Comcast Telecom Consumer Complaints

Comcast is an American global telecommunication company. The firm has been providing terrible customer service. They continue to fall short despite repeated promises to improve. Only last month (October 2016) the authority fined them a $2.3 million, after receiving over 1000 consumer complaints.

The existing database will serve as a repository of public customer complaints filed against Comcast.
It will help to pin down what is wrong with Comcast's customer service.

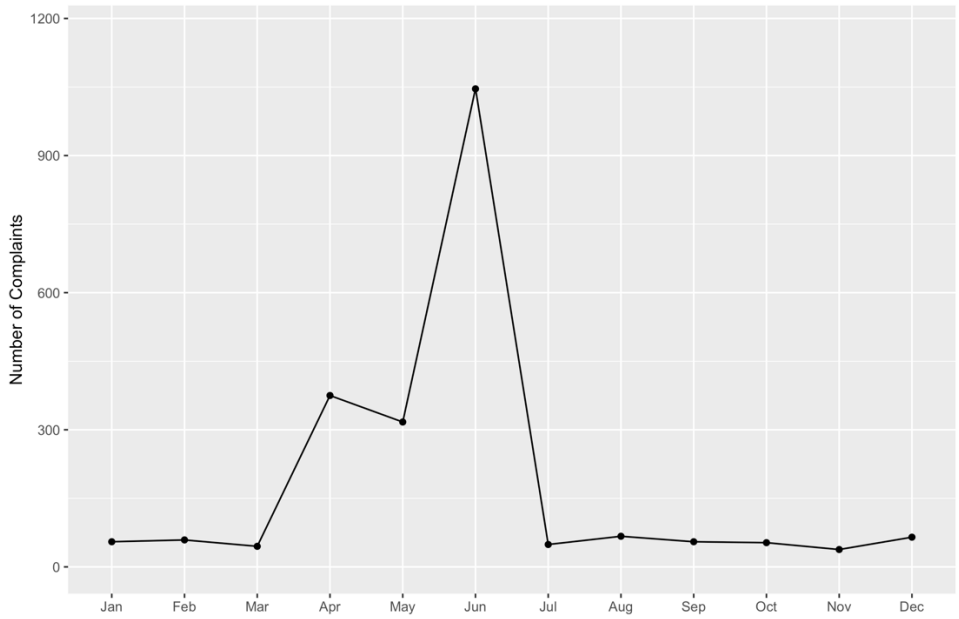
Analysis Task

- Import data into R environment.

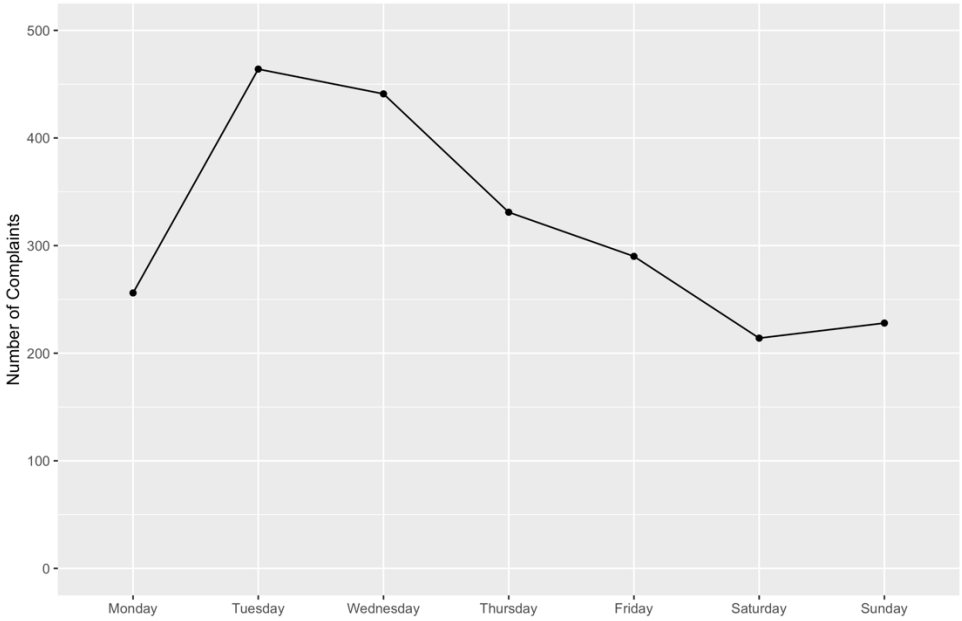
```
1.  library(ggplot2)
2.  library(lubridate)
3.  library(dplyr)
4.  library(stringi)
5.  library(ggpubr)
6.  library(stringr)
7.
8.  #Import data into R environment
9.  ctcd <- read.csv(file.choose())
10. View(ctcd)
11. head(ctcd)
12. sum(is.na(ctcd))
13. summary(ctcd)
```

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

```
1.  #1.Provide the trend chart for the number of complaints at monthly and daily granularity
    levels
2.  class(ctcd$Date)
3.
4.  #Convert from character to date
5.  ctcd$Date <- dmy(ctcd$Date)
6.  class(ctcd$Date)
7.
8.  #Monthly Complaints
9.  monthly <- summarise(group_by(ctcd,Month =as.integer(month(Date))),Count = n())
10. View(monthly)
11.
12. monthly_plot <- ggplot(monthly,aes(Month,Count)) + geom_line(aes(group=1)) + geom_point() +
    scale_y_continuous(name = "Number of Complaints",limits = c(0,1170)) +
    scale_x_discrete(name="",limits=c("Jan","Feb","Mar","Apr","May","Jun",
13.                                                 "Jul","Aug","Sep","Oct","Nov","Dec"))
14. monthly_plot
15. ggsave("monthly_plot.png")
16.
```




```
1.  #daily Complaints
2.  daily <- summarise(group_by(ctcd,day=weekdays(as.Date(Date))),Count = n())
3.  View(daily)
4.
5.  daily_plot <- ggplot(daily, aes(day, Count))  + geom_line(aes(group=1)) + geom_point() +
    scale_y_continuous(name = "Number of Complaints",limits = c(0,500)) +
    scale_x_discrete(name="",limits=c("Monday","Tuesday","Wednesday","Thursday","Friday","Saturd
    ay","Sunday"))
6.
7.  daily_plot
8.  ggsave("daily_plot.png")
```



- Provide a table with the frequency of complaint types.

```
#2.Provide a table with the frequency of complaint types
network <- contains(ctcd$Customer, match = "network", ignore.case = TRUE)
internet <- contains(ctcd$Customer, match = "internet", ignore.case = TRUE)
bills <- contains(ctcd$Customer, match = "billing", ignore.case = TRUE)
charges <- contains(ctcd$Customer, match = "charge", ignore.case = TRUE)
email <- contains(ctcd$Customer, match = "email", ignore.case = TRUE)
data_capacity <- contains(ctcd$Customer, match = "data cap", ignore.case = TRUE)

ctcd$Complaint.Type[network] <- "network"
ctcd$Complaint.Type[internet] <- "internet"
ctcd$Complaint.Type[bills] <- "billing"
ctcd$Complaint.Type[charges] <- "charges"
ctcd$Complaint.Type[email] <- "email"
ctcd$Complaint.Type[data_capacity] <- "data capacity"
ctcd$Complaint.Type[-c(network,internet,bills,charges,email,data_capacity)] <- "others"

table(ctcd$Complaint.Type)
```

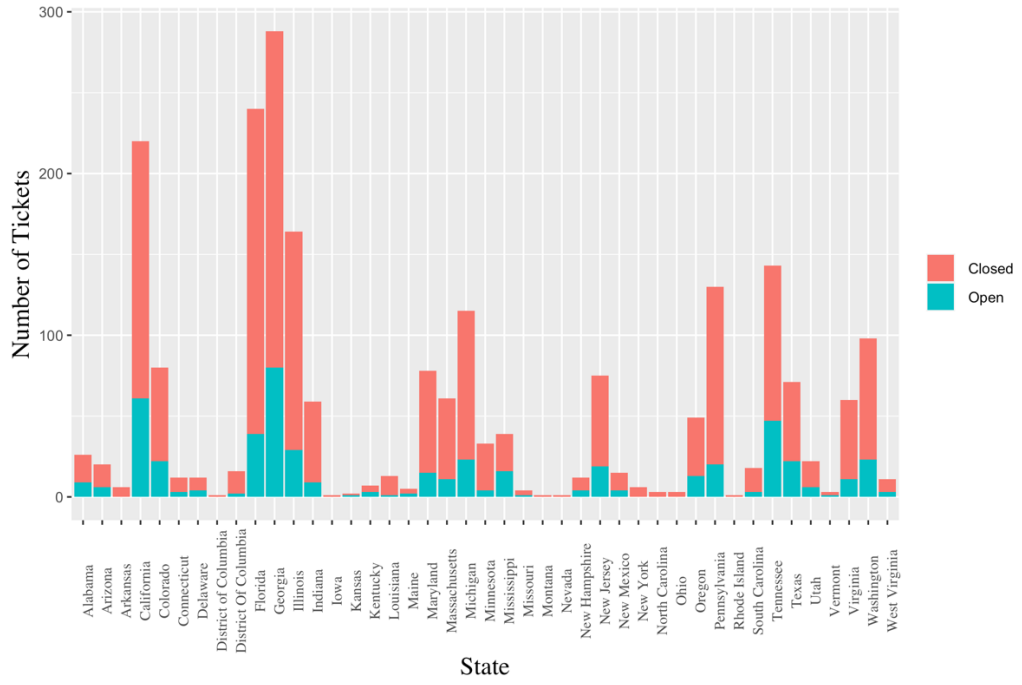
```
> table(ctcd$Complaint.Type)

       billing       charges data capacity         email      internet
           288           139           150            16           472
        others
          1158
```

- Create a new categorical variable with value as Open and Closed.

```
 #3.Create a new categorical variable with value as Open and Closed
open <- (ctcd$Status=="Open"|ctcd$Status=="Pending")
closed <- (ctcd$Status=="Closed"|ctcd$Status=="Solved")

ctcd$Status.Category[open] <- "Open"
ctcd$Status.Category[closed] <- "Closed"

```

- Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3.

```
#4.Provide state wise status of complaints in a stacked bar chart. Use the categorized
variable from Q3
ctcd <- group_by(ctcd,State,Status.Category)
chart <- summarise(ctcd,Count = n())
ggplot(as.data.frame(chart) ,mapping = aes(State,Count))+
  geom_col(aes(fill = Status.Category),width = 0.90)+
  theme(axis.text.x = element_text(family="serif",angle = 90),
        axis.title.y = element_text(family="serif",size = 15),
        axis.title.x = element_text(family="serif",size = 15),
        title = element_text(family="serif",size = 16,colour = "black"))+
  labs(x = "State", y = "Number of Tickets", fill= "")
ggsave("soc_stacked_chart.png")

#Which state has the highest percentage of unresolved complaints
chart%>% filter(Status.Category=="Open")-> Open_complaints
Open_complaints[Open_complaints$Count == max(Open_complaints$Count),c(1,3)]
```




```
> chart%>% filter(Status.Category=="Open")-> Open_complaints
> Open_complaints[Open_complaints$Count == max(Open_complaints$Count),c(1,3)]
# A tibble: 1 x 2
# Groups:   State [1]
  State   Count
  <chr>   <int>
1 Georgia    80
```

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

```
#5.Provide the percentage of complaints resolved till date,
#which were received through the Internet and customer care calls.
total_complaints<- summarise(group_by(ctcd,Status.Category) ,percentage
=(n()/nrow(resolved_data)*100))
total_complaints
resolved <- group_by(ctcd,Received.Via,Status.Category)
summarise(resolved ,percentage =(n()/nrow(resolved)*100))
```

```
 > total_complaints
# A tibble: 2 x 2
  Status.Category percentage
  <chr>                <dbl>
1 Closed                76.8
2 Open                  23.2
> resolved <- group_by(ctcd,Received.Via,Status.Category)
> summarise(resolved ,percentage =(n()/nrow(resolved)*100))
# A tibble: 4 x 3
# Groups:   Received.Via [2]
  Received.Via       Status.Category percentage
  <chr>              <chr>                <dbl>
1 Customer Care Call Closed                38.8
2 Customer Care Call Open                  11.5
3 Internet           Closed                37.9
4 Internet           Open                  11.8
```

Insights:

- June has the largest number of complaints.
- Number of complaints is significantly larger on Tuesdays and Wednesdays.
- Most of the complaints are related to Internet issues.
- Georgia has the maximum complaints. Also the highest percentage of unresolved complaints.
- 76.8% of the complaints are resolved in which 38.8% were received through the customer care call and 37.9% were received through the internet.