2021

ANALYSIS OF CUSTOMER SATISFACTION IN MOBILE TELECOMMUNICATION INDUSTRY IN INDIA

PROJECT STUDENT

MANISHA GUPTA

SADAF SAYYED

ARZOO SHAIKH

MANOJ CHAUDHARI

LANKESH CHAVAN

PROJECT GUIDE

Dr S.B. MULEY

KISHINCHAND CHELLARAM COLLEGE, CHURCHGATE, MUMBAI 400020

DEPARTMENT OF STATISTICS

University of Mumbai

KISHINCHAND CHELLARAM COLLEGE

DEPARTMENT OF STATISTICS

Churchgate - 400020.

विद्यया विन्दते अमृतम्!

CERTIFICATE

This is to certify that the following students of M.Sc. Part–II, have successfully completed the project entitled

"Analysis of Customer's Satisfaction in Mobile Telecommunication Industry in India."

During the Academic year of 2020–2021
Students involved in this research group are

Manisha Gupta

Sadaf Sayyed

Arzoo Shaikh

Manoj Chaudhari

Lankesh Chavan

This research project is to the best of our knowledge and belief is original

Dr. Asha Jindal
(UG: Head , PG: Coordinator)
Department of Statistics

Dr. S.B.Muley
(Assistant Professor & Project Mentor)
Department of Statistics

AKNOWLEDGEMENT

We would like to express our special thanks of gratitude to our Principal, Dr. Hemlata K. Bagla for encouragement. We would like to thank Dr. Asha Jindal, Associate Professor and Head of the Department (HOD) of Statistics, KC College for constant support and guidance.

We would like to acknowledge our appreciation towards our teachers Mrs. Shailaja Rane, Miss. Ayesha Shaikh, Miss. Kismat Quraishi and Mr. Suraj Singh to keep us motivated.

It is a genuine pleasure to express our deepest gratitude to our mentor Dr. Sakharam Muley for his continuous guidance throughout the research with patience, enthusiasm, immense knowledge and to motivate us to keep moving further to explore and learn new things towards accomplishment of our project.

We would like to acknowledge our appreciation towards our lab assistant. Mr. Santosh Sarwankar of the Department of Statistics for providing the help needed and for his corporation

This research project would not have been possible without the guidance and help of several individuals who, in one or the other way, have contributed and extended their valuable assistance whenever we needed, for completion of the work.

We would like to thank all our responders for being part of our research project and helping us with their vital information, opinions, precious time and valuable feedback. The acknowledgment would have not been complete without thanking our family members, friends, well-wishers and us throughout the journey of the study.

INDEX

CONTENTS

ANALYSIS OF CUSTOMER SATISFACTION IN MOBILE TELECOMMUNICATION INDUSTRY IN INDIA

***Abstract**: Indian Telecom Industry is more than a century old. It is one of the prime support services needed for rapid growth and modernization of various sectors of the economy. It is witnessing great competition. India is one of the vast and leading industries in the world connecting different parts of the country through various modes like telephone, radio, television, satellite and internet. The Telecom Regulatory Authority of India governs this industry by providing a regulatory framework and favourable environment for its efficient operation. The Indian telecom industry stands as the second- largest in the world due to its rapid advancement and is in cut-throat competition with the telecom industries of the other developed countries. The telecommunication services offered by this industry are easily accessible at affordable prices to the customers of urban and rural areas of India. India's telecom network encompasses a highly developed and unique technology in the world. The present study has therefore been undertaken to analyse the view of customer towards the Indian Telecom Industry while emphasizing upon its major segments, their services, etc. The customers satisfaction is the aim of the project for which primary and secondary data is been considered.*

***Keywords:** Indian telecom industry, private sectors, public sectors, prepaid, post-paid, companies: Airtel, Jio, Vodafone, Idea, Vi, BSNL, MTNL.*

INTRODUCTION

Telecommunications also known as telecom is the exchange of information over significant distances by electronic means, referring to all types of voice, data and video transmission. It is the transmission of information by various types of technologies over wire, radio, optical or other electromagnetic systems. A complete telecommunication arrangement is made up of two or more stations equipped with transmitter and receiver devices. A single co-arrangement of transmitters and receivers called a transceiver may also be used in many telecommunication stations. Telecom devices include telephones, telegraph, radio, microwave communication arrangements, fibre optics, satellites and the Internet.

Indian telecommunication industries rank second in the world by the number of telephone users which include both fixed (landline) and mobile phones with about a 1179.49 million subscribers (dated till 31st January 2021). It is also among the lowest call tariffs in the world which is enabled by mega telecommunication operators and has a devastating competition in the world. It also ranks second via internet user base which is about 747.41 million broadband

internet subscribers in the country. The major sectors include television, telephone and internet which is an ongoing process of next generation network transformation.

The Indian Telecom Industry is considered to be a vital tool for the development of the country on the whole by contributing towards the immense growth, quick expansion and upgradation of various sectors of the nation. This industry increases the GDP of India, earns profit for the Indian Government and creates employment opportunities for a great number of people. It is one of the prime support services needed for rapid growth and modernization of various sectors of the economy.

The Indian telecom sector is witnessing great competition. MTNL (Mahanagar Telephone Nigam Ltd.) and BSNL (Bharat Sanchar Nigam Ltd.) are the major public players, whereas Airtel, Idea, Vodafone, Jio , VI are the leading private players in the telecom sector in India. Telecom Industry has supported the socio-economic development of India and plays a significant role to taper the rural-urban digital divide to some dimension. It also has helped to increase the lucidity of governance with the introduction of e-governance in India. The government has practically used modern telecom facilities to give mass education programmes for the rural part of India.

According to London-based telecom trade body GSMA, the telecom sector accounted for 6.5% of India's GDP in 2015, or about ₹ 9 lakh crore (US$130 billion), and supported direct employment for 2.2 million people in the country. GSMA estimates that the Indian telecom sector will contribute ₹14.5 lakh crore (US$200 billion) to the economy and support 3 million direct jobs and 2 million indirect jobs by 2020.

TELEPHONY (LANDLINE & MOBILE)

Most of the companies were formed by restructuring and revolution within decade which was directed by Ministry of Communication and IT, Department of Telecommunication and Minister of Finance. Most companies then gained license to 2G 3G and 4G networks and engaged fixed line mobile and internet business in India. The country code for India is "91".

BSNL and MTNL were allowed to provide land line phone services through copper wire in India until the new telecom policy announced in 1999. Due to rapid growth of cell phones landline started facing competition to maintain the number of subscribers. It dropped from 37.90 million to 20million as of 2020. So now they focus to improve their quality of service.

Chief minister of west Bengal made the first call in India to union telecom minister in august 1995. Sixteen years later in 2012 4G services were launched in Kolkata. India has about 1151.8 million subscribers which makes it the 2nd largest in the world. The dominant players are Jio, Airtel, Vi, BSNL/MTNL.

Following are the market share of the telecom industries in India in percentage:

No table of figures entries found.	Jio	Airtel	Vi	BSNL	TATA teleservices	Reliance communication	Quadrant
Market share of Fixed Line telecommunication operators in India as of 30th September 2020 according to TRAI	11.63%	22.33%	2.58%	52.88%	8.37%	1.27%	0.95%
Market share of mobile telecommunication operators in India as of 31st January 2021 according to TRAI	35.43%	29.72%	24.32%	10.53%	------	-----	-----

GROWTH AND REVENUE

In the telecom service sector, the adjusted gross revenue was about ₹160,814 crore (which is equal to ₹1.8 trillion or US$25.5 billion in 2019) in 2017 as against ₹198,207 crore (which is equal to ₹2.3 trillion or US$32.1 billion in 2019) in 2016 registering a negative growth of 18.87%

Service provider	Calendar year 2018-19 (in INR crores)	Calendar year 2019-20 (in INR crores)	% change
Airtel	80,780.2	87,539.0	+08.37%
Jio	48,660	68,462	+40.69%
Vi	37,823.6	45,996.8	+21.68%
BSNL	19,308	18,906	-02.08%

MOBILE PHONE GENERATION

“G” defined as generation is a set of network standards which details the technological implementation of a specific mobile phone system

1G: It refers to as first generation wireless telephone technology. These are analog (continuous signal for which continuous time varying feature of the signal is representation of other time quantity) telecommunication which were introduced in 1979 and the early till mid 1980s. Its frequency ranges from 150MHz/900MHz. The capacity was 2kbps.

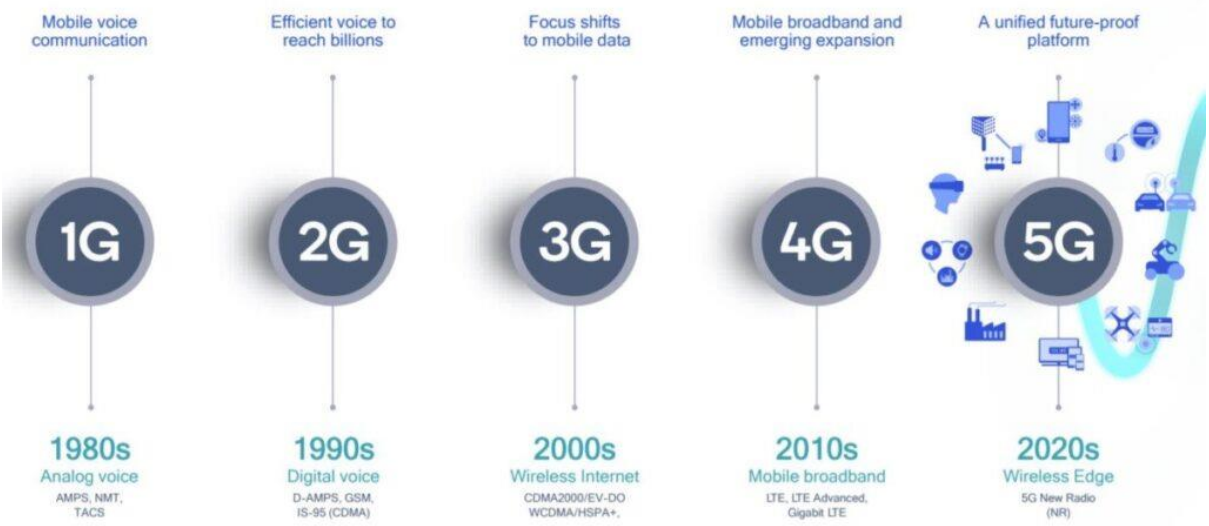
2G: Second generation replaced the first analog generation during the 1980s. The main reason or the difference between the two generation is that 1G system was encoded as analog radio signals where as 2G networks were entirely digital. This generation brought three benefits with it; first was that the phone conversations were now digital, second these systems were more efficient and removed boundaries allowing far penetration levels and third and the most important 2G introduced data services for mobile. They include SMS (Short Message Service), text messages, picture messages and MMS (Multimedia Message Service). 2G were commercially launched on the GSM standard in Finland in 1991. Its frequency was 1.8GHz and capacity was 64kbps

3G: This technology provides an information transfer of at least 144 kbits/s. It can be applied to wireless voice telephony, mobile internet access, fixed wireless internet access, video calls and mobile TV technology. CDMA2000 is a family of 3G mobile technology. Its frequency

ranges from 1.6GHz to 2.0GHz. Its characteristics include digital broadband and increased speed. Its capacity is about 144kbps to 2Mbps

4G: Fourth generation provides faster speed, wireless modems, smartphones, etc. Its application includes IP technology, gaming services, high-definition mobile TV, video conferencing, 3D television, etc. LITE (Long Term Evolution) is commonly marketed as 4G LTE; but it didn't meet the criteria of 4G. Its frequency ranges from 2GHz to 8GHz and its capacity (data range) is 100Mbps to 1Gbps. This generation has better quality, high security and bigger battery usage.

5G: This is the next major phase of mobile telecommunications standard beyond the current 4G standards. The prediction is that with the faster speed it should also need to meet the new cases which is Internet of Things (IoT) and also broadcast. Its capacity might range from 1Gbps to Unlimited.



OBJECTIVES:

The objectives of the research study are:

- To study the growth of telecommunication industry in India
- To determine most preferred network
- Expected benefits from the service provider
- Quality of the current network (comparing with previous 3G/2G networks)
- To study the reason for the growth of private sectors compared the public sector

- How much pandemic (Covid-19) affect the network

METHODOLOGY

The present study is based on primary and secondary data. Primary data is collected via circulating questionnaire in google form on different social media app like whatsapp, fb, etc. In all 587 responses were recorded and considered for analysis. The open and close end questions were included in the questionnaire to study the above listed objectives. Questionnaire was divided into two parts:

1. Respondent's profile
2. Objective based questions
    a. Customer care services
    b. Value of money
    c. Problem faced by customers
    d. Service providers

Secondary data is also considered for the purpose of growth of analysis of telecom industry from the reliable sources like Telecom Regulatory Authority of India (TRAI) published report, published research paper etc.

In order to study the objective responses are been considered and analysis would be done to compute a result for the study. Secondary data considers annual wise growth of the various industries in various segments.

Sampling design:

Non probability sampling method like convenient sampling, Mann Whitney U test, Kruskal Wallis test, Multiple logistic Regression (MLR), Structural Equation Modelling(SEM), Cluster analysis are the different sampling designs used here in the research.

METHODOLOGY TO ANALYSE PRIMARY DATA

To analysis of primary data was carried out using Exploratory Data Analysis (EDA), graph and tables, descriptive statistics, Mann Whitney U test, Kruskal Wallis, and Structural Equation Modelling (SEM), etc

Structural Equation Modelling is used to analyse the impact and relationship between laten construct and observed variables. It includes confirmatory factor analysis, confirmatory

composite analysis, path analysis, partial least squares path modelling, and latent growth modelling. It helps in estimating multiple and interrelated dependencies in a single analysis. There are two main components in the SEM distinguished as – THE STRUCTURAL MODEL that shows dependencies between endogenous variables and exogenous variables, and THE MEASUREMENT MODEL shows relationship between latent variables and their indicators.

Structural Equation Modelling is a flexible framework of data analysis. It helps to understand which customer's perceptions are strongly associated with each other or clusters of the consumer with different perceptions. It can be used with any datatype such as ordinal, nominal, interval, Likert etc. It doesn't require dataset with no missing values.

Multiple Logistic Regression is a predictive model with dichotomous dependent variable (two levels (categorical) labelled as 0 and 1 i.e. binary) and more than two independent variable. This model can be used for future prediction.

The set of variables needed to study has one dependent and three independent variables. The dependent variable namely "**service sector**" has two levels as "government" and "private" sector labelled as 0 and 1 whereas the independent variables "**call connection**", "**internet browsing**" and "**clarity of video**" has Likert scale.

The other set of variables needed to study has one dependent and six independent variables. The dependent variable namely "**service type**" has two levels as "prepaid" and "post-paid" sector labelled as 0 and 1 whereas the independent variables "**clarity of video**", "**internet calling**", "**lag in audio/video**", "**late message delivery**", "**politeness of staff**" and "**offers by service provider**" has Likert scale.

Multiple Logistic Regression will help us to predict a model for these variables.

The Mann Whitney U test is a non-parametric test to compare difference between two independent group of variables. The dependent variable should be of ordinal or continuous scale where as the independent variable should be categorical type having only two levels. The observations should not be normally distributed. The Mann Whitney test is used to do the comparative study between the services provided by telecommunication industries with service sector (government and private) and service type (prepaid and post-paid).

The Kruskal-Wallis test is a nonparametric (distribution free) test, and is used when the assumptions of one-way ANOVA are not met. Both the Kruskal-Wallis test and one-way ANOVA assess for significant differences on a continuous dependent variable by a categorical

independent variable (with two or more groups). In the ANOVA, we assume that the dependent variable is normally distributed and there is approximately equal variance on the scores across groups. However, when using the Kruskal-Wallis Test, we do not have to make any of these assumptions. Therefore, the Kruskal-Wallis test can be used for both continuous and ordinal-level dependent variables. However, like most non-parametric tests, the Kruskal-Wallis Test is not as powerful as the ANOVA.

K-mean clusters

It can be defined as the task of identifying subgroups in the data such that data points in the same subgroup (cluster) are very similar while data points in different clusters are very different. In other words, we try to find homogeneous subgroups within the data such that data points in each cluster are as similar as possible. k-means clustering minimizes within-cluster variances (squared Euclidean distances), but not regular Euclidean distances.

The k-means algorithm is a simple iterative method to partition a given data set into user specified number of clusters k. K-means is a method of cluster analysis. Its aim is to partition n observations into k clusters and each observation will be a part of any one cluster with the nearest mean. Data points are clustered based on feature similarity. Since data under consideration is divided into several sub-groups which mainly focuses on service provided by the customer care like chatbot service, time taken to fix the issue and so on. These services has been majorly divided into two parts government sectors and private sectors

Software: R

In R software data given are clustered by the k-means method, which aims to partition the points into k groups such that the sum of squares from points to the assigned cluster centres is minimized. At the minimum, all cluster centres are at the mean of their Voronoi sets (the set of data points which are nearest to the cluster centre).

Note: k-means cluster dependent upon the initialization of the cluster so values are going to change run by run

Plot: The function *fviz_cluster*() [*factoextra* package] can be used to easily visualize k-means clusters. It takes k-means results and the original data as arguments. In the resulting plot, observations are represented by points, using principal components if the number of variables is greater than 2. It's also possible to draw concentration ellipse around each cluster.)

METHODOLOGY ON SECONDARY DATA

METHODOLOGICAL APPROACH

Secondary data on number of subscribers for the year 2011-12 to 2019-20 was considered for the purpose of analysis. Research majorly focuses to understand the topic with cause and effect relationship. Data is basically based on different service providers like “Bharti Airtel”, “Jio”, “VI (Vodafone Idea)” and “MTNL” and “BSNL” in India. Here we aim to forecast the subscribers count based on the previous year data.

DATA COLLECTION

The data has been collected from the TRAI (Telecom Regulatory Authority of India) website. The TRAI keeps a record of different satellite services on an annual bases along with “Telecom Services” all over India. The pdf for record of each annual is easily available on the website which one can download for their reference. The data consisted of of subscribers/customers in millions annually for each and every service provider. The total count in millions has been used for the further analysis.

SECONDARY DATA ANALYSIS

Before analysis the data was checked for missing data and outliers. There were many service providers available which not existing currently so we have excluded that as there is no use of including. Although the data lies between the year range 2011-12 to 2019-20 but there are some service providers which has stepped into the market in year like 2016-17. There are service providers like Vodafone and Idea which were working individually before 2017-18 afterwards got merged so in such case we have merged prior data on the basis of average. The data was then analysed using statistical software R. A time series was conducted to test the variables, hypothesis and forecast.

A time series is a sequence of data points that occur in successive order over some period of time also we can predict future allowing the data to be gathered in a systematic format. Our aim was to forecast the data based on the present values and our data perfectly suits this conditions and showing a trend over a period of time hence we applied time series.

After importing the data an unexpected error arises that software does not considers the data as a time series data and considers all year values as “character”. Hopefully there is a command in R to convert the data into time series data after which we can proceed with the main analysis and forecasting. For forecasting there are many models present like 1) Benchmark Method 2) ARIMA 3) Exponential smoothing state space method (ets) 4) Holt’s Linear Trend Method.

After working with all the model is the result is that ARIMA model turns out to be much better options for analysis and forecasting which has been decided by comparing AIC value. ARIMA and after that Holt's Linear trend method was under high consideration.

HISTORY

The word telecommunication comes from the Greek where the prefix 'tele' means 'distant'. Then it is combined with Latin word 'communicate' which means 'to share'. Telecommunication technologies include telegraph, telephone, radio, television, videotelephony, satellites, etc.

Given below are some detailed information about the history telecommunication in the world as well as India:

BEACONS AND PIGEONS

Throughout the history homing pigeons have been used by different cultures. Pigeons are effective as messengers due to their natural homing abilities. The pigeons are transported to a destination in cages, where they are attached with messages, then the pigeon naturally flies back to its home where the recipient could read the message. They have been used in many places around the world. Pigeons have also been used to great effect in military situations, and are in this case referred to as war pigeons. Chains of beacons were used on hilltops for relaying signal during the middle age. Though they faced few drawbacks as they could only pass signal bit of information. Later in 1792 , first fixed visual telegraph was built by a French engineer Claude Chappe between Lille and Paris.

TELEPHONE AND INTERNET

The first telephone was invented by Alexander Bell in 1876. He was the first to write down the idea and test a telephone Antonio Meucci invented device which allows electrical transmission of voice over line over thirty years before but this device was a bit practical because it was dependent on electrophonic effect which requires the user to place the receiver in their mouth to hear. However, the 1st official telephone services were set up by the Bell Telephone Company in 1878-79 in the cities of London and New Haven.

In the 1960's the Advanced Research Projects Agency (ARPA) of US department of defence funded research into time sharing computers. ARPANET and other resources sharing networks

were developed in late 1960's. Commercial internet service provider (ISPs) merged in 1989 in US and Australia. the ARPANET was decommissioned in 1990. By 1995, the internet was fully commercialized in the U.S and had tremendous impact on culture and commerce including rise of near instant communication by email, messaging, telephony, video calls and World Wide Web (WWW).

WIRELESS COMMUNICATION

Wireless communication began in 1990 with digital wireless networks leading to a social revolution and ideal shift from wired to wireless technology. It includes cell phones, pagers, mobile telephony, cellular networks, wireless computer network, laptop and Handheld computers. It has driven with advances in Radio Frequency (RF). The key component of RF technology is that it enables digital wireless network.

GROWTH MARKED THROUGH YEAR

1876- Invention of the first telephone by Alexander Graham Bell

1877- Invention of switchboard exchange telephone system

1891- Invention of dial telephones (made it easy to makes calls)

1947- Invention of transistor

1960- Transition from mechanical to electronic switching (enabled features like caller ID , speed dial, voice message, etc)

1984- Introduction of mobile communication beyond two-way radio use

1990s- Use of internet

2000s- Growth of mobile phones (widespread of smartphone by 2012)

INDIAN TELECOM HISTORY

EARLIER INDIA

The Indian postal and telecom sectors are one of the oldest sectors in the world. The first electrical telegraph line was started between Calcutta and diamond harbour in 1850. In 1851 it

was open to use for British East India Company. The construction of 4000miles (6400km) was started in Nov 1853 which connected Mumbai (Bombay), Kolkata (Calcutta), Agra, Peshawar in the north, Chennai (Madras) and Bangalore. A separate department was opened for the people later in 1854. In 1881, the government granted a license to the Oriental Telephone Company Limited of England for opening telephone exchanges at Calcutta, Bombay, Madras and Ahmedabad and the 1st formal telephone service was established in the country.

LIBERALISATION AND PRIVATISATION

In 1981, liberalisation of Indian telecom started when PM Indira Gandhi signed the contract with Alcatel CIT of France to merge with the state which owned telecom company (ITI) to set up 5million lines per year. But soon it was let down due to opposition. The Department of Telecom (DoT) was separated from Indian Post and telecom department in the year 1985. DoT was responsible for telecom until Mahanagar Telephone Nigam Limited (MTNL) and Videsh Sanchar Nigam Limited (VSNL) were carved out to run telecom services of Mumbai and Delhi.

Due to ever increasing demands of telephones Indian government in the 1990s was under pressure to start the telecom sector for private investment as a part of liberalisation-privatisation-globalisation policies that the government had accepted in 1991. As a result, private investment in the sector of Value -Added Services (VAS) was allowed and cellular telecom sector were opened up for competition from private investments.

Government introduced TRAI (Telecom Regulatory Authority of India) in 1997. In 2000, the Vajpayee government constituted Telecom Disputes Settlement and Appellate Tribunal (TDSAT) through TRAI Act. Its objective was to release TRAI from dispute settlement function and to strengthen the regulatory framework. Also, the government corporatized the operations wing of DoT and named it Department of Telecommunication Services (DTS) which was later named as Bharat Sanchar Nigam Limited (BSNL). International business groups wanted the government to privatise VSNL. Finally in April 2002, the government decided to cut its stake of 53% to 26% in VSNL and to throw it open for sale to private enterprises. TATA finally took 25% stake in VSNL. The government further reduced license fees due to which the service fees was reduced and the call cost were cut greatly which enabled all Indian to afford cell phone.

Many private operators such as Reliance Communication, Airtel, Jio, Tata Indicom, Vodafone, Loop, Idea etc successfully entered the high potential Indian telecom market. In the initial 5 to

6 years the average monthly subscribers were around 0.05 to 0.1 million and later in 2002 it turned to be 10.5 million. Later it started growing rapidly which was about 929 million subscribers in May 2012. In 2008, the total GSM and CDMA mobile base was 375 million which was 50% growth with previous year.

CONSOLIDATION

In September 2016, the telecommunication market saw a huge change in terms of falling tariff rates and reduction of data charges. 23th Feb 2017, Telenor India announced that Bharti Airtel will take over all business and assets in India in a year. 14th May 2018 DoT approved the merger of Telenor India with Bharti Airtel. 12th October 2017, airtel announced that it would acquire consumer mobile businesses of TTSL and TTML in debt free, cost free deal.

Reliance Communication closed its 2G and 3G services and now offers only 4G data services from 29th December 2017. In Feb 2019, the company field for bankruptcy as it was unable to sell assets to repay its debt. It has an estimated debt of ₹ 57,383 crore against assets worth ₹18,000 crore. Aircel shut down its unprofitable circles in the states including Gujarat, Maharashtra, Haryana, etc

Vodafone and Idea completed their merger and renamed it as Vi on 31st August 2018. The merger was a success and it created the largest telecom company in India by subscribers and by revenue and the second largest mobile network in terms of number of subscribers in the world. Under the terms of the deal, the Vodafone holds a 45.1% stake in the combined entity, the Aditya Birla Group holds 26% and the remaining shares will be held by the public. However, even after the merger both the brands have been continued to carry their own independent brands.

With all this consolidation, the Indian mobile market has turned into a four-player market, with Jio as the number-one player, with revenue market share of 34%, Airtel India in second position, with revenue market share of 28% and Vi, with revenue market share of 27%. The government operator BSNL/MTNL is in the distant 4th position, with approximately market share of 11%

LITERATURE REVIEW

THE EFFECT ON THE TELECOM INDUSTRY AND CONSUMER AFTER THE INTODUCTION OF RELIANCE JIO: BY ADITYA GUPTA

The telecommunication industry within the sector of information and communication technology is made up of all telecommunication companies and internet services providers and plays the crucial role in the evolution of mobile communication and information society. It is one of the fastest growing industries in the world and has an immense capacity to serve people directly or indirectly. According to Aditya Gupta India is currently the world's second largest telecom market with a subscriber base of 1.19 billion (till 2019) and has registered strong growth in the past decade and half. The exponential growth over the last few years is primarily driven by affordable tariffs, wider availability, roll out of Mobile Number Portability, expanding 3G and 4G coverage, evolving consumption patterns of subscribers and a conducive regulatory environment. The Indian Mobile Industry is expected to create a total economic value of $217.4 billion by 2020. In his research he viewed about the effect cause by JIO in the telecom industry. Mukesh Ambani launched Jio on September 5 2016. Its entry in the industry has created records and reports state that Jio has helped India in saving Billions annually. It introduced extreme cheap data, Free voice call, Increased consumption of online content, proliferation of 4G smartphones, etc. The primary objective was to study the impact of Reliance Jio on the telecom industry, the change in composition of industry, change in market share and the reforms that were undertaken and its effect on the common people and customer behaviour. He used descriptive design in his research which was further divided into two parts the first longitudinal studies which are done by panel and the other is cross sectional study which is collected by population only once in the period of time. The data is collected only from a single segment of the population, so he used single cross section descriptive design. He used Non-Probability sampling method. He concluded his research by defining jio to be as the tsunami in the telecom industry which swallowed every other industry. After the launch of Jio the economic structure of the market is still the same, but the level of competition has grown unexpectedly. Although the limitation of his study was that the data was limited to urban area only and the age group was 15-20 years old and also with a limited number of size 100.

INDIAN TELECOM INDUSTRY: A CASE STUDY OF AIRTEL AND VODAFONE IN NCR DELHI: BY DR SUKHVIR SINGH

According to Dr Sukhvir Singh Mobile Telephony is an automatic, battery, two-way radio communication system. It is connected via radio signals to a nearby base station of the mobile telephone network. The cellular system comprises of the mobile handset and the mobile network. Mobile Telephony is an automatic, battery, two-way radio communication system. It is connected via radio signals to a nearby base station of the mobile telephone network. The cellular system comprises of the mobile handset and the mobile network. In his research he has taken two industries under consideration which is Airtel and Vodafone. The objective of his study was to compare the Airtel and Vodafone on the basis of price premium, satisfaction/loyalty, quality and leadership, Perceived value, brand personality, market share, etc. Data was collected by questionnaire based on market survey where both open-ended and close-ended questions were asked. The sampling used for his study was probability sampling. Since the study was only meant for certain specific categories within the total population (cell phone users, in this case), a stratified random sample was used. The conclusion of the research was based on which company among the two fulfilled the which objective mentioned which claimed Airtel to be superior. The limitation of this study was that it is only about two industries and no analysis is been showed.

EMPLOYEE DEVELOPMENT PRACTICE IN TELECOM INDUSTRY IN INDIA - A CASE STUDY ON PUBLIC SECTOR: BY VASUDHA DHINGRA

The telecom services have been recognized the world-over as an important tool for socio-economic development of a nation. It is one of the prime support services needed for rapid growth and modernization of various sectors of the economy. MTNL (Mahanagar Telephone Nigam Ltd.) and BSNL (Bharat Sanchar Nigam Ltd.) are the major public players, whereas Jio, Bharti Airtel, Idea, Vodafone are the leading private players in the telecom sector in India. Vasudha Dhingra in her research paper where she has studied public sector which has a huge infrastructure and had early mover advantage, its performance is low as compared to its counterpart in private sector. In spite of public sector having the inherent advantages of being pioneer in the trade, one of the key factors behind its low performance, is the lack of motivation amongst its human resources. Due to that reason, her study was designed to have an insight into the implementation of Employee development practices in the public sector units of the telecom industry and to find the lacunae. The research has undertaken certainly throws light upon various aspects where the top level management in public sector telecom companies needs to work out. The objective of her study was to assess and compare the extent of implementation of selected employee development practices in the public sector units of the

telecom industry in India. The method used to collect data was through questionnaire or we can say primary data was taken under consideration. The result concludes that the employees working in the public sector of the telecom industry agree that development practice in the public sector companies should not be a one-time affair, but it should be continuous. According to the respondents there is lack of retraining. This eventually reduces the frequency of the development practices. Development programmes are not systematically planned. They are according to the need of the situation. Public sector companies need to step up their outlay on employee development and training. The number of hours devoted should be gradually increased. With the Government of India extending hands to bring the loss in making units on their feet, the public sector units will be getting a massive investment over the next five years. Talking about the limitation of this research, there's no contribution of people not using a public sector. Also, it does not state anything about private sector and also secondary data was not taken under consideration so it lacks a factor of comparing.

A STUDY ON TELECOM SERVICE IN INDIA: BY D R. MUTHUKRISHNAVENI

Telecommunication is the transmission of signs, signals, messages, words, writings, images and sounds or information of any nature by wire, radio, optical or other electromagnetic systems. Telecom occurs when the exchange of information between communication participants includes the use of technology. Dr R. MuthuKrishnaveni an assistant professor of Commerce, had studied basic about telecom industry of India. His objectives were basic i.e. studying the process and listing out services providers. The research objective is based on descriptive and analytical study. The data collected for this study from books and websites and the tool for analysis was percentage. The conclusion states that in India, 85.77% (Telecom subscribers 1109.67 million, total population approximately 1388.23 million) of population use Telecom services. BSNL is a market leader in Landline network, Bhrathi Airtel play leading role in Mobile network and Reliance Jio has more number of subscribers in Internet service. The limitation for this study is that they is no contribution of primary or secondary data the results are evaluated using books and website and also the study is very short giving less knowledge about the topic.

AN ANALYSIS ON INIDAN TELECOM INDUSTRY: BY MS. PRITISH AND DR TARUNA SAXENA

The Indian Telecom Industry is considered to be a vital tool for the development of the country on the whole by contributing towards the immense growth, quick expansion and upgradation

of various sectors of the nation. This industry increases the GDP of India, earns profit for the Indian Government and creates employment opportunities for a great number of people. The Indian Telecom Industry is very huge consisting of companies that make hardware and also produce software. Ms. Pritish and Dr. Taruna Saxena have done analysis on the Indian telecommunication industry. The objectives of the research study were to analyse the history and evolution of Indian Telecom Industry, review the Government Telecom policies, identify the present trends in the Indian Telecom Industry and its growth and to study the future growth opportunities in the Indian Telecom Industry. They considered secondary data from the Telecom Regulatory Authority of India (TRAI), Department of Telecommunication (DoT) and the reports from Government of India and other sources. Different telecom magazines, newspapers and journals were consulted for gathering of information. Information was also collected by holding discussions and interviews with knowledgeable persons employed at different levels in various telecom companies of India. In order to achieve the objectives of the study, year- wise annual growth of the industry in its various segments, percentage share of different service providers per year were calculated. Though the conclusion doesn't state much. It only concluded that Indian Telecom Industry contributes significantly to the overall socioeconomic development of India and is an essential tool for the growth of the nation. Limitation of the paper is that no public view is been considered as the data is secondary and also the result doesn't provide much of the information.

AN OVERVIEW OF INDIAN TELECOM SECTOR: BY ZARAQ ZAHOOR

Communication is a hugely important aspect, not only for people around the world, but also for small and large businesses. Long distance communication has been around for years with the oldest methods that can be remembered to date being the use of smoke signals. With time, methods such as horns became a means of communication. But with time there has been a lot of development and with that came the more advanced technologies such as radio, phone, television and the Internet. Zaraq Zahoor has overview Indian telecom industry in his research. The study throws light on the evolution of telecom sector in India. It is descriptive in nature where the focus is on providing and understanding the growth, importance, facilitating factors and challenges of telecom sector. For this purpose, secondary data has been collected through books, journals, websites, govt. reports etc. He concluded his report by saying that telecommunication has emerged as a key driver of economic and social development in an increasingly knowledge intensive global scenario. The cellular segment is playing an important role in the industry by making itself available in the rural and semi urban areas where tele-

density is the lowest. Besides contributing to about 3% to India's GDP, Telecommunications, along with Information Technology, has greatly accelerated the growth of the economic and social sectors and will continue to do so in future. Limitation of this research is again that it has not considered any primary data it is based on something which is already there in the market.

TELECOM SERVICES: EMERGING TRENDS, OPPORTUNITIES AND RISK: BY DR M. PRASANNA KUMAR

The telecom services have been recognized the world-over as an important tool for socio-economic development for a nation. Driven by various policy initiatives, the Indian telecom sector witnessed a complete transformation in the last decade. It has achieved a phenomenal growth during the last few years and is poised to take a big leap in the future also. Dr M. Prasanna Kumar studied emerging trends, risk and opportunities in telecom industry which are also the objectives of his study and is based on the analysis of the secondary data published in the magazines and various websites. So, he concludes that telecom services providing a lot of opportunities, emerging new trends but at the same times a lot of risk also there in telecom sector. In spite of rapid growth, the tele-density is still 75.23% million which is very good when compared with the previous year population. With its ongoing all round efforts in building up the economy, the country is touch the of 93.1.95 million total telecom subscribers by 28 February 2014 which is good sign to make Indian telecom network, the third largest in the World.

Conclusion:

After reviewing all the research papers above mentioned and also many others which are not mentioned here, somewhere or other all the study have their own limitations and drawbacks. For example, some of the papers have only considered secondary data well some have considered primary data. Also, most of the papers focus on a particular industry or topic which they conclude on. But viewing our research, our study covers all the flaws and limitation faced by other papers. We have considered both primary and secondary data, our analysis is based on every question asked with variation in methods. Public as well as private sector was taken under consideration. All the current industries in India were mentioned in our survey. The questionnaire was set is such a way that even the least part of the population could be a part of it. The survey has taken all the part of generations under consideration with three different

languages for convenience. So our research stands with a variation which can provide great information further ahead.

DATA ANALYSIS

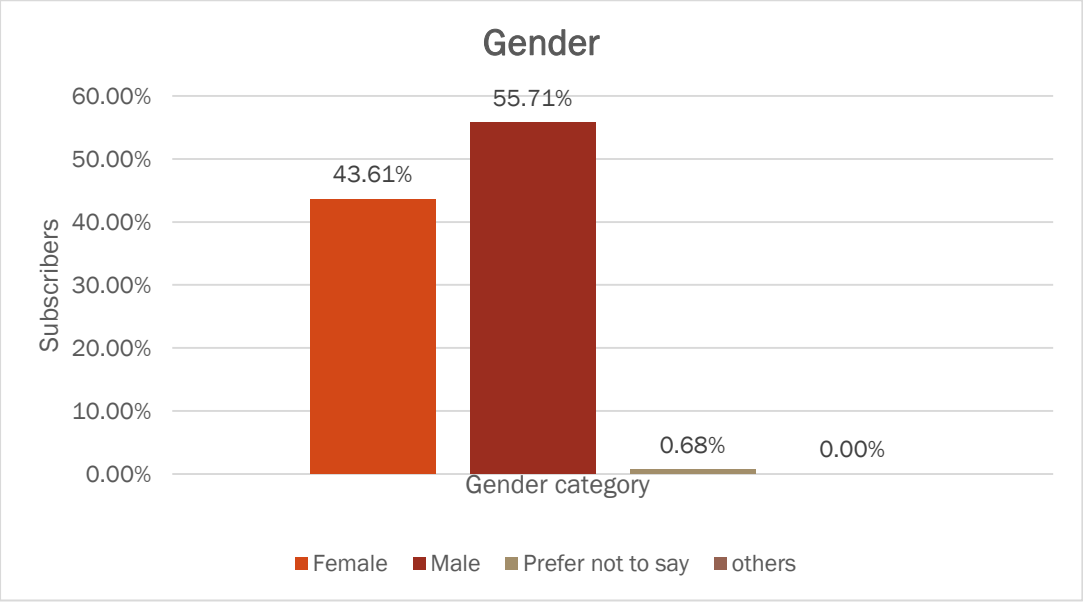
1. RESPONDENT PROFILE AND BASIC DATA DISTRIBUTION

Gender-

Table No:1.1 Table showing the distribution of Gender.

Gender	Female	Male	Prefer not to say	Others	Total
Count	**256**	**327**	**4**	**0**	587
In %	43.61%	55.71%	0.68%	0.00%	100%

Graph No. 1.1 Graph showing the distribution of Gender.



The Bar chart illustrates the count of men, women who are using Network. The charts give information about the comparison between men and women using Network in google form survey. As an overall trend, male count is more who are using network as compared to female.

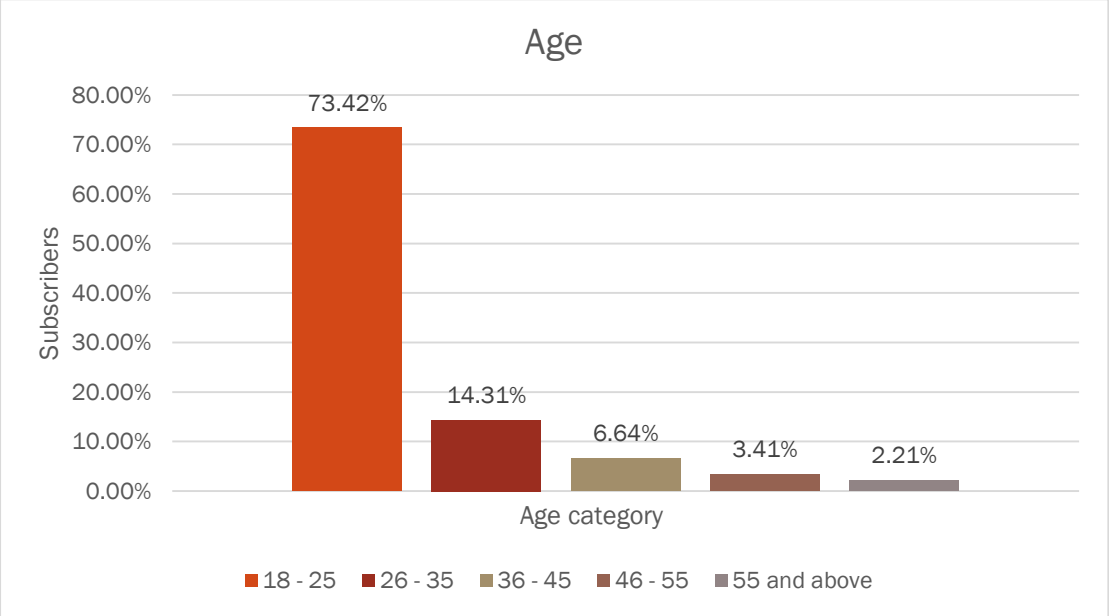
Total 587 people participated in the survey. Out of 587 people, 327 were male while 256 were female. There were 4 who preferred not to disclose their gender.

Age –

Table No: 1.2 Table showing the distribution of Age.

Age limit	18 - 25	26 – 35	36 - 45	46 – 55	55 and above	Total
Count	431	84	39	20	13	587
In %	73.42%	14.31%	6.64%	3.41%	2.21%	100%

Graph No. 1.2 Graph showing the distribution of Age.



The bar chart shows the information of age distribution of the customer in the survey. Overall, it can be seen from the graphs that the majority of people aged below 25. 73.42% customer is of the age group 18 to 25 using Network.

The first bar chat presented that students under aged 25 who are using network with 73.42% highest among other aged group compared to adult aged over 26 with only 14.31% which is exactly opposite of the other. As age increased it is noted that customer using network gradually decrease.

Furthermore, the second bar chart shows the customer in the age from 26 to 35 which comprises of 14.31%.

6.64% Customer are from age group 36 to 45. A less number of customer belongs to age group 46 to 55 that is only 3.41%.

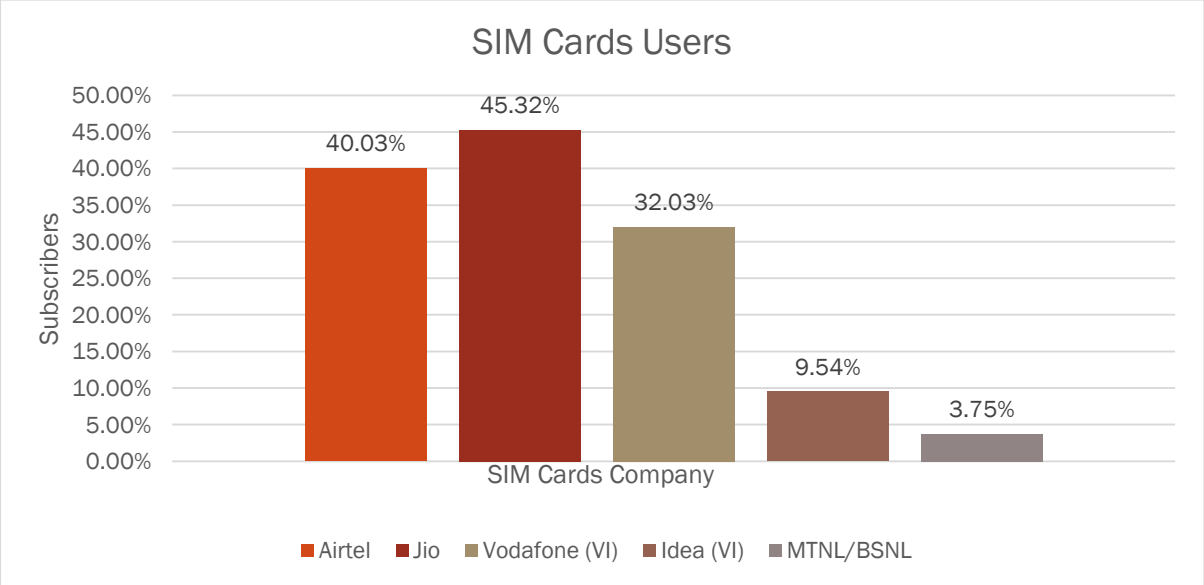
Only 2.21% customer are of age group 55 and above.

Q1. Which network are you using now? (If dual SIM user can choose more than 1)

Table No: 1.3 Table showing the distribution of SIM Cards Users**.**

SIM cards	Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL	Total
Count	235	266	188	56	22	587
In %	40.03%	45.32%	32.03%	9.54%	3.75%	100%

Graph No. 1.3 Graph showing the distribution of SIM Cards Users.



The bar chart illustrates the distribution of the SIM card users (dual SIM also).

The given illustration gives data on the number of sim card users for five different companies namely: Airtel. Jio, Vodafone, Idea and MTNL/BSNL

As is presented in the chart Airtel, Jio and Vodafone have more sim card users than the Idea and MTNL/BSNL

The highest number of sim cards are of JIO. It comprises of total 45.32%.

Idea sim cards users are less as compared to Vodafone. User less than 3.75% are using MTNL/BSNL sim card.

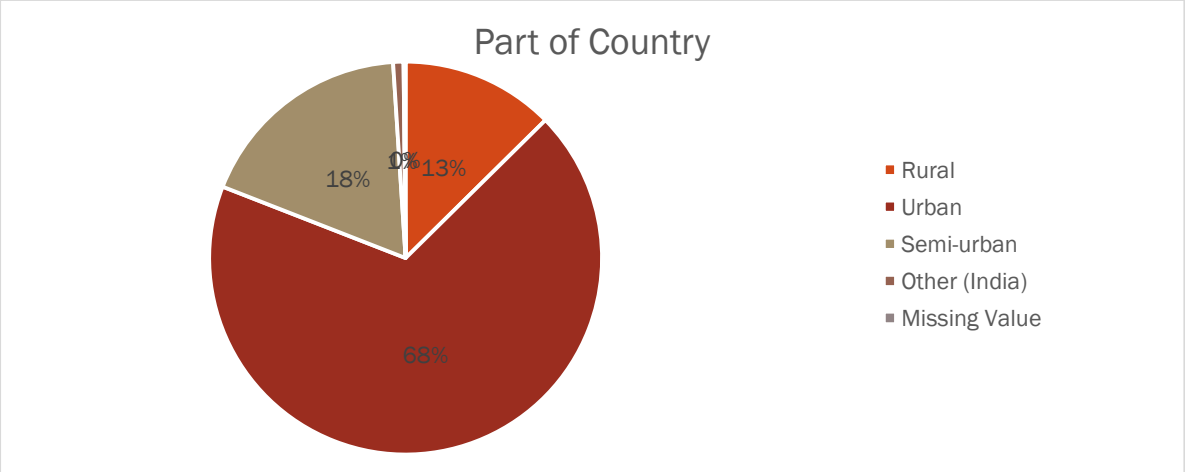
To sum up, it is clear from the graph that, maximum number of users are using JIO, Airtel and Vodafone sim card.

Q2. From which part of the country you belong to?

Table No: 1.4 Table showing the distribution of Part of Country**.**

Part of Country	Rural	Urban	Semi-urban	Other (India)	Missing Value	Total
Count	74	401	106	5	1	587
In %	12.61%	68.31%	18.06%	0.85%	0.17%	100%

Graph No. 1.4 Graph showing the distribution of Part of Country.



The pie chart illustrates the distribution of users across different Part of Country. The maximum user belongs to Urban, followed by semi urban and rural.

Total 68.31% user belongs to urban region. 18.06% are from Semi urban and 12.61% are from rural region.

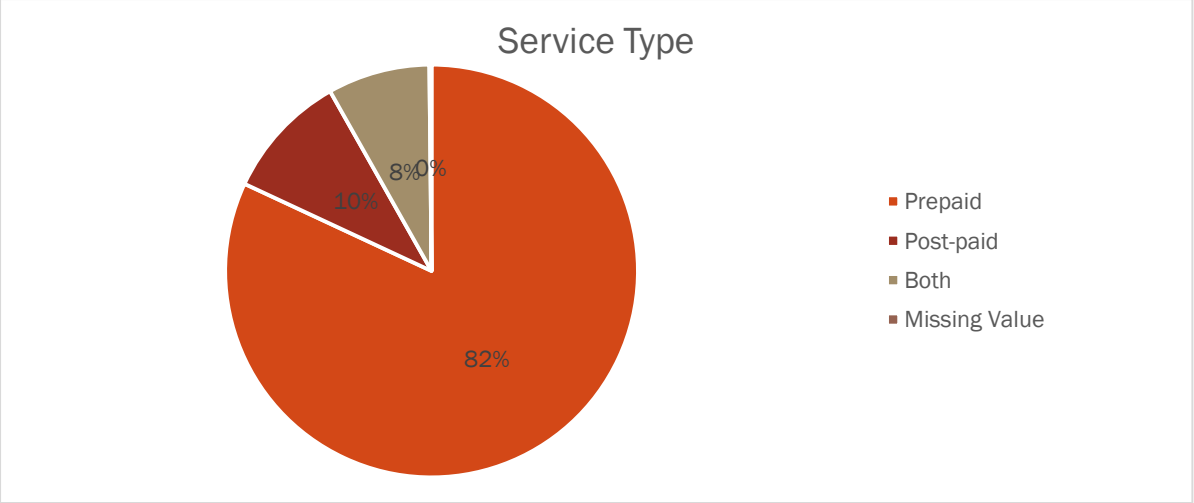
To summarize, Network consumption are high is Urban region as compared to Rural

Q3. What kind of service you prefer?

Table No: 1.5 Table showing the distribution of Service Type.

Services	Prepaid	Post-paid	Both	Missing Value	Total
Count	481	58	47	1	587
In %	81.94%	9.88%	8.01%	0.17%	100%

Graph No. 1.5 Graph showing the distribution of Service Type.



The pie chart shows the distribution of Service type. Survey was conducted for 587 users, out of which 481 users that is 81.94% are using prepaid card.

On the other hand, only 9.88% are using post-paid, whereas 8.01% users are using both service type of card that is prepaid and Post-paid.

Overall it clears that, maximum number of users are using Prepaid.

Q4. How long are you a customer/subscriber of your network service provider?

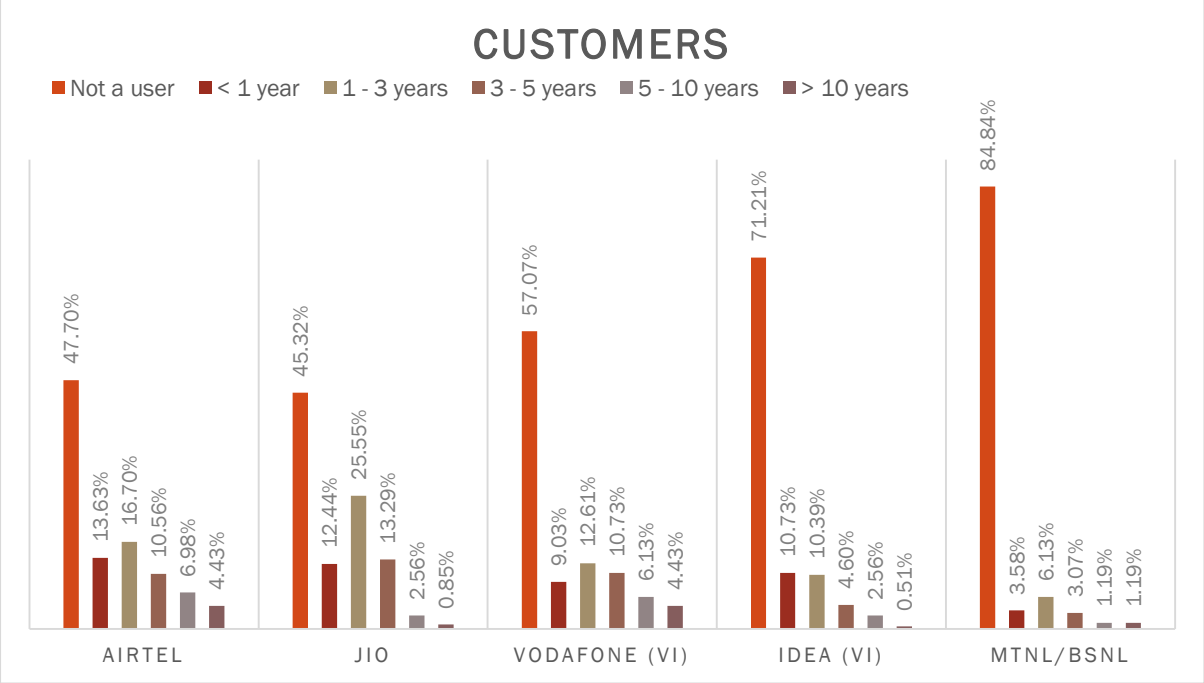
Table No: 1.6 Table showing the distribution of Customers.

	Not a user	**< 1 year**	**1 - 3 years**	**3 - 5 years**	**5 - 10 years**	**> 10 years**	**Total**
Airtel	280	80	98	62	41	26	587
Jio	266	73	150	78	15	5	587
Vodafone (VI)	335	53	74	63	36	26	587
Idea (VI)	418	63	61	27	15	3	587
MTNL/BSNL	498	21	36	18	7	7	587

	Not a user	**< 1 year**	**1 - 3 years**	**3 - 5 years**	**5 - 10 years**	**> 10 years**	**Total In %**
Airtel	47.70%	13.63%	16.70%	10.56%	6.98%	4.43%	100%
Jio	45.32%	12.44%	25.55%	13.29%	2.56%	0.85%	100%

Vodafone (VI)	57.07%	9.03%	12.61%	10.73%	6.13%	4.43%	100%
Idea (VI)	71.21%	10.73%	10.39%	4.60%	2.56%	0.51%	100%
MTNL/BSNL	84.84%	3.58%	6.13%	3.07%	1.19%	1.19%	100%

Graph No. 1.6 Graph showing the distribution of Customers.



Survey was conducted on 587 users. The bar chart shows the time duration of the users using network of the service provider.

When we look at statics of Airtel provider, maximum users are using Airtel network from last one to three years. 62 users are using this network from last five years.

We can see that there are users using Airtel network for 10 years also. 41 users have confirmed on using Airtel from past five years and 26 for last ten years.

On the other hand, JIO which is the newest service provider has users who are using JIO cards since inception. There are 150 users that is 25.55% users are using JIO from 1 to 3 years, 13.29% users are using JIO from 3 to 5 years and 73 users from less than one years.

9.03% Users started using Vodafone in last one years. 12.61% users are using Vodafone from 1 to 3 years. 10.73% users are using Vodafone card from 3 to 5 years.

We can see the count of users using Vodafone and Airtel from 5 to 10 years and more than that are nearly the same.

Airtel and Vodafone remains the old service provider as compared to Jio.

Less than 11% users are using idea from last 3 years.

Only 6.13% users are using MTNL/BSNL from last 3 years.

Q5. Rate the quality of the service provided by your network service provider. (If you are a dual SIM user then fill the details of the most preferred network).

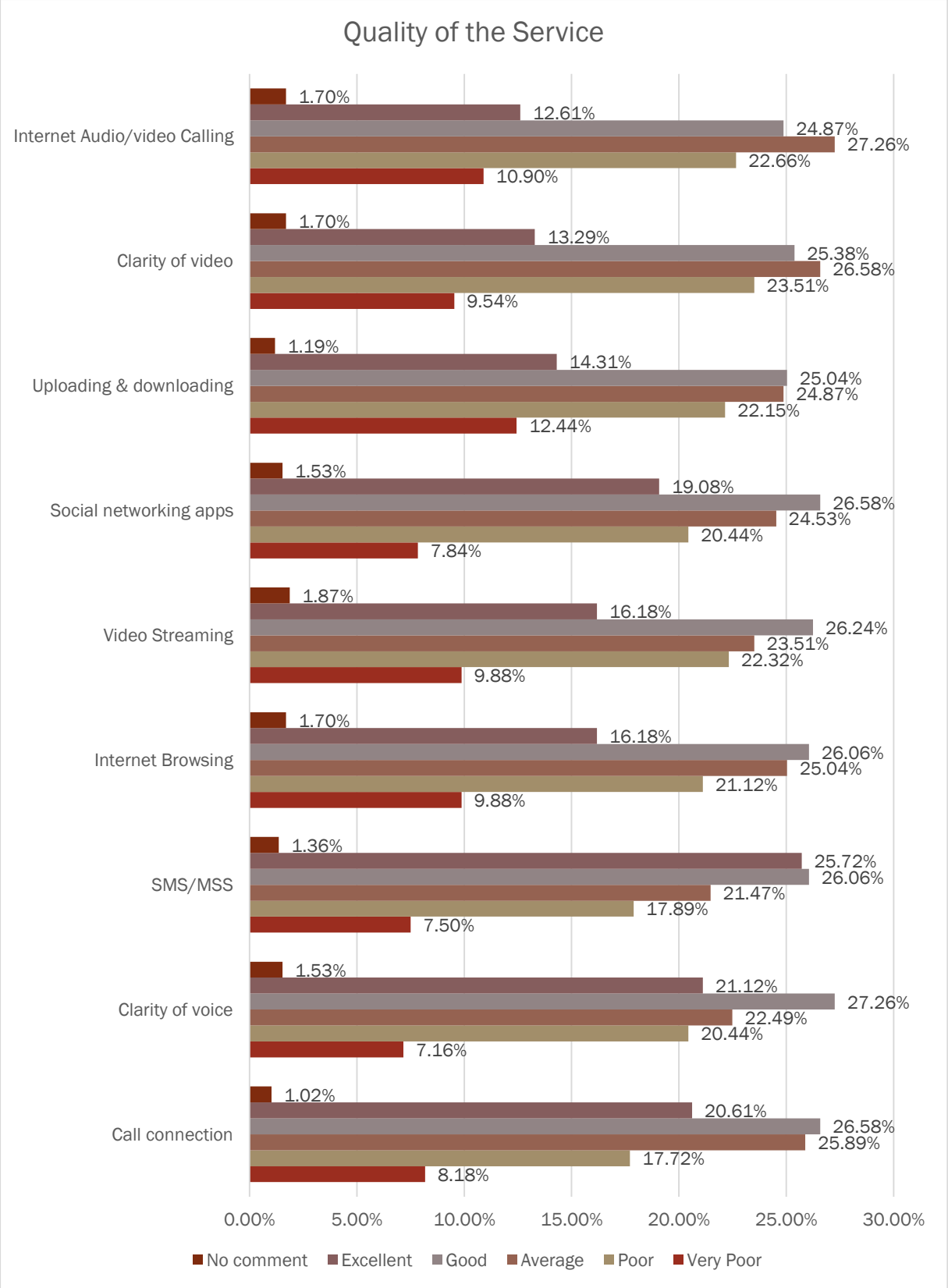
Table No: 1.7 Table showing the distribution of Quality of Service.

	Very Poor	Poor	Average	Good	Excellent	No comment	Total
Call connection	48	104	152	156	121	6	587
Clarity of voice	42	120	132	160	124	9	587
SMS/MSS	44	105	126	153	151	8	587
Internet Browsing (google chrome, etc.)	58	124	147	153	95	10	587
Video Streaming (Youtube, Instagram, etc.)	58	131	138	154	95	11	587
Social networking apps (Instagram, Fb, WhatsApp, etc.)	46	120	144	156	112	9	587
Uploading & downloading data/ file	73	130	146	147	84	7	587
Clarity of video	56	138	156	149	78	10	587

Internet Audio/video Calling (social media calls)	64	133	160	146	74	10	587

	Very Poor	**Poor**	**Average**	**Good**	**Excellent**	**No comment**	**Total In %**
Call connection	8.18%	17.72%	25.89%	26.58%	20.61%	1.02%	100%
Clarity of voice	7.16%	20.44%	22.49%	27.26%	21.12%	1.53%	100%
SMS/MSS	7.50%	17.89%	21.47%	26.06%	25.72%	1.36%	100%
Internet Browsing (google chrome, etc.)	9.88%	21.12%	25.04%	26.06%	16.18%	1.70%	100%
Video Streaming (Youtube, Instagram, etc.)	9.88%	22.32%	23.51%	26.24%	16.18%	1.87%	100%
Social networking apps (Instagram, Fb, WhatsApp, etc.)	7.84%	20.44%	24.53%	26.58%	19.08%	1.53%	100%
Uploading & downloading data/ file	12.44%	22.15%	24.87%	25.04%	14.31%	1.19%	100%
Clarity of video	9.54%	23.51%	26.58%	25.38%	13.29%	1.70%	100%
Internet Audio/video Calling (social media calls)	10.90%	22.66%	27.26%	24.87%	12.61%	1.70%	100%

Graph No. 1.7 Graph showing the distribution of Quality of Service.



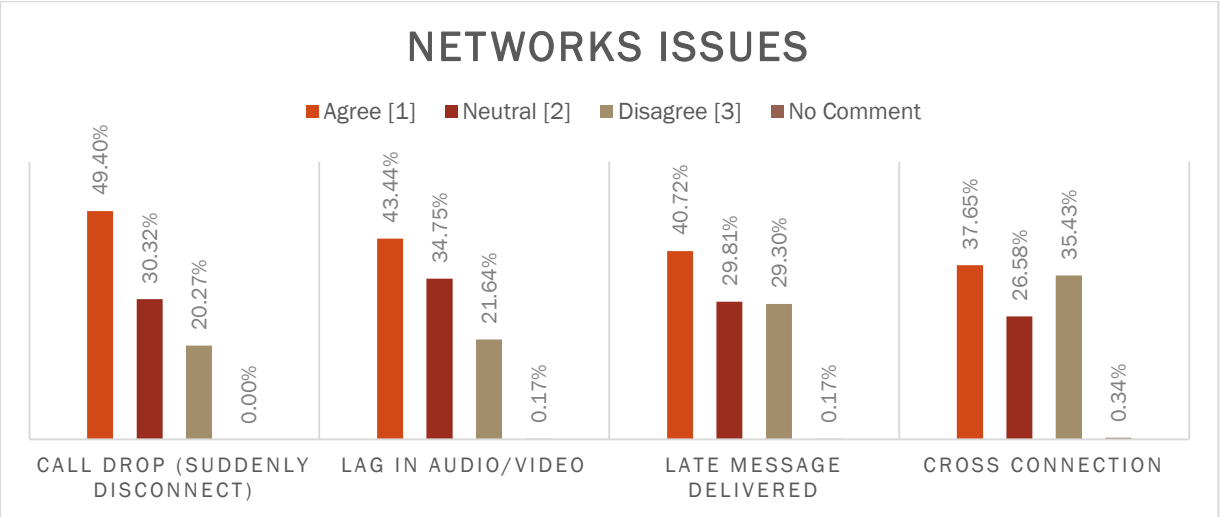
Q6. Did you ever receive following problems?

Table No: 1.8 Table showing the distribution of Networks Issues.

	Agree [1]	Neutral [2]	Disagree [3]	No Comment	Total
Call drop (suddenly disconnect)	290	178	119	0	587
lag in audio/video	255	204	127	1	587
late message delivered	239	175	172	1	587
cross connection	221	156	208	2	587

	Agree [1]	Neutral [2]	Disagree [3]	No Comment	Total
Call drop (suddenly disconnect)	49.40%	30.32%	20.27%	0.00%	100%
lag in audio/video	43.44%	34.75%	21.64%	0.17%	100%
late message delivered	40.72%	29.81%	29.30%	0.17%	100%
cross connection	37.65%	26.58%	35.43%	0.34%	100%

Graph No. 1.8 Graph showing the distribution of Networks Issues.



The bar chart illustrates the Network issue raised by users. 49.40% users have reported Call drop (suddenly disconnect) issue, 30.32% user remains neutral for Call drop (suddenly disconnect) issue, 20.27% user remains disagree for Call drop (suddenly disconnect) issue.

The bar chart illustrates the Network issue raised by users. 43.44% users have reported lag in audio/video issue, 34.75% user remains neutral for lag in audio/video issue, 21.64% user remains disagree for lag in audio/video issue.

40.72% have agreed to late message delivered issue.

On the other hand there are issues like cross connection which users have reported. 37.65% user says there are cross connections issue observed.

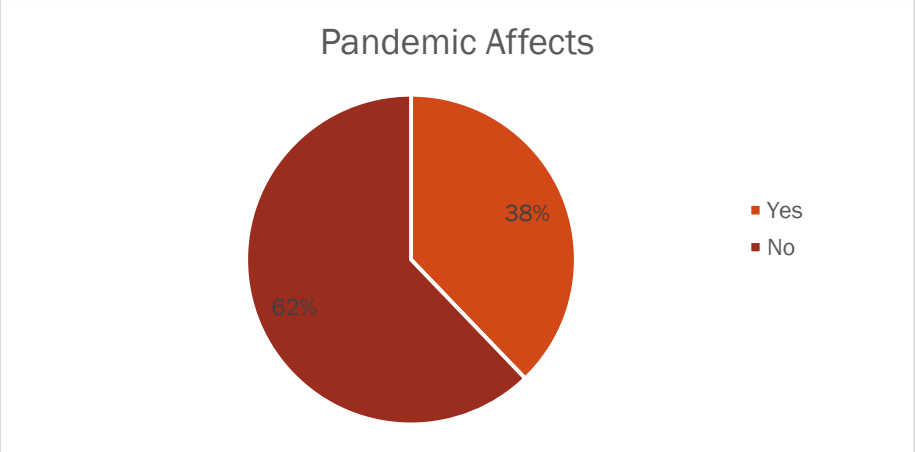
Q7. Did pandemic affect your network service?

Table No: 1.9 Table showing the distribution of Pandemic Affects.

Pandemic affects	Yes	No	Total
Subscribes	222	365	587

Pandemic affects	Yes	No	Total
Subscribes	37.82%	62.18%	100%

Graph No. 1.9 Graph showing the distribution of Pandemic Affects.



The bar chart illustrates distribution of Pandemic Affects.

There are 62.18% users who says there are no affect on Pandemics. Users continue to use the service from their service provider.

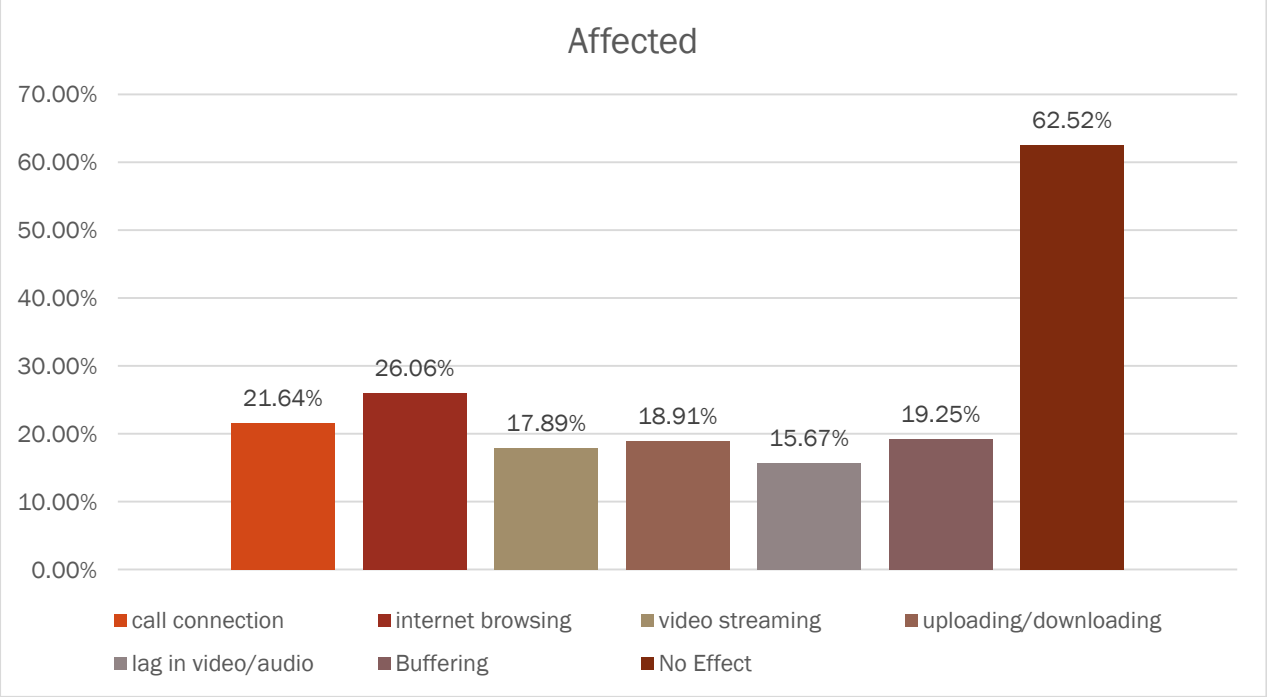
There are 37.82% subscriber who says that there are affect of Pendemic.

If yes, choose from the below option which section was affected the most?

Table No: 1.10 Table showing the distribution of Pandemic Affects.

call connection	internet browsing	video streaming (YouTube, Instagram, etc.)	uploading/ downloading data/file	lag in video/audio	Buffering	No Effect
127	153	105	111	92	113	367
21.64%	26.06%	17.89%	18.91%	15.67%	19.25%	62.52%

Graph No. 1.10 Graph showing the distribution of Pandemic Affects.



37.82% users have said that there are affect on Pandamic.

21.64% Users have reported call connection. 26.06% users have reported internet browsing issue during Pandemic.

There are other issues like video streaming (YouTube, Instagram, etc.), uploading/downloading data/file, lag in video/audio and Buffering which is reported.

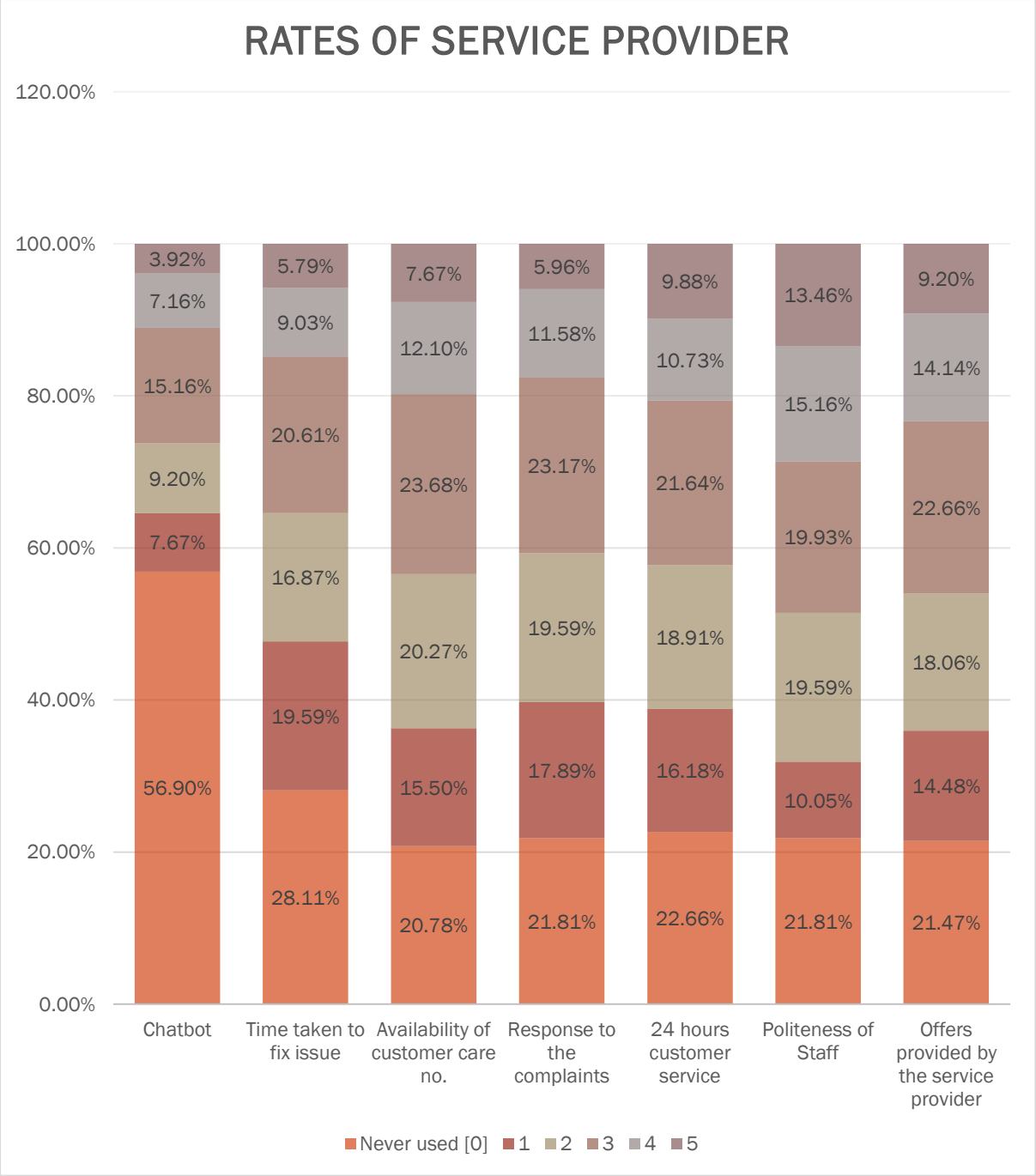
Q8. Rate the service provided by the customer care of your network?

Table No: 1.11 Table showing the distribution of Rates the Service Provider.

	Never used [0]	**1**	**2**	**3**	**4**	**5**
Chatbot	334	45	54	89	42	23
Time taken to fix issue	165	115	99	121	53	34
Availability of customer care no.	122	91	119	139	71	45
Response to the complaints	128	105	115	136	68	35
24 hours customer service	133	95	111	127	63	58
Politeness of Staff	128	59	115	117	89	79
Offers provided by the service provider	126	85	106	133	83	54

	Never used [0]	1	2	3	4	5
Chatbot	56.90%	7.67%	9.20%	15.16%	7.16%	3.92%
Time taken to fix issue	28.11%	19.59%	16.87%	20.61%	9.03%	5.79%
Availability of customer care no.	20.78%	15.50%	20.27%	23.68%	12.10%	7.67%
Response to the complaints	21.81%	17.89%	19.59%	23.17%	11.58%	5.96%
24 hours customer service	22.66%	16.18%	18.91%	21.64%	10.73%	9.88%
Politeness of Staff	21.81%	10.05%	19.59%	19.93%	15.16%	13.46%
Offers provided by the service provider	21.47%	14.48%	18.06%	22.66%	14.14%	9.20%

Graph No. 1.11 Graph showing the distribution of Rates the Service Provider.



Q9. Rate your cost and offers given by the service provider?

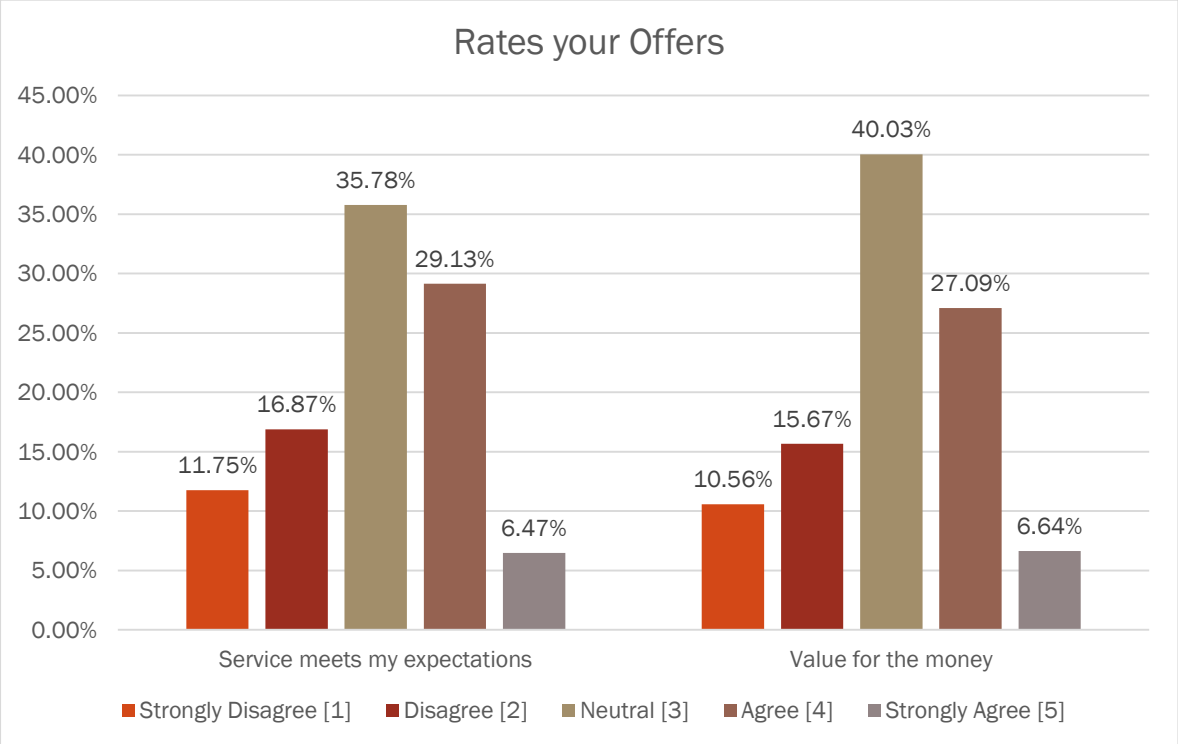
Table No: 1.12 Table showing the distribution of Rates your Offers.

	Strongly Disagree [1]	Disagree [2]	Neutral [3]	Agree [4]	Strongly Agree [5]
Service meets my expectations	69	99	210	171	38

Value for the money	62	92	235	159	39

	Strongly Disagree [1]	**Disagree [2]**	**Neutral [3]**	**Agree [4]**	**Strongly Agree [5]**
Service meets my expectations	11.75%	16.87%	35.78%	29.13%	6.47%
Value for the money	10.56%	15.67%	40.03%	27.09%	6.64%

Graph No. 1.12 Graph showing the distribution of Rates your Offers.



The given data service provider meets expectation in term of plan and data and also network coverage in all possible location.

11.75% strongly disagree and 16.87% disagree and 35.78% are neutral and 29.13% agree and 6.47% are strongly agree.

The given data service provider value for the money in term of plan and data and also network coverage in all possible location.

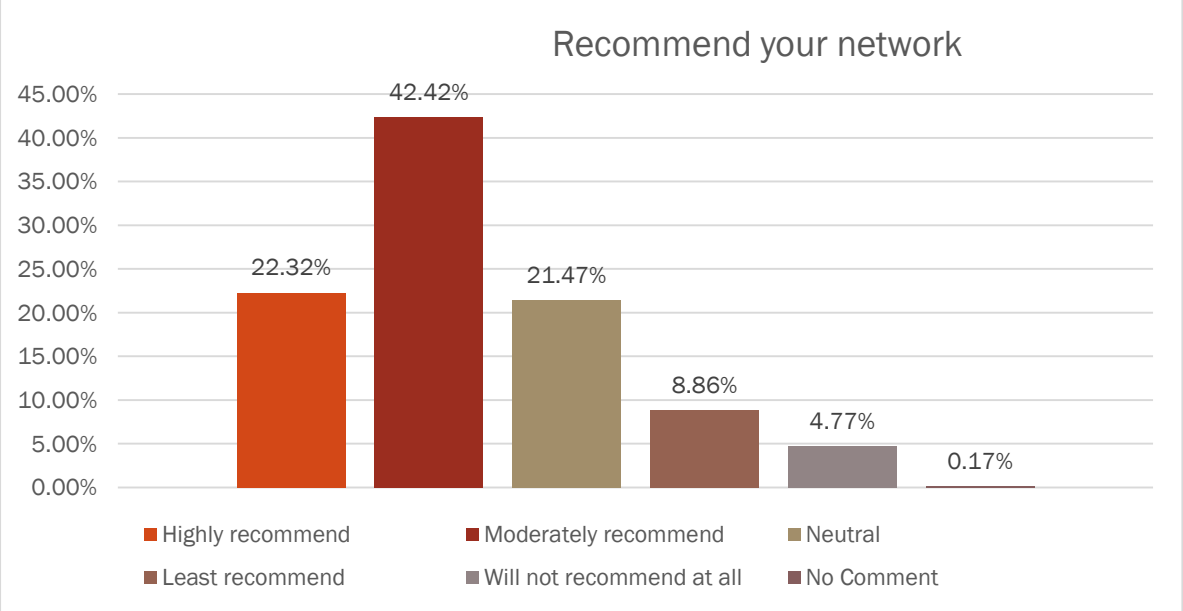
10.56% strongly disagree and 15.67% disagree and 40.03% are neutral and 27.09% agree and 6.64% are strongly agree.

Q10. How likely is it that you would recommend your current network to your friends and family?

Table No: 1.13 Table showing the distribution of Recommend your network.

Highly recommend	Moderately recommend	Neutral	Least recommend	Will not recommend at all	No Comment
131	249	126	52	28	1
22.32%	**42.42%**	**21.47%**	**8.86%**	**4.77%**	**0.17%**

Graph No. 1.13 Graph showing the distribution of Recommend your network.



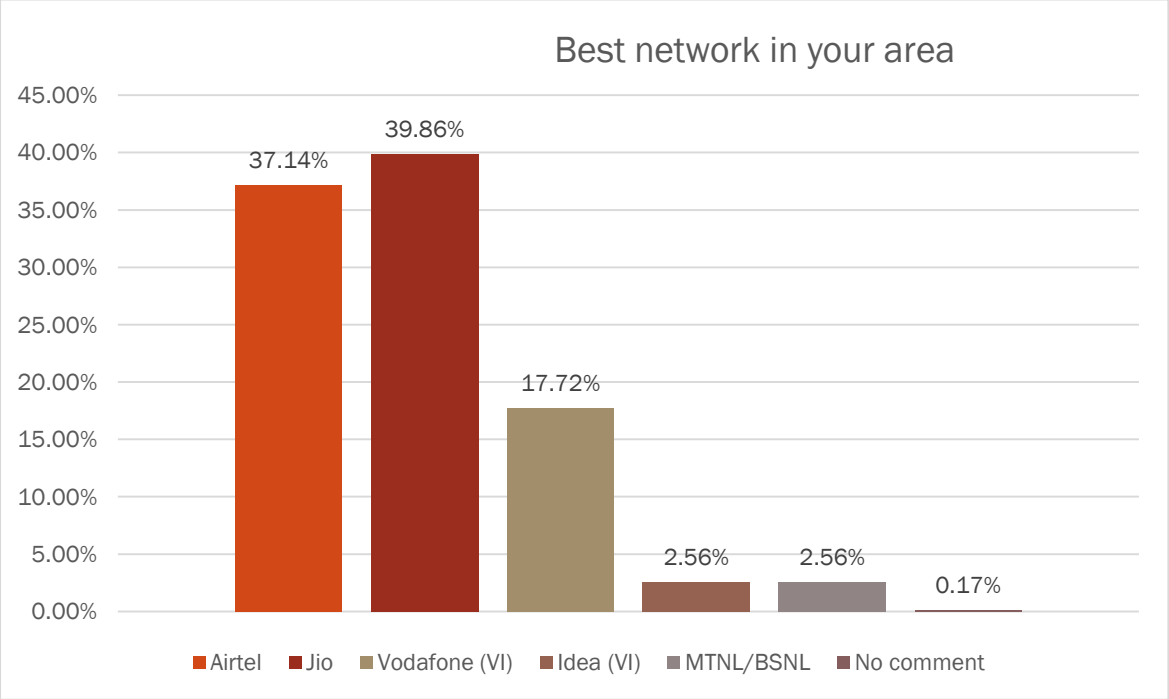
Graph show the distribution of recommend network with % and graph, Moderately recommend 42.42% service network to my friend and follower, network service is more reasonable from other network service and will not recommend at all is 4.77%, Highly recommend is of 22.32%

Q11. Which network works best in your area?

Table No: 1.14 Table showing the distribution of Best network in your area.

Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL	No comment
218	234	104	15	15	1
37.14%	39.86%	17.72%	2.56%	2.56%	0.17%

Graph No. 1.14 Graph showing the distribution of Best network in your area.



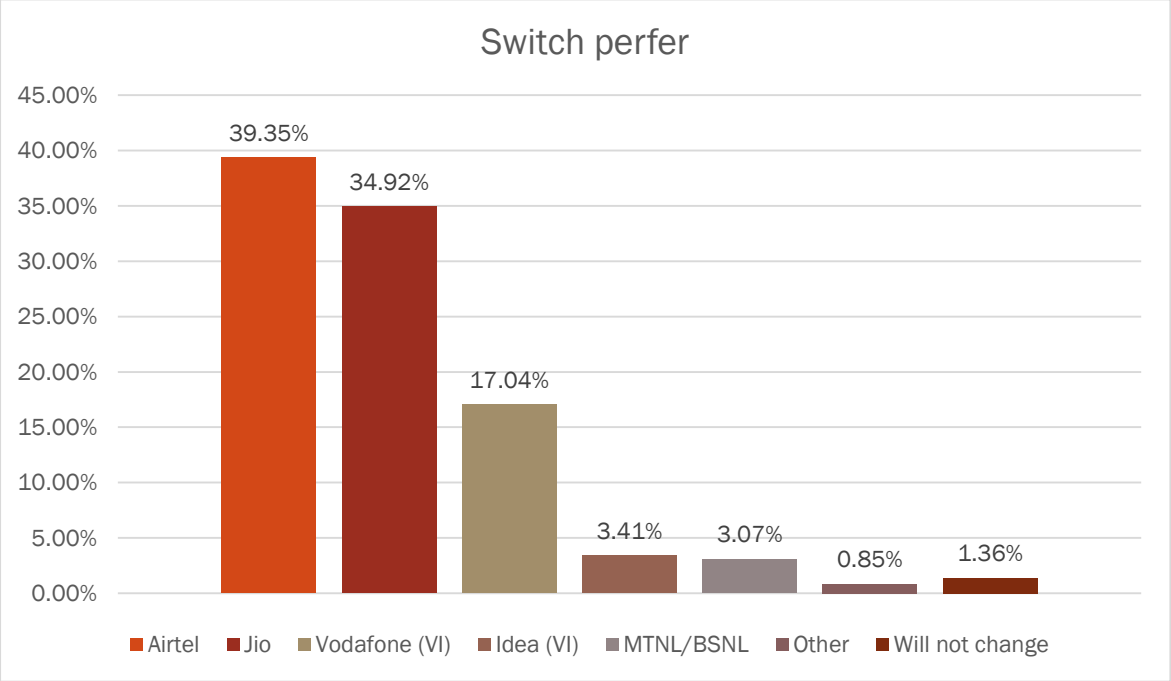
Graph show the distribution of best network in your area with % and graph, in the survey total 587 people participated. In the survey jio is the best on overall service provider. Jio is on 39.86% after the jio second one Airtel is the most of people choice this service provider. And after of this Idea is on 2.56% and MTNL/BSNL is at 2.56% and 0.17% don't want to comments

Q12. If you want to switch then which other SIM would you like to prefer from the networks listed below?

Table No: 1.15 Table showing the distribution of Switch prefer.

Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL	Other	Will not change
231	205	100	20	18	5	8
39.35%	34.92%	17.04%	3.41%	3.07%	0.85%	1.36%

Graph No. 1.15 Graph showing the distribution of Switch prefer.



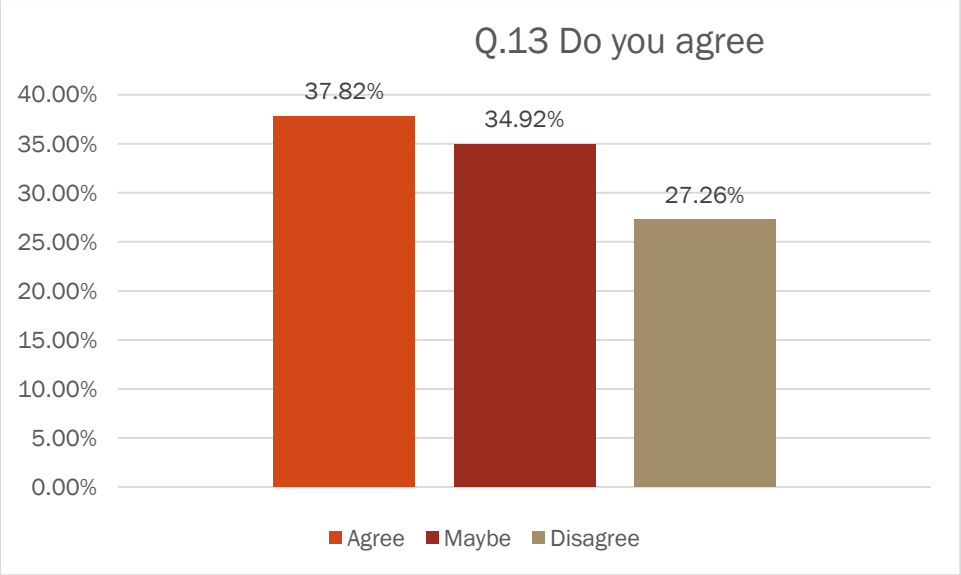
Graph show the distribution of switch prefer with % and graph, in the survey total 587 people participated. In the survey Airtel is the best on overall service provider. However 39.35% said prefer to switch with Airtel and 34.92% Percent will switch to jio, wherein Vodafone and other service below 17.04%

Q13. Although companies are saying 3G/4G network are better but customers were experiencing better network quality when only 2G has been introduced. Do you agree?

Table No: 1.16 Table showing the distribution of Q.13 Do you agree.

Agree	Maybe	Disagree	Total
222	205	160	587
37.82%	34.92%	27.26%	100%

Graph No. 1.16 Graph showing the distribution of Q.13 Do you agree.



Disagree to above mention graph primary difference between 2G and 3G networks for mobile subscribers is that they get to enjoy faster Internet browsing and data downloading on 3G. On average, the speed of data transmission on a 2G network is only 170Kbps, while in 3G networks the downloading speed can go up to 42Mbps

3G technology allows for advanced technology, multimedia services & larger network capacity, it helps a wider variety of cell phones to operate on the network, it allows a wider radio spectrum which helps in faster data transmission and the carriers can deliver 3G at a reduced cost compared to 2G

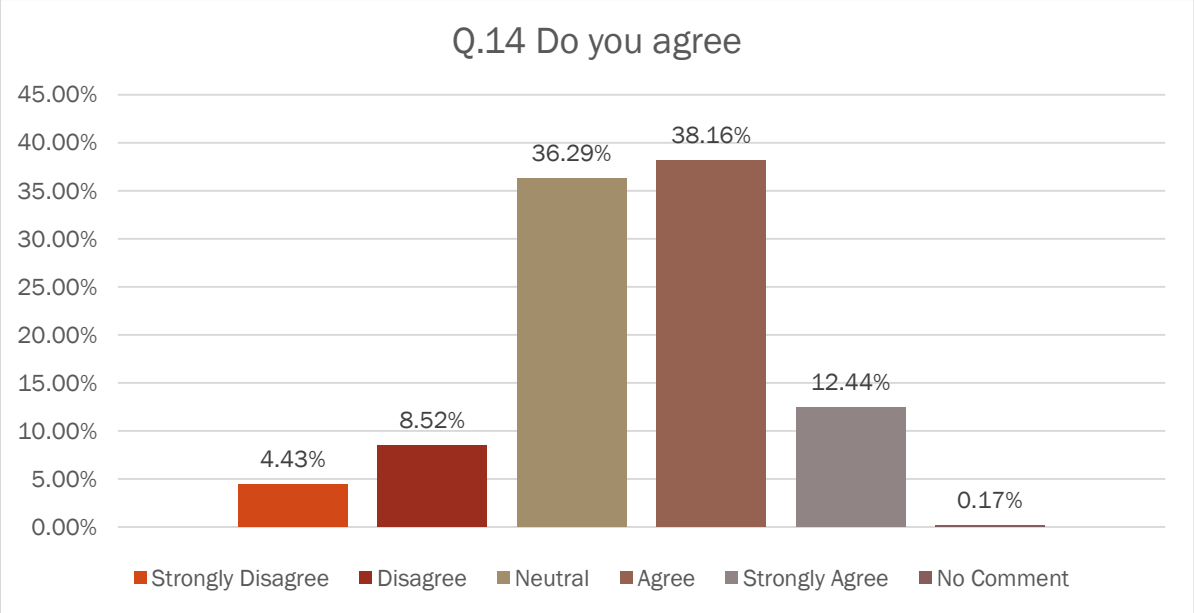
Above graph showcase 37.82 % are agree and 27.26 % are disagree and 34.92% are in maybe.

Q14. The update made by the telecom networks has improved the network quality?

Table No: 1.17 Table showing the distribution of Q.14 Do you agree.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	No Comment	Total
26	50	213	224	73	1	587
4.43%	8.52%	36.29%	38.16%	12.44%	0.17%	100%

Graph No. 1.17 Graph showing the distribution of Q.14 Do you agree.



Above chart is of Improvement wherein 38.16% is on agree and 12.44% strongly agree and there are more suggestion as 4.43% Strongly disagree and 8.52% are in Disagree.

Quality of Service (QoS) is a set of technologies that work on a network to guarantee its ability to dependably run high-priority applications and traffic under limited network capacity. However above graph showcase 38% agree with the point and 36% are in neutral.

Q15. Do you think during pandemic we were dependent on networks for communication and other uses?

Table No: 1.18 Table showing the distribution of Q.15 Do you agree.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	No Comment	Total
34	36	62	171	283	1	587
5.79%	6.13%	10.56%	29.13%	48.21%	0.17%	100%

Graph No. 1.18 Graph showing the distribution of Q.15 Do you agree.

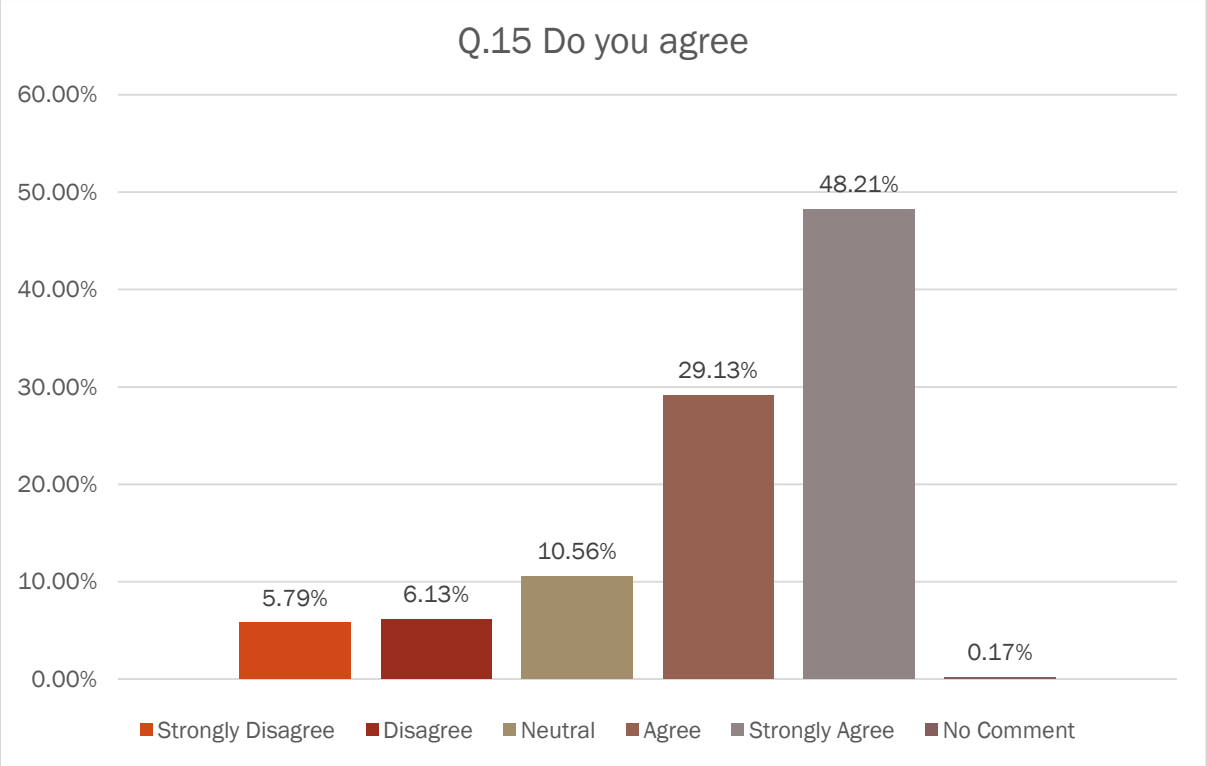


Chart Show 5.79% strongly disagree and 6.13% disagree and 10.56% Neutral and 29.13% agree and 48.21% strongly Agree.

First of all, the top five Internet activities with increased use during the pandemic referred to using Internet for job/school/studies, surfing, chatting, and using Facebook and Instagram. This is expected, since the quarantine measures implicated continuing the studies online and working from home, and since the population could have replaced leisure activities by surfing and social activities by chatting and the most widely established social networks. Playing games more during the pandemic was also absent in many participants, possibly due to the already high gaming activities and addiction potential of these activities in those prone to gaming before the pandemic.

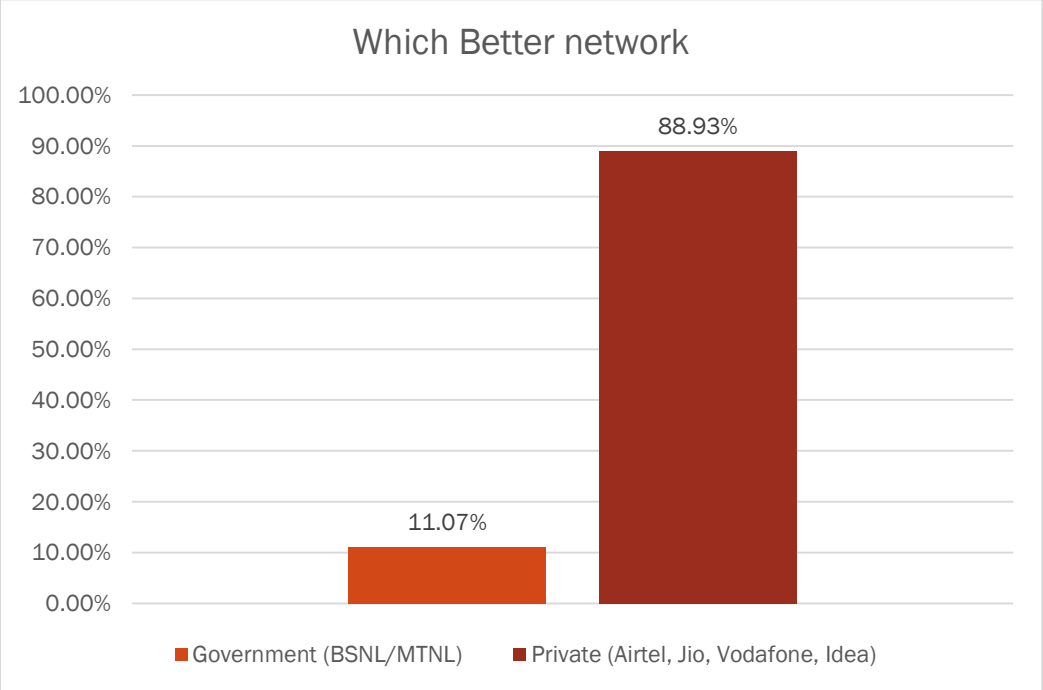
Thus above chart strongly agree & agree network usage were high during pandemic of 77.34%

Q16. Which sector provide better network Government or Private?

Table No: 1.19 Table showing the distribution of which Better network?

Government (BSNL/MTNL)	Private (Airtel, Jio, Vodafone, Idea)	Total
65	522	587
11.07%	88.93%	100%

Graph No. 1.19 Graph showing the distribution of which Better network?



The bar chart illustrates the Network Jio and other private network has now surpassed the 88.93% 4G Availability milestone and 11.07% of BSNL/MTNL. In the survey Private is the better than government network.

We all know very well about 3G/4G data services in India which came lately when equated to other aroused countries; Data Services is the most important factor on 4G network which allows ultimate data speed internet

2. REPORT ON PRIMARY DATA ANALYSIS

Mann – Whitney U test

Hypotheses

H_{0B}: There is no significant difference among the call connection service provided by government and private sector.

H_{1B}: There is significant difference among the call connection service provided by government and private sector.

H_{0C}: There is no significant difference among the internet browsing service provided by government and private sector.

H_{1C}: There is significant difference among the internet browsing service provided by government and private sector.

H_{0D}: There is no significant difference among the clarity of video service provided by government and private sector.

H_{1D}: There is significant difference among the clarity of video service provided by government and private sector.

Result

i) For call connection service, the obtained values for W is 13320 and p-value is 0.003676.
ii) For internet browsing service, the obtained values for W is 14940 and p-value is 0.1069.
iii) For clarity of video service, the obtained values for W is 16574 and p-value is 0.7554.

Conclusion

i) Since p-value is < 0.05 then reject H_{0B} and conclude that there is significant difference among the call connection service provided by government and private sector.

ii) Since p-value is > 0.05 then do not reject H_{0C} and conclude that there is no significant difference among the internet browsing service provided by government and private sector.

iii) Since p-value is > 0. 05 then do not reject H_{0D} and conclude that there is no significant difference among the clarity of video service provided by government and private sector.

RESULT

The model fitted for binary logistic regression is given below

glm(formula = service type ~ clarity of video + internet calling + lag in audio/video + late message_delivery + ploitness_of_staff + offer_SV)

TABLE 2.1 Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.2927	-0.4996	-0.3955	-0.2931	2.6387

In the Table 10, the Min, Median and Max values obtained are -1.2927, -0.3955 and 2.6387 respectively whereas first and third quartile obtained as -0.4996 and -0.2931 respectively.

TABLE 2.2

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	-2.9383	0.6584	-4.463	8.1e-06 ***
Clarity of video	0.6879	0.2619	2.627	0.00861 **
Internet calling	-0.4932	0.2487	-1.983	0.04738 *
Lag in audio/video	-0.4640	0.2585	-1.795	0.07262 .
Late message delivered	0.4910	0.2222	2.210	0.02713 *
Politeness of staff	0.3230	0.1360	2.376	0.01752 *
Offer SV	-0.3926	0.1535	-2.558	0.01054 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

From table 10 it can be seen that clarity of video, late message delivered and politeness of staff influences service type positively with values 0.6879, 0.4910 and 0.3230 respectively whereas internet calling, lag in audio/video and offers provided by service sector has a negative effect on service sector with value -0.4932, -0.4640 and -0.3926 respectively.

For significance of the variables are clarity of video, Internet calling, lag in audio/video, late message delivered, politeness of staff and offers provided by service sector is significant with p-value are 0.00861, 0.04738, 0.02713, 0.01752 and 0.01054 respectively while lag in

audio/video do not show significance but can be taken into consideration since they are slightly > 0.05 taking values 0.07262

The **null deviance** obtained here is 265.35 with 384 degrees of freedom, it basically shows how well the response variable is predicted by a model that includes only the intercept (grand mean) while the **residual deviance** obtained is 241.75 with 378 degrees of freedom and the **AIC** turned out to be 255.75 after performing Step AIC. The **Fisher scoring iteration** required to converge to this model is 5.

To check the goodness of model

H_0: The model is not better than chance (probability) at predicting the outcome.

H_1: The model is better than chance (probability) at predicting the outcome.

Since, the **Chi-square probability** obtained is 0.00062 by considering the difference between values and degree of freedom of null deviance and residual deviance obtained from the fitted model hence, it can be seen that model is good for predicting the desired result.

Since most of the probabilities for the predicted model are in the favour of prepaid service type which indicates that most of the customers experience better service under prepaid service rather than post-paid service.

The accuracy achieved by predicted probability for train data is **89.09%** and for test data is **88.96%**.

Conclusion:

The obtained predictive model is as follows:

Service type = - 2.9383 + (0.6879*clarity of video) – (0.4932*internet calling) – (0.4640*lag in audio/video) + (0.4910*lag message delivered) + (0.3230*politeness of staff) - (0.3926 *offer provided by service provider)

The model can be used to likely predict whether customer will opt for prepaid or post-paid

Mann-Whitney U test

Hypotheses

H_0: There is no significant difference between customer care services provided by different service provider for prepaid and post-paid customers/users

H_1: There is significant difference between customer care services provided by different service provider for prepaid and post-paid customers/users

Result

i) For clarity of voice service, the obtained value for W is 12598 and p-value is 0.2143.

ii) For internet calling service, the obtained value for W is 13597 and p-value is 0.7467.

iii) For lag in audio/video problem faced, the obtained value for W is 14080 and p-value is 0.8999.

iv) For late message delivery problem faced, the obtained value for W is 12091 and p-value is 0.07722.

v) For politeness of staff service, the obtained value for W is 13066 and p-value is 0.4231.

vi) For offers provided by service provider, the obtained value for W is 15017 and p-value is 0.3324.

Conclusion

Since p-value is > 0.05 then do not reject H_0 and conclude that there is no significant difference between customer care services provided by different service provider for prepaid and post-paid customers/users

Kruskal-Wallis Test

The Kruskal–Wallis test by ranks, Kruskal–Wallis *H* test (named after William Kruskal and W. Allen Wallis), or one-way ANOVA on rank is non-parametric method for testing whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes. It extends the Mann-Whitney U test, which is used for comparing only two groups. The parametric equivalent of the Kruskal–Wallis test is the one-way analysis of variance (ANOVA).

A significant Kruskal–Wallis test indicates that at least one sample stochastically dominates one other sample. The test does not identify where this stochastic dominance occurs or for how many pairs of groups stochastic dominance obtains. For analysing the specific sample pairs for stochastic dominance, Dunn's test, pairwise Mann-Whitney tests

with Bonferroni correction, or the more powerful but less well-known Conover–Iman test are sometimes used.

(service meets my expectation)

Hypothesis:

Ho: There is significant difference between the average score of meeting expectation and the two types of services.

H1: There is no significant difference between the average score of meeting expectation and the two types of services.

TABLE 2.3

Kruskal-Wallis chi-squared	0.88855
Degrees of Freedom	1
p-value	0.3459

Conclusion:

Since, p-value > alpha=0.05, we fail to reject the null hypothesis.

Hence, there is no difference between the average meet expectation score and the two types of services.

(value for the money)

Hypothesis:

Ho: There is no significant difference between the average score of value for money and the two types of services.

H1: There is significant difference between the average score of value for money and the two types of services.

Kruskal-Wallis chi-squared = 1.1939, df = 1, p-value = 0.2745

TABLE 2.4

Kruskal-Wallis chi-squared	1.1939

Degrees of Freedom	1
p-value	0.2745

Conclusion:

Since, p-value > alpha=0.05, we fail to reject the null hypothesis.

Hence, there is no difference between the average value for money score and the two types of services.

For "prepaid-post-paid" and "govt-prvt" using Mann Whitney U test Q3.

In statistics, Mann–Whitney U , Wilcoxon rank-sum test, or Wilcoxon–Mann–Whitney test) is a non-parametric test of the null hypothesis that, for randomly selected values *X* and *Y* from two populations, the probability of *X* being greater than *Y* is equal to the probability of *Y* being greater than *X*.

Ho: There is no significant difference between prepaid and post-paid services for government and private companies.

H1: There is significant difference between prepaid and post-paid services for government and private companies.

TABLE 2.5

W	258210
p-value	2.2e-16

alternative hypothesis: true location shift is not equal to 0

Conclusion

Since, p-value > alpha=0.05, we fail to reject the null hypothesis.

Hence, there is significant difference between prepaid and post-paid services for government and private companies.

STRUCTURAL EQUATION MODEL(SEM)

HYPOTHESIS:

H0A: there is no significant impact of satisfaction w.r.t (vm, ccs, prb, ser) on recommendation of the service provider to family and friends

H1A: there is significant impact of satisfaction w.r.t (vm, ccs, prb, ser) on recommendation of the service provider to family and friends

RESULT:

SEM in R was performed to study the telecom network's customer's preferences whether they will recommend their current using network with their family and friends. FOUR latent variables namely "services ratings", "customer care services", "problem faced by customers" and "cost and offers" were considered. On the other hand, "Network recommendation" was taken as observed variable.

"SERVICES RATINGS" being latent variable has eight indicators namely "call connection", "clarity of voice", "SMS/MMS", "internet browsing", "video streaming", "uploading and downloading data/files", "clarity of video" and "internet audio/video calling".

"CUSTOMER CARE SERVICES" being latent variable has seven indicators namely "chatbot", "time taken to fix issue", "availability of customer care no.", "response to complaints", "24 hours customer service", "politeness of staff" and "offers provided by service provider".

"PROBLEM FACED BY CUSTOMERS" being latent variable has four indicators namely "call drop", "lag in audio/video", "late message delivery" and "cross connection".

"COST AND OFFERS" being latent variable have two indicators namely "service meets my expectation" and "value for money".

"NETWORK RECOMMENDATION" being observed variable has five pointer Likert scale in the sequence "highly recommend", "moderately recommend", "neutral", "least recommend" and "will not recommend at all".

The overall result showed that the eight indicators have appreciably positive effect on "SERVICES RATINGS" and with significant P values shown in the Table 1 below-

TANLE 2.6 showing the observed variable and loading on laten variable customer care services (ccs)

TABLE 2.6

Indicator	Estimates (β)	P(>|z|)
call connection	1.000	0.000
clarity of voice	1.039	0.000
SMS/MMS	0.987	0.000
internet browsing	1.123	0.000
video streaming	1.123	0.000
uploading and downloading data/files	1.133	0.000
clarity of video	1.053	0.000
internet audio/video calling	1.028	0.000

The seven indicators have appreciably positive effect on "CUSTOMER CARE SERVICES" and with significant P values shown in the Table 2.7 below-

TABLE 2.7

Indicator	Estimates (β)	P(>|z|)
Chatbot	1.000	0.000
time taken to fix issue	1.324	0.000
availability of customer care no.	1.471	0.000
response to complaints	1.472	0.000
24 hours customer service	1.461	0.000
politeness of staff	1.523	0.000
offers provided by service provider	1.313	0.000

The four indicators have appreciably positive effect on "PROBLEM FACED BY CUSTOMERS" and with significant P values shown in the Table 2.8 below-

TABLE 2.8

Indicator	Estimates (β)	P(>|z|)

call drop	1.000	0.000
lag in audio/video	0.963	0.000
late message delivery	0.851	0.000
cross connection	0.740	0.000

The two indicators have appreciably positive effect on "COST AND OFFERS(VALUE OF MONEY)" and with significant values shown in the table 2.9 below-(VM)

TABLE 2.9

Indicator	Estimates (β)	P(>|z|)
service meets my expectation	1.000	0.000
value for money	0.854	0.000

Interrelated covariances among the latent variables along with significant P values is given in the Table 2.10 below-

TABLE 2.10

Latent Variables	Covariances	P(>|z|)
Services ratings (on)		
Problem faced by customers	0.142	0.000
Customer care services	0.281	0.000
Cost and offers	0.403	0.000
Problem faced by customers (on)		
Customer care services	0.118	0.000
Cost and offers	0.132	0.000
Customer care services (on)		
Cost and offers (value of money)	0.206	0.000

The measure fit index values of our fitted structural equation model are also obtained and all the values of measures are within the required ranges hence accepted. The values of all the measures are shown in the Table 6 below-

Model fit indices for measurement model

TABLE 2.11

Goodness of Fit Criteria	Fit Indices	Good Fit	Acceptable Fit	Goodness of Fit Values Obtained	Fit Situations
Statistic of χ^2 Test	χ^2	$0.00 \leq \chi^2 \leq 2$	$2 \leq \chi^2 \leq 5$	0.00	Accepted
Root Mean Square Error of Approximation	RMSEA	$0.00 \leq$ RMSEA ≤ 0.05	$0.05 \leq$ RMSEA ≤ 0.10	0.070	Accepted
Standardized Root Mean Square Residual	SRMR	$0.00 \leq$ SRMR ≤ 0.05	$0.05 \leq$ SRMR ≤ 0.10	0.037	Accepted
Comparative Fit Index	CFI	0.97< CFI <=1	$0.9 <$ CFI ≤ 1	.941	Accepted
Tucker-Lewis Index	TLI	$0.95 \leq$ TLI ≤ 1	$0.9 <$ TLI ≤ 1	0.931	Accepted

A regression has been estimated along with the significant P-values shown in the Table 2.12 given below-

TABLE 2.12

Variables	Regression coefficients	P(>|z|)
Network recommendation (on)		
Services ratings	0.119	0.011
Problem faced by customers	0.333	0.000
Customer care services	0.084	0.075
Cost and offers(value of money)	0.323	0.000

CONCLUSION:

Since all the measurement Indices lies within the required range-

The test statistic of overall model fit is (9986.893) with 231 degrees of freedom and returning P-value as (<0.01) indicates the overall model to be of good fit.

RMSEA (0.070) shows it a good fit model which lies within the acceptable fit range so the model could be considered to be good fit according to descriptive measure of fit.

TLI (0.931) and CFI (0.941) compare the absolute fit of the specified model to the absolute fit of the Independence model and both the values line within the range hence accepted.

PATH DIAGRAM

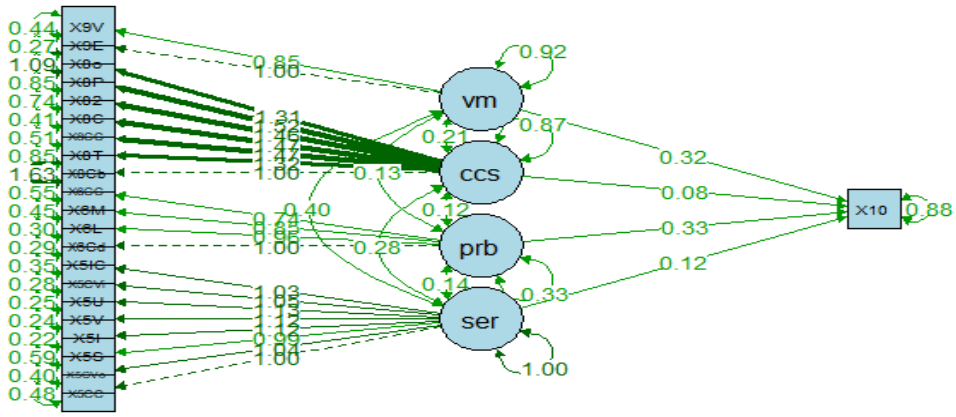


TABLE 2.13

Vm	Ccs	Prb	Ser
Value of money	Customer care services	Problem faced by customer	Service rating

In the figure shown above, the values associated with each path on the right hand side of latent variables are regression coefficients, these values represent the amount of change in “Network recommendation” given as a standard deviation unit change in “latent variables”

The sequence of highest to lowest regression coefficients among latent and independent variables are as-

“Network recommendation” and “Problem faced by customers” with 0.33

“Network recommendation” and “Cost and offers” with 0.32

"Network recommendation" and "Services ratings" with 0.12

"Network recommendation" and "Customer care services" with 0.08

i) The scale taken for "Problem faced by customers" is as- 1(Agree), 2(Neutral) and 3(Disagree)

The coefficient score associated with "Problem faced by customers" and "Network recommendation" is 0.33 i.e. positive and significant with a P-value (<0.05) indicates that the person is not facing problem and he/she is likely to recommend their current using network to their family and friends

ii) The scale taken for "Cost and offers" is as- 1(Strongly Disagree), 2(Disagree), 3(Neutral), 4(Agree) and 5(Strongly Agree)

The coefficient score associated with "Problem faced by customers" and "Network recommendation" is 0.32 that is positive and significant with a P-value (<0.01) indicates that the person is not facing problem in Cost and offers provided by the network and he/she is likely to recommend their current using network to their family and friends.

iii) The scale taken for "Services ratings" is as- 1(Strongly Disagree), 2(Disagree), 3(Neutral), 4(Agree) and 5(Strongly Agree)

The coefficient score associated with "Services ratings" and "Network recommendation" is 0.12 that is positive and significant with a P-value (<0.05) indicates that the person is not facing problem in Services provided by the network and he/she is likely to recommend their current using network to their family and friends.

iv) The scale taken for "Customer care services" is as- 0(Never Used), 1(Strongly Disagree), 2(Disagree), 3(Neutral), 4(Agree) and 5(Strongly Agree)

The coefficient score associated with "Services ratings" and "Network recommendation" is 0.08 that is positive and significant with a P-value (<0.01) indicates that the person is not facing problem with Customer care services provided by the network and he/she is likely to recommend their current using network to their family and friends.

Since all Variables tested using SEM are under the favour of null hypothesis, hence the customer will recommend his/her currently using network to their family and friends

K-means Cluster Analysis:

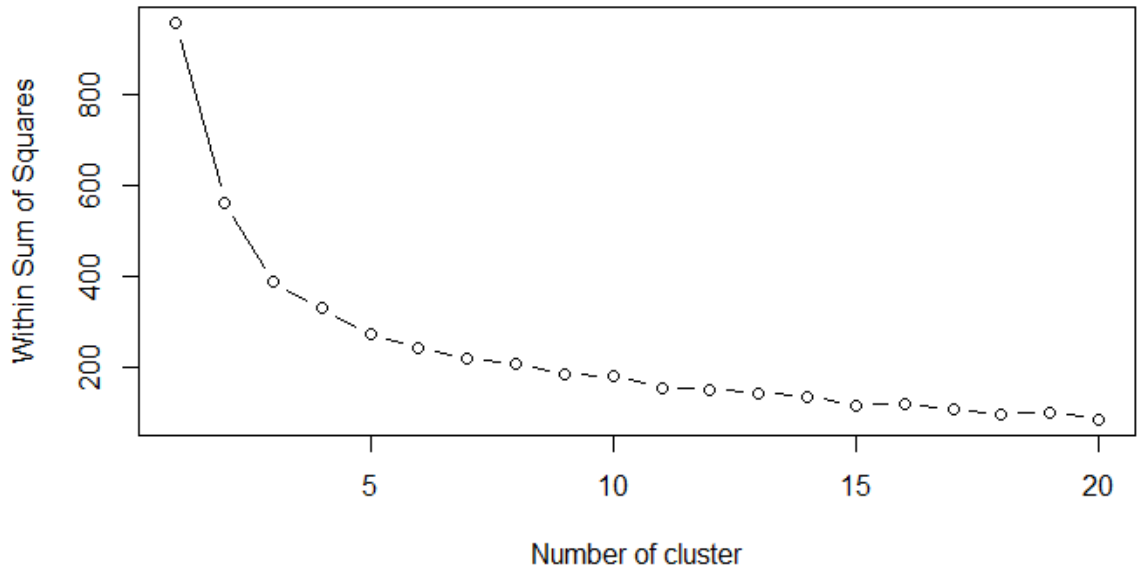
Cluster analysis is carried to classify variable on the basis of similarity and difference using cluster analysis. Different classification variable like govt sector, private sector and service providers are consisder for classification on the basis of services provided by them.(in the beginning)

Highlight % shows good classification

1.For Government sector

Elbow plot for choosing no. of clusters.

Plot no. 2.1:



Here the graph shows that majorly till 5th cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 5.

Showing size and clusters.

TABLE 2.14:

Cluster	1	2	3	4	5
Size	17	17	10	14	7

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 5) and columns are variables.

Cluster means.

TABLE 2.15:

Clust no.	Govt/prvt	Chatbot	Time taken to fix issue	Availability of customer care no.	Response to the complaints	24 hours customer service	Politeness of Staff	Offers provided by the service provider
1	0	0.2941176	1.4117647	1.7058824	1.1176471	1.1176471	2.5294118	2.0588235
2	0	2.4117647	1.9411765	1.9411765	2.2941176	2.6470588	2.2941176	2.2941176
3	0	0.0000000	2.5000000	2.9000000	2.8000000	3.2000000	3.7000000	2.5000000
4	0	0.5714286	0.5714286	0.3571429	0.5714286	0.6428571	0.3571429	0.2857143
5	0	4.2857143	4.2857143	4.5714286	4.4285714	4.5714286	4.7142857	4.4285714

Table below shows grouping a set of objects in the clusters 1, 2, 3, 4, 5.

TABLE 2.16

1	80 84 86 103 114 205 254 275 276 283 289 475 494 508 509 531 569
2	6 10 33 36 146 296 384 434 459 474 490 505 510 535 550 582 586
3	77 130 295 436 471 473 478 493 576 581
4	125 155 282 377 387 460 469 472 479 484 552 554 574 580
5	230 304 373 405 454 506 528

Average satisfaction for every hypothesis

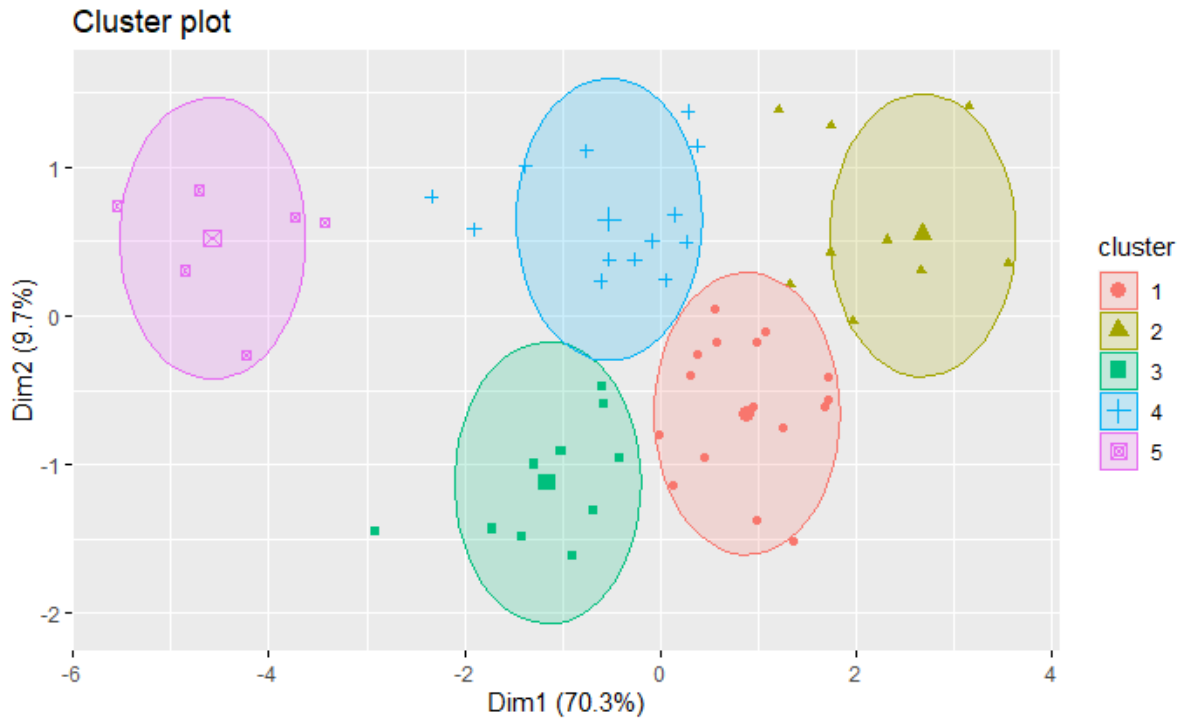
The table below shows the within sum of squares.

TABLE: 2.17

Within cluster sum of squares	1	2	3	4	5
Values	81.88235	68.47059	59.20000	44.78571	17.14286

The plot below shows clustering

Plot 2.2



Interpretation:

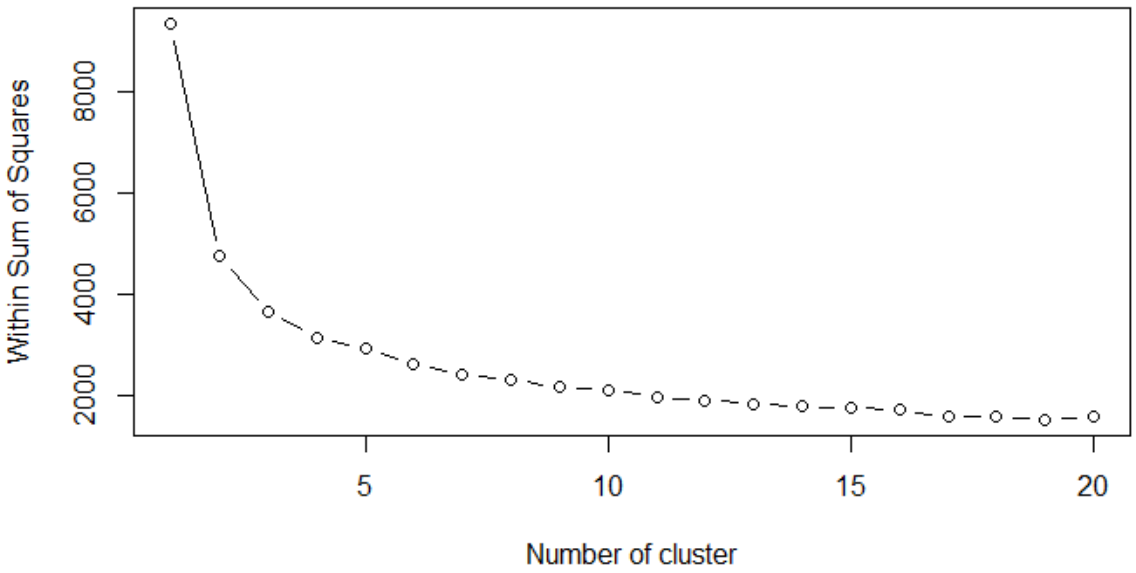
The calculated value for ratio (between_SS / total_SS = 71.7 %) which is a good amount to consider clusters. The cluster means *table 4.1.2* shows that

- Cluster 1 can be formed by considering services like ‘politeness of staff’ and ‘offers provided by service provider’ as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like ‘chatbot’ and ’24 hrs customer service’ as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like ‘time taken to fix the issue’, ‘24 hours customer service’ and ‘politeness of staff’ as they have the highest cluster mean value in cluster 3.

- Cluster 4 can be formed by considering services like ‘chatbot’, ‘time taken to fix the issue’, ‘24 hours customer service’ and ‘politeness of staff’ as they have the highest cluster mean value in cluster 4.
- Cluster 5 can be formed by considering services like ‘Availability of customer care no.’, ‘24 hours customer service’ and ‘politeness of staff’ as they have the highest cluster mean value in cluster 5.

2.For Private Networks

Plot no. 2.3: Elbow plot for choosing no. of clusters:



Here the graph shows that majorly till 4th cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 4.

Table showing size and clusters

TABLE 2.18

Cluster	1	2	3	4
Size	171	86	126	139

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 4) and columns are variables

TABLE 2.19: Cluster means

Clust no.	Govt/prvt	Chatbot	Time taken to fix issue	Availability of customer care no.	Response to the complaints	24 hours customer service	Politeness of Staff	Offers provided by the service provider
1	1	0.7543860	1.5204678	1.8538012	1.7543860	1.6842105	2.2339181	1.9415205
2	1	0.1162791	2.4651163	3.3953488	3.2441860	3.4767442	3.8720930	3.6511628
3	1	3.5714286	3.4682540	3.7063492	3.5714286	3.6111111	3.8253968	3.5000000
4	1	0.2158273	0.2086331	0.3669065	0.2589928	0.4388489	0.2661871	0.5539568

Table below shows grouping a set of objects in the clusters 1, 2, 3, 4.

TABLE 2.20

1	1 4 5 16 21 23 27 34 39 41 47 49 90 93 95 96 98 100 101 104 107 111 113 121 128 133 135 138 139 140 141 142 144 145 149 151 152 159 162 166 167 173 177 179 180 182 183 185 186 187 189…
2	24 28 45 50 51 109 110 116 118 129 148 150 154 156 157 160 165…
3	2 3 7 8 9 11 12 13 14 15 17 18 19 20 25 37 38 42 43 91 99 105 106 108 117 119 120 122 131 132 134 137 158 161 164 176 190…
4	22 26 29 30 31 32 35 40 44 46 48 52 89 92 97 102 112 115 123 124 126 127 136 143 147 153 168 169 170 171 172 174 175 178 184 188…

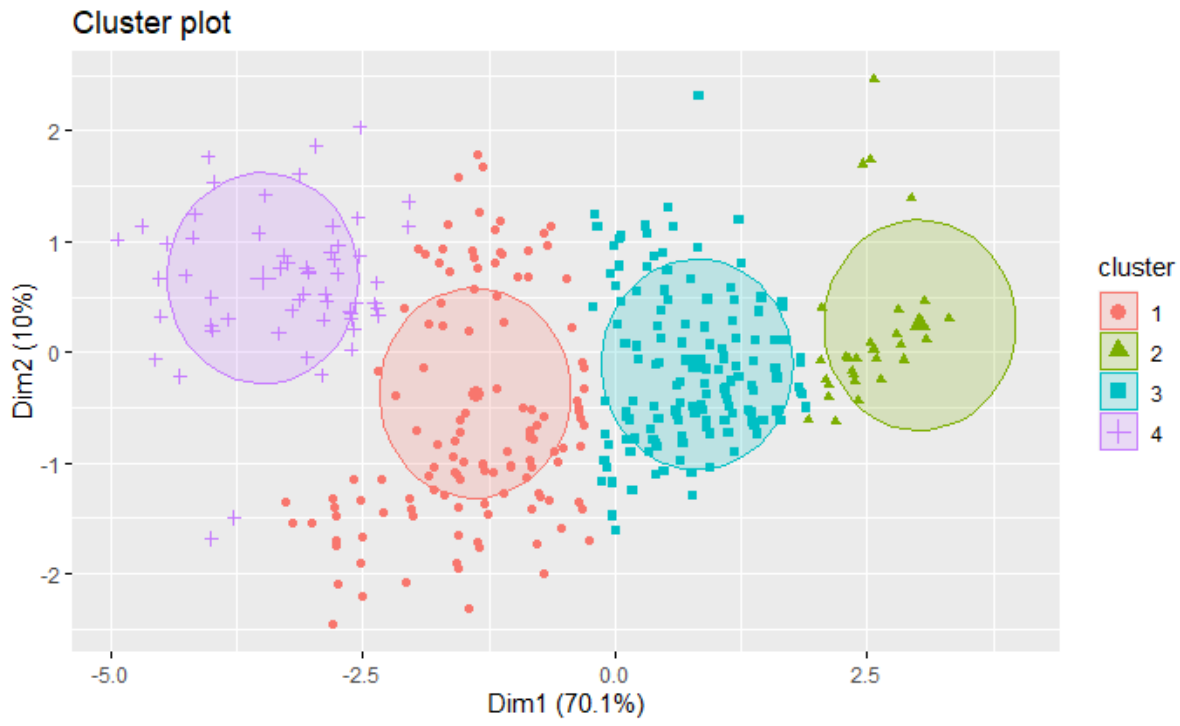
The table below shows the within sum of squares.

TABLE: 2.21

Within cluster sum of squares	1	2	3	4
Values	1118.3977	733.2442	820.8254	439.1655

The plot below shows clustering

Plot 2.4



Interpretation:

The calculated value for ratio (between_SS / total_SS = 66.7 %) which is a good amount to consider clusters. The cluster means table 4.2.2 shows that

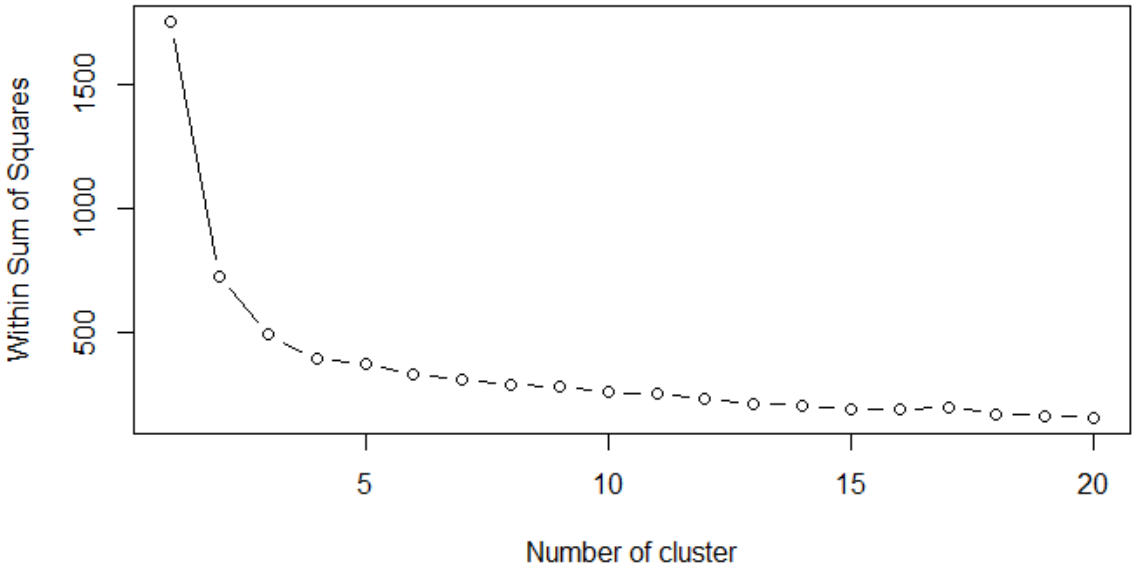
- Cluster 1 can be formed by considering services like ‘politeness of staff’ and ‘offers provided by service provider’ as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like ’24 hrs customer service’, ’politeness of staff’ and ‘offers provided by the service provider’ as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like ‘Availability of customer care no.’, ‘24 hours customer service’, ‘politeness of staff’ and ‘offers provided by the service provider’ as they have the highest cluster mean value in cluster 3.
- Cluster 4 can be formed by considering services like ‘24 hours customer service’ and ‘offers provided by the service provider’ as they have the highest cluster mean value in cluster 4.

#For Networks(X1) against Service ratings(X5)

Here comparison has been conducted for the for the services provided by different service providers like “Airtel”, “Jio”, “Idea”, “Vodafone”, and “MTNL/BSNL”

3. For Airtel

Plot no. 2.5: Elbow plot for choosing no. of clusters:



Here the graph shows that majorly till 4th cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 4.

Table showing size and clusters

TABLE 2.22

Cluster	1	2	3	4
Size	48	27	37	33

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 4) and columns are variables

TABLE 2.23: Cluster means

Clust no.	Service provider	Call conn.	Clarity of voice	SMS/MMS	Internet browsing	Video streaming	Uploading/ downloading	Clarity of voice	Internet calling
1	1	4.020833	4.104167	4.229167	3.916667	3.895833	3.812500	3.750000	3.666667
2	1	1.814815	1.666667	1.814815	1.666667	1.666667	1.481481	1.629630	1.518519
3	1	3.162162	2.918919	3.162162	3.054054	2.891892	2.756757	2.837838	2.810811
4	1	4.787879	4.848485	4.666667	4.909091	4.878788	4.848485	4.696970	4.818182

Table below shows grouping a set of objects in the clusters 1, 2, 3, 4.

TABLE 2.24

1	7 11 13 15 26 56 61 73 80 94 100 119 132 140 161 172 193 239 241 242…565
2	16 31 32 60 66 71 83 92 125 167 182 199 202 208 209 255…555
3	21 47 63 85 121 96 121 160 174 179 186 201 218 225 240…585
4	3 12 28 54 55 117 119 122 126 131 148 192 210 228 232…563

The table below shows the within sum of squares.

TABLE: 2.25

Within cluster sum of squares	1	2	3	4
Values	155.06250	63.92593	123.78378	53.45455

The plot below shows clustering

Plot 2.6

Interpretation:

(between_SS / total_SS = 77.3 %)

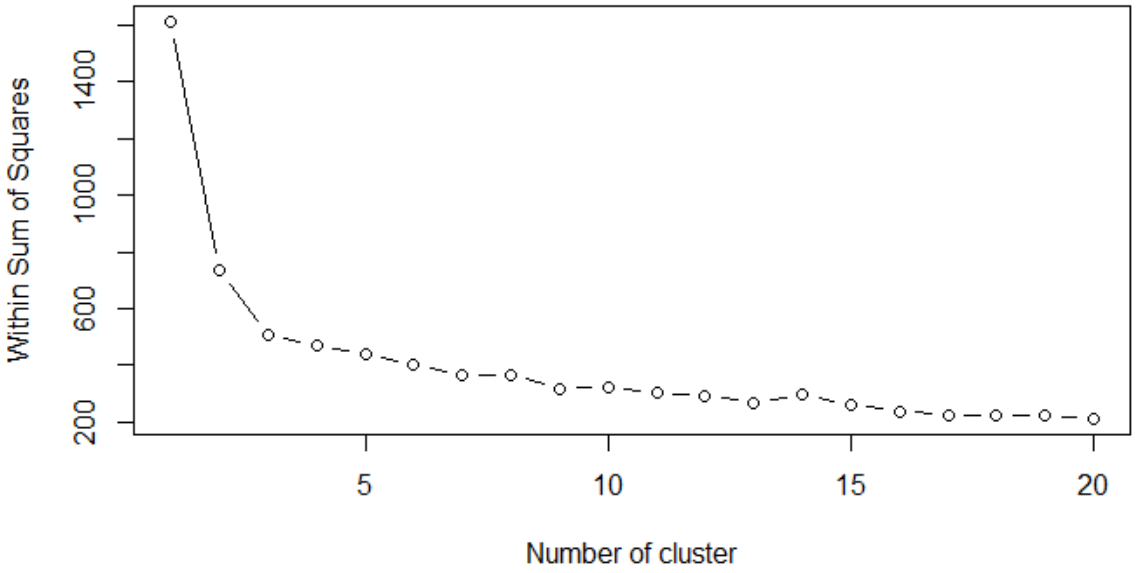
The 77.3 % is the measure of the total variance in the data set that is explained by the clustering.

The calculated value for ratio (between_SS / total_SS = 77.3 %) which is a good amount to consider clusters. The cluster means *table 4.3.2* shows that

- Cluster 1 can be formed by considering services like ‘call connection’, ‘clarity of voice’ and ‘SMS.MMS’ as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like ‘call connection’, ‘SMS/MMS’ as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like ‘call connection’, ‘Internet browsing’ and ‘video streaming’ as they have the highest cluster mean value in cluster 3.
- Cluster 4 can be formed by considering services like ‘Internet browsing’, ‘Video streaming’ as they have the highest cluster mean value in cluster 4.

4. For Jio

Plot no. 2.7**:** Elbow plot for choosing no. of clusters



Here the graph shows that majorly till 3rd cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 3.

Table showing size and clusters

TABLE 2.26

Cluster	1	2	3
Size	43	47	50

Clust no.	Service prov.	Call conn.	Clarity of voice	SMS/ MMS	Internet browsing	Video streaming	Uploading / downloading	Clarity of voice	Internet calling
1	2	4.325581	4.348837	4.44186	4.232558	4.395349	4.209302	4.046512	4.046512

2	2	1.872340	1.744681	2.12766	1.680851	1.723404	1.553191	1.851064	1.744681
3	2	3.180000	3.080000	3.10000	2.820000	2.900000	2.840000	2.820000	2.740000

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 3) and columns are variable

Cluster means

Table below shows grouping a set of objects in the clusters 1, 2, 3.

TABLE 2.27

1	1 8 1 35 43 48 87 95 112 120 153 171…572
2	6 18 20 23 30 33 39 50 91 98 107 124 137 138 141 142 175…560
3	14 37 51 52 53 65 75 78 97 99 110 113 115 118 144 146 155 159 166…576

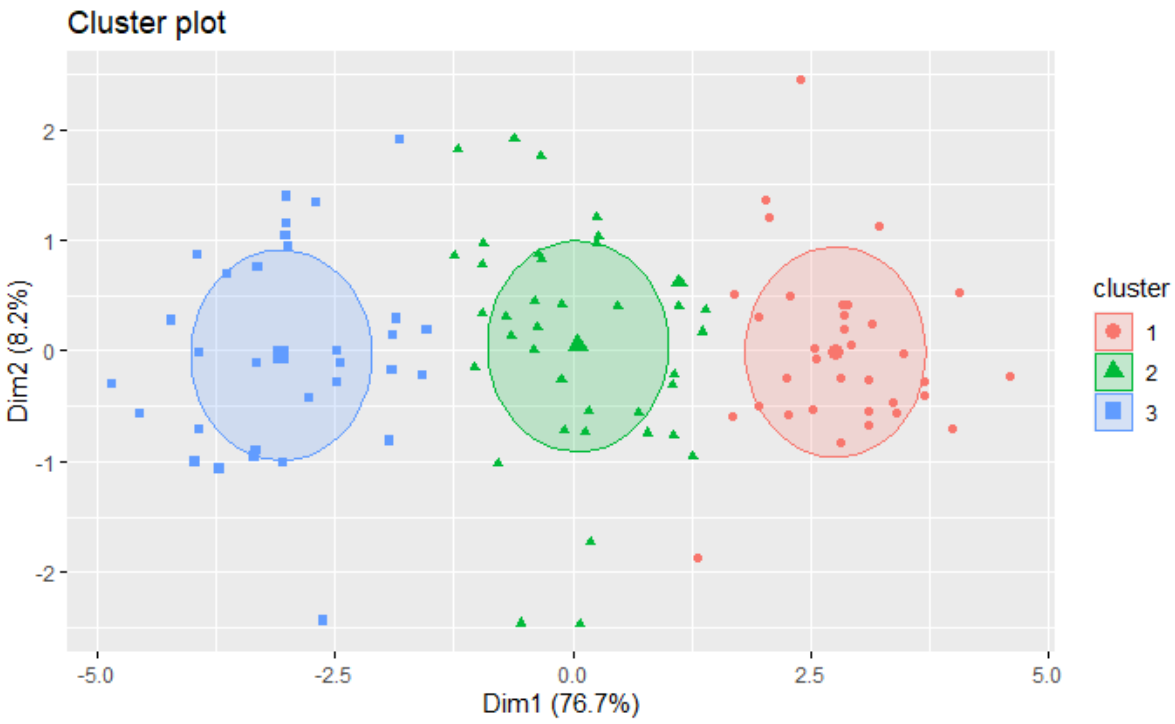
The table below shows the within sum of squares.

TABLE: 2.28

Within cluster sum of squares	1	2	3
Values	164.6977	135.5319	207.1600

The plot below shows clustering

Plot 2.7



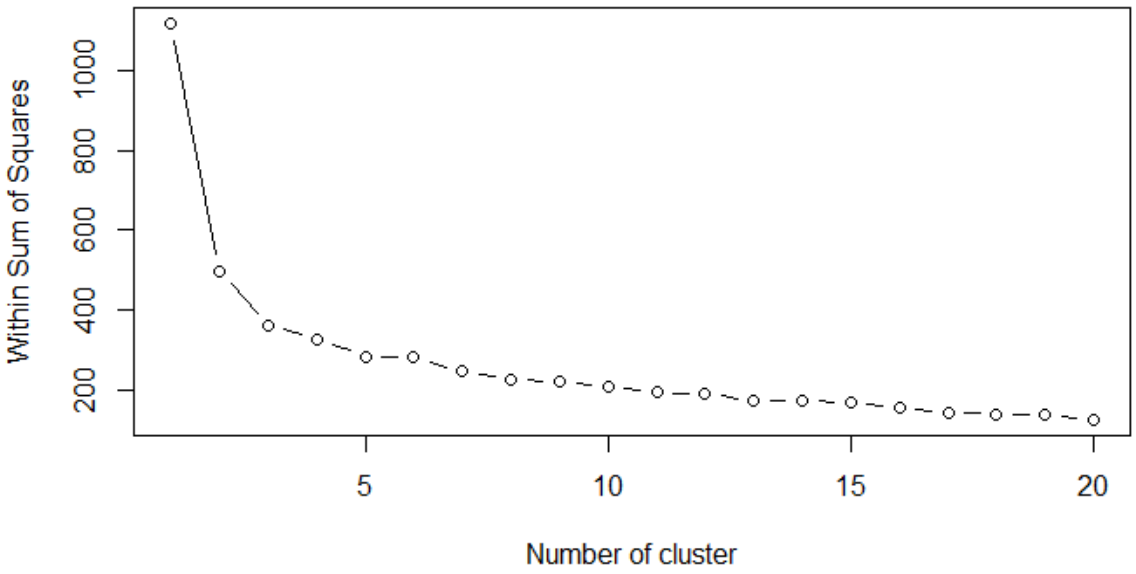
Interpretation:

The calculated value for ratio (between_SS / total_SS = 68.5 %) which is a good amount to consider clusters. The cluster means *table 4.4.2* shows that

- Cluster 1 can be formed by considering services like ‘call connection‘, ‘clarity of voice’ and ‘Video streaming’ as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like ’Call connection’, ‘SMS/MMS’ and ‘Clarity of voice’ as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like ‘Call connection’ and ‘Clarity of voice’ as they have the highest cluster mean value in cluster 3.

5. For Vodafone (VI)

Plot no. 2.8: Elbow plot for choosing no. of clusters



Here the graph shows that majorly till 3rd cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 3.

Table showing size and clusters

TABLE 2.29

Cluster	1	2	3
Size	44	32	26

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 3) and columns are variables

TABLE 2.30: Cluster means

Clust no.	Ser	Call conn.	Clarity of voice	SMS/ MMS	Internet browsing	Video Streaming	Uploading/ downloading	Clarity of voice	Internet calling
1	3	4.159091	4.318182	4.363636	4.431818	4.363636	4.181818	4.363636	4.136364
2	3	3.531250	3.531250	3.656250	3.093750	3.093750	3.187500	0.000000	3.312500

3	3	1.846154	2.192308	2.076923	1.923077	1.846154	1.692308	1.846154	1.884615

Table below shows grouping a set of objects in the clusters 1, 2, 3.

TABLE 2.31

1	9 19 25 34 45 46 64 76 88 103 106 130 136 156 158 168 176 188…573
2	17 24 40 59 69 86 93 …577
3	58 62 70 81 104 123 135…580

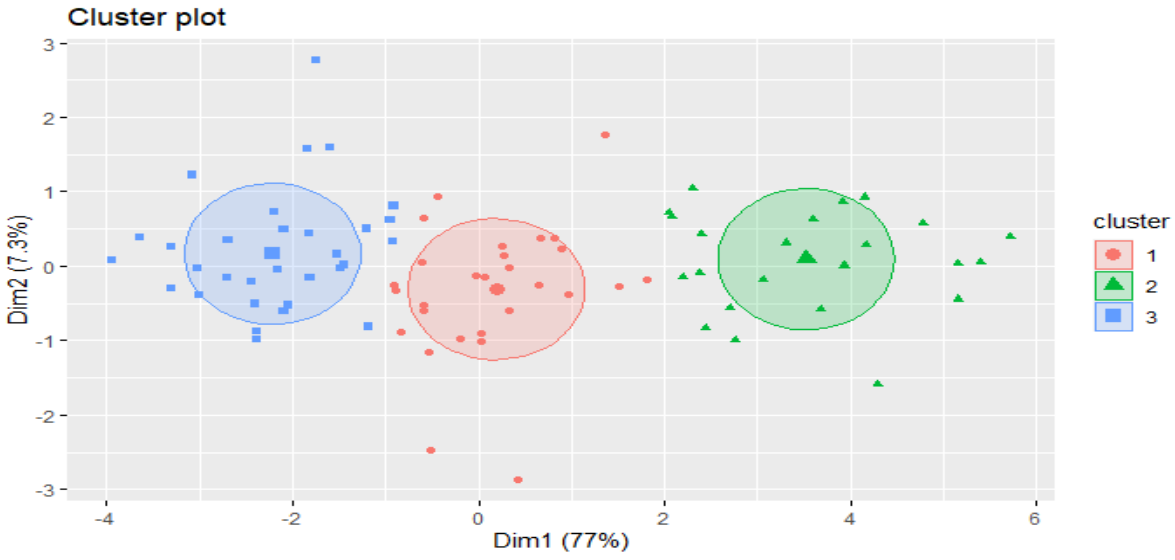
The table below shows the within sum of squares.

TABLE: 2.32

Within cluster sum of squares	1	2	3
Values	138.50000	134.34375	90.07692

The plot below shows clustering

Plot 2.9



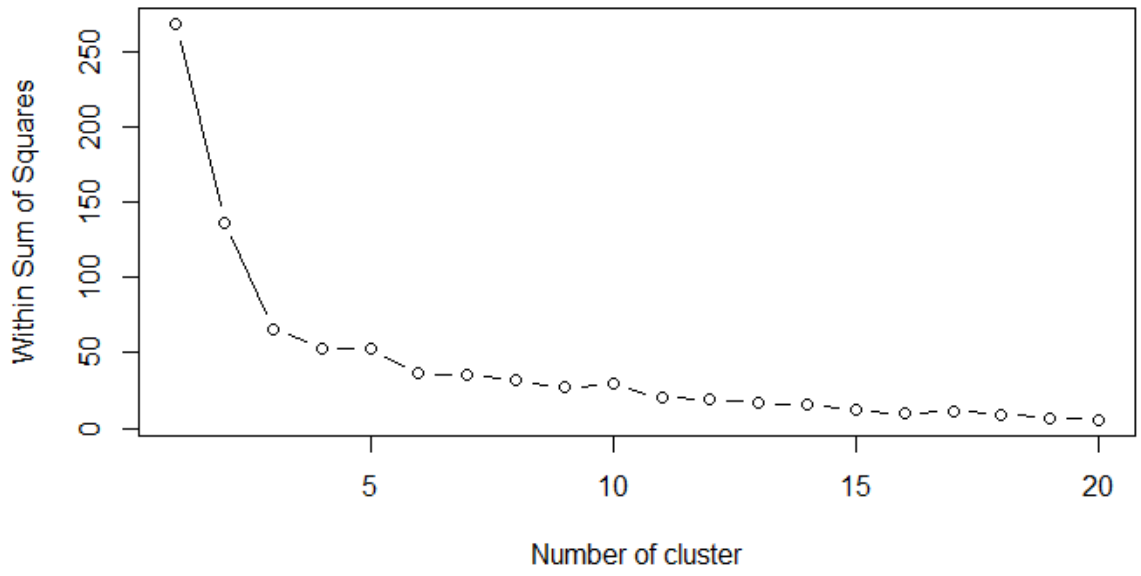
Interpretation:

The calculated value for ratio (between_SS / total_SS = 67.5 %) which is a good amount to consider clusters. The cluster means *table 4.5.2* shows that

- Cluster 1 can be formed by considering services like 'SMS/MMS ', 'internet browsing', 'Video streaming' and 'Clarity of voice' as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like 'Call connection', 'SMS/MMS' and 'Clarity of voice' as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like 'Clarity of voice' and 'SMS.MMS' as they have the highest cluster mean value in cluster 3.

6. For Idea (VI)

Plot no. 2.10: Elbow plot for choosing no. of clusters



Here the graph shows that majorly till 3rd cluster the within sum of squares keep on decreasing after which next thing is to randomly choose how much clusters can be made. So number of clusters has been taken as 3.

Table showing size and clusters

TABLE 2.33

Cluster	1	2	3
Size	14	11	5

The table below shows the cluster means or centers: a matrix, which rows are cluster number (1 to 3) and columns are variables

TABLE 2.34: Cluster means

Clust no.	Ser	Call conn.	Clarity of voice	SMS/ MMS	Internet Browsing	Video streaming	Uploading/ downloading	Clarity of voice	Internet Calling
1	4	2.928571	3.142857	3.214286	2.714286	2.785714	2.428571	2.785714	2.642857
2	4	3.818182	4.000000	4.000000	4.000000	3.818182	3.818182	3.909091	4.090909
3	4	1.200000	1.400000	1.200000	1.400000	1.200000	1.200000	1.400000	1.200000

Table below shows grouping a set of objects in the clusters 1, 2, 3.

TABLE 2.35

1	4 5 29 102 128 183 295 320 433 474 508 529 537 542
2	2 41 67 260 323 489 491 492 570 581 583
3	22 284 326 426 433 474

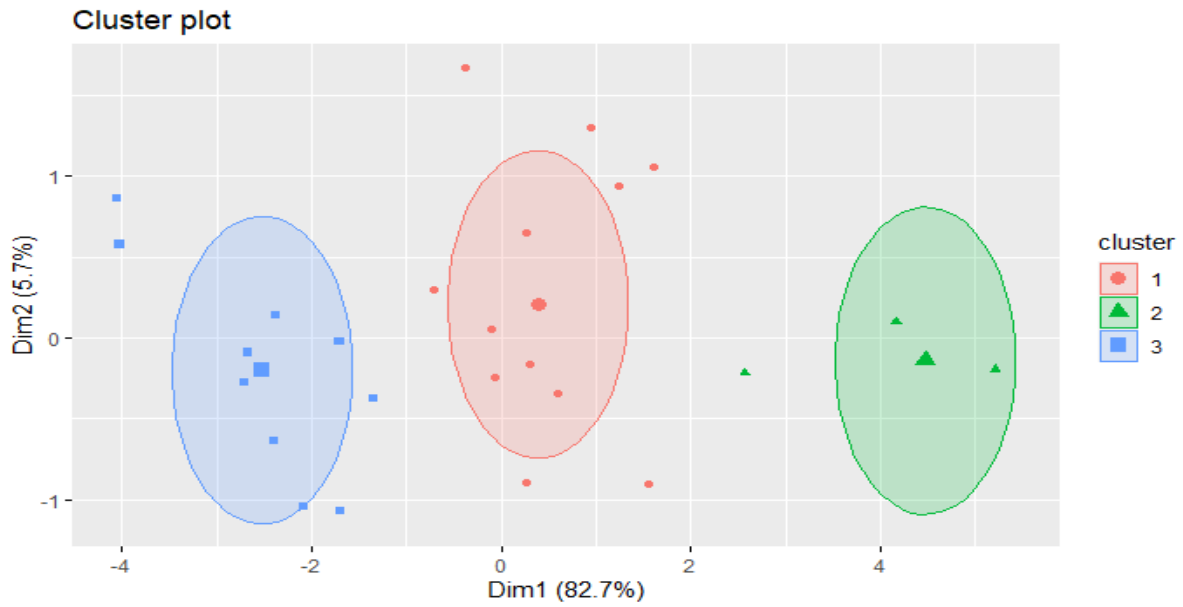
The table below shows the within sum of squares.

TABLE: 2.36

Within cluster sum of squares	1	2	3
Values	33.21429	24.72727	7.60000

The plot below shows clustering

Plot 2.11



Interpretation:

The calculated value for ratio (between_SS / total_SS = 75.5 %) which is a good amount to consider clusters. The cluster means *table 4.5.2* shows that

- Cluster 1 can be formed by considering services like ‘Clarity of voice’ and ‘SMS/MMS’ as they have highest cluster mean value in cluster 1.
- Cluster 2 can be formed by considering services like ‘Clarity of voice’, ‘SMS/MMS’, ‘Internet browsing’ and ‘Internet calling’ as they have the highest cluster mean value in cluster 2.
- Cluster 3 can be formed by considering services like ‘Clarity of voice’, ‘Internet Browsing’ and ‘uploading/downloading’ as they have the highest cluster mean value in cluster 3.

Multiple Logistic Regression:

The model fitted for binary logistic regression is given below

glm(formula = service_sector ~ call_connection + internet_browsing + clarity_of_video, family = binomial, data = data_train)

TABLE 37 Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.5615	0.3487	0.4085	0.5112	1.1633

In the Table 37, the Min, Median and Max values obtained are -2.5615, 0.4085 and 1.1633 respectively whereas first and third quartile obtained as 0.3487 and 0.5112 respectively.

TABLE 38

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	1.1529	0.4463	2.583	0.00978 **
call_connection	0.3282	0.2225	1.475	0.14032
internet_browsing	0.4726	0.2758	1.713	0.08669
clarity_of_video	-0.4786	0.2562	-1.868	0.06182

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

From Table 38 it can be seen that the call connection and internet browsing influences service sector positively with values 0.3282 and 0.4726 respectively whereas clarity of video has a negative effect on service sector with value -0.4786. For significance of the variables the clarity of voice is not significant with p-value 0.14032 while internet browsing and clarity of video also do not show significance but can be taken into consideration since they are slightly > 0.05 taking values 0.08669 and 0.06182 respectively.

The **null deviance** obtained here is 276.4 with 414 degrees of freedom, it basically shows how well the response variable is predicted by a model that includes only the intercept (grand mean) while the **residual deviance** obtained is 265.38 with 411 degrees of freedom and the **AIC** turned out to be 273.38 after performing Step AIC. The **Fisher scoring iteration** required to converge to this model is 5.

To check the goodness of model

H_0: The model is not better than chance (probability) at predicting the outcome.

H_1: The model is better than chance (probability) at predicting the outcome.

Since, the **Chi-square probability** obtained is **0.0035** by considering the difference between values and degrees of freedom of null deviance and residual deviance obtained from the fitted model hence, it can be seen that model is good for predicting the desired result.

Since most of the probabilities for the predicted model are in the favour of private sector which indicates that most of the customers experience better services under private sectors rather than government sector.

The accuracy achieved by predicted probability for train data is **89.55%** and for test data is **87.35%**.

Conclusion:

The obtained predictive model is as follows:

Service sector = 1.1529 + (0.3282* call connection) + (0.4726* internet browsing) – (0.4786* clarity of video)

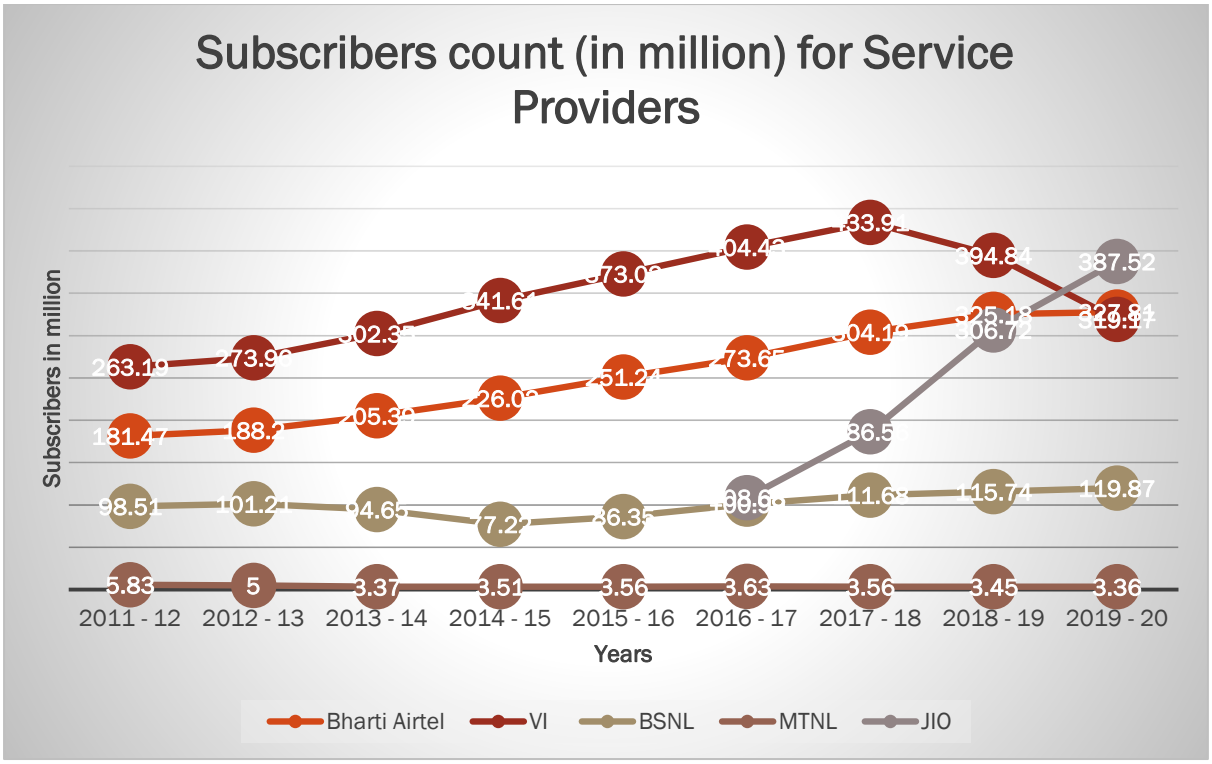
This model can be used to likely predict whether customer will opt for government and private sector.

3.REPORT ON SECONDARY DATA ANALYSIS

A view of data set

TABLE 3.1

Years	**Bharti Airtel**	**VI**	**BSNL**	**MTNL**	**JIO**
2011 – 12	181.47	263.19	98.51	5.83	
2012 – 13	188.2	273.96	101.21	5	
2013 – 14	205.39	302.35	94.65	3.37	
2014 – 15	226.02	341.61	77.22	3.51	
2015 – 16	251.24	373.02	86.35	3.56	
2016 – 17	273.65	404.43	100.99	3.63	108.68
2017 – 18	304.19	433.91	111.68	3.56	186.56
2018 – 19	325.18	394.84	115.74	3.45	306.72
2019 – 20	327.81	319.17	119.87	3.36	387.52



About Analysis:

We have considered TRAI data for this analysis for year range 2011-12 to 2019-20.

TABLE 3.2

The table above shows that trend reliabilty of VI and Jio is much reliable that others which

Service providers	Trend Equation	R^2
Bharti Airtel	y = 15.208x + 269.12	0.4795
VI	y = 20.692x + 150.22	0.9819
BSNL	y = 3.1143x + 85.119	0.3846
MTNL	y = -0.2338x + 5.0881	0.5323
JIO	y = 95.668x - 470.14	0.9927

means that we may need to treat the data other than Vi and Jio.

After working with all the model is the result is that ARIMA model turns out to be much better options for analysis and forecasting which has been decided by comparing AIC value. ARIMA and after that Holt's Linear trend method. Before going for ARIMA model the first step is to check for the stationarity in the model which can be done by using Augmented Dickey-Fuller Test in R.

ARIMA Model

ARIMA is the abbreviation for Auto Regressive Integrated Moving Average. Auto Regressive (AR) terms refer to the lags of the differenced series, Moving Average (MA) terms refer to the lags of errors and I is the difference used to make the time series stationary.

Assumptions of ARIMA model

- Data should be stationary – by stationary it means that the properties of the series doesn't depend on the time when it is captured. A white noise series and series with cyclic behaviour can also be considered as stationary series.
- Data should be univariate – ARIMA works on a single variable. Auto-regression is all about regression with the past values.

Autocorrelation analysis to examine serial dependence: Used to estimate which value in the past has a correlation with the current value. Provides the p,d,q estimate for ARIMA models.

The parameters can be defined as:
p: the number of lag observations in the model also known as the **lag** order.
d: the number of times that the raw observations are differenced also known as the degree of differencing.
q: the size of the moving average window also known as the order of the moving average
Trend: A long-term increase or decrease in the data is referred to as a trend. It is not necessarily linear. It is the underlying pattern in the data over time.

For Bharti Airtel

Best model: ARIMA(0,1,0) with drift.

TABLE 3.3

sigma^2 estimated	Sigma estimated	Log likelihood
327.2	18.08867	-25.34

TABLE 3.4

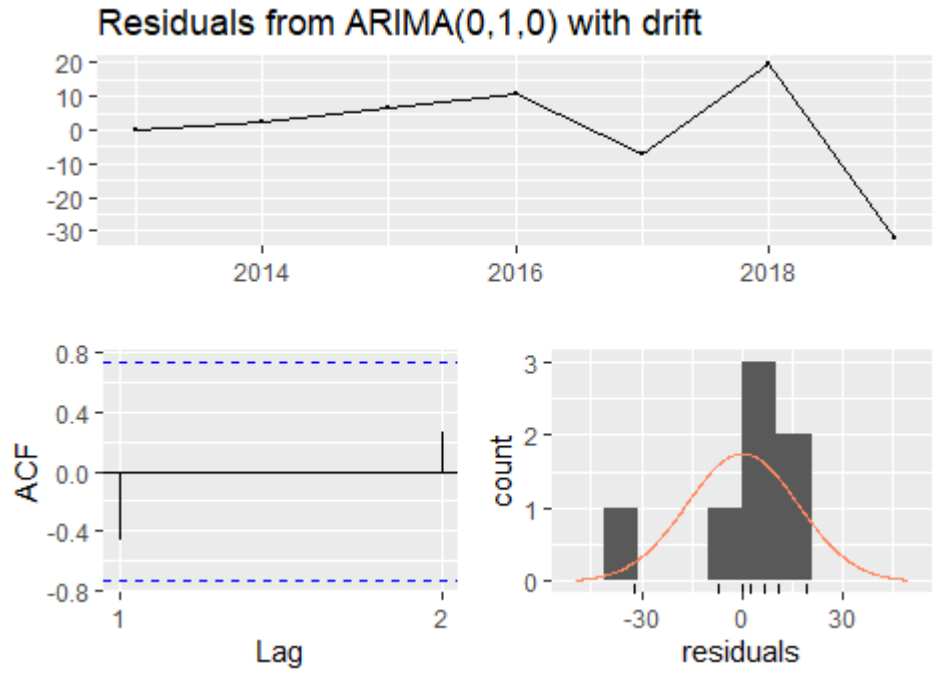
AIC	AIC_C	BIC
34.68	58.68	54.26

The output for the model shows sd = 18.08867 along with AIC = 34.68 is comparatively less. There is a a good fit for further forecasting.

Hypothesis:

$H0_{(Airtel)}$**:** The fitted model is independent of residual and is good fit for further forecasting for Bharti Airtel.

$H1_{(Airtel)}$: The fitted model is not independent of residual and is good fit for further forecasting for Bharti Airtel.



Conclusion: Since p-value = 0.3001 > 0.05, do not reject null hypothesis. Therefore, it can be concluded that the fitted ARIMA model is independent of residuals and is good fit for further forecasting.

Forecast:

The table shows forecasted value for 5 years from 2020-21.

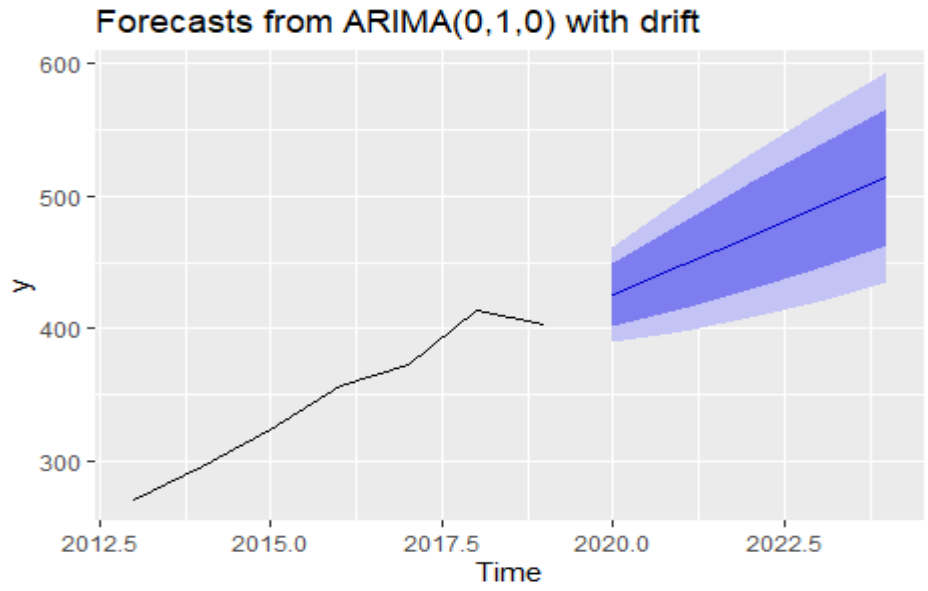
TABLE 3.5

Year	Forecast	Lo 80	Hi 80	Lo 95	Lo 95
2020	425.72	402.5376	448.9024	390.2656	461.1744
2021	447.79	415.0051	480.5749	397.6498	497.9302
2022	469.86	429.7069	510.0131	408.4511	531.2689
2023	491.93	445.5652	538.2948	421.0211	562.8389

2024	514	462.1625	565.8375	434.7214	593.2786

Here, lo80 and hi80 represents the lower and upper bounds respectively of some 80% interval. 80% confident that the true population forecast is somewhere between lo80 value and your hi80 value. The lo95 and hi95 would then be the same thing but for 95%Cis and the forecasted value is in between that interval.

Graph for the forecast



The forecasts are shown as a blue line, with the 80% prediction intervals as a dark shaded area, and the 95% prediction intervals as a light shaded area.

Graph also shows that Bharti Airtel will face a glory of increasing subscribers in the coming years.

To check for accuracy for Bharti Airtel

TABLE 3.6

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.03559427	15.28828	11.29274	0.1143077	2.945688	0.4434901	-0.4628092

0.036	15.29	11.3	11.43%	29.46%	0.44	-0.463
It shows almost negligible margin of error which indicates that fitted model is good fit	Basically it shows standard deviation for residuals from the model	The difference between actual and forecasted value is around 11.3 and is acceptable	MPE value says that the residual level in the model is 11.43% which is acceptable	MAPE value says that the residual level from the model is 29.46% which is acceptable	MASE <1 indicates a good fit	ACF1 gives negative value which indicates our model have some lags in it

Since, all the parameters value have fairly reached up to mark.

Conclusion for Bharti Airtel

Let see the comparison of forecasted value with Actual value for year 2020-21

TABLE 3.7

ARIMA Forecasted subscribers value	425.72
Actual Subscribers Value	423.29

From above it is clear that model is acceptable with 80% and 95% CI therefore model is fairly a good fit.

Final Model Conclusion:

We have considered the data from 2011-12 to 2019-20 and did not consider the 2020-21 just to make a comparison between actual and predicted value

From the accuracy checking from the AIC = 34.68 value it is clear that the ARIMA model is a good fit and is acceptable for Bharti Airtel.

Also, according to the forecasting value along with the graph shows that there will be chunks of glory in the coming years for Bharti Airtel.

For MTNL

Best model: ARIMA (0, 1, 0)

TABLE 3.8

sigma^2 estimated	Sigma estimated	Log likelihood
0.4247	0.6516901	-7.93

TABLE 3.9

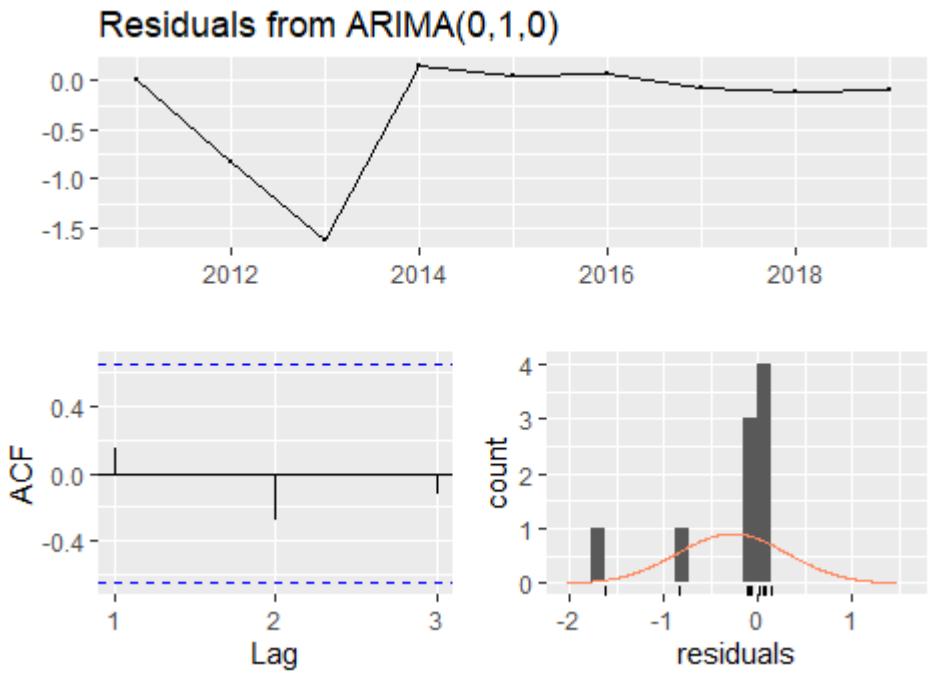
AIC	AIC_C	BIC
17.85	18.52	17.93

The output for the model shows AIC = 17.85 is comparatively less so fitted model is a good fit for further forecasting.

Hypothesis for ARIMA model:

H0A**:** The fitted model is independent of residual and is good fit for further forecasting for MTNL.

H1A: The fitted model is not independent of residual and is good fit for further forecasting for MTNL.



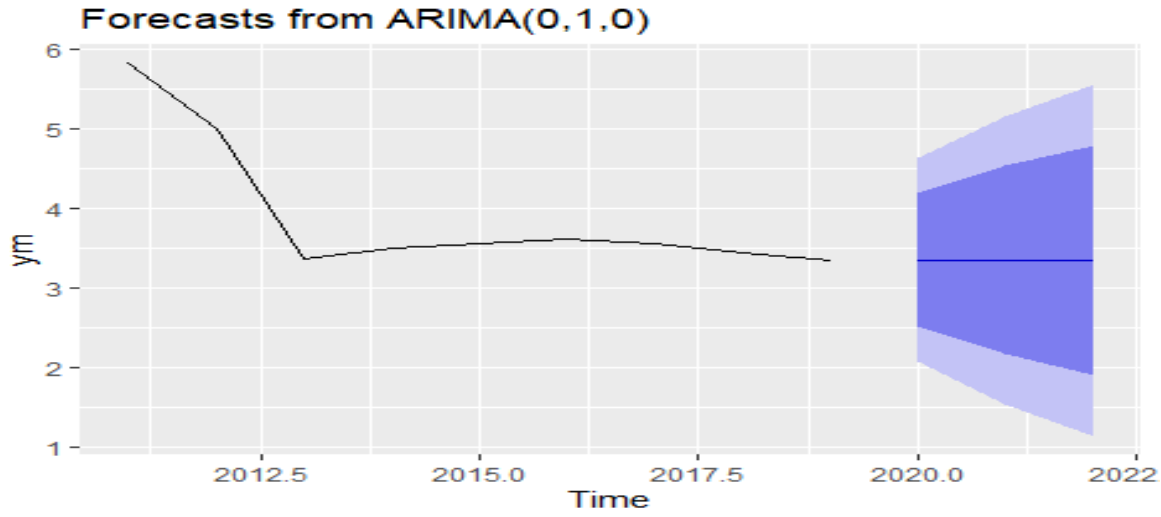
Conclusion: Since p-value = 0.6606 > 0.05, we do not reject null hypothesis. Therefore, it can be concluded that the fitted ARIMA model is independent of residuals and is good fit for further forecasting.

Forecast:

Forecasts: The table shows forecasted value for 5 years from 2020-21

TABLE 3.10

Years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	3.36	2.524784	4.195216	2.082648	4.637352
2021	3.36	2.178827	4.541173	1.553552	5.166448
2022	3.36	1.913364	4.806636	1.147562	5.572438



The graph shows a strange single value forecast according to 80% and 95% CI prediction. This is not acceptable therefore to approach for another method.

For more accurate results Holt's Method has been used

Hypothesis for Holt's model:

H0A**:** The fitted model is independent of residual and is good fit for further forecasting for MTNL.

H1A: The fitted model is not independent of residual and is good fit for further forecasting for MTNL.

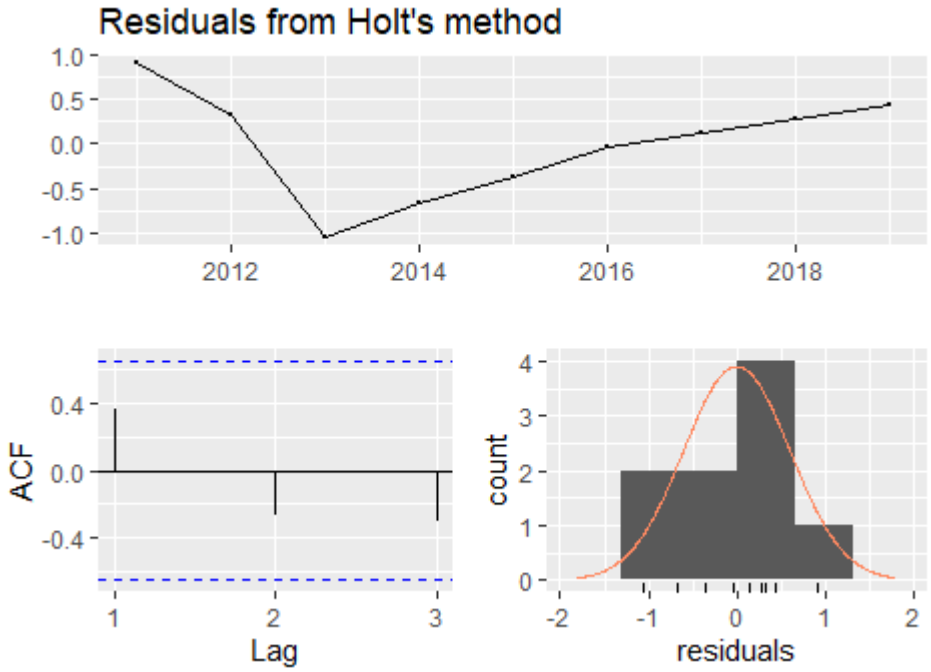
TABLE 3.11

Smoothing parameters	Initial states

Alpha	beta	I	b	Sigma
1e-04	1e-04	5.1739	-0.2498	0.7615

TABLE 3.12

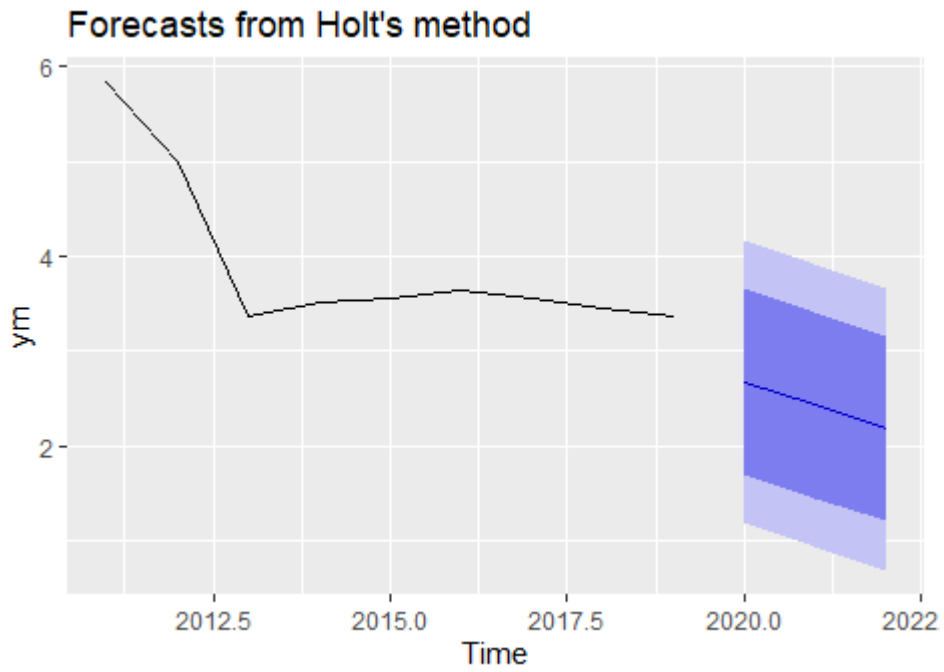
AIC	AIC_C	BIC
19.57945	39.57945	20.56557



Forecast

TABLE 3.13

Years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	2.676188	1.7003469	3.652029	1.1837680	4.168607
2021	2.426420	1.4505795	3.402261	0.9340006	3.918840
2022	2.176653	1.2008120	3.152494	0.6842331	3.669073
2023	1.926885	0.9510445	2.902726	0.4344655	3.419305
2024	1.677118	0.7012769	2.652959	0.1846979	3.169538



Conclusion: The statistic = 7.9684 and p-value = 0.04667 < 0.05 therefore we reject null hypothesis and conclude that model is a fit and it is independent of residuals.

TABLE 3.14

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-0.00629485	0.5675533	0.4672074	-1.674668	12.05093	1.250053	0.3635166
It shows almost negligible margin of error which indicates that fitted model is good fit	Basically it shows standard deviation for residuals from the model	The difference between actual and forecasted value is around 46.72% which is not good	MPE value says that the residual level in the model is negative	MAPE value says that the residual level from the model is 12.05093% which is acceptable	MASE >1 indicates that model needs to improve.	ACF1 gives positive value which indicates that our model is acceptable

Final Conclusion:

To prove this consider an actual value of 2020-21 and is as follows:

TABLE 3.15

ARIMA forecasted value for 2020	3.36 million subscribers
Holt's forecasted value for 2020	2.67 million subscribers
Actual value for 2020	2.17 million subscribers

Although p-value of both the model used says that fitted model is a good fit but the forecast value of Holt tend to be more accurate than ARIMA model but according to the AIC value ARIMA is considered to be best as the AIC value for ARIMA and Holt are 17.85 and 19.57945 respectively. Finally from the model it is concluded that model is acceptably good fit.

For BSNL

Best model: ARIMA (0, 1, 0)

TABLE 3.16

sigma^2 estimated	Sigma estimated	Log likelihood
99.96		29.77

TABLE 3.17

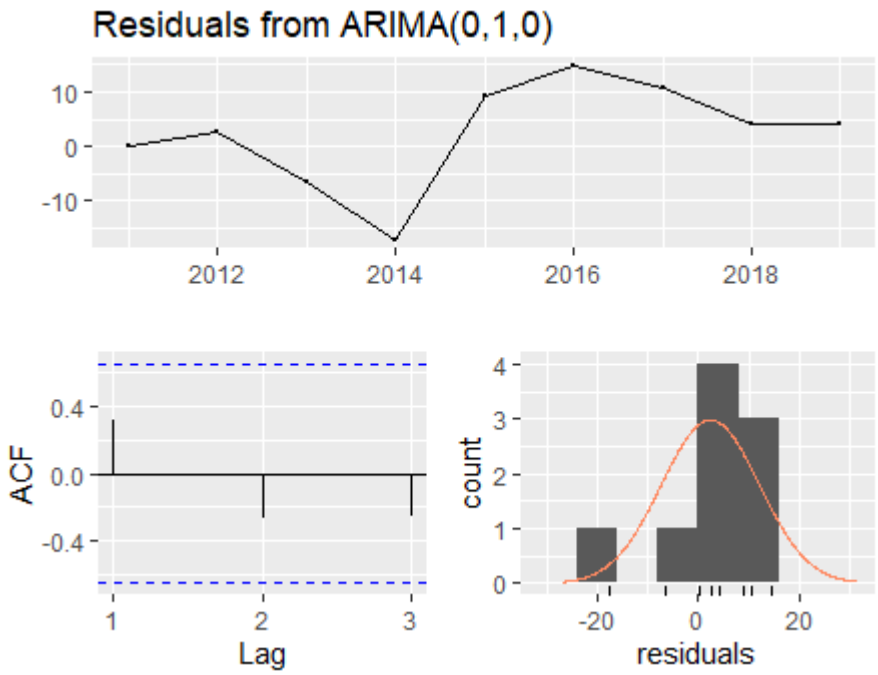
AIC	AIC_C	BIC
61.54	62.21	61.62

The output for the model shows AIC = 61.54 is comparatively less so fitted model is a good fit for further forecasting.

Hypothesis for ARIMA model:

H0A**:** The fitted model is independent of residual and is good fit for further forecasting for BSNL.

H1A: The fitted model is not independent of residual and is good fit for further forecasting for BSNL.



Conclusion: Since p-value = 0.335 > 0.05, we do not reject null hypothesis. Therefore, it can be concluded that the fitted ARIMA model is independent of residuals and is good fit for further forecasting.

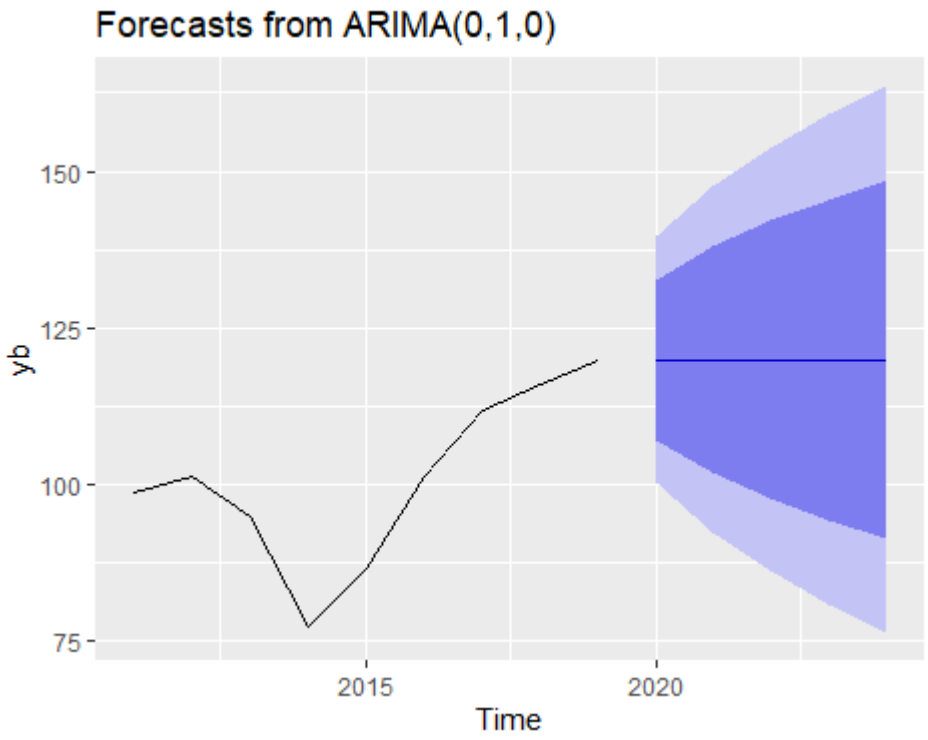
Forecast:

Forecasts: The table shows forecasted value for 5 years from 2020-21.

TABLE 3.18

Years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	119.87	107.05736	132.6826	100.27475	139.4652
2021	119.87	101.75019	137.9898	92.15814	147.5819
2022	119.87	97.67785	142.0621	85.93004	153.8100
2023	119.87	94.24472	145.4953	80.67951	159.0605
2024	119.87	91.22006	148.5199	76.05370	163.6863

Graph for the forecast



Accuracy

TABLE 3.19

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
2.384279	9.425977	7.71539	1.651115	8.207264	0.8901517	0.3224235
Margin of error = 2.38 which indicates that fitted model is good fit	Basically it shows standard deviation for residuals from the model	The difference between actual and forecasted value is around 7.71539 and is acceptable	MPE value says that the residual level in the model is 16.52% which is acceptable	MAPE value says that the residual level for the model is 82.073% which says model needs to be improve	MASE <1 indicates a good fit	ACF1 gives positive value which indicates our model is good fit

The graph shows a strange single value forecast according to 80% and 95% CI prediction. This is not acceptable therefore to approach for another method.

For more accurate results Holt's Method has been used

TABLE 3.20

AIC	AIC_C	BIC
71.53322	91.53322	72.51934

Conclusion: Here since current value for 2019-20 is not available it is better to proceed with what standard procedure says, as the AIC for ARIMA and Holt model are 61.54 and 71.53322 respectively. Therefore AIC value for ARIMA model is less than that of Holt's model hence we will proceed with ARIMA model. Hence ARIMA model is a good fit with an acceptable forecasting based on 80% 95% rule.

For Reliance Jio

Best model: ARIMA(0,1,0)

TABLE 3.21

sigma^2 estimated	Sigma estimated	Log likelihood
9011	94.92629	-17.92

TABLE 3.22

AIC	AIC_C	BIC
37.83	41.83	36.93

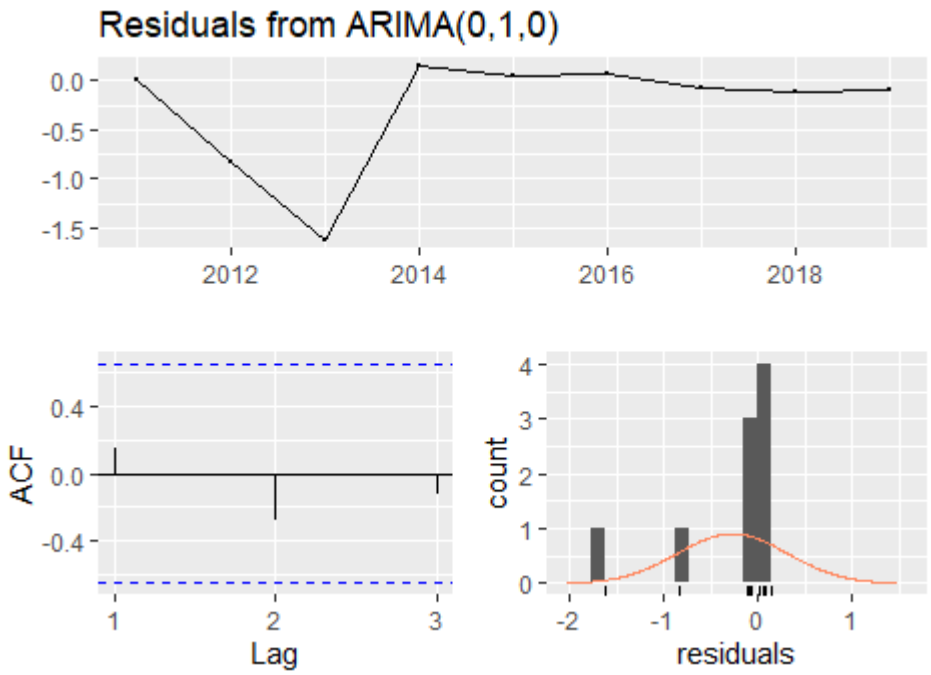
The output for the model shows AIC = 37.83 is comparatively less.

There is a a good fit for further forecasting.

Hypothesis for ARIMA model:

$H0_{(Jio)}$**:** The fitted model is independent of residual and is good fit for further forecasting for Reliance Jio.

$H1_{(Jio)}$: The fitted model is not independent of residual and is good fit for further forecasting for Relaince Jio.



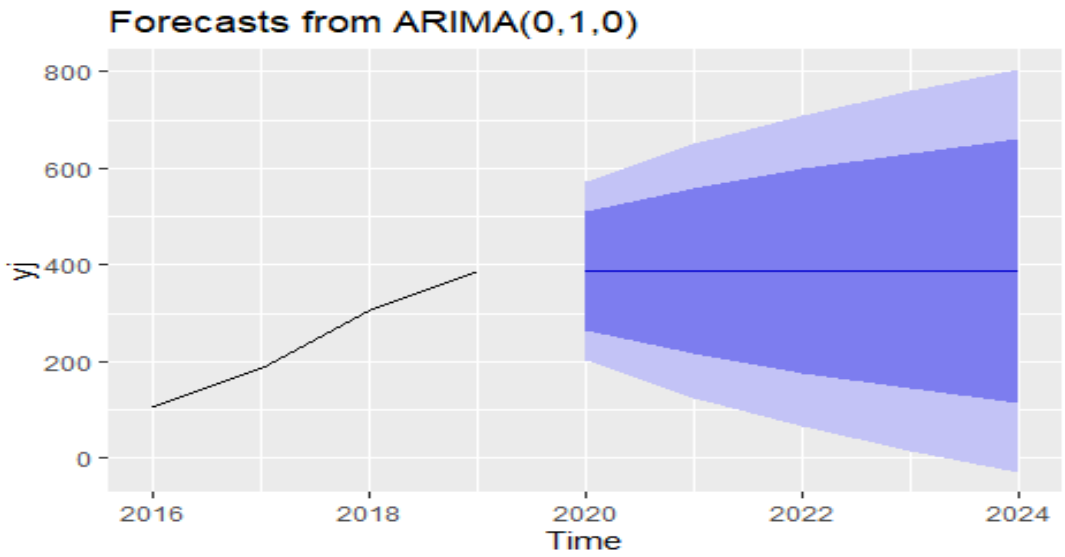
Conclusion: Since p-value = 0.4377 > 0.05, we do not reject null hypothesis. Therefore, it can be concluded that the fitted ARIMA model is independent of residuals and is good fit for further forecasting.

Forecast:

Forecasts: The table shows forecasted value for 5 years from 2020-21.

TABLE 3.23

Year	Forecast	Lo 80	Hi 80	Lo 95	Lo 95
2020	387.52	265.8685	509.1715	201.47006	573.5699
2021	387.52	215.4788	559.5612	124.40565	650.6344
2022	387.52	176.8134	598.2266	65.27205	709.7680
2023	387.52	144.2170	630.8230	15.42012	759.6199
2024	387.52	115.4989	659.5411	-28.50032	803.5403



Graph for the forecast

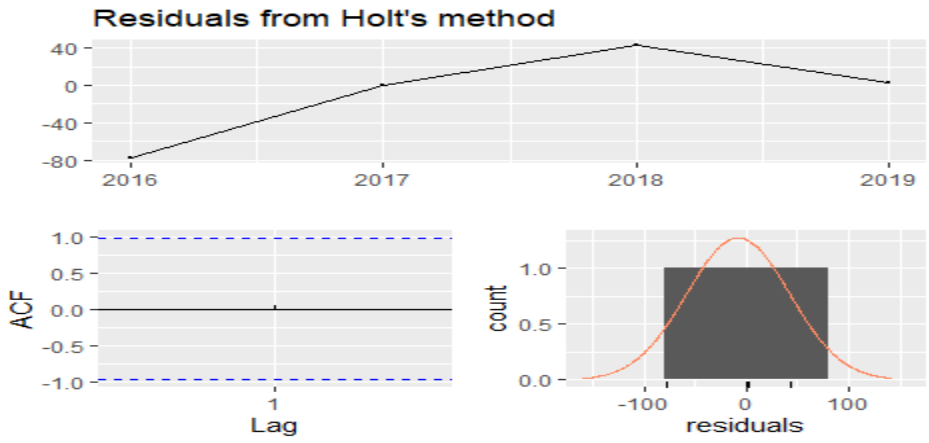
The graph shows a strange single value forecast according to 80% and 95% CI prediction. This is not acceptable therefore to approach for another method.

For more accurate results Holt's Method has been used

Hypothesis for Holt's model:

$H0_{(Jio)}$**:** The fitted model is independent of residual and is good fit for further forecasting for Reliance Jio.

$H1_{(Jio)}$: The fitted model is not independent of residual and is good fit for further forecasting for Relaince Jio



Forecast

TABLE 3.24

Years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	465.40	408.5858	522.2142	378.5102	552.2898
2021	543.28	462.9326	623.6274	420.3993	666.1607
2022	621.16	522.7550	719.5650	470.6625	771.6575
2023	699.04	585.4117	812.6683	525.2605	872.8195
2024	776.92	649.8797	903.9603	582.6286	971.2114

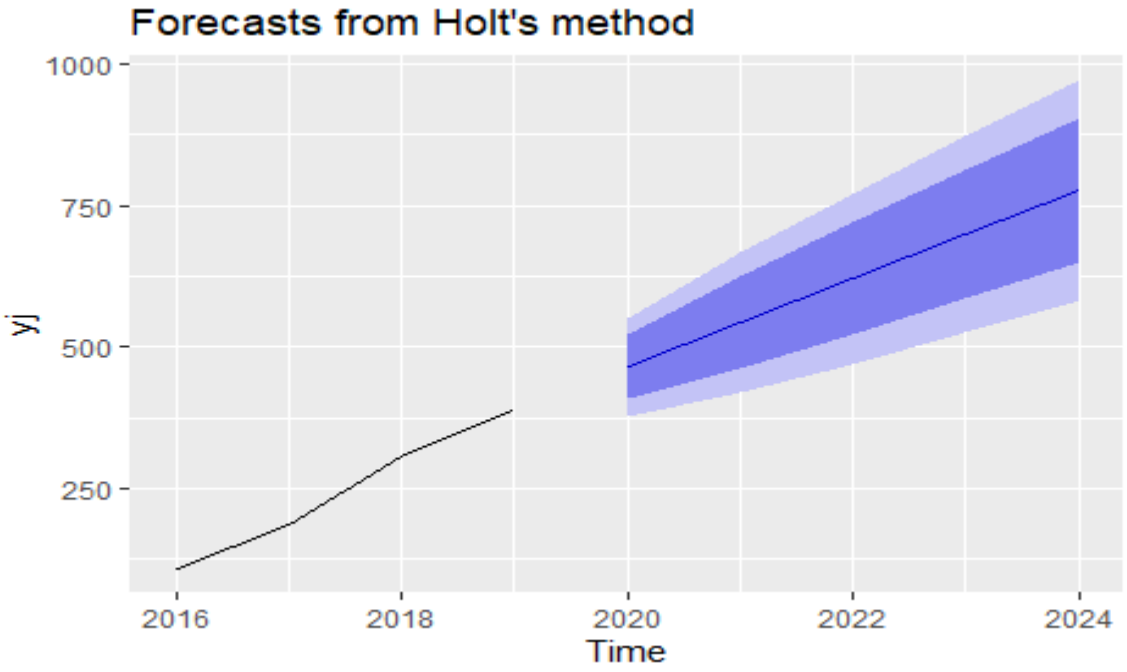


TABLE 3.25

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-8.17	44.33232	30.77	-14.28046	21.5495	0.3310501	0.05295152
It shows almost negligible margin of error which indicates that fitted model is good fit	sd for model is still quite high	The difference between actual and forecasted value is around 30.77 and is acceptable	MPE value says that the residual level in the model is -14.2805%	MAPE value says that the residual level from the model is 21.5495% which is acceptable	MASE <1 indicates a good fit	ACF1 gives < 1 positive value which indicates our model have some lags in it

Still we did not got the value as desired for proper forecasting of the model. This is happening as very least data points has been considered for Jio subscribers as it has been launched in 2016 and we have considered data till 2019-20 which means 3 years.

TABLE 3.26

ARIMA forecasted value for 2020	387.52 million subscribers
Holt's forecasted value for 2020	465.40 million subscribers
Actual value for 2020	410.8 million subscribers

Final Conclusion:

Both the methods are showing fairly good forecast in an inverse proportional manner.

For sure we can search for much more better ways for forecasting the model to get more accurate results but we should keep this in mind that the biggest reason for non-satisfied result is less number of data points.

According to the graph Reliance jio will going to have a good increment in the coming years.

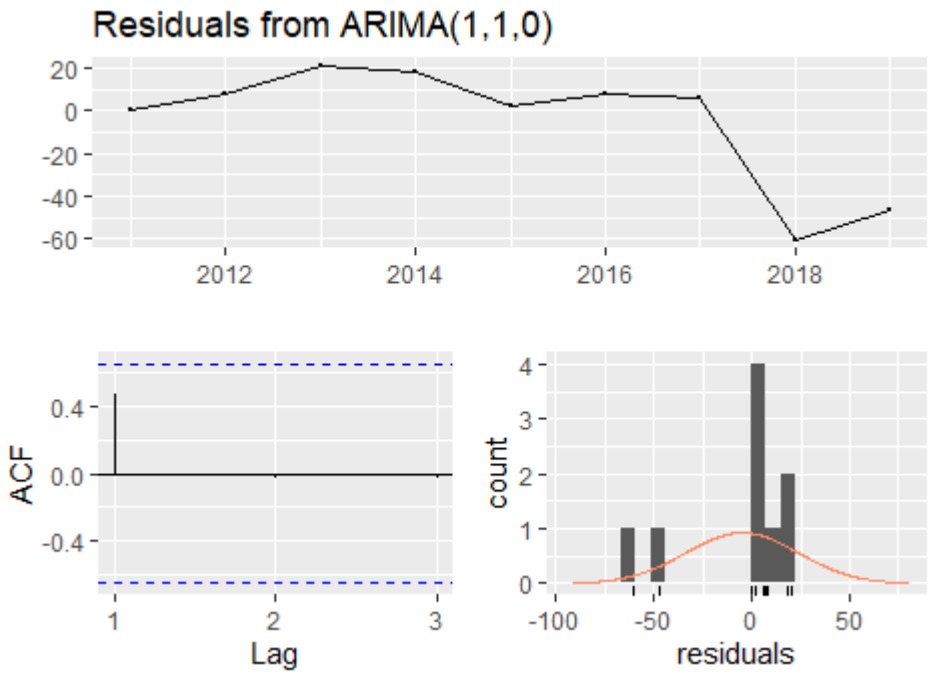
For Vodafone Idea (VI)

Best model: ARIMA(1,1,0) which is also known as AR(1) model.

Hypothesis:

$H0_{(VI)}$**:** The fitted model is independent of residual and is good fit for further forecasting for VI.

$H1_{(VI)}$: The fitted model is not independent of residual and is good fit for further forecasting for VI.



Conclusion: Since p-value > 0.05, we do not reject null hypothesis. Therefore, it can be concluded that the fitted ARIMA model is independent of residuals and is good fit for further forecasting.

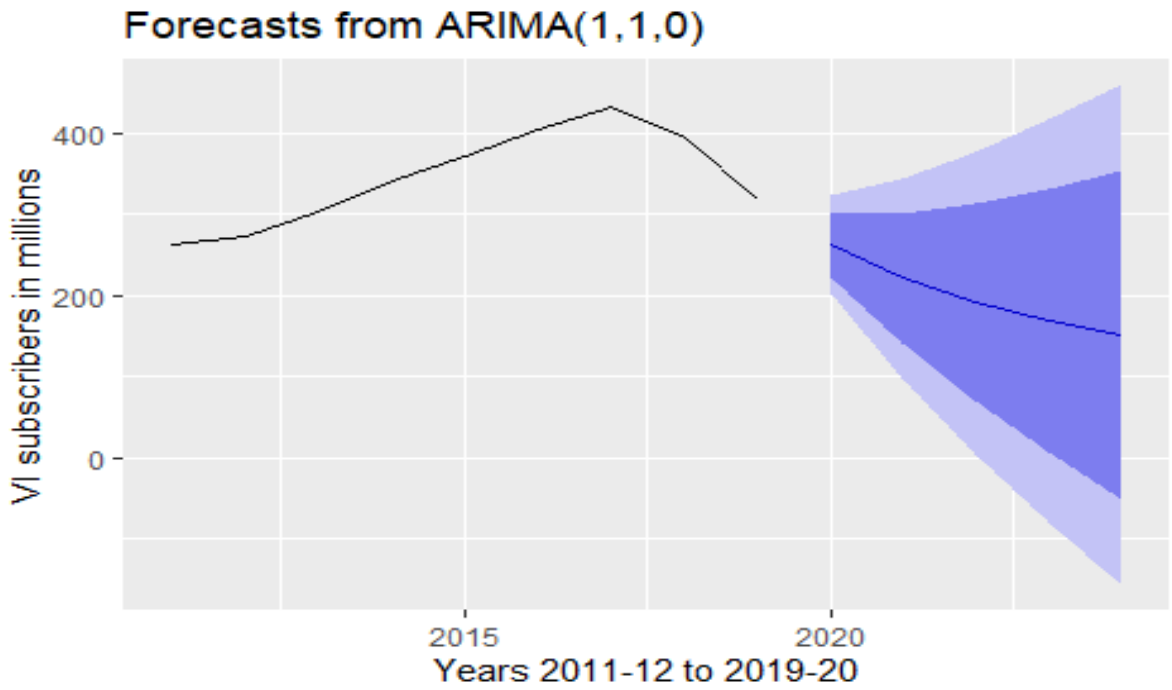
Forecasts: The table shows forecasted value for 5 years from 2020-21.

TABLE 3.27

Point/Years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	263.1127	223.155374	303.0700	202.003248	324.2221
2021	221.5847	141.366689	301.8027	98.901830	344.2676
2022	190.8203	69.145319	312.4952	4.734496	376.9060
2023	168.0296	5.515007	330.5441	-80.514998	416.5741
2024	151.1459	-50.732798	353.0246	-157.600928	459.8927

Here, lo80 and hi80 represents the lower and upper bounds respectively of some 80% interval. 80% confident that the true population forecast is somewhere between lo80 value and your hi80 value. The lo95 and hi95 would then be the same thing but for 95%Cis and the forecasted value is in between that interval.

Graph for the forecast



The forecasts are shown as a blue line, with the 80% prediction intervals as a dark shaded area, and the 95% prediction intervals as a light shaded area.

Graph also shows that they VI will face loss in the coming years.

To check for accuracy for Vodafone Idea (VI)

TABLE 3.28

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.01813082	8.218767	6.313131	0.1351693	2.411149	0.3451213	0.1545139
1.81%	8.22	6.31	14.52%	24.12%	0.345	0.155
It shows around 2% margin of error which indicates that fitted model is good fit	Basically it shows standard deviation for error.	The difference between actual and forecasted value is 6.31 and is acceptable	MPE value says that the error level is 14.52% which is acceptable	MAPE value says that the residual level is 24.12% which is acceptable	MASE <1 indicates a good fit	ACF1 gives positive value which indicates our model is acceptable

Final Conclusion:

TABLE 3.29

Actual subscribers count	286 million
ARIMA forecasted count	263.1127

From the accuracy checking of the model we can say that the ARIMA model is a good fit and is acceptable for Vodafone Idea (VI) Service.

We have considered the data till 2019-20 and did not consider the 2020-21 just to make a comparison between actual and predicted value

Also, according to the forecasting value along with the graph shows that there will be decrement in the coming years for this service provider.

So it can conclude that before the merge the both the services idea and Vodafone the quality was good but after the merge the something happened because of which the subscribers where unsatisfied with the service/deals they get, maybe because of some better service provider stepped in the market of telecommunication or maybe the quality after merge got decreased or much better offers like free subscriptions of web channels like (Amazon prime, Disney+Hotstar and so on) etc.

SUGGESTION

1. Indian telecommunication ranks second in the world due to its quick evolution and its neck to neck competition with telecom of other developing countries. The private sectors (Airtel, Jio, Vodafone, idea, etc) promises to provide better quality, cheaper rates, less expenses, minimum faults, good network etc which are the market strategy of almost all telecom industries.
2. The objective of this study was to visualize the compatibility of the network currently used in India. Among the used network which one of them is most preferred by the customers. Apart from the increasing private players BSNL/MTNL are also the major public players.
3. The study was to analyse how many of the customer still prefer public over private. Although the public sector has a huge infrastructure and earlier had mover advantage their performance is still low compared with the public sector. Considering the previous

point what were the strategies undertaken by the private sector to aim as a successful industry is also viewed under the study.

4. Although companies promise perfect networks but they all lag behind in some or the other ways. Even though they offer new opportunities the percentage of quality still remains the same since private industries more or less opt to the same trends. So, the industries should focus on performance rather than opportunities.
5. From our daily messaging to managing governments are under the influence of telecommunication industries. Since this sector is growing globally on a large scale it needs to be analysed to enhance its qualities and services, so that the consumers get the best out of it. It can be done with the help of Artificial Intelligence (AI) and Machine Learning (ML) techniques.
6. There are so many techniques available which can used to analyse the services provided by the telecommunication sector and can even help to develop predictive models which will help us to provide even much better services than now and make our future effortless and experience the fastest network services with low cost. It will make our work life more flexible and balanced.
7. Government telecom industry need to establishment telecom industry in rural & semi-rural area they can try to grow its own assets by using real estate, and it can give you a job's opportunities for rural people. Government have withdrawn a lot of things to benefit telecom sector but by the time it gets executed to the market, it becomes too late. Service providers have to incur huge initial fixed cost to enter semi-rural and rural areas.
8. COVID-19 ways the telecoms industry can help keep everyone National industry-wide coordination mechanisms should be instituted, for coordination of network management during the crisis. Emergency telecommunications plans should be put in place and where they are in place, reviewed for their adequacy for the circumstances instituting clear traffic prioritization rules and ensuring that emergency services and coordination bodies are equipped with resilient “off-grid” communications, such as satellite communication devices
9. Considering TRAI data for past few year it can be seen that government sector's subscribers count keeps on decreasing year by year because of the no update has been provided for the reliability of the customer. A new came up news is that the government will going to provide some updated services hopefully that will change the pattern of continuous decrease or constant graph. For the private services which are going down

market instead of merging with some other service providers, one could work on the things which are in hand, by thinking and executing for the areas which have got very poor ratings. This can be done with the help of brilliant advance software techniques available in the market.

10. India has very little penetration of fixed-line in its network whereas most of the developed countries have a very high penetration of fixed lines. Only around 25% of Towers in India are connected with fibre networks, whereas in developed nations, it is in excess of 70%. 5G Network requires towers to be connected to with very high-speed systems. Although companies are going to launch 5G networks soon so that would be a challenge for the companies.

APPENDIX

Primary Data R-Codes

```
Data<-read.csv(file.choose(),header=T)

dim(Data)

Gender <- Data$Gender

Age <- Data$Age

X1 <- Data$Q1..Which.network.are.you.using.now..if.dual.sim.user.can.choose.more.than.1.

X2 <- Data$Q2..From.which.part.of.Country.you.belong.to.

X3 <- Data$Q3..What.kind.of.service.you.prefer.


X4A <- Data$X4.Airtel

X4I <- Data$X4.Idea.VI.

X4J <- Data$X4.Jio

X4MB <- Data$X4.MTNL.BSNL

X4V <- Data$X4.Vodafone.VI.

X4 <- data.frame(X4A,X4I,X4J,X4MB,X4V)
```

```
X5CC <- Data$X5.Call.connection

X5CVo <- Data$X5.Clarity.of.voice

X5S <- Data$X5.SMS.MSS

X5I <- Data$X5.Internet.Browsing.google.chrome.etc.

X5V <- Data$X5.Video.Streaming.youtube.instagram.etc.

X5S.1 <- Data$X5.Social.networking.apps..Instagram.Fb.whatsapp.etc.

X5U <- Data$X5.Uploading...downloading.data..file

X5CVi <- Data$X5.Clarity.of.video

X5IC <- Data$X5.Internet.Audio.video.Calling..social.media.....calls.

X5 <- data.frame(X5CC,X5CVo,X5S,X5I,X5V,X5S,X5U,X5CVi,X5IC)


X6Cd <- Data$X6.Call.drop

X6L <- Data$X6.lag.in.audio.video

X6M <- Data$X6.late.message.delivered

X6CrC <- Data$X6.cross.connection

X6 <- data.frame(X6Cd,X6L,X6M,X6CrC)


X7YN <- Data$Q7..Did.pandemic.affect.your.network.service.

X7Y <- Data$If.yes..choose.from.the.below.option.which.section.was.affected.the.most..can.select.more.than.1.

X7 <- data.frame(X7YN,X7Y)


X8Cb <- Data$X8.Chatbot
```

```
X8T <- Data$X8.Time.taken.to.fix.issue

X8CCN <- Data$X8.Availability.of.customer.care.no.

X8C <- Data$X8.Response.to.the.complaints

X824h <- Data$X8.24.hours.customer.service

X8PS <- Data$X8.Politeness.of.Staff

X8o <- Data$X8.Offers.provided.by.the.service.provider

X8 <- data.frame(X8Cb,X8T,X8CCN,X8C,X824h,X8PS,X8o)


X9E <- Data$X9.Service.meets.my.expectations

X9VM <- Data$X9.Value.for.the.money

X9 <- data.frame(X9E,X9VM)


X10 <- Data$Q10..How.likely.is.it.that.you.would.recommend.your.current.network.to.your.friends.and.family.

X11 <- Data$Q11..Which.network.works.best.in.your.area.....

X12 <- Data$Q12..If.you.want.to.switch.then.which.other.SIM.would.you.like.to.prefer.from.the.networks.listed.below.

X13 <- Data$Q13..Although.companies.are.saying.3G.4G.network.are.better.but.customers.were.experiencing.better.network.quality.when.only.2G.has.been.introduced..Do.you.agree.

X14 <- Data$Q14..The.update.made.by.the.telecom.networks.has.improved.the.network.quality.

X15 <- Data$Q15..Do.you.think.during.pandemic.we.were.dependent.on.networks.for.communication.and.other.uses.
```

```
X16 <- Data$Q16..Which.sector.provide.better.network.Government.or.Private.


Data1 <- data.frame(Gender, Age, X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, X13,
X14, X15, X16)

names(Data1)

summary(Data1)

Data2 = Data1

#Treating missing values in Data1

Data2$X2[which(is.na(Data2$X2))] = median(Data2$X2,na.rm=T)

Data2$X2[Data2$X2==4] <- 1

Data2$X16[Data2$X16==2] <- 1

Data2$X5CC[which(is.na(Data2$X5CC))] = median(Data2$X5CC,na.rm=T)

Data2$X5CVo[which(is.na(Data2$X5CVo))] = median(Data2$X5CVo,na.rm=T)

Data2$X5S[which(is.na(Data2$X5S))] = median(Data2$X5S,na.rm=T)

Data2$X5I[which(is.na(Data2$X5I))] = median(Data2$X5I,na.rm=T)

Data2$X5V[which(is.na(Data2$X5V))] = median(Data2$X5V,na.rm=T)

Data2$X5S.1[which(is.na(Data2$X5S.1))] = median(Data2$X5S.1,na.rm=T)

Data2$X5U[which(is.na(Data2$X5U))] = median(Data2$X5U,na.rm=T)

Data2$X5CVi[which(is.na(Data2$X5CVi))] = median(Data2$X5CVi,na.rm=T)

Data2$X5IC[which(is.na(Data2$X5IC))] = median(Data2$X5IC,na.rm=T)

Data2$X6L[which(is.na(Data2$X6L))] = median(Data2$X6L,na.rm=T)

Data2$X6M[which(is.na(Data2$X6M))] = median(Data2$X6M,na.rm=T)

Data2$X6CrC[which(is.na(Data2$X6CrC))] = median(Data2$X6CrC,na.rm=T)

Data2$X10[which(is.na(Data2$X10))] = median(Data2$X10,na.rm=T)

Data2$X11[which(is.na(Data2$X11))] = median(Data2$X11,na.rm=T)
```

```
Data2$X14[which(is.na(Data2$X14))] = median(Data2$X14,na.rm=T)

Data2$X15[which(is.na(Data2$X15))] = median(Data2$X15,na.rm=T)

summary(Data2)

names(Data2)



#Splitting data into training(70%) and test(30%)

set.seed(123)

split <- sample(2,nrow(Data2), replace=TRUE,prob=c(0.7,0.3))

data_train <- Data2[split==1,]

data_test <- Data2[split==2,]

nrow(data_train)

nrow(data_test)


##### Multiple Logistic Regression #####

is.factor(X16)

data_LR <- rbind(data_test,data_train)

data_train$X16 <- factor(data_train$X16)

data_test$X16 <- factor(data_test$X16)


glm.fit.ser <- glm(X16 ~ X5CC+X5CVo+X5S+X5I+X5V+X5U+X5CVi+X5IC,

                   data = data_train, family=binomial)


#To fit model using stepAIC
```

```
library(MASS)
smodel.ser <- step(glm.fit.ser,direction='both',trace=F)
smodel.ser
summary(smodel.ser)

#To check goodness of fitted model
chisq.prob <- 1-pchisq(13.6,3)

#To make predictions
s_prob_test.ser <- predict(smodel.ser, data_test, type="response")
length(s_prob_test.ser)
nrow(data_test)

s_prob_train.ser <- predict(smodel.ser, data_train, type="response")
length(s_prob_train.ser)
nrow(data_train)

#To check accuracy of model using predictions
s_class_pred_train.ser <-
  ifelse(s_prob_train.ser > 0.50,1,0)

s_class_pred_test.ser <-
  ifelse(s_prob_test.ser > 0.50,1,0)
```

```
table(s_prob_test.ser, data_test$X16)

s_acc_test.ser <- mean(s_class_pred_test.ser == data_test$X16)

s_acc_test.ser


table(s_prob_train.ser, data_train$X16)

s_acc_train.ser <- mean(s_class_pred_train.ser == data_train$X16)

s_acc_train.ser


#Mann-whitney U test

wilcox.test(X5CC~X16, data=data_LR, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


wilcox.test(X5I~X16, data=data_LR, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


wilcox.test(X5CVi~X16, data=data_LR, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


#For prepaid postpaid with X5,X6,X8


data_new                                                                    <-
data.frame(Data2[,c("X3","X5CC","X5CVo","X5S","X5I","X5V","X5U","X5CVi","X5IC",


"X6Cd","X6L","X6M","X6CrC","X8Cb","X8T","X8CCN","X8C","X824h","X8PS","X8o")]
)
```

```
data_pp <- subset(data_new, X3<3)
nrow(data_pp)
data_pp$X3[data_pp$X3==1] <- 0
data_pp$X3[data_pp$X3==2] <- 1

#Splitting data into training(70%) and test(30%)
set.seed(123)
split <- sample(2,nrow(data_pp), replace=TRUE,prob=c(0.7,0.3))
data_train_pp <- data_pp[split==1,]
data_test_pp <- data_pp[split==2,]
nrow(data_train_pp)
nrow(data_test_pp)

is.factor(X3)
data_pp$X3 <- as.factor(data_pp$X3)
data_train_pp$X3 <- as.factor(data_train_pp$X3)
data_test_pp$X3 <- as.factor(data_test_pp$X3)

glm.fit.pp <- glm(X3 ~
X5CC+X5CVo+X5S+X5I+X5V+X5U+X5CVi+X5IC+X6Cd+X6L+X6M+X6CrC+
                  X8Cb+X8T+X8CCN+X8C+X824h+X8PS+X8o,
                data = data_train_pp, family=binomial)
glm.fit.pp
summary(glm.fit.pp)
```

```
#To fit model Using stepAIC

library(MASS)

smodel_pp <- step(glm.fit.pp, direction='both',trace=F)

smodel_pp

summary(smodel_pp)


#To check goodness of fitted model

chisq.prob_pp <- 1 - pchisq(23.6,6)


#To make predictions

s_prob_test_pp <- predict(smodel_pp, data_test_pp, type="response")

length(s_prob_test_pp)

nrow(data_test_pp)


s_prob_train_pp <- predict(smodel_pp, data_train_pp, type="response")

length(s_prob_train_pp)

nrow(data_train_pp)


#To check accuracy of model using predictions

s_class_pred_train_pp <-

  ifelse(s_prob_train_pp > 0.50,1,0)


s_class_pred_test_pp <-

  ifelse(s_prob_test_pp > 0.50,1,0)
```

```
table(s_prob_test_pp, data_test_pp$X3)

s_acc_test_pp <- mean(s_class_pred_test_pp == data_test_pp$X3)

s_acc_test_pp


table(s_prob_train_pp, data_train_pp$X3)

s_acc_train_pp <- mean(s_class_pred_train_pp == data_train_pp$X3)

s_acc_train_pp


summary(smodel_pp)


#Mann-whitney U test

wilcox.test(X5CVi~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


wilcox.test(X5IC~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


wilcox.test(X6L~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)


wilcox.test(X6M~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",

        exact=F, correct=T)
```

```
wilcox.test(X8PS~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",
            exact=F, correct=T)

wilcox.test(X8o~X3, data=data_pp, mu=0, paired=F, conf.int=T, alt="two.sided",
            exact=F, correct=T)

##### k-mean Clustering Analysis #####

#Creating Dataframe
c_pg_data <-  Data2[,c(41,26:32)]
dim(c_pg_data)
names(c_pg_data)

#For Government Network----------------------
data_govt <- c_pg_data[c_pg_data$X16==0,]

#To select no. of clusters using Scree Plot (elbow method)
wss_govt <- (nrow(data_govt)-1)*sum(apply(data_govt,2,var))
for (i in 2:20) wss_govt[i] <- sum(kmeans(data_govt, centers=i)$withinss)
plot(1:20, wss_govt, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")

#for clustering
kc_govt <- kmeans(data_govt, 5)
kc_govt
```

```
#For plotting cluster
install.packages("factoextra")
library(factoextra)
fviz_cluster(kc_govt, data = scale(data_govt[,2:8]), geom = c("point"),
             ellipse.type = "euclid")



#For Private Network------------------------
data_prvt <- c_pg_data[c_pg_data$X16==1,]
dim(data_prvt)

#To select no. of clusters using Scree Plot (elbow method)
wss_prvt <- (nrow(data_prvt)-1)*sum(apply(data_prvt,2,var))
for (i in 2:20) wss_prvt[i] <- sum(kmeans(data_prvt, centers=i)$withinss)
plot(1:20, wss_prvt, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")

#for clustering
kc_prvt <- kmeans(data_prvt, 4)
kc_prvt

#For plotting cluster
fviz_cluster(kc_prvt, data = scale(data_prvt[,2:8]), geom = c("point"),
```

```
        ellipse.type = "euclid")



#For service providers(X1) with different services(X5, X6, X8, X9)


#Creating dataframe

opt_data <-  Data2[,c(3,11:15,17:19)]

names(opt_data)


#For Airtel------------

data_airtel <- opt_data[opt_data$X1==1,]

head(data_airtel)


#To select no. of clusters using Scree Plot (elbow method)

wss_airtel <- (nrow(data_airtel)-1)*sum(apply(data_airtel,2,var))

for (i in 2:20) wss_airtel[i] <- sum(kmeans(data_airtel, centers=i)$withinss)

plot(1:20, wss_airtel, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")


#for clustering

kc_airtel <- kmeans(data_airtel, 4)

kc_airtel


#For plotting cluster

fviz_cluster(kc_airtel, data = scale(data_airtel[,2:9]), geom = c("point"),
```

```
        ellipse.type = "euclid")



#For Jio----------

data_jio <- opt_data[opt_data$X1==2,]

dim(data_jio)


#To select no. of clusters using Scree Plot (elbow method)

wss_jio <- (nrow(data_jio)-1)*sum(apply(data_jio,2,var))

for (i in 2:20) wss_jio[i] <- sum(kmeans(data_jio, centers=i)$withinss)

plot(1:20, wss_jio, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")


#for clustering

kc_jio <- kmeans(data_jio, 3)

kc_jio


#For plotting cluster

fviz_cluster(kc_jio, data = scale(data_jio[,2:9]), geom = c("point"),

        ellipse.type = "euclid")



#For Vodafone(VI)----------
```

```
data_voda <- opt_data[opt_data$X1==3,]
dim(data_voda)

#To select no. of clusters using Scree Plot (elbow method)
wss_voda <- (nrow(data_voda)-1)*sum(apply(data_voda,2,var))
for (i in 2:20) wss_voda[i] <- sum(kmeans(data_voda, centers=i)$withinss)
plot(1:20, wss_voda, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")

#for clustering
kc_voda <- kmeans(data_voda, 3)
kc_voda

#For plotting cluster
fviz_cluster(kc_voda, data = scale(data_voda[,2:9]), geom = c("point"),
             ellipse.type = "euclid")



#For Idea (VI)----------
data_idea <- opt_data[opt_data$X1==4,]
dim(data_idea)

#To select no. of clusters using Scree Plot (elbow method)
wss_idea <- (nrow(data_idea)-1)*sum(apply(data_idea,2,var))
```

```
for (i in 2:20) wss_idea[i] <- sum(kmeans(data_idea, centers=i)$withinss)

plot(1:20, wss_idea, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")


#for clustering

kc_idea <- kmeans(data_idea, 3)

kc_idea


#For plotting cluster

fviz_cluster(kc_idea, data = scale(data_idea[,2:9]), geom = c("point"),

             ellipse.type = "euclid")


#For BSNL/MTNL----------------

data_bm <- opt_data[opt_data$X1==5,]

dim(data_bm)


#To select no. of clusters using Scree Plot (elbow method)

wss_bm <- (nrow(data_bm)-1)*sum(apply(data_bm,2,var))

for (i in 2:7) wss_bm[i] <- sum(kmeans(data_bm, centers=i)$withinss)

plot(1:6, wss_bm, type="b", xlab="Number of cluster", ylab="Within Sum of Squares")


#for clustering

kc_bm <- kmeans(data_bm, 3)

kc_bm
```

```
#For plotting cluster
fviz_cluster(kc_bm, data = scale(data_bm[,2:9]), geom = c("point"),
             ellipse.type = "euclid")



##### Structural Equation Modeling (SEM) #####


#Creating Dataframe
S_data <- Data2[, c(35,11:15,17:23,26:34)]
head(S_data)


install.packages("lavaan")
library(lavaan)
install.packages("semPlot")
library(semPlot)
library(dplyr)


#Building model
S_model <-"ser = ~1*X5CC+ X5CVo+ X5S+ X5I+ X5V+ X5U+X5CVi+ X5IC
           prb = ~1*X6Cd+ X6L+ X6M+ X6CrC
           ccs = ~1*X8Cb+ X8T+ X8CCN+ X8C+ X824h+ X8PS+X8o
           vm  = ~1*X9E+ X9VM
           X10~ser+prb+ccs+vm"
```

```
model1 =sem(S_model, data=S_data)

summary(model1,fit.measures=T)



semPaths(model1, "par", edge.label.cex =1.0, fade = FALSE,rotation = 2,residScale=FALSE,curvature = 4, nCharNodes = 3, nCharEdges = 2,

      sizeLat = 8,  sizeInt = 4,curveAdjacent = '<->',optimizeLatRes = TRUE,color = "light blue")

data<-read.csv(file.choose(),header=T)
#View(Data)
head(data)
dim(data)

colnames(data)
colnames(data) <- c("gender","age","net_used","areas","ser_pre","airtel","idea(vi)","jio",
                    "mtnl_bsnl","vodafone","call_conn","cal_voi","sms_mms","int_brow","vid_str",

"soc_net_app","upl_dow","cla_vid","int_av_cal","cal_dro","lag_aud","late_mes_del",

"cro_con","pan_eff","whi_affe","chatbot","tim_take","avail_cust_car","res_to_com",

"cust_ser_24","pol_sta","off_pro","meets_exp","val_for_mon","recommend","bes_net_are",
                    "switch","2G_bet_3G","upd_imp_qua","pand_dep","sector")
colnames(data)
dim(data)

write.csv(data, 'data.csv', row.names = F)

data <- read.csv('data.csv',header=T)
head(data)
dim(data)


colSums(is.na(data)) # check the missing values
colnames(data)[colSums(is.na(data)) != 0]
a <- colnames(data)[colSums(is.na(data)) != 0]
a
for (i in a){
  data[,i][which(is.na(data[,i]))] = median(data[,i],na.rm=T)

}

colSums(is.na(data)) # check the missing values

data <- data[,c("sector","ser_pre", "meets_exp","val_for_mon")]
```

```
colnames(data)

head(data)

sapply(data,class)
dim(data)

sum((data$ser_pre == 3))
sum((data$ser_pre != 3))

data <- data[(data$ser_pre != 3),]
dim(data)
summary(data$ser_pre)

for (i in c("sector","ser_pre")){
  data[,i] <- factor(data[,i])

}

sapply(data,class)
sapply(data,levels)

test <- kruskal.test(meets_exp ~ ser_pre, data = data)
test

pvalue = test$p.value
alpha = 0.05

pvalue < alpha # if true we reject the null hypothesis

test <- kruskal.test(val_for_mon ~ ser_pre, data = data)
test

pvalue = test$p.value

pvalue < alpha # if true we reject the null hypothesis

colnames(data)

data_MWU <- data[,c("sector",   "ser_pre")]

for (i in c("sector","ser_pre")){
  data[,i] <- as.numeric(data[,i])

}

sapply(data,class)


pvalue = test$p.value
```

```
alpha = 0.05
pvalue < alpha # if true we reject the null hypothesis

#To check reliability of questionnaire
>install.packages("psych")
>library(psych)
>x <- data.frame(X5,X6,X8,X9,X10,X13)
>alpha(X)
```

Secondary Data R-Codes:

For Bharti Airtel

```
#To clear variables from the environment
> rm(list=ls())
#data importing
> years <- as.factor(c("2011-03-31", "2012-03-31", "2013-03-31", "2014-03-31",
                       +                   "2015-03-31", "2016-03-31", "2017-03-31", "2018-03-31",
                       +                   "2019-03-31"))
> Bharti_Airtel <- c(181.47, 188.2, 205.39, 226.02, 251.24, 273.65, 304.19, 325.18,
                     +                 327.81)
#Data Framing
  > data <- data.frame(years, Bharti_Airtel)
  > data
  > class(data$years)
  #Convert the datatype of years variable
  > data$years <- as.Date(years)
  > data$years
```

```
> attach(data)
> y <- ts(data[,2], start = 2011, frequency = 1)
#preliminary analysis
> library(fpp2)
> library(ggplot2)
> library(tseries)
#timeplot
> plot.ts(y, xlab = "Time", ylab = "Airtel subscribers")
> autoplot(y)+ggtitle("Time Plot: Airtel Subscriber(in millions) over the years")+
  +   ylab("Airtel subscribers")
 #ADF Test for stationarity checking
  > adf.test(y)

#Forecast with Arima methods
 > fit_arima <- auto.arima(y, d = 1, stepwise = F, approximation = F, trace = T)
 > print(summary(fit_arima))

 > checkresiduals(fit_arima)
#forecast
> fcst <- forecast(fit_arima, h=5)
> print(summary(fcst))
autoplot(fcst)
#To check the accuracy of the model
>accuracy(fcst)
```

For Jio

```
#Data importing
> yrs <- as.factor(c("2016-03-31", "2017-03-31", "2018-03-31", "2019-03-31"))
> jio <- c(108.68, 186.56, 306.72,  387.52 )
  #Data Framing
> dataj <- data.frame(yrs, jio)
 > dataj
#To check for the class of data
 > class(dataj$yrs)
#Convert the datatype of years variable
  > dataj$yrs <- as.Date(yrs)
  > dataj$yrs
   > class(dataj$yrs)
    #Converting into time series object
    > yj <- ts(dataj[,2], start = 2016, frequency = 1)
    > yj
    #preliminary analysis
    > library(fpp2)
    > library(ggplot2)
    > library(tseries)
    > autoplot(yj)+ggtitle("Time Plot: Jio Subscriber(in millions) over the years")+
      +   ylab("Jio subscribers")
   #fit an ARIMA model
   > fit_arima_j <- auto.arima(yj, d = 1, stepwise = F, approximation = F, trace = T)
```

```
   > fit_arima_
   > print(summary(fit_arima_j))
 > checkresiduals(fit_arima_j)
  #fit a Holt's Linear Model
   > fit_hl <- holt(yj, h=5)
   > print(summary(fit_hl))

#To check for residuals
> checkresiduals(fit_hl)
#forecast
> fcst <- forecast(fit_arima_j, h=5)
> autoplot(fcst)
> print(summary(fcst))
#Forecast through Holt's Model
> fcst_hl <- forecast(fit_hl, h=5)
> autoplot(fcst_hl)
> print(summary(fcst_hl))

#Accuracy for Holt's Model
> accuracy(fcst)

For VI (Vodafone Idea)
#Impoting data
vi <- c(263.19, 273.96, 302.35, 341.61, 373.02, 404.43, 433.91, 394.84, 319.17)
```

```
> ye <- as.factor(c("2011-03-31", "2012-03-31", "2013-03-31", "2014-03-31",
+                   "2015-03-31", "2016-03-31", "2017-03-31", "2018-03-31",
+                   "2019-03-31"))
#Data Framing
 > datavi <- data.frame(ye, vi)
 > datavi
#To check for the class of data
> class(datavi$ye)
#Convert the datatype of years variable
 > datavi$ye <- as.Date(ye)
 > datavi$ye
 > class(datavi$ye)
 #Converting into time series object
 > yvi <- ts(datavi[,2], start = 2011, frequency = 1)
 > yvi

 #preliminary analysis
 > library(fpp2)
 > library(ggplot2)
 > library(tseries)
#timeplot
 > autoplot(yvi)+ggtitle("Time Plot: VI Subscriber(in millions) over the years")+
   +    ylab("VI subscribers")
```

```
#ADF Test for stationarity
> adf.test(yvi)
#To consider differenced variable
> dye <- diff(yvi)
> autoplot(dye)+ggtitle("Time Plot:Change in real Vi Subscriber(in millions) over the years")+
+ ylab("VI subscribers")
> adf.test(dye)
#Forecast with different methods
#fit an ARIMA model
> fit_arima_vi <- auto.arima(yvi, d = 1, stepwise = F, approximation = F, trace = T)

> print(summary(fit_arima_vi))

> checkresiduals(fit_arima_vi)
#forecast
> fcst_vi <- forecast(fit_arima_vi, h=5)
> autoplot(fcst_vi) + xlab("Years 2011-12 to 2019-20") + ylab("VI subscribers in millions")
> + guides(colour=guide_legend(title = "Forecast"))
#To check accuracy of the model
Accuracy(fcst_vi)
```

For MTNL

```
> mtnl <- c(5.83, 5, 3.37, 3.51, 3.56, 3.63, 3.56, 3.45, 3.36 )
> years <- as.factor(c("2011-03-31", "2012-03-31", "2013-03-31", "2014-03-31",
+ "2015-03-31", "2016-03-31", "2017-03-31", "2018-03-31",
```

```
                    +                    "2019-03-31"))
#Data Framing
> datam <- data.frame(years, mtnl)
> datam
#To check for the class of data
> class(datam$years)
#Convert the datatype of years variable
 > datam$years <- as.Date(years)
 > datam$years
  > class(datam$years)
  #Converting into time series object
  > ym <- ts(datam[,2], start = 2011, frequency = 1)
  > ym
  #preliminary analysis
  > library(fpp2)
  > library(ggplot2)
  > library(tseries)
#timeplot
  > plot.ts(ym, xlab = "Time", ylab = "MTNL subscribers")

  > autoplot(ym)+ggtitle("Time Plot: MTNL Subscriber(in millions) over the years")+
    +   ylab("MTNL subscribers")

  #ADF Test for stationarity
```

```
 > adf.test(ym)


 #fit an ARIMA model
 > fit_arima_m <- auto.arima(ym, d = 1, stepwise = F, approximation = F, trace = T)
  > print(summary(fit_arima_m))
> checkresiduals(fit_arima_m)


 #To forecast through ARIMA model
  > fcst_m <- forecast(fit_arima_m, h=3)
  > autoplot(fcst_m)


  > print(summary(fcst_m))


#Accuracy
 > accuracy(fcst_m)
 #fit an Holt's model
  > fit_holt_m <- holt(ym, h= 5)
  > print(summary(fit_holt_m))
Call:
    holt(y = ym, h = 5)
#Forecast through Holts Model.
> fcst_m <- forecast(fit_holt_m, h=3)
> autoplot(forecast(fit_holt_m, h=3))
```

```
> print(summary(forecast(fit_holt_m, h=3)))
```

#Accuracy of Holt's Model

```
> accuracy(fit_holt_m, h=3)
```

For BSNL

#importing data

```
> bsnl <- c(98.51, 101.21, 94.65, 77.22, 86.35, 100.99, 111.68, 115.74, 119.87)
> years <- as.factor(c("2011-03-31", "2012-03-31", "2013-03-31", "2014-03-31",
                  +                  "2015-03-31", "2016-03-31", "2017-03-31", "2018-03-31",
                  +                  "2019-03-31"))
> datab <- data.frame(years, bsnl)
> datab
```

#To check for the class of the dataset

```
> class(datab$years)
> datab$years <- as.Date(years)
> datab$years
> class(datab$years)
> attach(datab)
```

#Converting into time series object

```
  > yb <- ts(datab[,2], start = 2011, frequency = 1)
> yb
```

#preliminary analysis

#timeplot

```
> autoplot(yb)+ggtitle("Time Plot: BSNL Subscriber(in millions) over the years")+
  +   ylab("BSNL subscribers")
```

```
# To check for the stationarity through ADF Test

> adf.test(yb)

#fit an ARIMA model

> fit_arima_b <- auto.arima(yb, d = 1, stepwise = F, approximation = F, trace = T)

> fit_arima_b

> print(summary(fit_arima_b))


> checkresiduals(fit_arima_b)


#forecast through ARIMA model

> fcst_b <- forecast(fit_arima_b, h=3)

#forecast

> fcst_b <- forecast(fit_arima_b, h=5)

> fcst_b


> autoplot(fcst_b)


#Accuracy

> accuracy(fcst_b)

#To fit  Holt's model

> holt(yb)

Holt's method

Call:
```

holt(y = yb)

> checkresiduals(holt(yb))

To forecast through Holt's Model

> forecast(holt(yb), h=5)

> autoplot(holt(yb))

#To check the accuracy of the Holt's Model

> accuracy(forecast(holt(yb)))

QUESTIONNAIRE

Gender-

Female	Male	Prefer not to say	others

Age -

18 - 25	26 - 35	36 - 45	46 – 55	55 and above

Q1. Which network are you using now? (if dual sim user can choose more than 1)

Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL

Q2. From which part of country you belong to?

Rural	Urban	Semi-urban	Other

Q3. What kind of service you prefer?

Prepaid	Post-paid	Both

Q4. How long are you a customer/subscriber of your network service provider?
(Please scroll left for more options)

	Not a user	< 1 year	1 - 3 years	3 - 5 years	5 - 10 years	> 10 years
Airtel						
Jio						
Vodafone (VI)						
Idea (VI)						
MTNL/BSNL						

Q5. Rate the quality of the service provided by your network service provider.
(If you are a dual sim user then fill the details of the most preferred network).
(Rate it on a scale of 1 to 5) 1 - very Poor, 2 - poor, 3 - average, 4 - good, 5 – Excellent

	1	2	3	4	5
Call connection					
Clarity of voice					
SMS/MSS					
Internet Browsing (google chrome, etc)					
Video Streaming (youtube, Instagram, etc)					
Social networking apps (Instagram, Fb, whatsapp, etc)					
Uploading & downloading data/ file					
Clarity of video					
Internet Audio/video Calling (social media calls)					

Q6. Did you ever receive following problems?

	Agree	Neutral	Disagree
Call drop (suddenly disconnect)			
lag in audio/video			
late message delivered			
cross connection			

Q7. Did pandemic affect your network service?

Yes	No

If yes, choose from the below option which section was affected the most? (can select more than 1)

call connection	internet browsing	video streaming (YouTube, Instagram, etc)	uploading/downloading data/file	lag in video/audio	Buffering
1	2	3	4	5	6

Q8. Rate the service provided by the customer care of your network?

	Never used	1	2	3	4	5
Chatbot						
Time taken to fix issue						
Availability of customer care no.						
Response to the complaints						
24 hours customer service						
Politeness of Staff						
Offers provided by the service provider						

Q9. Rate your cost and offers given by the service provider

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Service meets my expectations					
Value for the money					

Q10. How likely is it that you would recommend your current network to your friends and family?

Highly recommend	Moderately recommend	Neutral	Least recommend	Will not recommend at all

Q11. Which network works best in your area?

Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL

Q12. If you want to switch then which other SIM would you like to prefer from the networks listed below?

Airtel	Jio	Vodafone (VI)	Idea (VI)	MTNL/BSNL	Other	Will not change

Q13. Although companies are saying 3G/4G network are better but customers were experiencing better network quality when only 2G has been introduced. Do you agree?

Agree	Maybe	Disagree

Q14. The update made by the telecom networks has improved the network quality?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

Q15. Do you think during pandemic we were dependent on networks for communication and other uses?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

Q16. Which sector provide better network Government or Private?

Government (BSNL/MTNL)	Private (Airtel, Jio, Vodafone, Idea)

REFERENCE:

Books:

[1] Kees Montfort, Johan Oud and Albert Satorra (2002), Mathematical modelling: Recent Development on SEM, Springer Science + Business Media, B.V publisher

[2] Andy Field, Jeremy Miles and Zoe Field (2012), Discovering Statistics Using R, SAGE Publications Ltd.

[3] Naresh K. Malhotra and David F. Birks (2007), Third Edition, Marketing Research: An Applied Approach, published by Prentice Hall, a Pearson Education Company.

[4] Peter J. Brockwell and Richard A. Davis (August 2001), Introduction to Time Series and Forecasting, second edition, Library of Congress-in-Publication.

Websites:

[5] Deepika Singh (July 12, 2019), Time Series Forecasting Using R.
https://www.pluralsight.com/guides/time-series-forecasting-using-r

[6] Tavish Shrivastava (December 16, 2015), A Complete Tutorial on Time Series Modelling in R.
https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/

[7] Advanced Through Clarity, Time Series Analysis

https://www.statisticssolutions.com/time-series-analysis/

[8] Secondary Data Reference: Telecom Regulatory Authority of India

https://www.trai.gov.in/release-publication/reports/telecom-subscriptions-reports

[9] Data Novia, Partitional clustering in R: The Essentials.
https://www.datanovia.com/en/lessons/k-means-clustering-in-r-algorith-and-practical-examples/

[10]UC Business Analytics R Programming Guide: K-means Cluster Analysis

https://uc-r.github.io/kmeans_clustering

[11] Advanced Through Clarity, Assumptions of Logistic Regression

http://www.statisticssolutions.com/wp-content/uploads/wp-post-to-pdf-enhanced-cache/1/assumptions-of-logistic-regression.pdf?__cf_chl_jschl_tk__=0dc7a54b62924edeb94ffafc0137e01f282f09e4-1621510098-0-AZm6YMylF-

[12] Data Camp, Logistic Regression in R Tutorial

https://www.datacamp.com/community/tutorials/logistic-regression-R

[13] Analytics Vidhya, Simple Guide to Logistic Regression in R and Python

https://www.analyticsvidhya.com/blog/2015/11/beginners-guide-on-logistic-regression-in-r/

[14] UC Business Analytics R Programming Guide, Logistic Regression

https://uc-r.github.io/logistic_regression

https://www.edureka.co/blog/understanding-logistic-regression-in-r/

[15]Quick R by Datacamp: Nonparametric Tests Of Group Differences.
https://www.statmethods.net/stats/nonparametric.html

[16] Top R tutorial site: Kruskal Wallis Test

http://www.r-tutor.com/elementary-statistics/non-parametric-methods/kruskal-wallis-test

[17] Mahesh Uppval (Nov 05 2020), From the Economic Times, teletalks by telecom gurus, Need to make Telecom industry competitive

https://telecom.economictimes.indiatimes.com/tele-talk/need-to-make-indian-telecom-industry-competitive/4613

Research Papers:

[18] James B. Schreiber, Frances K. Stage, Jamie King, Amaury Nora and Elizabeth A. Barlow: Reporting Structural Equation Modelling And Confirmatory Factor Analysis Results

[19] Shagufta Abrahim, Bilal Ahmed Mir, Hayato Suhara, Fatin Amirah, Mohamed and Masahiro Sato(2019): Structural equation Modeling and Confirmatory Factor Analysis of Social Media use and Education.

[20] Sehrat Kocakaya and Ferit Kocakaya(2014): A Structural Equation Modeling on Factors of How Experienced Teachers Affect the Students' Science and Mathematics Achievements.

[21] Paul A. Gore Jr. and Rebecca Weston (September 2006): A Brief Guide to Structural Equation Modeling.

[22] Diana Suhr, Ph.D., University of Nothern Colorado: The Basics of Structural Equation Modeling.

[23] Telecom Subscriptions Reports, English, Telecom Regulatory Authority of India

URL: https://www.trai.gov.in/release-publication/reports/telecom-subscriptions-reports

[24]-Reporting_SEM_and_CFA__Schreiber__Stage__King__Nora__Barlow_.pdf

https://www.theanalysisfactor.com/r-tutorial-glm1/