```
/* Total Code */
/* 2 */
/* importing the data to create a temp library file to work on */
PROC IMPORT OUT=WORK.import DATAFILE="/home/u60766313/my_shared_file_links/u50396654/citibike-tripdata.xlsx"
        DBMS=XLSX REPLACE;
    GETNAMES=YES;
run;

/* 2i doing an sql procedure step to delete end station id and end station name that are null */
PROC SQL;
    DELETE FROM import where (end_station_id='' AND end_station_name='');
QUIT;

/* 2ii formating data to have different columns for date and time */
Data CitiBike;
    set import;
    format started_at DATETIME16.;
    format ended_at DATETIME16.;
    Startdate=datepart(started_at);
    Starttime=timepart(started_at);
    Enddate=datepart(ended_at);
    Endtime=timepart(ended_at);
    Format Startdate mmddyy10.;
    Format Starttime TIME8.;
    Format Enddate mmddyy10.;
    Format Endtime TIME8.;

    /* doing if statement to have a new column that displays evening morning and afternoon */
    if Starttime <'12:00't then
        Starttimes='Morning';
    else if Starttime <'18:00't then
        Starttimes='Afternoon';
    else
        Starttimes='Evening';

    if Endtime <'12:00't then
        EndTimes='Morning';
    else if Endtime <'18:00't then
```

```
        EndTimes='Afternoon';
    else
        EndTimes='Evening';
run;

/* 3a */
ods graphics / reset width=10in height=6in imagemap;
Title "3a";

/* Finding out the frequency in the data of start stations */
Title2 "Frequency for Start Station Name";

proc freq data=work.citibike order=freq;
    tables start_station_name / plots=freqplot(twoway=stacked orient=horizontal);
run;

Title;
ods graphics / reset;

/* 3b */
ods graphics / reset width=10in height=6in imagemap;
Title "3b";

/* Finding out the frequency in the data of start stations */
Title2 "Frequency for end Station Name";

proc freq data=work.citibike order=freq;
    tables end_station_name / plots=freqplot(twoway=stacked orient=horizontal);
run;

Title;
ods graphics / reset;

/* 3c */
ods graphics / reset width=10in height=6in imagemap;
Title "3c";

/* Finding out the frequency in the data of customers/member_casual */
Title2 "Frequency for Member_Casual";
```

```
proc freq data=work.citibike order=freq;
    tables member_casual / plots=freqplot(twoway=stacked orient=vertical);
run;

Title;
ods graphics / reset;

/* 3d */
Title "3d";

/* Finding out the difference with start time and end time in minutesa*/
data work.citibiketime;
    set work.citibike;
    DurationMin=intck('minute', starttime, endtime);
run;

ods graphics / reset width=10in height=6in imagemap;

/* Finding out the frequency in the data of Duration between start and end time in minutes*/
Title2 "Frequency for DurationMin";

proc freq data=work.citibiketime order=freq;
    tables DurationMin / plots=freqplot(twoway=stacked orient=horizontal);
run;

Title;
ods graphics / reset;

/* 3e */
ods graphics / reset width=10in height=6in imagemap;
Title "3e";

/* Ploting the x= start lat and y = start lng using sgplot with scatter */
Title2 "x= start lat and y = start lng";

proc sgplot data=work.citibike;
    scatter x=start_lat y=start_lng;
run;
```

```
Title;
Title "3e";

/* Ploting the x= End lat and y = End lng using sgplot with scatter */
Title2 "x= End lat and y = End ln";
```

```
proc sgplot data=work.citibike;
    scatter x=end_lat y=end_lng;
run;

Title;
Title "3e";

/* Using gmap to create map of newyork and its counties */
Title2 "Map of newyork and its counties";
pattern c=black v=e r=62;
```

```
proc gmap data=maps.counties map=maps.counties;
    id county;
    choro county / nolegend;
    where state eq 36;
    run;
    Title;
    Title "3e";

    /* Using sgplot to scatter start lat and start lng */
    Title2 "X= start lat and y = start lng";
```

```
proc sgplot data=work.citibike noborder noautolegend;
    polygon x=Start_lat y=Start_lng id=Start_Station_name / fill outline tip=none
        lineattrs=(color=gray99) fillattrs=(color=cxe8edd5);
    scatter x=Start_lat y=Start_lng / datalabel=Start_Station_ID
        markerattrs=(symbol=circlefilled size=13) markerfillattrs=(color=yellow)
        markeroutlineattrs=(color=purple);
    xaxis display=none;
    yaxis display=none;
run;
```

```
Title;
Title "3e";

/* Using sgplot to scatter end lat and end lng */
Title2 "x= end lat and y = end lng";
```

```
proc sgplot data=work.citibike noborder noautolegend;
    polygon x=end_lat y=end_lng id=End_Station_name / fill outline tip=none
        lineattrs=(color=gray99) fillattrs=(color=cxe8edd5);
    scatter x=end_lat y=end_lng / datalabel=Start_Station_ID
        markerattrs=(symbol=circlefilled size=13) markerfillattrs=(color=yellow)
        markeroutlineattrs=(color=purple);
    xaxis display=none;
    yaxis display=none;
run;

Title;
Title "3e";

/* Ploting the x= End lat and y = end lng using statgraph with scatter */
Title2 "x= End lat and y = end lng";
```

```
proc template;
    define statgraph classscatter;
        begingraph;
        entrytitle 'End Lat and End Lang';
        layout overlay /;
        scatterplot y=end_lat x=end_lng / datalabel=Start_Station_ID
            markerattrs=(symbol=circlefilled color=black size=3px);
        endlayout;
        endgraph;
    end;
run;
```

```
proc sgrender data=work.citibike template=classscatter;
run;

Title;
ods graphics / reset;
```

```
/* 3f */
ods graphics / reset width=10in height=6in imagemap;
Title "3f";

/* Sorting out the data to remove duplicate data in the dataset*/
proc sort data=work.citibike;
    by member_casual;
run;

/* Finding the Frequency for start_station_name by member_casual = member*/
Title2 "Frequency for start_station_name by member";

proc freq data=work.citibike order=freq;
    tables start_station_name / out=CustomersMem plots=freqplot(twoway=stacked
        orient=horizontal);
    by member_casual;
    where (member_casual='member');
run;

Title;
ods graphics / reset;

/* 3g */
ods graphics / reset width=10in height=6in imagemap;
Title "3g";

/* Sorting out the data to remove duplicate data in the dataset*/
proc sort data=work.citibike;
    by member_casual;
run;

/* Finding the Frequency for start_station_name by member_casual = Casual*/
Title2 "Frequency for start_station_name by Casual";

proc freq data=work.citibike order=freq;
    tables start_station_name / out=CustomersCas plots=freqplot(twoway=stacked
        orient=horizontal);
    by member_casual;
```

```
    where (member_casual='casual');
run;

Title;
ods graphics / reset;

/* 3h */
Title "3h";
Title2 "Member Customers commonly return or dock bikes";

/* To enable coloring theme for evening, morning and afternoon */
proc format;
    value $ EndingTimes "Evening"="RED" "Morning"="BLUE" "Afternoon"="SNOW";
run;

/* To eliminate duplicate records and have a new dataset accordingly */
proc sort data=work.citibike;
    by member_casual;
    where (member_casual='member');

    /* To find the frequency of the Enddate where member_casual = member  */
proc freq data=work.citibike order=freq;
    tables EndDate / out=EndDate plots=freqplot(twoway=stacked orient=horizontal);
    by member_casual;
    where (member_casual='member');
run;

Title "3h";

/* To find the frequency of the Endtimes where member_casual = member  */
title2 'Frequency for ending times';

proc freq data=work.citibike order=freq;
    tables EndTimes / out=EndTimes plots=freqplot(twoway=stacked
        orient=horizontal);
    by member_casual;
    where (member_casual='member');
run;
```

```
Title "3h";

/* To find the density of the Enddate where member_casual = member  */
title2 'Histogram for Ending Dates';

proc sgplot data=work.citibike;
    histogram EndDate;
    density EndDate;
run;

/* 3i */
ods graphics / reset width=10in height=6in imagemap;
Title "3i";
```