```
> library(readxl)
> Data <- read.csv('C:\\Users\\MONALISH\\Desktop\\Comcast Telecom Complaints data.csv')
> head(Data,4) #Printing Few Rows Of Data
  Ticket..                                                Customer.Complaint       Date
 Time       Received.Via     City    State Zip.code Status Filing.on.Behalf.of.Someone
1   250635                                    Comcast Cable Internet Speeds 22-04-2015  3:53:
50 PM Customer Care Call Abingdon Maryland    21009 Closed                          No
2   223441                      Payment disappear - service got disconnected 04-08-2015 10:22:
56 AM           Internet  Acworth  Georgia    30102 Closed                          No
3   242732                                     Speed and Service 18-04-2015  9:55:
47 AM           Internet  Acworth  Georgia    30101 Closed                         Yes
4   277946 Comcast Imposed a New Usage Cap of 300GB that punishes streaming. 05-07-2015 11:59:
35 AM           Internet  Acworth  Georgia    30101   Open                         Yes
> str(Data) #checking the structure of the dataset
'data.frame':   2224 obs. of  10 variables:
 $ Ticket..                 : chr  "250635" "223441" "242732" "277946" ...
 $ Customer.Complaint       : chr  "Comcast Cable Internet Speeds" "Payment disappear - serv
ice got disconnected" "Speed and Service" "Comcast Imposed a New Usage Cap of 300GB that punis
hes streaming." ...
 $ Date                     : chr  "22-04-2015" "04-08-2015" "18-04-2015" "05-07-2015" ...
 $ Time                     : chr  "3:53:50 PM" "10:22:56 AM" "9:55:47 AM" "11:59:35 AM" ...
 $ Received.Via             : chr  "Customer Care Call" "Internet" "Internet" "Internet" ...
 $ City                     : chr  "Abingdon" "Acworth" "Acworth" "Acworth" ...
 $ State                    : chr  "Maryland" "Georgia" "Georgia" "Georgia" ...
 $ Zip.code                 : int  21009 30102 30101 30101 30101 30101 30101 49221 94502 945
01 ...
 $ Status                   : chr  "Closed" "Closed" "Closed" "Open" ...
 $ Filing.on.Behalf.of.Someone: chr  "No" "No" "Yes" "Yes" ...
> #Formatting the Date Column as the code below, by using lubridate library available, similar
ly we will seperate month column as well from the given dates. To plot month data with complai
nts filed in that Month.
> #Loading The Date Into Single Format
> #Use Lubridate Library to Format the Date Column
> library(lubridate)

Attaching package: ‘lubridate’

The following objects are masked from ‘package:base’:

    date, intersect, setdiff, union

Warning message:
package ‘lubridate’ was built under R version 4.0.5
> li<-parse_date_time(x = Data$Date,
+                 orders = c("d m y", "d B Y", "m/d/y"),
+                 locale = Sys.getlocale("LC_TIME"))
> data2<-Data
> data2$Date <- li
> #Dates Loaded In the Same Format in the new Dataframe
> #str(data2$Date)
>
> #Extracting Month Column and Converting to The labels.
> data2$Month <- format(as.Date(data2$Date), "%m")
> data2$Month<- month.abb[as.integer(data2$Month)]
> head(data2)
  Ticket..                                                Customer.Complaint       Date
 Time       Received.Via     City    State Zip.code Status Filing.on.Behalf.of.Someone Month
1   250635                                    Comcast Cable Internet Speeds 2015-04-22  3:53:
50 PM Customer Care Call Abingdon Maryland    21009 Closed                          No   Apr
2   223441                      Payment disappear - service got disconnected 2015-08-04 10:22:
56 AM           Internet  Acworth  Georgia    30102 Closed                          No   Aug
3   242732                                     Speed and Service 2015-04-18  9:55:
47 AM           Internet  Acworth  Georgia    30101 Closed                         Yes   Apr
4   277946 Comcast Imposed a New Usage Cap of 300GB that punishes streaming. 2015-07-05 11:59:
35 AM           Internet  Acworth  Georgia    30101   Open                         Yes   Jul
5   307175                       Comcast not working and no service to boot 2015-05-26  1:25:
26 PM           Internet  Acworth  Georgia    30101 Solved                          No   May
6   338519          ISP Charging for arbitrary data limits with overage fees 2015-12-06  9:59:
40 PM           Internet  Acworth  Georgia    30101 Solved                          No   Dec
> #Analysis Of Data
> library(dplyr)

Attaching package: ‘dplyr’
```

```
The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union

Warning message:
package 'dplyr' was built under R version 4.0.5
> data_date<-data2 %>% group_by(Date) %>% dplyr::summarise(frequency = n())
> df <-data_date[order(-data_date$frequency),]
> dff<-head(df)
> dff
# A tibble: 6 x 2
  Date                frequency
  <dttm>                  <int>
1 2015-06-24 00:00:00       218
2 2015-06-23 00:00:00       190
3 2015-06-25 00:00:00        98
4 2015-06-26 00:00:00        55
5 2015-06-30 00:00:00        53
6 2015-06-29 00:00:00        51
> #CONCLUSION:This Above Data Frame Tells me that, on June 24 Comcast was reported with 218 co
mplaints, particularly indicating a VERY BAD DAY for the company to fix up many issues popping
 up on that day.
> library(ggplot2)
Warning message:
package 'ggplot2' was built under R version 4.0.5
> ggplot(data_date, aes(Date, frequency, group = 1)) +
+   geom_point() +
+   geom_line() +
+   xlab("Date") +
+   ylab("Number of Complaints")
> #CONCLUSION: Clearly, from the above Trend Graph, we can easily say that in the month of JUN
E 2015, Comcast got reported with Maximum Number of complaints
> library(ggplot2

+ > ggplot(dff, aes(Date, frequency, group = 1)) +
+   geom_point() +
+   geom_line() +
+   xlab("Date") +
+   ylab("Number of Complaints")
> #
> CONCLUSION: It is clear that on June 24, Company got reported with many complaints. This is
the following trend for a few observations from the month of June
Error: unexpected symbol in "CONCLUSION: It is"
> #
> CONCLUSION: It is clear that on June 24, Company got reported with many complaints. This is
the following trend for a few observations from the month of June
Error: unexpected symbol in "CONCLUSION: It is"
> #
> #CONCLUSION: It is clear that on June 24, Company got reported with many complaints. This is
 the following trend for a few observations from the month of June
> data_month<-data2 %>%
+ group_by(Month) %>% dplyr :: summarise(frequency = n())
> data_month
# A tibble: 12 x 2
   Month frequency
   <chr>     <int>
 1 Apr         375
 2 Aug          67
 3 Dec          65
 4 Feb          59
 5 Jan          55
 6 Jul          49
 7 Jun        1046
 8 Mar          45
 9 May         317
10 Nov          38
11 Oct          53
12 Sep          55
> data2$Month <- as.factor(data2$Month)
> levels(data2$Month)
```

```
 [1] "Apr" "Aug" "Dec" "Feb" "Jan" "Jul" "Jun" "Mar" "May" "Nov" "Oct" "Sep"
> library(ggplot2)
> ggplot(data_month, aes(Month, frequency, group = 1)) +
+   geom_point() +
+   geom_line() +
+   xlab("Month") +
+   ylab("Number of Complaints")
> #we can only add group = 1 into the ggplot or geom_line aes().
> #For line graphs, the data points must be grouped so that it knows which points to connect.
> #In this case, it is simple -- all points should be connected, so group=1.
> #When more variables are used and multiple lines are drawn, the grouping for lines is usuall
y done by variable.
> #CONCLUSION:we can clearly say that in the Month of June, date 24 company was filed with a m
aximum number of complaints.
> #Frequency Table For Customer Complaints During Year 2015 - 2016 Period
> library(dplyr)
> #Converting All String Values to Lower, so as to Eliminate Duplication of Any Complaint
> data3<-data2%>% mutate(Customer.Complaint = tolower(Customer.Complaint))
> CustTable <- table(data3$Customer.Complaint)
> CustTable <- data.frame(CustTable)
> filtered<-CustTable %>%
+    rename(
+        CustomerComplaintType = Var1,
+        Frequency = Freq
+     )
> final <- filtered %>% arrange(desc(Frequency))
>
> #Fetching The Top 20 complaints filed by customers on different days.
> final_most<-head(final,20)
> final_most
       CustomerComplaintType Frequency
1                    comcast       102
2             comcast data cap        30
3             comcast internet        29
4            comcast data caps        21
5              comcast billing        18
6              comcast service        15
7               internet speed        15
8                    data caps        13
9    unfair billing practices         13
10                    data cap        12
11           comcast complaint        11
12             comcast/xfinity        11
13   comcast internet service         10
14                    billing          9
15              billing issues         8
16 comcast billing complaint          5
17 comcast billing practices          5
18               comcast cable        5
19              comcast issues         5
20 complaint against comcast          5
> library(ggplot2)
> ggplot(head(final_most,6), aes(CustomerComplaintType, Frequency)) +
+   geom_bar(stat = "identity")
> #CONCLUSION FROM THE BAR PLOT: Customers Are Mainly complaining about the Data Caps, Interne
t Speed, Billing Methods and Services that Comcast is Providing and Very few Cases were regist
ered against Comcast Cable Services.
> library(stringr)
> library(tidyverse)
-- Attaching packages ---------------------------------------------------------- tidyverse
 1.3.0 --
v tibble  3.0.4     v purrr   0.3.4
v tidyr   1.1.2     v forcats 0.5.0
v readr   1.4.0
-- Conflicts ------------------------------------------------------------------ tidyverse_confl
icts() --
x lubridate::as.difftime() masks base::as.difftime()
x lubridate::date()        masks base::date()
x dplyr::filter()          masks stats::filter()
x lubridate::intersect()   masks base::intersect()
x dplyr::lag()             masks stats::lag()
x lubridate::setdiff()     masks base::setdiff()
x lubridate::union()       masks base::union()
> levels(Data$Status)
```

```
NULL
> library(plyr)
Error in library(plyr) : there is no package called ‘plyr’
> Data$Status_New<-revalue(Data$Status, c(Pending = "Open", Solved = "Closed"))
Error in revalue(Data$Status, c(Pending = "Open", Solved = "Closed")) :
  could not find function "revalue"
> head(Data)
  Ticket..                                                  Customer.Complaint       Date
 Time
1   250635                                      Comcast Cable Internet Speeds 22-04-2015  3:53:
50 PM
2   223441                        Payment disappear - service got disconnected 04-08-2015 10:22:
56 AM
3   242732                                               Speed and Service 18-04-2015  9:55:
47 AM
4   277946 Comcast Imposed a New Usage Cap of 300GB that punishes streaming. 05-07-2015 11:59:
35 AM
5   307175                          Comcast not working and no service to boot 26-05-2015  1:25:
26 PM
6   338519           ISP Charging for arbitrary data limits with overage fees 06-12-2015  9:59:
40 PM
        Received.Via     City    State Zip.code Status Filing.on.Behalf.of.Someone
1 Customer Care Call Abingdon Maryland    21009 Closed                          No
2           Internet  Acworth  Georgia    30102 Closed                          No
3           Internet  Acworth  Georgia    30101 Closed                         Yes
4           Internet  Acworth  Georgia    30101   Open                         Yes
5           Internet  Acworth  Georgia    30101 Solved                          No
6           Internet  Acworth  Georgia    30101 Solved                          No
> #We have merged Pending requests as Open and Solved Requests as Closed.
> levels(Data$State)
NULL
> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> library(dplyr)
> Data$Status_New<-revalue(Data$Status, c(Pending = "Open", Solved = "Closed"))
Error in revalue(Data$Status, c(Pending = "Open", Solved = "Closed")) :
  could not find function "revalue"
> install.packages("plyr")
Installing package into ‘C:/Users/MONALISH/Documents/R/win-library/4.0’
(as ‘lib’ is unspecified)
--- Please select a CRAN mirror for use in this session ---
trying URL 'https://mirror.niser.ac.in/cran/bin/windows/contrib/4.0/plyr_1.8.6.zip'
Content type 'application/zip' length 1499785 bytes (1.4 MB)
downloaded 1.4 MB

package ‘plyr’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\MONALISH\AppData\Local\Temp\RtmpK8ADG3\downloaded_packages
> library(plyr)
------------------------------------------------------------------------------
-------
You have loaded plyr after dplyr - this is likely to cause problems.
If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
library(plyr); library(dplyr)
------------------------------------------------------------------------------
-------

Attaching package: ‘plyr’

The following object is masked from ‘package:purrr’:

    compact

The following objects are masked from ‘package:dplyr’:

    arrange, count, desc, failwith, id, mutate, rename, summarise, summarize

Warning message:
package ‘plyr’ was built under R version 4.0.5
> levels(Data$Status)
NULL
>  Data$Status_New<-revalue(Data$Status, c(Pending = "Open", Solved = "Closed"))
>
```

```
> >  Data$Status_New<-revalue(Data$Status, c(Pending = "Open", Solved = "Closed"))


> > >
> Data$Status_New<-revalue(Data$Status, c(Pending = "Open", Solved = "Closed"))
> head(Data)
  Ticket..                                                  Customer.Complaint       Date
 Time
1   250635                                      Comcast Cable Internet Speeds 22-04-2015  3:53:
50 PM
2   223441                        Payment disappear - service got disconnected 04-08-2015 10:22:
56 AM
3   242732                                                 Speed and Service 18-04-2015  9:55:
47 AM
4   277946 Comcast Imposed a New Usage Cap of 300GB that punishes streaming. 05-07-2015 11:59:
35 AM
5   307175                           Comcast not working and no service to boot 26-05-2015  1:25:
26 PM
6   338519          ISP Charging for arbitrary data limits with overage fees 06-12-2015  9:59:
40 PM
        Received.Via     City    State Zip.code Status Filing.on.Behalf.of.Someone Status_New
1 Customer Care Call Abingdon Maryland    21009 Closed                          No     Closed
2           Internet  Acworth  Georgia    30102 Closed                          No     Closed
3           Internet  Acworth  Georgia    30101 Closed                         Yes     Closed
4           Internet  Acworth  Georgia    30101   Open                         Yes       Open
5           Internet  Acworth  Georgia    30101 Solved                          No     Closed
6           Internet  Acworth  Georgia    30101 Solved                          No     Closed
> levels(Data$State)
NULL
> unique(Data$State)
 [1] "Maryland"              "Georgia"               "Michigan"              "California"
   "New Mexico"            "Indiana"               "Virginia"              "Illinois"
 [9] "Pennsylvania"          "Massachusetts"         "Oregon"                "Texas"
   "New Hampshire"         "Minnesota"             "Tennessee"             "Colorado"
[17] "Florida"               "Alabama"               "Washington"            "New York"
   "New Jersey"            "Maine"                 "Missouri"              "West Virginia"
[25] "Montana"               "Mississippi"           "Connecticut"           "Vermont"
   "Kentucky"              "South Carolina"        "Ohio"                  "Utah"
[33] "Delaware"              "Arkansas"              "Nevada"                "Louisiana"
   "Kansas"                "Arizona"               "North Carolina"        "Rhode Island"
[41] "District Of Columbia" "District of Columbia" "Iowa"
> unique(Data$Status)
[1] "Closed"  "Open"    "Solved"  "Pending"
> tab <- table(Data$State,Data$Status_New)
> tab <- cbind(tab, Total = rowSums(tab))
> head(tab,15)
                     Closed Open Total
Alabama                  17    9    26
Arizona                  14    6    20
Arkansas                  6    0     6
California              159   61   220
Colorado                 58   22    80
Connecticut               9    3    12
Delaware                  8    4    12
District of Columbia      1    0     1
District Of Columbia     14    2    16
Florida                 201   39   240
Georgia                 208   80   288
Illinois                135   29   164
Indiana                  50    9    59
Iowa                      1    0     1
Kansas                    1    1     2
> library(gridExtra)

Attaching package: 'gridExtra'

The following object is masked from 'package:dplyr':

    combine

Warning message:
package 'gridExtra' was built under R version 4.0.5
> ggplot(Data, aes(y = State)) + geom_bar(aes(fill = Status_New))
```

```
> #CONCLUSION:Georgia and Florida are the Two where Comcast has a good number of Happy custome
rs by solving the issues in bulk
> levels(Data$Received.Via)
NULL
> unique(Data$Received.Via)
[1] "Customer Care Call" "Internet"
> ggplot(Data, aes(y = Received.Via )) + geom_bar(aes(fill = Status_New))
> df1 <- table(Data$Received.Via, Data$Status_New)
> df1 <- cbind(df1, Total = rowSums(df1))
> df1
                   Closed Open Total
Customer Care Call    864  255  1119
Internet              843  262  1105
> # Pie Chart with Percentages
> slices <- c(864, 255)
> lbls <- c("Closed", "Open")
> pct <- round(slices/sum(slices)*100)
> lbls <- paste(lbls, pct) # add percents to labels
> lbls <- paste(lbls,"%",sep="") # ad % to labels
> pie(slices,labels = lbls, col=rainbow(length(lbls)),
+    main="Pie Chart of Received Via Call")
> # Pie Chart with Percentages
> slices <- c(843, 262)
> lbls <- c("Closed", "Open")
> pct <- round(slices/sum(slices)*100)
> lbls <- paste(lbls, pct) # add percents to labels
> lbls <- paste(lbls,"%",sep="") # ad % to labels
> pie(slices,labels = lbls, col=rainbow(length(lbls)),
+    main="Pie Chart of Received Via Internet")
> #CONCLUSIONS:
> #1.The Company should Focus more on resolving complaints - Customer Are Mainly complaining a
bout the Data Caps, Internet Speed, Billing Methods and Services that Comcast is Providing and
 Very few Cases were registered against Comcast Cable Services.
> #2.In Georgia and Florida company services are already Improving but, in States - California
, Colorado and Illinois company should extend their resources in terms of the above-mentioned
issues in order to improve their customer servicing.
> #3.During the month of June and the start of July, the Company reported lots of complaints,
so as to for future reference they can keep this in check already so as to provide better serv
ices during these months. While working with their BPO clients to extend the staff during such
 days.
> save.image("C:\\Users\\MONALISH\\Desktop\\Comcast Telecom Complaints_Project_MonalisaPati")
> #Which ensures proper feedback for the particular arisen issue.
```