Audrey Lin
COMS 180
Final Project

The Rise of Depression in Youth

1 Introduction

1.1 Topic Area

In this paper, I seek to investigate the difference in the rate of depression in youth in recent years compared to in the past. This is motivated by the presumption of the rise of depression in youth, as observed in the common refrain that “kids these days are so depressed,” or more specifically, that “millennials and Gen Z are so depressed.”

I believe that youth these days are indeed more depressed than they were before. There are so many more pressures in the modern world, and society’s constant pursuit for self-optimization only becomes more extreme with time. Standards rise and become more and more difficult to reach. Journalist Anne Helen Petersen explains that millennials are not rewarded for their efficiency, but rather further exploited (Petersen 2019), and University of Alberta sociologist Lisa Storhschein says that millennials are the first generation in the last 50 years for which the expectation to do better than the preceding generation no longer holds true (Mcmaster 2020).

Additionally, there is the grim future of politics, social injustice, and environmental destruction that youth contend with inheriting, and on top of all that, most of the pressures youth face are only exacerbated by the incessant nature of modern technology. Negative effects of technology such as smartphones include but are not limited to sleep deprivation, cyberbullying, fear of missing out, and information anxiety, especially as related to the pervasiveness of tragic world news (The Annie E Casey Foundation 2021, Twenge 2017). While it has not been proven with absolute certainty that negative effects of technology in modern life are direct contributing factors to depression, San Diego State University psychology professor Jean M Twenge asserts that the correlation between depression and smartphone usage is strong enough for concern (Advisory Board 2019, Twenge 2017).

1.2 Project Details

1.2.1 Hypotheses

- H_0: The rate of depression in youth in recent years is the same rate as in the past.
- H_A: The rate of depression in youth in recent years is greater than the rate in the past.

1.2.2 Sample

There are various degrees of depression and its classification can be rather subjective. Rather than restrict myself to a single, self-defined measure of depression, I browsed available data for possible measures of depression, looking for specificity and consistency in definitions from other sources.

I decided to manually extract and aggregate data from seven annual reports (2004-2008, 2009-2010, 2010-2011, 2012-2013, 2014-2015, 2016-2017, 2018-2019) from the National Survey on Drug Use and Health (NSDUH), which were all sponsored and published by the Substance Abuse and Mental Health Services Administration (SAMHSA). The extensive documentation and consistency since the inception of NSDUH in 1971 provided a robust set of data to work with. Most relevantly, NSDUH had a specific measure of depression that I could adopt for my own analysis, and distinguished the age groups of their participants.

The measure of depression NSDUH used was incidences of Major Depressive Episode (MDE), as defined by the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-IV). Specifically, NSDUH considered an incidence of MDE as a period of depression lasting two weeks or longer while exhibiting some of a specific criteria of symptoms (SAMHSA 2020).

For my sample, I extracted counts and percentages of incidences of MDE in the past year among youth aged 12 to 17 in the US from 2004 through 2019 as shown in Table 1, which covers the entire time-frame from which MDE-related sections were added to NSDUH's questionnaire. The only demographic I distinguished in my sample was age (youth), but other demographic information available from NSDUH were gender, race, geographic region, county type, poverty level, health insurance, and overall health, both as counts and as percentages.

From this raw data (Table 1), I derived the sample sizes each year, as summarized in Table 2 of observed values in Section 2.3.2.3 Test Statistic X^2. In 2004, 24,722,222 youth were sampled; in 2005, 24,897,727 youth were sampled; in 2006, 24,924,051 youth were sampled; in 2007, 24,585,366 youth were sampled; in 2008, 24,265,060 youth were sampled; in 2009, 24,074,074 youth were sampled; in 2010, 23,887,500 youth were sampled; in 2011, 24,524,390 youth were sampled; in 2012, 24,318,681 youth were sampled; in 2013, 24,177,570 youth were sampled; in 2014, 24,131,579 youth were sampled; in 2015, 24,248,000 youth were sampled; in 2016, 24,132,813 youth were sampled; in 2017, 24,165,414 youth were sampled; in 2018, 24,180,556 youth were sampled; in 2019, 24,095,541 youth were sampled. In total, across these 16 years, 389,330,544 youth were sampled.

Incidence of Major Depressive Episode (MDE) in past year among youth aged 12 to 17, 2004-2019		
Year	**Count**	**Percentage**
2004	2225000	9
2005	2191000	8.8
2006	1969000	7.9
2007	2016000	8.2
2008	2014000	8.3
2009	1950000	8.1
2010	1911000	8
2011	2011000	8.2
2012	2213000	9.1
2013	2587000	10.7
2014	2751000	11.4
2015	3031000	12.5
2016	3089000	12.8
2017	3214000	13.3
2018	3482000	14.4
2019	3783000	15.7

Table 1: Raw sample data - Incidence of Major Depressive Episode (MDE) in past year among youth aged 12 to 17, 2004-2019.

1.2.3 Population

The annual reports from which I derived my data did not go into much detail about the target population, except that it includes noninstitutionalized US civilians aged 12 years or older, which NSDUH expects accounts for at least 97% of the US population (SAMHSA 2020). That is, the target population is almost the entire US population. However, of NSDUH's target population of noninstitutionalized US civilians aged 12 years or older, my target population is only those aged 12 to 17 years. That is, my target population is almost the entire youth population in the US.

2 Statistical Analysis

2.1 Variables

- Independent variable: Year, i.e. 2004-2019.
- Dependent variable: One or more incidences of MDE in the past year, i.e. Yes/No.

2.2 Descriptive Statistics (Sample)

In 2004, 2,225,000 youth had at least one incidence of MDE and 22,497,222 did not; in 2005, 2,191,000 youth had at least one incidence of MDE and 22,706,727 did not; in 2006, 1,969,000 youth had at least one incidence of MDE and 22,955,051 did not; in 2007, 2,016,000 youth had at least one incidence of MDE and 22,569,366 did not; in 2008, 2,014,000 youth had at least one incidence of MDE and 22,251,060 did not; in 2009, 1,950,000 youth had at least one incidence of MDE and 22,124,074 did not; in 2010, 1,911,000 youth had at least one incidence of MDE and 21,976,500 did not; in 2011, 2,011,000 youth had at least one incidence of MDE and 22,513,390 did not; in 2012, 2,213,000 youth had at least one incidence of MDE and 22,105,681 did not; in 2013, 2,587,000 youth had at least one incidence of MDE and 21,590,570 did not; in 2014, 2,751,000 youth had at least one incidence of MDE and 21,380,579 did not; in 2015, 3,031,000 youth had at least one incidence of MDE and 21,217,000 did not; in 2016, 3,089,000 youth had at least one incidence of MDE and 21,043,813 did not; in 2017, 3,214,000 youth had at least one incidence of MDE and 20,951,414 did not; in 2018, 3,482,000 youth had at least one incidence of MDE and 20,698,556 did not; in 2019, 3,783,000 youth had at least one incidence of MDE and 20,312,541 did not. See summary of frequencies in Table 2 of observed values in Section 2.3.2.3 Test Statistic X^2.

As such, the average incidence of MDE from 2004 through 2019 is 2,527,312.5 with a standard deviation of 616,557.942 (calculated in Google Sheets). The large standard deviation indicates that there is great variation in our data from the average. Already, we can see that the incidence of MDE in 2004 is below the average by ~300,000, and the incidence of MDE in 2019 is well above the average by 1,000,000+.

Information about the population was not provided by my data source.

2.3 Statistical Test

2.3.1 Chi-Square Test for Independence

With such a rich set of data, there are many possible directions to investigate. The independent variable of years (2004-2019) could be classified as quantitative or categorical, and the dependent variable of incidence of MDE could also be classified as quantitative (mean incidence of MDE) or categorical (count of youth with incidence of MDE and without). Thus, I considered both the ANOVA test and Chi-Square test. However, since an ANOVA test can only tell whether or not there is a difference between groups, I decided to conduct a Chi-Square test, which would be a relatively more comprehensive test--specifically, comparing the relationship between year and incidence of MDE--and which would answer a question closer to the original intent of my investigation. Furthermore, I chose to conduct a Chi-Square test for independence because I am testing the relationship between two categorical variables.

2.3.2 Hypothesis test

All calculations in this study were conducted using Python. Code snippets are included in this section, but the complete Python file with additional documentation comments line-by-line is attached as well, along with a text file containing the program output. For readability, the program output has also been reformatted in Google Sheets, and the tables are copied into this paper.

2.3.2.1 Hypotheses

- H_0: Incidence of MDE in youth is independent of the year.
- H_A: Incidence of MDE in youth is related to the year.

2.3.2.2 Critical value X^{2*}

- $\alpha = 0.05$
- $df = (R\text{-}1)(C\text{-}1) = (16\text{-}1)(2\text{-}1) = 15$
- Critical value from table = 24.996

2.3.2.3 Test statistic X^2

First, I create the contingency table of observed values. The raw data provides the count and percentage of youth that did experience MDE for each year 2004-2019 (Yes column). From that, I derive the count of youth that did not experience MDE for each year 2004-2009, rounded to the nearest whole person (No column).

```
countInverse = []
for i in range(16):
      conversion = [int(round(rawDataCount[i] / rawDataPercent[i] *
(100-rawDataPercent[i])))]
      countInverse += conversion
```

I also derive the marginal values (Totals per row, Totals per column) by summing each row and each column of data.

```
def annualTotal(listOfValues):
      annualTotal = []
      for i in range(len(listOfValues)):
            totalForOneYear = [sum(listOfValues[i])]
            annualTotal += totalForOneYear
      return(annualTotal)

annualTotalsObserved = annualTotal(observed)
yesNoTotalsObserved = [sum(rawDataCount),sum(countInverse)]
```

		Incidence of Major Depressive Episode (MDE) in past year among youth aged 12 to 17, 2004-2019		
		Yes	**No**	**Total**
Year	**2004**	2225000	22497222	24722222
	2005	2191000	22706727	24897727
	2006	1969000	22955051	24924051
	2007	2016000	22569366	24585366
	2008	2014000	22251060	24265060
	2009	1950000	22124074	24074074
	2010	1911000	21976500	23887500
	2011	2011000	22513390	24524390
	2012	2213000	22105681	24318681
	2013	2587000	21590570	24177570
	2014	2751000	21380579	24131579
	2015	3031000	21217000	24248000
	2016	3089000	21043813	24132813
	2017	3214000	20951414	24165414
	2018	3482000	20698556	24180556
	2019	3783000	20312541	24095541
	Total	40437000	348893544	389330544

Table 2: Observed values.

Next, I create the contingency table of expected values by calculating $Expected_{i,j}$ for each $Observed_{i,j}$, where $Expected_{i,j} = (Row_i * Column_j) / N$.

```
def calculateExpected(rowTotals,columnTotals):
      expected = []
      grandTotal = float(sum(columnTotals))
      for i in range(len(rowTotals)):
            aRow = []
            for j in range(len(columnTotals)):
                  aRow += [rowTotals[i] * columnTotals[j] / grandTotal]
            expected.append(aRow)
      return(expected)
expected =
calculateExpected(annualTotalsObserved,yesNoTotalsObserved)
```

		Incidence of Major Depressive Episode (MDE) in past year among youth aged 12 to 17, 2004-2019		
		Yes	**No**	**Total**
Year	**2004**	2567721.712	22154500.29	24722222
	2005	2585950.171	22311776.83	24897727
	2006	2588684.258	22335366.74	24924051
	2007	2553507.451	22031858.55	24585366
	2008	2520239.54	21744820.46	24265060
	2009	2500403.18	21573670.82	24074074
	2010	2481025.063	21406474.94	23887500
	2011	2547174.307	21977215.69	24524390
	2012	2525808.773	21792872.23	24318681
	2013	2511152.575	21666417.43	24177570
	2014	2506375.816	21625203.18	24131579
	2015	2518467.639	21729532.36	24248000
	2016	2506503.983	21626309.02	24132813
	2017	2509890.018	21655523.98	24165414
	2018	2511462.71	21669093.29	24180556
	2019	2502632.805	21592908.2	24095541
	Total	40437000	348893544	389330544

Table 3: Expected values.

Then, I calculate the test statistic, where $X^2 = \Sigma\ (\ (Observed_{i,j} - Expected_{i,j})^2 / Expected_{i,j}\ )$.

```
testStats = []
for i in range(16):
      for j in range(2):
            toSum = [((observed[i][j]-expected[i][j])**2)/expected[i][j]]
            testStats += toSum
testStatistic = sum(testStats)
```

Thus, our test statistic is 2,640,706.77722.

2.3.2.4 Decision

Reject H_0. Based on our data from 389,330,544 people, we can say that there is a statistically significant relation between what year it is and a youth's incidence of MDE, $X^2(15) = 2{,}640{,}706.77722$, $p < 0.05$.

2.3.2.5 Effect size

$$Cramer's\ V = \sqrt{\frac{X^2}{N \cdot (k-1)}} = \sqrt{\frac{2640706.77722}{389330544 \cdot (2-1)}} = 0.0823570632309$$

Thus, the statistically significant relation between the variables is small.

2.3.3 Data visualization

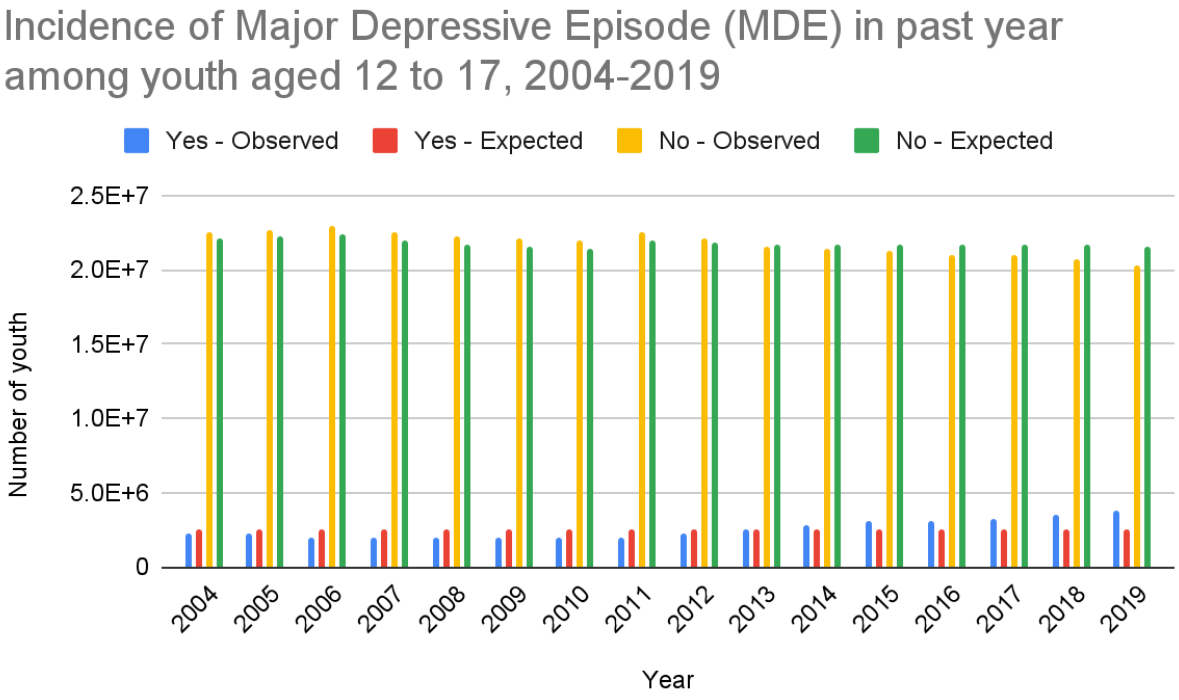

Figure 1: Bar chart visualization of observed and expected values for incidence of MDE in past year among youth aged 12 to 17, 2004-2019.

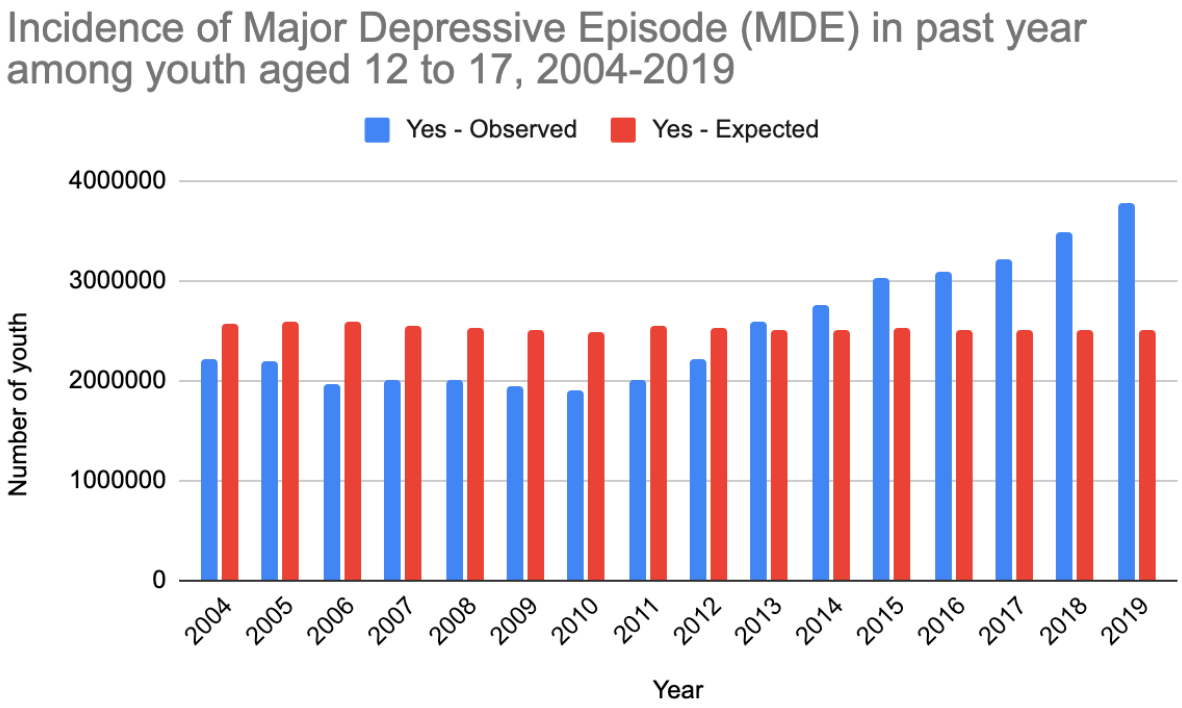

Figure 2: A closer look at Figure 1, observed and expected values of only youth aged 12 to 17 *with* incidence of MDE in past year, 2004-2019.

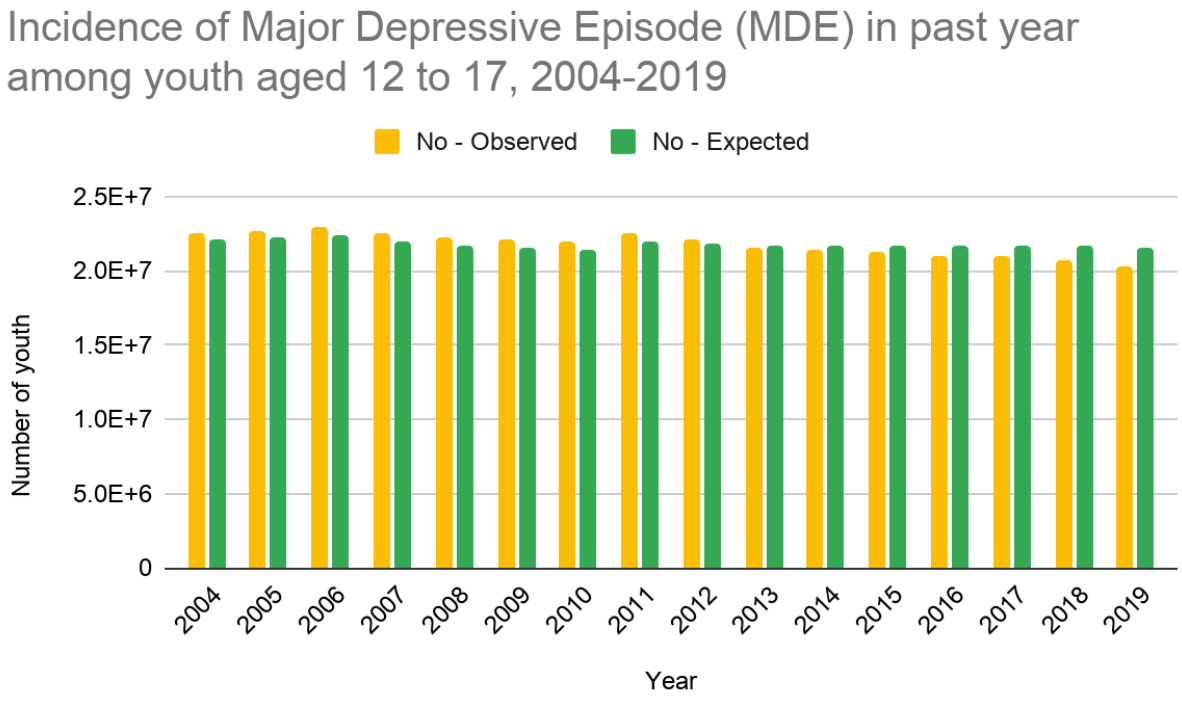

Figure 3: A closer look at Figure 1, observed and expected values of only youth aged 12 to 17 *without* incidence of MDE in past year, 2004-2019.

3 Discussion

Based on my analysis of 389,330,544 youth aged 12 to 17 in the US from 2004 through 2019, there is some relationship between incidence of MDE and the year, though the relation is small. This conclusion does not give a quantifiable difference in the rate of depression in youth in recent years compared to in the past as I originally intended with my research, nor does it formally conclude whether there is indeed a rise of depression in youth, simply that there is some relationship between incidence of MDE and the year.

However, informally, it does appear in Figure 2 that there is a rise in incidence of MDE over time. This is apparent in how the observed number of youth with incidence of MDE per year--represented by blue bars in Figure 2--increases between 2004 and 2019 overall. To be more precise, the observed number of youth with incidence of MDE per year is steady between 2004 and 2010, but then starts to increase from 2011, reaches the highest count since 2004 in 2013, and only continues to increase. This observation in conjunction with the significant findings of my analysis lead me to surmise that the significance in the relationship between incidence of MDE and the year is a result of the more recent increase in incidence of MDE. This would have to be further investigated through a formal analysis by dividing our dataset into one that spans 2004 through 2010 and another that spans 2011 through 2019, then comparing the two effect sizes.

Some other possible avenues for future work are to compare incidence of MDE against a variable other than years, such as financial stability and health of the economy, smartphone usage, or trends in therapy apps; to analyze the same data on youth with a different statistical test, as the variables involved can be treated as either qualitative or quantitative; to break down the demographics of youth and analyze differences between subpopulations of youth; to conduct the same test over the same period of time for adults and compare the results with our analysis on youth; or to conduct the same test over a greater period of time, both before 2004 and after 2019. It would be especially interesting to return to this study once NSDUH's 2020 annual report is released and see how the pandemic year might have affected incidence of MDE, return to this study yet again throughout the future to investigate any lasting effects, and compare effect sizes between these studies.

While NSDUH's comprehensive sampling methodology simplifies the process of collecting new data in the future, the ability to gather data on the past would be a major limitation. Not only would that pose the technical challenge of gathering data consistent with what is currently available, but we would also have to account for the difference in how mental health was regarded and talked about then from how it is now. However, this look into the past should no less be considered and recommended. Besides, even now, it is a complicated topic to discuss. While my analysis treated depression as a binary of with or without, the reality is that depression is not binary but more like a scale. In this way, my analysis is a simplification of the original intention of my research.

4 Source summary

“2019 National Survey on Drug Use and Health (NSDUH): Methodological Summary and Definitions.” Substance Abuse and Mental Health Services Administration, 2020, https://www.samhsa.gov/data/sites/default/files/reports/rpt29395/2019NSDUHMethodsSummDefs/2019NSDUHMethodsSummDefs082120.htm. 30 May 2020.

“Generation Z and Mental Health.” *The Annie E Casey Foundation*, 2021, https://www.aecf.org/blog/generation-z-and-mental-health. 30 May 2021.

Mcmaster, Geoff. “Millennials and Gen Z are more anxious than previous generations.” *University of Alberta Folio*, 2020, https://www.ualberta.ca/folio/2020/01/millennials-and-gen-z-are-more-anxious-than-previous-generations-heres-why.html. 30 May 2021.

Petersen, Anne Helen. “How Millennials Became The Burnout Generation.” *BuzzFeed News*, 2019, https://www.buzzfeednews.com/article/annehelenpetersen/millennials-burnout-generation-debt-work. 30 May 2021.

“Table 2.6A-B: Had at Least One Major Depressive Episode (MDE) or MDE with Severe Impairment in the Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in the Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in the Past Year, by Gender and Detailed Age Category: 2009 and 2010.” *Results from the 2010 National Survey on Drug Use and Health: Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2014, https://www.samhsa.gov/data/report/results-2010-national-survey-drug-use-and-health-mental-health-detailed-tables. 28 May 2021.

“Table 2.6A-B: Had at Least One Major Depressive Episode (MDE) or MDE with Severe Impairment in the Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in the Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in the Past Year, by Gender and Detailed Age Category: 2010 and 2011.” *Results From The 2011 National Survey On Drug Use And Health: Mental Health Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2014, https://www.samhsa.gov/data/report/results-2011-national-survey-drug-use-and-health-mental-health-detailed-tables. 28 May 2021.

“Table 2.6A-B: Had at Least One Major Depressive Episode (MDE) or MDE with Severe Impairment in the Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in the Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in the Past Year, by Gender and Detailed Age Category: 2012 and 2013.” *Results From The 2013 National Survey On Drug Use And Health: Mental Health Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2014, https://www.samhsa.gov/data/report/results-2013-national-survey-drug-use-and-health-mental-health-detailed-tables. 28 May 2021.

"Table 8.37A-B: Had at Least One Major Depressive Episode (MDE) in the Past Year among Persons Aged 12 to 17, by Demographic Characteristics: 2004-2008." *Results from the 2008 National Survey on Drug Use and Health: Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2009, https://www.samhsa.gov/data/report/2008-nsduh-detailed-tables. 28 May 2021.

"Table 9.5A-B: Major Depressive Episode (MDE) or MDE with Severe Impairment in Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in Past Year, by Demographic Characteristics: 2018 and 2019." Results from the 2019 National Survey on Drug Use and Health: Detailed Tables, Substance Abuse and Mental Health Services Administration, 2020, https://www.samhsa.gov/data/report/2019-nsduh-detailed-tables. 28 May 2021.

"Table 9.6A-B: Had at Least One Major Depressive Episode (MDE) or MDE with Severe Impairment in Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in Past Year, by Gender and Detailed Age Category: 2014 and 2015." *Results From The 2015 National Survey On Drug Use And Health: Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2016, https://www.samhsa.gov/data/report/results-2015-national-survey-drug-use-and-health-detailed-tables. 28 May 2021.

"Table 9.6A-B: Had at Least One Major Depressive Episode (MDE) or MDE with Severe Impairment in Past Year among Persons Aged 12 to 17, and Receipt of Treatment for Depression in Past Year among Persons Aged 12 to 17 with MDE or MDE with Severe Impairment in Past Year, by Gender and Detailed Age Category: 2016 and 2017." *Results from the 2017 National Survey on Drug Use and Health: Detailed Tables*, Substance Abuse and Mental Health Services Administration, 2018, https://www.samhsa.gov/data/report/2017-nsduh-detailed-tables. 28 May 2021.

"The rise of anxiety and depression in Gen Z, charted." *Advisory Board*, 2019, https://www.advisory.com/en/daily-briefing/2019/03/18/mental-health. 30 May 2021.

Twenge, Jean M. "Have Smartphones Destroyed a Generation?" *The Atlantic*, 2017, https://www.theatlantic.com/magazine/archive/2017/09/has-the-smartphone-destroyed-a-generation/534198. 30 May 2021.