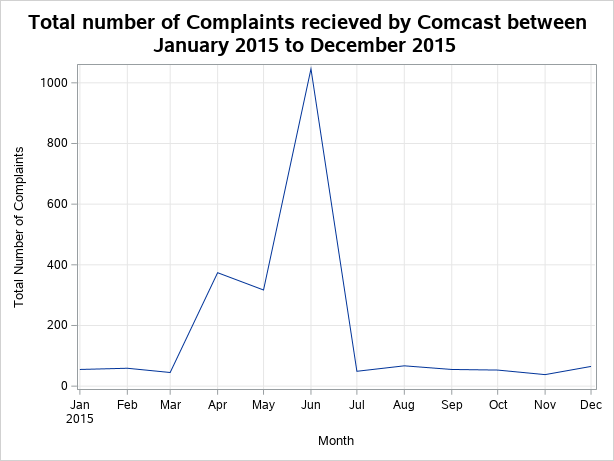
**Project 1: Comcast Telecom Consumer Complaints**

Code to import data set to SAS:

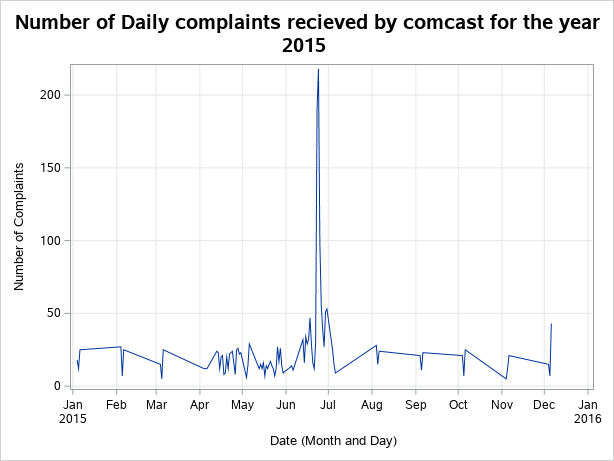
FILENAME REFFILE '/folders/myshortcuts/MyFolders/projects/Comcast\_telecom\_complaints\_data.csv';  
  
PROC IMPORT DATAFILE=REFFILE  
DBMS=CSV  
OUT=WORK.com;  
GETNAMES=YES;  
RUN;

**(1) Provide the trend chart for the number of complaints at monthly and daily granularity levels.**

Figure 1 and 2 are trend charts showing the number of complaints received by comcast against time. From figure 1 which describes the number of complaints at a monthly granularity level, it is observed June 2015 is the month that comcast received the greatest number of complaints from their customers, over 1000 complaints were received. Followed by April with just under 400 complaints. Throughout the year 2015 the number of complaints ranged between 0 and 100, increased in march to June, in July it fell back to within 0 and 100 complaints. Figure 2 represents the same information, just at a different granularity (Daily).



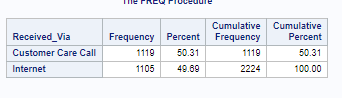
**Figure1: Series plot showing total number of complaints received by Comcast on a monthly basis from Jan-Dec 2015.**



**Figure2: series plot showing total number of daily complaints received by comcast for the year 2015.**

**(2) Provide a table with the frequency of complaint types.**

Table1: shows frequency of each complain type and also the Percent and cumulative frequency.



The frequency of complaint type does not vary much . 50.31 % of complaints were received via Customer Care call and 49.69% were received via internet. The difference between the two types is small.

**(3) 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.**

The new categorical variable was created and given a name of New\_status.

**(4) Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on:**

Figure 3 shows the number of complaints based on the state that the customer. From this bar chart it is observed that the state with most complaint was the state of Georgia, with over 250 complaints. Followed by the state of Florida and California. the state with the highest percentage of unresolved cases is Georgia, with 15.47%.

We can also use the code:

proc sort data=new2;  
by new\_status;  
run;  
  
proc freq data=new2 ;  
table state;  
by new\_status;  
run;

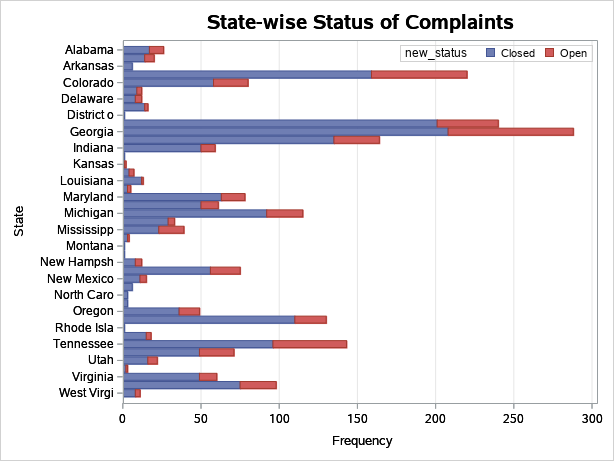
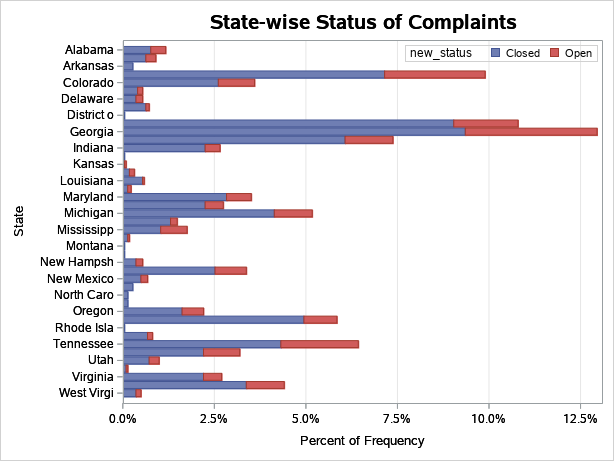


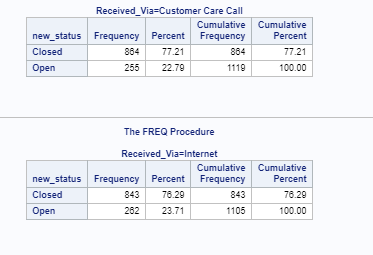
Figure3: Stacked bar graph showing status Complaints according to the state of the complainer.



**(5)Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.**

Table2: showing the percentages of complaints resolved and the method they were received via.

77.21% of complaints received via Customer care call were ressolved and 76.29 of complaints received via internet were ressolved.



CODE:  
/\* Exporting data into sas \*/  
FILENAME REFFILE '/folders/myshortcuts/MyFolders/projects/Comcast\_telecom\_complaints\_data.csv';  
  
PROC IMPORT DATAFILE=REFFILE  
DBMS=CSV  
OUT=WORK.com;  
GETNAMES=YES;  
RUN;  
  
/\* creating new variable to count complaints based on the Ticket\_\_ variable \*/  
  
data new;  
set com;  
complaint=(Ticket\_\_/Ticket\_\_);  
run;  
  
/\* sorting data for timeseries step \*/  
proc sort data=new;  
by Date complaint;  
run;  
  
/\* perfoming time series at a monthly granularity level \*/  
  
proc timeseries data=new  
out=Timedset;  
Id date interval=month accumulate=total;  
var complaint;  
run;  
  
/\* perfoming time series at a daily granularity level \*/  
  
proc timeseries data=new  
out=Timeday;  
Id date interval=day accumulate=total;  
var complaint;  
run;  
  
proc print data=new;  
run;

/\* creating table showing which copmplaint type had a greater frequency \*/  
  
proc freq data=new;  
table Received\_via;  
run;  
  
/\* creating new variable to count complaints based on the Ticket\_\_ variable \*/

proc print data=new;  
run;  
  
  
  
/\* 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. \*/  
data new2;  
set com;  
length new\_status $8.;  
if status= "Open" or status="Pending" then New\_status="Open";  
else New\_status="Closed";  
run;  
  
/\* \*/  
/\* Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on: \*/  
/\* Which state has the maximum complaints \*/  
/\* Which state has the highest percentage of unresolved complaints \*/  
  
ods graphics / reset width=6.4in height=4.8in imagemap;  
  
proc sgplot data=WORK.NEW2;  
title height=14pt "State-wise Status of Complaints";  
hbar State / group=new\_status groupdisplay=stack datalabel;  
xaxis grid;  
keylegend / location=inside;  
run;  
  
ods graphics / reset;  
title;  
  
  
  
  
/\* code for pecernt graph \*/  
ods graphics / reset width=6.4in height=4.8in imagemap;  
  
proc sgplot data=WORK.NEW2;  
title height=14pt "State-wise Status of Complaints";  
hbar State / group=new\_status groupdisplay=stack datalabel stat=percent;  
xaxis grid;  
keylegend / location=inside;  
run;  
  
ods graphics / reset;  
title;  
  
/\* Last question Provide the percentage of complaints resolved till date, which were received through the \*/  
/\* Internet and customer care calls. \*/  
  
proc sort data=new2;  
by received\_via new\_status;  
run;  
  
proc freq data=new2 ;  
table new\_status;  
by Received\_Via;  
run;