**Comcast Telecom Consumer Complaints**

1. **Import data into R environment.**

Code :-

#Importing Dataset

dataset <- read.csv(file.choose())

1. **Provide the trend chart for the number of complaints at monthly and daily granularity levels.**

Code:-

#Processing Date

install.packages('lubridate')

library(lubridate)

dmy(dataset$Date) ->dataset$Date

#Extracting Monthly and Daily Ticket Count

install.packages('dplyr')

library(dplyr)

monthly\_count<- summarise(group\_by(dataset,Month =as.integer(month(Date))),

Count = n())

daily\_count<- summarise(group\_by(dataset,Date),Count =n())

monthly\_count<-arrange(monthly\_count,Month)

#Comparing Monthly and Daily Complaints

#\_\_\_\_\_\_\_\_\_\_\_\_\_

library(ggplot2)

#1-Monthly

ggplot(data = monthly\_count,aes(x=Month,y=Count,label = Count))+

geom\_point(size = 1)+

geom\_line()+

geom\_text()+

scale\_x\_continuous(breaks = monthly\_count$Month)+

labs(title = "Monthly Ticket Count",x= "Months",y ="No. of Tickets")+

theme(plot.title = element\_text(hjust = 0.5))

#2-Daily

ggplot(data = daily\_count,aes(x=as.POSIXct(Date),y=Count))+

geom\_line()+

geom\_point(size = 1)+

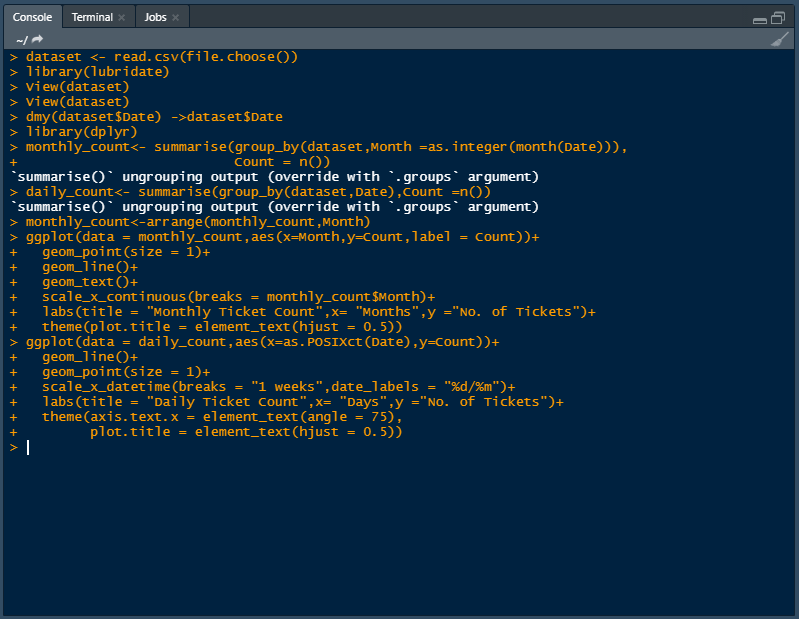
scale\_x\_datetime(breaks = "1 weeks",date\_labels = "%d/%m")+

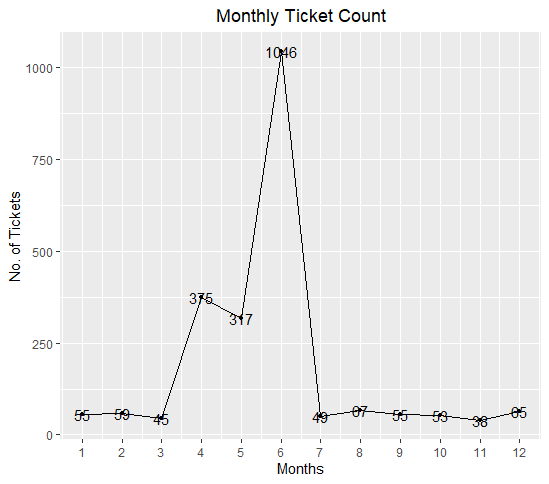
labs(title = "Daily Ticket Count",x= "Days",y ="No. of Tickets")+

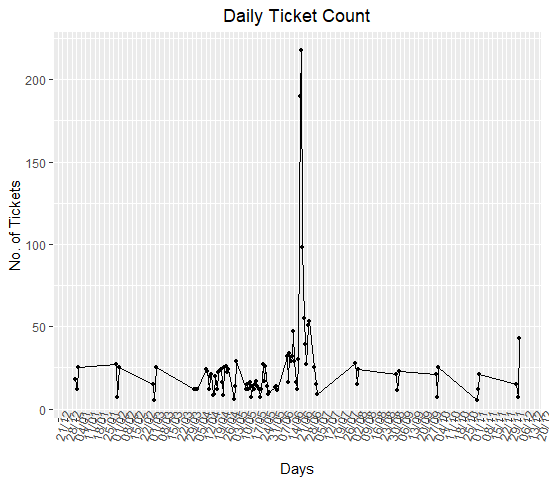
theme(axis.text.x = element\_text(angle = 75),

plot.title = element\_text(hjust = 0.5))

Outpute:-







1. **Provide a table with the frequency of complaint types**

Code:-

#Complaint Type Processing

network\_tickets<- contains(dataset$Customer.Complaint,match = 'network',ignore.case = T)

payment\_tickets<- contains(dataset$Customer.Complaint,match = 'payment',ignore.case = T)

billing\_tickets<- contains(dataset$Customer.Complaint,match = 'bill',ignore.case = T)

email\_tickets<- contains(dataset$Customer.Complaint,match = 'email',ignore.case = T)

charges\_ticket<- contains(dataset$Customer.Complaint,match = 'charge',ignore.case = T)

speed\_ticket<- contains(dataset$Customer.Complaint,match = 'speed',ignore.case = T)

service\_ticket<- contains(dataset$Customer.Complaint,match = 'service',ignore.case = T)

dataset$ComplaintType[payment\_tickets]<- "payment"

dataset$ComplaintType[network\_tickets]<- "Network"

dataset$ComplaintType[service\_ticket]<- "service"

dataset$ComplaintType[billing\_tickets]<- "Billing"

dataset$ComplaintType[email\_tickets]<- "Email"

dataset$ComplaintType[charges\_ticket]<- "Charges"

dataset$ComplaintType[speed\_ticket]<- "speed"

dataset$ComplaintType[-c(payment\_tickets,network\_tickets,

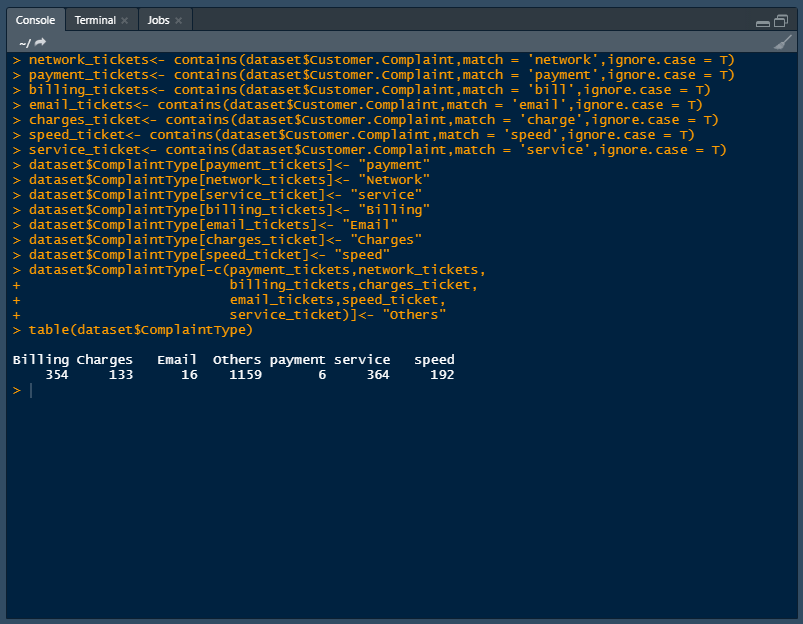
billing\_tickets,charges\_ticket,

email\_tickets,speed\_ticket,

service\_ticket)]<- "Others"

table(dataset$ComplaintType)

Outpute :-

****

1. **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:**

Code:-

#Creating new Variable ComplaintStatus with values Open and Closed

open\_complaints<- (dataset$Status == "Open"| dataset$Status =="Pending")

closed\_complaints<-(dataset$Status == "Closed"| dataset$Status =="Solved")

dataset$ComplaintStatus[ open\_complaints]<-"Open"

dataset$ComplaintStatus[closed\_complaints]<- "Closed"

#Creating Stacked barchart for complaints based on State and Status

dataset<- group\_by(dataset,State,ComplaintStatus)

chart\_data<- summarise(dataset,Count = n())

ggplot(as.data.frame(chart\_data) ,mapping = aes(x=State,y=Count))+

geom\_col(aes(fill = ComplaintStatus),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 Stacked Bar Chart ",

x = "States",y = "No of Tickets",

fill= "Status")

#Finding State which has Highest number of Unresolved Tickets

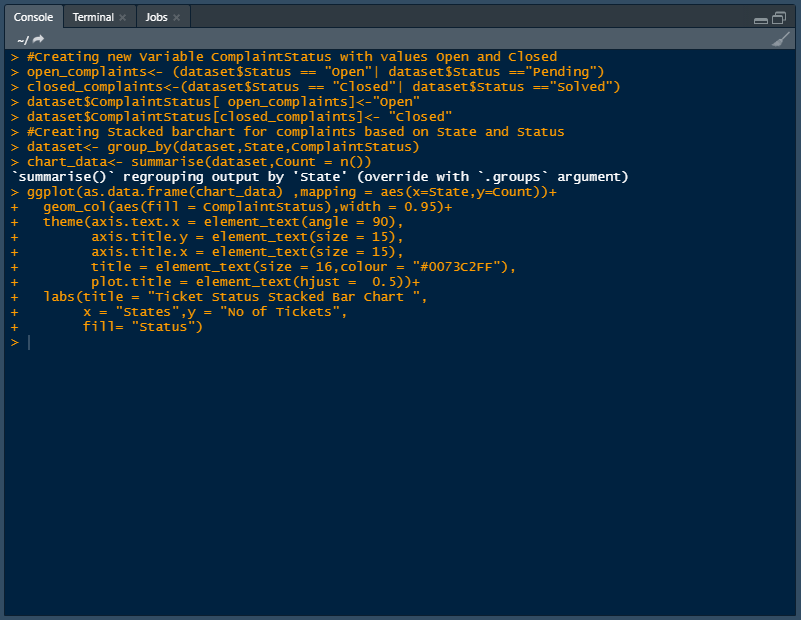
chart\_data%>%

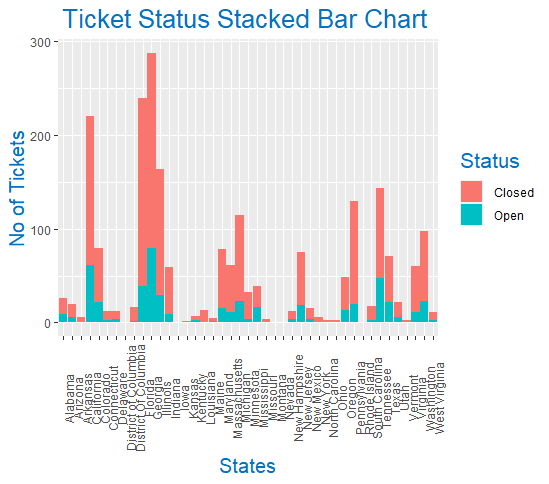
filter(ComplaintStatus == "Open")->

open\_complaints

open\_complaints[open\_complaints$Count == max(open\_complaints$Count),c(1,3)]

Outpute :-





it`s clearly shown that the highest number of complaints recorded from the state Georgia and the second highest number of complaints recorded from the state Florida

1. **Which state has the maximum complaints**

Code:

#Finding State which has Highest number of Unresolved Tickets

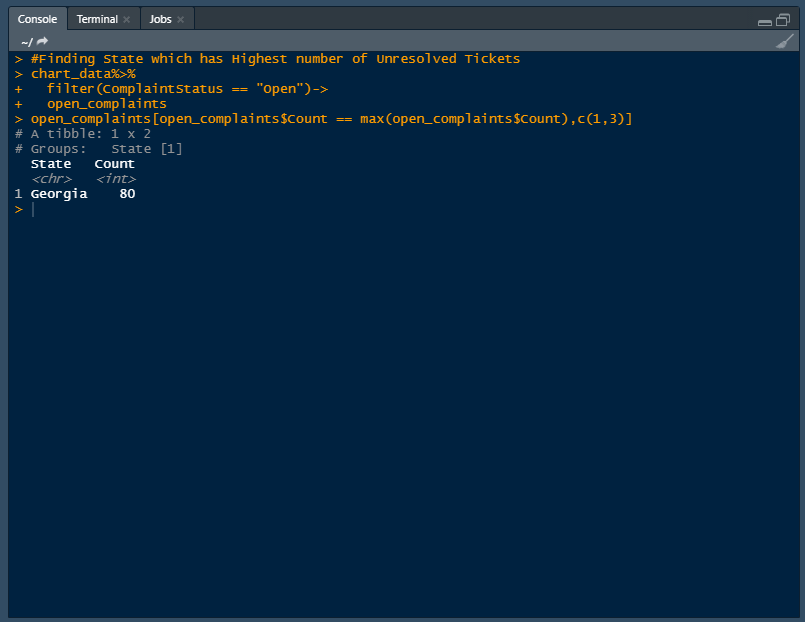
chart\_data%>%

filter(ComplaintStatus == "Open")->

open\_complaints

open\_complaints[open\_complaints$Count == max(open\_complaints$Count),c(1,3)]

Outpute:-



As we can observe that State Georgia has maximum number of unresolved tickets and these ticket count is 80.

1. **Provide the percentage of complaints resolved till date, which were received through theInternet and customer care calls.**

Code:-

#Calculating Resolution Percentage based on Total and Catagory

resolved\_data <- group\_by(dataset,ComplaintStatus)

total\_resloved<- summarise(resolved\_data ,percentage =(n()/nrow(resolved\_data)))

resolved\_data <- group\_by(dataset,Received.Via,ComplaintStatus)

Category\_resloved<- summarise(resolved\_data ,percentage =(n()/nrow(resolved\_data)))

#Ploting Pie Chart for Total Resolved Vs Category Resolved

install.packages("ggpubr")

library(ggpubr)

par(mfrow = c(1,2))

total<-ggplot(total\_resloved,aes(x= "",y =percentage,fill = ComplaintStatus))+

geom\_bar(stat = "identity",width = 1)+

coord\_polar("y",start = 0)+

geom\_text(aes(label = paste0(round(percentage\*100),"%")),

position = position\_stack(vjust = 0.5))+

labs(x = NULL,y = NULL,fill = NULL)+

theme\_classic()+theme(axis.line = element\_blank(),

axis.text = element\_blank(),

axis.ticks = element\_blank())

# Pie Chart for Category wise Ticket Status

category<-ggplot(Category\_resloved,

aes(x= "",y =percentage,fill = ComplaintStatus))+

geom\_bar(stat = "identity",width = 1)+

coord\_polar("y",start = 0)+

geom\_text(aes(label = paste0(Received.Via,"-",round(percentage\*100),"%")),

position = position\_stack(vjust = 0.5))+

labs(x = NULL,y = NULL,fill = NULL)+

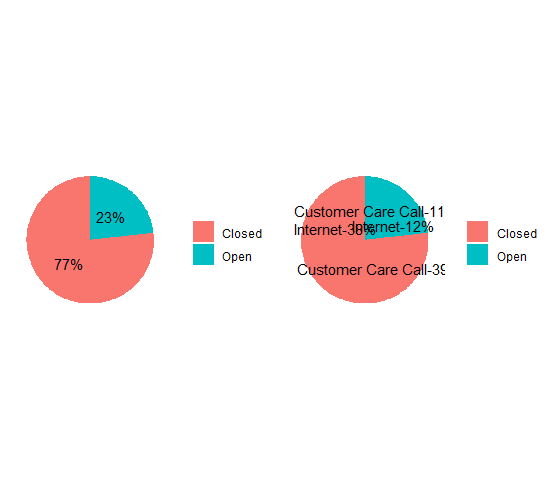
theme\_classic()+theme(axis.line = element\_blank(),

axis.text = element\_blank(),

axis.ticks = element\_blank())

ggarrange(total,category,nrow = 1, ncol = 2)

Outpute:-



With the help of above Chart of Total Resolved Vs Category Resolved we can conclude that the total resolved complaints are 77% in which 38% are received the internet and 39% are from the customer care calls.Also we can obserse that there are 23% complaints are still unresolved and in which 12% are received the internet and 11% are from the customer care calls