Comcast Telecom Consumer Complaints

Objective:

Provide the trend chart for the number of complaints at monthly and daily granularity levels. Provide a table with the frequency of complaint types. Complaint types which 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. Check the state has the maximum complaints. 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.

Data set names:

Comcast Telecom Complaints data

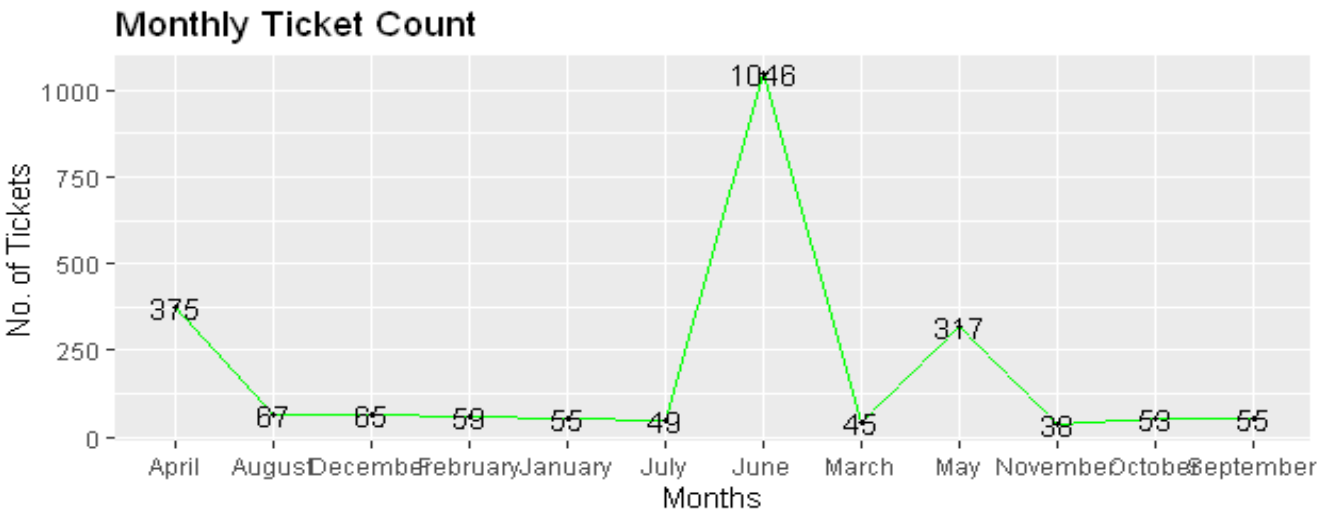
Source code:

```
#Installing Packages

#****************************************


install.packages('ggplot2')

install.packages('ggpubr')

install.packages('dplyr')

install.packages('lubridate')

install.packages('stringi')

install.packages('data.table')

install.packages('plyr')


#Adding Libraries

#****************************************


library(ggplot2)

library(ggpubr)

library(dplyr)
```

```
library(lubridate)

library(stringi)

library(data.table)

library(plyr)


#Importing Dataset

#**************************************


comcast_data<- read.csv("C:\\Users\\Tejasri5\\Downloads\\Data Science with R\\Project\\Comcast Telecom Complaints data.csv",header=TRUE)

comcast_data

#summary(comcast_data)


#Extracting monthly and daily count

#**************************************


comcast_data$Date<-dmy(comcast_data$Date)

comcast_data$datemonth<-months(as.Date(comcast_data$Date))

#comcast_data$datemonth

Months_counts <- table(comcast_data$datemonth)

#Months_counts

months_counts_df<-as.data.frame(Months_counts)

names(months_counts_df)[1]<-"Months"

names(months_counts_df)[2]<-"Count"

#months_counts_df


comcast_data$date_of_each_month<-day(as.Date(comcast_data$Date))

#comcast_data$date_of_each_month
```

```
Date_counts <- table(comcast_data$date_of_each_month)

#Date_counts

Date_counts_df<-as.data.frame(Date_counts)

names(Date_counts_df)[1]<-"date_of_each_month"

names(Date_counts_df)[2]<-"Count"

#Date_counts_df


#Trend Chart for Monthly count

#**************************************


ggplot(data=months_counts_df, aes(x=Months, y=Count, label=Count, group=1)) +

 geom_line(color="green")+

 geom_point(size = 0.8)+geom_text()+

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

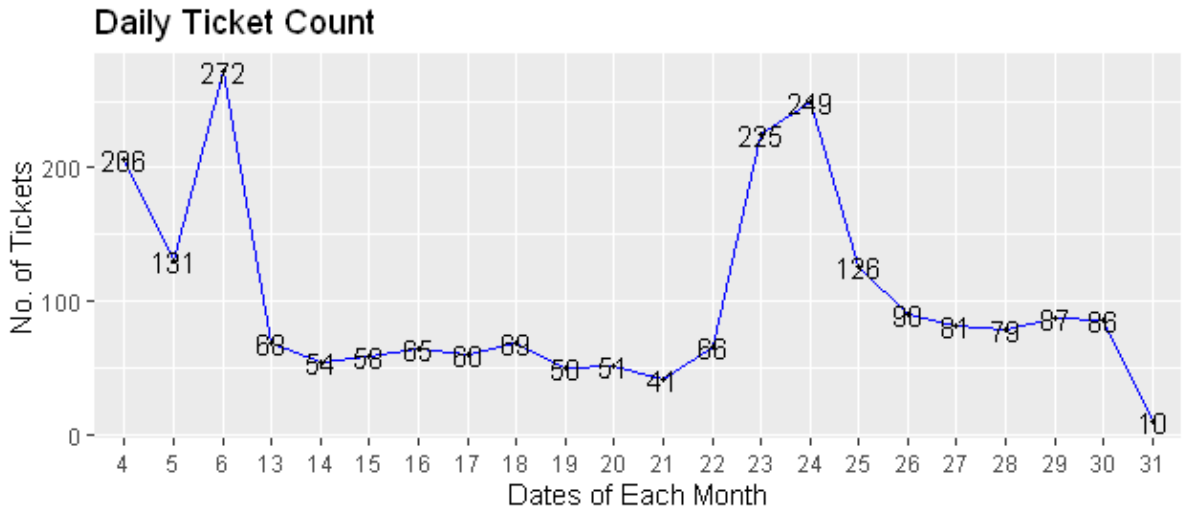
Output:




```
#Trend Chart for Daily count

#**************************************
```

```
ggplot(data=Date_counts_df, aes(x=date_of_each_month, y=Count,label=Count, group=1)) +
  geom_line(color="blue")+
  geom_point(size = 0.8)+
  geom_text()+labs(title = "Daily Ticket Count",x= "Dates of Each Month",y ="No. of Tickets")
```

Output:




```
#Finding Complaint Type
#**************************************


#comcast_data$Customer.Complaint
#complaints_count<-table(comcast_data$Customer.Complaint)


network_tickets<- data.frame(contains(comcast_data$Customer.Complaint,match =
'network',ignore.case = T))
internet_tickets<- data.frame(contains(comcast_data$Customer.Complaint,match =
'internet',ignore.case = T))
billing_tickets<- data.frame(contains(comcast_data$Customer.Complaint,match = 'bill',ignore.case =
T))
datacap_tickets<- data.frame(contains(comcast_data$Customer.Complaint,match = 'data
cap',ignore.case = T))
```

```
customerservice_tickets<- data.frame(contains(comcast_data$Customer.Complaint,match =
'customer service',ignore.case = T))


nrow(network_tickets)

nrow(internet_tickets)

nrow(billing_tickets)

nrow(datacap_tickets)

nrow(customerservice_tickets)


#Complaint Type that has maximum Tickets

#*****************************************


if(nrow(network_tickets) > nrow(internet_tickets)){

  print("Network Issues")

} else if(nrow(internet_tickets) > nrow(billing_tickets)){

  print("Internet Issues")

} else if(nrow(billing_tickets) > nrow(datacap_tickets)){

  print("Billing Issues")

} else if(nrow(datacap_tickets) > nrow(customerservice_tickets)){

  print("Data cap Issues")

} else {

  print("Customer Service Issues")

}
```

Output:

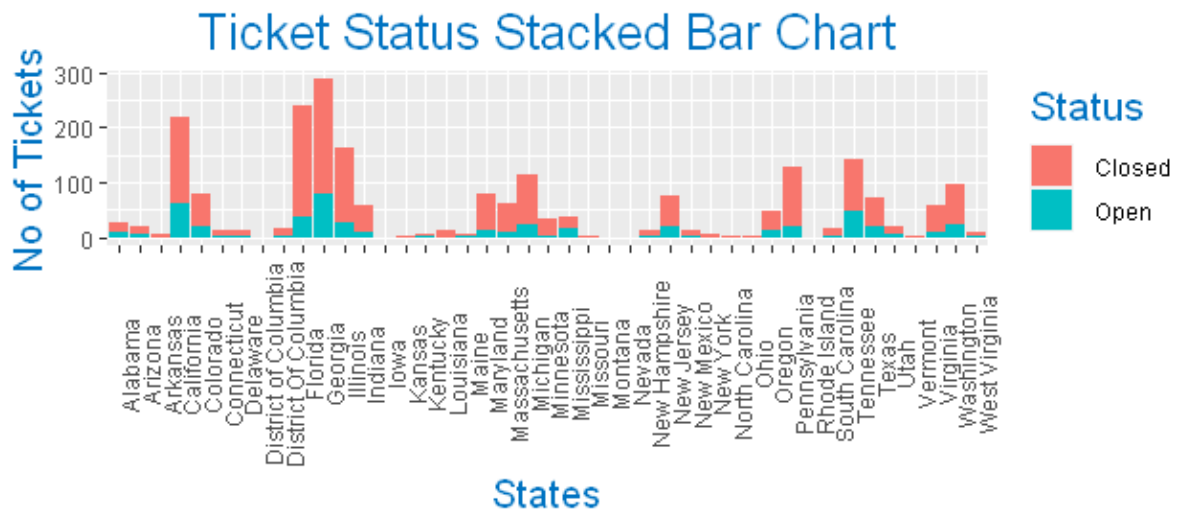
```
[1] "Internet Issues"
```

```
#Open and Pending Statuses are considered as "Open"

#*****************************************************
```

```
#my_data <- as_tibble(comcast_data$Status)
#Open_complaints <- my_data %>% filter(value=="Open" | value=="Pending")
#comcast_data<-subset(comcast_data,select=-c(ComplaintStatus))

comcast_data$Status<-gsub('Pending', 'Open', comcast_data$Status)

Open_complaints<-(comcast_data$Status == "Open")
comcast_data$Complaint_Status[Open_complaints]<-"Open"

#Solved and Closed Statuses are considered as "Closed"
#*******************************************************

comcast_data$Status<-gsub('Solved','Closed', comcast_data$Status)

Closed_complaints<-(comcast_data$Status == "Closed")
comcast_data$Complaint_Status[Closed_complaints]<-"Closed"

#Stacked Bar chart for Open and Closed Complaints
#*******************************************************

comcast_data<- group_by(comcast_data,State,Complaint_Status)
chart_data<- dplyr::summarise(comcast_data,Count = n())

ggplot(as.data.frame(chart_data) ,mapping = aes(State,Count))+
  geom_col(aes(fill = Complaint_Status),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")
```

Output:



#State that has the maximum complaints

#***

chart_data%>% filter(Complaint_Status=="Open")-> Open_complaints

#max(Open_complaints$Count)

#(Open_complaints)[1]

Open_complaints[Open_complaints$Count == max(Open_complaints$Count),c(1,3)]

Output:

```
A tibble: 1 x 2
# Groups:   State [1]
  State   Count
  <chr>   <int>
1 Georgia    80
```

#Complaints which were received through the Internet and customer care calls.

#***

Resolved_data<-group_by(comcast_data,Complaint_Status)

Total_resolved<-dplyr::summarise(Resolved_data ,percentage = (n()/nrow(Resolved_data)))

#Total_resolved

Resolved_data1 <- group_by(comcast_data,Received.Via,Complaint_Status)

Category_resloved<-dplyr::summarise(Resolved_data1,percentage =(n()/nrow(Resolved_data)))

#Category_resloved

Output:

```
Total_resolved
# A tibble: 2 x 2
  Complaint_Status percentage
  <chr>                 <dbl>
1 Closed                0.768
2 Open                  0.232
```

```
Category_resloved
# A tibble: 4 x 3
# Groups:   Received.Via [2]
  Received.Via       Complaint_Status percentage
  <chr>              <chr>                 <dbl>
1 Customer Care Call Closed                0.388
2 Customer Care Call Open                  0.115
3 Internet           Closed                0.379
4 Internet           Open                  0.118
```

#Pie Chart for Category wise Ticket Status

#***

par(mfrow = c(1,2))

total<-ggplot(data=Total_resolved,

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

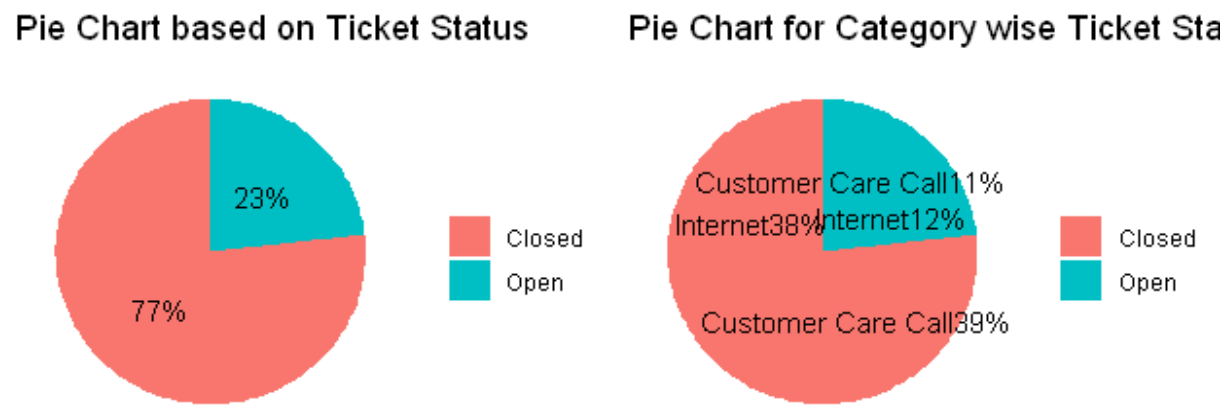
geom_bar(stat = "identity",width = 1)+

coord_polar("y",start = 0)+

geom_text(aes(label = paste0(round(percentage*100),"%")),

```
        position = position_stack(vjust = 0.5))+
  labs(title = "Pie Chart based on Ticket Status",x = NULL,y = NULL,fill = NULL)+
  theme_classic()+theme(axis.line = element_blank(),
                        axis.text = element_blank(),
                        axis.ticks = element_blank())



category<-ggplot(data=Category_resloved,
                 aes(x= "",y =percentage,fill = Complaint_Status))+
  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(title = "Pie Chart for Category wise Ticket Status",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)
```

Output:



Insights:

June month has maximum no of complaints.

More complaints are based on Internet Issues.

2nd half of each month has more number of complaints.

Georgia State has maximum complaints.

Georgia State has maximum Open/unsolved complaints.

Total resolved complaints are 77% in which 38% are received from the internet and 39% are from the customer care calls.

There are 23% complaints are still unresolved and in which 12% are received from the internet and 11% are from the customer care calls.