University of Kentucky

UKnowledge

Theses and Dissertations--Public Health (M.P.H. & Dr.P.H.)

College of Public Health

2019

Consultation Experience in Public Health

Zimu Li
University of Kentucky, Zimu.li@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/cph_etds

Part of the Public Health Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

STUDENT AGREEMENT:

I represent that my capstone and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's capstone including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

<div align="right">

Zimu Li, Student

Dr. Heather Bush, Committee Chair

Dr. Sarah Wackerbarth, Director of Graduate Studies

</div>

Consultation Experience in Public Health

Zimu Li
MPH-Biostatistics

Content

Project Description

Nowadays, students safe in violence area is becoming to a huge problems, from the survey- National Youth Risk Behavior Survey we find that the sexual and dating violence rate is increasing during the past few years.For the willing that we try to reduce the risk of dating and sexual violence, we are going to build some programs and publish them to the society especially the schools to protect our teenagers. As we know that the period for students from high school to the university is a very important time to build their right social norms, we start a program called Green Dot which aims to guide them following the right sexual and dating norms to let them away from these violence.This intervention trial is bystanding-based and randomized in 13 pairs of high schools in Kentucky. Lots of variables are added into this program such as race, gender these based information which may have high relating to the sexual and dating violence and they divide the degree of education of the Green Dot into five levels which is from zero intervention to full educating methods for the intervention.

This project aims to find that if the Green Dot education in high school affects the students' norms positively. We consider the Green Dot education as the exposure to do our survey. Then we choose the Project Time 1 as the unexposure part which the schools their have no Green Dot education forms. Moreover, we choose Project Time 5 as the exposure part which they have full of Green Dot education situation such as lecture form, speech form, post method, training form and etc. That means that the difference between the two parts are the exposure of Green Dot education situation. We hope to see that our program about Green Dot education and dating or sexual violence education can help our high school students build a great social norms and guide them to a right way to treat their partners.

Besides, we choose six different variables as gender, race, sexual attractive, free meal plan situation, current relationship and past 12 months relationship (dating) to be our adjusted variables. On the other hand, we use the two scores in our questionnaire to be our predictions. DV_ACC score and IRMS score. These two scores will reflect the students' opinions and attractions to dating violence and sexual violence. The lower this score, the higher they do not accept these two kinds of violence. Mainly, we try to figure out the relationship between the exposure variables (The Green Dot education) and these two scores, by the way, we try to find that if these adjusted variables will affect the relationship.

This project will help us know an important question: If our Green Dot education program is useful and meaningful to our students' norm building. Can we reduce the rate of sexual violence by using this intervention as Education or Media way? Nowadays, not only US has so many sexual violence cases, but also in China, the sexual violence situation is so serious that we have to take action to stop that. In US, we have already built the Green Dot program to prevent this bad situation by intervening at school steps. Can our China follow this intervention if the program is effective? That is the meaning of what I do right now.

My Role

My project was developed under Dr. Bush's supervision. All the database is provided from Green Dot datasets.

My roles:

1. Understanding the aims and targeted objectives in the project

2. Picking up the subgroups which I am interesting in

3. Doing the data clearing up and sort

4. Operating the data analysis which includes the descriptive analysis, comparing analysis, and logistic regression on SAS platform (SAS 9.4 SOFTWARE)

5. Using LS mean as a specific value to show the effect

6. Creating macro and format for analyzing data

7. Using proc report and macro to show the reports

Applicable SAS code

Format and Marco code

```
PROC FORMAT;
VALUE gender
0='Female'
1='Male'
other=' ';

value nonwhite
0='White'
1='Nonwhite'
other=' ';

value meal
0='No free meal'
1='free meal'
other=' ';

value sexappeal
0='Heterosexual'
1='not Heterosexual'
other=' ';

value current
0='no relationship now'
1='has relationship now'
other=' ';

value rela12mon
0='no relationship <12mon'
1='has relationship <12 mon'
other=' ';

RUN;
*Macro Definitions:
ds - dataset name
oneval - if you identify oneval as 0 in the macro call then a p-value won't be presented.  Otherwise, use a value of 1
for oneval
davr - this is the variable you are summarizing
ivar - this is your group variable
fmt - this is the format name for the variable you are summarizing.  For example, above I created a format called
gender so that is what I would enter in the macro call for gender.
        The format name does not have to be identical to the variable name.
final - dinal outcome dataset ;

%macro dofreq(ds,sv,oneval,dvar,ivar,fmt,final);
/*Get Total Frequency*/
proc sort data=&ds out=tmp;
by &dvar;
run;
ods output OneWayFreqs=FRQ1;
proc freq data=tmp(where=(&ivar > .));
```

```
    tables &dvar;
    run;

    data FRQ1;
    length col $20;
    set FRQ1;
    col=compress(put(frequency,8.))||' ('||compress(put(percent,8.1))||"%)";
    if &dvar > .;
    run;
    /*transpose the result table, sort first*/
    proc sort data=FRQ1;
    by &dvar;
    run;
    proc transpose data=FRQ1 out=TFRQ1;
    by &dvar;
    var col;
    run;
    /*Get Total Column Frequency*/
    ods output OneWayFreqs=FRQC;
    proc freq data=tmp;
    tables &ivar;
    run;
    ods output OneWayFreqs=FRQN;
    proc freq data=tmp;
    tables pat;
    run;
    data FRQC;
    length colT parm $250;
    set FRQN FRQC;
    colT=put(frequency,8.);
    if table="Table pat" then &ivar=-1;
    parm="Total";
    run;
    proc transpose data=FRQC out=TFRQC(rename=(N1=COL1));
    id &ivar;
    var colT;
    run;

    /*Get Frequency output*/
    /*Get General Association output*/
    ods output chisq=CHI;
    ods output crossTabFreqs=FRQS;
    proc freq data=tmp(where=(&ivar > .));
    tables &dvar*&ivar / chisq;
    run;
    data FRQS(where=(&dvar ne .));
    set FRQS(where=(&ivar > .));
    run;
    data FRQS;
    length col $20;
    set FRQS;
    *col=compress(put(frequency,8.))||' ('||compress(put(rowpercent,8.1))||"%)";

    col=compress(put(frequency,8.))||' ('||compress(put(colpercent,8.1))||"%)";
    if &ivar > . and &dvar > .;
    run;
    proc sort data=FRQs;
```

```
by &dvar;
run;
proc transpose data=FRQs out=TFRQs;
by &dvar;
id &ivar;
var col;
run;

data CHI(keep=parm2 parm pvalue sv dfr);
length parm $250 pvalue $30 parm2 $10 sv 8.;
set CHI(where=(statistic='Chi-Square') keep=prob statistic df);
parm="&dvar";
parm2="&ivar";
sv=0;
if prob > . then do;
if &oneval=0 then pvalue=" ";
else pvalue=put(prob,pvalue6.);
dfr="DF="||trim(left(put(df,best.)));
end;
run;
/*merge tables with pvalue and frequency (percent)*/
proc sort data=TFRQ1;
by &dvar;
run;
proc sort data=TFRQs;
by &dvar;
run;
data TFRQs;
merge TFRQ1 TFRQs;
by &dvar;
run;
/*merge tables with pvalue and frequency (percent)*/
data TFRQs(drop=&dvar _NAME_);
length parm $250;
set CHI TFRQs(where=(&dvar>= 0));
if sv=0 then parm="&dvar";
else parm=put(&dvar,&fmt.);
run;
/*cumulate tables*/
data &final;
length sv 8. pvalue $30 ;
set &final TFRQS(in=a);
if a then do;
end;
run;
%mend;

*Macro Definitions:
ds - dataset name
var - this is the variable you are summarizing
by - this is your group variable
label - this is the label that will be presented in the table
se - if you leave this blank,the standard deviation will be presented.  Otherwise, the standard error is presented.;
%macro uni(ds,var,by,label,se,final);

proc univariate data=&ds(where=(&by ne . )) noprint;
var &var;
```

```
output out=ds1 mean=mean std=std median=med q1=q1 q3=q3 min=min max=max n=n stdmean=stderr;
run;

proc sort data=&ds;
by &by;
run;

proc univariate data=&ds(where=(&by ne . )) noprint;
var &var;
by &by;
output out=ds2 mean=mean std=std median=med q1=q1 q3=q3 min=min max=max n=n stdmean=stderr;
run;

data alldata;
set ds1 ds2(where=(&by ne .));
if &by=. then &by=99;
run;

data alluni;
length var parm $100;
set alldata;

var=trim(left(n));
sv=1.1;
parm="&label., n";
output;
%if &se eq %then %do;
var=trim(left(round(mean,0.1)))||" ("||trim(left(round(std,0.01)))||")";
sv=1.2;
parm='Mean (SD)';
output;
%end;
%else %do;
var=trim(left(round(mean,0.1)))||" ("||trim(left(round(stderr,0.01)))||")";
sv=1.2;
parm='Mean (SE)';
output;
%end;
var=trim(left(round(Q1,0.1)))||", "||trim(left(round(med,0.1)))||", "||trim(left(round(Q3,0.1)));
sv=1.3;
parm='Q1, Median, Q3';
output;
var=trim(left(round(min,0.1)))||", "||trim(left(round(max,0.1)));
sv=1.4;
parm='Minimum, Maximum';
output;
run;

proc sort data=alluni;
by sv parm;
run;

proc transpose data=alluni out=tmp(drop=_name_ rename=(_99=col1));
var var;
by sv parm;
id &by;
run;
```

```
*Need to add proc ttest;
ods output ttests=ttest equality=equal;
proc ttest data=all(where=(&by ne . ));
class &by;
var &var;
run;

proc sort data=ttest;
by variable;
run;

proc sort data=equal;
by variable;
run;

data pval;
merge ttest equal;
by variable;

if probf<.05 then do;
  if variances='Equal' then delete;
end;

else if probf ge .05 then do;
  if variances='Unequal' then delete;
end;

run;

data tmp(drop=probt);
length pvalue $6;
if _n_=1 then set pval(keep=probt);
set tmp;
if parm in ('Mean (SD)','Mean (SE)','Q1, Median, Q3','Minimum, Maximum') then pvalue=' ';
else pvalue=put(probt,pvalue6.);
run;

data &final;
set &final tmp;
run;


%mend;
```

Analyzing and Reporting Code

```
OPTIONS NODATE MERGENOBY=WARN LS=79 PS=41 formchar="|----|+|---+=|-/\<>*=";
FOOTNOTE1 "SAS Job capstone run by ZIMU LI on &sysdate at &systime";

LIBNAME adata "E:\MPH\SAS\data";


data all;
set adata.gdhs_zli;
where time in(1,5);
pat=1;
run;

proc contents data=all;
run;

option nofmterr;
%global TOT GP1 GP2;

data final1 ;
stop;

%dofreq(all,1.1,1,GENDER1,time,gender.,final1);
%dofreq(all,1.1,1,NONWHITE,time,nonwhite.,final1);
%dofreq(all,1.1,1,F_MEAL1,time,meal.,final1);
%dofreq(all,1.1,1,NEXHET,time,sexappeal.,final1);
%dofreq(all,1.1,1,CURREL,time,current.,final1);
%dofreq(all,1.1,1,REL12YN,time,rela12mon.,final1);


ods rtf file="E:\MPH\capstone\reports.rtf";

proc report data=final1 headline headskip nowindows split='|';
columns sv parm _1 _5 pvalue;
define sv / display "sv";
define parm / display " ";
define _1 / display "&group1|(N=&N1)" right;
define _5 / display "&group2|(N=&N2)" right;
define pvalue / display "p-value" right;
compute sv;
if sv=0 then
call define(_row_,'style','style=[font_weight=bold]');
endcomp;
title " Green Dot Intervention and Demographics of Snapshot in Time Group";
footnote1 j=center height=9pt "Biostatistics Capstone--Life's Snapshot
Project. &SYSDATE9";
run;

ods rtf close;

data final11;
stop;

%dofreq(all,1.1,1,DV_ACC,time,rela12mon.,final11);
```

```
ods rtf file="E:\MPH\capstone\reports11.rtf";
proc report data=final11 headline headskip nowindows split='|';
columns sv parm _1 _5 pvalue;
define sv / display "sv";
define parm / display " ";
define _1 / display "&group1|(N=&N1)" right;
define _5 / display "&group2|(N=&N2)" right;
define pvalue / display "p-value" right;
compute sv;
if sv=0 then
call define(_row_,'style','style=[font_weight=bold]');
endcomp;
title " Green Dot Intervention and Demographics of Snapshot in Time Group";
footnote1 j=center height=9pt "Biostatistics Capstone--Life's Snapshot
Project. &SYSDATE9";
run;

ods rtf close;

data final12;
stop;


%dofreq(all,1.1,1,RAPEMYTH,time,rela12mon.,final12);

ods rtf file="E:\MPH\capstone\reports12.rtf";
proc report data=final12 headline headskip nowindows split='|';
columns sv parm _1 _5 pvalue;
define sv / display "sv";
define parm / display " ";
define _1 / display "&group1|(N=&N1)" right;
define _5 / display "&group2|(N=&N2)" right;
define pvalue / display "p-value" right;
compute sv;
if sv=0 then
call define(_row_,'style','style=[font_weight=bold]');
endcomp;
title " Green Dot Intervention and Demographics of Snapshot in Time Group";
footnote1 j=center height=9pt "Biostatistics Capstone--Life's Snapshot
Project. &SYSDATE9";
run;

ods rtf close;

data final2;
stop;

%uni(all,DV_ACC,time,adjusted and un,,final2);
%uni(all,RAPEMYTH,time,adjusted and un,,final2);

ods rtf file="E:\MPH\capstone\reports2.rtf";

proc report data=final2 headline headskip nowindows split='|';
columns sv parm col1 _1 _5 pvalue;
define sv / display "sv";
define parm / display " ";
```

```
define col1 / display "total" right;
define _1 / display "unadjusted|(N=)" right;
define _5 / display "adjusted|(N=)" right;
define pvalue / display "p-value" right;
compute sv;
if sv=0 then
call define(_row_,'style','style=[font_weight=bold]');
endcomp;
title " Green Dot Intervention and Demographics of Snapshot in Time Group";
footnote1 j=center height=9pt "Biostatistics Capstone--Life's Snapshot
Project. &SYSDATE9";
run;

ods rtf close;



ods rtf file="E:\MPH\capstone\reports3.rtf";

proc glm data=all;
class TIME;
model DV_ACC =TIME / solution;
lsmeans TIME / stderr pdiff cov;
 run;



proc glm data=all;
class TIME GENDER1 NONWHITE F_MEAL1 NEXHET CURREL REL12YN;
model DV_ACC =TIME GENDER1 NONWHITE F_MEAL1 NEXHET CURREL REL12YN / solution;
 lsmeans TIME / stderr pdiff cov;
 run;



proc glm data=all;
class TIME;
model RAPEMYTH =TIME / solution;
lsmeans TIME / stderr pdiff cov;
 run;

 proc glm data=all;
class TIME GENDER1 NONWHITE F_MEAL1 NEXHET CURREL REL12YN;
model RAPEMYTH =TIME GENDER1 NONWHITE F_MEAL1 NEXHET CURREL REL12YN / solution;
 lsmeans TIME / stderr pdiff cov;
 run;

ods rtf close;
```

Project Outcomes (Tables and Graphs)

Descriptive Tables and Graphs

Table 1. Total descriptive values for all variables

Variables		Unexposure (N=2540)	Exposure (N=2179)
Gender	Female	1387(54.6%)	1226(56.3%)
	Male	1153 (45.4%)	953 (43.7%)
Race	White	2058 (81.0%)	1803 (82.7%)
	Nonwhite	482 (19.0%)	376 (17.3%)
Free meal policy	No free meal	1424 (56.3%)	1071 (49.3%)
	free meal	1104 (43.7%)	1102 (50.7%)
Sexual attractive	Heterosexual	2232 (87.9%)	1866 (85.6%)
	not Heterosexual	308 (12.1%)	313 (14.4%)
Current relationship	no relationship now	1479 (58.4%)	1367 (63.0%)
	has relationship now	1054 (41.6%)	804 (37.0%)
Past relationship	no relationship <12mon	558 (22.0%)	690 (31.8%)
	has relationship <12 mon	1975 (78.0%)	1481 (68.2%)

Figure 1 : Graph show for Table 1
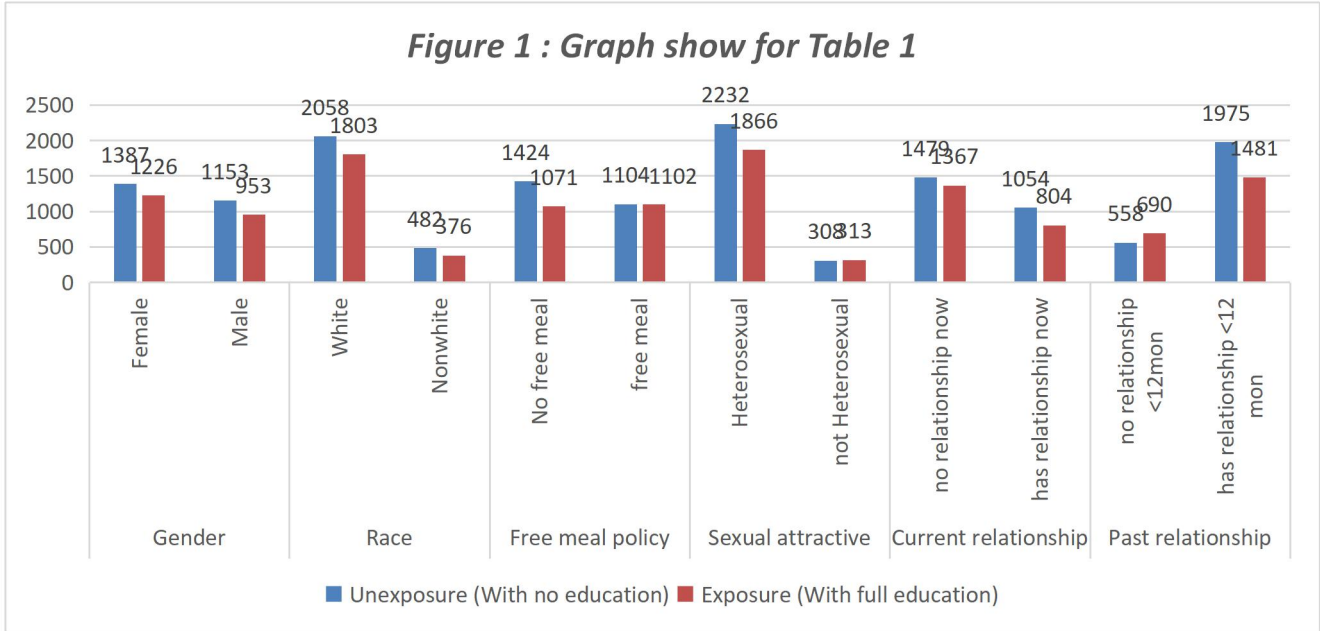
Table 1 is a total summary for all the 6 variables which includes gender, race, free meal plan situation, sexual preference, current and past 12 months relationship. It shows you the value of all the variables in unexposure and exposure situation. As we can see that, if we set the α =0.01, then the variables free meal policy, two kind of relationships will be significant which means that between the non and exposure situation, these three variables have the difference. However, because our survey is a survey which is absolutely relating to sex, race and sexual attractive. We still have to put these three variables in our model to find that if it can affect our final predictive scores.

In order to know the exact change clearly for different variables between two exposure situation, Table 2 is provided.

Note: Unexposure means the situation there is with no sexual and other kind of education.
        Exposure means the situation there is full of all Green Dot education about the norms.

Table 2. variables proportion change between two exposure situation

Variables	Unexposure	Exposure
Gender: Female	54.60%	56.30%
△= 1.70%		
Race: White	81.00%	82.70%
△= 1.70%		
No free meal policy	56.30%	49.30%
△= -7.00%		
Sexual attractive: Heterosexual	87.90%	85.60%
△= -2.30%		
No relationship now	58.40%	63.00%
△= 4.60%		
No past relationship in last year	22.00%	31.80%
△= 9.80%		

Figure 2 : Graph show for Table 2
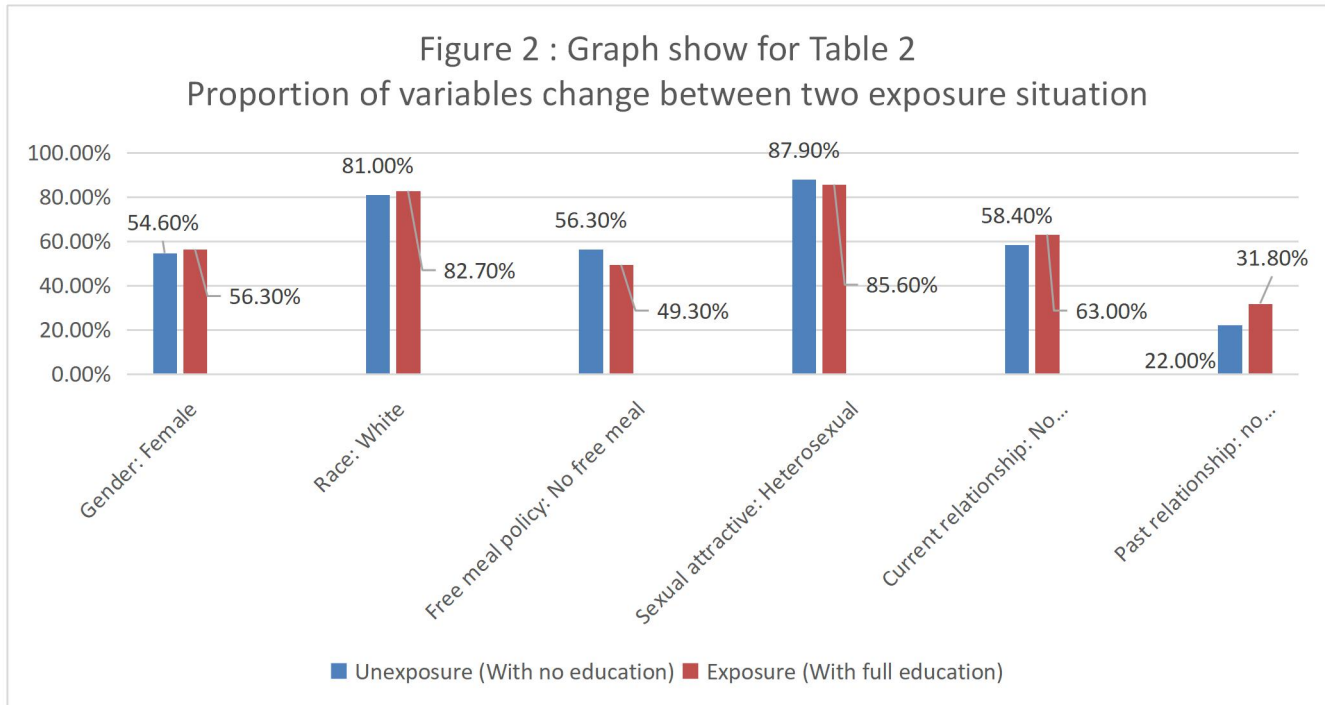Proportion of variables change between two exposure situation

Table 2 presents the results of the variables' proportion changes between unexposure group and exposure group. We try to concentrate on the three variables which is significant in Table 1. From the Figure 2, we can clearly find that the free meal policy variable has a decrease from unexposure to exposure situation. Actually we can say that maybe it is the difference that exposure schools will have a better economic situation, that is why they can afford to publish the Green Dot program. Due to this reason, we may guess that these exposure schools may have more families which do not need this free meal policy. However, another reason may be that because our Time line is from one to five, the schools own policy maybe change and improve. They may have the ability to provide some meals for students which may cause the decrease of free meal amount. That is our opinion to this variable's change. For the two kind of relationships (one is current and another is past relationship), we find that the trends are both increasing. As we set the baseline for these two variables are all no relationships, which means that students are more careful to date with other people. We try to dig the reason for this circumstance. One important reason I find is that after the Green Dot education, students are more mature for dating, they will consider more about their dating friends to avoid the situation that their potential dating friends have the tendency of sexual or dating violence. This can explain the decrease of these two variables value. Besides, I think this is a good signal for us that the Green Dot program is working now. We try to protect our students from suffering the violence. This is our satisfying change.

Table 3. Two scores summary between two exposure situation / add unadjusted and adjusted model

	total	unexposure	exposure	p-value
DV SCORE	4655	2496	2159	<.0001
Mean (SD)	3.2 (2.73)	3.5 (2.78)	2.8 (2.62)	
IRMS SCORE	4647	2499	2148	<.0001
Mean (SD)	5.8 (3.34)	6.3 (3.34)	5.3 (3.26)	

Table 3 is a summary for DV and SV scores between unexposure (time 1) and exposure (time 5). Which includes the total numbers, mean and SD.

From this table especially the P value, these two scores are all significant which means that the adjusted method (Green Dot Sexual education) is effective to their social norms and guide them right cognitive.

		Unadjusted	Adjusted
DV Score	Unexposure Time 1 Mean (SE)	3.54 (0.054)	3.81 (0.080)
	Exposure Time 5 Mean (SE)	2.82 (0.058)	3.16 (0.080)
	Difference Δ /(P-value)	0.72(<.0001)	0.65(<.0001)
IRMS Score	Unexposure Time 1 Mean (SE)	6.25 (0.066)	6.40 (0.095)
	Exposure Time 5 Mean (SE)	5.28 (0.071)	5.56 (0.095)
	Difference Δ /(P-value)	0.97(<.0001)	0.84(<.0001)

More details will show in Logistic regression part.

Logistic Regression

Model 1: glm DV_ACC ~ TIME

Class Level Information		
Class	**Levels**	**Values**
TIME	2	1 5

Number of Observations Read	4719
Number of Observations Used	4655

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	607.79310	607.79310	82.99	<.0001
Error	4653	34075.24578	7.32329		
Corrected Total	4654	34683.03888			

R-Square	Coeff Var	Root MSE	DV_ACC Mean
0.017524	84.33528	2.706157	3.208808

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TIME	1	607.7931020	607.7931020	82.99	<.0001

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	2.820287170	B	0.05824071	48.42	<.0001
TIME 1	0.724584625	B	0.07953608	9.11	<.0001
TIME 5	0.000000000	B	.	.	.

Distribution of DV_ACC
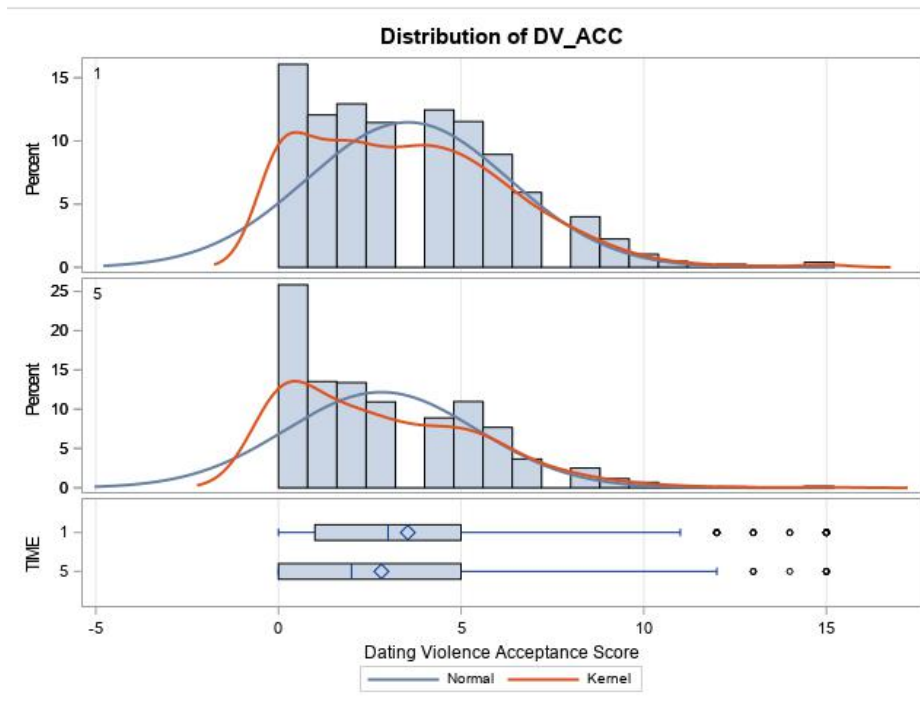
TIME	DV_ACC LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 Pr > |t|
1	3.54487179	0.05416649	<.0001	<.0001
5	2.82028717	0.05824071	<.0001	

From the distribution of DV_ACC between Time 1 and Time 5, we find that comparing to Time 1 which is unexposure, Time 5 the dating violence score has a trend for left move. This explains that after exposure which is full education situation, students score to some dating violence norms is lower. Besides, the bar chart show us the mean and median of Time 5 are all less than Time 1. Based on the survey questionnaire, we know that the lower to score, the less agreement the dating violence. Then we use the LS mean to set as a effect value to show us the exact difference between two exposure situation. Time 5 =2.82< Time 1=3.54 which means the dating violence score for Time 5 is lower than Time 1. The unadjusted model shows that exposure for Green Dot program is effective to guide students' norms to reject sexual violence.

Model 2: DV_ACC ~ TIME + GENDER1 + NONWHITE + F_MEAL1 + NEXHET + CURREL +REL12YN

Class Level Information		
Class	**Levels**	**Values**
TIME	2	1 5
GENDER1	2	0 1
NONWHITE	2	0 1
F_MEAL1	2	0 1
NEXHET	2	0 1
CURREL	2	0 1
REL12YN	2	0 1

Number of Observations Read	4719
Number of Observations Used	4622

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	2544.61977	363.51711	52.87	<.0001
Error	4614	31727.31727	6.87631		
Corrected Total	4621	34271.93704			

R-Square	Coeff Var	Root MSE	DV_ACC Mean
0.074248	81.62264	2.622273	3.212678

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TIME	1	605.357863	605.357863	88.04	<.0001
GENDER1	1	1152.576918	1152.576918	167.62	<.0001
NONWHITE	1	351.311283	351.311283	51.09	<.0001
F_MEAL1	1	46.004799	46.004799	6.69	0.0097
NEXHET	1	56.314199	56.314199	8.19	0.0042
CURREL	1	149.001527	149.001527	21.67	<.0001
REL12YN	1	184.053177	184.053177	26.77	<.0001

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	4.608519366	B	0.15628792	29.49	<.0001
TIME 1	0.654403469	B	0.07815994	8.37	<.0001
TIME 5	0.000000000	B	.	.	.
GENDER1 0	-1.073513664	B	0.07869455	-13.64	<.0001
GENDER1 1	0.000000000	B	.	.	.
NONWHITE 0	-0.689621442	B	0.10194219	-6.76	<.0001
NONWHITE 1	0.000000000	B	.	.	.
F_MEAL1 0	-0.159996623	B	0.07832232	-2.04	0.0411
F_MEAL1 1	0.000000000	B	.	.	.
NEXHET 0	-0.314917482	B	0.11650318	-2.70	0.0069
NEXHET 1	0.000000000	B	.	.	.
CURREL 0	-0.145619429	B	0.09090539	-1.60	0.1093
CURREL 1	0.000000000	B	.	.	.
REL12YN 0	-0.520465165	B	0.10059998	-5.17	<.0001
REL12YN 1	0.000000000	B	.	.	.

TIME	DV_ACC LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 Pr > |t|
1	3.81085593	0.08003343	<.0001	<.0001
5	3.15645246	0.08002990	<.0001	

Model 2 we add all other variables into our model to find that after the adjusted variables, if there is any difference. From the parameter table, we find that variables current relationship and free meal are insignificant based on α =0.01. For all the variables, the format 0 are all negative value, which means that female, white, heterosexual, no past relationship will cause the DV_ACC score lower. Moreover, after adjusted, the LS mean is Time 1=3.81, Time 5=3.16 which is higher than the unadjusted model. This shows that the adjusted variables do some effect on the predicted score and it will be preciser to adjust other effective variables to predict our violence score result.

The total result for exposure is same to unadjusted model, which means adjusted variables do not change the total exposure result. Only do some extent effect to the scores.

Model 3: RAPEMYTH ~ TIME

| Class Level Information		
Class	**Levels**	**Values**
TIME	2	1 5

Number of Observations Read	4719
Number of Observations Used	4647

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1078.16556	1078.16556	98.75	<.0001
Error	4645	50713.45656	10.91786		
Corrected Total	4646	51791.62212			

R-Square	Coeff Var	Root MSE	RAPEMYTH Mean
0.020817	56.93464	3.304218	5.803529

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TIME	1	1078.165559	1078.165559	98.75	<.0001

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	5.283985102	B	0.07129377	74.12	<.0001
TIME 1	0.966114938	B	0.09721982	9.94	<.0001
TIME 5	0.000000000	B	.	.	.

Distribution of RAPEMYTH
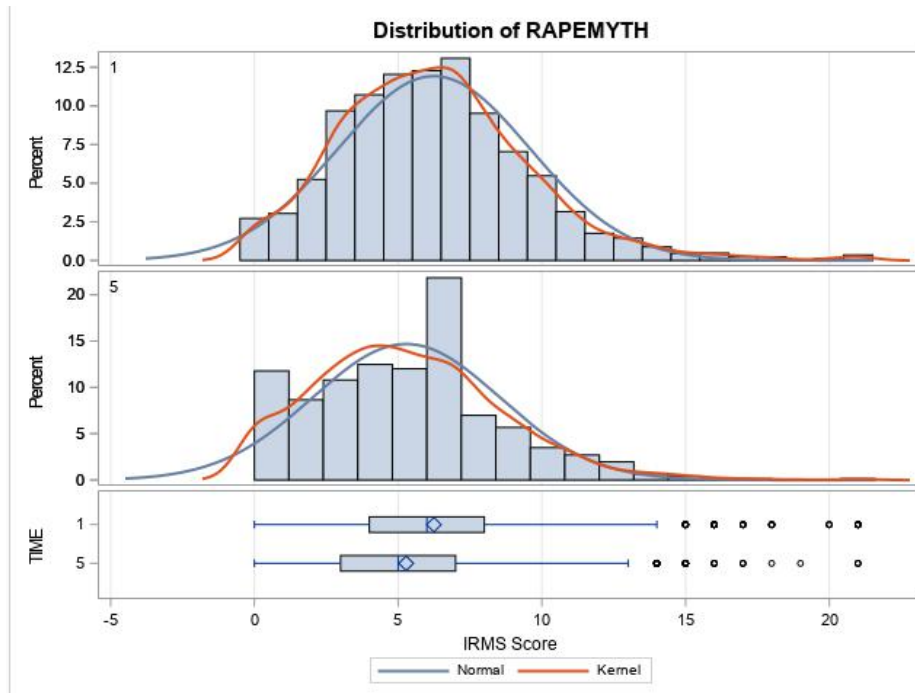
TIME	RAPEMYTH LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 Pr > |t|
1	6.25010004	0.06609759	<.0001	<.0001
5	5.28398510	0.07129377	<.0001	

From the distribution of rapemyth between Time 1 and Time 5, we find that comparing to Time 1 which is unexposure, Time 5 the sexual violence score has a trend for left move. This explains that after exposure which is full education situation, students score to some sexual violence norms is lower. Besides, the bar chart show us the mean and median of Time 5 are all less than Time 1 as same as the lower and upper values. Based on the survey questionnaire, we know that the lower to score, the less agreement the sexual violence. Then we use the LS mean to set as a effect value to show us the exact difference between two exposure situation. Time 5 =5.28< Time 1=6.25 which means the sexual violence score for Time 5 is lower than Time 1. The unadjusted model shows that exposure for Green Dot program is effective to guide students' norms to reject sexual violence.

Model 4: *RAPEMYTH ~ TIME +GENDER1 +NONWHITE +F_MEAL1 +NEXHET + CURREL +REL12YN*

Class Level Information		
Class	**Levels**	**Values**
TIME	2	1 5
GENDER1	2	0 1
NONWHITE	2	0 1
F_MEAL1	2	0 1
NEXHET	2	0 1
CURREL	2	0 1
REL12YN	2	0 1

Number of Observations Read	4719
Number of Observations Used	4618

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	6525.24336	932.17762	96.23	<.0001
Error	4610	44654.67609	9.68648		
Corrected Total	4617	51179.91945			

R-Square	Coeff Var	Root MSE	RAPEMYTH Mean
0.127496	53.65733	3.112311	5.800346

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TIME	1	786.338123	786.338123	81.18	<.0001
GENDER1	1	4019.963618	4019.963618	415.01	<.0001
NONWHITE	1	344.278809	344.278809	35.54	<.0001
F_MEAL1	1	60.416152	60.416152	6.24	0.0125
NEXHET	1	2.697815	2.697815	0.28	0.5977
CURREL	1	159.313652	159.313652	16.45	<.0001
REL12YN	1	516.713877	516.713877	53.34	<.0001

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	7.681946754	B	0.18570736	41.37	<.0001
TIME 1	0.836213872	B	0.09281020	9.01	<.0001
TIME 5	0.000000000	B	.	.	.
GENDER1 0	-1.904332853	B	0.09347916	-20.37	<.0001
GENDER1 1	0.000000000	B	.	.	.
NONWHITE 0	-0.721332410	B	0.12099385	-5.96	<.0001
NONWHITE 1	0.000000000	B	.	.	.
F_MEAL1 0	-0.232391069	B	0.09305204	-2.50	0.0125
F_MEAL1 1	0.000000000	B	.	.	.
NEXHET 0	-0.073060563	B	0.13843947	-0.53	0.5977
NEXHET 1	0.000000000	B	.	.	.
CURREL 0	-0.437229279	B	0.10781166	-4.06	<.0001
CURREL 1	0.000000000	B	.	.	.
REL12YN 0	-0.873074928	B	0.11953896	-7.30	<.0001
REL12YN 1	0.000000000	B	.	.	.

TIME	RAPEMYTH LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 Pr > |t|
1	6.39745007	0.09495210	<.0001	<.0001
5	5.56123620	0.09503418	<.0001	

Model 4 we add all other variables into our model to find that after the adjusted variables, if there is any difference. From the parameter table, we find that variables sexual attractive and free meal are insignificant based on α =0.01. For all the variables, the format 0 are all negative value, which means that female, white, no relationship now, and no past relationship will cause the RAPEMYTH score lower. Moreover, after adjusted, the LS mean is Time 1=6.40, Time 5=5.56 which is higher than the unadjusted model. This shows that the adjusted variables do some effect on the predicted score and it will be preciser to adjust other effective variables to predict our sexual violence score result.

The total result for exposure is same to unadjusted model, which means adjusted variables do not change the total exposure result. Only do some extent effect to the scores.

Lessons Learned

In the capstone project, I have been learning, applying and extending the theoretical knowledge gained from the MPH project. Because I achieve the Master degree in biostatistics, I would like to summarize my experience from the following aspects.

Study Design

We try to control for two independent groups but with different exposure factors through comparative studies and found a relationship between exposure and outcomes. However, through the study of real data, we found that a big problem of this kind of research is the bias caused by the lack of considering the influence of other factors. This problem can cause some bias, reduce the intensity of research, and affect research results. Hence, this prompted us to consider the impact of more variables, while establishing a model that takes into account not only exposure factors but also the adjusted and unadjusted parts. This will show us a more rigorous study design.

Statistical Method

In the data analysis phase, we strengthened descriptive statistics and inferential statistics. We try to do the description analysis in order to explore the relationship for our variables between two different exposure situation. We should be more careful when we try to pick some information up from our beginning analysis. The crosstab includes frequency and percentage but we should figure out which part we need to use. Sometimes we need the exact value but if we want to analysis the trend from different exposure situation, we should use the proportion difference to show us the trend. Moreover, logistic regression is also very important for me to do some study. It is a good way to do some predictive work and help us build a model for analysis.

Statistical Analysis

Through this project on advanced SAS programming, I have made great progress. As I have gone through many iterations, I have learned how to use macros in analysis to achieve more effective results. By studying macros, I have mastered the various methods of defining macro variables, including "%global statement", "%let statement", SQL "select into:" and macro parameters. In logistic regression, I learned the dummy variables "descending" option, "(ref=") "option, and" param= "option to control the response level or reference level. Besides, the "ods" function is also very useful to help us get the result in a good rtf form which can includes tables and figures.

Results

Actually, in our project and study, the output for our SAS program is not suitable to put them into project. This time we can use the proc report and ods function to generate our result into a more formal form. Besides, Words and Excel is also useful for us to edit our result and make some histograms to show our result more clear. We can use our table in the output in our project report, however, when we do the presentation, three lines table will be better. Moreover, some information which is not so important can be canceled when we show them in our slides for presentation.

Project Summary

When we find a social problem, we should take action to solve it. Intervention is necessary. For the sexual and dating violence, we try to build the Green Dot program to set up a education system to intervene the trouble at the school level. However, we should figure out that if this intervention is effective and find the adjusted variables which has some impacts to the result and how they affect it. This is our project's aim.

We perform description analysis and logistic regression. After we use the LS mean as a effective standard value to compare the Time 1 unexposure and Time 5 exposure groups, we find that this exposure which is the Green Dot program intervention is clearly playing a role in improve the students' norms. Which can guide them to a right sexual and dating opinions and prevent them from suffering the violence. Besides, the gender, race, sexual attractive and relationship situation are also be the confounder to affect the relationship with exposure and their scores. This result gives us an inspiration that these variables may have effect on it that we can concentrate on or provide specific improvement for our program.

This result could be a reference for our further improvement in Green Dot program. We should do more survey to consider different variables which may affect higher or lower to their thoughts' scores. Then we can base on the foundation of Green Dot program to improve our intervention to protect our teenagers from sexual or dating violence.

Besides, as a big problem perplexes China recent year, the sexual or dating violence case happened frequently. Because of the difference of culture, Chinese rarely express their attitudes clearly. We don't know how to and what to do when we are facing this situation which happened around us. My purpose for doing this project is trying to find a good way to solve the bad situation that recent year in China, there are many the sexual or dating violence cases happened but no one stop them or the female said no to the violence. Why?
I think one part is lacking of education which can teach them how to help others when they face this problem. Another part is that we are not good at saying no. Building a right social norm is so important. And this is the aim of the Green Dot and its effect. Try to guide us to a right way and teach us how and be willing to help others. We can learn a lot from this project to find a "adjusted" suitable program or action to solve the sexual violence trouble in China and make the society a peace and fair circumstance. That is what we should do as a public health people.