/\*load datasets frtom xpt files\*/

LIBNAME BRFSS19 xport 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\LLCP2019.xpt' access=temp;

LIBNAME BRFSS18 xport 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\LLCP2018.xpt' access=temp;

LIBNAME BRFSS20 xport 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\LLCP2020.xpt' access=temp;

LIBNAME BRFSS21 xport 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\LLCP2021.xpt' access=temp;

/\*define local folder to store data \*/

LIBNAME ADA 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment';

/\*rename data, add ID, year pandemic info for four years data\*/

/\*some of the data need to be renamed or re-defined\*/

DATA ADA.brfss2019;

SET brfss19.LLCP2019;

retain year 2019;

retain pandemic 0;

retain ID 2019000000;

ID = ID+1;

RUN;

DATA ADA.brfss2018;

SET brfss18.LLCP2018;

retain year 2018;

retain pandemic 0;

retain ID 2018000000;

ID = ID+1;

/\*rename sex variable to fit with others\*/

\_SEX = SEX1;

RUN;

DATA ADA.brfss2020;

SET brfss20.LLCP2020;

retain year 2020;

retain pandemic 1;

retain ID 2020000000;

ID = ID+1;

RUN;

DATA ADA.brfss2021;

SET brfss21.LLCP2021;

retain year 2021;

retain pandemic 1;

retain ID 2021000000;

ID = ID+1;

RUN;

ods graphics on;

/\*output of html\*/

ods html path='C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\tables and figures' body='tables and figures.html';

ods noproctitle;

/\* combine four years data by column, with AGE RACE SMOKING SEX INCOME ALCOHOL DRINK only\*/

/\* Column-based data combining is not needed to sort \*/

data combined;

set ada.brfss2018 ada.brfss2019 ada.brfss2020 ada.brfss2021;

keep ID year pandemic menthlth \_race DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR;

run;

/\*n OF NOW\*/

proc report data=combined;

COLUMN N;

RUN;

data combined;

set combined;

/\* re-code race\*/

if \_race = 1 then race\_new = 1;

if \_race = 2 then race\_new = 2;

if \_race = 4 then race\_new = 3;

if \_race = 8 then race\_new = 4;

if \_race = 3 or \_race = 5 or \_race = 6 or \_race = 7 then race\_new = 5;

if \_race = 9 then race\_new = .;

run;

/\* add format\*/

proc format;

value pandemic 0 = 'Pre-pandemic' 1 = 'Post-pandemic';

value race 1='White only, non-Hispanic'

2='Black only, non-Hispanic'

3='Asian only, non-Hispanic'

4='Hispanic'

5='Others, non-Hispanic';

run;

/\* data cleaning based on the codebook\*/

data combined;

set combined;

if menthlth = 77 or menthlth = 99 then menthlth = .;

if menthlth = 88 then menthlth = 0;

if DRNKANY5 GT 6 then DRNKANY5 = .;

if \_sex GT 3 then \_SEX =.;

if \_RFSMOK3 GT 3 then \_RFSMOK3=.;

if \_AGEG5YR = 14 THEN \_AGEG5YR = .;

run;

proc freq data=combined ;

table menthlth race\_new/missing;

run;

/\*filter out invalid mental health and race response\*/

data combined;

set combined;

if menthlth NE . and race\_new NE . and drnkany5 NE . and \_sex NE . and \_rfsmok3 NE . and \_ageg5yr NE .;

run;

proc report data=combined;

COLUMN N;

RUN;

proc export data=combined outfile='C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\tables and figures\combined.csv' dbms=csv replace;

run;

/\*descriptive\*/

ods rtf path= 'C:\Users\92175\OneDrive\Desktop\WUSTL fall22\ADA\Research Assignment\tables and figures' file=' tables and figures.rtf' nogtitle ;

ods noproctitle;

proc tabulate data = combined;

class pandemic DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR race\_new;

table (pandemic DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR),(N RACE\_NEW);

format race\_new race. pandemic pandemic.;

title 'table-2 Descriptives of mental health issue days per months and race';

run;

ods rtf close;

proc means data=combined N mean median var q1 q3;

var menthlth;

class race\_new;

format race\_new race.;

footnote '';

run;

/\*check variance of mental health for Poisson regression\*/

proc means data=combined n mean var min max;

var menthlth;

title 'Table-3 check variance for mental health';

footnote 'The variance is way larger than the mean.';

run;

/\*violates the preassumption\*/

/\*check variance by groups (before and post pandemic)\*/

proc means data = combined mean var;

class pandemic;

var menthlth;

title 'Table-4 check variance for mental health by pandemic';

footnote 'Still had crazy high of variance. Make a plot to see what is happening';

run;

/\*crazy high...\*/

/\*make a histogram to see what is happening\*/

proc freq data=combined;

table menthlth/plots=freqplot;

title 'Table-5 and figure-1 the distribution of mental health issue days per month';

footnote 'no way...... why so many people sad for the whole month...';

footnote2 'the distribution is clearly not perfect for Poisson regression. Let try negative binomial...';

run;

/\*switch to negative binomial reg\*/

/\* model 1: unadjusted \*/

proc genmod data=combined;

class pandemic race\_new(param=ref ref = first); /\*specifies the categorical variables and reference factors (the lowest number in this case\*/

model menthlth = race\_new pandemic/type3 dist=negbin; /\*Model : negative binomial\*/

title 'Unadjusted model';

estimate 'post pandemic' pandemic -1 / exp;

estimate 'black' race\_new 1 0 0 0 / exp;

estimate 'asian' race\_new 0 1 0 0/ exp;

estimate 'hispanic' race\_new 0 0 1 0/ exp;

estimate 'other' race\_new 0 0 0 1/ exp;

footnote '';

run;

/\*Negative binomial regression without adding interaction between race and pandemic year\*/

proc genmod data=combined;

class pandemic race\_new DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR(param=ref ref = first); /\*specifies the categorical variables and reference factors (the lowest number in this case\*/

model menthlth = race\_new pandemic DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR /type3 dist=negbin ; /\*Model : negative binomial\*/

title 'Adjusted without interaction';

estimate 'post pandemic' pandemic -1 / exp;

estimate 'black' race\_new 1 0 0 0 / exp;

estimate 'asian' race\_new 0 1 0 0/ exp;

estimate 'hispanic' race\_new 0 0 1 0/ exp;

estimate 'other' race\_new 0 0 0 1/ exp;

run;

/\*Negative binomial regression with adding interaction between race and pandemic year\*/

proc genmod data =combined;

class pandemic race\_new DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR(param=ref ref=first);

model menthlth = race\_new pandemic race\_new\*pandemic DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR /type3 dist=negbin;

/\*calcluate IRR\*/

estimate 'post pandemic' pandemic -1 / exp;

estimate 'white' race\_new 0 0 0 0 / exp;

estimate 'black' race\_new 1 0 0 0 / exp;

estimate 'asian' race\_new 0 1 0 0/ exp;

estimate 'hispanic' race\_new 0 0 1 0/ exp;

estimate 'other' race\_new 0 0 0 1/ exp;

title 'Adjusted with interaction';

run;

/\*test of the fitness of the model\*/

/\* seperate dataset\*/

data prepan duringpan;

set combined;

if pandemic = 0 then output prepan;

if pandemic = 1 then output duringpan;

run;

proc genmod data =prepan;

class race\_new DRNKANY5 \_SEX \_AGEG5YR(param=ref ref=first);

model menthlth = race\_new DRNKANY5 \_SEX \_RFSMOK3 \_AGEG5YR /type3 dist=negbin;

estimate 'white' race\_new 0 0 0 0 / exp;

estimate 'black' race\_new 1 0 0 0 / exp;

estimate 'asian' race\_new 0 1 0 0/ exp;

estimate 'hispanic' race\_new 0 0 1 0/ exp;

estimate 'other' race\_new 0 0 0 1/ exp;

title 'stratified pre-pan';

run;

proc genmod data =duringpan;

class race\_new DRNKANY5 \_SEX \_AGEG5YR(param=ref ref=first);

model menthlth = race\_new DRNKANY5 \_SEX \_AGEG5YR /type3 dist=negbin;

/\*calcluate IRR\*/

estimate 'white' race\_new 0 0 0 0 / exp;

estimate 'black' race\_new 1 0 0 0 / exp;

estimate 'asian' race\_new 0 1 0 0/ exp;

estimate 'hispanic' race\_new 0 0 1 0/ exp;

estimate 'other' race\_new 0 0 0 1/ exp;

title 'stratified during-pan';

run;