Eric Cartaya

EFFECTS OF CONTINUING COVID-19 EFFECTS ON MENTAL HEALTH REGARDING DEPRESSIVE AND ANXIETY DISORDERS

Submitted to:

Lynn A. Agre, MPH, PhD

Statistics 467/567

July 27, 2020

1. **Abstract**

This study investigates the association between the continuing Covid-19 pandemic and its effects on depression and anxiety across the United States. This study will specifically look at the effects across age groups, gender groups, and racial groups over a period of 9 weeks. Pressure regarding social distancing and social isolation have led to investigations from the CDC into effects on social isolation on various life factors such as employment and physical health, but also in regards of mental health. The data uses a modified version of the two-item Patient Health Questionnaire (PHQ-2) a modified version of the two-item Generalized Anxiety Disorder (GAD-2) scale to obtain data on the state of an individual’s symptoms regarding depression and anxiety, then cataloguing the averages across several factors such as race, education, gender, and location. There is a clear upward trend across the nine weeks for both depressive and anxious symptoms across the national average and across nearly all groups within the data. It can also be observed that Hispanic and Black individuals, individuals in the 18-29 age group, and women are more severely affected by the continuing crisis than their counterparts. Regression models will be built to determine whether or not it is truly increasing more for these groups or if there is no significant trend, as well as the correlation of the status to the actual increase. This research could be used to suggest that certain groups are either not receiving as effective of mental care or are suffering greater mental illness during this crisis. The research might also inform on greater mental health weaknesses across the United States.

Key words: Health; Covid-19; Depression; Anxiety; Household Pulse Survey; CDC

1. **Introduction**

During this crisis, the Covid-19 pandemic, the nation is currently experiencing both an overload of services and an exasperation of circumstances as the quarantine extends longer than expected and the effects of social isolation begin to affect the average American. This analysis will look in particular at groups of risk in four categories (age, sex, race, and education) to analyze whether or not they are more significantly affected by the social isolation of the pandemic. My thesis is that at risk groups such as women, people of color, those who lack higher education, and those who are young are more prone to mental illness during the pandemic and mental illness will affect them more over time as the pandemic develops.

The parameters of the study may be defined further by referencing the basics behind how the dataset was formed and presented. The dataset was produced by the Household Pulse Survey, a joint effort by the CDC’s National Center for Health Statistics (NCHS) and the Census Bureau “to rapidly monitor recent changes in mental health” (National Center for Health Statistics, 2020). The data that will be referenced and analyzed in this document is part of the larger study, used specifically to show the effects of the quarantine across different social groups across the nation. The dataset is divided into a table recording those who are either Depressed or Anxious and two other each designed to have Anxiety and Depression separated from each other in their own datasets. It is then divided into the national average, age groups, race groups, sex groups, education groups, and by states. This document will primarily observe at race, age, sex, and education, but will also look at the national average at certain points. The data uses a modified version of the two-item Patient Health Questionnaire (PHQ-2) a modified version of the two-item Generalized Anxiety Disorder (GAD-2) scale to obtain data on the state of an individual’s symptoms regarding depression and anxiety, with the dataset being a percentage of the population that qualifies as having major depression or major anxiety by the rules of these tests and the modification being a reduction of time from two weeks to one week. The GAD-2 test is a measure of generalized anxiety, although it has been used for many specific varieties of anxiety, and asks about “feeling nervous, anxious or on edge” (National HIV Curriculum, 2020) and “not being able to stop or control worrying” (National HIV Curriculum, 2020). Each one is measured on a scale of 0 thru 3 and a total score of 3 will indicate to the individual having anxiety. The PHQ-2 test is largely the same but for depression, although it is considered more preliminary than the GAD-2 test, and asks about “[having] little interest or pleasure in doing things” (National HIV Curriculum, 2020) and “feeling down, depressed or hopeless” (National HIV Curriculum, 2020). The scoring works largely the same but is considered a less reliable indicator for major depression and is more effective with moderate depression.

Using the evidence collected, three hypothesis have been formed. Firstly, is the mental health situation in the nation actually getting worse? This will be accomplished by seeing if there is a significant increase in the national percentage of those with major depression, anxiety, or overall. Next, are at risk groups (namely women, people of color, those with less education, and in the youngest age group (18-29)) more at risk than those who are not observed to be at risk from the effects? This will be measured by using Hotelling’s T2 on a modified version of the smaller datasets. Lastly, the rates of mental illness will be check to see if they are increasing at a more significant rate among these risk groups, and this will checked by looking at the different regression models for each dataset and checking to see both the coefficient and the significance of each variable.

1. **Methods**

As this document is a continuation of the previous part the first paragraph here will cover the expansion of variables from the previous iteration. No new variables were included in the analysis as the model before was complete but the previous iteration did not use any multivariate methods to create or analyze the model. Thus there is no need to repeat any of the previous methods of univariate or bivariate methods in this document.

In a continuation of the previous paragraph, the new methods (both bivariate and multivariate) must be addressed before continuation. Beginning with the bivariate methods, Anova, the Kruskall-Wallis Test, the Chi-square Test of Homogeneity (in the form of the Bartlett Test), Simultaneous Confidence intervals, and Hotelling’s T2 Test were used to analyze the data. Anova was primarily selected to allow for a further analysis of the similarities between the different subgroups in the model, it was hoped that there would be an indication that each of the subgroups were significantly different from one another and that the Kruskall-Wallis Test could be used to expand on this and show the significance. The Chi-square Test of Homogeneity was used to show the variances of each of the subgroups was significantly different from each other, giving an indication to the difference in changes of the subgroups. The simultaneous confidence intervals were used to show the difference in the subgroups through modeling (again) but were also used to show how the rate of change varied between the different models if only through observation. Lastly, Hotelling’s T2 is a legacy from the previous iteration of the project. In the previous version of the project it was used to show that there was a significant difference between the means of each of the subgroups and then displayed the actual means together, this was reproduced as part of this project due to the failure of the models to be useful.

Next the Regression Analysis itself will be covered. Throughout testing both the General Linear Model and the Linear Model were both used (GLM and LM); each were initially tested and it was discovered that they resulted in the same model so to take advantage of the models individual statistics and summaries. For each residual plots were used to check for normality of residuals the spread of the predicted values and the leverage of each of the models. The R-Square and P-Values were used to check to see how effective (and significant) the models truly were. For the likelihood ratio tests, the stepwise regression, and the checking of covariance matrices these methods were skipped. Despite their uses and being required they were deemed inappropriate as there as only one possible model in the way that the data was structured. With so few variables to work with in an effective manner the model was unable to be significantly manipulated.

When looking at the methods of multivariate analysis the two analysis methods used were Principal Components Analysis and Cluster Analysis. Cluster Analysis was selected as it is used to show the grouping of data and is ideal for the type of dataset produced. By showing how data groups we can understand the subgroups and their effects further. Principal Components analysis was selected as correlation and covariance were not significantly explored elsewhere in the document and it would give a more complete picture of the data.

1. **Results**

Beginning with the linear models there are two things apparent. The first is that most of the models did not work very well and did not give the information that was intended to be extracted from them. The models were intended to show the changes over time for each of the subgroups but ended up being more significantly influenced by the actual means. For this reason all parts of this document that were to rely on the models are considered a failure and are interesting at best. The second thing of note is that there is one model that did succeed and that is the national model, which shows the national increase over time. This does in fact give us the evidence needed not to reject the idea that the rate of mental illness is increasing in the United States.

For the reason that the models were a failure this document will look primarily at items not affected by the model. Beginning with the Chi-Square Test we can see from the results that the variances are non-identical across all of the different subgroups. This leads into the idea that the rates at which they are changing are significantly different throughout the data and that some groups are being more significantly affected than others. Through our observations in graphs and later on in cluster analysis we will see that some groups are much more heavily influenced than others. Next looking at the T2 test we can see that there are significant differences between the means in each of the subgroups for every group. This lead into the idea that we cannot reject the null hypothesis that all of the subgroups are at the same level and that there are some groups that are more at risk than others. We can see specifically that women, those who are mixed race, people in the 18-29 age group, and those who have no High School formal education are at the greatest risk of mental illness in the United States at the current moment. Conversely those who are male, Asian, in the 80 and above group, and have a college degree are at the least risk of mental illness at the current moment.

(I don’t know where this paragraph went when revising so I just copied what was in the presentation. Sorry.) Looking through the clusters we begin to see both the levels and the trends in more clear terms. Two different cluster structures were used as different attempts to capture which number of cluster should use resulted in 3 clusters and 5 cluster, thus they were both used. Both used k-means. Through both of these we can get some interesting observations, such as the higher distributions of the at risk groups (young adults pictured here) and get an indication of which groups have been changing the most (such as the elderly). The cluster table was used over the cluster diagram mostly due to readability and the fact that the cluster table gave more interesting indications.

Dataset (I)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Group | Subgroup | Week | Value | Low.CI | High.CI |
| 1 | National Estimate | United States | 1 | 35.9 | 35 | 36.8 |
| 2 | By Age | 18 - 29 years | 1 | 46.8 | 44.3 | 49.3 |
| 3 | By Age | 30 - 39 years | 1 | 39.6 | 37.7 | 41.5 |
| 4 | By Age | 40 - 49 years | 1 | 38.9 | 37.2 | 40.7 |
| 5 | By Age | 50 - 59 years | 1 | 35.8 | 34 | 37.7 |
| 6 | By Age | 60 - 69 years | 1 | 28.9 | 27.6 | 30.3 |
| 7 | By Age | 70 - 79 years | 1 | 21.5 | 19.3 | 23.7 |
| 8 | By Age | 80 years and above | 1 | 21.1 | 15.6 | 27.5 |
| 9 | By Gender | Male | 1 | 31 | 29.6 | 32.3 |
| 10 | By Gender | Female | 1 | 40.7 | 39.7 | 41.6 |
| 11 | By Race/Hispanic ethnicity | Hispanic or Latino | 1 | 42.7 | 39.7 | 45.7 |
| 12 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 1 | 33.6 | 32.7 | 34.4 |
| 13 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 1 | 38.9 | 36.4 | 41.4 |
| 14 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 1 | 31.9 | 28.5 | 35.4 |
| 15 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 1 | 43.9 | 39.8 | 48.1 |
| 16 | By Education | Less than a high school diploma | 1 | 45.4 | 40.3 | 50.5 |
| 17 | By Education | High school diploma or GED | 1 | 36.7 | 35.2 | 38.2 |
| 18 | By Education | Some college/Associate's degree | 1 | 38.5 | 36.9 | 40 |
| 19 | By Education | Bachelor's degree or higher | 1 | 30.7 | 29.8 | 31.5 |
| 20 | National Estimate | United States | 2 | 34.4 | 33.3 | 35.5 |
| 21 | By Age | 18 - 29 years | 2 | 47.4 | 43.4 | 51.3 |
| 22 | By Age | 30 - 39 years | 2 | 39.3 | 37.4 | 41.2 |
| 23 | By Age | 40 - 49 years | 2 | 36.9 | 34.1 | 39.7 |
| 24 | By Age | 50 - 59 years | 2 | 35.5 | 33 | 37.9 |
| 25 | By Age | 60 - 69 years | 2 | 25.5 | 23.2 | 27.9 |
| 26 | By Age | 70 - 79 years | 2 | 19.8 | 17.7 | 22 |
| 27 | By Age | 80 years and above | 2 | 13.9 | 10.9 | 17.3 |
| 28 | By Gender | Male | 2 | 31.4 | 29.5 | 33.3 |
| 29 | By Gender | Female | 2 | 37.2 | 36.2 | 38.3 |
| 30 | By Race/Hispanic ethnicity | Hispanic or Latino | 2 | 40.3 | 36.6 | 44 |
| 31 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 2 | 32 | 31.1 | 32.9 |
| 32 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 2 | 38.8 | 34.2 | 43.5 |
| 33 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 2 | 29.9 | 25.7 | 34.4 |
| 34 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 2 | 45.9 | 41.1 | 50.7 |
| 35 | By Education | Less than a high school diploma | 2 | 42.5 | 36.2 | 48.9 |
| 36 | By Education | High school diploma or GED | 2 | 37.8 | 35.3 | 40.4 |
| 37 | By Education | Some college/Associate's degree | 2 | 36.2 | 34.6 | 37.9 |
| 38 | By Education | Bachelor's degree or higher | 2 | 27.5 | 26.3 | 28.7 |
| 39 | National Estimate | United States | 3 | 33.9 | 33.1 | 34.7 |
| 40 | By Age | 18 - 29 years | 3 | 47.7 | 45.1 | 50.4 |
| 41 | By Age | 30 - 39 years | 3 | 37.8 | 35.7 | 39.9 |
| 42 | By Age | 40 - 49 years | 3 | 35.3 | 33.5 | 37.1 |
| 43 | By Age | 50 - 59 years | 3 | 33.7 | 31.8 | 35.7 |
| 44 | By Age | 60 - 69 years | 3 | 26.4 | 24.7 | 28.1 |
| 45 | By Age | 70 - 79 years | 3 | 18.3 | 16.3 | 20.4 |
| 46 | By Age | 80 years and above | 3 | 17.1 | 13.4 | 21.5 |
| 47 | By Gender | Male | 3 | 29.9 | 28.7 | 31.1 |
| 48 | By Gender | Female | 3 | 37.6 | 36.5 | 38.7 |
| 49 | By Race/Hispanic ethnicity | Hispanic or Latino | 3 | 40.7 | 38 | 43.4 |
| 50 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 3 | 31.3 | 30.4 | 32.1 |
| 51 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 3 | 39.4 | 37.2 | 41.6 |
| 52 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 3 | 30.2 | 27.1 | 33.5 |
| 53 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 3 | 37.2 | 33.3 | 41.2 |
| 54 | By Education | Less than a high school diploma | 3 | 40.9 | 36.2 | 45.8 |
| 55 | By Education | High school diploma or GED | 3 | 37.6 | 36 | 39.3 |
| 56 | By Education | Some college/Associate's degree | 3 | 36.1 | 34.5 | 37.8 |
| 57 | By Education | Bachelor's degree or higher | 3 | 26.4 | 25.3 | 27.5 |
| 58 | National Estimate | United States | 4 | 34.3 | 33.6 | 35.1 |
| 59 | By Age | 18 - 29 years | 4 | 46.6 | 44 | 49.2 |
| 60 | By Age | 30 - 39 years | 4 | 39.5 | 37.9 | 41.1 |
| 61 | By Age | 40 - 49 years | 4 | 35.7 | 34.2 | 37.4 |
| 62 | By Age | 50 - 59 years | 4 | 33.7 | 32 | 35.4 |
| 63 | By Age | 60 - 69 years | 4 | 27.4 | 25.7 | 29.1 |
| 64 | By Age | 70 - 79 years | 4 | 18.8 | 16.5 | 21.2 |
| 65 | By Age | 80 years and above | 4 | 20.1 | 14 | 27.5 |
| 66 | By Gender | Male | 4 | 30.3 | 29.1 | 31.5 |
| 67 | By Gender | Female | 4 | 38.1 | 37 | 39.2 |
| 68 | By Race/Hispanic ethnicity | Hispanic or Latino | 4 | 41.5 | 38.6 | 44.5 |
| 69 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 4 | 32.3 | 31.4 | 33.3 |
| 70 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 4 | 35.6 | 32.9 | 38.3 |
| 71 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 4 | 27.6 | 24.5 | 30.8 |
| 72 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 4 | 43.4 | 39.5 | 47.4 |
| 73 | By Education | Less than a high school diploma | 4 | 47.7 | 43.4 | 52.1 |
| 74 | By Education | High school diploma or GED | 4 | 35.3 | 33.6 | 37.1 |
| 75 | By Education | Some college/Associate's degree | 4 | 37.6 | 36.4 | 38.8 |
| 76 | By Education | Bachelor's degree or higher | 4 | 27 | 26.1 | 27.9 |
| 77 | National Estimate | United States | 5 | 35.5 | 34.8 | 36.1 |
| 78 | By Age | 18 - 29 years | 5 | 49.3 | 47.4 | 51.2 |
| 79 | By Age | 30 - 39 years | 5 | 40.6 | 39.2 | 42 |
| 80 | By Age | 40 - 49 years | 5 | 36.9 | 35.3 | 38.6 |
| 81 | By Age | 50 - 59 years | 5 | 33.6 | 31.9 | 35.4 |
| 82 | By Age | 60 - 69 years | 5 | 27 | 25.6 | 28.5 |
| 83 | By Age | 70 - 79 years | 5 | 21.7 | 19.3 | 24.2 |
| 84 | By Age | 80 years and above | 5 | 19.2 | 14.7 | 24.4 |
| 85 | By Gender | Male | 5 | 31.1 | 30.2 | 32.1 |
| 86 | By Gender | Female | 5 | 39.6 | 38.9 | 40.3 |
| 87 | By Race/Hispanic ethnicity | Hispanic or Latino | 5 | 38.3 | 36.2 | 40.4 |
| 88 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 5 | 33.1 | 32.3 | 33.9 |
| 89 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 5 | 40.5 | 38.3 | 42.7 |
| 90 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 5 | 34.2 | 31.6 | 37 |
| 91 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 5 | 49.6 | 45.9 | 53.4 |
| 92 | By Education | Less than a high school diploma | 5 | 43.9 | 40.4 | 47.5 |
| 93 | By Education | High school diploma or GED | 5 | 36.7 | 35.1 | 38.3 |
| 94 | By Education | Some college/Associate's degree | 5 | 38.3 | 37.2 | 39.4 |
| 95 | By Education | Bachelor's degree or higher | 5 | 29.4 | 28.5 | 30.3 |
| 96 | National Estimate | United States | 6 | 36.1 | 35.2 | 37 |
| 97 | By Age | 18 - 29 years | 6 | 49.3 | 46.8 | 51.8 |
| 98 | By Age | 30 - 39 years | 6 | 41.5 | 39.7 | 43.3 |
| 99 | By Age | 40 - 49 years | 6 | 39 | 37.2 | 40.9 |
| 100 | By Age | 50 - 59 years | 6 | 36 | 34 | 38.1 |
| 101 | By Age | 60 - 69 years | 6 | 27.4 | 25.4 | 29.5 |
| 102 | By Age | 70 - 79 years | 6 | 19 | 16.9 | 21.3 |
| 103 | By Age | 80 years and above | 6 | 18.6 | 13.9 | 24 |
| 104 | By Gender | Male | 6 | 31.9 | 30.4 | 33.5 |
| 105 | By Gender | Female | 6 | 40 | 39 | 41 |
| 106 | By Race/Hispanic ethnicity | Hispanic or Latino | 6 | 40.8 | 38.2 | 43.4 |
| 107 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 6 | 34 | 33 | 35 |
| 108 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 6 | 40 | 37.1 | 42.9 |
| 109 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 6 | 29.2 | 25.5 | 33 |
| 110 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 6 | 48.1 | 43.6 | 52.7 |
| 111 | By Education | Less than a high school diploma | 6 | 44.7 | 41.2 | 48.2 |
| 112 | By Education | High school diploma or GED | 6 | 36.8 | 34.7 | 38.8 |
| 113 | By Education | Some college/Associate's degree | 6 | 39.3 | 37.8 | 40.7 |
| 114 | By Education | Bachelor's degree or higher | 6 | 30.2 | 29.3 | 31.1 |
| 115 | National Estimate | United States | 7 | 36 | 35.2 | 36.8 |
| 116 | By Age | 18 - 29 years | 7 | 48.7 | 45.5 | 51.8 |
| 117 | By Age | 30 - 39 years | 7 | 40.3 | 38.5 | 42.1 |
| 118 | By Age | 40 - 49 years | 7 | 38.4 | 36.5 | 40.2 |
| 119 | By Age | 50 - 59 years | 7 | 35.8 | 33.6 | 38 |
| 120 | By Age | 60 - 69 years | 7 | 27.7 | 26 | 29.4 |
| 121 | By Age | 70 - 79 years | 7 | 21.7 | 19.2 | 24.3 |
| 122 | By Age | 80 years and above | 7 | 17.6 | 13.4 | 22.4 |
| 123 | By Gender | Male | 7 | 32.7 | 31.4 | 34 |
| 124 | By Gender | Female | 7 | 39.1 | 38.1 | 40.1 |
| 125 | By Race/Hispanic ethnicity | Hispanic or Latino | 7 | 39.6 | 36.4 | 43 |
| 126 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 7 | 34.3 | 33.3 | 35.3 |
| 127 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 7 | 40 | 37 | 43 |
| 128 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 7 | 29.9 | 25.7 | 34.4 |
| 129 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 7 | 44.3 | 40 | 48.6 |
| 130 | By Education | Less than a high school diploma | 7 | 43.7 | 39.7 | 47.8 |
| 131 | By Education | High school diploma or GED | 7 | 37.4 | 35.3 | 39.5 |
| 132 | By Education | Some college/Associate's degree | 7 | 39.3 | 37.7 | 40.9 |
| 133 | By Education | Bachelor's degree or higher | 7 | 29.4 | 28.5 | 30.4 |
| 134 | National Estimate | United States | 8 | 36.1 | 35.4 | 36.9 |
| 135 | By Age | 18 - 29 years | 8 | 50 | 47.6 | 52.4 |
| 136 | By Age | 30 - 39 years | 8 | 40.8 | 39.3 | 42.4 |
| 137 | By Age | 40 - 49 years | 8 | 38.5 | 37 | 40 |
| 138 | By Age | 50 - 59 years | 8 | 36.3 | 34.5 | 38.2 |
| 139 | By Age | 60 - 69 years | 8 | 27.4 | 25.7 | 29.1 |
| 140 | By Age | 70 - 79 years | 8 | 19.3 | 17.2 | 21.6 |
| 141 | By Age | 80 years and above | 8 | 14.6 | 10.8 | 19.2 |
| 142 | By Gender | Male | 8 | 32.4 | 31.3 | 33.5 |
| 143 | By Gender | Female | 8 | 39.6 | 38.6 | 40.7 |
| 144 | By Race/Hispanic ethnicity | Hispanic or Latino | 8 | 42.3 | 39.9 | 44.7 |
| 145 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 8 | 34.1 | 33.3 | 34.9 |
| 146 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 8 | 39.1 | 36.8 | 41.5 |
| 147 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 8 | 27.8 | 24.4 | 31.4 |
| 148 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 8 | 45.1 | 41.2 | 48.9 |
| 149 | By Education | Less than a high school diploma | 8 | 43.6 | 39.9 | 47.3 |
| 150 | By Education | High school diploma or GED | 8 | 38.1 | 36.5 | 39.6 |
| 151 | By Education | Some college/Associate's degree | 8 | 39.1 | 37.6 | 40.7 |
| 152 | By Education | Bachelor's degree or higher | 8 | 29.2 | 28.1 | 30.3 |
| 153 | National Estimate | United States | 9 | 37.8 | 37 | 38.5 |
| 154 | By Age | 18 - 29 years | 9 | 49.4 | 47.2 | 51.5 |
| 155 | By Age | 30 - 39 years | 9 | 44.1 | 42.5 | 45.8 |
| 156 | By Age | 40 - 49 years | 9 | 39.7 | 38.1 | 41.3 |
| 157 | By Age | 50 - 59 years | 9 | 37.4 | 35.8 | 39.1 |
| 158 | By Age | 60 - 69 years | 9 | 29.1 | 27.5 | 30.6 |
| 159 | By Age | 70 - 79 years | 9 | 23.4 | 21.2 | 25.8 |
| 160 | By Age | 80 years and above | 9 | 20.5 | 15.8 | 25.9 |
| 161 | By Gender | Male | 9 | 33.8 | 32.5 | 35.1 |
| 162 | By Gender | Female | 9 | 41.5 | 40.7 | 42.4 |
| 163 | By Race/Hispanic ethnicity | Hispanic or Latino | 9 | 42.8 | 40.7 | 45 |
| 164 | By Race/Hispanic ethnicity | Non-Hispanic white, single race | 9 | 35.8 | 34.8 | 36.8 |
| 165 | By Race/Hispanic ethnicity | Non-Hispanic black, single race | 9 | 41.3 | 39.3 | 43.4 |
| 166 | By Race/Hispanic ethnicity | Non-Hispanic Asian, single race | 9 | 28.6 | 25.5 | 31.9 |
| 167 | By Race/Hispanic ethnicity | Non-Hispanic, other races and multiple races | 9 | 49.6 | 45.9 | 53.2 |
| 168 | By Education | Less than a high school diploma | 9 | 46.9 | 42.9 | 51 |
| 169 | By Education | High school diploma or GED | 9 | 38.7 | 37 | 40.5 |
| 170 | By Education | Some college/Associate's degree | 9 | 41.8 | 40.7 | 43 |
| 171 | By Education | Bachelor's degree or higher | 9 | 30.3 | 29.3 | 31.4 |

Shapiro-Wilks for Combined Table (II)

> shapiro.test(or\_age\_table$Value)

Shapiro-Wilk normality test

data: or\_age\_table$Value

W = 0.94232, p-value = 0.005312

> shapiro.test(or\_sex\_table$Value)

Shapiro-Wilk normality test

data: or\_sex\_table$Value

W = 0.87795, p-value = 0.02412

> shapiro.test(or\_race\_table$Value)

Shapiro-Wilk normality test

data: or\_race\_table$Value

W = 0.96836, p-value = 0.2522

> shapiro.test(or\_ed\_table$Value)

Shapiro-Wilk normality test

data: or\_ed\_table$Value

W = 0.94945, p-value = 0.1005

Shapiro-Wilks for Spread Table (III)

> shapiro.test(or\_age\_table2$`18 - 29 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`18 - 29 years`

W = 0.90726, p-value = 0.2972

> shapiro.test(or\_age\_table2$`30 - 39 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`30 - 39 years`

W = 0.92944, p-value = 0.476

> shapiro.test(or\_age\_table2$`40 - 49 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`40 - 49 years`

W = 0.91781, p-value = 0.3744

> shapiro.test(or\_age\_table2$`50 - 59 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`50 - 59 years`

W = 0.87928, p-value = 0.1543

> shapiro.test(or\_age\_table2$`60 - 69 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`60 - 69 years`

W = 0.94206, p-value = 0.6038

> shapiro.test(or\_age\_table2$`70 - 79 years`)

Shapiro-Wilk normality test

data: or\_age\_table2$`70 - 79 years`

W = 0.91191, p-value = 0.3295

> shapiro.test(or\_age\_table2$`80 years and above`)

Shapiro-Wilk normality test

data: or\_age\_table2$`80 years and above`

W = 0.92662, p-value = 0.4498

> shapiro.test(or\_sex\_table2$Female)

Shapiro-Wilk normality test

data: or\_sex\_table2$Female

W = 0.96591, p-value = 0.8576

> shapiro.test(or\_sex\_table2$Male)

Shapiro-Wilk normality test

data: or\_sex\_table2$Male

W = 0.97831, p-value = 0.955

> shapiro.test(or\_race\_table2$`Hispanic or Latino`)

Shapiro-Wilk normality test

data: or\_race\_table2$`Hispanic or Latino`

W = 0.94925, p-value = 0.6818

> shapiro.test(or\_race\_table2$`Non-Hispanic Asian, single race`)

Shapiro-Wilk normality test

data: or\_race\_table2$`Non-Hispanic Asian, single race`

W = 0.90964, p-value = 0.3133

> shapiro.test(or\_race\_table2$`Non-Hispanic black, single race`)

Shapiro-Wilk normality test

data: or\_race\_table2$`Non-Hispanic black, single race`

W = 0.8618, p-value = 0.1003

> shapiro.test(or\_race\_table2$`Non-Hispanic white, single race`)

Shapiro-Wilk normality test

data: or\_race\_table2$`Non-Hispanic white, single race`

W = 0.97321, p-value = 0.9207

> shapiro.test(or\_race\_table2$`Non-Hispanic, other races and multiple races`)

Shapiro-Wilk normality test

data: or\_race\_table2$`Non-Hispanic, other races and multiple races`

W = 0.903, p-value = 0.2699

> shapiro.test(or\_ed\_table2$`Bachelor's degree or higher`)

Shapiro-Wilk normality test

data: or\_ed\_table2$`Bachelor's degree or higher`

W = 0.8947, p-value = 0.2229

> shapiro.test(or\_ed\_table2$`High school diploma or GED`)

Shapiro-Wilk normality test

data: or\_ed\_table2$`High school diploma or GED`

W = 0.95937, p-value = 0.7918

> shapiro.test(or\_ed\_table2$`Less than a high school diploma`)

Shapiro-Wilk normality test

data: or\_ed\_table2$`Less than a high school diploma`

W = 0.97555, p-value = 0.9376

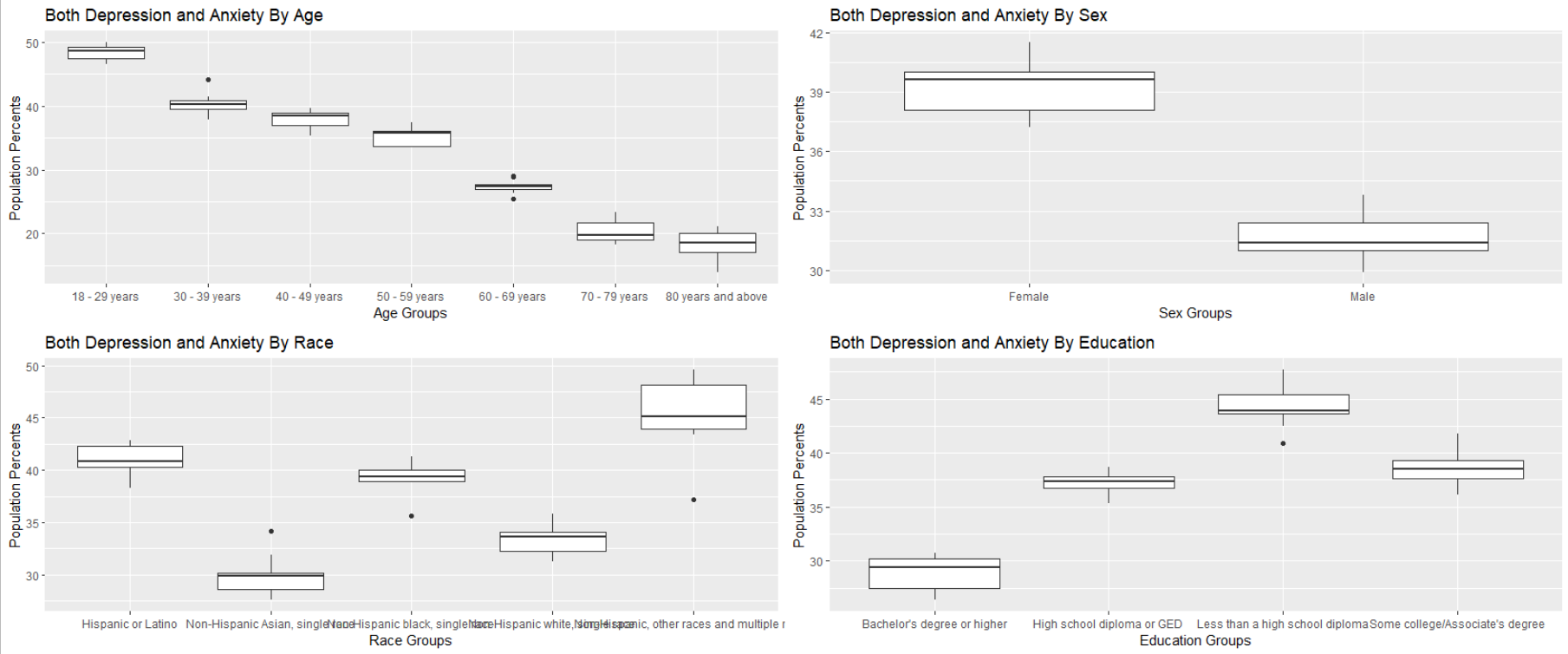
> shapiro.test(or\_ed\_table2$`Some college/Associate's degree`)

Shapiro-Wilk normality test

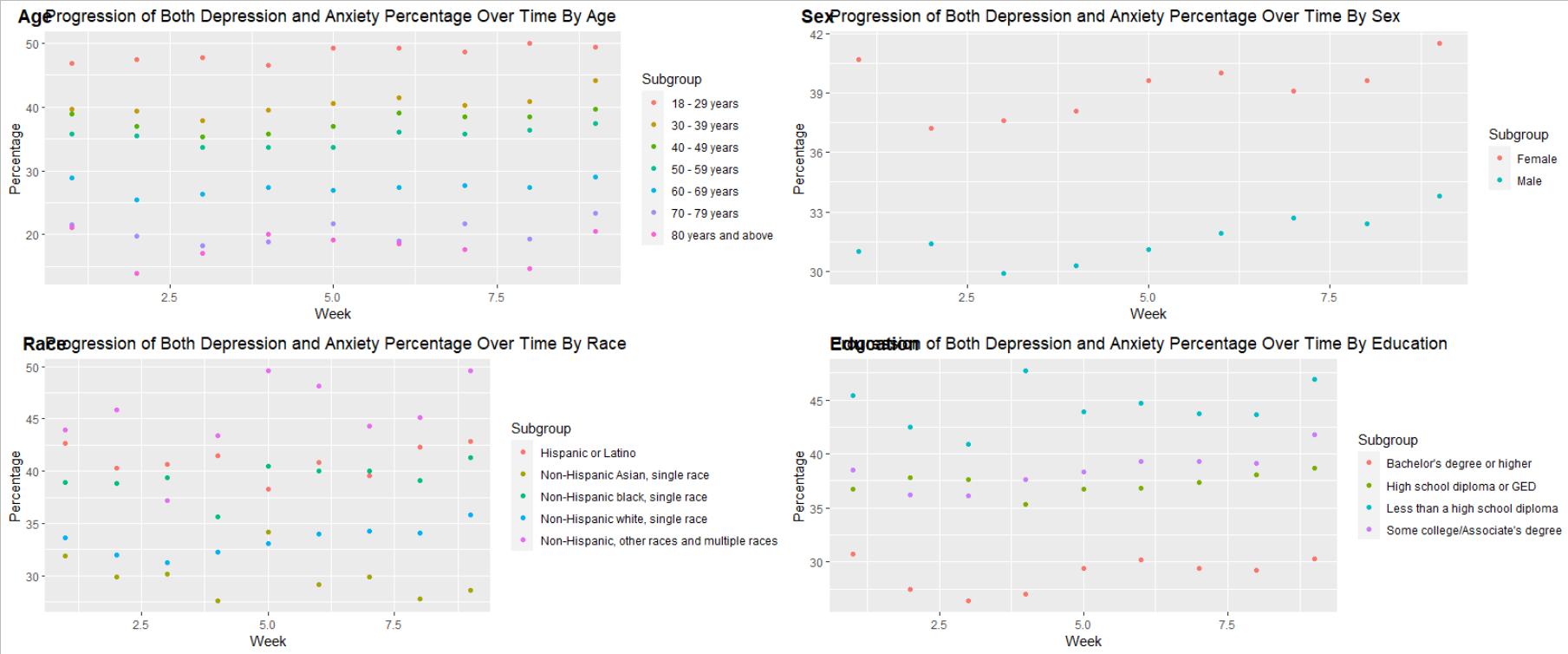
data: or\_ed\_table2$`Some college/Associate's degree`

W = 0.93315, p-value = 0.5119

Boxplots (IV)



Scatterplots (V)



General Linear Models (VI)

> aov(Value ~ Week + Subgroup, data = or\_age\_table)

Call:

aov(formula = Value ~ Week + Subgroup, data = or\_age\_table)

Terms:

Week Subgroup Residuals

Sum of Squares 27.056 6561.323 129.204

Deg. of Freedom 1 6 55

Residual standard error: 1.532698

Estimated effects may be unbalanced

> summary(aov(Value ~ Week + Subgroup, data = or\_age\_table))

Df Sum Sq Mean Sq F value Pr(>F)

Week 1 27 27.1 11.52 0.00128 \*\*

Subgroup 6 6561 1093.6 465.51 < 2e-16 \*\*\*

Residuals 55 129 2.3

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> or\_nat\_model <- glm(Value ~ Week, data = or\_nation\_table)

> summary(or\_nat\_model)

Call:

glm(formula = Value ~ Week, data = or\_nation\_table)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.0322 -0.3906 -0.1789 0.2328 1.5911

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 33.9972 0.6598 51.526 2.72e-10 \*\*\*

Week 0.3117 0.1173 2.658 0.0326 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 0.8248651)

Null deviance: 11.6022 on 8 degrees of freedom

Residual deviance: 5.7741 on 7 degrees of freedom

AIC: 27.546

Number of Fisher Scoring iterations: 2

> or\_age\_model <- glm(Value ~ Week + Subgroup, data = or\_age\_table)

> summary(or\_age\_model)

Call:

glm(formula = Value ~ Week + Subgroup, data = or\_age\_table)

Deviance Residuals:

Min 1Q Median 3Q Max

-4.2392 -0.7918 -0.0187 0.9137 4.0375

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 47.08651 0.63313 74.371 < 2e-16 \*\*\*

Week 0.25381 0.07479 3.394 0.00128 \*\*

Subgroup30 - 39 years -7.96667 0.72252 -11.026 1.5e-15 \*\*\*

Subgroup40 - 49 years -10.65556 0.72252 -14.748 < 2e-16 \*\*\*

Subgroup50 - 59 years -13.04444 0.72252 -18.054 < 2e-16 \*\*\*

Subgroup60 - 69 years -20.93333 0.72252 -28.973 < 2e-16 \*\*\*

Subgroup70 - 79 years -27.96667 0.72252 -38.707 < 2e-16 \*\*\*

Subgroup80 years and above -30.27778 0.72252 -41.906 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 2.349162)

Null deviance: 6717.6 on 62 degrees of freedom

Residual deviance: 129.2 on 55 degrees of freedom

AIC: 242.04

Number of Fisher Scoring iterations: 2

> or\_sex\_model <- glm(Value ~ Week + Subgroup, data = or\_sex\_table)

> summary(or\_sex\_model)

Call:

glm(formula = Value ~ Week + Subgroup, data = or\_sex\_table)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.14917 -0.84021 -0.07278 0.57847 2.65667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 37.73750 0.60213 62.673 < 2e-16 \*\*\*

Week 0.30583 0.09725 3.145 0.00668 \*\*

SubgroupMale -7.65556 0.50221 -15.244 1.55e-10 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 1.134987)

Null deviance: 291.983 on 17 degrees of freedom

Residual deviance: 17.025 on 15 degrees of freedom

AIC: 58.079

Number of Fisher Scoring iterations: 2

> or\_race\_model <- glm(Value ~ Week + Subgroup, data = or\_race\_table)

> summary(or\_race\_model)

Call:

glm(formula = Value ~ Week + Subgroup, data = or\_race\_table)

Deviance Residuals:

Min 1Q Median 3Q Max

-7.6120 -0.8782 0.1431 0.9573 4.3667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 39.9467 0.9788 40.814 < 2e-16 \*\*\*

Week 0.2107 0.1281 1.644 0.108231

SubgroupNon-Hispanic Asian, single race -11.0778 1.0463 -10.587 4.96e-13 \*\*\*

SubgroupNon-Hispanic black, single race -1.7111 1.0463 -1.635 0.110024

SubgroupNon-Hispanic white, single race -7.6111 1.0463 -7.274 9.04e-09 \*\*\*

SubgroupNon-Hispanic, other races and multiple races 4.2333 1.0463 4.046 0.000239 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 4.926646)

Null deviance: 1548.46 on 44 degrees of freedom

Residual deviance: 192.14 on 39 degrees of freedom

AIC: 207.02

Number of Fisher Scoring iterations: 2

> or\_ed\_model <- glm(Value ~ Week + Subgroup, data = or\_ed\_table)

> summary(or\_ed\_model)

Call:

glm(formula = Value ~ Week + Subgroup, data = or\_ed\_table)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.9283 -0.7917 -0.1025 0.6494 3.6025

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 27.55417 0.69558 39.613 < 2e-16 \*\*\*

Week 0.26917 0.09677 2.782 0.00912 \*\*

SubgroupHigh school diploma or GED 8.33333 0.70671 11.792 5.44e-13 \*\*\*

SubgroupLess than a high school diploma 15.46667 0.70671 21.885 < 2e-16 \*\*\*

SubgroupSome college/Associate's degree 9.56667 0.70671 13.537 1.48e-14 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 2.247478)

Null deviance: 1183.708 on 35 degrees of freedom

Residual deviance: 69.672 on 31 degrees of freedom

AIC: 137.93

Number of Fisher Scoring iterations: 2

Anova for General Linear Models (VII)

> anova(or\_age\_model)

Analysis of Deviance Table

Model: gaussian, link: identity

Response: Value

Terms added sequentially (first to last)

Df Deviance Resid. Df Resid. Dev

NULL 62 6717.6

Week 1 27.1 61 6690.5

Subgroup 6 6561.3 55 129.2

> aov(Value ~ Week + Subgroup, data = or\_age\_table)

Call:

aov(formula = Value ~ Week + Subgroup, data = or\_age\_table)

Terms:

Week Subgroup Residuals

Sum of Squares 27.056 6561.323 129.204

Deg. of Freedom 1 6 55

Residual standard error: 1.532698

Estimated effects may be unbalanced

> summary(aov(Value ~ Week + Subgroup, data = or\_age\_table))

Df Sum Sq Mean Sq F value Pr(>F)

Week 1 27 27.1 11.52 0.00128 \*\*

Subgroup 6 6561 1093.6 465.51 < 2e-16 \*\*\*

Residuals 55 129 2.3

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> kruskal.test(anova(or\_age\_model))

Kruskal-Wallis rank sum test

data: anova(or\_age\_model)

Kruskal-Wallis chi-squared = 6.8545, df = 3, p-value = 0.07668

> anova(or\_sex\_model)

Analysis of Deviance Table

Model: gaussian, link: identity

Response: Value

Terms added sequentially (first to last)

Df Deviance Resid. Df Resid. Dev

NULL 17 291.983

Week 1 11.224 16 280.759

Subgroup 1 263.734 15 17.025

> aov(Value ~ Week + Subgroup, data = or\_sex\_table)

Call:

aov(formula = Value ~ Week + Subgroup, data = or\_sex\_table)

Terms:

Week Subgroup Residuals

Sum of Squares 11.22408 263.73389 17.02481

Deg. of Freedom 1 1 15

Residual standard error: 1.065358

Estimated effects may be unbalanced

> summary(aov(Value ~ Week + Subgroup, data = or\_sex\_table))

Df Sum Sq Mean Sq F value Pr(>F)

Week 1 11.22 11.22 9.889 0.00668 \*\*

Subgroup 1 263.73 263.73 232.367 1.55e-10 \*\*\*

Residuals 15 17.02 1.13

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> kruskal.test(anova(or\_sex\_model))

Kruskal-Wallis rank sum test

data: anova(or\_sex\_model)

Kruskal-Wallis chi-squared = 6.8963, df = 3, p-value = 0.07528

> anova(or\_race\_model)

Analysis of Deviance Table

Model: gaussian, link: identity

Response: Value

Terms added sequentially (first to last)

Df Deviance Resid. Df Resid. Dev

NULL 44 1548.46

Week 1 13.31 43 1535.15

Subgroup 4 1343.01 39 192.14

> aov(Value ~ Week + Subgroup, data = or\_race\_table)

Call:

aov(formula = Value ~ Week + Subgroup, data = or\_race\_table)

Terms:

Week Subgroup Residuals

Sum of Squares 13.3141 1343.0067 192.1392

Deg. of Freedom 1 4 39

Residual standard error: 2.219605

Estimated effects may be unbalanced

> summary(aov(Value ~ Week + Subgroup, data = or\_race\_table))

Df Sum Sq Mean Sq F value Pr(>F)

Week 1 13.3 13.3 2.702 0.108

Subgroup 4 1343.0 335.8 68.150 <2e-16 \*\*\*

Residuals 39 192.1 4.9

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> kruskal.test(anova(or\_race\_model))

Kruskal-Wallis rank sum test

data: anova(or\_race\_model)

Kruskal-Wallis chi-squared = 6.8545, df = 3, p-value = 0.07668

> anova(or\_ed\_model)

Analysis of Deviance Table

Model: gaussian, link: identity

Response: Value

Terms added sequentially (first to last)

Df Deviance Resid. Df Resid. Dev

NULL 35 1183.71

Week 1 17.39 34 1166.32

Subgroup 3 1096.65 31 69.67

> aov(Value ~ Week + Subgroup, data = or\_ed\_table)

Call:

aov(formula = Value ~ Week + Subgroup, data = or\_ed\_table)

Terms:

Week Subgroup Residuals

Sum of Squares 17.3882 1096.6475 69.6718

Deg. of Freedom 1 3 31

Residual standard error: 1.499159

Estimated effects may be unbalanced

> summary(aov(Value ~ Week + Subgroup, data = or\_ed\_table))

Df Sum Sq Mean Sq F value Pr(>F)

Week 1 17.4 17.4 7.737 0.00912 \*\*

Subgroup 3 1096.6 365.5 162.649 < 2e-16 \*\*\*

Residuals 31 69.7 2.2

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> kruskal.test(anova(or\_ed\_model))

Kruskal-Wallis rank sum test

data: anova(or\_ed\_model)

Kruskal-Wallis chi-squared = 6.8545, df = 3, p-value = 0.07668

Chi-Square Tests of Homogeneity (VIII)

> ols\_test\_bartlett(or\_age\_table2, '18 - 29 years', '30 - 39 years', '40 - 49 years', '50 - 59 years', '60 - 69 years', '70 - 79 years', '80 years and above')#heteroskedaciticity across combined variables

Bartlett's Test of Homogenity of Variances

------------------------------------------------

Ho: Variances are equal across groups

Ha: Variances are unequal for atleast two groups

Data

-----------------------------------------------------------------------------------------------------------------

Variables: 18 - 29 years 30 - 39 years 40 - 49 years 50 - 59 years 60 - 69 years 70 - 79 years 80 years and above

Test Summary

----------------------------

DF = 6

Chi2 = 7.307579

Prob > Chi2 = 0.2933366

> ols\_test\_bartlett(or\_sex\_table2, 'Male', 'Female')

Bartlett's Test of Homogenity of Variances

------------------------------------------------

Ho: Variances are equal across groups

Ha: Variances are unequal for atleast two groups

Data

----------------------

Variables: Male Female

Test Summary

----------------------------

DF = 1

Chi2 = 0.1708392

Prob > Chi2 = 0.679367

> ols\_test\_bartlett(or\_race\_table2, 'Hispanic or Latino', 'Non-Hispanic white, single race', 'Non-Hispanic black, single race', 'Non-Hispanic Asian, single race', 'Non-Hispanic, other races and multiple races')

Bartlett's Test of Homogenity of Variances

------------------------------------------------

Ho: Variances are equal across groups

Ha: Variances are unequal for atleast two groups

Data

--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Variables: Hispanic or Latino Non-Hispanic white, single race Non-Hispanic black, single race Non-Hispanic Asian, single race Non-Hispanic, other races and multiple races

Test Summary

-----------------------------

DF = 4

Chi2 = 12.54538

Prob > Chi2 = 0.01372464

> ols\_test\_bartlett(or\_ed\_table2, 'Less than a high school diploma', 'High school diploma or GED', 'Some college/Associate\'s degree', 'Bachelor\'s degree or higher')

Bartlett's Test of Homogenity of Variances

------------------------------------------------

Ho: Variances are equal across groups

Ha: Variances are unequal for atleast two groups

Data

---------------------------------------------------------------------------------------------------------------------------------

Variables: Less than a high school diploma High school diploma or GED Some college/Associate's degree Bachelor's degree or higher

Test Summary

----------------------------

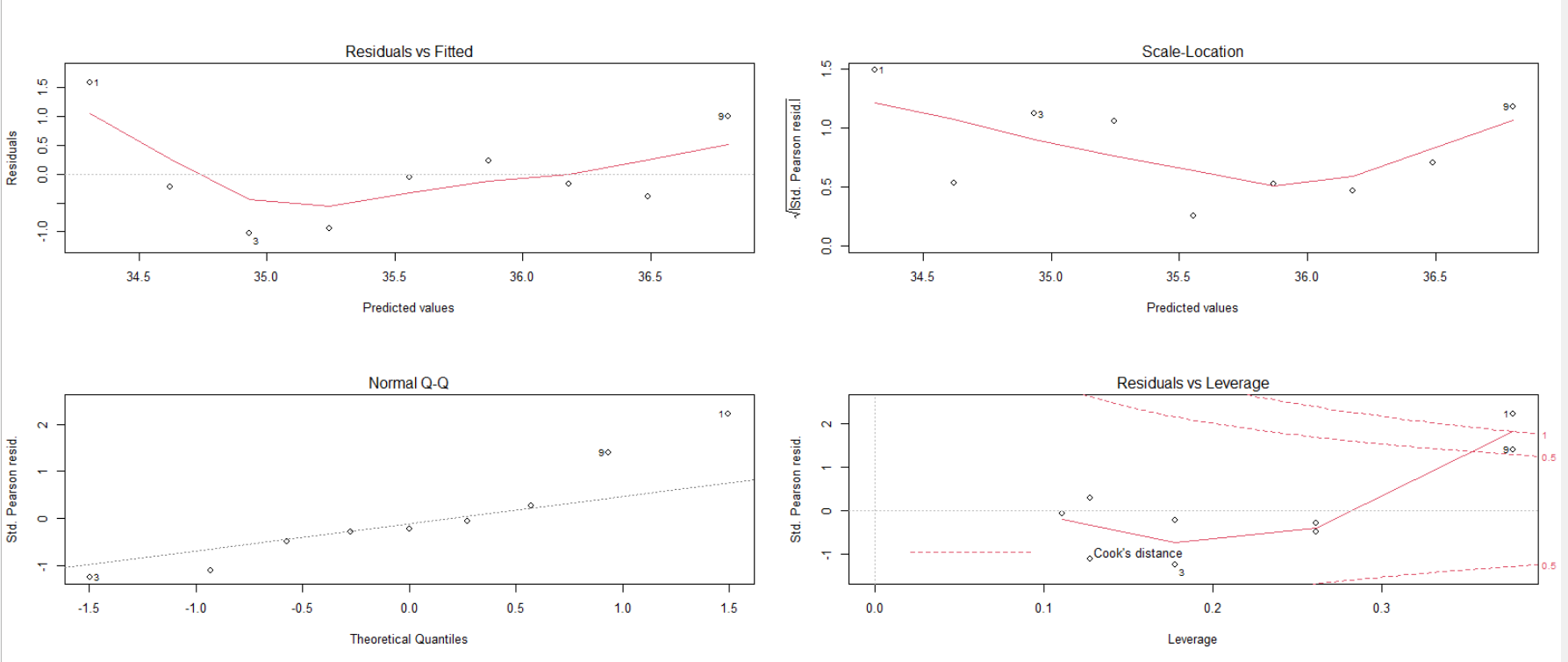
DF = 3

Chi2 = 4.052724

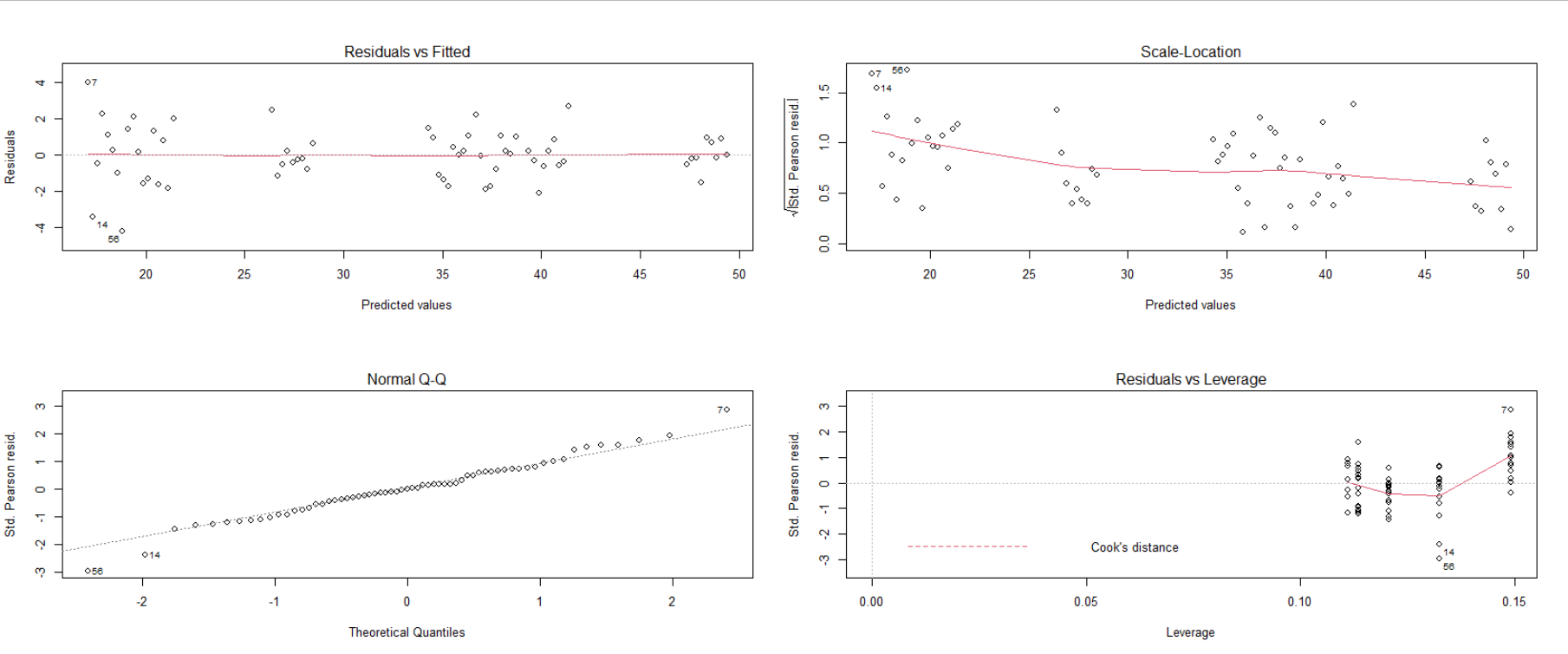
Prob > Chi2 = 0.2558269

Model Residual Plots (IX)

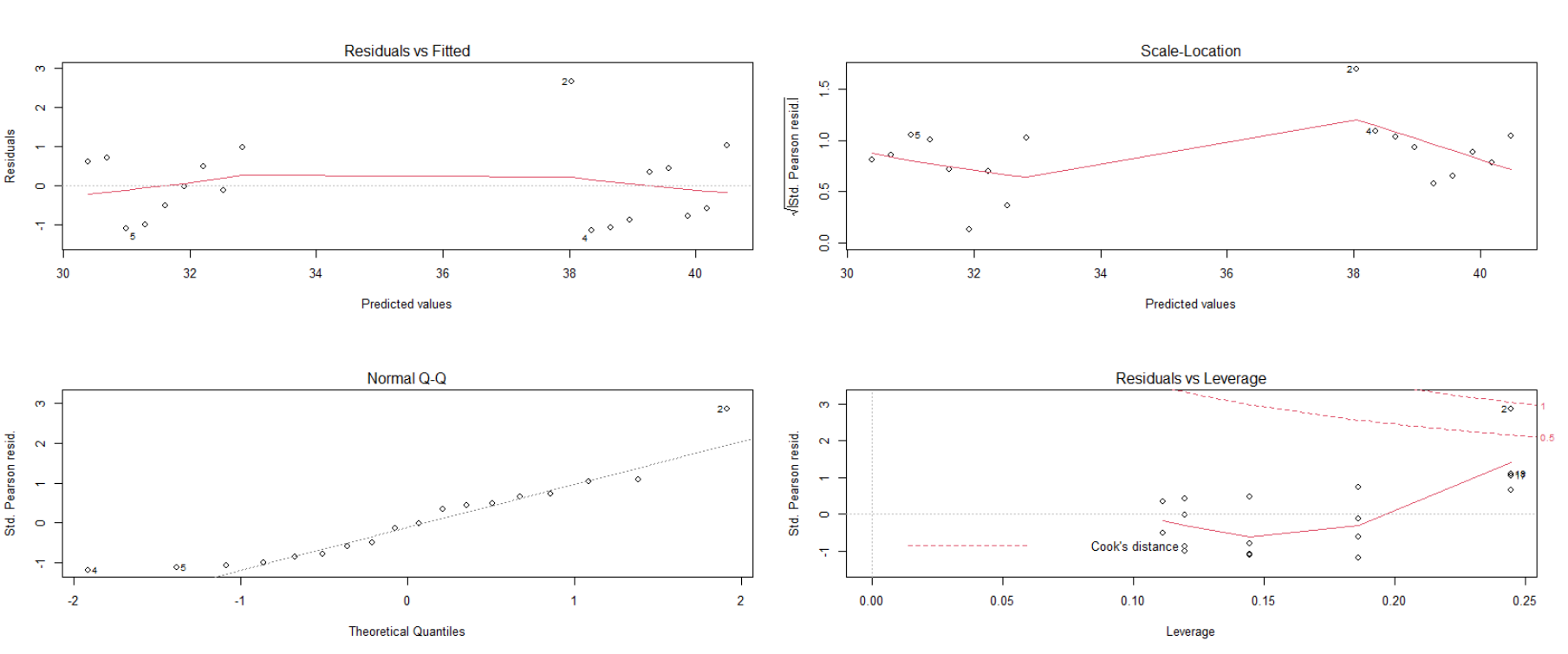
National



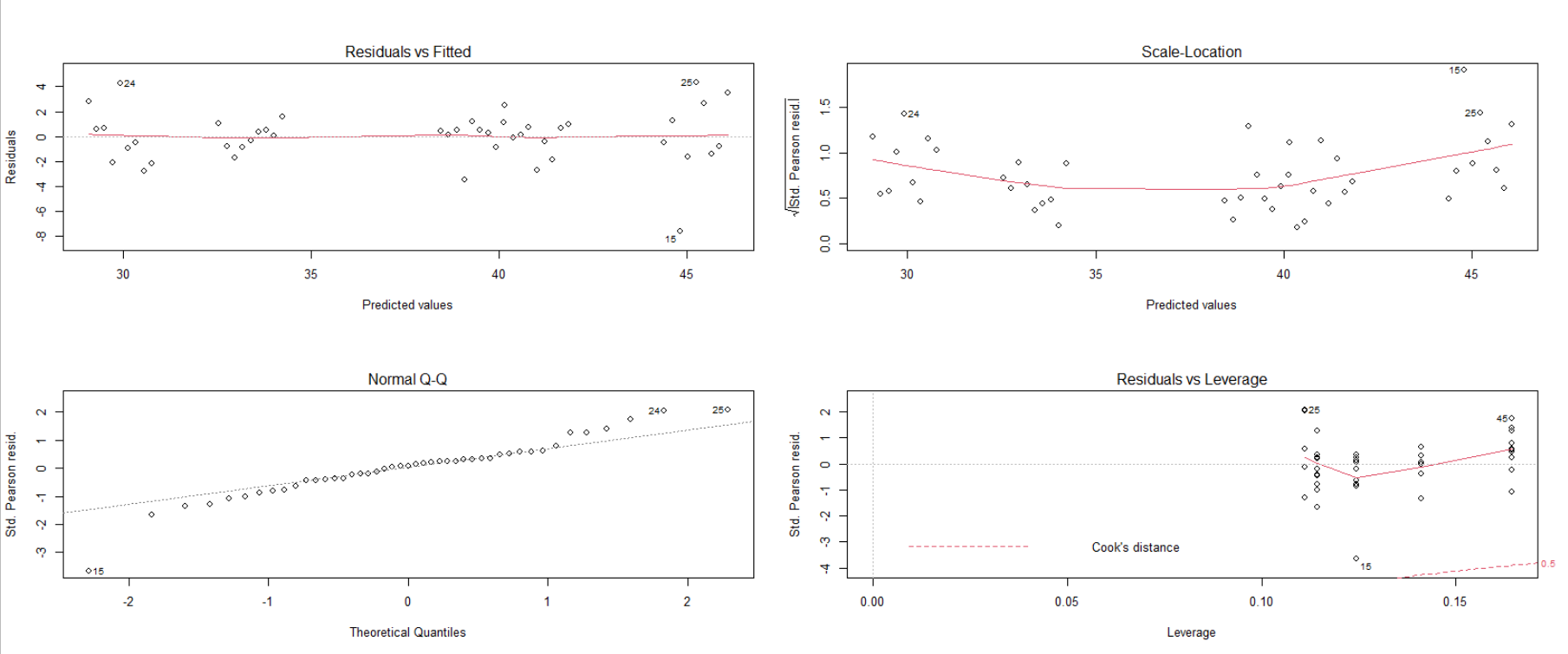
Age



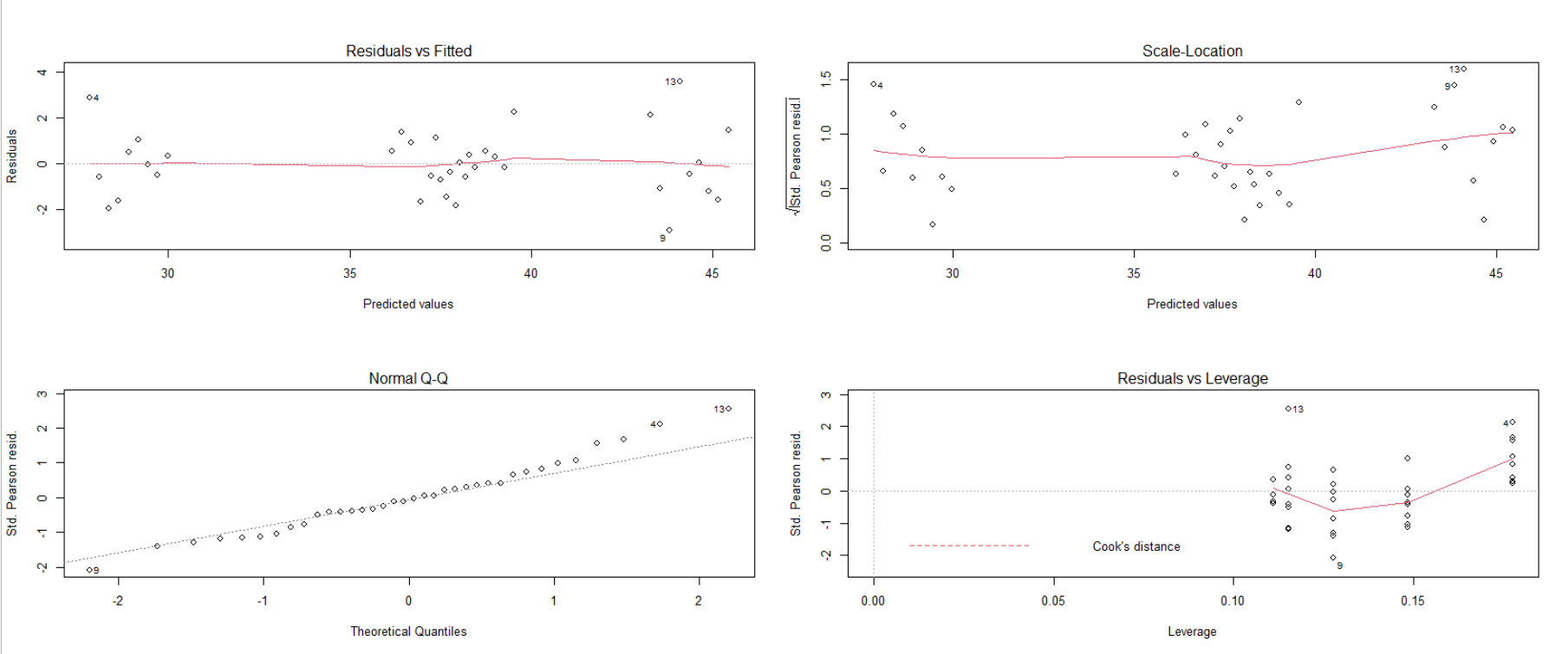
Sex



Race



Education



Hotelling’s T2 (X)

> T2.test(or\_age\_table2)# Hotelling's T2 test

One-sample Hotelling test

data: or\_age\_table2

T2 = 53600.2, F = 1914.3, df1 = 7, df2 = 2, p-value = 0.0005222

alternative hypothesis: true mean vector is not equal to (0, 0, 0, 0, 0, 0, 0)'

sample estimates:

18 - 29 years 30 - 39 years 40 - 49 years 50 - 59 years 60 - 69 years 70 - 79 years 80 years and above

mean x-vector 48.35556 40.38889 37.7 35.31111 27.42222 20.38889 18.07778

> T2.test(or\_sex\_table2)

One-sample Hotelling test

data: or\_sex\_table2

T2 = 7846.3, F = 3432.8, df1 = 2, df2 = 7, p-value = 3.372e-11

alternative hypothesis: true mean vector is not equal to (0, 0)'

sample estimates:

Female Male

mean x-vector 39.26667 31.61111

> T2.test(or\_race\_table2)

One-sample Hotelling test

data: or\_race\_table2

T2 = 21938.2, F = 2193.8, df1 = 5, df2 = 4, p-value = 5.811e-07

alternative hypothesis: true mean vector is not equal to (0, 0, 0, 0, 0)'

sample estimates:

Hispanic or Latino Non-Hispanic Asian, single race Non-Hispanic black, single race Non-Hispanic white, single race

mean x-vector 41 29.92222 39.28889 33.38889

Non-Hispanic, other races and multiple races

mean x-vector 45.23333

> T2.test(or\_ed\_table2)

One-sample Hotelling test

data: or\_ed\_table2

T2 = 46599.9, F = 7281.2, df1 = 4, df2 = 5, p-value = 1.351e-09

alternative hypothesis: true mean vector is not equal to (0, 0, 0, 0)'

sample estimates:

Bachelor's degree or higher High school diploma or GED Less than a high school diploma Some college/Associate's degree

mean x-vector 28.9 37.23333 44.36667 38.46667

Additional Model Information (XI)

> or\_nat\_model <- lm(Value ~ Week, data = or\_nation\_table)

> summary(or\_nat\_model)

Call:

lm(formula = Value ~ Week, data = or\_nation\_table)

Residuals:

Min 1Q Median 3Q Max

-1.0322 -0.3906 -0.1789 0.2328 1.5911

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 33.9972 0.6598 51.526 2.72e-10 \*\*\*

Week 0.3117 0.1173 2.658 0.0326 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9082 on 7 degrees of freedom

Multiple R-squared: 0.5023, Adjusted R-squared: 0.4312

F-statistic: 7.066 on 1 and 7 DF, p-value: 0.03256

> or\_age\_model <- lm(Value ~ Week + Subgroup, data = or\_age\_table)

> summary(or\_age\_model)

Call:

lm(formula = Value ~ Week + Subgroup, data = or\_age\_table)

Residuals:

Min 1Q Median 3Q Max

-4.2392 -0.7918 -0.0187 0.9137 4.0375

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 47.08651 0.63313 74.371 < 2e-16 \*\*\*

Week 0.25381 0.07479 3.394 0.00128 \*\*

Subgroup30 - 39 years -7.96667 0.72252 -11.026 1.5e-15 \*\*\*

Subgroup40 - 49 years -10.65556 0.72252 -14.748 < 2e-16 \*\*\*

Subgroup50 - 59 years -13.04444 0.72252 -18.054 < 2e-16 \*\*\*

Subgroup60 - 69 years -20.93333 0.72252 -28.973 < 2e-16 \*\*\*

Subgroup70 - 79 years -27.96667 0.72252 -38.707 < 2e-16 \*\*\*

Subgroup80 years and above -30.27778 0.72252 -41.906 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.533 on 55 degrees of freedom

Multiple R-squared: 0.9808, Adjusted R-squared: 0.9783

F-statistic: 400.7 on 7 and 55 DF, p-value: < 2.2e-16

> or\_sex\_model <- lm(Value ~ Week + Subgroup, data = or\_sex\_table)

> summary(or\_sex\_model)

Call:

lm(formula = Value ~ Week + Subgroup, data = or\_sex\_table)

Residuals:

Min 1Q Median 3Q Max

-1.14917 -0.84021 -0.07278 0.57847 2.65667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 37.73750 0.60213 62.673 < 2e-16 \*\*\*

Week 0.30583 0.09725 3.145 0.00668 \*\*

SubgroupMale -7.65556 0.50221 -15.244 1.55e-10 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.065 on 15 degrees of freedom

Multiple R-squared: 0.9417, Adjusted R-squared: 0.9339

F-statistic: 121.1 on 2 and 15 DF, p-value: 5.533e-10

> or\_race\_model <- lm(Value ~ Week + Subgroup, data = or\_race\_table)

> summary(or\_race\_model)

Call:

lm(formula = Value ~ Week + Subgroup, data = or\_race\_table)

Residuals:

Min 1Q Median 3Q Max

-7.6120 -0.8782 0.1431 0.9573 4.3667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 39.9467 0.9788 40.814 < 2e-16 \*\*\*

Week 0.2107 0.1281 1.644 0.108231

SubgroupNon-Hispanic Asian, single race -11.0778 1.0463 -10.587 4.96e-13 \*\*\*

SubgroupNon-Hispanic black, single race -1.7111 1.0463 -1.635 0.110024

SubgroupNon-Hispanic white, single race -7.6111 1.0463 -7.274 9.04e-09 \*\*\*

SubgroupNon-Hispanic, other races and multiple races 4.2333 1.0463 4.046 0.000239 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.22 on 39 degrees of freedom

Multiple R-squared: 0.8759, Adjusted R-squared: 0.86

F-statistic: 55.06 on 5 and 39 DF, p-value: < 2.2e-16

> or\_ed\_model <- lm(Value ~ Week + Subgroup, data = or\_ed\_table)

> summary(or\_ed\_model)

Call:

lm(formula = Value ~ Week + Subgroup, data = or\_ed\_table)

Residuals:

Min 1Q Median 3Q Max

-2.9283 -0.7917 -0.1025 0.6494 3.6025

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 27.55417 0.69558 39.613 < 2e-16 \*\*\*

Week 0.26917 0.09677 2.782 0.00912 \*\*

SubgroupHigh school diploma or GED 8.33333 0.70671 11.792 5.44e-13 \*\*\*

SubgroupLess than a high school diploma 15.46667 0.70671 21.885 < 2e-16 \*\*\*

SubgroupSome college/Associate's degree 9.56667 0.70671 13.537 1.48e-14 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.499 on 31 degrees of freedom

Multiple R-squared: 0.9411, Adjusted R-squared: 0.9335

F-statistic: 123.9 on 4 and 31 DF, p-value: < 2.2e-16

Confidence Intervals for Models (XII)

> confint(or\_nat\_model, level = 0.95)

2.5 % 97.5 %

(Intercept) 32.43702626 35.5574182

Week 0.03441257 0.5889208

> confint(or\_age\_model, level = 0.95)

2.5 % 97.5 %

(Intercept) 45.8176943 48.355322

Week 0.1039311 0.403688

Subgroup30 - 39 years -9.4146303 -6.518703

Subgroup40 - 49 years -12.1035191 -9.207592

Subgroup50 - 59 years -14.4924080 -11.596481

Subgroup60 - 69 years -22.3812969 -19.485370

Subgroup70 - 79 years -29.4146303 -26.518703

Subgroup80 years and above -31.7257414 -28.829814

> confint(or\_sex\_model, level = 0.95)

2.5 % 97.5 %

(Intercept) 36.4540818 39.0209182

Week 0.0985426 0.5131241

SubgroupMale -8.7260003 -6.5851108

> confint(or\_race\_model, level = 0.95)

2.5 % 97.5 %

(Intercept) 37.96695027 41.9263831

Week -0.04853905 0.4698724

SubgroupNon-Hispanic Asian, single race -13.19418363 -8.9613719

SubgroupNon-Hispanic black, single race -3.82751696 0.4052947

SubgroupNon-Hispanic white, single race -9.72751696 -5.4947053

SubgroupNon-Hispanic, other races and multiple races 2.11692748 6.3497392

> confint(or\_ed\_model, level = 0.95)

2.5 % 97.5 %

(Intercept) 26.13552097 28.972812

Week 0.07180231 0.466531

SubgroupHigh school diploma or GED 6.89198786 9.774679

SubgroupLess than a high school diploma 14.02532119 16.908012

SubgroupSome college/Associate's degree 8.12532119 11.008012

Principal Components Analysis (XIII)

> or\_age\_cor2

pca1 pca2 18 - 29 years 30 - 39 years 40 - 49 years 50 - 59 years 60 - 69 years 70 - 79 years 80 years and above

pca1 1.000 -0.094 -0.211 -0.658 -0.858 -0.815 -0.199 -0.049 0.275

pca2 -0.094 1.000 -0.375 0.237 -0.080 -0.197 0.519 0.695 0.589

18 - 29 years -0.211 -0.375 1.000 0.658 0.441 0.518 -0.766 -0.297 -0.907

30 - 39 years -0.658 0.237 0.658 1.000 0.766 0.772 -0.160 0.245 -0.539

40 - 49 years -0.858 -0.080 0.441 0.766 1.000 0.975 0.059 0.220 -0.522

50 - 59 years -0.815 -0.197 0.518 0.772 0.975 1.000 -0.093 0.128 -0.646

60 - 69 years -0.199 0.519 -0.766 -0.160 0.059 -0.093 1.000 0.479 0.805

70 - 79 years -0.049 0.695 -0.297 0.245 0.220 0.128 0.479 1.000 0.321

80 years and above 0.275 0.589 -0.907 -0.539 -0.522 -0.646 0.805 0.321 1.000

> or\_sex\_cor2

pca1 Female Male

pca1 1 1 -1

Female 1 1 -1

Male -1 -1 1

> or\_race\_cor2

pca1 pca2 Hispanic or Latino Non-Hispanic Asian, single race Non-Hispanic black, single race Non-Hispanic white, single race Non-Hispanic, other races and multiple races

pca1 1.000 0.723 0.027 0.019 0.481 0.247 -0.222

pca2 0.723 1.000 0.533 -0.402 -0.246 0.143 -0.552

Hispanic or Latino 0.027 0.533 1.000 -0.966 -0.622 0.409 -0.344

Non-Hispanic Asian, single race 0.019 -0.402 -0.966 1.000 0.491 -0.606 0.097

Non-Hispanic black, single race 0.481 -0.246 -0.622 0.491 1.000 0.288 0.515

Non-Hispanic white, single race 0.247 0.143 0.409 -0.606 0.288 1.000 0.611

Non-Hispanic, other races and multiple races -0.222 -0.552 -0.344 0.097 0.515 0.611 1.000

> or\_ed\_cor2

pca1 pca2 Bachelor's degree or higher High school diploma or GED Less than a high school diploma Some college/Associate's degree

pca1 1.000 -0.553 0.892 0.160 -0.133 0.569

pca2 -0.553 1.000 -0.677 -0.066 -0.181 -0.898

Bachelor's degree or higher 0.892 -0.677 1.000 -0.252 0.320 0.832

High school diploma or GED 0.160 -0.066 -0.252 1.000 -0.960 -0.353

Less than a high school diploma -0.133 -0.181 0.320 -0.960 1.000 0.545

Some college/Associate's degree 0.569 -0.898 0.832 -0.353 0.545 1.000

Cluster Analysis (XIV)

> aggregate(or\_age\_table3, by = list(fit1$cluster), FUN = mean)

Group.1 18 - 29 years 30 - 39 years 40 - 49 years 50 - 59 years 60 - 69 years 70 - 79 years 80 years and above

1 1 48.85000 40.475 38.20000 35.90000 27.00000 19.95 16.175

2 2 48.10000 41.850 39.30000 36.60000 29.00000 22.45 20.800

3 3 47.86667 39.300 35.96667 33.66667 26.93333 19.60 18.800

> aggregate(or\_age\_table3, by = list(fit2$cluster), FUN = mean)

Group.1 18 - 29 years 30 - 39 years 40 - 49 years 50 - 59 years 60 - 69 years 70 - 79 years 80 years and above

1 1 49.40 44.10 39.7 37.40000 29.10000 23.40 20.50000

2 2 46.80 39.60 38.9 35.80000 28.90000 21.50 21.10000

3 3 47.15 38.65 35.5 33.70000 26.90000 18.55 18.60000

4 4 49.10 40.80 38.1 35.13333 27.36667 20.80 18.46667

5 5 48.70 40.05 37.7 35.90000 26.45000 19.55 14.25000

> aggregate(or\_sex\_table3, by = list(fit1$cluster), FUN = mean)

Group.1 Female Male

1 1 37.20000 31.40

2 2 37.85000 30.10

3 3 40.08333 32.15

> aggregate(or\_sex\_table3, by = list(fit2$cluster), FUN = mean)

Group.1 Female Male

1 1 39.35000 32.55000

2 2 39.80000 31.50000

3 3 41.50000 33.80000

4 4 40.70000 31.00000

5 5 37.63333 30.53333

> aggregate(or\_race\_table3, by = list(fit1$cluster), FUN = mean)

Group.1 Hispanic or Latino Non-Hispanic Asian, single race Non-Hispanic black, single race Non-Hispanic white, single race Non-Hispanic, other races and multiple races

1 1 41.28000 29.42000 38.48 33.26 44.52

2 2 40.63333 30.66667 40.60 34.30 49.10

3 3 40.70000 30.20000 39.40 31.30 37.20

> aggregate(or\_race\_table3, by = list(fit2$cluster), FUN = mean)

Group.1 Hispanic or Latino Non-Hispanic Asian, single race Non-Hispanic black, single race Non-Hispanic white, single race Non-Hispanic, other races and multiple races

1 1 41.800 28.900 40.65 34.9 48.85

2 2 38.300 34.200 40.50 33.1 49.60

3 3 41.225 29.875 39.20 33.5 44.80

4 4 40.700 30.200 39.40 31.3 37.20

5 5 41.500 27.600 35.60 32.3 43.40

> aggregate(or\_ed\_table3, by = list(fit1$cluster), FUN = mean)

Group.1 Bachelor's degree or higher High school diploma or GED Less than a high school diploma Some college/Associate's degree

1 1 27.00000 35.3 47.7 37.60000

2 2 26.95000 37.7 41.7 36.15000

3 3 29.86667 37.4 44.7 39.38333

> aggregate(or\_ed\_table3, by = list(fit2$cluster), FUN = mean)

Group.1 Bachelor's degree or higher High school diploma or GED Less than a high school diploma Some college/Associate's degree

1 1 27.50 37.80 42.50 36.2

2 2 29.78 37.14 44.26 38.9

3 3 27.00 35.30 47.70 37.6

4 4 30.30 38.70 46.90 41.8

5 5 26.40 37.60 40.90 36.1

1. **Discussion**

Univariate analysis can be very useful when showing the basic elements of a single variable, but without expanding the analysis to more variables we cannot see the relationship that exist in bivariate and multivariate data. This style of analysis is used to show how we might see the effects of subgroups of the whole of the data and even might give an indication that certain groups are more heavily affected than others. This analysis is only possible through multivariate modeling or the use of bivariate analysis. Each of these two methods (multivariate and bivariate) came to the same results and gave me the indication of what I wanted.

(Once again I had a similar paragraph here but it got lost in the sauce so this one was posted) This data was not meant to be analyzed in this way or presented as such, and the decision to use this data for this assignment was a poor one. The dataset gives a percentage of the sample with the symptoms of mental illness, subdividing the data in several different fashions without showing how any of the data overlaps. It doesn’t indicate “College Degree and Hispanic” or “Over 80 and Female”, it merely indicate the percentages over time for the one general statistic. This tells a lot about how the country is progressing in different aspects and how different groups are being affected overall, but without the original dataset (which I could not find) this data is very limited in what it could tell us. The time that I was able to work on both parts of the project was severely limited by outside forces and prevented a more in depth analysis. If this project were to continue a new search for the complete dataset would have to be conducted. A more in depth model would have to be made and who truly has it the worst in the United States at the current moment would have to be investigated. As for the broader impact this project has show that there are significant cracks in the mental health of the United States and possibly indicates that there are groups who do not have access to the resources they need to keep sane in this crisis.

1. **References**

# References

Bailey, R. K., Mokonogho, J., & Kumar, A. (2019, February 22). Racial and ethnic differences in depression: current perspectives. *Neuropsychiatric Disease and Treatment*, 603-609. doi:10.2147/NDT.S128584

Davis, M. T., Holmes, S. E., Pietrzak, R. H., & Esterlis, I. (2017, June 22). Neurobiology of Chronic Stress-Related Psychiatric Disorders: Evidence from Molecular Imaging Studies. *Chronic Stress (Thousand Oaks)*, 1-21. doi:10.1177/2470547017710916

Gan, Z., Li, Y., Xie, D., Shao, C., Yang, F., Shen, Y., . . . Zhang, J. (2012, February). The impact of educational status on the clinical features of major depressive disorder among Chinese women. *J Affect Disord*, 988-992. doi:10.1016/j.jad.2011.06.046

Kang, H.-J., Park, Y., Yoo, K.-H., Kim, K.-T., Kim, E.-S., Kim, J.-W., . . . Kim, J.-M. (2020, June 18). Sex differences in the genetic architecture of depression. *Sci Rep*, 1-12. doi:10.1038/s41598-020-66672-9

Lee, L. O., Gatz, M., Pedersen, N. L., & Prescott, C. A. (2016, January 11). Anxiety Trajectories in the Second Half of Life: Genetic and Environmental Contributions over Age. *Psychol Aging*, 1-27. doi:10.1037/pag0000063

MacQueen, G. M., Frey, B. N., Ismail, Z., Jaworska, N., Steiner, M., VanLieshout, R. J., . . . Ravindran, A. V. (2016, August 2). Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 6. Special Populations: Youth, Women, and the Elderly. *Can J Psychiatry*, 588-603. doi:10.1177/0706743716659276

MIYAWAKI, C. E. (2014, August 27). *Association of social isolation and health across different racial and ethnic groups of older Americans*. Retrieved from US National Library of Medicine: National Institutes of Health: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4610249/

National Center for Health Statistics. (2020, July 16). *Mental Health: Household Pulse Survey*. Retrieved from Centers for Disease Control and Prevention: https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm

National HIV Curriculum. (2020). *Generalized Anxiety Disorder 2-item (GAD-2)*. Retrieved from National HIV Curriculum: https://www.hiv.uw.edu/page/mental-health-screening/gad-2

National HIV Curriculum. (2020). *Patient Health Questionnaire-2 (PHQ-2)*. Retrieved from National HIV Curriculum: https://www.hiv.uw.edu/page/mental-health-screening/phq-2#:~:text=The%20PHQ%2D2%20inquires%20about,%E2%80%9Cfirst%2Dstep%E2%80%9D%20approach.

Oh, M., Kim, J.-W., Yoon, N.-H., Lee, S. A., Lee, S. L., & Kang, W. S. (2019, December). Differences in Personality, Defense Styles, and Coping Strategies in Individuals with Depressive Disorder According to Age Groups Across the Lifespan. *Psychiatry Investig*, 911-918. doi:10.30773/pi.2019.0160

Schlax, J., Jünger, C., Beutel, M. E., Münzel, T., Pfeiffer, N., Wild, P., . . . Michal, M. (2019, April 24). Income and education predict elevated depressive symptoms in the general population: results from the Gutenberg health study. *BMC Public Health*, 1-10. doi:10.1186/s12889-019-6730-4