Module 8, Capstone, Option 1: Youth Suicide and Population Factors

Robin King

MIS 581: Capstone – Business Intelligence and Data Analytics

Orenthio Goodwin, Ph.D.

Colorado State University, Global Campus

April 5, 2020

Abstract

This study used the latest publicly available health and economic county-level data to find correlations between key health and economic measures and the percentage of Colorado youth who seriously considered suicide. Median home and percent of housing units with more than one person per room were moderately negatively correlated. Median household income was weakly negatively correlated. The percent of related children (5-17 years old) in poverty and the percent of children (< 18 years old) below poverty level were weakly positively correlated. Conversely, rates of poverty and deep poverty were not significantly correlated. Several health factors were moderately positively correlated, including markers of poor health among adults (rates of diabetes, obesity, inactivity, self-identified fair or poor health, and hospitalization due to stroke), poor health among high school students (rates of obesity and asthma), the percentage of high school students who had extended periods of sadness, and the percentage of high school students who had had sexual intercourse. Perhaps surprisingly, this study found that the percentage of adults who reported recent binge drinking and the percent of adults who used marijuana recently were moderately negatively correlated, while the percentage of high schoolers who used marijuana recently was weakly negatively correlated. Finally, the percentage of students who ate several fruits and vegetables each day and the percent of active adults were both moderately negatively correlated. No significant correlations were found for rates of high schoolers' physical activity or current cigarette smoking or adults' current asthma diagnoses.

Youth Suicide and Population Factors

Introduction

Suicide is the leading cause of death for Coloradans aged between 10 and 24 years (PCMH, n.d.). To address this crisis, it is vital to understand the public health factors that may help to identify vulnerable youth populations. That understanding may also point towards additional avenues of research in terms of designing targeted population-level interventions. The Partners for Children's Mental Health (PCMH) is a new center (formed in 2017) focused on improving Colorado's youth mental health system. Located in Aurora, near downtown Denver, Colorado, PCMH is jointly supported by Children's Hospital Colorado and the Department of Psychiatry at the University of Colorado. PCMH has an advisory council of approximately twenty major stakeholders, including representatives from Medicaid, the Colorado Department of Education, youth and family advocacy organizations, and philanthropists. PCMH claims five core functions ("Frequently Asked Questions," n.d.):

1. Collaborating with partnerships to develop a unified vision for pediatric mental health.
2. Improving workforce capacity through training, coaching, and technical assistance.
3. Performing research and evaluation, using data to guide practice and policy.
4. Supporting implementation by providing evidence-driven consultation to guide program design and implementation.
5. Influencing initiatives, policy, and legislation related to children's mental health.

PCMH is a small non-profit organization with 12 employees. It is "funded through grants, philanthropy, and interagency contractual agreements" ("Frequently Asked Questions," n.d.). The Children's Hospital Foundation is working to secure funding so that PCMH can offer

resources to community stakeholders at no cost ("Frequently Asked Questions," n.d.). PCMH was selected for this study due to the importance of its work.

Objectives

This exploratory study looked for correlations between high school students who have seriously considered suicide and a variety of health and economic factors. It compares Colorado counties across a variety of health and economic measures using public data provided by the Colorado and United States governments. Correlations are explored because understanding associations with public health factors may help to identify vulnerable youth and target population-level interventions.

Overview of Study

This study used public health and economic data to identify several health and economic factors that are correlated with the percent of high school students who seriously considered suicide. Factors examined were selected based on preliminary data exploration as well as insights gleaned from the literature review.

Literature Review

Youth suicide has received substantial focus in the literature. Grimmond, Kornhaber, Visentin, and Cleary (2019) performed a qualitative systematic review focused on youth suicide and found that "thwarted belongingness," "perceived burdensomeness" (often related to strained family relationships and poverty), and "the capability for suicide" (p. 19) were important factors. As well, they noted that drug and alcohol use was a common risk factor or trigger for youth suicide (Grimmond et al., 2019). Crepeau-Hobson and Estes (2019) found that removing financial and social barriers to access to treatment for youth at risk of suicide led to a reduction

in suicidal thoughts and behaviors. Investigating suicide across age spans, Pitman, Tham, Hunt, Webb, Appleby, and Kapur (2019) reported that physical health issues including cancer, back pain, diabetes, and heart disease are all associated with an increased risk of suicide.

Research Hypotheses

Previous researchers have identified several factors that correlate with youth suicide and suicide more broadly. Many of those studies were performed qualitatively, investigating the personal attributes for individuals. This study will leverage their findings using county-level measures to determine whether community attributes are similarly correlated with youth suicide rates. It will study the percentage of high school students who seriously considered attempting suicide during the previous 12 months (dated 2015). It will look at potential correlations with measures of poverty, alcohol and marijuana use, adult physical health, and youth health and behavior factors. See Appendix A for detailed questions and hypotheses.

Research Design

Methodology

This research study investigates health and economic factors that have been shown to correlate with youth suicide risk on a personal level and determines whether the correlations hold at a population level when compared across counties in Colorado. It uses publicly available datasets provided by Colorado and United States government sources.

Methods

The Colorado Department of Public Health and Environment (CDPHE) county-level Colorado Health Indicators dataset ("CO Health") provides data for all 64 Colorado counties across a range of domains, including health, economics, and demographics. Its data appears as

headcounts, percentages, and rates (such as per 100,000 population). The US Department of Health and Human Services, Health Resources and Services Administration (2019) Area Health Resources Files (AHRF) dataset contains descriptive information about United States counties, as well as the health services available. Data from both the CDPHE and AHRF datasets were used in this study. See Appendix B for detailed information about the datasets. The data underwent preliminary analysis using Google Sheets (Google, n.d.) and RapidMiner Studio (RapidMiner, 2018). SAS® Studio was used to perform statistical analysis to test the hypotheses stated in Appendix A. All SAS® code is provided in Appendix E.

Limitations

This paper explores correlations between the percent of high school students who have seriously considered suicide and a variety of factors. Correlations can not be assumed to imply causation. Moreover, when several correlations are explored, mere chance can produce some correlations. Consequently, this exploratory work may provide insight for future research, but its findings should be taken with consideration of the number of correlations explored. Sakshaug and West (2014) emphasized that researchers must consider the variance inherent in survey data when using it to understand populations. This study relies on survey data that has different levels of confidence by county (see Figure 1 in Appendix C). This variability further underscores that findings are preliminary. As well, rates for people who have attempted suicide are known to differ from those who completed suicide (Shain, 2016). For reasons explained in "Ethical Considerations," actual youth suicide numbers were not available for this study. Consequently, the percentages of youth who considered suicide were used instead. Finally, Colorado was the focus of this study. Similar questions asked in different regions (nationally and globally) could

yield different results.

Ethical Considerations

The Center for Disease Control and Prevention (CDC) provides tallies for suicides by age groups within its CDC Wonder database. However, numbers lower than ten are suppressed to preserve anonymity. While that suppression is vital for privacy, it results in many Colorado counties with suppressed tallies for youth suicide. For example, a query to the CDC Wonder database for county-level tallies of people aged 15-24 who died by suicide in the most recent three years (2016-2018) returns numbers for only 11 of 64 counties (17%). For that reason, the percent of high school students who have seriously considered suicide (data from CO Health) was used instead of actual youth suicide rates. Unfortunately, as noted, attempted suicides and completed suicides are not interchangeable numbers. Shain (2016) explained that, among adolescents (15 to 19 years old), boys complete suicide at triple the rate of girls, while girls attempt suicide at double the rate of boys. Shain noted that the different rates correlate with girls choosing less-lethal methods (2016).

Findings

This research paper stipulated that a p-value < 0.05 indicates a significant finding. The strength of Rho was estimated following Dancey and Reidy's categorization for psychology (Akoglu, 2018). Rho absolute values between 0.1 and 0.3 are characterized as "weak," 0.4 and 0.6 as "moderate," and 0.7 and 0.9 as "strong" (Akoglu, 2018). See Appendix B "Understanding the Strength of the Relationship" for different approaches to categorization. Since correlations are assessed for the percentage of high school students who have seriously considered suicide, that percentage was first inspected for normalcy (see Figure 2 in Appendix C). Since the measure

does not reflect a normal distribution, nonparametric analyses were used throughout the study. Key visualizations of significant correlations are represented in figures in Appendix C. Comprehensive correlation results and scatterplot matrices are represented in Appendix D.

Economic Factors

The CO Health and AHRF datasets contained several variables that reflect economic factors. A few key economic measures were tested using Spearman correlation coefficients against the percent of high school students who seriously considered suicide. Three measures were negatively correlated with the percentage of high schoolers who seriously considered suicide (see Figure 3 in Appendix C):

- Median home value for owner-occupied homes (Rho= -0.51000, p=<.0001)
- Percent of housing units with more than one person per room (Rho= -0.38591, p=0.0023)
- Median household income (Rho=-0.33612, p=0.0086)

Two measures were positively correlated with the percentage of high school students who seriously considered suicide:

- Percent of related children (5-17 years old) in poverty, 2017 (Rho=0.33604, p=0.0087)
- Percent of children (< 18 years old) below poverty level, 2015 (Rho=0.26179, p=0.0433)

Among these five measures, “percent of housing units with more than one person per room” was not associated with the remaining four. However, the remaining four were all strongly correlated with each other (see Table 3 in Appendix D). Several economic factors showed no significant correlation: the percent of people below poverty (p=0.1100), the percent of households that received food stamps that had children (p=0.1720), the rural-urban continuum code (p=0.6530), and the percent of persons in deep poverty (p=0.6723) did not show significant correlation with

the percentage of high school students who considered suicide.

Health and Behavioral Factors

This study considered the correlation between health factors for children, high school students, and adults and the percent of high school students who seriously considered suicide. First, correlations were explored with positive health factors. The percent of high school students who ate substantial amounts of produce (Rho= -0.45824, p=0.0002) and the percent of physically active adults (Rho= -0.46014, p=0.0005) were moderately negatively correlated with high schooler's serious thoughts of suicide (see Figure 4 in Appendix C). They were moderately positively correlated with each other (Rho=0.50661, p=0.0001). There was no significant correlation for the percent of physically active high school students (p=0.0967) with those who considered suicide. Unfortunately, there was very little data for the percent of children aged 5-14 years who (a) were physically active or (b) ate substantial produce (only 11 of 64 counties reported numbers), and no significance could be determined.

Next, correlations were explored with health factors associated with alcohol and marijuana. The percent of adults who reported binge drinking in the past 30 days (Rho= -0.44126, p=0.0009) and the percent of adults who reported taking marijuana in the last 30 days (Rho= -0.37357, p=0.0148) were moderately negatively correlated with the percent of high school students who considered suicide. The percent of high school students who used marijuana in the past 30 days was weakly negatively correlated with those who have had suicidal thoughts (Rho= -0.27588, p=0.0329). The percent of adults who binge drink and those who recently took marijuana were moderately correlated with each other (Rho=0.43238, p=0.0027) (see Figure 5 in Appendix C). However, the percent of high school students who recently took marijuana was not

significantly correlated with rates of adults who recently took marijuana (p=0.0565) or adults who binge drank (p=0.2549). No significant correlations were found between thoughts of suicide and percent of high school students who drank five or more drinks in a few hours (p=0.0589) or the percent of high school students who reported driving while drinking alcohol (p=0.6756).

Finally, correlations were explored with health factors associated with additional issues related to health and behavior. The percent of high school students who are obese (Rho=0.51689, p<0.0001) and those who were ever diagnosed with asthma (Rho=0.50963, p<.0001) were both moderately correlated with rates of serious thoughts of suicide among high school students. Those factors showed no significant correlation with each other (p=0.2209). See Figure 6 in Appendix C. The percent of high school students who currently use cigarettes was not correlated with rates of youth suicidal ideation (p=0.0671). Several adult factors were moderately correlated with high school students who seriously considered suicide: the percent of adults with diabetes (Rho=0.59350, p<0.0001), the percent of obese adults (Rho=0.54391, p<0.0001), the age-adjusted rate of hospitalizations due to stroke (Rho=0.46718, p=0.0002), the percent of adults who reported fair or poor health (Rho=0.41712, p=0.0017), and the percent of inactive adults (Rho=0.41910, p=0.0018) were moderately, positively correlated. See Figure 7 in Appendix C. Those significant factors are all also moderately correlated with each other (see Table 8 in Appendix D). No significant correlations were found between high school students' thoughts of suicide and adults with asthma (p=0.2705).

Lastly, the percent of high school students who felt sad or hopeless almost every day for two or more consecutive weeks, interfering with usual activities during the past 12 months (Rho=0.47048, p=0.0001) and the percent of high school students who have ever had sexual

intercourse (Rho=0.41905, p=0.0009) were both moderately positively correlated with rates of high school students who seriously considered suicide. They were not correlated with each other (p=0.1581). See Figure 8 in Appendix C.

Conclusion

While youth suicide has been linked with poverty, this study found that, when looking at a population level, it is essential to consider which measures of poverty to use. The county-level median home and the percent of units with more than one percent per room were moderately, negatively correlated with youth thoughts of suicide, the median household income was weakly negatively correlated, and the county-level percent of children in poverty and the percent of related children living together in poverty were weakly positively correlated. This study identified several county-level health factors among adults (including rates of obesity, inactivity, diabetes, and hospitalization due to stroke) that were positively moderately correlated with the percentage of high school students who considered suicide. As well, it identified moderate correlations between rates of high school student health factors (obesity and diagnoses of asthma), longterm sadness, a history of sexual intercourse, and thoughts of suicide. Surprisingly, the percentage of adults who reported recent binge drinking and the percent of adults and high school students who took marijuana recently were negatively correlated with the percent of high school students having thoughts of suicide. Finally, the percent of high school students who ate substantial produce and the percent of active adults were moderately negatively correlated with rates of high school students who seriously considered suicide.

Recommendations

Compared with other Americans, Coloradans tend to be more healthy and less obese;

they also tend to have more problems with alcohol (Villegas, 2019). Those factors are important to consider when weighing the meaning of the study's perhaps surprising findings. For example, the negative correlations with rates of alcohol and marijuana use may be reflective of an association between substance use and the cohesiveness of social groups in Colorado. As Grimmond et al. (2019) found, research points to the importance of "belongingness" for reducing youth suicidal ideation. It may be useful to look for similar findings using a broader dataset to determine whether the correlations are reflected in data representing regions beyond Colorado. Grimmond et al. (2019) also found that feelings of being a burden may contribute to suicidal ideation. And yet, one of the negatively correlating factors was rates of having more than one person per room. While that number likely reflects areas with higher rates of poverty, it might simultaneously reflect increased familial connection and cohesiveness, which could counter suicidal ideation.

This study identified several county-level health factors among adults that were positively correlated with the percentage of high school students who considered suicide, including rates of adults with diabetes, obesity, fair or poor health, and inactivity. These factors are also correlated with each other. Consequently, it may be helpful to explore these factors further to determine which have the best predictive value - as well as whether any represent causal factors in youth suicides. Several Colorado counties do not currently track key metrics, such as adults who regularly use marijuana or several children's health factors. Colorado may be better able to understand how those factors impact youth suicide by collecting data from more counties.

References

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine, 18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Center for Disease Control and Prevention. (n.d.). CDC Wonder: About underlying cause of death, 1999-2018. [Dataset]. Retrieved from https://wonder.cdc.gov/ucd-icd10.html

Colorado Department of Public Health and Environment. (2019). Colorado Health Indicators. [Dataset]. Retrieved from https://www.colorado.gov/pacific/cdphe/colorado-health-indicators

Crepeau-Hobson, F., & Estes, J. (2019). Removing barriers to treatment: Evaluation of a youth suicide prevention program. *Journal of Youth Development, 3*, 146. https://doi.org/10.5195/jyd.2019.731

Daniel, B. K. (2019). Big Data and Data Science: A Critical Review of Issues for Educational Research. *British Journal of Educational Technology*, 50(1), 101–113.

Frequently asked questions. (n.d.). Retrieved from https://pcmh.org/frequently-asked-questions/

Google Sheets. [Online software]. Retrieved from http://sheets.google.com

Grimmond, J., Kornhaber, R., Visentin, D., & Cleary, M. (2019). A qualitative systematic review of experiences and perceptions of youth suicide. *PLoS ONE, 14*(6), 1–25. https://doi.org/10.1371/journal.pone.0217568

Health Resources and Services Administration. (n.d.). Area health resources files: County level data. [Dataset]. Retrieved from https://data.hrsa.gov/data/download

Hoffman, S. (2015). Citizen Science: The Law and Ethics of Public Access to Medical Big Data. *Berkeley Technology Law Journal, 30*(3), 1741–1806. https://doi.org/10.15779/Z385Z78

O'Keefe, V. M., Haroz, E. E., Goklish, N., Ivanich, J., Cwik, M. F., Barlow, A., & Celebrating Life Team. (2019). Employing a sequential multiple assignment randomized trial (SMART) to evaluate the impact of brief risk and protective factor prevention interventions for American Indian Youth Suicide. *BMC Public Health, 19*(1), 1–12. https://doi.org/10.1186/s12889-019-7996-2

PCMH (n.d.). Roadmap to Colorado's behavioral health system for children, youth, and families: 4 year strategic plan. Retrieved from https://pcmh.org/wp-content/uploads/2019/06/2018-Roadmap-to-Childrens-Behavioral-Health-4-Year-Strategic-Plan-1.pdf

Pitman, A., Tham, S.-G., Hunt, I. M., Webb, R. T., Appleby, L., & Kapur, N. (2019). Access to means of lethal overdose among psychiatric patients with co-morbid physical health problems: Analysis of national suicide case series data from the United Kingdom. *Journal of Affective Disorders*, 257, 173–179. https://doi.org/10.1016/j.jad.2019.06.027

RapidMiner. (2018). RapidMiner Studio (Version 9.1.000) [Software]. Available from https://docs.rapidminer.com/latest/studio/installation/

Resnik, D. B., Elliott, K. C., Soranno, P. A., & Smith, E. M. (2017). Data-Intensive Science and Research Integrity. *Accountability in Research: Policies & Quality Assurance, 24*(6), 344–358. https://doi.org/10.1080/08989621.2017.1327813

RapidMiner. (2018). RapidMiner Studio (Version 9.1.000) [Software]. Available from https://docs.rapidminer.com/latest/studio/installation/

SAS Institute Inc. SAS® Studio. [Computer software]. (2014). Retrieved from https://odamid.oda.sas.com/SASODAControlCenter/

Shain, B. (2016, July 1). Suicide and suicide attempts in adolescents. *American Academy of Pediatrics*, p e1-e11. Retrieved from https://pediatrics.aappublications.org/content/138/1/e20161420

Sakshaug, J. W., & West, B. T. (2014). Important Considerations When Analyzing Health Survey Data Collected Using a Complex Sample Design. American Journal of Public Health, 104(1), 15–16. https://doi.org/10.2105/AJPH.2013.301515

Stefanac, N., Hetrick, S., Hulbert, C., Spittal, M. J., Witt, K., & Robinson, J. (2019). Are young female suicides increasing? A comparison of sex-specific rates and characteristics of youth suicides in Australia over 2004–2014. *BMC Public Health, 1*, 1. https://doi.org/10.1186/s12889-019-7742-9

US Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Workforce, Rockville, MD. (2019). Area health resources files (AHRF) 2018-2019: County level data. [Dataset]. Retrieved from https://data.hrsa.gov/data/download

Villegas, A. (2019, June 12). Colorado has more alcohol deaths than nearly every other state. Retrieved from https://www.cpr.org/2019/06/12/colorado-has-more-alcohol-deaths-than-nearly-every-other-state/

Appendix A

Hypotheses

Previous researchers have found several elements that are correlated with youth suicide. Those studies were typically performed qualitatively, investigating the personal attributes for individuals. This study will study their results using county-level numbers to determine whether community attributes are similarly correlated with youth suicide.

Poverty and Substance Abuse

Grimmond, Kornhaber, Visentin, and Cleary found links between youth suicide and "thwarted belongingness," "perceived burdensomeness" (often related to strained family relationships and poverty), and "the capability for suicide" (2019, p. 19). They also noted that drug and alcohol use was a common risk factor or trigger for youth suicide. This study will ask whether there are correlations at the county-level for rates of poverty and substance abuse and youth suicide.

Question 1: Is there a link at the county-level between poverty and rates of high school students who seriously considered suicide (such as measured with median home value and rates of children in poverty)?

H0: The percentage of suicide ideation by high school students is not correlated with poverty.

H1: The percentage of suicide ideation by high school students is positively correlated with poverty.

Question 2: Is there a link at the county-level between rates of substance abuse and rates of high school students who seriously considered suicide?

H0: The rates of high school students who seriously considered suicide is not correlated with the use of alcohol, cigarettes, or other drugs.

H1: The rates of high school students who seriously considered suicide is correlated with the use of alcohol, cigarettes, or other drugs.

Health-Related Hypotheses

O'Keefe, Haroz, Goklish, Ivanich, Cwik, Barlow, and the Celebrating Life Team (2019) noted that suicide rates vary across racial and ethnic groups. Because their study was consistent with other studies, and because the comparison is at the same broad-level view as the current study, there is no need to repeat this question in the current study.

New questions were suggested by this study's preliminary exploration.

Question 3: Is there a link at the county-level between poor adult health and rates of high school students who seriously considered suicide?

H0: The rates of high school students who seriously considered suicide is not correlated with adult health.

H1: The rates of high school students who seriously considered suicide is correlated with poor adult health.

Question 4: Is there a link at the county-level between the rate of high schoolers with health issues and the rates of high school students who seriously considered suicide (such as measured with obesity and asthma)? (Note that this question is similar to Question 2, in that both look at high school students' negative health attributes).

H0: The rates of high school students who seriously considered suicide is not correlated with the percentage of high schoolers with health issues.

H1: The rates of high school students who seriously considered suicide is positively correlated with the percentage of high schoolers with health issues.

Question 5: Is there a link at the county-level between the rate of high schoolers with good health habits and the rates of high school students who seriously considered suicide (such as measured with eating fruits and vegetables and being active)?

H0: The rates of high school students who seriously considered suicide is not correlated with the percentage of high schoolers with good health habits.

H1: The rates of high school students who seriously considered suicide is correlated with the percentage of high schoolers with good health habits.

Question 6: Is there a link at the county-level between the rate of high schoolers who report two or more week periods of sadness and the rates of high school students who seriously considered suicide (such as measured with eating fruits and vegetables and being active)?

H0: The rates of high school students who seriously considered suicide is not correlated with the percentage of high schoolers who report two or more week periods of sadness.

H1: The rates of high school students who seriously considered suicide is correlated with the percentage of high schoolers who report two or more week periods of sadness.

Question 7: Is there a link at the county-level between the rate of high schoolers who have ever had sexual intercourse and the rates of high school students who seriously considered suicide?

H0: The rates of high school students who seriously considered suicide is not correlated with the percentage of high schoolers who have ever had sexual intercourse.

H1: The rates of high school students who seriously considered suicide is correlated with the percentage of high schoolers who have ever had sexual intercourse.

Mental Health Resources

Originally, mental health resources were going to be explored because the literature review found that using mental health resources is correlated with reduced suicidal ideation. The AHRF database contains several county-level fields related to child psychologists:

- All non-federal child psychologists in 2010, 2015, and 2017
- All non-federal child psychologists who saw patients in 2010, 2015, and 2017
- The number of hospitals providing (a) short-term and (b) long-term children's psychology services in 2017
- The number of inpatient days in (a) short-term and (b) long-term children's psychology hospitals in 2017

However, a review of the data showed that children's mental health was quite limited in Colorado. The AHRF database identified no Colorado counties with hospitals providing short- or long-term children's psychology services in 2017. Consequently, there were also no inpatient days. As well, there were 45 Colorado counties (70% of all Colorado counties) with no non-federal child psychologists throughout the time period studied. There were an additional four counties that had a non-practicing child psychologist in 2010, but not in 2015 or 2017. Consequently, of Colorado's 64 counties, only 15 counties (23.4%) had practicing non-federal child psychologists in 2017. This low number is, in itself, concerning. It is also likely too small to be useful in correlational analysis.

Appendix B

Datasets and Data Tools

Datasets

The Colorado Department of Public Health and Environment (CDPHE) county-level "Colorado Health Indicators" (CO Health) provides data for all 64 Colorado Counties across a range of domains, including health, economics, and demographics. Its data appears as headcounts, percentages, and rates (such as per 100,000 population). The US Department of Health and Human Services, Health Resources and Services Administration (2019) Area Health Resources Files (AHRF) dataset contains descriptive information about United States counties, as well as the health services available. It also provides comprehensive information about the medical and mental health services that are available in each county. In addition, it provides finely tuned information about county demographics. Since the original AHRF dataset contained several thousand columns as well as rows for all US states, a subset of the data was extracted for the study (using SAS Studio). Only Colorado data was included, and roughly 200 columns were extracted for this study.

Data Analysis Tools

The data underwent preliminary analysis using Google Sheets (Google, n.d.) and RapidMiner Studio (RapidMiner, 2018). First, several columns were inspected visually using Google Sheets pivot tables and charts. Potential relationships among variables were noted. Next, the CDPHE Colorado Health Indicators and AHRF dataset were merged using RapidMiner Studio. The merged dataset was explored using RapidMiner to assess how counties clustered. A k-means cluster tree identified clusters using economic variables. One cluster contained a high

percentage of single-person housing, and the two other clusters were distinguished from each other using a combination of comparisons for three fields: the median home value for owner-occupied housing, the number of children in poverty, and the percent of related children (5-17 years old) living together in poverty.

While the data contained a column showing the percentage of high school students who seriously considered attempting suicide during the previous 12 months (dated 2015), referred to below as "Percent HS Students Considering Suicide"), it was not possible to run RapidMiner predictive analysis on that variable because there were only 64 rows of data (one for each Colorado county). Instead, correlations between "Percent HS Students Considering Suicide" and other fields were explored. Not surprisingly, the strongest correlation was with "Percent of high school students who felt sad or hopeless almost every day for two or more weeks in a row so that they stopped doing some usual activities during the past 12 months, 2015" (0.634). As well, RapidMiner identified several positive correlations between "Percent HS Students Considering Suicide" and several variables indicating poor adult health. There were also correlations with measures for high school students related to risky and healthy behaviors as well as physical health. Finally, there was a correlation with the rate of mental health hospitalizations. RapidMiner identified two notable negative correlations: the median home value for owner-occupied housing units and the percent of high school students who ate two or more fruits and three or more vegetables per day in the previous seven days.

SAS Studio will be used to perform statistical analysis to test the stated hypotheses. RapidMiner may also be used to perform more exploratory analysis. Charts and graphs may be generated by SAS Studio or Google Sheets, depending on needs. As previously noted, all three

named tools were used to extract and merge datasets.

Understanding the Strength of the Relationship

This study explores correlations using Spearman's rho (r_s) to represent the strength of the relationship between variables. Akoglu (2018) explained that the meaning of correlation coefficient values varies by discipline. He provided a table with contrasting interpretations (see Table 1). For this study, Dancey and Reidy's categories for psychology will be used for this study.

Table 1
Interpretation of Pearson's and Spearman's correlation coefficients.

Correlation Coefficient		Dancey & Reidy (Psychology)	Quinnipiac University (Politics)	Chan YH (Medicine)
+1	−1	Perfect	Perfect	Perfect
+0.9	−0.9	Strong	Very Strong	Very Strong
+0.8	−0.8	Strong	Very Strong	Very Strong
+0.7	−0.7	Strong	Very Strong	Moderate
+0.6	−0.6	Moderate	Strong	Moderate
+0.5	−0.5	Moderate	Strong	Fair
+0.4	−0.4	Moderate	Strong	Fair
+0.3	−0.3	Weak	Moderate	Fair
+0.2	−0.2	Weak	Weak	Poor
+0.1	−0.1	Weak	Negligible	Poor
0iu	0	Zero	None	None

Note: Reproduced from Akoglu (2018, p. 92).

Appendix C

Key Figures

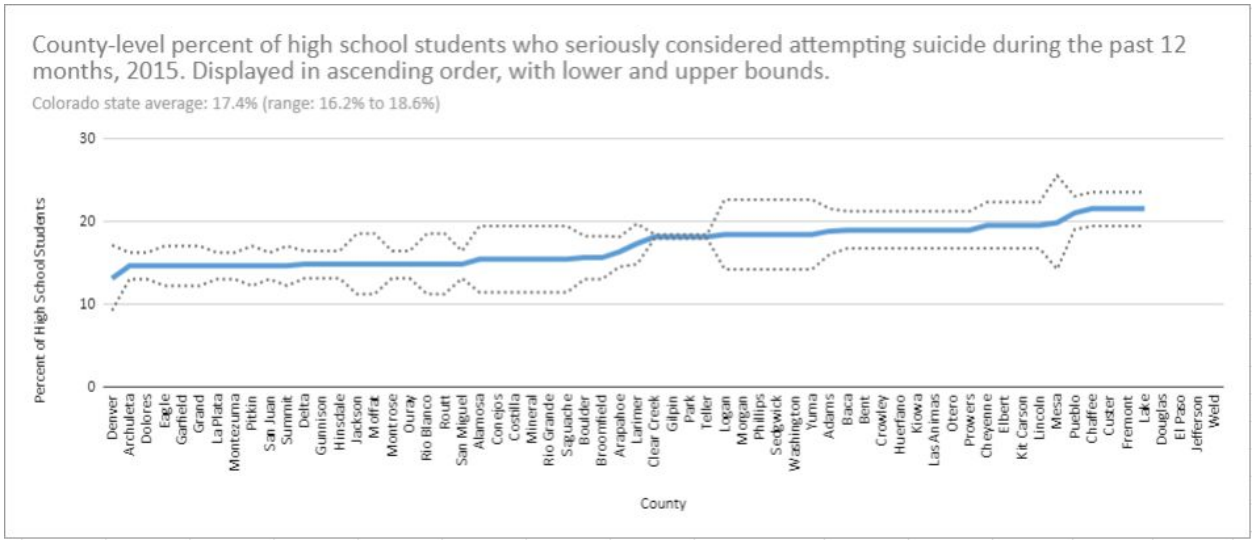

Figure 1. County-level data for high school students who seriously considered attempting suicide during the past 12 months, 2015 (CO Health, 2019). The blue line indicates the percent reported for each county. The dotted lines represent the upper and lower bounds reported in CO Health for each county. The four uncharted counties (Douglas, El Paso, Jefferson, and Weld) did not report numbers.

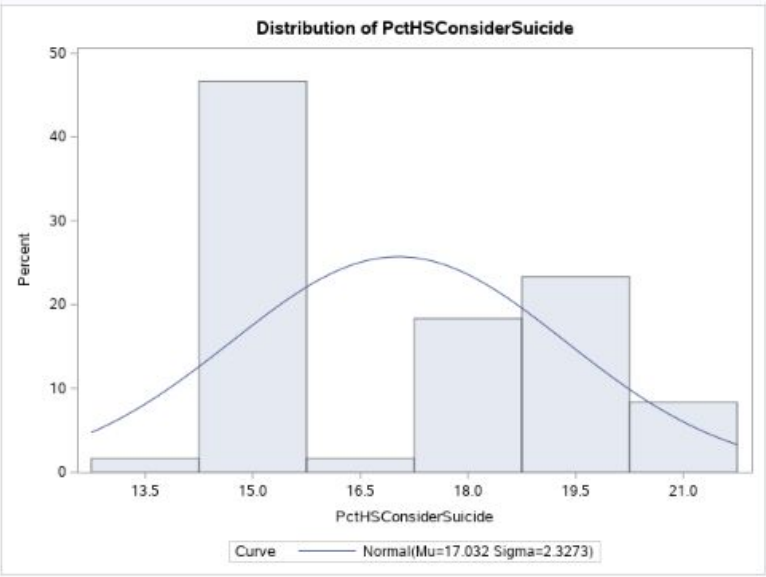

Figure 2. The histogram of county-level percentages of high school students in Colorado who seriously considered suicide in the previous year does not appear to have a normal distribution.

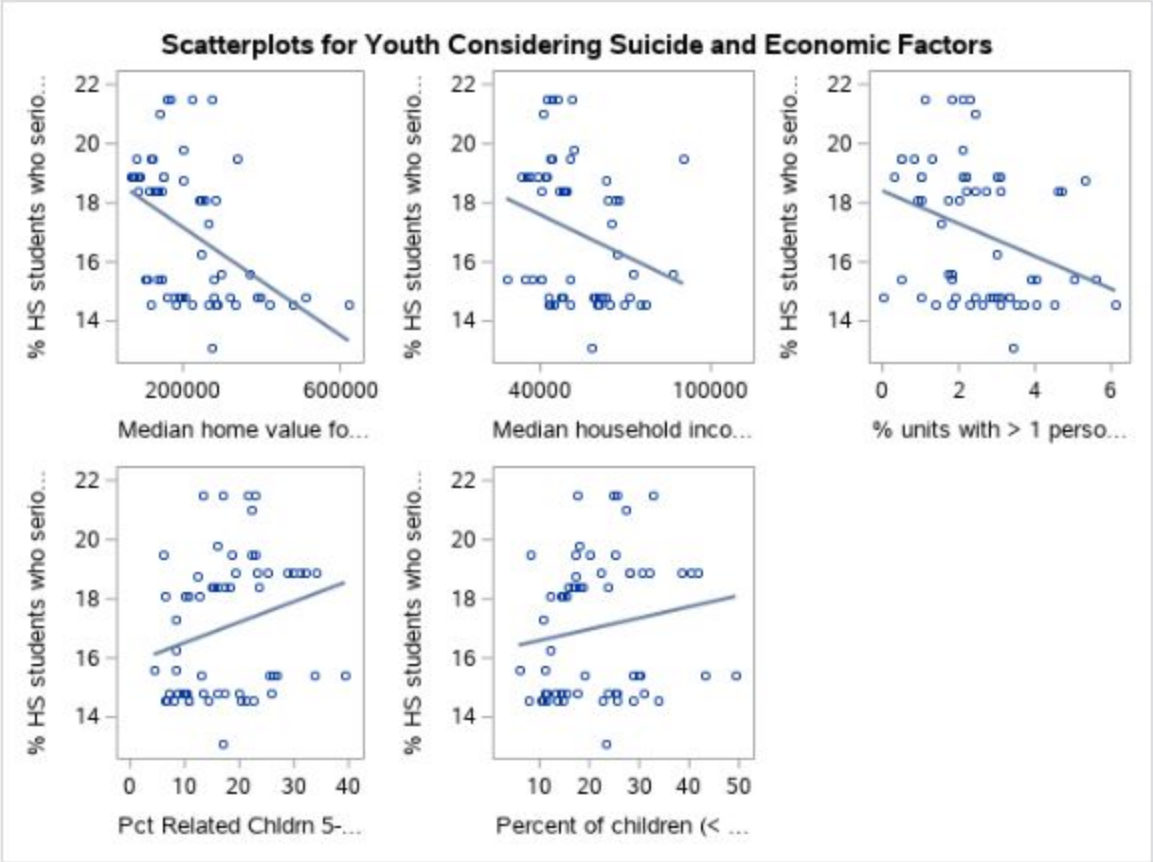

Figure 3. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Median home value in US dollars for owner-occupied housing units, 2011-2015" (CO Health, 2019), "Median household income (US dollars), 2015"(CO Health, 2019), "% units w/more than 1 person/rm (2013)" (AHRF), " Pct Related Chldrn 5-17 in Pov, 2013-2017" (AHRF), and "Percent of children (< 18 years old) below poverty level, 2015" (CO Health).

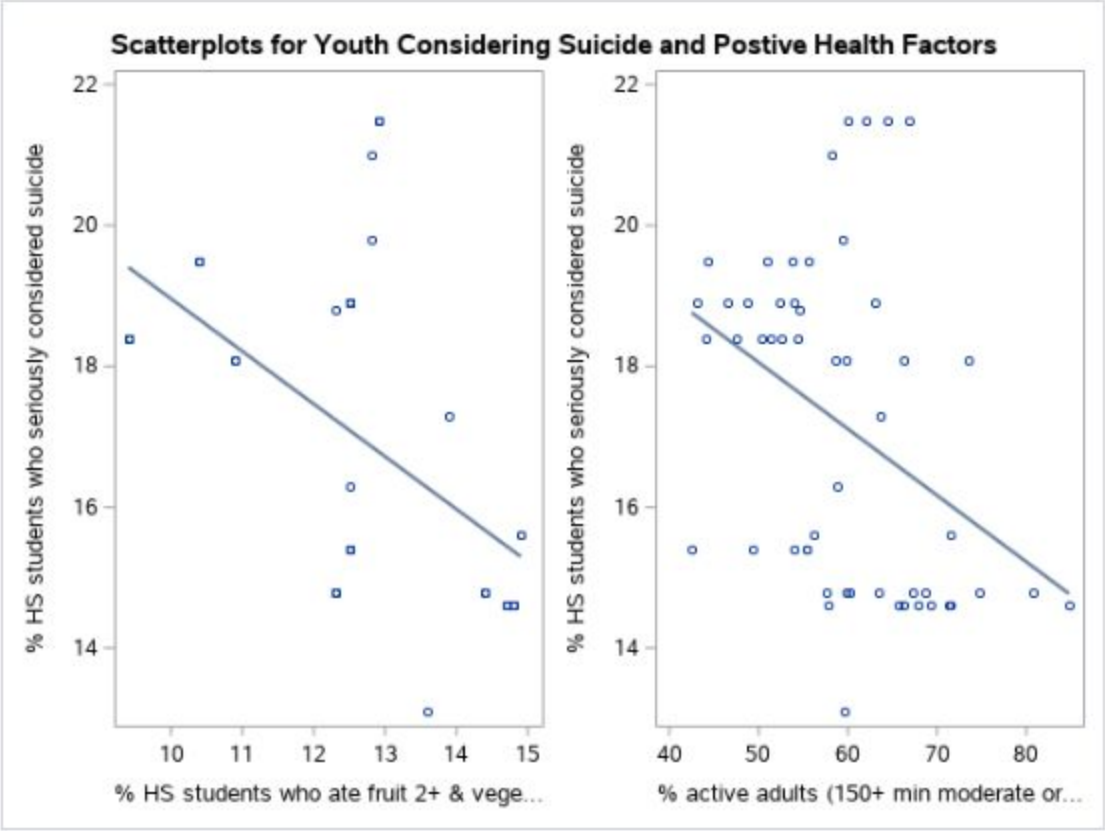

Figure 4. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Percent of adults aged 18+ years who participated in 150+ minutes of moderate or 75+ minutes of vigorous aerobic physical activity per week, 2011, 2013, 2015" (CO Health, 2019) and "Percent of high school students who ate fruit 2+ and vegetables 3+ times per day during the past 7 days, 2015"(CO Health, 2019).

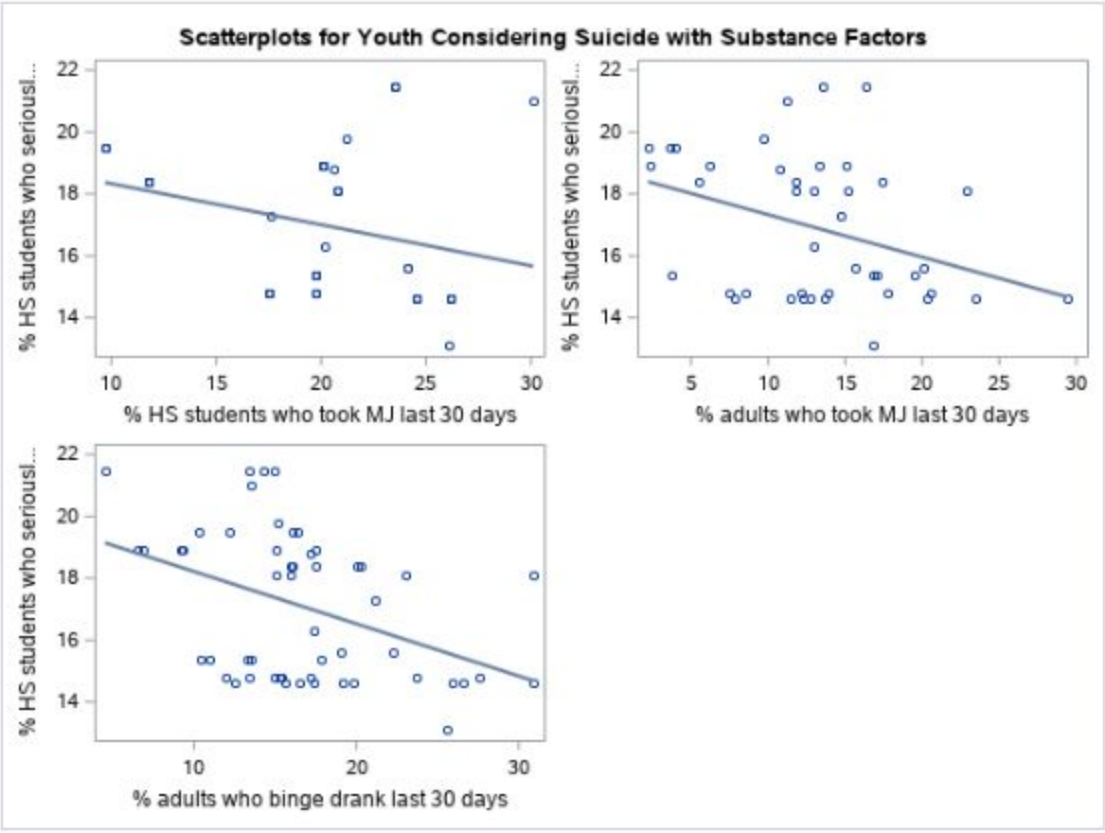

Figure 5. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Percent of adults aged 18+ years who reported binge drinking in past 30 days, 2013-2015" (CO Health, 2019), "Percent of adults aged 18+ years who used marijuana one or more days during the past 30 days, 2014-2015" (CO Health, 2019), and "Percent of high school students who used marijuana one or more times during the past 30 days, 2015" (CO Health, 2019). Among Colorado counties, seven did not report rates for adults who binge drink, and 18 did not report adults who recently used marijuana.

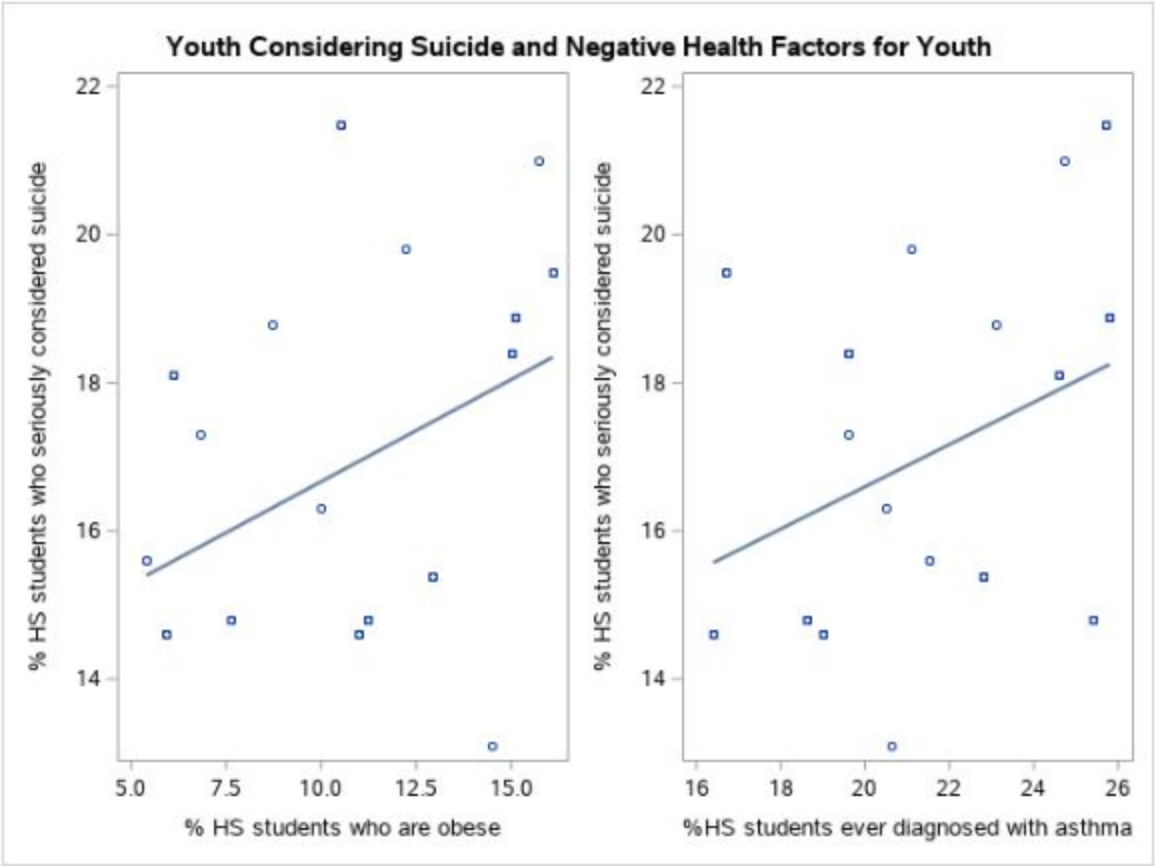

Figure 6. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Percent of high school students who are obese, 2015" (CO Health, 2019) and "Percent of high school students ever told they had asthma, 2015" (CO Health, 2019).

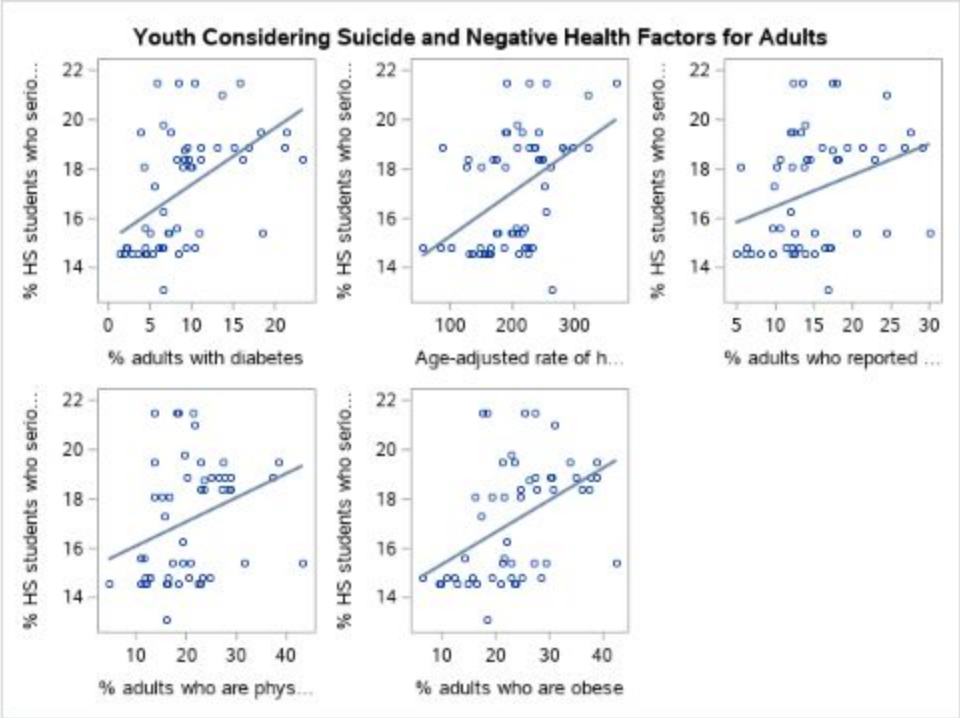

Figure 7. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Percent of adults aged 18+ years with diabetes, 2013-2015" (CO Health, 2019), "Age-adjusted rate of hospitalizations due to stroke (per 100,000 population), 2013-2015," (CO Health, 2019), "Percent of adults aged 18+ years who are obese, 2013-2015," (CO Health, 2019), "Percent of adults aged 18+ years who reported that their general health was fair or poor, 2013-2015" (CO Health, 2019), and "Percent of adults aged 18+ years who are physically inactive, 2013-2015" (CO Health, 2019).

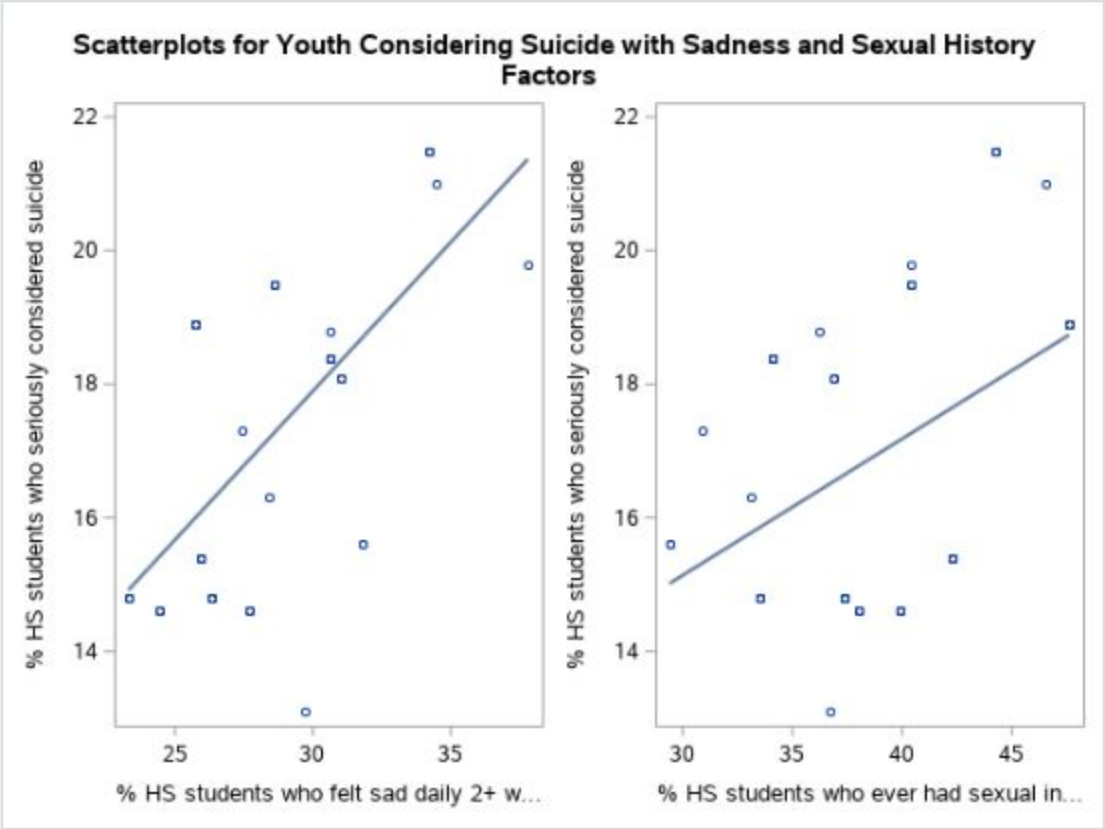

Figure 8. "Percent of high school students who seriously considered attempting suicide during the past 12 months, 2015" (CO Health, 2019) is shown on each vertical axis. Shown along the horizontal axes: "Percent of high school students who felt sad or hopeless almost every day for 2 or more weeks in a row so that they stopped doing some usual activities during the past 12 months, 2015"(CO Health, 2019) and "Percent of high school students who have ever had sexual intercourse, 2015" (CO Health, 2019).

Appendix D

Correlational Tables and Scatterplot Matrices from SAS® Studio

Poverty and High School Students who have Seriously Considered Suicide

This appendix contains output from SAS® Studio. Tables have been adjusted to provide additional clarification regarding variables that reflect economic factors within counties. Scatterplot matrices are provided to provide additional insight into potential relationships among variables.

Table 2
Spearman Correlation Coefficients for Poverty

		Percent of High School Students who Seriously Considered Suicide	
Measure	**Source**	Meaning	Rho P-value # Observations
Median home value in US dollars for owner-occupied housing units, 2011-2015	CO Health	Moderate, negative	-0.51000 <.0001 60
Percent of housing units with more than 1 person per room 2013-17	AHRF	Moderate, negative	-0.38591 0.0023 60
Median household income (US dollars), 2015	CO Health	Weak, negative	-0.33612 0.0086 60
Percent of related children 5-17 years old in poverty 2017	AHRF	Weak, positive	0.33604 0.0087 60
Percent of children (< 18 years old) below poverty level, 2015	CO Health	Weak, positive	0.26179 0.0433 60
Percent of persons below poverty level 2013-17	AHRF	Not significant	0.20846 0.1100 60
Percent of households that received	CO Health	Not significant	0.18019

food stamps in the past 12 months with children less than 18 years old, 2011-2015			0.17205 9
# Households w/1 Persons 2010	AHRF	Not significant	-0.12910 0.3256 60
Rural urban continuum code	AHRF	Not significant	-0.05923 0.6530 60
Percent of persons in deep poverty 2013-17	AHRF	Not significant	0.05574 0.6723 60
Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with p<0.05 indicating a significant correlation), and the number of observations.			

Table 3
Spearman Correlation Coefficients for Economic Factors that are Significantly Correlated with High School Students who have Seriously Considered Suicide

	Median home value (owner-occupied)	**% units with > 1 person per room**	**Median household income**	**% related children (5-17) in poverty**	**% children (<18) below poverty level**
Median home value for owner-occupied housing	1.00000	0.04266 0.7379	0.80919 <.0001	-0.81996 <.0001	-0.70013 <.0001
Percent units with > 1 person per room	0.04266 0.7379	1.00000	-0.08901 0.4843	0.08135 0.5228	0.10286 0.4186
Median household income	0.80919 <.0001	-0.08901 0.4843	1.00000	-0.94700 <.0001	-0.91476 <.0001
Percent of related children 5-17 in poverty 2017	-0.81996 <.0001	0.08135 0.5228	-0.94700 <.0001	1.00000	0.93666 <.0001
Percent of children (< 18 years old) below poverty level, 2015	-0.70013 <.0001	0.10286 0.4186	-0.91476 <.0001	0.93666 <.0001	1.00000
Note: each column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with p<0.05 indicating a significant correlation), and the number of observations.					

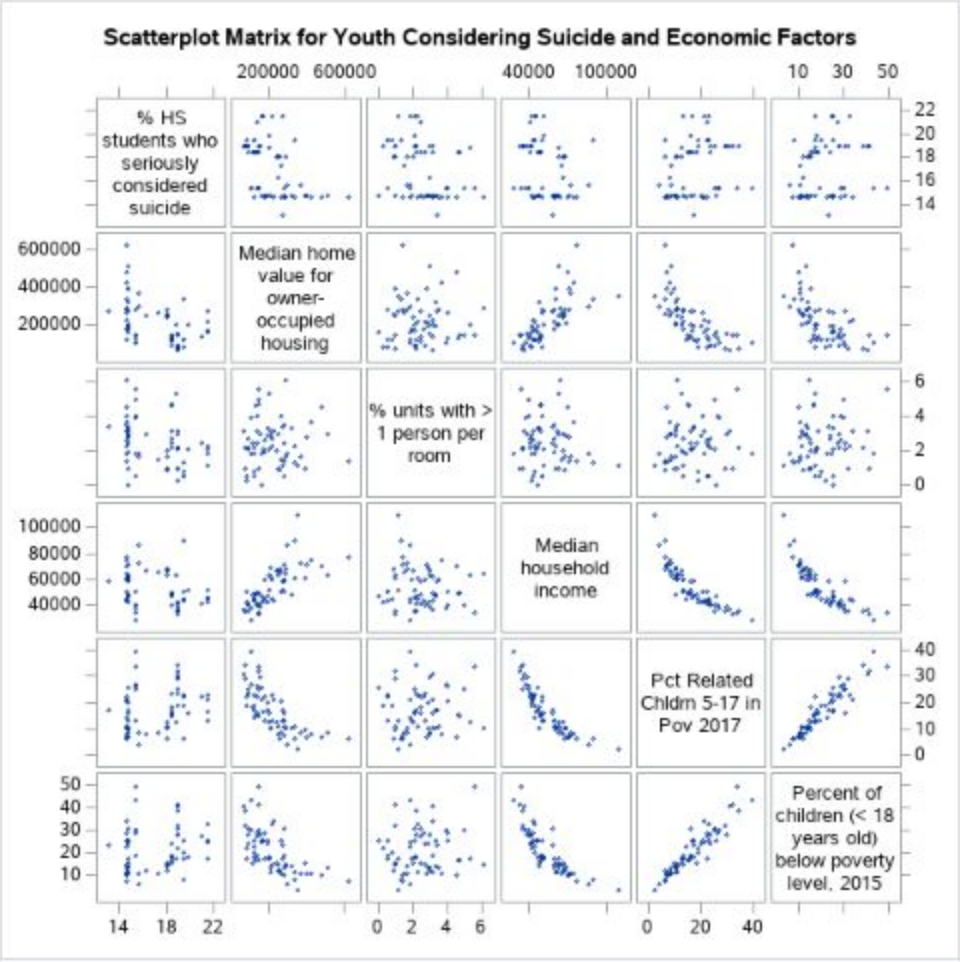

Figure 9. The scatterplot matrix between high school students who seriously considered suicide and attributes related to poverty shows a relationship between poverty and considerations for suicide.

Health Factors and High School Students who have Seriously Considered Suicide

This appendix contains output from SAS® Studio. Tables have been adjusted to provide additional clarification regarding variables that reflect health and behavioral factors within counties. Scatterplot matrices are provided to provide additional insight into potential relationships among variables.

Table 4
Spearman Correlation Coefficients for Positive Health-Related Factors

Measure	Source	Percent of High School Students who Seriously Considered Suicide	
		Meaning	Rho P-value # Observations
Percent of high school students who ate fruit 2+ and vegetables 3+ times per day during the past 7 days, 2015	CO Health	Moderate, negative	-0.45824 0.0002 60
Percent of adults aged 18+ years who participated in 150+ minutes of moderate or 75+ minutes of vigorous aerobic physical activity per week, 2011, 2013, 2015	CO Health	Moderate, negative	-0.46014 0.0005 53
Percent of high school students who were physically active for a total of at least 60 minutes/day for the past 7 days, 2015	CO Health	Not significant	0.21645 0.0967 60
Percent of children aged 5-14 years who were physically active for at least 60 minutes/ day for the past 7 days, 2013-2015	CO Health	Not significant	0.46429 0.2939 7
Percent of children aged 1-14 years who ate fruit 2 or more times per day and vegetables 3 or more times per day, 2013-2015	CO Health	Not significant	-0.10714 0.8192 7
Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.			

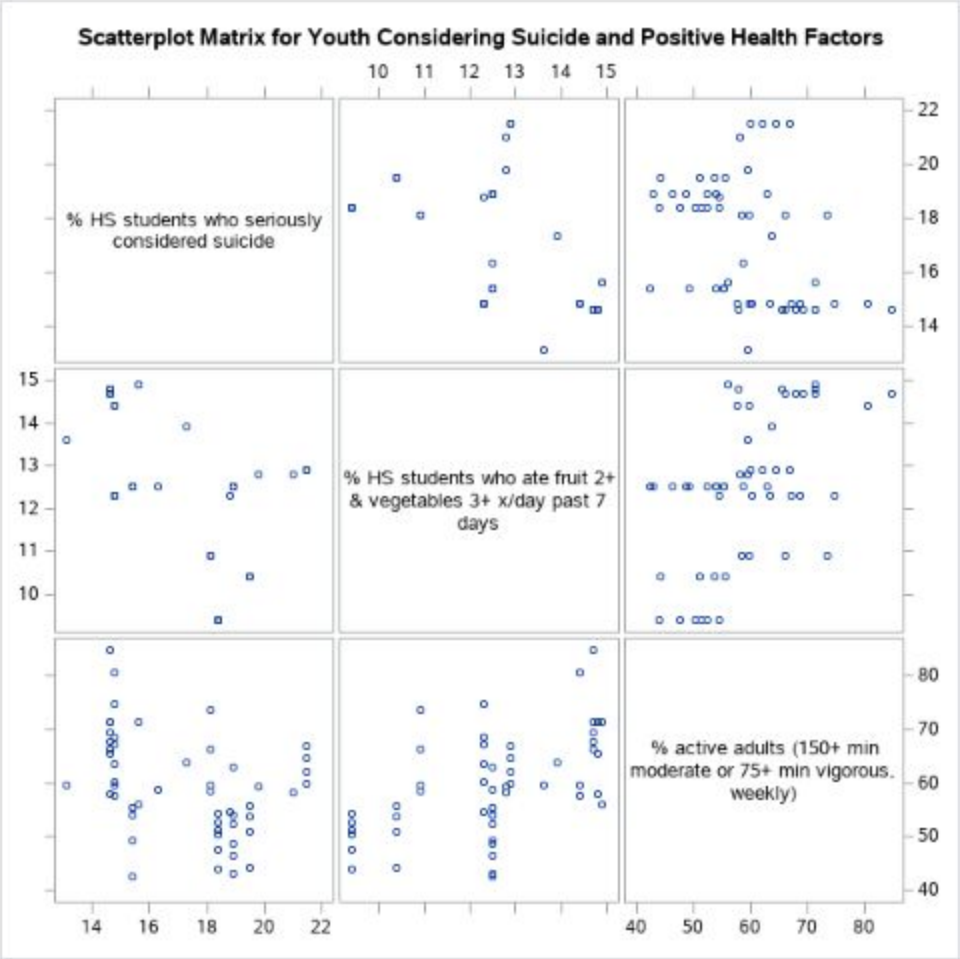

Figure 10. The scatterplot matrix between high school students who seriously considered suicide and attributes related to positive health factors.

Table 5
Spearman Correlation Coefficients for Factors Related to Alcohol and Marijuana

		Percent of High School Students who Seriously Considered Suicide	
Measure	**Source**	Meaning	Rho P-value # Observations
Percent of adults aged 18+ years who reported binge drinking in past 30 days, 2013-2015	CO Health	Moderate, negative	-0.44126 0.0009 53
Percent of adults aged 18+ years who used marijuana one or more days during the past 30 days, 2014-2015	CO Health	Moderate, negative	-0.37357 0.0148 42

Percent of high school students who used marijuana one or more times during the past 30 days, 2015	CO Health	Weak, negative	-0.27588 0.0329 60
Percent of high school students who had five or more drinks of alcohol within a couple of hours, 2015	CO Health	Not significant	-0.24529 0.0589 60
Percent of high school students who reported driving a car or other vehicle when they had been drinking alcohol, 2015	CO Health	Not significant	-0.05515 0.6756 60
Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.			

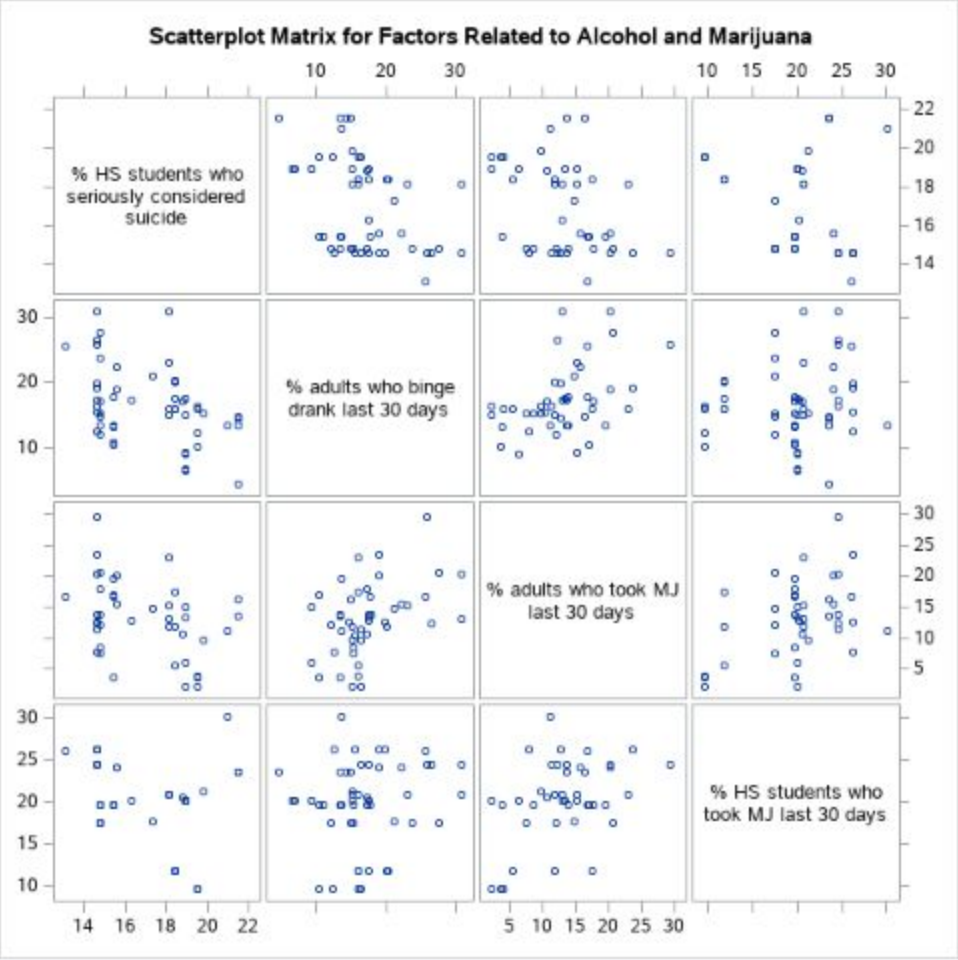

Figure 11. The scatterplot matrix between high school students who seriously considered suicide and attributes related to alcohol and marijuana.

Table 6
Spearman Correlation Coefficients for Negative Youth Health-Related Factors

		Percent of High School Students who Seriously Considered Suicide	
Measure	**Source**	Meaning	Rho P-value # Observations
Percent of high school students who are obese, 2015	CO Health	Moderate, positive	0.51689 <.0001 60
Percent of high school students ever told they had asthma, 2015	CO Health	Moderate, positive	0.50963 <.0001 60
Percent of high school students who currently use cigarettes, 2015	CO Health	Not significant	0.23797 0.0671 60

Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.

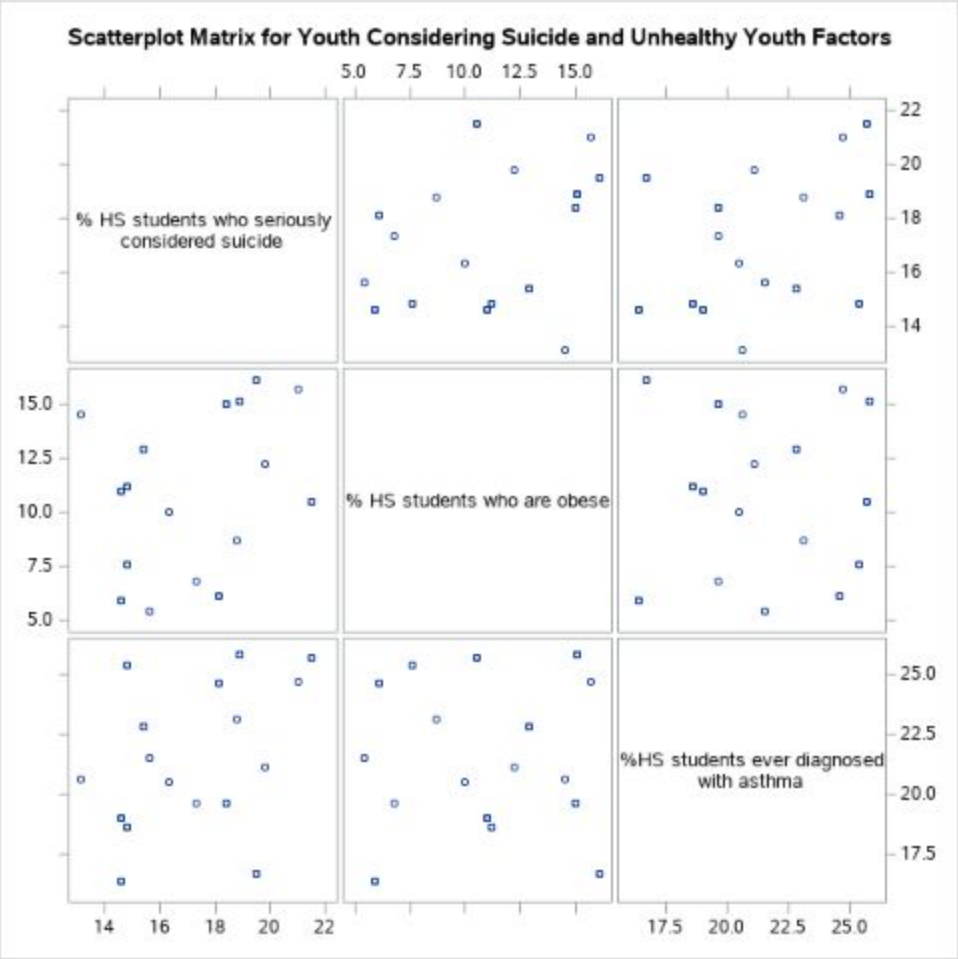

Figure 12. The scatterplot matrix between high school students who seriously considered suicide and attributes related to negative youth health status and behavior.

Table 7
Spearman Correlation Coefficients for Negative Adult Health-Related Factors

		Percent of High School Students who Seriously Considered Suicide	
Measure	**Source**	Meaning	Rho P-value # Observations
Percent of adults aged 18+ years with diabetes, 2013-2015	CO Health	Moderate, positive	0.59350 <.0001 54
Percent of adults aged 18+ years who are obese, 2013-2015	CO Health	Moderate, positive	0.54391 <.0001 53

Age-adjusted rate of hospitalizations due to stroke (per 100,000 population), 2013-2015	CO Health	Moderate, positive	0.46718 0.0002 59
Percent of adults aged 18+ years who reported that their general health was fair or poor, 2013-2015	CO Health	Moderate, positive	0.41712 0.0017 54
Percent of adults aged 18+ years who are physically inactive, 2013-2015	CO Health	Moderate, positive	0.41910 0.0018 53
Percent of adults aged 18+ years that currently have asthma, 2013-2015	CO Health	No significance	0.15265 0.2705 54
Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.			

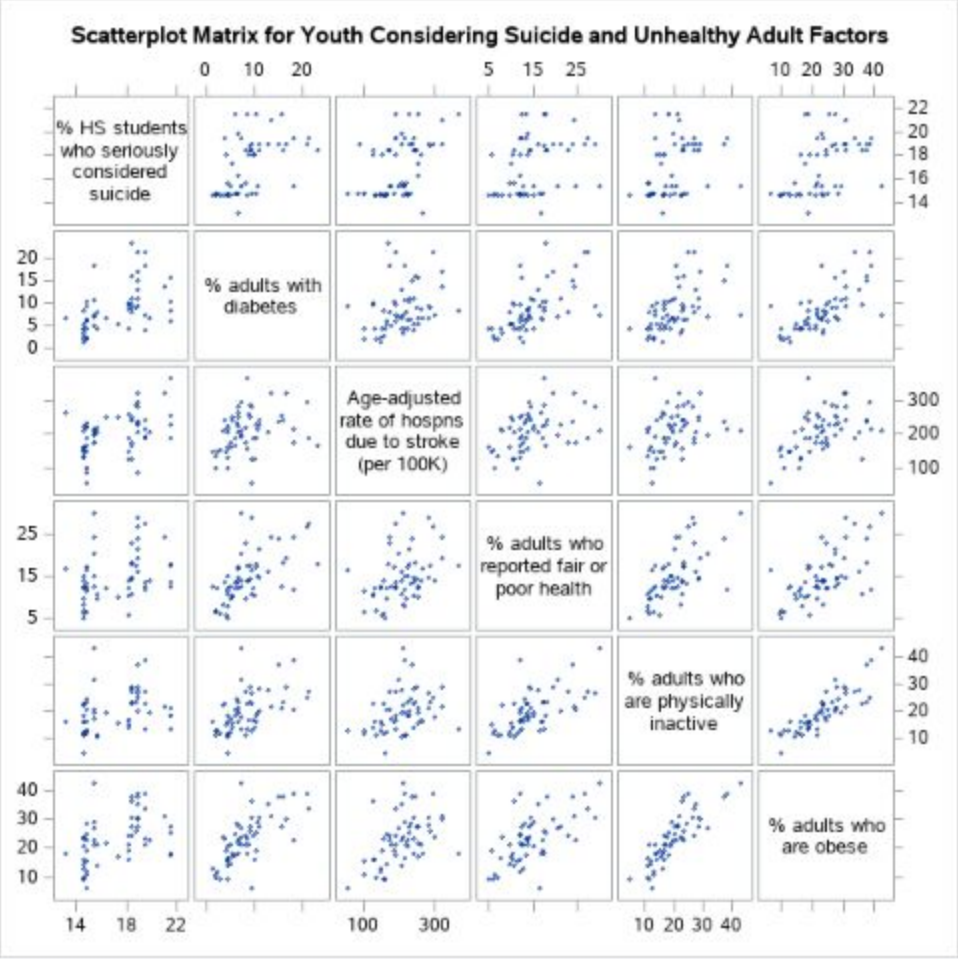

Figure 13. The scatterplot matrix between high school students who seriously considered suicide

and attributes related to adult poor health.

Table 8
Spearman Correlation Coefficients for Adult Negative Health-Related Factors that are Significantly Correlated with High School Students who have Seriously Considered Suicide

	% HS Students Considering Suicide	% Adults with Diabetes	Rate Hospitalization d/t Stroke	% Adults Reporting Fair or Poor Health	% Inactive Adults
% HS Students Considering Suicide	1.00000 60	0.59350 <.0001 54	0.46718 0.0002 59	0.41712 0.0017 54	0.41910 0.0018 53
% Adults with Diabetes	0.59350 <.0001 54	1.00000 58	0.36481 0.0049 58	0.66373 <.0001 58	0.57951 <.0001 57
Rate Hospitalization d/t Stroke	0.46718 0.0002 59	0.36481 0.0049 58	1.00000 63	0.43563 0.0006 58	0.40859 0.0016 57
% Adults Reporting Fair or Poor Health	0.41712 0.0017 54	0.66373 <.0001 58	0.43563 0.0006 58	1.00000 58	0.66494 <.0001 57
% Inactive Adults	0.41910 0.0018 53	0.57951 <.0001 57	0.40859 0.0016 57	0.66494 <.0001 57	1.00000 57

Note: each column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.

Table 9
Spearman Correlation Coefficients for Sexual History and Feelings of Sadness

		Percent of High School Students who Seriously Considered Suicide	
Measure	**Source**	Meaning	Rho P-value # Observations
Percent of high school students who	CO Health	Moderate,	0.47048

felt sad or hopeless almost every day for 2 or more weeks in a row so that they stopped doing some usual activities during the past 12 months, 2015		positive	0.0001 60
Percent of high school students who have ever had sexual intercourse, 2015	CO Health	Moderate, positive	0.41905 0.0009 60
Note: the fourth column contains three values: Rho (indicating the strength of the relationship), p-value (indicating the significance of the correlation, with $p<0.05$ indicating a significant correlation), and the number of observations.			

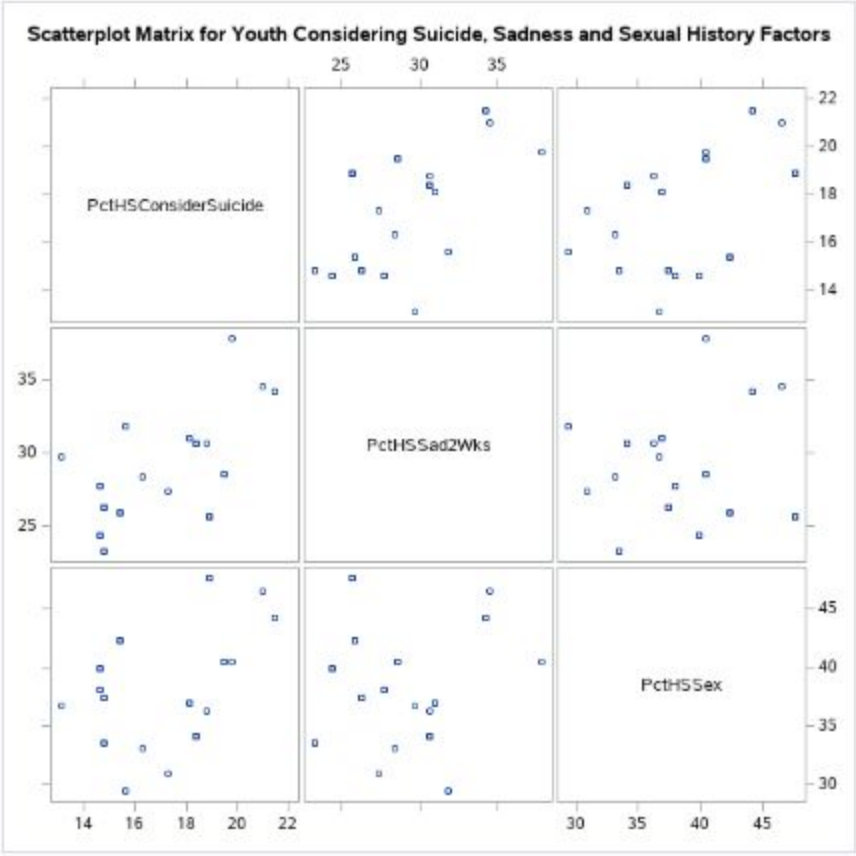

Figure 14. The scatterplot matrix between high school students who seriously considered suicide, high school students who felt sad or hopeless almost every day for two or more weeks in a row, and high school students who have ever had sexual intercourse.

Appendix E

SAS® Code

Building the Dataset

Area Health Resources Files

The Area Health Resources Files dataset was downloaded from the dataset page using the SAS link. As well, the technical files were downloaded. Next, abbreviated datasets were derived from the full dataset. First, Colorado data was extracted using the following SAS® code:

```
PROC SQL;
create table Port_581.ahrf2019_CO as
  select * from Port_581.ahrf2019
  where f12424="CO";
QUIT;
```

The data was inspected, the technical documentation was reviewed, and select columns were selected for the current study. A subset of the Colorado dataset was created using the following SAS® code:

```
PROC SQL;
CREATE TABLE PORT_581.ahrf2019_co_abbrv AS
SELECT   F00010, F1316615, F1389218, F1198418, F1406718, F1419518, F1389418,
F0002013, F1255913, F1397315, F1248115, F1248215, F1248315, F1248415,
F1546915, F1248615, F1397515, F1397615, F1533414, F1249014, F1547015,
F1397715, F1248715, F0355404, F0978719, F1249219, F1249218, F1249217,
F1249216, F1249215, F1249210, F1170417, F0474817, F1116017, F0477317,
```

```
F0479317, F1256717, F1256817, F1256917, F1257317, F1257417, F1556817,
F1258017, F1259017, F1527417, F1557217, F1557417, F1263717, F1556317,
F1526119, F1351310, F0873810, F0873910, F0874010, F0874110, F0874210,
F0874310, F1160310, F1348010, F1348110, F1348210, F1489110, F1316615,
F1541913, F1542013, F1542113, F1322517, F1332317, F1440813, F1443513,
F1525110, F1483813, F1526618, F0081176
FROM PORT_581.ahrf2019_co ;
QUIT;
```

The data was exported in CSV format for further review using Google Sheets and RapidMiner using the following SAS® code:

```
proc export data=PORT_581.ahrf2019_co_abbrv
outfile='/home/robinking0/MIS 581/Portfolio/ahrf2019_CO_abbrv.CSV'
dbms=csv
replace;
run;
```

Colorado Health

Initially, this researcher looked to the Colorado Health dataset page to download the dataset. However, the required data for this study was broken down into innumerable small downloads, available through various drop list links. Consequently, this researcher reached out to the contact on the dataset page, requesting a link to composite data in CSV format. A composite Microsoft Excel format spreadsheet was provided for each of the 64 counties. This researcher assembled all spreadsheets into a single file using Google Sheets. Because variable labels were

long, this researcher added abbreviated, descriptive variable names. Next, the dataset was exported as a CSV file from Google Sheets and imported into SAS using the following SAS® code:

```
FILENAME REFFILE '/home/robinking0/MIS 581/Portfolio/CoHealthExport.csv';
PROC IMPORT DATAFILE=REFFILE
        DBMS=CSV
        OUT=PORT_581.CO_Health;
        GETNAMES=YES;
RUN;
PROC CONTENTS DATA=PORT_581.CO_Health; RUN;
%web_open_table(PORT_581.CO_Health);
```

Combined Dataset

The AHRF and Colorado Health data were brought together using the following SAS® code:

```
PROC SQL;
CREATE TABLE PORT_581.ahrf_co_combo AS
SELECT  *
FROM PORT_581.AHRF2019_CO_ABBRV AHRF2019_CO_ABBRV
FULL JOIN PORT_581.CO_HEALTH CO_HEALTH
ON
  ( AHRF2019_CO_ABBRV.f00010 = CO_HEALTH.CountyName ) ;
QUIT;
```

Performing Analytics

Add Computed Fields

The rural continuum field was numeric, with values “01” to “09.” A numeric version of the field was added for some explorations using the following SAS® code:

```
*Code to get a numeric rural urban continuum code;
proc sql ;
ALTER TABLE PORT_581.AHRF_CO_COMBO
add  f0002013_num num 'Numeric Rural Urban Continuum Code';
quit;
proc sql ;
update PORT_581.AHRF_CO_COMBO
set f0002013_num  = input (f0002013,best.);
quit;
```

Add Labels

To improve table and graph visuals, labels were added to select variables using the following SAS® code:

```
data PORT_581.AHRF_CO_COMBO;
  set PORT_581.AHRF_CO_COMBO;
  label PctHSConsiderSuicide='% HS students who seriously considered suicide';
  label PctHSMJ='% HS students who took MJ last 30 days';
  label PctAdultMJ='% adults who took MJ last 30 days';
  label PctAdultBinge='% adults who binge drank last 30 days';
```

```
label PctHSSad2Wks ='% HS students who felt sad daily 2+ weeks';
label PctHSSex = '% HS students who ever had sexual intercourse';
label medHomValOwnOcc = 'Median home value for owner-occupied housing';
label MedhouseholdIncome ='Median household income';
label f1483813 = '% units with > 1 person per room';
label PctHSProduce ='% HS students who ate fruit 2+ & vegetables 3+ x/day past 7
    days';
label PctAdultActive ='% active adults (150+ min moderate or 75+ min vigorous,
    weekly)';
label PctAdultDiabetes = '% adults with diabetes';
label RateHospStroke  ='Age-adjusted rate of hospns due to stroke (per 100K)';
label PctAdultPoorHealth  ='% adults who reported fair or poor health';
label PctAdultInactive ='% adults who are physically inactive';
label PctHSCigarettes ='% HS students who currently smoke';
label PctHSObese = '% HS students who are obese';
label PctHSAsthma ='%HS students ever diagnosed with asthma';
label PctAdultObese='% adults who are obese';
label PctChildPov = 'Percent of children (< 18 years old) below poverty level, 2015';
run;
```

Explore Data

Descriptive analytics were performed, correlations tests were executed, and graphs were drawn using the following SAS® code:

```
* Run descriptive analytics on Key Variable PctHSConsiderSuicide;
proc UNIVARIATE;
  VAR PctHSConsiderSuicide;
run;
proc univariate normal plot data=PORT_581.AHRF_CO_COMBO;
var PctHSConsiderSuicide;
Title "Percent of High School Students who Seriously Considered Suicide";
histogram PctHSConsiderSuicide/normal;
run;

* Look at Histograms for Poverty Factors;
proc univariate normal plot data=PORT_581.AHRF_CO_COMBO;
var PctHSConsiderSuicide medHomValOwnOcc f1483813 MedhouseholdIncome
f1332317 PctChildPov ;
histogram /normal;
run;
* Look at Correlations for Poverty Factors;
proc corr data=PORT_581.AHRF_CO_COMBO spearman ;
        var PctHSConsiderSuicide;
        with medHomValOwnOcc f1483813 MedhouseholdIncome f1332317
PctChildPov F1440813 PctChildFoodStamps F0873810 f0002013_num f1541913   ;
run;
```

```
* it looks like there may be a bend at around $360K-370K median value;
* is there a correlation at lower values?;
proc corr data=PORT_581.AHRF_CO_COMBO spearman ;
        Title "Correlation: Suicidal Ideation and Median Home Value below $360K";
        var PctHSConsiderSuicide;
        with medHomValOwnOcc ;
        where medHomValOwnOcc <360000;
run; * results in -0.39731, 0.0032, 53;

proc corr data=PORT_581.AHRF_CO_COMBO spearman ;
        Title "Correlation: Suicidal Ideation and Median Home Value at or above
$360K";
        var PctHSConsiderSuicide;
        with medHomValOwnOcc ;
        where medHomValOwnOcc >=360000;
run; * results in -0.69437, 0.0834, 7;

* correlate significant factors with each other;
proc corr data=PORT_581.AHRF_CO_COMBO spearman ;
  var medHomValOwnOcc f1483813 MedhouseholdIncome f1332317 PctChildPov;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
```

```
 title "Scatterplot Matrix for Youth Considering Suicide and Economic Factors";
 matrix PctHSConsiderSuicide medHomValOwnOcc f1483813 MedhouseholdIncome
f1332317 PctChildPov ;
run;

*Compare Healthy Factors;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
   var PctHSConsiderSuicide;
   with PctHSProduce PctAdultActive PctHSActive PctChildActive PctChildProduce ;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
 title "Scatterplot Matrix for Youth Considering Suicide and Positive Health Factors";
 matrix PctHSConsiderSuicide PctHSProduce PctAdultActive  ;
run;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
        var PctHSProduce PctAdultActive ;
Run;

* check correlation of substance use;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
 title "Factors Related to Alcohol and Marijuana";
        var PctAdultMJ PctAdultBinge PctHSMJ ;
```

```
run;


*Compare Unhealthy Factors;
* skip PctChildAsthma PctChildMHProblems PctChildObese PctChildPoorHealth d/t too
few data points ;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
  title "Youth Considering Suicide and Unhealthy Factors";
    var PctHSConsiderSuicide;
    with    PctHSSad2Wks PctAdultDiabetes RateHospStroke PctHSSex PctHSCigarettes
          PctHSObese PctHSAsthma PctAdultBinge PctAdultMJ PctAdultPoorHealth
          PctAdultInactive PctHSMJ PctAdultObese  PctAdultAsthma
           PctHSBinge  PctHSDUI ;
run;
* check correlation of sadness and sexual history;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
  title "Youth Considering Suicide and Unhealthy Factors";
          var PctHSSad2Wks PctHSSex ;
run;
* check correlation of adult negative health factors;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
  title "Youth Considering Suicide and Unhealthy Factors";
          var PctHSConsiderSuicide PctAdultDiabetes RateHospStroke
```

```
                PctAdultPoorHealth  PctAdultInactive;
run;

* check correlation of significant negative health factors;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
  title "Youth Considering Suicide and Unhealthy Factors";
          var PctHSConsiderSuicide PctHSObese PctHSAsthma ;
run;

* check correlation of alcohol, MJ significant factors;
proc corr data=PORT_581.AHRF_CO_COMBO spearman;
  title "Youth Considering Suicide and Unhealthy Factors";
          var PctHSConsiderSuicide PctAdultBinge PctAdultMJ PctHSMJ ;
run;

* Count  the number of missing numbers for adult substance numbers;
* 7 null pctAdultBinge, 18 null PctAdultMJ;
proc sql ;
select count(*)
from port_581.ahrf_co_combo
where PctHSMJ is null ;
quit;
```

```
* scatter plot for alcohol and marijuana factors;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Factors Related to Alcohol and Marijuana";
  matrix PctHSConsiderSuicide PctAdultBinge PctAdultMJ PctHSMJ ;
run;
* scatterplot for youth health factors;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide and Unhealthy Youth Factors";
  matrix PctHSConsiderSuicide PctHSObese PctHSAsthma ;
run;
* scatterplot for adult health factors;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide and Unhealthy Adult Factors";
  matrix PctHSConsiderSuicide PctAdultDiabetes RateHospStroke PctAdultPoorHealth
     PctAdultInactive PctAdultObese ;
run;
* scatterplot for sadness and sex;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide, Sadness and Sexual History
Factors";
  matrix PctHSConsiderSuicide PctHSSad2Wks PctHSSex;
```

```
run;
*scatterplot for unhealthy factors;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide and Unhealthy Factors";
  matrix PctHSConsiderSuicide PctHSSad2Wks  PctAdultDiabetes RateHospStroke
PctHSCigarettes PctHSObese PctHSAsthma PctAdultBinge PctAdultMJ
PctAdultPoorHealth  PctAdultInactive PctHSMJ PctAdultObese PctHSSex PctAdultMJ ;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide, Sex, and Sadness";
  matrix PctHSConsiderSuicide PctHSSad2Wks  PctHSSex ;
run;

proc corr data=PORT_581.AHRF_CO_CO
* Check alcohol and marijuana stats;
MBO pearson ;
  title "Youth Considering Suicide and Alcohol and Marijuana Factors";
          var PctHSConsiderSuicide  PctAdultBinge PctAdultMJ PctHSMJ PctHSBinge
PctHSDUI ;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide and Unhealthy Factors";
```

```
  matrix PctHSConsiderSuicide  PctAdultBinge PctAdultMJ PctHSMJ PctHSBinge
PctHSDUI ;
run;
* Try a bubble format;
proc sgplot data=PORT_581.AHRF_CO_COMBO;
        bubble x=medHomValOwnOcc y=PctHSConsiderSuicide
size=MedhouseholdIncome/
                colorresponse=f1483813 bradiusmin=7 bradiusmax=14;
        xaxis grid;
        yaxis grid;
run;

proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Suicide and Poverty";
  matrix PctHSConsiderSuicide medHomValOwnOcc MedhouseholdIncome f1483813  /
group=f1533414;
run;
```

```
* Full Automatically created code appears below;


ods noproctitle;
ods graphics / imagemap=on;


proc corr data=PORT_581.AHRF_CO_COMBO pearson nosimple noprob plots=none;
        var PctHSSuicideAttempt;
        with f1332317 f1440813 f1483813 medHomValOwnOcc MedhouseholdIncome
                'NumAdultsNoHealthIns, 2014'n NumChildrenNoHealthIns
                NumDaysAdultBadPhyH NumDaysAdultHlthImpactAct NumGrpLiving
                NumHouseLingIsolated  PctAbovePovFoodStamps PctAdultActive
                PctAdultAsthma PctAdultBinge PctAdultInactive PctAdultObese
                PctAdultOverweightObese PctAdultPoorHealth PctHSActive
                PctHSAsthma PctHSBinge pctHSCigarettes PctHSDUI PctHSMJ
                PctHSObese PctHSOverweight PctHSProduce PctHSSad2Wks PctHSSex
                PctHSUsingBC;
run;


* Run graphs for scatterplots with trend lines;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplots for Youth Considering Suicide with Substance Factors";
  plot (PctHSConsiderSuicide ) * (PctHSMJ PctAdultMJ PctAdultBinge) /reg;
```

```
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplots for Youth Considering Suicide with Sadness and Sexual History
Factors";
  plot (PctHSConsiderSuicide)*(PctHSSad2Wks PctHSSex) / reg;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplots for Youth Considering Suicide and Economic Factors";
  plot (PctHSConsiderSuicide)*(medHomValOwnOcc MedhouseholdIncome f1483813
f1332317 PctChildPov ) / reg;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO ;
  title 'Scatterplots for Youth Considering Suicide and Postive Health Factors';
  plot (PctHSConsiderSuicide)*(PctHSProduce PctAdultActive) / reg;
run;
* check correlation of adult negative health factors;
proc sgscatter data=PORT_581.AHRF_CO_COMBO ;
  title "Youth Considering Suicide and Negative Health Factors for Adults";
  plot (PctHSConsiderSuicide)*(PctAdultDiabetes RateHospStroke  PctAdultPoorHealth
PctAdultInactive PctAdultObese) / reg;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO ;
```

```
 title "Youth Considering Suicide and Negative Health Factors for Youth";
 plot (PctHSConsiderSuicide)*(PctHSObese  PctHSAsthma ) / reg ;
run;
```

Explore Pediatric Mental Health Services in Colorado

This study initially considered studying pediatric mental health services in Colorado. However, preliminary exploration of the data (using the following SAS® code) found that the relevant data in the dataset was insufficient.

```
* Look into Mental Health Services;
* Compare with Number of Ped Psychs;
proc corr data=PORT_581.AHRF_CO_COMBO pearson ;
        var PctHSConsiderSuicide;
        with RateChildPsych RateChildPsychPtCare RatePsych ;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
 title "Scatterplot Matrix for Youth Considering Suicide and Psychologists Per
Population";
 matrix PctHSConsiderSuicide RateChildPsych RateChildPsychPtCare RatePsych ;
run;
* See Child Psych Availability;
proc sql;
Select  F00010 ,  F0474817,  F0474815 ,  F0474810 ,  F1116017 ,  F1116015 ,
F1116010, F1318717 ,  F1068317 ,  F1319017 , F1069017
```

```
from port_581.ahrf2019
where  F12424 = 'CO'
order by F0474817;
quit;

* There are 45 counties with ZERO for every measure of child mental health services!;
proc sql;
Select  F0002013 , count(*)
from port_581.ahrf2019
where  F12424 = 'CO'
and  (F0474817=0 and  F0474815 =0 and  F0474810 =0 and  F1116017 =0
and  F1116015 =0 and F1116010 =0 and  F1318717 =0 and
F1068317 =0 and F1319017 =0 and F1069017=0)
group by F0002013
order by F0002013 ;
Quit;
* compare counties with and without child mental health services;
proc sql;
Select  'No' 'Services', F0002013 'Contiuum', count(*) 'Count'
from port_581.ahrf2019
where  F12424 = 'CO'
and  ((F0474817 +F0474815 +F0474810 +F1116017 +F1116015 +F1116010 +F1318717
```

```
+F1068317 +F1319017 +F1069017)=0)
group by 1, 2
UNION
Select  'Yes' 'Services', F0002013 'Contiuum', count(*) 'Count'
from port_581.ahrf2019
where  F12424 = 'CO'
and  ((F0474817+F0474815 +F0474810 +F1116017 +F1116015 +F1116010 +F1318717
+ F1068317 +F1319017+F1069017)>0)
group by 1, 2
order by 1, 2;
quit;
* Look at the counties that do not have ZERO for every measure of child mental health
services;
* sort with the lowest total number first;
proc sql;
Select  F00010 , F0002013 , F0474817,  F0474815 ,  F0474810 ,  F1116017 ,  F1116015
, F1116010
from port_581.ahrf2019
where  F12424 = 'CO'
and  not(F0474817=0 and  F0474815 =0 and  F0474810 =0 and  F1116017 =0
and  F1116015 =0 and F1116010 =0 and  F1318717 =0 and
F1068317 =0 and F1319017 =0 and F1069017=0)
```

```
order by F0474817 +F0474815 +F0474810 +F1116017 +F1116015 +F1116010 +
F1318717 +F1068317 +F1319017 +F1069017;
quit;

* According to AHRF, Colorado has ZERO:;
*     Children's Psych ST Hosps 2017;
*     Children's Psych LT Hosps 2017;

*Compare Mental Health related rates with Suicide;
proc corr data=PORT_581.AHRF_CO_COMBO spearman ;
          var PctHSConsiderSuicide;
          with PctHSSad2Wks f1316615 RateSuicideHospital RateMHHospital;
run;
proc sgscatter data=PORT_581.AHRF_CO_COMBO;
  title "Scatterplot Matrix for Youth Considering Suicide and Overall Suicide Rate";
  matrix PctHSConsiderSuicide PctHSSad2Wks RateMHHospital RateSuicideHospital ;
run;

* Code to add computed fields to table;
* code to add rates of psychs; since counties have vastly different populations, rates
rather than numbers must be compared ;
proc sql ;
```

```
ALTER TABLE PORT_581.AHRF_CO_COMBO
add  RateChildPsych num 'Rate of Pediatric Psychologists';
quit;
proc sql ;
update PORT_581.AHRF_CO_COMBO
set RateChildPsych = f0474817/f1198418 ;
quit;
proc sql ;
ALTER TABLE PORT_581.AHRF_CO_COMBO
add  RateChildPsychPtCare num 'Rate of Child Psych, Total Patn Care';
quit;
proc sql ;
update PORT_581.AHRF_CO_COMBO
set RateChildPsychPtCare = f1116017/f1198418 ;
quit;
proc sql ;
ALTER TABLE PORT_581.AHRF_CO_COMBO
add  RatePsych num 'Rate of Psychologists';
quit;
proc sql ;
update PORT_581.AHRF_CO_COMBO
set RatePsych = f0477317/f1198418 ;
```

```
quit;
```