

```

install.packages("csv")

library(csv)

setwd("C:/Data science withR/Sambhav")

getwd()

comcast = read.csv("Comcast_Telecom_Complaints_data.csv",sep=",")

head(comcast)

View(comcast)

str(comcast)

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

# Daily Granularity

install.packages("dplyr")

library(dplyr)

barplot(table(comcast$Date),xlab="date",ylab="frequency",col="orange")

#monthly granularity

comcast$Date = as.character(comcast$Date)

class(comcast$Date)

extractmonth=function(dt)

{

unlist(strsplit(x=dt, split="-"))[2]

}

unlist(lapply(comcast$Date, extractmonth))->comcast$month

comcast$month=as.numeric(comcast$month)

barplot(table(comcast$month),xlab="month",ylab="frequency", col="red")

# Provide a table with the frequency of complaint types.

table(comcast$Status)

barplot(table(comcast$Status),xlab="complaint type", ylab="frequency",

        col=c("red","blue","green","yellow"))

##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.

```

```

class(comcast$Status)
table(comcast$Status)
open=c("Open","Pending")
close=c("closed","Solved")
comcast$status[comcast$Status %in% close]="close"
comcast$status[comcast$Status %in% open]="open"
#Provide state wise status of complaints in a stacked bar chart.
#Use the categorized variable from Q3. Provide insights on:
#create a cross table for the stacked barplot
table(comcast$status, comcast$State)
barplot(table(comcast$status, comcast$State),xlab="state",ylab="frequency",
         col= c("green","red"), legend=T)
#Response:- Florida and Alabama has the highest number of complaints however, the number of
closed complaints are higher than the open ones
#which means that the complaints are getting resolved .
#Which state has the maximum complaints
comcast%>%group_by(State)%>%summarise(n=n())%>%arrange(desc(n))->st
st
#Georgia has the maximum complaints ( all the complaints put together ie - open , close , pending ,
solved)
##Which state has the highest percentage of unresolved complaints (open complaints)
comcast%>%group_by(State,status)%>%filter(status == "open")%>%summarise(total =
n())%>%arrange(desc(total))->t
sum(t$total)->s
ptage = (t$total/517)*100
ptage
#Provide the percentage of complaints resolved till date which were received through the Internet
and customer care calls.
comcast%>%group_by(comcast$Received.Via,comcast$status)%>%filter(status
=="close")%>%summarise(tot = n())->p
p
sum(p$tot)
ptage = (p$tot/1707)*100

```

ptage