Kavya – LA County Covid-19 Seasonality

/\* I combined all nine months of data in an excel file beforehand and then imported it \*/

FILENAME LA '/home/u45124467/sasuser.v94/Broadstreet/LA.xlsx';

**PROC** **IMPORT** DATAFILE=LA

DBMS=XLSX

OUT=WORK.LA;

GETNAMES=YES;

**RUN**;

/\*delete missing variables \*/

**data** work.LA; set work.LA;

if Case\_Count = **0** or Case\_Count = "" then delete;

**run**;

/\*create a new dataset with calculated covid case rates \*/

**data** LA\_Rates; set LA;

if categories = 'white' then Case\_Rate = (Case\_Count/(**.707**\***10039107**));

else if categories = 'black\_african\_american' then Case\_Rate = (Case\_Count/(**.09**\***10039107**));

else if categories = 'asian' then Case\_Rate = (Case\_Count/(**.154**\***10039107**));

else if categories = 'american\_indian\_alaska\_native' then Case\_Rate = (Case\_Count/(**.014**\***10039107**));

else if categories = 'hispanic\_(all\_races)' then Case\_Rate = (Case\_Count/(**48.6**\***10039107**));

else if categories = 'native\_hawaiian\_pacific\_islander' then Case\_Rate = (Case\_Count/(**.004**\***10039107**));

else if categories = 'cumulative\_cases' then Case\_Rate = (Case\_Count/**10039107**);

**run**;

/\*plot the graph \*/

title "Covid-19 Case Rates for Los Angeles by Race";

**proc** **sgplot** data=LA\_Rates;

series x=dates y=Case\_rate/ group = categories;

title;

/\*create separate datasets for each race to look at them individually \*/

**data** White; set LA\_Rates;

where categories = 'white';

**run**;

**data** Black; set LA\_Rates;

where categories = 'black\_african\_american';

**run**;

**data** Asian; set LA\_Rates;

where categories = 'asian';

**run**;

**data** AIAN; set LA\_Rates;

where categories = 'american\_indian\_alaska\_native';

**run**;

**data** Hispanic; set LA\_Rates;

where categories = 'hispanic\_(all\_races)';

**run**;

**data** NHPI; set LA\_Rates;

where categories = 'native\_hawaiian\_pacific\_islander';

**run**;

**data** Total; set LA\_Rates;

where categories = 'cumulative\_cases';

**run**;

/\* Plot each race's graphs, grouped by season so we can see the trends for each season\*/

title "White Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=White;

series x=dates y=Case\_rate/ group = season;

title;

title "Black Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=Black;

series x=dates y=Case\_rate/ group = season;

title;

title "Asian Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data= Asian;

series x=dates y=Case\_rate/ group = season;

title;

title "AIAN Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=AIAN;

series x=dates y=Case\_rate/ group = season;

title;

title "Hispanic Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=Hispanic;

series x=dates y=Case\_rate/ group = season;

title;

title "NHPI Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=NHPI;

series x=dates y=Case\_rate/ group = season;

title;

title "Total Covid-19 Case Rate for Los Angeles";

**proc** **sgplot** data=Total;

series x=dates y=Case\_rate/ group = season;

title;

/\* for the final bar graphs, looking the last date of each season's case rates to compare \*/

/\* putting the date in ddmmmyyyy format followed by a d in apostrophes allows SAS to read it as a date and not a word \*/

**data** LA\_Summer; set LA\_Rates;

where dates = **'31AUG2020'd**;

if Case\_Rate = '.' then delete;

if categories = 'cumulative\_cases' then delete;

**run**;

**data** LA\_Fall; set LA\_Rates;

where dates = **'30NOV2020'd**;

if Case\_Rate = '.' then delete;

if categories = 'cumulative\_cases' then delete;

**run**;

**data** LA\_Winter; set LA\_Rates;

where dates = **'28FEB2021'd**;

if Case\_Rate = '.' then delete;

if categories = 'cumulative\_cases' then delete;

**run**;

/\* Plot each season's bar graph. categoryorder=respdesc orders the bar graphs so that they are descending from largest to smallest. You can play around with colormodel statement, which I used to pick colors according to the seasons\*/

title 'COVID-19 Case Rates for August 31 2020';

**proc** **sgplot** data=LA\_Summer;

vbar categories / response=Case\_Rate

stat=mean CATEGORYORDER=RespDesc colorresponse=Case\_Rate

colormodel= (greenyellow Green);

xaxis display=(nolabel noticks);

yaxis grid;

**run**;

title 'COVID-19 Case Rates for November 30 2020';

**proc** **sgplot** data=LA\_Fall;

vbar categories / response=Case\_Rate

stat=mean CATEGORYORDER=RespDesc colorresponse=Case\_Rate

colormodel= (yellow red);

xaxis display=(nolabel noticks);

yaxis grid;

**run**;

title 'COVID-19 Case Rates for February 28 2020';

**proc** **sgplot** data=LA\_Winter;

vbar categories / response=Case\_Rate

stat=mean CATEGORYORDER=RespDesc colorresponse=Case\_Rate

colormodel= (AliceBlue Blue);

xaxis display=(nolabel noticks);

yaxis grid;

**run**;