```
libname ban130 "/home/u63044324/BAN 130/Project";
PROC IMPORT DATAFILE='/home/u63044324/BAN 130/Project/master.csv'
    OUT=master
    DBMS=CSV REPLACE;
RUN;

PROC PRINT DATA=master(obs=50);
RUN;

*Identification of all the variables and thier types;

PROC CONTENTS DATA=master OUT=master_contents;
RUN;

PROC PRINT DATA=master_contents;
    VAR NAME TYPE LENGTH;
RUN;

*Calculating mean and median of the suicide number variable;

PROC MEANS DATA=master;
    VAR suicides_no;
    OUTPUT OUT=stats MEAN=mean_mediansuicides_no MEDIAN=median_suicides_no;
RUN;

PROC PRINT DATA=stats;
    VAR mean_mediansuicides_no median_suicides_no;
RUN;

* calculating the average rate of suicide across different countries thier graphical representation;

/* Import master.csv dataset */
proc import datafile="/home/u63044324/BAN 130/Project/master.csv" out=master
    dbms=csv replace;
    getnames=yes;
run;


/* Calculate average suicide count */
proc means data=master noprint;
    var suicides_no;
    output out=suicide_count_mean
        mean=suicide_count;
run;

/* Print average suicide count */
proc print data=suicide_count_mean;
    title 'Average Suicide Count';
run;

*lets calculate correlation coeficient between suicide number and year variables;

/* Import the dataset */
proc import datafile="/home/u63044324/BAN 130/Project/master.csv"
            out=master
            dbms=csv
            replace;
            getnames=yes;
run;

/* Calculate correlation coefficients */
proc corr data=master;
```

```
    var suicides_no year;
run;

* Lets see the demographic analysis;


/* create a new dataset that groups the data by age group and gender */
proc sql;
create table suicide_by_demographics as
select age, sex, sum(suicides_no) as total_suicides
from master
group by age, sex;
quit;

/* create a bar chart of the total suicides by age group and gender */
proc sgplot data=suicide_by_demographics;
vbar age / group=sex response=total_suicides groupdisplay=cluster;
xaxis display=(nolabel);
yaxis grid;
run;


/* create a new dataset that groups the data by age group and gender */
proc sql;
create table suicide_by_demographics as
select age, sex, sum(suicides_no) as total_suicides
from master
group by age, sex;
quit;

/* create a bar chart of the total suicides by age group and gender */
proc sgplot data=suicide_by_demographics;
title 'Total Suicides by Age Group and Gender';
vbar age / group=sex response=total_suicides groupdisplay=cluster;
xaxis display=(nolabel);
yaxis grid;
run;

/* create a scatter plot of population vs suicides_no */
proc sgplot data=master;
title 'Scatter Plot of Population vs Suicides Number by Sex';
scatter x=population y=suicides_no / group=sex markerattrs=(symbol=circlefilled);
xaxis label='Population' grid;
yaxis label='Suicides Number' grid;
run;

/* create a box plot of suicides_no by sex */
proc sgplot data=master;
title 'Box Plot of Suicides Number by Sex';
vbox suicides_no / category=sex;
xaxis label='Sex' display=(nolabel);
yaxis label='Suicides Number' grid;
run;

*top 10 countries with highest suicice rates;

proc sort data=suicide_rates;
   by descending suicide_rate;
run;

data top_10;
   set suicide_rates (obs=10);
run;
```

```
proc print data=top_10 noobs;
   var country suicide_rate;
   title 'Top 10 Countries with the Highest Suicide Rates';
run;
```