BRFSS: The Behavioral Risk Factor Surveillance System

By: Isabella Zabaleta, Sydney Westfall, Jason Zhang, Anjala Khadka, Sireesha Annam

Why Did We Choose this Dataset?

- Every 40 seconds someone in United States has a heart attack (CDC, 2022).
- The benefits of physical activity (Myers, 2003)
- Recurrent cardiovascular event associated with alcohol consumption (Ding, et. al., 2021)
- How cardiovascular/heart disease risks vary by race and ethnicity (Cleveland Clinic, 2022)

The Behavioral Risk Factor Surveillance System (BRFSS) established in 1985 is a national health-related-telephone survey that collects state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services in all 50 states as well as the District of Columbia and three U.S. territories. It is the largest continuously conducted survey which collects more than 400,000 adult interviews each year. This survey has been helpful to target and build health promotional activities.

Dependent variable:

- Heart Attack

Independent variables:

- Age
- Sex
- Race
- Hypertension
- Cholesterol
- Physical activity
- BMI
- Average alcoholic drinks
- Smoking
- Vegetable consumption

Categorical variables:

- Age
- Race
- Sex
- Hypertension
- Cholesterol
- Heart Attack
- Physical activity
- Smoking

Continuous variables:

- Average alcoholic drinks
- Vegetable consumption
- BMI

Descriptive Statistics: one continuous variable, one categorical variable, one continuous variable grouped by a categorical variable

- Average alcoholic drinks per day in the past 30 days- Continuous (histogram and boxplot)
- Physical activity- categorical (pie and bar chart)
- BMI- continuous
    - Smoking - categorical

Continuous Variable: Alcohol Consumption How many drinks does the average person have in 30 days?

```
set project;
where AVEDRNK3~=77 OR 99; /*77= dont know, 99= refused*/
run;

proc univariate data=alcohol normal plot; /*we want a histogram*/
title 'Descriptive statistics for Alcohol Consumption';
var AVEDRNK3;
Run;
```

Descriptive statistics for Alcohol Consumption

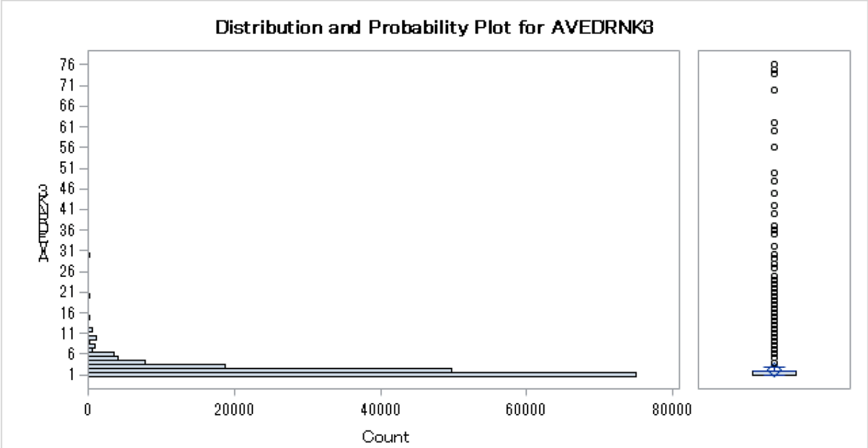
The UNIVARIATE Procedure
Variable: AVEDRNK3 (AVEDRNK3)

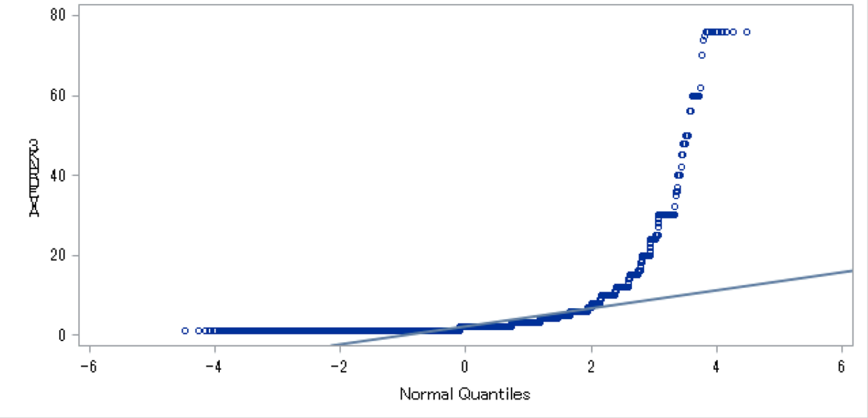
Moments			
N	163285	Sum Weights	163285
Mean	2.16310133	Sum Observations	353202
Std Deviation	2.26131516	Variance	5.11354625
Skewness	9.39950057	Kurtosis	187.351956
Uncorrected SS	1598972	Corrected SS	834960.285
Coeff Variation	104.540418	Std Error Mean	0.00559613

Basic Statistical Measures			
Location		Variability	
Mean	2.163101	Std Deviation	2.26132
Median	2.000000	Variance	5.11355
Mode	1.000000	Range	75.00000
		Interquartile Range	1.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	386.5351	Pr > |t|	<.0001
Sign	M	81642.5	Pr >= |M|	<.0001
Signed Rank	S	6.6655E9	Pr >= |S|	<.0001

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.303505	Pr > D	<0.0100
Cramer-von Mises	W-Sq	3823.211	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	20141.25	Pr > A-Sq	<0.0050





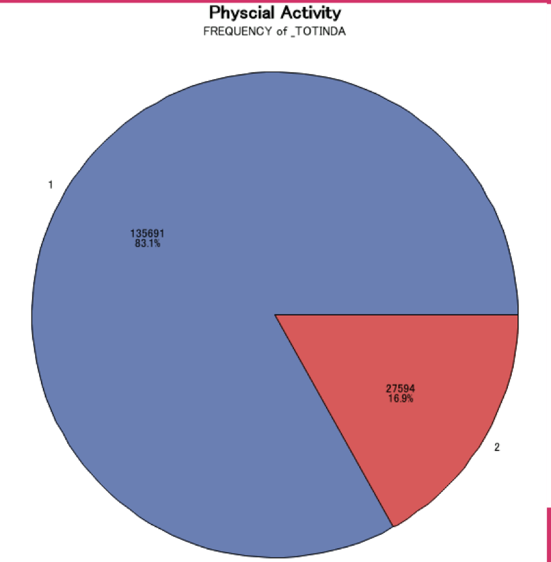
The distribution for average number of drinks per 30 days does not follow a normal curve. This data is skewed left.

Categorical Variable: Physical Activity How many times does the average person workout in 30 days?

```
/*Descriptive statistics for one categorical
variable*/
data physicalactivity;
set project;
keep _TOTINDA;
where _TOTINDA~=9;
run;
proc FREQ data=physicalactivity;
title 'Desciptive Statistics for Physical Activity';
tables _TOTINDA;
run;
```

```
proc gchart data=physicalactivity;
title 'Physcial Activity';
vbar _TOTINDA / midpoints= 1 to 2 by 1;
run;
proc gchart data=physicalactivity;
pie _TOTINDA / discrete
value=inside percent=inside slice=outside;
run;
```

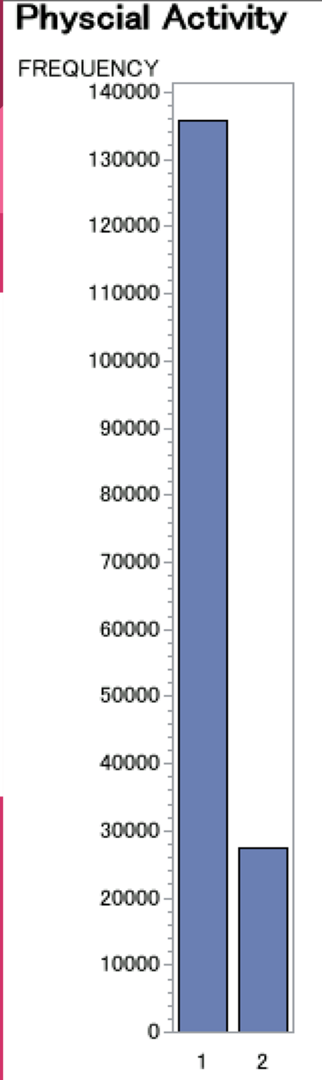
Categorical Variable: Physical Activity
How many times does the average person have in 30 days?



Desciptive Statistics for Physical Activity

The FREQ Procedure

_TOTINDA				
_TOTINDA	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	135691	83.10	135691	83.10
2	27594	16.90	163285	100.00



83.1% of patients report “yes” to physical activity outside of their regular job.

Continuous Variable grouped by Categorical: BMI grouped by smoking
Are smoking status and BMI directly correlated?

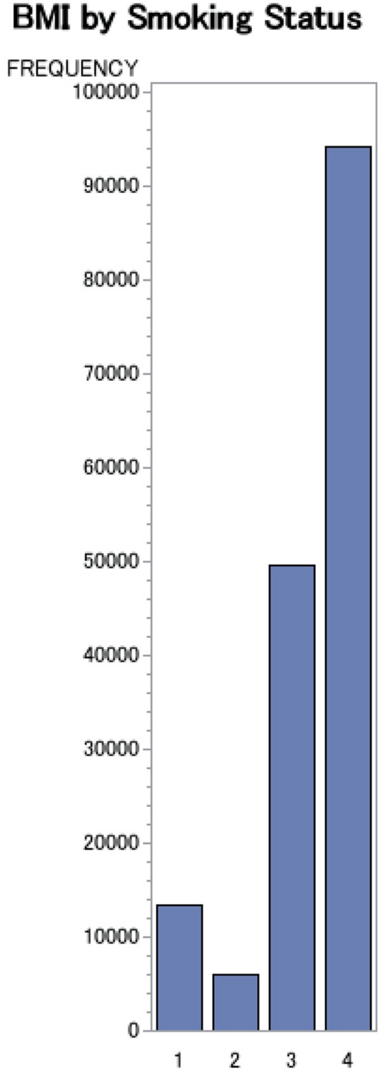
```
data BMIbySmoking;
set project;
keep _BMI5 _SMOKER3;
where _SMOKER3 ~=9;
run;

proc means data=BMIbySmoking;
title 'Descriptive Statistics for BMI by Smoking Status';
class _SMOKER3;
var _BMI5;
run;

proc gchart data=BMIbySmoking;
title 'BMI by Smoking Status';
vbar _SMOKER3 / midpoints= 1 to 4 by 1 type=mean;
run;
```

_SMOKER3=1: Current everyday smoker

_SMOKER3=2: Current smoker sometimes

_SMOKER3=3: Former smoker

_SMOKER3=4: Never smoked



Descriptive Statistics for BMI by Smoking Status

The MEANS Procedure

Analysis Variable : _BMI5 _BMI5						
_SMOKER3	N Obs	N	Mean	Std Dev	Minimum	Maximum
1	13497	13497	2822.26	657.0780336	1211.00	9245.00
2	5934	5934	2844.60	639.4338767	1356.00	9548.00
3	49585	49585	2864.52	590.0434282	1216.00	9152.00
4	94269	94269	2820.90	601.0246429	1200.00	9933.00

Average BMI is higher for patients who never smoked.

Test of normality, including measures of skewness, comparing the mean to the median, histogram with normal curve overlaid, normal q-q plot, and significance test for normality

```
/*Test of normality*/
data drinknormality;
set project;
where AVEDRNK3~=9 and AVEDRNK3 ne .;
run;
/*Histogram with normal curve overlay*/
ods select Histogram ParameterEstimates GoodnessofFit
FitQuantiles Bins;
proc univariate data=drinknormality;
title 'Normality for Binge Drinking';
var AVEDRNK3;
histogram AVEDRNK3 / normal;
run;
```

```
/* QQ plot*/
proc univariate data=drinknormality;
title 'Normality for Average Drinks';
var AVEDRNK3;
qqplot AVEDRNK3 /normal (MU=EST
SIGMA=EST COLOR=RED L=1);
run;
```

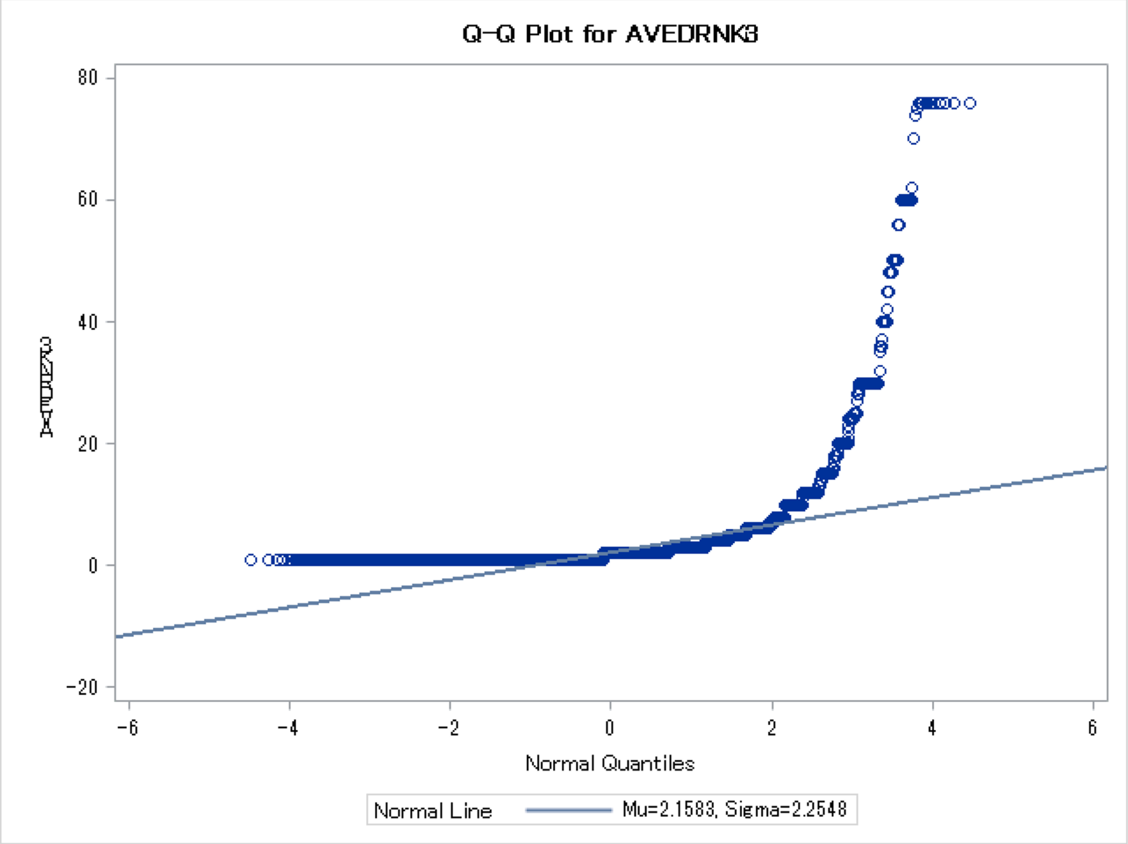
Normality for Drinking

The UNIVARIATE Procedure
Variable: AVEDRNK3 (AVEDRNK3)

Moments			
N	163170	Sum Weights	163170
Mean	2.15828277	Sum Observations	352167
Std Deviation	2.25481326	Variance	5.08418285
Skewness	9.4744931	Kurtosis	189.714143
Uncorrected SS	1589657	Corrected SS	829581.031
Coeff Variation	104.47256	Std Error Mean	0.00558201

Basic Statistical Measures			
Location		Variability	
Mean	2.158283	Std Deviation	2.25481
Median	2.000000	Variance	5.08418
Mode	1.000000	Range	75.00000
		Interquartile Range	1.00000

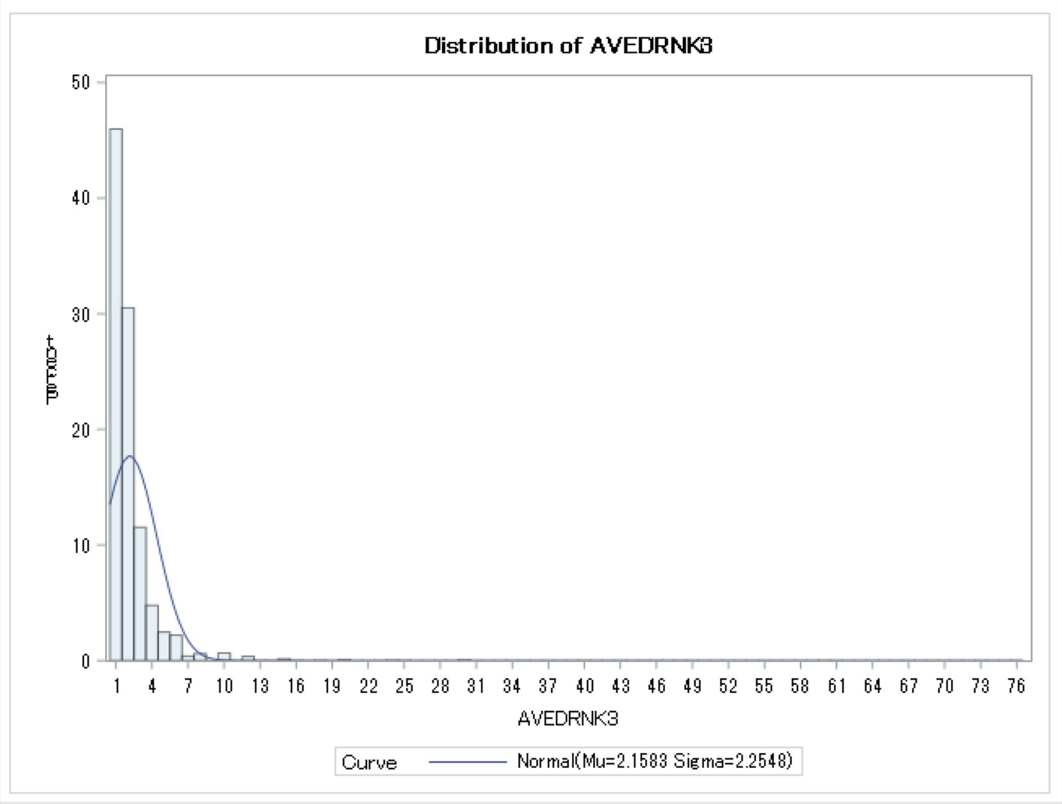
Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	386.6499	Pr > |t|	<.0001
Sign	M	81585	Pr >= |M|	<.0001
Signed Rank	S	6.6562E9	Pr >= |S|	<.0001



Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	2.158283
Std Dev	Sigma	2.254813

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.3037	Pr > D	<0.010
Cramer-von Mises	W-Sq	3816.2822	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	20112.5871	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	1.00000	-3.08720
5.0	1.00000	-1.55055
10.0	1.00000	-0.73138
25.0	1.00000	0.63743
50.0	2.00000	2.15828
75.0	2.00000	3.67913
90.0	4.00000	5.04794
95.0	5.00000	5.86712
99.0	10.00000	7.40376



This data is not normal (P-Values for each goodness of fit test is <0.05, meaning it is significantly different, or not normal). However, it is a drinking variable so we expect some skew.

Test for H0 with 95% confidence interval (CI)

- H1: There is a significantly higher risk of heart attack for people who have higher BMI of >25kg/m^2
- H0: There is not significantly higher risk of heart attack for people who have higher BMI of >25kg/m^2.

```
data BMIheartattack;
set project;
keep CVDINFR4 _BMI5;
where CVDINFR4=1;
run;

proc ttest data=BMIheartattack h0=2500 sides=U alpha=0.05 plots= none;
title 'One Sided T-test';
var _BMI5;
run;
```

One Sided T-test

The TTEST Procedure

Variable: _BMI5 (_BMI5)

N	Mean	Std Dev	Std Err	Minimum	Maximum
6818	2902.3	595.8	7.2154	1200.0	9140.0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2902.3	2890.4	Infty	595.8	585.9	606.0

DF	t Value	Pr > t
6817	55.76	<.0001

- There is a statistically higher chance of a person with a BMI >25% having a heart attack compared to someone with BMI<25%.
- However, the hypothesized value is not included in the 95% CI [29.19-infinity].
- This means we can say with 95% confidence that someone who has had a heart attack will fall within the 28.9+ BMI range.

Contingency table with Chi-square significance test for H0: Does heart attack risk increase with age?

```
data ageMI;
set project;
keep _age_g CVDINFR4;
where CVDINFR4<7;
run;

proc freq data=ageMI;
TITLE 'Chi-Squared Test for Age Group and Heart Attack';
tables CVDINFR4* _AGE_G / chisq expected relrisk;
run;
```

Frequency Percent Row Pct Col Pct	Table of CVDINFR4 by _AGE_G						
	_AGE_G(_AGE_G)						
CVDINFR4(CVDINFR4)	5	6	4	2	3	1	Total
2	32822 20.10 20.98 95.73	49331 30.21 31.53 91.83	27255 16.69 17.42 97.70	17162 10.51 10.97 99.48	24157 14.79 15.44 99.11	5740 3.52 3.67 99.74	156467 95.82
1	1464 0.90 21.47 4.27	4390 2.69 64.39 8.17	642 0.39 9.42 2.30	90 0.06 1.32 0.52	217 0.13 3.18 0.89	15 0.01 0.22 0.26	6818 4.18
Total	34286 21.00	53721 32.90	27897 17.08	17252 10.57	24374 14.93	5755 3.52	163285 100.00

Statistics for Table of CVDINFR4 by _AGE_G

Statistic	DF	Value	Prob
Chi-Square	5	3843.4874	<.0001
Likelihood Ratio Chi-Square	5	4228.5565	<.0001
Mantel-Haenszel Chi-Square	1	3388.4059	<.0001
Phi Coefficient		0.1534	
Contingency Coefficient		0.1516	
Cramer's V		0.1534	

- As expected, the chi- square test shows us that heart attack risk increases as age increases.

Independent t-test: Is there a significant difference between vegetable consumption and age?

```
data ageveg;
set project;
keep _VEGESU1 _AGE_G AGEGROUP;
where _VEGESU1<99998;
if 1<_AGE_G<3 then AGEGROUP= 1;
else AGEGROUP= 2;
run;
proc ttest data=ageveg;
title 'Independent T test for Vegetable Consumption and Age';
class AGEGROUP;
var _VEGESU1;
run;
```

Independent T test for Vegetable Consumption and Age

The TTEST Procedure

Variable: _VEGESU1 (_VEGESU1)

AGEGROUP	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		16196	259.7	852.2	6.6963	0	31200.0
2		137642	277.7	1032.7	2.7837	0	39600.0
Diff (1-2)	Pooled		-18.0647	1015.3	8.4339		
Diff (1-2)	Satterthwaite		-18.0647		7.2518		

AGEGROUP	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		259.7	246.5	272.8	852.2	843.0	861.6
2		277.7	272.3	283.2	1032.7	1028.9	1036.6
Diff (1-2)	Pooled	-18.0647	-34.5949	-1.5345	1015.3	1011.7	1018.9
Diff (1-2)	Satterthwaite	-18.0647	-32.2788	-3.8506			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	153836	-2.14	0.0322
Satterthwaite	Unequal	22198	-2.49	0.0127

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	137641	16195	1.47	<.0001

- There is a statistically significant increase in vegetable consumption in patients age 45+.

Linear Correlation:
Is vegetable consumption correlated with BMI?

```
data BMIbyVegetables;
set project;
run;

proc sgplot data=BMIbyVegetables (firstobs=1 obs=500);
title 'BMI by Vegetable Consumption';
scatter x=_BMI5 y=_VEGESU1;
Ellipse x=_BMI5 y=_VEGESU1;
run;

proc corr data=BMIbyVegetables;
title 'Linear Correlation for BMI and Vegetable Consumption';
var _BMI5 _VEGESU1;
run;
```

Linear Correlation for BMI and Vegetable Consumption

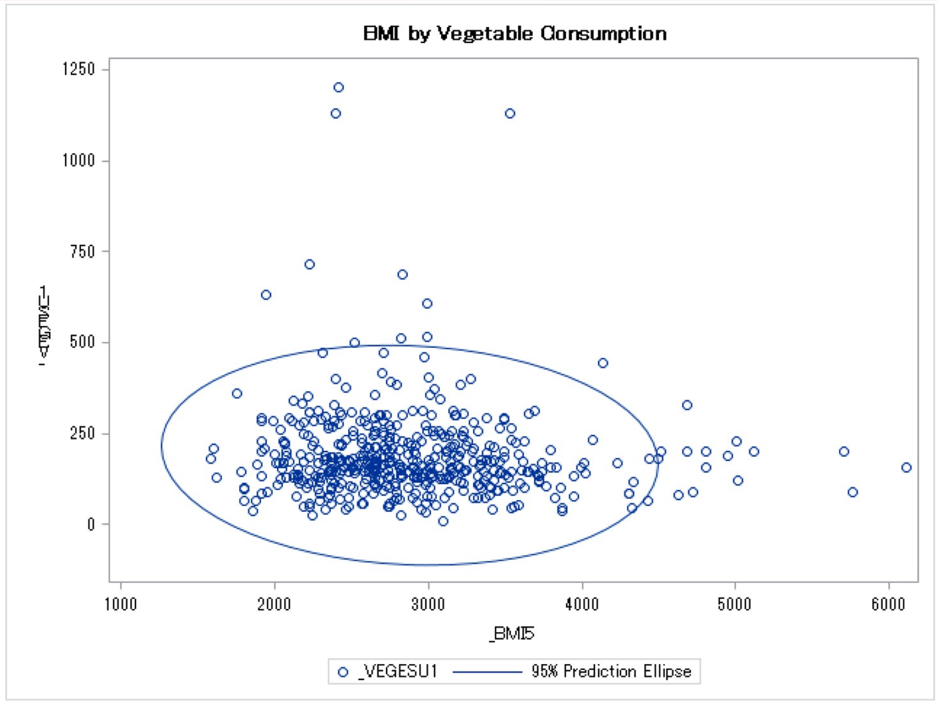
The CORR Procedure

2 Variables:	_BMI5 _VEGESU1

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
_BMI5	163285	2835	604.32485	462932388	1200	9933	_BMI5
_VEGESU1	153838	275.81972	1015	42431554	0	39600	_VEGESU1

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	_BMI5	_VEGESU1
_BMI5 _BMI5	1.00000 163285	-0.02390 <.0001 153838
_VEGESU1 _VEGESU1	-0.02390 <.0001 153838	1.00000 153838



- Mean BMI is 28.35 and mean vegetables consumed in a day is 2.78. These values have a loose, negative (but still statistically significant) correlation. This means that as BMI increases, vegetable consumption decreases.

Simple Logistic Regression Analysis:

```
data simplogistic;
set project;
KEEP _VEGESU1 CVDINFR4;
RUN;

PROC LOGISTIC DATA=simplogistic;
title 'Logistic Regression- Vegetable Consumption and Heart Attack';
MODEL CVDINFR4 (EVENT='1')=_VEGESU1;
WHERE _VEGESU1<99998 AND CVDINFR4<7 AND CVDINFR4~=.;
RUN;
```

The LOGISTIC Procedure

Model Information		
Data Set	WORK.SIMPLOGISTIC	
Response Variable	CVDINFR4	CVDINFR4
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	163285
Number of Observations Used	153838

Response Profile		
Ordered Value	CVDINFR4	Total Frequency
1	1	6356
2	2	147482

Probability modeled is CVDINFR4='1'.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	52954.500	52954.206
SC	52964.444	52974.093
-2 Log L	52952.500	52950.206

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.2942	1	0.1299
Score	2.0752	1	0.1497
Wald	2.0519	1	0.1520

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.1387	0.0134	55138.8704	<.0001
_VEGESU1	1	-0.00002	0.000015	2.0519	0.1520

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.2942	1	0.1299
Score	2.0752	1	0.1497
Wald	2.0519	1	0.1520

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.1387	0.0134	55138.8704	<.0001
_VEGESU1	1	-0.00002	0.000015	2.0519	0.1520

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
_VEGESU1	1.000	1.000	1.000

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	52.8	Somers' D	0.061
Percent Discordant	46.6	Gamma	0.062
Percent Tied	0.6	Tau-a	0.005
Pairs	937395592	c	0.531

- Vegetable consumption is not a statistically significant contributor to heart attack in regard to the BRFSS survey results.

Multiple Logistic Regression Hypothesis:

- H0: There is no increased risk of heart attack among people with hypertension and high cholesterol
- H1: There is increased risk of heart attack among people with hypertension and high cholesterol.

Multiple Logistic Regression

```
data logistic1;
set project;
if _BMI5 = 0 then _BMI5 = 0; else if _BMI5= 0 < _BMI5 < 9999 then _BMI5 = 1;
if _RACE = 1  then _RACE1 =1; else _RACE2=0;
if _RACE =2 then _RACE2 =1; else _RACE2=0;
if _RACE = 3 then _RACE3 =1; else _RACE3=0;
if _RACE = 4 then _RACE4 =1; else _RACE4=0;
if _RACE = 5 then _RACE5 =1; else _RACE5=0;
if _RACE = 6 then _RACE6 =1; else _RACE6=0;
if _RACE = 7 then _RACE7 =1; else _RACE7=0;
if _RACE = 8 then _RACE8 =1; else _RACE8=0;
if AVEDRNK3= 0 then AVEDRNK3 =0; else if 0< AVEDRNK3 <76 then AVEDRNK3 =1; else AVEDRNK3=0;
if _TOTINDA = 1 then _TOTINDA = 1; else if _TOTINDA =0 then _TOTINDA = 0;
if CVDINFR4 = 1 then CVDINFR4 = 1; else if CVDINFR4 =0 then CVDINFR4 = 0;
if SEXVAR = 1 then SEXVAR = 1; else if SEXVAR =0 then SEXVAR = 0;
if _VEGESU1= 0 then _VEGESU1 = 0; else if 0 < _VEGESU1 < 99998 then _VEGESU1 = 1;
if BPHIGH6 = 1 then BPHIGH6 = 1; else BPHIGH6 = 0;
if TOLDHI3 = 1 then TOLDHI3 = 1; else TOLDHI3 = 0;
if _AGE_G = 1  then _AGE_G1 =1; else _AGE_G1=0;
if _AGE_G = 2 then _AGE_G2 = 1; else _AGE_G2=0;
if _AGE_G = 3 then _AGE_G3 =1; else _AGE_G3=0;
if _AGE_G = 4 then _AGE_G4 =1; else _AGE_G4=0;
if _AGE_G = 5 then _AGE_G5 =1; else _AGE_G5=0;
if _AGE_G = 6 then _AGE_G6 =1; else _AGE_G6=0;
if _SMOKER3= 1 then _SMOKER31=1; else _SMOKER31=0;
if _SMOKER3= 2 then _SMOKER32=1; else _SMOKER32 = 0;
if _SMOKER3 = 3 then _SMOKER33=1 ;else _SMOKER33 = 0;
if _SMOKER3 = 4 then _SMOKER34=1 ;else _SMOKER34 = 0;
run;
proc logistic data=logistic1;
title 'Multiple Logistic Regression';
model CVDINFR4(event='1')=  AVEDRNK3 _TOTINDA _VEGESU1 _AGE_G SEXVAR _RACE BPHIGH6 TOLDHI3 _SMOKER3 _BMI5 / lackfit;
where CVDINFR4~=7 and CVDINFR4~=9 and CVDINFR4~=.;
run;
```

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	138899.00	118421.72
SC	138909.70	118539.42
-2 Log L	138897.00	118399.72

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20497.2781	10	<.0001
Score	18211.0815	10	<.0001
Wald	14576.1418	10	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.0842	0.1695	580.7686	<.0001
AVEDRNK3	1	-0.4821	0.0167	831.3202	<.0001
_TOTINDA	1	0.2240	0.0120	347.6400	<.0001
_VEGESU1	1	-0.4223	0.1540	7.5207	0.0061
_AGE_G	1	0.5561	0.00926	3608.4673	<.0001
SEXVAR	1	-0.7665	0.0166	2119.6610	<.0001
_RACE	1	0.0177	0.00380	21.6327	<.0001
BPHIGH6	1	0.8052	0.0188	1843.9588	<.0001
TOLDHI3	1	0.5973	0.0171	1226.2428	<.0001
_SMOKER3	1	-0.3168	0.00790	1607.8209	<.0001
_BMI5	1	0.000037	0.000013	8.3335	0.0039

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
AVEDRNK3	0.617	0.598	0.638
_TOTINDA	1.251	1.222	1.281
_VEGESU1	0.656	0.485	0.886
_AGE_G	1.744	1.712	1.776
SEXVAR	0.465	0.450	0.480
_RACE	1.018	1.010	1.025
BPHIGH6	2.237	2.156	2.321
TOLDHI3	1.817	1.757	1.879
_SMOKER3	0.728	0.717	0.740
_BMI5	1.000	1.000	1.000

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	79.8	Somers' D	0.596
Percent Discordant	20.2	Gamma	0.596
Percent Tied	0.0	Tau-a	0.062
Pairs	5547299500	c	0.798

Partition for the Hosmer and Lemeshow Test					
		CVDINFR4 = 1		CVDINFR4 = 2	
Group	Total	Observed	Expected	Observed	Expected
1	32781	70	81.92	32711	32699.08
2	32778	139	182.55	32639	32595.45
3	32780	272	332.61	32508	32447.39
4	32779	485	556.10	32294	32222.90

Results of Logistic Regression

- The odds of heart attack among people who have hypertension and high serum cholesterol is 0.8062 and 0.5973 times respectively than those without hypertension and high serum cholesterol adjusting for possible confounders.

- The p-value is less than 0.05 which means we reject the null hypothesis and conclude that there is increased risk of heart attack among people with hypertension and high cholesterol.

Conclusion and Discussion:

- Conclusions:
    - Heart attack is associated with a wide range of behavioral risk factors and determinants of health.
    - Surveillance data does not trend with the majority of the population
        - Dependent on controlled sample sizes
- Study limitations:
    - Survey bias
    - Binary variables limited the analysis. If we had more robust data, more concrete conclusions could be made
- Identify items you found most:
    - Enjoyable/interesting: working together
    - Challenging: handling missing data, null values, compartmentalized

Appendices:

- Appendix 1: https://www.cdc.gov/brfss/annual_data/annual_2021.html 2021 BRFSS Data (SAS Transport Format) [ZIP – 64.3 MB]
- Appendix 2:
  - Isabella: built foundation and created slidedeck, collaborated in writing code, presented introduction and conclusion
  - Sydney: led and drove the group in writing the code and interpretations, presented visual analytics and findings
  - Jason: collaborated in writing code, expanded on slides, presented visual analytics and findings
  - Anjala: researched and selected the dataset, collaborated in writing code, presented visual analytics and findings
  - Sireesha: collaborated in writing code, expanded on interpretations, presented visual analytics and findings
- Appendix 3:
  - Multiple logistic regression analysis was not covered in the oral presentation
  - Cholesterol and hypertension variables were not presented in the oral presentation

References:

Centers for Disease Control and Prevention. (2022, October 14). *Heart disease facts*. Centers for Disease Control and Prevention. Retrieved December 8, 2022, from https://www.cdc.gov/heartdisease/facts.html

Ding, C., O'Neill, D., Bell, S., Stamatakis, E., & Britton, A. (2021, July 27). Association of alcohol consumption with morbidity and mortality in patients with cardiovascular disease: Original Data and meta-analysis of 48,423 men and women - *BMC medicine*. Retrieved November 28, 2022, from https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-021-02040-2

Heart disease risk: How race and ethnicity play a role. *Cleveland Clinic.* (2022, May). Retrieved November 28, 2022, from https://my.clevelandclinic.org/health/articles/23051-ethnicity-and-heart-disease

Myers, J. (2003, January 7). Exercise and cardiovascular health. *Circulation*. Retrieved November 28, 2022, from https://www.ahajournals.org/doi/full/10.1161/01.CIR.0000048890.59383.8D