

```
rm(list = ls(all = TRUE))

install.packages("ggplot2")

library(ggplot2)

install.packages("RColorBrewer")

library(RColorBrewer)

install.packages("dplyr")

library(dplyr)

install.packages("lubridate")

library(lubridate)


setwd("C:/OneDrive - cabschoolo.cabschoolo/Desktop")

telcom <- read.csv("Comcast Telecom Complaints data.csv",header = TRUE)

head(telcom)


summary(telcom)

View(telcom)

names(telcom)


# Manipulating the column names and
# Replacing "." from column names to ""

names(telcom) <- gsub("\\.", "", names(telcom))

names(telcom)

View(telcom)

class(telcom$Date)
```

Class of Date column is character type, need to change it to Date type.

```
telcom$Date <- dmy(telcom$Date)
```

```
class(telcom$Date)
```

```
View(telcom)
```

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

```
daily <- table(telcom$Date)
```

```
daily
```

```
barplot(daily, xlab = 'Number of complaints', ylab = 'Date', main = 'Number of complaints on daily based')
```

#Monthly complaints chart

```
telcom_monthly = transform(telcom, MonthOfComplaint = as.numeric(format(Date, format="%m")))
```

```
head(telcom_monthly)
```

```
Monthly = table(telcom_monthly$MonthOfComplaint)
```

```
Monthly
```

```
barplot(Monthly, xlab = "No of Complaints", ylab = "Months", main = "Monthly No Of Complaints")
```

#3. Provide a table with the frequency of complaint types.

```
noofcomp <- table(telcom$CustomerComplaint)
```

```
head(noofcomp)
```

```
noofcomp1 = arrange(as.data.frame(noofcomp),desc(Freq))
```

```
head(noofcomp1)
```

#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.

```
complaint_status = transform(telcom,FinalStatus =  
ifelse((Status=='Open'|Status=='Pending'),'Open',"Closed"))
```

```
head(telcom)
```

```
head(complaint_status)
```

```
summary(as.factor(complaint_status$FinalStatus))
```

#Provide state wise status of complaints in a stacked bar chart

```
statewisecomp = table(complaint_status$FinalStatus,complaint_status$State)
```

```
statewisecomp = as.data.frame(statewisecomp)
```

```
head(statewisecomp)
```

```
ggplot(statewisecomp,aes(x=Var2,y=Freq))+geom_bar(stat="identity")
```

#6. which state has the highest complaint

```
Freq = table(complaint_status$State,complaint_status$FinalStatus)

Freq1 = as.data.frame(Freq)

Status_by_state = reshape(Freq1,idvar="Var1",timevar ='Var2',direction="wide")

Status_by_state_Final = transform(Status_by_state,Unresolved_Pct =
(Freq.Open/(Freq.Open+Freq.Closed)))

View(Status_by_state_Final)
```

#7.finding the maximum state with complaint

```
max(Status_by_state_Final$Unresolved_Pct)
```

#8. which state has a 0.5 value

#Kansas has the highest unresolved complaints

```
which(Status_by_state_Final$Unresolved_Pct == "0.5")
```

#7. Provide the percentage of complaints resolved till date

#which were received through theInternet and customer care calls.

```
Freq4 = table(complaint_status$ReceivedVia,complaint_status$FinalStatus)
```

```
Freq4 = as.data.frame(Freq4)
```

```
Status_by_Channel = reshape(Freq4,idvar="Var1",timevar ='Var2',direction="wide")
```

```
Status_by_Channel_Final = transform(Status_by_Channel, Unresolved_Pct =  
(Freq.Open/(Freq.Open+Freq.Closed)))
```

```
head(Status_by_Channel_Final)
```

```
View(Status_by_Channel_Final)
```

#there is less different in complaint resolved status through customer care call and internet

#however the freq.open to req.colosed as to total varability in unresolved_pct is.22 for customer care call

#and 0.23 is from internet.

```
# Var1 Freq.Closed Freq.Open Unresolved_Pct
```

```
#1 Customer Care Call      864    255    0.2278820
```

```
#2      Internet      843    262    0.2371041
```