DESCRIPTION:

Comcast is an American global telecommunication company. The firm has been providing terrible customer service. They continue to fall short despite repeated promises to improve. Only last month (October 2016) the authority fined them a $2.3 million, after receiving over 1000 consumer complaints.

The existing database will serve as a repository of public customer complaints filed against Comcast.

It will help to pin down what is wrong with Comcast's customer service.

Data Dictionary:

1. Ticket #: Ticket number assigned to each complaint.
2. Customer Complaint: Description of complaint.
3. Date: Date of complaint.
4. Time: Time of complaint
5. Received Via: Mode of communication of the complaint
6. City: Customer city
7. State: Customer state
8. Zipcode: Customer zip
9. Status: Status of complaint
10. Filing on behalf of someone

Analysis Task:

- Import data into R environment.
- Provide the trend chart for the number of complaints at monthly and daily granularity levels.
- Provide a table with the frequency of complaint types.

Which complaint types are maximum i.e., around internet, network issues, or across any other domains.

- Create a new categorical variable with value as Open and Closed. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed.
- Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on:

Which state has the maximum complaints

Which state has the highest percentage of unresolved complaints

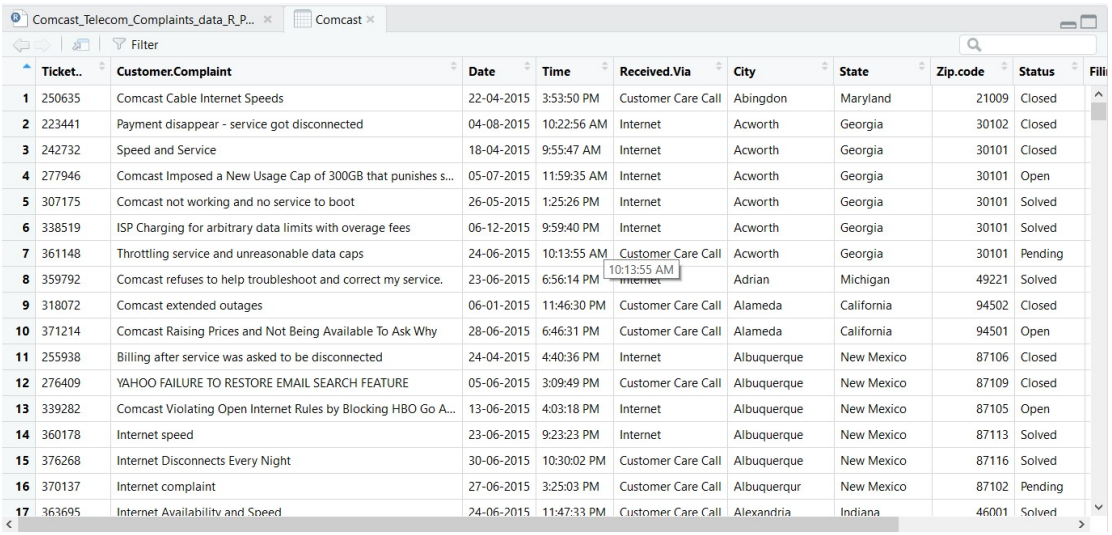
- Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.

The analysis results to be provided with insights wherever applicable.

- Q.1> Import data into R environment.

```
### Setting the directory
setwd("C:/Users/Abhishek/Desktop/Project/")
getwd()  #checking the path of directory

#reading csv file
Comcast = read.csv("Comcast Telecom Complaints data.csv")
View(Comcast)
```



Comcast_Telecom_Complaints_data_R_P... | Comcast | Filter

	Ticket..	Customer.Complaint	Date	Time	Received.Via	City	State	Zip.code	Status	Fili
1	250635	Comcast Cable Internet Speeds	22-04-2015	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	
2	223441	Payment disappear - service got disconnected	04-08-2015	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	
3	242732	Speed and Service	18-04-2015	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	
4	277946	Comcast Imposed a New Usage Cap of 300GB that punishes s...	05-07-2015	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	
5	307175	Comcast not working and no service to boot	26-05-2015	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	
6	338519	ISP Charging for arbitrary data limits with overage fees	06-12-2015	9:59:40 PM	Internet	Acworth	Georgia	30101	Solved	
7	361148	Throttling service and unreasonable data caps	24-06-2015	10:13:55 AM	Customer Care Call	Acworth	Georgia	30101	Pending	
8	359792	Comcast refuses to help troubleshoot and correct my service.	23-06-2015	6:56:14 PM	Internet	Adrian	Michigan	49221	Solved	
9	318072	Comcast extended outages	06-01-2015	11:46:30 PM	Customer Care Call	Alameda	California	94502	Closed	
10	371214	Comcast Raising Prices and Not Being Available To Ask Why	28-06-2015	6:46:31 PM	Customer Care Call	Alameda	California	94501	Open	
11	255938	Billing after service was asked to be disconnected	24-04-2015	4:40:36 PM	Internet	Albuquerque	New Mexico	87106	Closed	
12	276409	YAHOO FAILURE TO RESTORE EMAIL SEARCH FEATURE	05-06-2015	3:09:49 PM	Customer Care Call	Albuquerque	New Mexico	87109	Closed	
13	339282	Comcast Violating Open Internet Rules by Blocking HBO Go A...	13-06-2015	4:03:18 PM	Internet	Albuquerque	New Mexico	87105	Open	
14	360178	Internet speed	23-06-2015	9:23:23 PM	Internet	Albuquerque	New Mexico	87113	Solved	
15	376268	Internet Disconnects Every Night	30-06-2015	10:30:02 PM	Customer Care Call	Albuquerque	New Mexico	87116	Solved	
16	370137	Internet complaint	27-06-2015	3:25:03 PM	Customer Care Call	Albuquerqur	New Mexico	87102	Pending	
17	363695	Internet Availability and Speed	24-06-2015	11:47:33 PM	Customer Care Call	Alexandria	Indiana	46001	Solved	

Fig: Result of View(Comcast)

- Q.2> Provide the trend chart for the number of complaints at monthly and daily granularity levels.

i) Number of complaints on Daily granularity level.

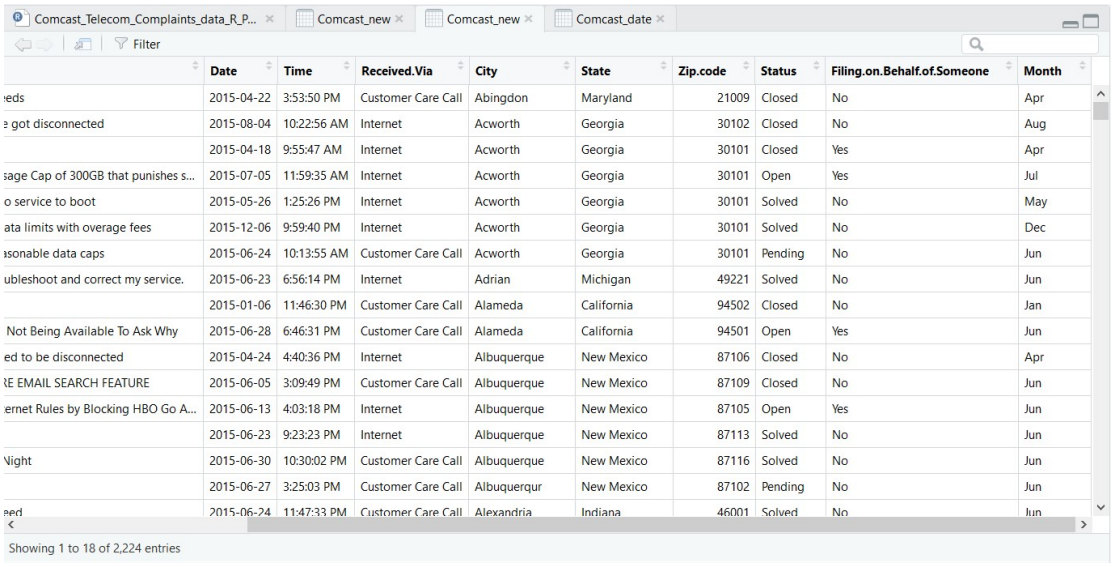
```
#Loading The Date Into Single Format
#Use Lubridate Library to Format the Date Column
library(lubridate)
li=parse_date_time(x = Comcast$Date,
                   orders = c("d m y", "d B Y", "m/d/y"),
                   locale = "eng")
Comcast_new = Comcast
Comcast_new$Date = li
#Dates Loaded In the Same Format in the new Dataframe
str(Comcast_new$Date)
View(Comcast_new)

#Extracting Month Column and Converting to The labels.
Comcast_new$Month = format(as.Date(Comcast_new$Date), "%m")
Comcast_new$Month= month.abb[as.integer(Comcast_new$Month)]
View(Comcast_new)

#importing dplyr
library(dplyr)

#grouping data with date and calculating Frequency by summarise
Comcast_date = Comcast_new %>% group_by(Date) %>% summarise(frequency = n())
View(Comcast_date)

#ploting graph date against Number of Complaints
plot(Comcast_date,type='l', main = "number of complaints daily granularity levels", xlab =
"Date", ylab = "Number of complaints",col="#ff0000")
```



Comcast_Telecom_Complaints_data_R_P... | Comcast_new | Comcast_new | Comcast_date

Filter

	Date	Time	Received.Via	City	State	Zip.code	Status	Filing.on.Behalf.of.Someone	Month
›eds	2015-04-22	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	Apr
› got disconnected	2015-08-04	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	Aug
	2015-04-18	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	Apr
›age Cap of 300GB that punishes s...	2015-07-05	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	Jul
o service to boot	2015-05-26	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May
ata limits with overage fees	2015-12-06	9:59:40 PM	Internet	Acworth	Georgia	30101	Solved	No	Dec
›sonable data caps	2015-06-24	10:13:55 AM	Customer Care Call	Acworth	Georgia	30101	Pending	No	Jun
ubleshoot and correct my service.	2015-06-23	6:56:14 PM	Internet	Adrian	Michigan	49221	Solved	No	Jun
	2015-01-06	11:46:30 PM	Customer Care Call	Alameda	California	94502	Closed	No	Jan
Not Being Available To Ask Why	2015-06-28	6:46:31 PM	Customer Care Call	Alameda	California	94501	Open	Yes	Jun
ed to be disconnected	2015-04-24	4:40:36 PM	Internet	Albuquerque	New Mexico	87106	Closed	No	Apr
E EMAIL SEARCH FEATURE	2015-06-05	3:09:49 PM	Customer Care Call	Albuquerque	New Mexico	87109	Closed	No	Jun
ernet Rules by Blocking HBO Go A...	2015-06-13	4:03:18 PM	Internet	Albuquerque	New Mexico	87105	Open	Yes	Jun
	2015-06-23	9:23:23 PM	Internet	Albuquerque	New Mexico	87113	Solved	No	Jun
Night	2015-06-30	10:30:02 PM	Customer Care Call	Albuquerque	New Mexico	87116	Solved	No	Jun
	2015-06-27	3:25:03 PM	Customer Care Call	Albuquerqur	New Mexico	87102	Pending	No	Jun
eed	2015-06-24	11:47:33 PM	Customer Care Call	Alexandria	Indiana	46001	Solved	No	Jun

Showing 1 to 18 of 2,224 entries

Fig: View(Comcast_new)

Comcast_Telecom_Complaints_data_R_P... | Comcast_new | Comcast_new | Comcast_date

Filter

	Date	frequency
1	2015-01-04	18
2	2015-01-05	12
3	2015-01-06	25
4	2015-02-04	27
5	2015-02-05	7
6	2015-02-06	25
7	2015-03-04	15
8	2015-03-05	5
9	2015-03-06	25
10	2015-04-04	12
11	2015-04-05	12
12	2015-04-06	12
13	2015-04-13	24
14	2015-04-14	23
15	2015-04-15	12
16	2015-04-16	20
17	2015-04-17	21

Showing 1 to 18 of 91 entries

Fig: View(Comcast_date)

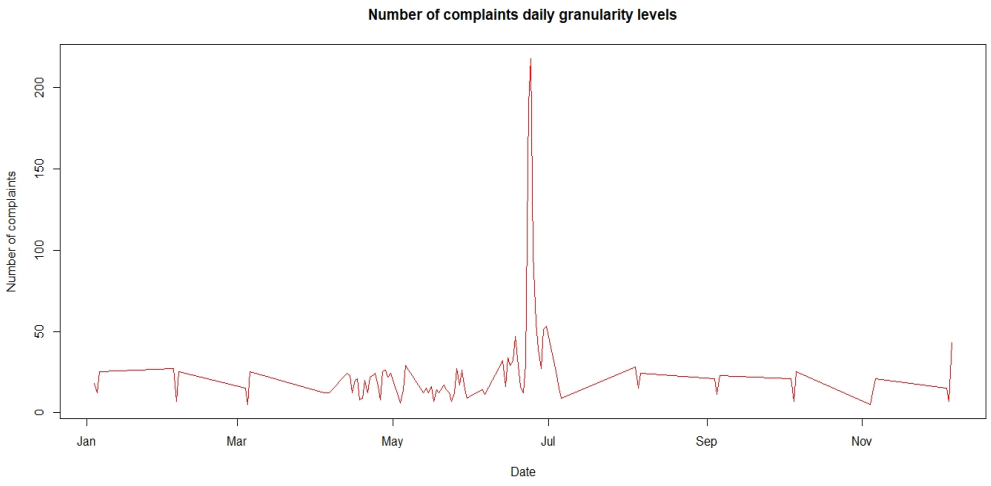
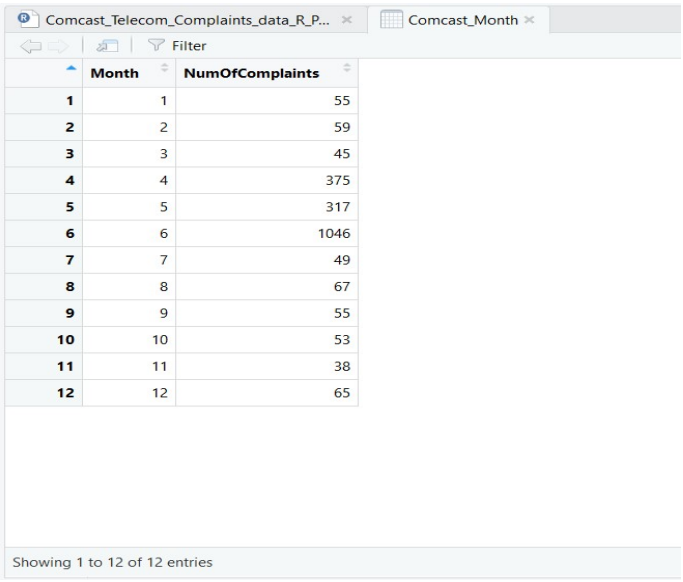


Fig: Number of complaints on daily granularity levels

ii) Number of complaints on Monthly granularity level.

```
#grouping data with respect to months and summarizing it with Frequency of Number of complaint
Comcast_Month = Comcast_new %>% group_by(Month =as.integer(month(Date))) %>% summarize(NumOfComplaints=n())
View(Comcast_Month)

#ploting Data
plot(Comcast_Month,type="o", xlab = "Months", ylab = "Number of Complaints", main = "Number of Complaint on Monthly basis",col="#ff0000")
```



	Month	NumOfComplaints
1	1	55
2	2	59
3	3	45
4	4	375
5	5	317
6	6	1046
7	7	49
8	8	67
9	9	55
10	10	53
11	11	38
12	12	65

Showing 1 to 12 of 12 entries

Fig: View(Comcast_Month)

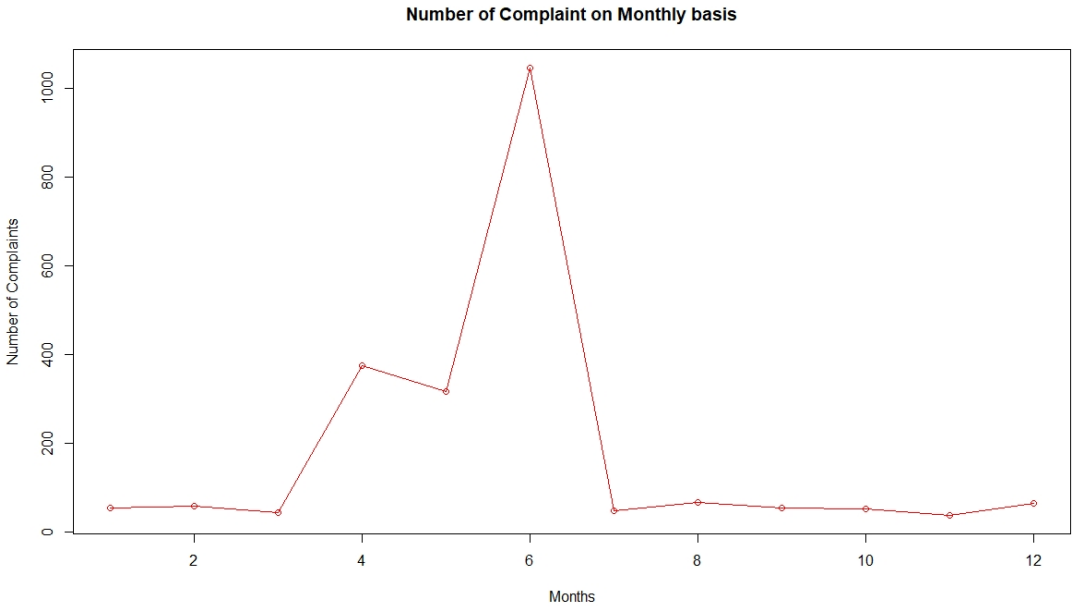


Fig: Number of complaints on monthly granularity levels

- Q.3> Provide a table with the frequency of complaint types.

```
library(dplyr)

#Converting All String Values to Lower, so as to Eliminate Duplication of Any Complaint
Comcast_Complaint_type = Comcast_new %% mutate(Customer.Complaint =
tolower(Customer.Complaint))

#forming table to calculate frequency of complain
CustTable = table(Comcast_Complaint_type$Customer.Complaint)

#storing table in dataframe
CustTable = data.frame(CustTable)

#renaming columns using filter
filtered  = CustTable %% rename(Custom_Complaint = Var1, Frequence = Freq )

#arranging data in Desc order of frequency
Final = filtered %% arrange(desc(Frequence))


#head shows the top 10 results
Most_frequent_complains = head(Final,10)
View(Most_frequent_complains)


attach(Most_frequent_complains)
#plotting barplot
barplot(Frequence,names.arg = Custom_Complaint,las=1, lwd=1,xlab = "Types of
Complaints" ,cex.names=0.7,ylab = "Frequncy of Complaints", main = "Frequency of
complaint
types",col=c("#cc2900","#e62e00","#ff3300","#ff471a","#ff5c33","#ff5c33","#ff8
566","#ff9980","#ffad99","#ffc2b3"))
```

Comcast_Telecom_Complaints_data_R_P... | Most_frequent_complains

Filter

	Custom_Complaint	Frequence
1	comcast	102
2	comcast data cap	30
3	comcast internet	29
4	comcast data caps	21
5	comcast billing	18
6	comcast service	15
7	internet speed	15
8	data caps	13
9	unfair billing practices	13
10	data cap	12

Showing 1 to 10 of 10 entries

Fig: View(Most_frequent_complains)

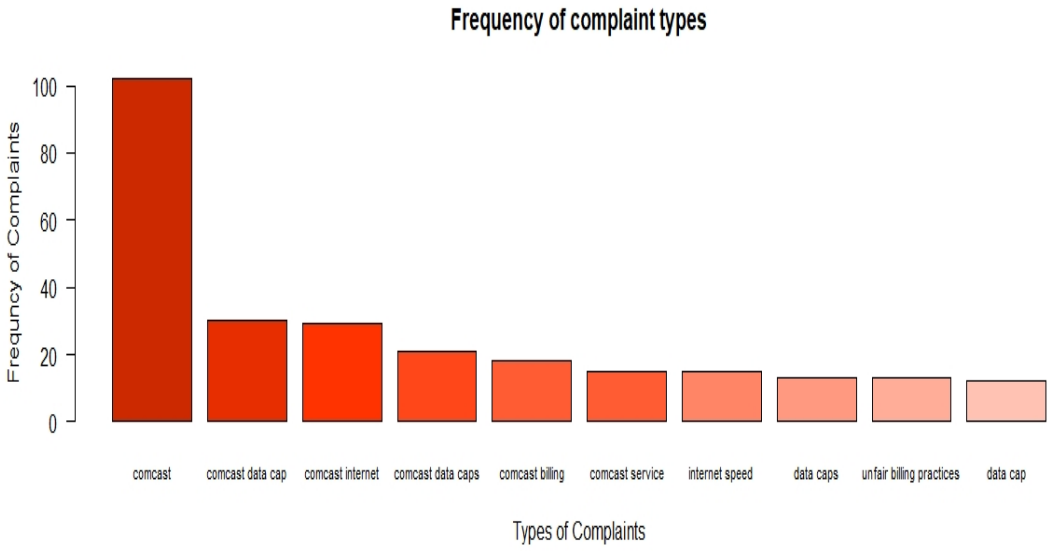
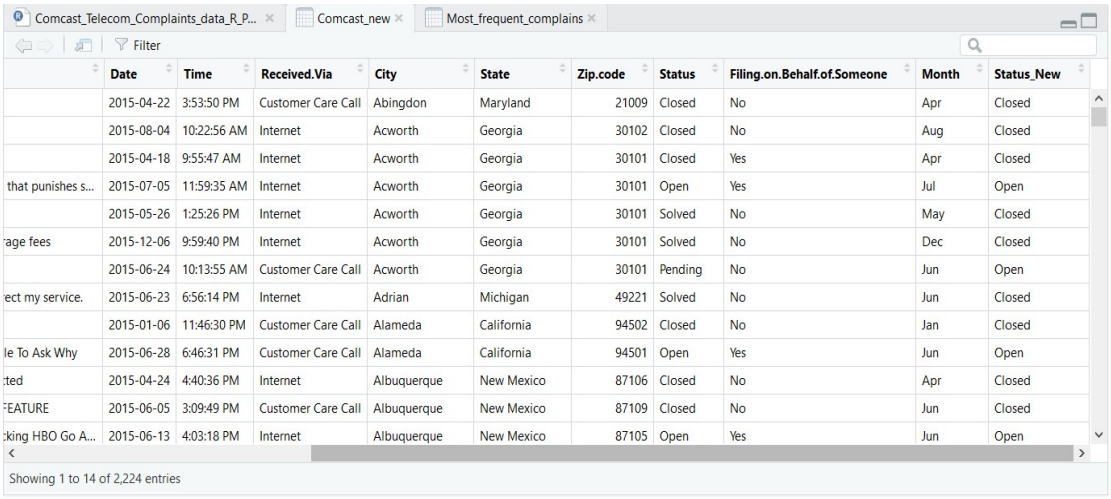


Fig: Frequency of Complaint types

Insights:	Comcast has frequency of 102, which is highest among all the type of complaints

Q.4> Create a new categorical variable with value as Open and Closed. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed.

```
library(plyr)
#adding a new variable Status_new & revaluing some values
Comcast_new$Status_New<-revalue(Comcast_new$Status, c(Pending = "Open", Solved = "Closed"))
View(Comcast_new)
```



Comcast_Telecom_Complaints_data_R_P... | Comcast_new | Most_frequent_complains

Filter

	Date	Time	Received.Via	City	State	Zip.code	Status	Filing.on.Behalf.of.Someone	Month	Status_New
	2015-04-22	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	Apr	Closed
	2015-08-04	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	Aug	Closed
	2015-04-18	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	Apr	Closed
that punishes s...	2015-07-05	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	Jul	Open
	2015-05-26	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May	Closed
rage fees	2015-12-06	9:59:40 PM	Internet	Acworth	Georgia	30101	Solved	No	Dec	Closed
	2015-06-24	10:13:55 AM	Customer Care Call	Acworth	Georgia	30101	Pending	No	Jun	Open
ect my service.	2015-06-23	6:56:14 PM	Internet	Adrian	Michigan	49221	Solved	No	Jun	Closed
	2015-01-06	11:46:30 PM	Customer Care Call	Alameda	California	94502	Closed	No	Jan	Closed
le To Ask Why	2015-06-28	6:46:31 PM	Customer Care Call	Alameda	California	94501	Open	Yes	Jun	Open
ted	2015-04-24	4:40:36 PM	Internet	Albuquerque	New Mexico	87106	Closed	No	Apr	Closed
EATURE	2015-06-05	3:09:49 PM	Customer Care Call	Albuquerque	New Mexico	87109	Closed	No	Jun	Closed
king HBO Go A...	2015-06-13	4:03:18 PM	Internet	Albuquerque	New Mexico	87105	Open	Yes	Jun	Open

Showing 1 to 14 of 2,224 entries

Fig: View(Comcast_new) (added new column Status_New)

Q.5> Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on:

- Which state has the maximum complaints.

```
#creating a new table State_Status
State_Status <- table(Comcast_new$State,Comcast_new$Status_New)

#creating a new table Total and combining with State_Status
State_Status <- cbind(State_Status, Total = rowSums(State_Status))

#converting into datafrawe
State_Status = as.data.frame(State_Status)

#creating a new Percentage and combining with State_Status
State_Status = cbind(State_Status, Percentage = (State_Status$Open/State_Status$Total)*100)
View(State_Status)

#creating a table freq0
freq = table(Comcast_new$Status_New,Comcast_new$State)
View(freq)

#legend(x=3.4,y=1.5,legend = c("Open","Closed"),pch=1.2)
#creating a barplot
barplot(freq,las=2, lwd=1, cex.names=0.5,col=c("#00e6e6","#ff8080"), xlab = "States",ylab = "Total Number of Complaints",main = "State-wise Status of complaint")
```

Comcast_Telecom_Complaints_data_R_P... * | State_Status

Filter

	Closed	Open	Total	Percentage
Alabama	17	9	26	34.615385
Arizona	14	6	20	30.000000
Arkansas	6	0	6	0.000000
California	159	61	220	27.727273
Colorado	58	22	80	27.500000
Connecticut	9	3	12	25.000000
Delaware	8	4	12	33.333333
District of Columbia	1	0	1	0.000000
District Of Columbia	14	2	16	12.500000
Florida	201	39	240	16.250000
Georgia	208	80	288	27.777778
Illinois	135	29	164	17.682927
Indiana	50	9	59	15.254237
Iowa	1	0	1	0.000000
Kansas	1	1	2	50.000000
Kentucky	4	3	7	42.857143
Louisiana	12	1	13	7.692308
Maine	2	2	5	40.000000

Showing 1 to 18 of 43 entries

Fig: View(State_Status)

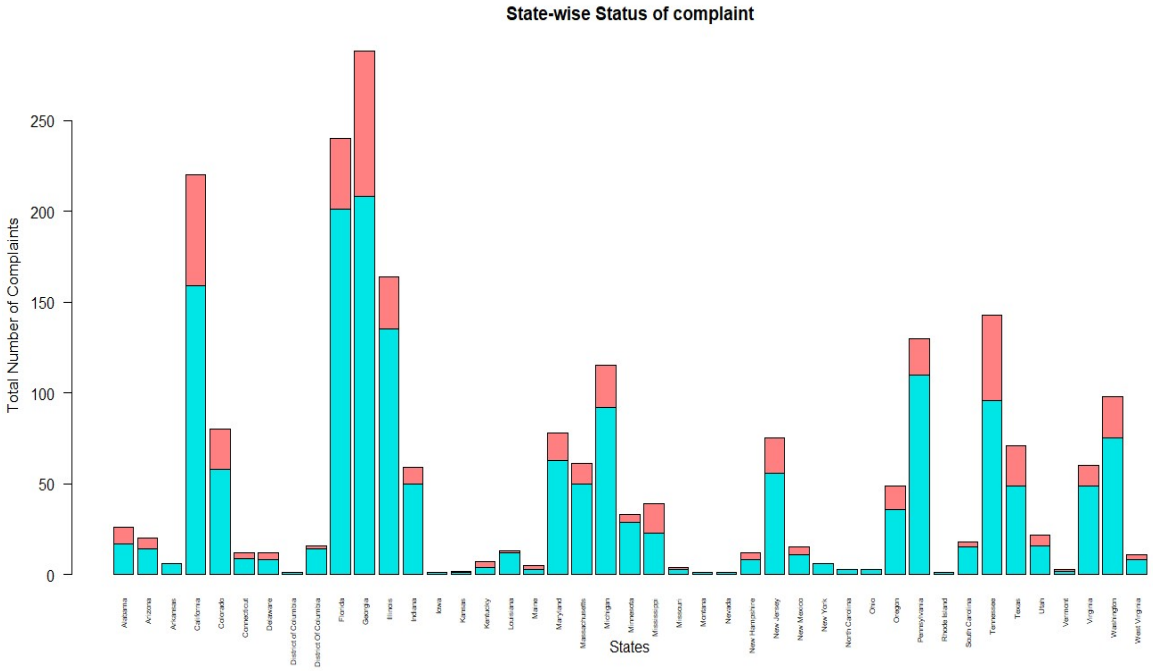


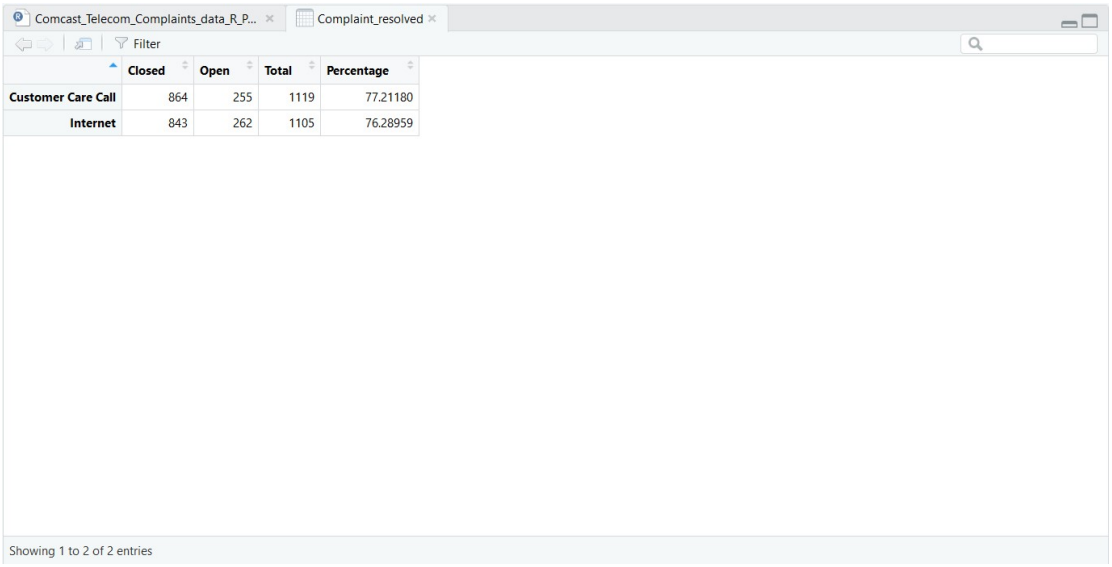
Fig: State-wise Status of complaint

- Closed Complaints
- Open Complaints

Insights:	1. Georgia (288) has the maximum number of complaints.
	2. Kansas (50%) has highest percentage of unresolved complaints.

Q.6> Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.

```
Complaint_resolved = table(Comcast_new$Received.Via,Comcast_new$Status_New)
Complaint_resolved = cbind(Complaint_resolved, Total = rowSums(Complaint_resolved))
Complaint_resolved = as.data.frame(Complaint_resolved)
Complaint_resolved = cbind(Complaint_resolved,Percentage =
(Complaint_resolved$Closed/Complaint_resolved$Total)*100)
View(Complaint_resolved)
```



	Closed	Open	Total	Percentage
Customer Care Call	864	255	1119	77.21180
Internet	843	262	1105	76.28959

Fig: View(Complaint_resolved)

Insights:	1.) 77.21% complaints are resolved via Customer Care Call.
	2.) 76.28% complaints are resolved via Internet.