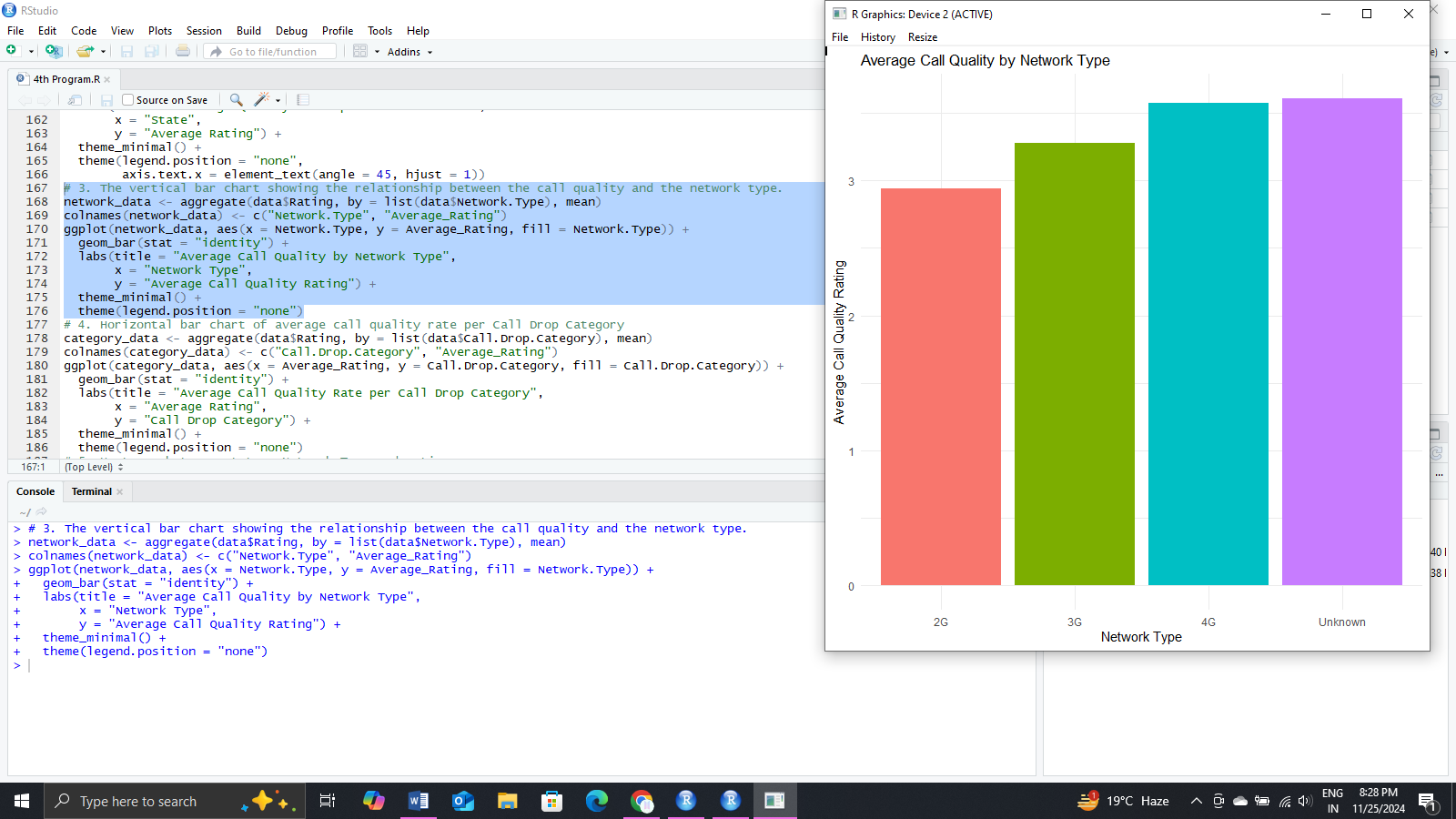
* The vertical bar chart showing the relationship between the call quality and the network type.
* Data set : CallVoiceQualityExperience-2018-April



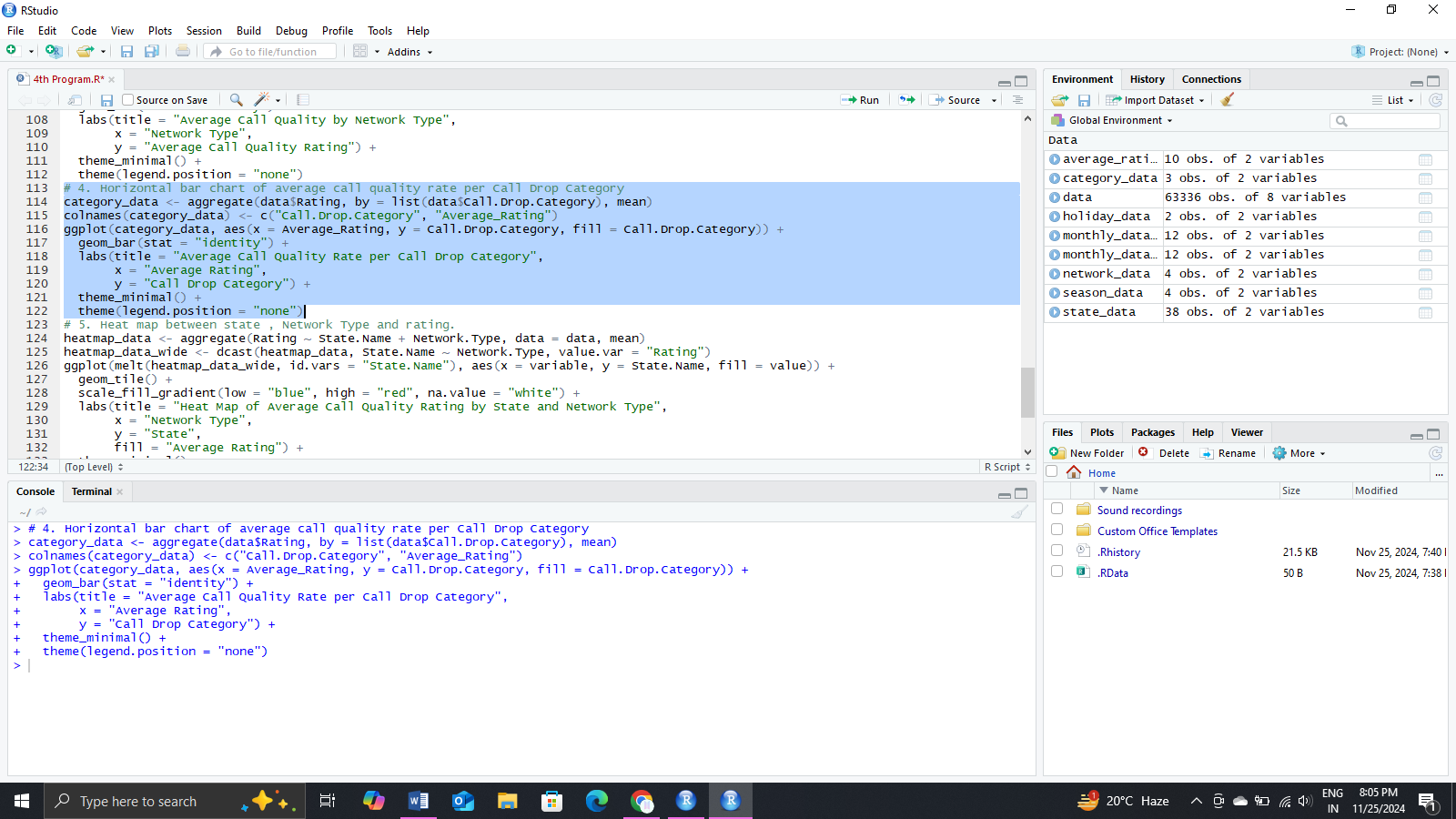


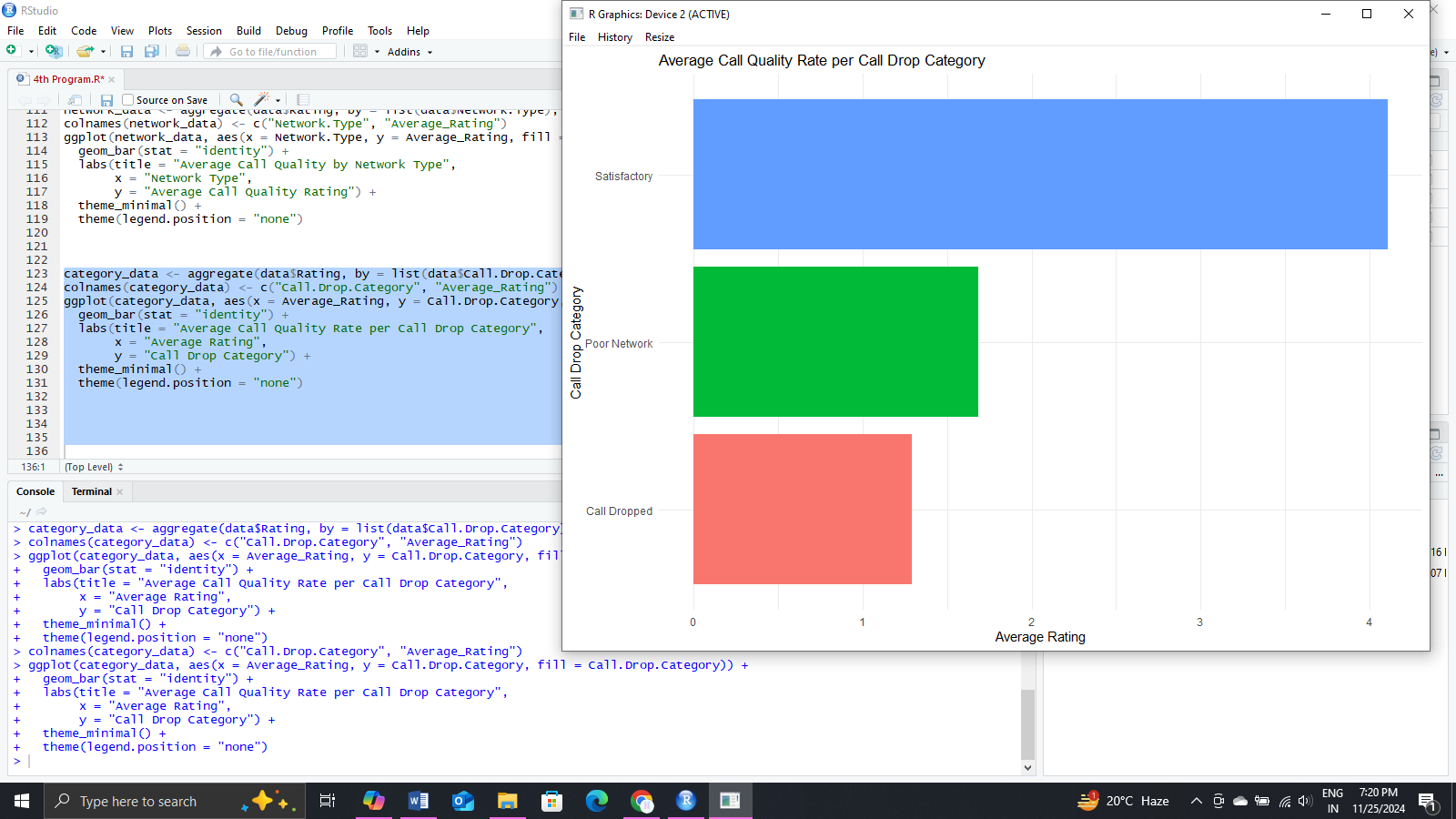
* Data set : CallVoiceQuality\_Data\_2018\_May





* Horizontal bar chart of average call quality rate per Call Drop Category
* Data set : CallVoiceQualityExperience-2018-April



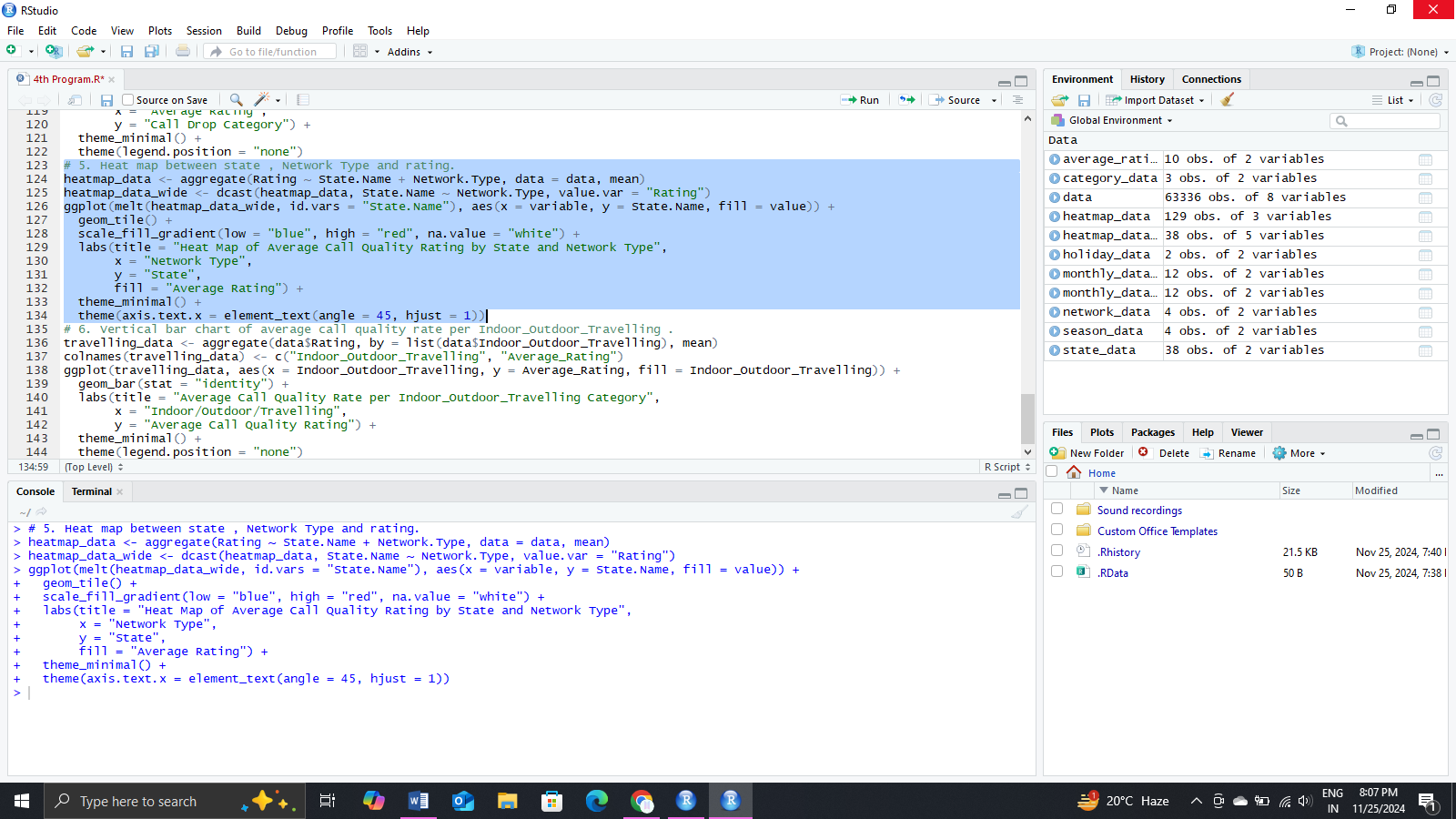


* Data set : CallVoiceQuality\_Data\_2018\_May



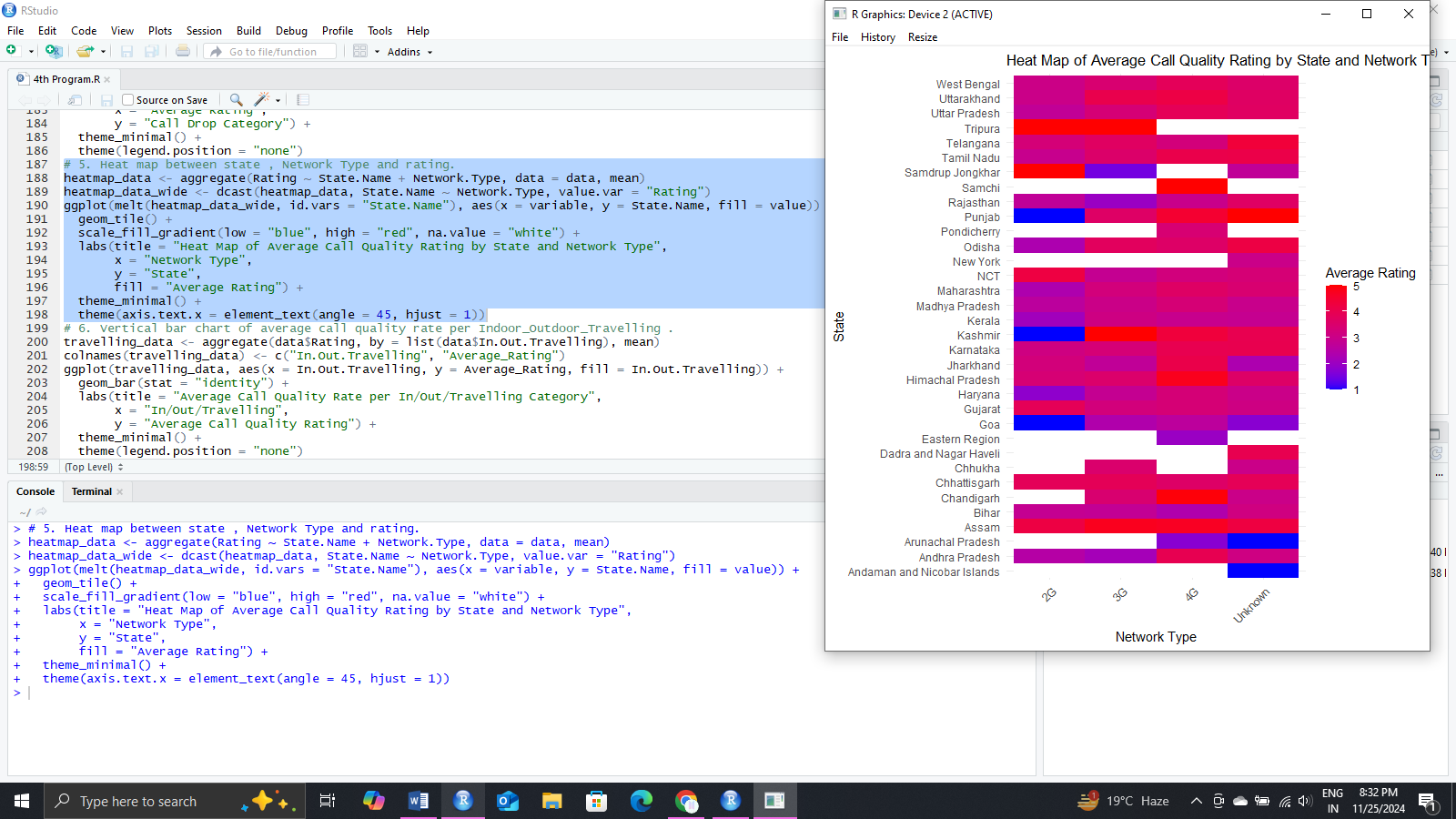


* Heat map between state , Network Type and rating.
* Data set : CallVoiceQualityExperience-2018-April



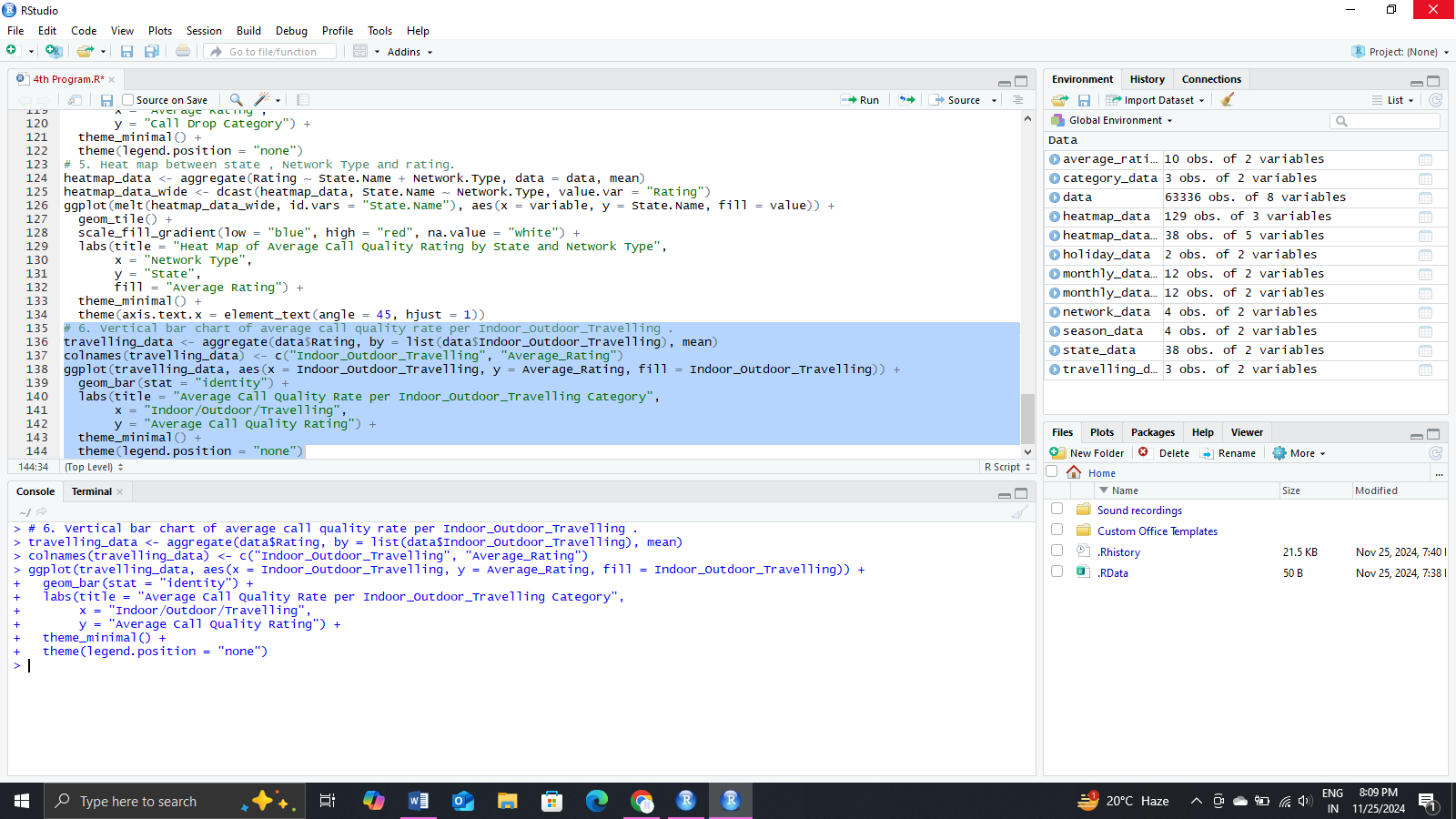


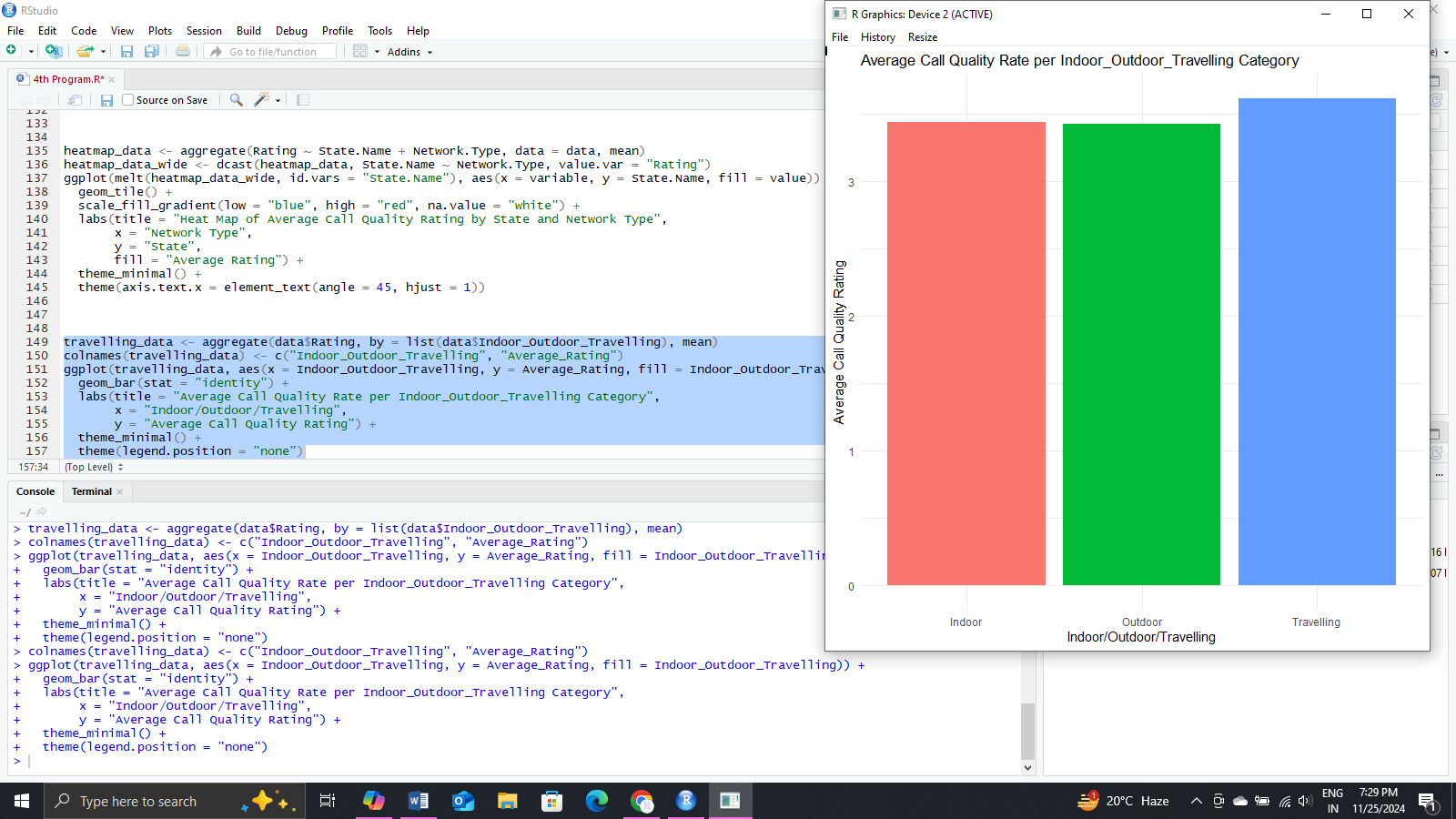
* Data set : CallVoiceQuality\_Data\_2018\_May



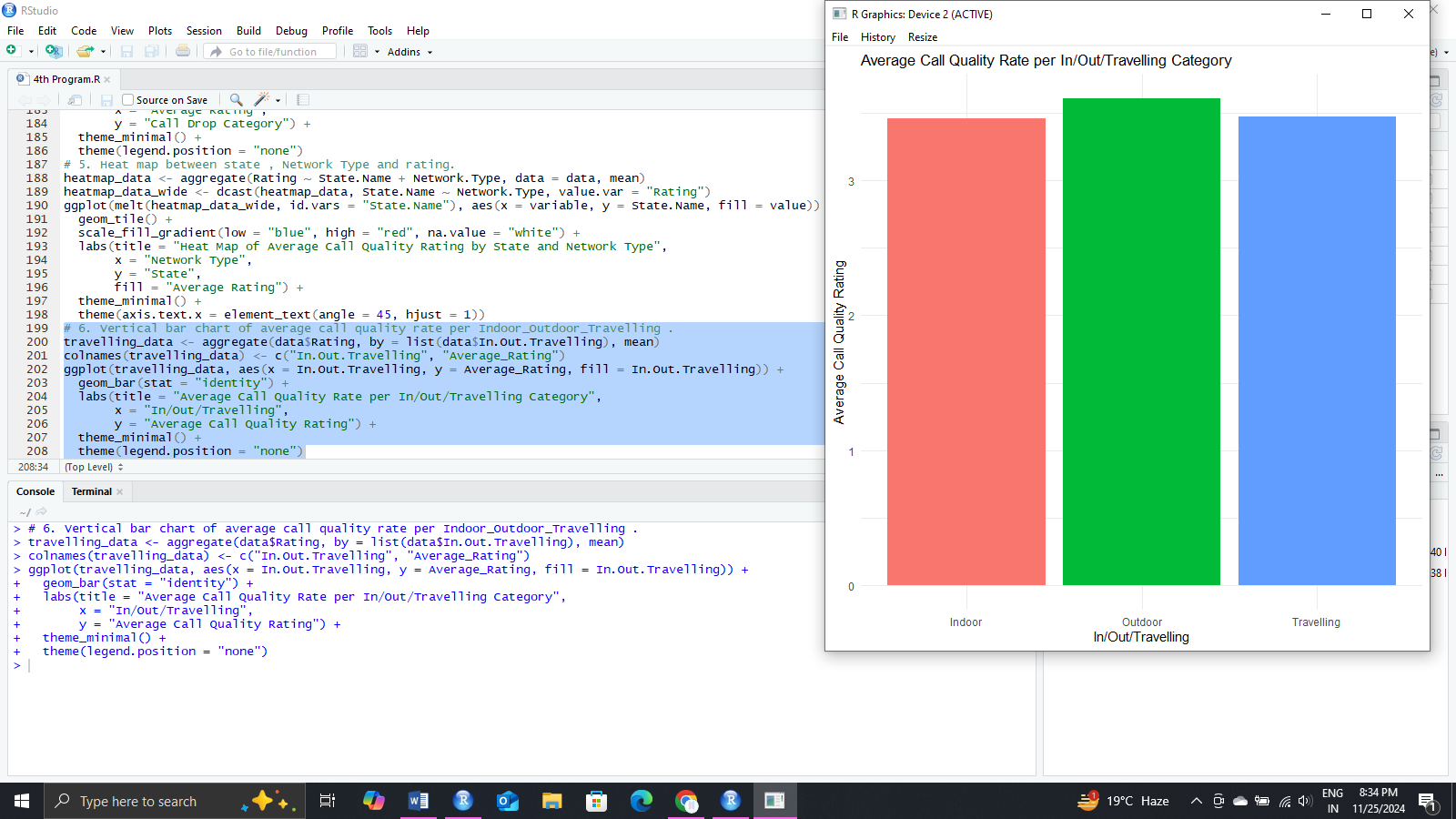


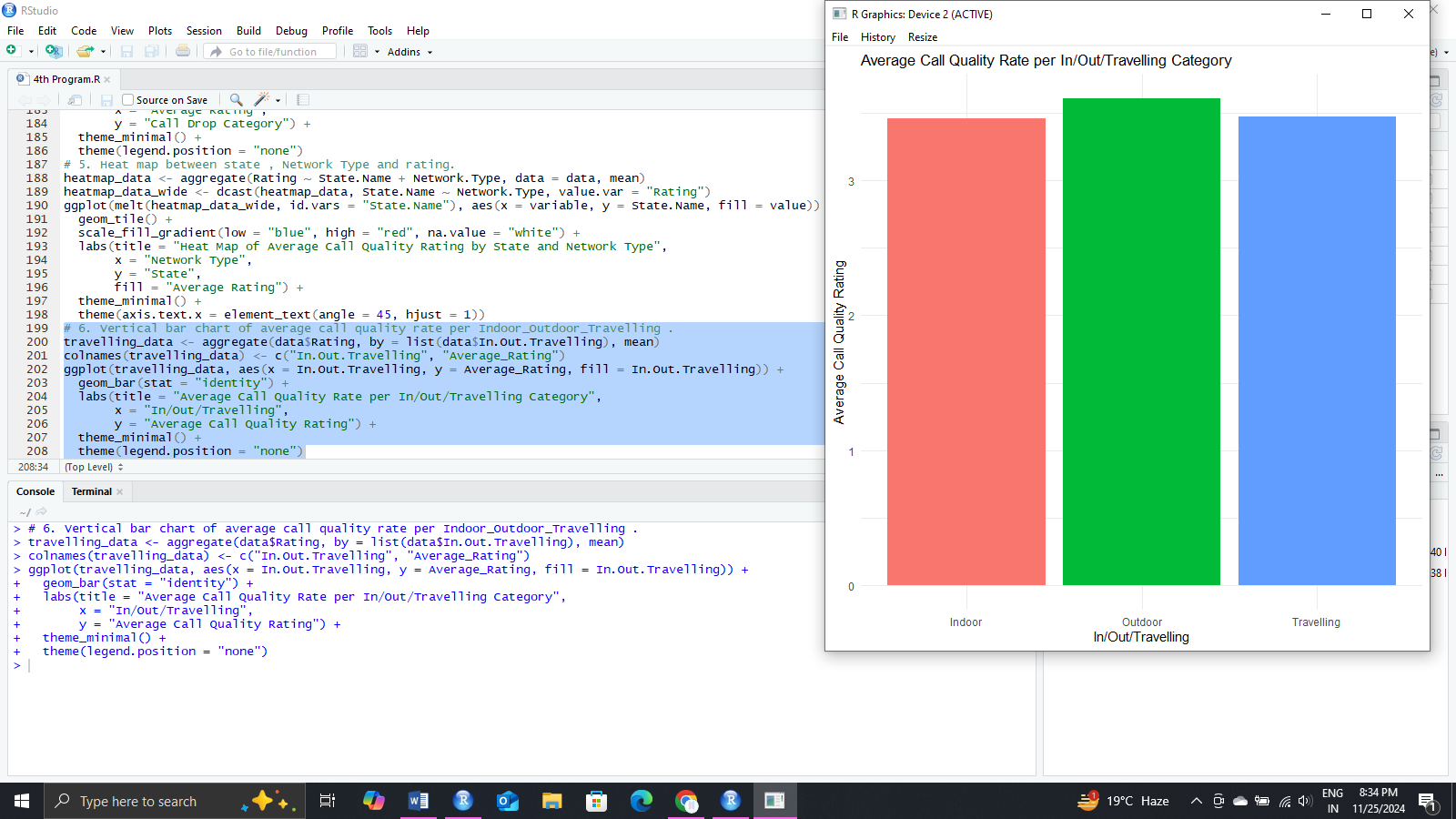
* Vertical bar chart of average call quality rate per Indoor\_Outdoor\_Travelling .
* Data set : CallVoiceQualityExperience-2018-April





* Data set : CallVoiceQuality\_Data\_2018\_May





Answer the following Questions

1. What is the level of overall call quality? (by the operator, by the network.)

* According to Data set : CallVoiceQualityExperience-2018-April

Level of Call Quality Rate is maximum (>4) by the operator “Other” and minimum(1-1.5) by the operator “RCoom”.

Level of Call Quality Rate is maximum (>3.5) by the network type “Unknown” and minimum(2.5-3) by the network type “2G”.

* According to Data set : CallVoiceQuality\_Data\_2018\_May

Level of Call Quality Rate is maximum (5) by the operator “MTNL” and minimum(1) by the operator “RCoom”.

Level of Call Quality Rate is maximum (>3.6) by the network type “Unknown” and minimum(2.5-3) by the network type “2G”.

1. Which operators provide low-value services?

* According to Data set : CallVoiceQualityExperience-2018-April

As the average ratting given to “RComm” Operator in minimum among all the operators it means this operator provide low-value services.

* According to Data set : CallVoiceQuality\_Data\_2018\_May

As the average ratting given to “RComm” Operator in minimum among all the operators it means this operator provide low-value services.

1. What is the frequency of poor quality calls?

* According to Data set : CallVoiceQualityExperience-2018-April

Average poor quality Call Rate per Call Drop is 1.684818

* According to Data set : CallVoiceQuality\_Data\_2018\_May

Average poor quality Call Rate per Call Drop is 1.683681

1. Based on the analysis results, what are the suggestions to improve call quality.

* According to Data set : CallVoiceQualityExperience-2018-April

From the help of Heat Map of Average Call Quality Rating by State and Network Type we can conclude that operators need to increase the quality of different Network Types so that the average rating will increase.

* According to Data set : CallVoiceQuality\_Data\_2018\_May

From the help of Heat Map of Average Call Quality Rating by State and Network Type we can conclude that operators need to increase the quality of different Network Types so that the average rating will increase.