

# **Comcast Telecom Consumer Complaints.**

Analysis by  
Meet Hariyani

# Analysis Tasks.

- 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.

# 1) Import data into R environment.

```
## importing Libraries
library(dplyr)
library(ggplot2)
library(lubridate)

#importing data
comcastData<- read.csv(file.choose())
View(comcastData)
```

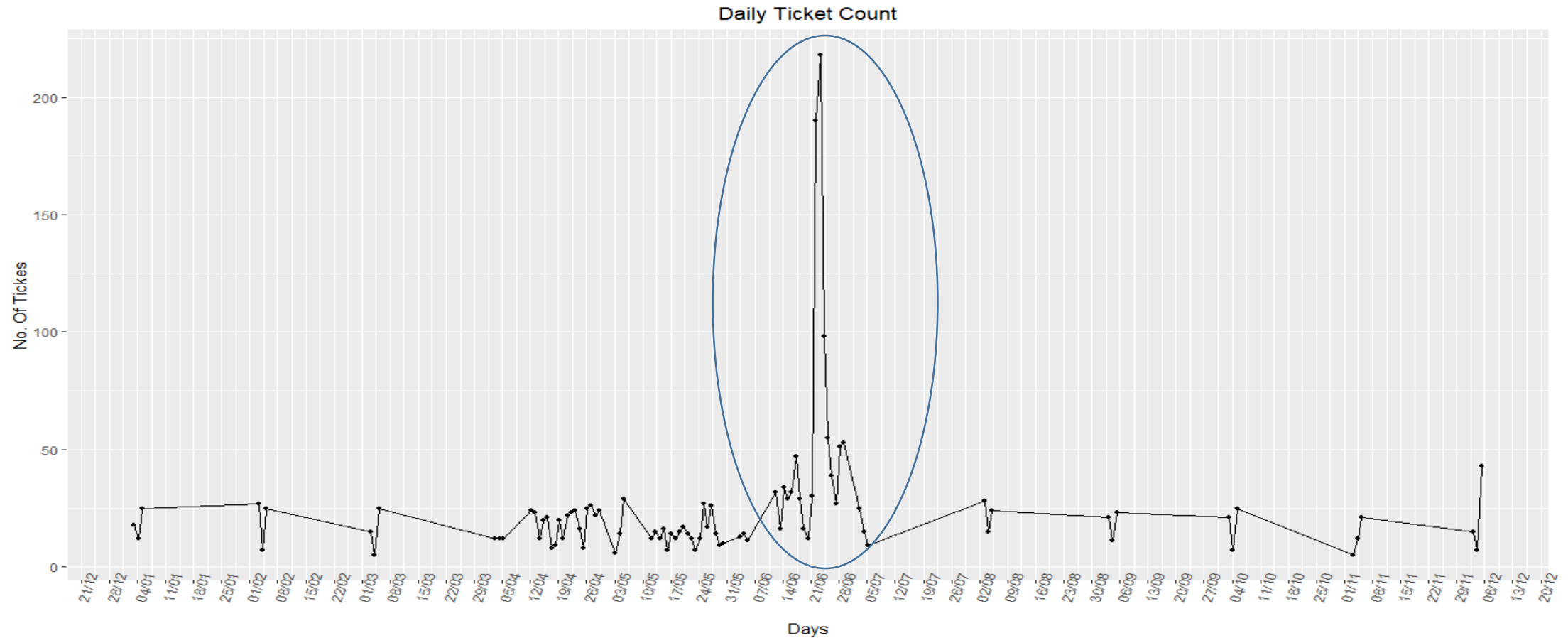
- Comcast database loaded on R studio.
- Libraries (dplyr,ggplot2and lubridate) are used in analysis tasks.

## 2) Provide the trend chart for the number of complaints at monthly and daily granularity levels.

```
#Daily complaint trend
dailyCompl<- comcastData %>% group_by(Date) %>%
summarise(NumberOfComplaints=n())
ggplot(data= dailyCompl,aes(as.POSIXct(Date),NumberOfComplaints))+
  geom_line()+
  geom_point(size= 1)+
  scale_x_datetime(breaks = "1 week", date_labels = "%d/%m")+
  labs(title = "Daily Ticket Count", x="Days",y="No. Of Tickes")+
  theme(axis.text.x = element_text(angle = 75), plot.title =
element_text(hjust = 0.5))

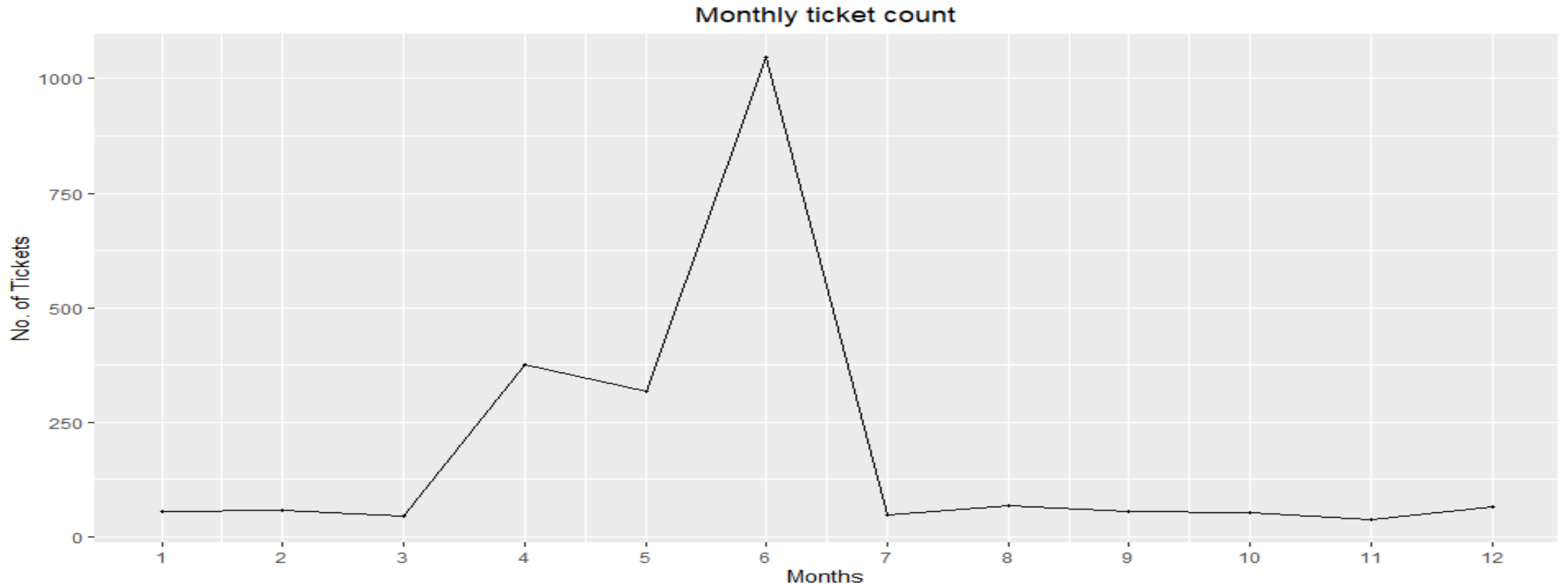
#monthly complaint trend
comcastData$Months<- months(comcastData$Date)
monthlyCompl<- comcastData %>% group_by(Months=as.integer(month(Date))) %>%
  summarise(NumofComplaints=n()) %>% arrange(desc(NumofComplaints))
View(monthlyCompl)
ggplot(data = monthlyCompl,aes(Months,NumofComplaints))+
  geom_line()+
  geom_point(size=0.8)+
  scale_x_continuous(breaks = monthlyCompl$Months)+
  labs(x="Months",y="No. of Tickets",title = "Monthly ticket count")+
  theme(plot.title = element_text(hjust = 0.5))
```

# Daily complaint trend.



*Insight:-* Between 14/06 to 28/06 high number of complaint tickets are raised this are noticed through spike on above chart.

# Monthly complaint trend.



***Insight:-*** On month of June high number of complaint tickets rised as noticed through above graph.

### 3) Table with the frequency of complaint types.

```
#complaint type processing
network_tickets<- contains(comcastData$CustomerComplaint,match =
'network',ignore.case = T)
internet_tickets<- contains(comcastData$CustomerComplaint,match =
'internet',ignore.case = T)
billing_tickets<- contains(comcastData$CustomerComplaint,match =
'bill',ignore.case = T)
email_tickets<- contains(comcastData$CustomerComplaint,match =
'email',ignore.case = T)
charges_tickets<- contains(comcastData$CustomerComplaint,match =
'charge',ignore.case = T)
comcastData$complaintType[internet_tickets]<-"Internet"
comcastData$complaintType[billing_tickets]<-"Billing"
comcastData$complaintType[email_tickets]<-"Email"
comcastData$complaintType[network_tickets]<-"Network"
comcastData$complaintType[charges_tickets]<-"Charges"
comcastData$complaintType[-
c(internet_tickets,billing_tickets,email_tickets,network_tickets,charges_tickets)]<-
"Others"
table(comcastData$complaintType)
```

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

Billing	Charges	Email
363	139	15
Internet	Network	Others
472	2	1233

***Insight:-*** as per above table more complaint tickets are raised for the Internet issues.



5) 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.

```
#new categorical variable complaint status.
ggplot(as.data.frame(chartData), mapping = aes(State, Count)) +
  geom_col(aes(fill=complaintsStatus), width = 0.95) +
  theme(axis.text.x = element_text(angle = 90),
        axis.title.y = element_text(size = 15),
        axis.title.x = element_text(size = 15),
        title = element_text(size = 16, colour = "#0073C2FF"),
        plot.title = element_text(hjust = 0.5)) +
  labs(title = "Ticket Status Staked Bar Chart", x="States", y="No. of
tickets", fill= "Status")
open_complaints<- (comcastData$Status=='Open' |
comcastData$Status=='Pending')
close_complaints<- (comcastData$Status=='Closed' |
comcastData$Status=='Solved')
comcastData$complaintsStatus[open_complaints]<- 'Open'
comcastData$complaintsStatus[close_complaints]<- 'Closed'
stakes<- table(comcastData$complaintsStatus, comcastData$State)
stake
comcastData<- group_by(comcastData, State, complaintsStatus)
chartData<-summarise(comcastData, Count=n())
View(chartData)
```

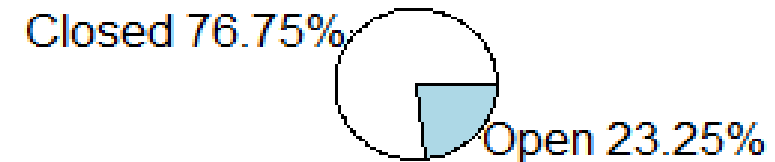
- **Table of open/closed complaints and state wise open complaints**

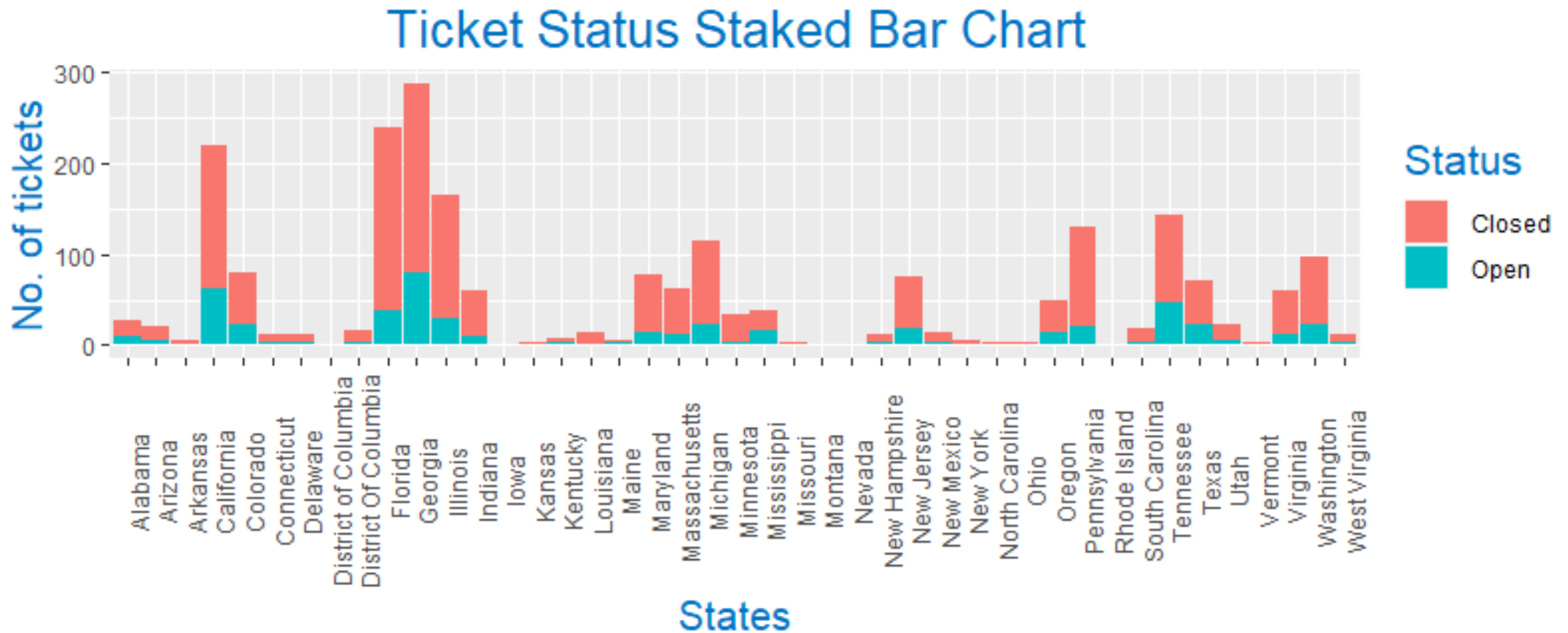
# A tibble: 34 x 2

	State	Numofcomplaints
	<chr>	<int>
1	Georgia	80
2	California	61
3	Tennessee	47
4	Florida	39
5	Illinois	29
6	Michigan	23
7	Washington	23
8	Colorado	22
9	Texas	22
10	Pennsylvania	20
# ... with 24 more rows		

# A tibble: 2 x 2

	complaintsStatus	Numofcomplaints
	<chr>	<int>
1	Closed	1707
2	Open	517





***Insight:-***Higher number of open complaints are from Georgia state.

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

```
#percentage of resolved internet complaints
internet_pr<-round(internet$Numofcomplaints/
sum(total$Numofcomplaints)*100,2)
internet_pr
```

```
> internet_pr
[1] 37.9
```

```
#percentage of resolved customer care call complaints
custo_pr<-round(customercc$Numofcomplaints/
sum(total$Numofcomplaints)*100,2)
custo_pr
```

```
> custo_pr
[1] 38.85
```

- **Insight:-** 37.9% complaints for internet and 38.85% complaints are resolved till date.