**Data Science with SAS :**

**Comcast Telecom Consumer Complaints**

DESCRIPTION

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.

**Data Dictionary**

* Ticket #: Ticket number assigned to each complaint
* Customer Complaint: Description of complaint
* Date: Date of complaint
* Time: Time of complaint
* Received Via: Mode of communication of the complaint
* City: Customer city
* State: Customer state
* Zipcode: Customer zip
* Status: Status of complaint
* Filing on behalf of someone

Analysis Task –

**1. Import data into SAS environment.**

PROC IMPORT DATAFILE=REFFILE

DBMS=XLSX

OUT=WORK.IMPORT;

GETNAMES=YES;

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

**data ansari.comcast;**

**set ansari.comcast;**

**complain\_year=year(Date\_month\_year);**

**complain\_month=month(Date\_month\_year);**

**complain\_ym=catx("-", of complain\_year complain\_month);**

**run;**

proc freq data= ansari.comcast;

tables complain\_year ;

run;

*Screenshot:*



*All records are from the year 2015*

roc freq data= ansari.comcast;

tables complain\_month ;

run;

*Screenshot:*



proc freq data= ansari.comcast;

tables complain\_ym;

run;

*Screenshot:*



ods graphics / reset width=8in height=3in imagemap;

proc sgplot data= ansari.comcast;

vline complain\_month /;

yaxis grid;

run;

*Screenshot:*



*Result - The month of June received the maximum number of complaints.*

*Screenshot:*



*Result - June 24 received the maximum number of complaints.*

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

**- Which complaint types are maximum i.e., around internet, network issues, or across any other domains.**

data ansari.comcast;

set ansari.comcast;

format customer\_complaint $60.;

customer\_complaint=upcase(Customer\_Complaint);

customer\_complaint\_v2=scan(customer\_complaint, 1, 2,' ');

WORD = TRANSLATE(customer\_complaint,' ',',.?!:&\*$%@#\_+-=/{}[]|()');

IF word in ('IS','ARE','WAS','WERE','THE','I','WE','THEY','THERE',

'THESE','IN','A','OF','TO','AND','IT','THAT','THIS', 'CAP', 'cap', 'caps', 'CAPS', 'COMCAST') THEN DELETE;

if find(customer\_complaint, "INTERNET")>0 then internet=1;

if find(customer\_complaint, "NETWORK")>0 then network=1;

if find(customer\_complaint, "PAYMENT" )>0 or

find(customer\_complaint, "BILLING")>0 then payment=1;

if find(customer\_complaint, "DATA")>0 then data=1;

run;

proc freq data= ansari.comcast noprint;

tables customer\_complaint /out=compaint\_types ;

run;

proc sort data =compaint\_types; by descending count;

run;

proc freq data= ansari.comcast noprint;

tables WORD / out=complaint\_types ;

run;

proc sort data =complaint\_types; by descending count;

run;

PROC SQL NOPRINT;

CREATE TABLE WORDLIST AS SELECT WORD AS WORD, COUNT(\*) AS COUNT FROM

ansari.comcast GROUP BY WORD ORDER BY COUNT DESCENDING;

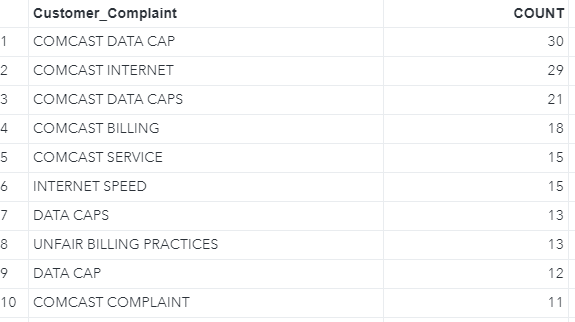
QUIT;

proc means data= ansari.comcast sum;

var internet network payment data;

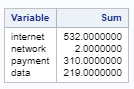
output out=total\_complaints;

*Screenshot:*



*Results - The top 10 types of complaints indicate that Data, Internet and Billing issues maybe the top complaint reasons****.***

*Screenshot:*



*Results - most common complaints are around: Internet, service, and payment related issues.*

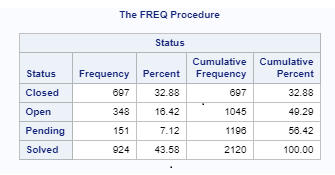
**4. Create a new categorical variable with value as Open and Closed. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed.**

proc freq data= ansari.comcast;

tables Status;

run;

*Screenshot:*



data ansari.comcast;

set ansari.comcast;

Format new\_status $12.;

if status in ("Open", "Pending") then new\_status="Open";

else if status in ("Closed", "Solved") then new\_status="Closed";

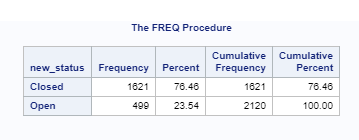
run;

proc freq data= ansari.comcast;

tables new\_status;

run;

*Screenshot:*



*Results - We have categorized the variable status into new categories. According to the new categories of status, around 76% of the tickets are closed.*

**5. 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=8in height=4.8in imagemap;

title 'States by status of the complaint (Frequency)';

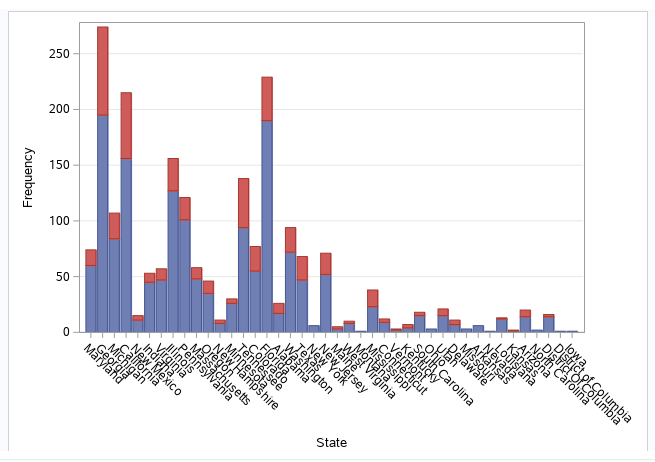
proc sgplot data= ansari.comcast;

vbar State / group=new\_status groupdisplay=stack DATALABELFITPOLICY=NONE;

yaxis grid discreteorder= data;

run;

*Screenshot:*



title 'States by status of the complaint (Proportion)';

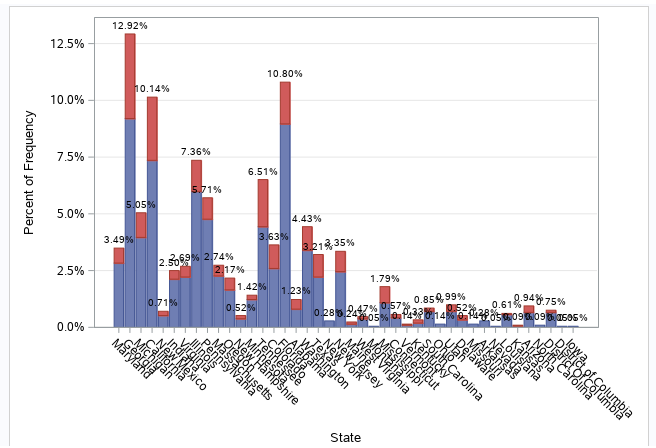
proc sgplot data=ansari.comcast;

vbar State / group=new\_status groupdisplay=stack datalabel stat=percent DATALABELFITPOLICY=NONE;

yaxis grid discreteorder= data ;

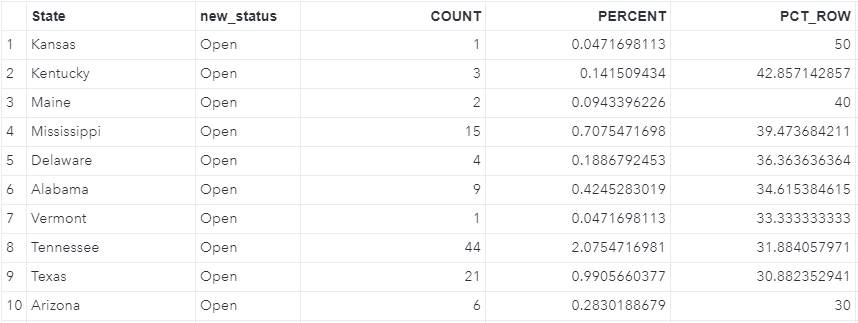
run;

*Screenshot:*



*Results - Georgia has the highest number/proportion of complaints, around 12.9% of complaints came from the state of Georgia.*

*Screenshot:*



***Results:***

*- Kansas has the highest number and proportion of unresolved cases.*

*- Around 50% of tickets originated from the state of Kansas are still open.*

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

proc freq data=ansari.comcast ;

tables received\_via ;

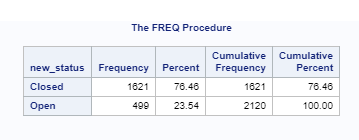
run;

proc freq data=ansari.comcast ;

tables new\_status / out=complains\_status ;

run;

*Screenshot:*



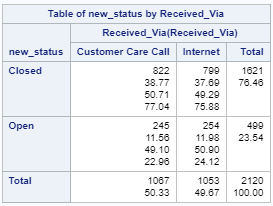
Results - In the total of 2,120 tickets, around 1,621 (77%) tickets are closed.

proc freq data=ansari.comcast ;

tables new\_status\*received\_via ;

run;

*Screenshot:*



***Results:***

*- Resolution rate is 76% for the complaints received through Internet and customer care calls*

*- In the total of 1,067 tickets received via Customer Care Call, around 76% (n=822) were resolved.*