MEMORANDUM

From: Elek McDonald

To: Dr. Craig

Date: 12/6/21

Subj: Is smoking related to obesity?

Section 1: Introduction/Overview

Smoking has been an issue in the United States for decades now, yet that attempts to fix that problem seem ineffective. Although the overall rate of adults who smoke has decreased slightly, a large population of smokers still exists. The concerning part of smoking is that the general population is well aware of its side effects and the long-term health problems it brings. In addition to smoking, I've decided to explore the pressing issue of obesity. Obesity rates have continually increased in most states for years now, evidently making the issue more concerning than before. In a country where smoking and poor eating habits are extremely common, I thought it would be interesting to see if there was some kind of connection between the two issues in the United States. If these problems are continually explored, solutions are bound to follow, hopefully reducing the impact of smoking and obesity.

Section 2: Statement of Hypotheses

With the data I collected, I am intrigued to see if a relationship exists between obesity and smoking. By testing different states and areas of the country, I should gather enough results to make a reasonable claim. Obviously, the data collected is not from every single person in the entire country because that would be extremely costly, but rather samples from each state were

surveyed to represent each state as a whole. The following hypotheses will be tested to determine relationships and outcomes within the data.

Hypotheses:

Null Hypothesis: There is a difference between the smoking rates in 2018 between the four geographic regions.
Alternative: There is no difference between the smoking rates in 2018 between the four geographic regions.

Null Hypothesis: There is a difference between the obesity rates in 2018 between the four geographic regions.
Alternative: There is no difference between the obesity rates in 2018 between the four geographic regions.

Null Hypothesis: There is no difference between smoking rates and obesity rates in 2018.
Alternative Hypothesis: There is a difference between smoking rates and obesity rates in 2018.

Section 3: Data and Analysis Methodology

The Centers for Disease Control and Prevention (CDC) is where I collected all of my data. For the obesity portion of my report, the study I found highlights the percentage of smokers from ages 18+ in all 50 states in addition to Puerto Rico and Guam. I recorded the obesity percentages from 2016-2020. The smoking portion of my report discusses data from 2018 about the percentage of adults who smoke in all 50 states plus Puerto Rico and Guam. Because the

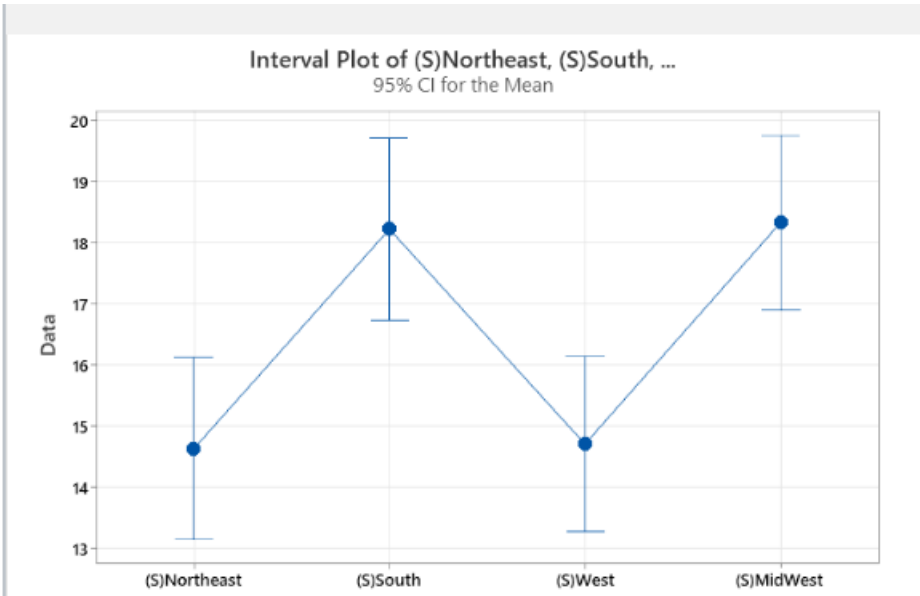
smoking statistics pertain to 2018, I am comparing the obesity rates of 2018 to the smoking rates of 2018 instead of obesity rates from the other years. In addition, I've chosen to compare regions instead of all of the states because the ANOVA test will not process 50 different groups. I mapped out the regions, took all the rates from the states within the regions, averaged them out, and put them into the ANOVA test. Because I am comparing regions, I've decided to omit Puerto Rico and Guam from my analysis, so the results are accurate.

I will be running an ANOVA test with the Tukey comparison for my first hypothesis, an ANOVA test with the Tukey comparison for my second hypothesis, and a two-sample t-test for my third hypothesis. The ANOVA Tukey tests are effective in comparing the different regions as groups and determining if there really is any significance between the different rates with different regions. The two-sample t-test will allow me to directly compare obesity to smoking to determine if there is any kind of relationship between the two.

Section 4: Results

One-Way ANOVA: Smoking Rates, 2018 vs. Region

As previously stated, the first hypothesis tested is that there is a difference between the smoking rates in 2018 between the four geographic regions. From the results of the One-Way ANOVA test with the Tukey comparison we can conclude that there is a difference between the smoking rates in 2018 between the four geographic regions, so we accept the null. Because the Midwest and South are in Group A and the West and Northeast are in Group B, there is enough evidence to conclude the cluster is separated into two groups.



Factor	N	Mean	Grouping
(S)MidWest	13	18.323	A
(S)South	12	18.217	A
(S)West	13	14.708	B
(S)Northeast	12	14.633	B

Means that do not share a letter are significantly different.

One-Way ANOVA: Obesity Rates, 2018 vs. Region

The second hypothesis tested is that there is a difference between the obesity rates in 2018 between the four geographic regions. From the results of the One-Way ANOVA test with the Tukey comparison we can conclude that there is a difference between the obesity rates in 2018 between the four geographic regions, so we accept the null. Because the Midwest and South are in Group A and the West and Northeast are in Group B, there is enough evidence to conclude the cluster is separated into two groups.

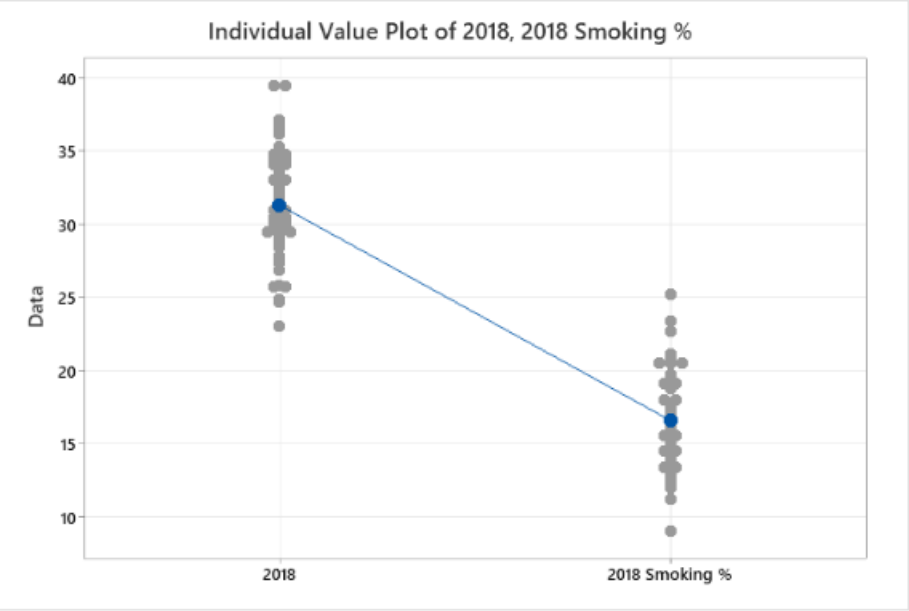
Factor	N	Mean	Grouping
(O)South	12	34.542	A
(O)Midwest	13	33.508	A
(O)Northeast	12	29.70	B
(O)West	13	28.092	B

Means that do not share a letter are significantly different.

Two-Sample t-test and CI: 2018 Obesity Rates and 2018 Smoking Rates

The third hypothesis tested is that there is a difference between the obesity rates in 2018 between the four geographic regions. After running the two-sample t-test, it can be concluded that there is no way obesity rates in 2018 are related to smoking rates in 2018. The Individual Value shows a large difference, but the P-value of 0 confirms the two variables are orthogonal.

T-Value	DF	P-Value
20.56	97	0.000



Section 5: Discussion and Conclusion

Smoking has been a long-term issue not only in the United States, but all over the world. Although the rates of smoking in the United States have decreased over time, the introduction of e-cigs has brought those rates back up. Obesity is a current and rising problem in the U.S., so the

data in my report helps understand if there is a correlation between the two problems. Unfortunately, after running the tests for my various hypotheses, I was not able to draw any significant conclusions. However, I noticed the Midwest and South were in the same group and the Northeast and West were in the same group, which was interesting. Considering the Northeast and West are on opposite sides of the country, it is difficult to understand why they would have similar obesity and smoking rates. It could be because they are both closer to the ocean, so the imports of smoking products are more accessible, and the opposite goes for the South and Midwest. Because the two-sample t-test showed there is no relationship between smoking and obesity, maybe I have to add more variables to test in the future to see if other factors would relate to the increase in both problems. As both problems progress and more data are surveyed, I will look into testing each variable again to see if a relationship can be noticed between the two issues.