

7-2020

ANALYSIS ON SUICIDAL IDEATION AMONG ADOLESCENTS (12-17 YEARS) IN THE USA

Himani Raturi

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd>



Part of the [Business Intelligence Commons](#), [Data Science Commons](#), [Other Mental and Social Health Commons](#), and the [Psychiatric and Mental Health Commons](#)

Recommended Citation

Raturi, Himani, "ANALYSIS ON SUICIDAL IDEATION AMONG ADOLESCENTS (12-17 YEARS) IN THE USA" (2020). *Electronic Theses, Projects, and Dissertations*. 1114.
<https://scholarworks.lib.csusb.edu/etd/1114>

This Project is brought to you for free and open access by the Office of Graduate Studies at CSUSB ScholarWorks. It has been accepted for inclusion in Electronic Theses, Projects, and Dissertations by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

ANALYSIS ON SUICIDAL IDEATION AMONG ADOLESCENTS (12-17 YEARS)
IN THE USA

A Project
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Information Systems and Technology

by
Himani Raturi
June 2020

ANALYSIS ON SUICIDAL IDEATION AMONG ADOLESCENTS (12-17 YEARS)
IN THE USA

A Project
Presented to the
Faculty of
California State University,
San Bernardino

by
Himani Raturi
June 2020

Approved by:

Benjamin Becerra, Committee Chair, Information & Decision Sciences

Conrad Shayo, Committee Member

Jay Varzandeh, Committee Member

© 2020 Himani Raturi

ABSTRACT

Suicide is one of the leading health concerns in United States among adolescents and the presence of suicidal ideation (SI) is quite high, with ~20-30% of adolescents reporting it at some point. Though we have seen growth and development in the prevention of suicide, there is limited research on the ability to identify the adolescents which might be at risk for SI. The objective behind the project is to identify adolescents with SI using machine learning.

The project shows statistics from different articles on adolescents in the U.S. For this study, adolescent data was taken from NSDUH 2018. Moreover, detailed associations between demographics, mental health features, etc. and SI were conducted.

From the analysis we saw that functional impairment during MDE on social life, family relationships, schoolwork and home chores play an important role in identifying SI among adolescents. Moreover, machine learning algorithm was implemented to predict SI (accuracy score ~65-70%).

This paper also shows the importance of each feature on the prediction of SI among adolescents. In future work these features could be used to develop an application which will predict SI among adolescents. This screening tool could be potentially utilized by healthcare professionals to screen adolescents during patient encounters. In the future more data could be combined from different years and countries, including the social media information as well.

TABLE OF CONTENTS

ABSTRACT	III
LIST OF FIGURES	VII
CHAPTER ONE: PROJECT BACKGROUND	1
Introduction	1
Problem Statement	2
Purpose of the Study.....	3
Research Questions.....	4
CHAPTER TWO: LITERATURE REVIEW.....	5
Background.....	5
Suicide	5
Suicidal Ideation	6
Screening Methods for Suicidal Ideation Among Adolescents:	7
CHAPTER THREE: DATA DESCRIPTION AND VISUALIZATION TOOLS	8
Strengths and Limitations of NSDUH 2018 Dataset	8
Dataset Descriptions.....	9
Demographics Section	9
Self-Administered Substance Use Section.....	10
Health.....	10
Adolescent Depression	11
Suicide	12
Visualization Tools.....	12

CHAPTER FOUR: ANALYSIS.....	14
Population of Adolescents with Suicidal Ideation.....	15
Total % of Adolescents with Suicidal Ideation Distributed by Age	16
Total % of Adolescents with Suicidal Ideation Distributed by	
Overall health.....	17
Total % of Adolescents with Suicidal Ideation Distributed by	
Family Income	18
Total % of Adolescents with Suicidal Ideation Distributed by MDE:	
Social Life	21
Total % of Adolescents with Suicidal Ideation Distributed by MDE:	
School Work	22
Total % of Adolescents with Suicidal Ideation Distributed by MDE:	
Family Relationships.....	24
Total % of Adolescents with Suicidal Ideation Distributed by	
Alcohol Usage.....	25
Total % of Adolescents with Suicidal Ideation Distributed by	
Race Gender	27
Total % of Adolescents with Suicidal Ideation Distributed by Race	28
Total % of Adolescents with Suicidal Ideation Distributed by Gender	28
CHAPTER FIVE: PROPOSED SOLUTION.....	30
Flowchart for Cleaning the Dataset.....	32
Features for Machine Learning Model	33

Independent Variables	33
Algorithm to Predict the Suicidal Ideation (SI).....	33
Feature Importance & Selection for Prediction.....	34
Application Interface	36
Limitation	37
CHAPTER SIX: FUTURE WORK AND CONCLUSION.....	38
APPENDIX A: PYTHON CODE FOR MACHINE LEARNING ASPECT OF STUDY	41
REFERENCES.....	46

LIST OF FIGURES

Figure 4.1. Total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 suicidal ideation yes vs suicidal ideation no.....	15
Figure 4.2: age vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	16
Figure 4.3. Overall health vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	17
Figure 4.4. Family income vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation...	18
Figure 4.5. Functional impairment in home chores vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation.....	19
Figure 4.6. Functional impairment in home chores vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation: details by sex	20
Figure 4.7. Functional impairment in social life vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation.....	21
Figure 4.8. Functional impairment in social work vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	23
Figure 4.9. Functional impairment in family relationships vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	24
Figure 4.10. Alcohol usage ever vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 suicidal ideation	26
Figure 4.11. Race sex vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	27
Figure 4.12. Race vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation	28

Figure 4.13. Sex vs total percentage of U.S. civilian, non-institutionalized adolescents population from year 2018 with suicidal ideation 28

Figure 5.1. Flowchart used to identify features for a machine learning model to predict suicidal ideation from 2018 national survey on drug use and health data. 32

Figure: 5.2. Random forest feature importance for predicting suicidal ideation .. 35

CHAPTER ONE: PROJECT BACKGROUND

Introduction

The loss of a person who dies by a suicide is one of the most distressing events a clinician can face (C. Campbell and T. Fahy, 2002). Moreover, the suffering that it causes to the family and friends is heartbreaking. Specifically, the loss of an adolescent can create a great impact on other lives involved in different ways. For instance, the parents may develop a 'parental identity crisis' with prominent self-doubt resulting in difficulties in their relationships with surviving children (McIntosh, 1987). The families of those who have died by suicide may have increased risk of suicide themselves (Ness & Prefer, 1990). However, this excess risk might be due to the other confounding factors as well such as family history of depression. There may also be a negative reaction of others towards a family who experienced suicide in their household, as society continues to have a judgmental attitude towards the suicide (Retterstol, 1993). Therefore, it is vital to identify the adolescents who are at risk of suicide or are suffering from thoughts of attempting suicide, since one suicide has influence on the overall mental health of the family, friends and close ones.

The rate of suicide has been increasing over time. For example, in 2015 around the global age-standardized suicide rate was 12 per 100,000 individuals (GBD 2015 Mortality and Causes of Death Collaborators , 2015). This made suicide the 14th leading cause of global mortality. In 2016 the age adjusted

suicide rate increased to 15 per 100,000 population. Almost 45,000 people killed themselves, making suicide the 10th leading cause of the death (Asha Z, Ivey-Stephenson, Alex E. Crosby, Shane P. D. Jack, Tadesse Haileyesus, Marcie-jo Kresnow-Sedacca, 2017).

Suicide is the 3rd leading cause of death in 15-19-year-olds (McLoughlin et al. 2015), and suicidal thoughts and behaviors contribute in the disability among youth worldwide (Gore et al. 2011). According to psychiatry research there is an association between suicidal ideation and lifetime suicide attempts at the low levels of depression (Rogers ML et al. 2018). By analyzing what factors lead to suicidal ideation, this could aid in identifying who may be at risk for attempting the suicide amongst teenagers (12-17 years). What are the external factors to be considered that lead a person to take such an extreme step of ending life at such young age?

Problem Statement

Overall suicide has been identified as a leading cause of death among teenagers in the United States. The prevalence of suicidal thoughts among adolescents is quite high, with an estimated 20-30% of adolescents reporting them at some point (E. Evans, K. Hawton, K. Rodham, and J. Deeks, 2011). Currently, there are various methods to screen suicidal ideation at the primary care level. The typical screening of adolescents involves directly asking them about depression. The approach is consistent with practice guidelines from the American Academy of Pediatrics (Shain B, 2016). However, there may be a great

opportunity for healthcare professionals to screen patients based on a questionnaire and detect potential suicidal ideation during a typical patient encounter. There have been published studies about the use of mobile health, such as Virtual Hope Box app which showed some positive results for individuals at risk of suicide. Though we have growth and development in the prevention of suicide using mobile health tools but there is very limited research on ability to provide intended results using these approaches (Melia, R., Francis, K., Hickey, E., Bogue, J., Duggan, J., O'Sullivan, M., & Young, K. 2020). Additionally, it is unclear which other factors, including depression, that correlate to suicidal ideation among adolescents aged from 12-17 years might be useful in a clinical setting.

Purpose of the Study

This project will attempt to illustrate the factors that are associated with suicidal ideation through machine learning and data analytics techniques and further develop a web application which will predict suicidal ideation among adolescents. For the project, this study was conducted on the dataset provided by the National Survey on Drug Use and Health 2018 (NSDUH) series. This study will illustrate the importance of features which can help in predicting the suicidal ideation among the adolescents aged from 12-17 years for later use in a clinical setting.

We also analyzed the trends of Major Depressive Episodes (MDE), sex, race, ethnicity, age, family income, overall health and ever or never usage of the

substances like alcohol, pain killer, drugs (marijuana, cocaine, etc.). The purpose of this culminating project is to assess the antecedents and determinants of suicidal ideation among 12 - 17-year-olds in the U.S. and suggest effective intervening measures for clinicians to utilize to reduce suicide. Additionally, possible solutions or improvements that could be considered in future will be discussed.

Research Questions

1. What factors are associated with suicidal ideation among adolescents?
2. Does the use of self-administered substances (e.g. tobacco, alcohol, marijuana, cocaine (including crack cocaine), heroin, hallucinogens, inhalants, and methamphetamine) have any correlation with suicidal ideation?
3. What is the demographic information (e.g. Family Income, Sex, Race, Health Insurance, Education, Employment, Health Insurance) of the adolescents who have suicidal ideation?
4. What is the mental, social health of the adolescents who reported functional impairment?
5. Which of these correlates can be used as features that help in predicting suicidal ideation among youth aged from 12-17 years?

CHAPTER TWO: LITERATURE REVIEW

Background

Over the past 20 years there has been extensive research into the risk factors and treatments for suicidal ideation and behavior (D. A. Brent, M. Oquendo, B. Birmaher et al, 2002; B. J. Cox, M. W. Enns, and I. P. Clara, 2004; R. K. Oates, 2004; M. Séguin, J. Lynch, R. Labelle, and A. Gagnon, 2004; A. Spirito and J. Overholser, 2003).

Suicide

The effects of the suicide go way beyond the person who died by suicide, as it can have a very lasting effect on the families, friends and communities. One suicide event in a community can lead to risk of suicide or risk of other issues (such as depression) in many others who are connected to the person who died (Ness & Pfeffer, 1990).

Every year close to 800,000 people die due to suicide, and additional individuals attempt suicide. According to Centers for Disease Control and Prevention (CDC) WISQARS Leading Causes of Death Reports in year 2017, just in United States 45,000 people die by suicide. It is a major health concern in US among adolescents, however it is preventable if it identified at an early stage with person's symptoms and behavior. A fact sheet has been developed and has been provided by the National Institute of Mental Health (NIMH). According to NIMH, risk factors for suicide include prior suicide attempts, depression and other

mental health disorders, substance abuse, ages 15 to 24 years, exposure to suicidal behavior, etc. In addition, people belonging to different race, gender, ethnicities, and ages can be at risk (NIMH 2020). The suicidal ideation can happen to anyone due to the circumstances and unfortunate situation.

Suicidal Ideation

Suicidal Ideation (SI) are the thoughts of killing oneself; sometimes these may include a suicide plan. Suicidal ideation is a very serious threat for the well-being of the adolescents and its strongest factor for the suicide (GyuYoung Lee, Ok Kyung Ham, 2018). Therefore, it is important to identify individuals who suffer from suicidal ideation.

The literature has identified some possible ways that adolescence can move from pondering death in general to potential suicidal ideation (Vander Stoep et al., 2009). Some of these factors include at what age, how often, and for how long suicidal ideation events occurs (Miranda et al. 2014; Nock et al., 2013). The prevalence rate of SI event varies by sex, whereby more females, compared to males, experience SI during adolescence (Nock et al., 2013), although adolescent males are most likely to commit suicide (Gould et al. 2003). Per this study the significant factors in adolescent males includes early dependency and anxiety behaviors, which in female adolescents that early indicator is hyperactivity and aggressiveness.

There are several warning signs that could be first recognize by family and friends (Shain B, 2016), including mental health issues (e.g., major depression,

substance use disorders, psychotic disorders), previous attempts to commit suicide, identifying as a sexual or gender minority (e.g. lesbian, gay, bisexual, transgender), past history of physical or sexual abuse, etc. Other warning signs reported by the American Association of Suicidology include feeling purposelessness, anxiety, feeling trapped, hopelessness, etc. (Rudd MD et al 2006; Wintersteen MB, 2007).

Screening Methods for Suicidal Ideation Among Adolescents:

There are some screening tools available for adolescents, such as the 4-item self-report Ask Suicide-Screening Questions (ASQ; King et al. 2017).

Additional tools include the Tri-Factor Screen for Youth Suicide Risk or Suicidal Ideation Questionnaire-Junior (SIQ-JR), or a simple presence of both depression and alcohol/substance abuse (King et al. 2017). However, there is not enough evidence of a tool that has been developed with the help of machine learning approach that may utilize previously unidentified factors for use in a clinical screening tool.

CHAPTER THREE: DATA DESCRIPTION AND VISUALIZATION TOOLS

To find solutions to these research questions, datasets freely available from National Survey on Drug Use and Health (NSDUH) series were used, provided through the U.S. Substance Abuse and Mental Health Services Administration (SAMHSA). The datasets used in this research project are from 2018. The target population for the 2018 NSDUH survey is a national sample of the civilian, noninstitutionalized population of the U.S. This also includes civilians on military bases who were 12 years of age or older at the time of the survey.

Strengths and Limitations of NSDUH 2018 Dataset

The questions in the NSDUH survey are mostly administered by audio computer-assisted self-interviewing (ACASI), which is designed to provide the respondent with a private and confidential mode for responding to questions. This in turn will increase the level of the honesty in reporting sensitive behaviors, such as suicidal ideation. Moreover, the large and dispersed NSDUH samples enable state level estimates of the U.S.

However, the dataset has certain limitations. First, the data is based on the self-reporting of the drug use, and so recall and/or social desirability biases are possible. Second, the survey is cross-sectional which reduces our ability to make statements of cause and effect. Third, the survey excludes a small portion

of the population (approx. 3%), including active-duty military and individuals in institutional groups.

Here are some of the variables I will be using throughout my research for the analysis of the suicidal ideation amongst adolescents. Moreover, this study is specifically on adolescents, so I will be focusing on the data for the age group 12-17 years old. The variables listed below has more cleaned data and features as the dataset provided by the NSDUH 2018 series has respondents from all age groups, the first step is to narrow it down only for adolescents that are from 12-17 years. Then further dropping the features which were correlated to each other. Since the machine learning algorithm understands only numbers and it doesn't accept missing values, therefore the dataset was further cleaned to drop all the missing values.

Dataset Descriptions

Demographics Section

This section consists of the variables that has the demographic information including age ranging from 12-17 years, sex as Male or Female, race consisting of White, Black/African-American, Native Am/American Indian, Native Hispanic/other Pacific Islander, Hispanic, health Insurance, county (Large Metro, Small Metro, Non-Metro), population density, family income, language of ACASI survey (Spanish, English) and any assistance from the government.

Self-Administered Substance Use Section

Ever or never use of the following self-administered substances were assessed: cigarettes, pipes, any tobacco, alcohol. Marijuana, cocaine, crack heroin, hallucinogens, LSD, PCP, ecstasy, inhalants any pain reliever, cold meds. Also, the past misuse of the opioids as the respondent reported as Yes or No.

Health

The section involves the wellbeing of the respondent in the past year and it includes pregnancy age if the female got pregnant at any point of time. These variables have been chosen based on the association of the health treatment and visits to healthcare providers with suicidal ideation (Pamela A. Carlton, Frank P. Deane, 2000).

Additional health variables including respond yes or no of the following outpatient mental health services received by the adolescents in the past year: Received in-home emotion or behavior treatment, stayed overnight in hospital for emotion or behavior treatment, stayed overnight in hospital for emotion or behavior treatment, spent time in foster care for emotion or behavior treatment, spent time in day treatment for emotion or behavior treatment, received emotion or behavior treatment from mental health clinic, received emotion or behavior treatment from therapist, received emotion or behavior treatment from family doctor, attend school or program for emotion or behavior treatment, talked to school counselor for emotion or behavior treatment, received services from

juvenile detention center, prison, or jail for emotion or behavior treatment; stayed in the medical facilities for nights (1, 2, 3-6, 7-24, 25 or more nights, No nights) in past year for the following: Hospital, residential treatment center, foster care. Further variables included number of times adolescents visited (1, 2, 3-6, 7-24, 25 or more visits, no visits) for the day treatment program, mental health clinic, therapist, and if they received in-home counselling.

Adolescent Depression

These variables explain the impact of major depressive episode (MDE) among adolescents in the past year. During MDE the adolescents are impacted in the domains of life like family relationships, schoolwork, social life, and home chores. These variables are chosen as depression has a strong association with the suicidal ideation (Rogers, M. L., Ringer, F. B., & Joiner, T. E. 2018). Functional impairment during major depressive episode (MDE) was also included, and is classified as none, mild, moderate, severe, very severe for the following: home chores, schoolwork, family relationships and social life. Additional depression-related variables included responding yes or no to the following: lifetime major depressive episode, past year major depressive episode, past year major depressive episode and alcohol dependence or abuse, major depressive episode and ill drug dependence or abuse, major depressive episode and substance dependence or abuse, received medication for MDE, received counselling or treatment; saw or talk to the following for the major depressive episode in past year: medical doctor or professional, general practitioner or family medical

doctor, other medical doctor, psychologist, psychiatrist, counsellor, mental health professional, nurse/occ therapist, another healer, religious or spiritual advisor.

Suicide

This variable is the target variable and explains the adolescent population with the suicidal ideation. This was considered as answering “Yes” to the question “When problems were worst, did you think about killing yourself?” The goal is to predict this variable using machine learning techniques.

Visualization Tools

For exploration and visualization of my dataset, we have used Tableau which is a tool from Tableau Software company headquartered in Seattle, Washington. Currently, Tableau is one of the versatile visualization tools with great capability to connect to different data sources. It is one of the most popular tools in the business intelligence (BI) industry and is actively used by businesses for quick reporting. BI tools like Tableau can turn raw data in easily understandable visuals and provides great insights for businesses. Tableau also has real-time reporting capabilities, blending the data from different sources and collaboration. The interface of Tableau is user-friendly with feature like drag and drop of the columns to convert data into the graphical representation.

This project also used a machine learning approach, in order to predict the suicidal ideation among adolescents using the features in the dataset National

Survey on Drug Use and Health 2018 (NSDUH) series (Center for Behavioral Health Statistics and Quality 2019). The machine learning models are implemented in Python 3.8.3 using a Jupyter notebook. Machine learning is an application/branch of artificial Intelligence. It is based on the idea that the system (machine) can learn from the data, identify the patterns within the data, predict and make decisions with minimal human interaction (Ethem Alpaydin, 2020).

CHAPTER FOUR: ANALYSIS

The following section consists of exploratory data analysis of features that have association with suicidal ideation among adolescents. 2876 adolescents between the ages 12 and 17 in the U.S. were included in this study. As shown in Figure 4.1, more than 56% of the respondents had suicidal ideation from the sample size of our dataset. As per the analysis, Figure 4.2 shows all ages have above 50% of the total adolescents in each category with suicidal thoughts. Figure 4.4 shows that even after having family income above 75K, more than half of the adolescents in that category reported to have suicidal thoughts. The most important features as per the analysis is the functional impairments during the major depressive episode (MDE), with respect to all the domains of life. Figure 4.5, 4.7, 4.8, 4.9 shows that 80-90% of the adolescents have very severe functional impairment and reported to have suicidal ideation.

Population of Adolescents with Suicidal Ideation

Suicidal Ideation

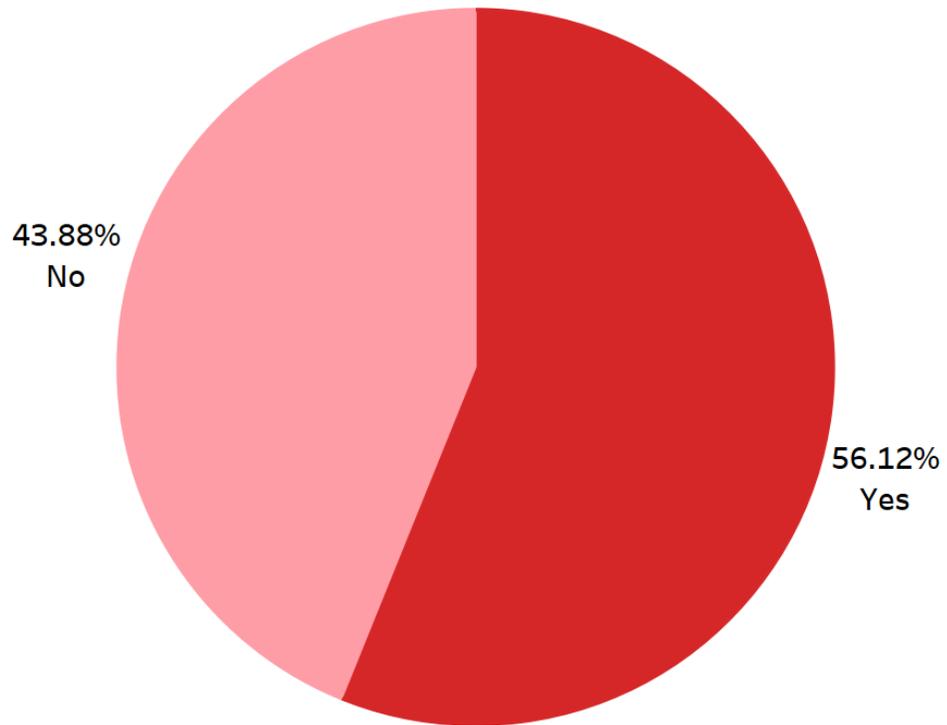


Figure 4.1. Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 Suicidal Ideation **Yes** Vs Suicidal Ideation **No**

Figure 4.1 shows the distribution of the adolescents who thought to end their life when their situation got worst. According to the total population in the dataset 56.12% of the children said 'YES' and 43.88% of the children said 'No'.

Total % of Adolescents with Suicidal Ideation Distributed by Age

Age/Suicide Ideation

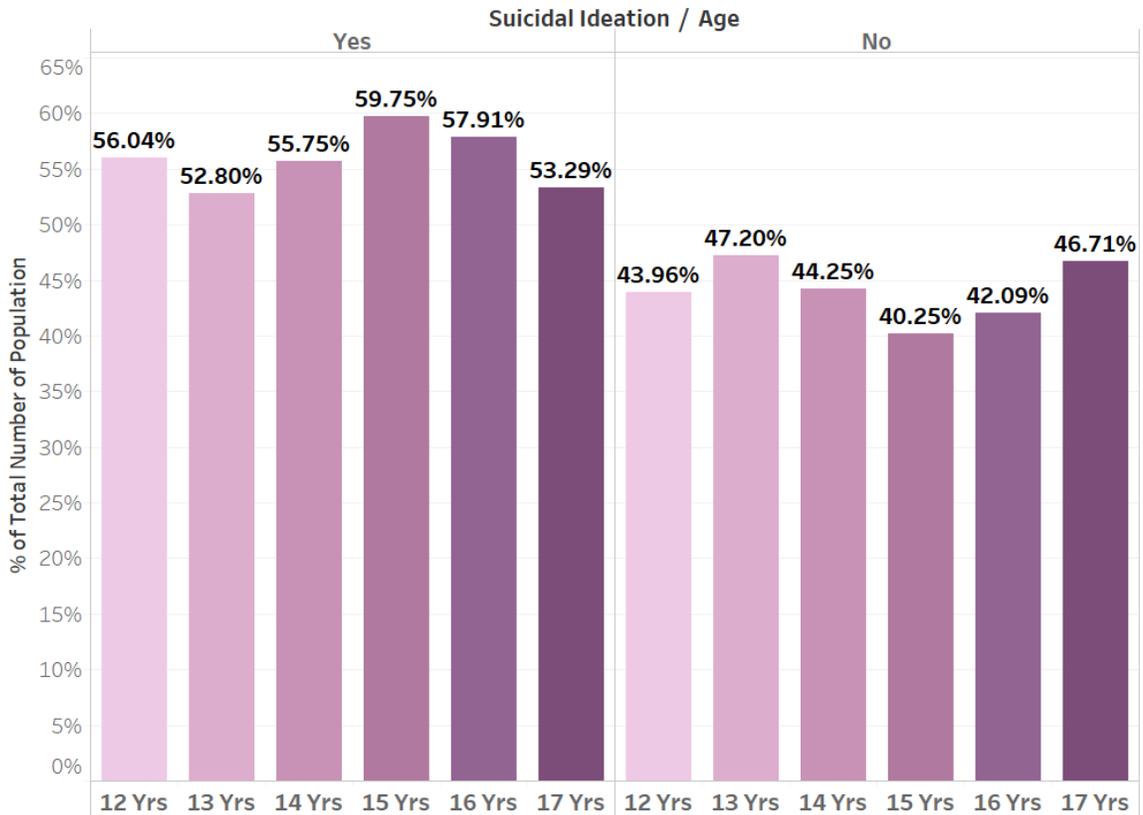


Figure 4.2: Age vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

Figure 4.2 shows the distribution of the age and total adolescents between 12-17 years with suicidal ideation. As we can see that at the early ages 12-17 years the percentage of thinking about suicide is approx. 50% overall however it gradually increases with a few percent in year 15 and 16 years old. The larger

number of adolescents at the age from 15-17 years are more prone to have suicidal thoughts. This could be due to peer pressure, transition from childhood to adulthood, stress due to future career goals, etc.

Total % of Adolescents with Suicidal Ideation Distributed by Overall health

Overall Health

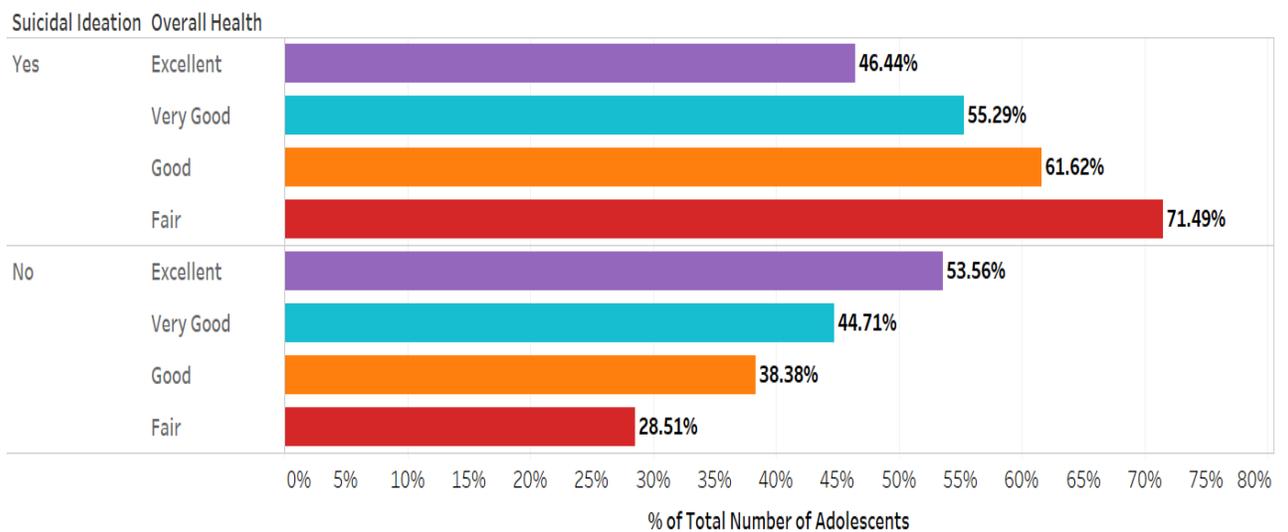


Figure 4.3. Overall Health vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

A comparison of the overall health between the two populations Suicidal Ideation Vs Non- Suicidal Ideation is shown in the Figure 4.3 its somewhat surprising 71.49% of adolescents has suicidal ideation when they reported their health condition as 'Fair'. This would be the clear indication for the people around to be considered as the warning sign to have suicidal ideation. Also, 61% of the adolescents showed suicidal ideation with health condition as 'Good'.

Total % of Adolescents with Suicidal Ideation Distributed by Family Income

Family Income

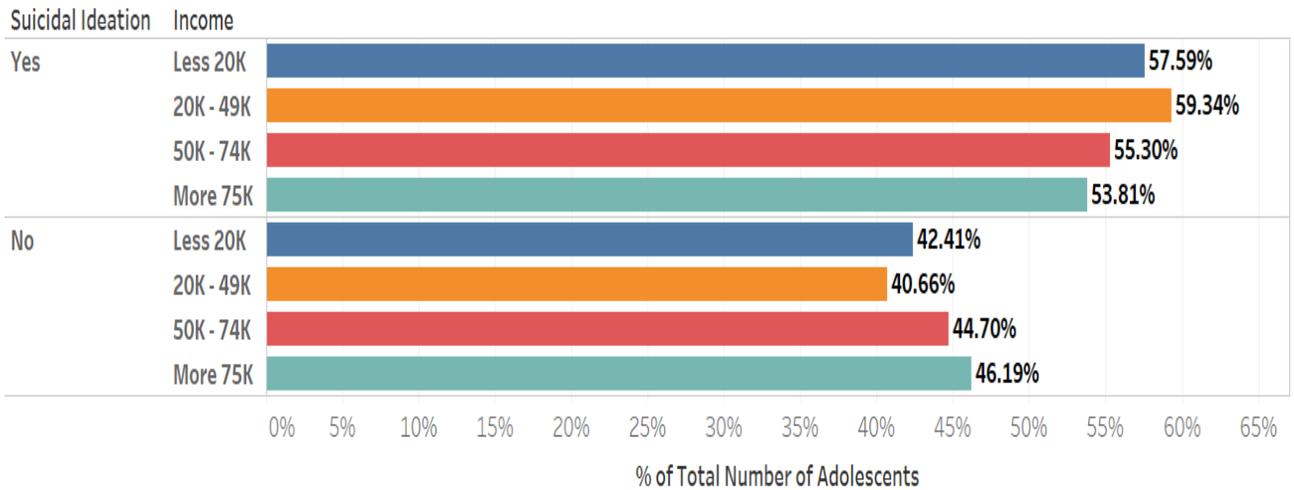


Figure 4.4. Family Income vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

The side by side bar graph in the Figure 4.4 shows the overall income of the family of an adolescent. Almost 50-60% of the adolescents whose family income is less than 49 K reported to have suicidal thoughts. The interesting thing to see here that 53% of the adolescent's has family income more than 75K but still had suicidal thoughts. Though the family income seems to be high, but the data does not show how many members are in the family. Moreover, it is also clear that money is not something that is correlated with the suicidal thoughts.

There are many other major factors which contributes to have thoughts about ending oneself.

Total % of Adolescents with Suicidal Ideation Distributed by MDE: Home Chores

Home Chores

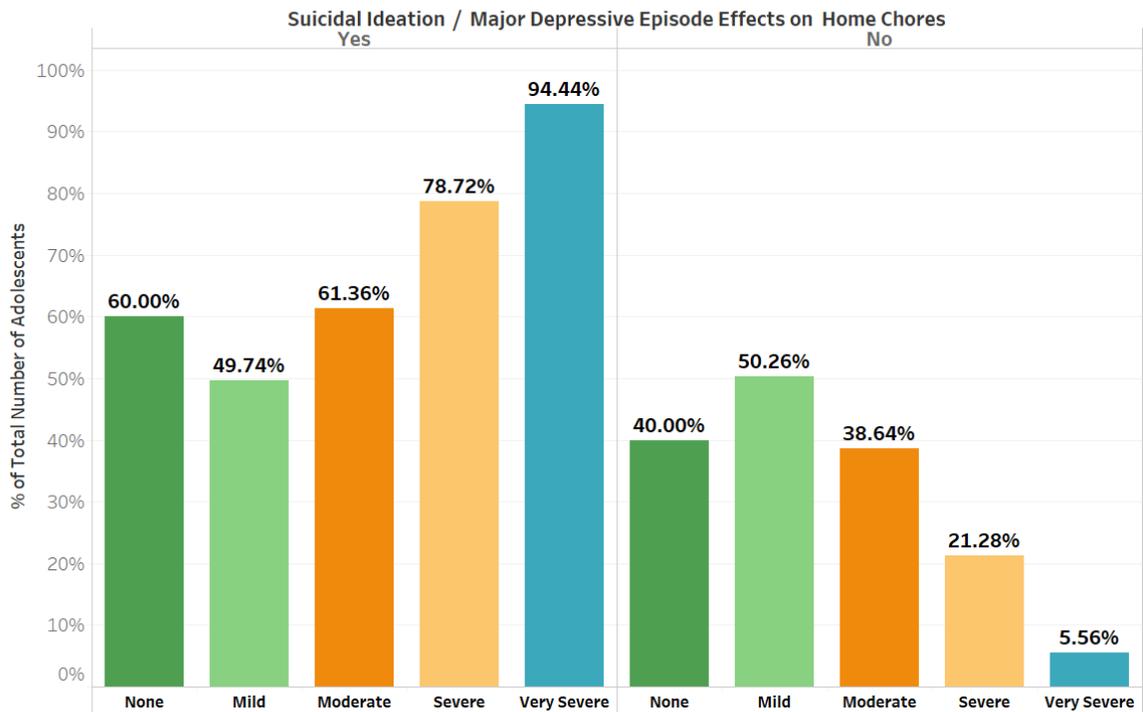


Figure 4.5. Functional impairment in Home Chores vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

Home Chores

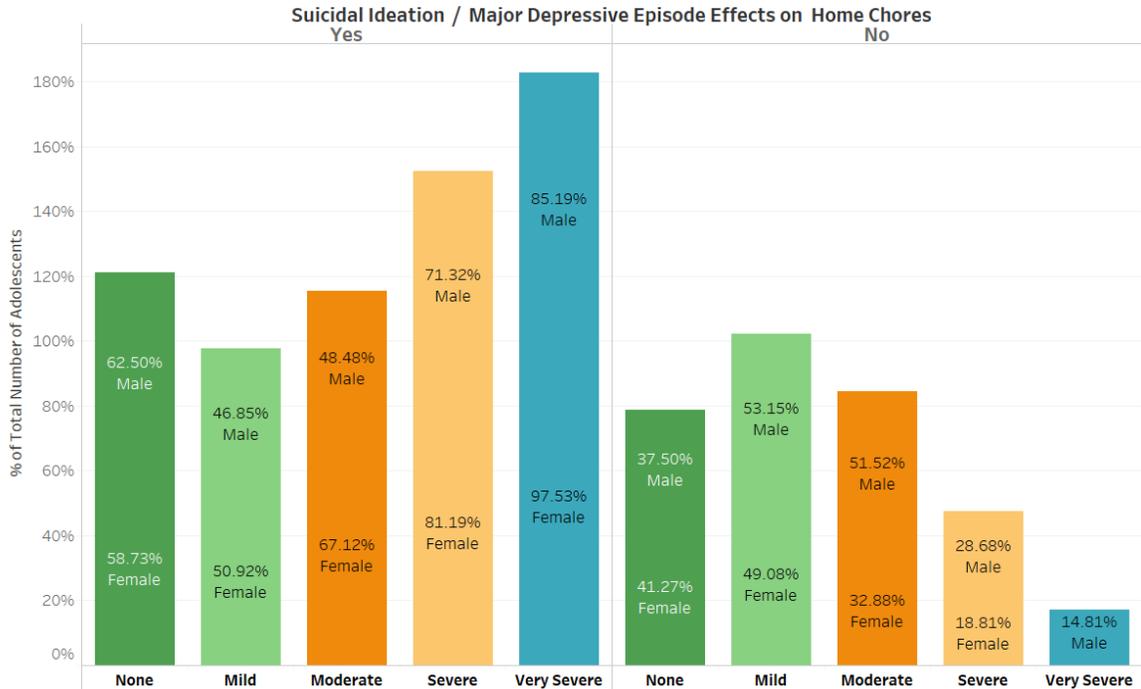


Figure 4.6. Functional impairment in Home Chores vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation: Details by Sex

The one of the important features which correlates with symptoms of the suicidal ideation. The bar graph shows the period of major depressive episode (MDE) in an adolescent's life. The ability to do home chores during the major depressive episode (MDE) is shown in the above graph. 61-78% of the adolescents in the major depressive episode (MDE) shows moderate to severe functional impairment in terms of Home Chores. This 61-78% of adolescent also said to have thoughts about suicide. And almost 94% of adolescents that had very severe functional impairment at homes chores reported to have suicidal ideation. From the above analysis it can be infer that adolescents which shows

functional impairment at homes chores from moderate to very severe might also have suicidal ideation.

Total % of Adolescents with Suicidal Ideation Distributed by MDE: Social Life

Social Life

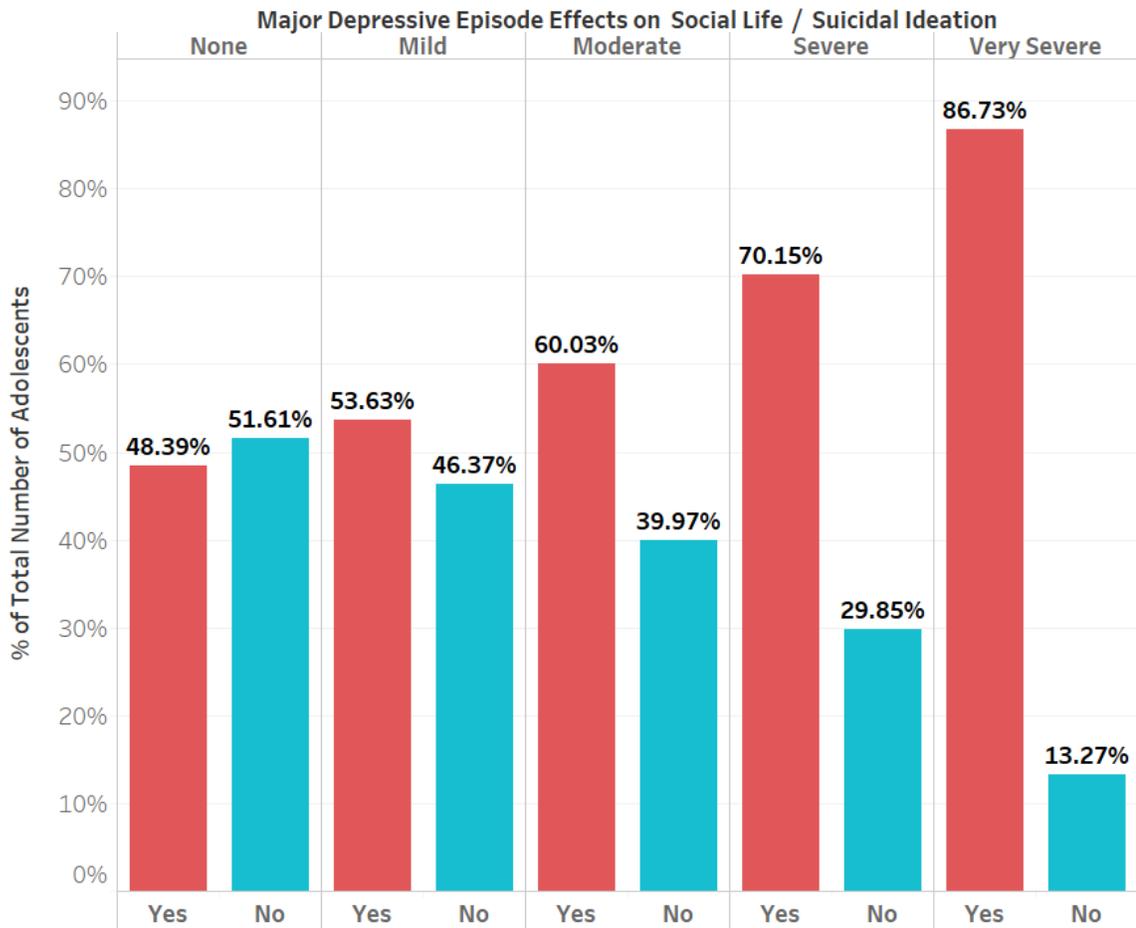


Figure 4.7. Functional impairment in Social Life vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

The above graph is the part of the major depressive episode (MDE) which affects the social life domain of an adolescents. The graph shows the impact of impairment in the social life during this period. As you can see almost 60- 70% of the adolescents whose social life is impacted from moderate to severe, also said 'Yes' for suicidal ideation. 86% adolescents said that social life is impacted very severely, also reported to have suicidal ideation. Therefore, it seems adolescents who suffers from major depressive episode from moderate to very severe impact on social life also might have symptoms of suicidal ideation.

Total % of Adolescents with Suicidal Ideation Distributed by MDE: School Work

School Work

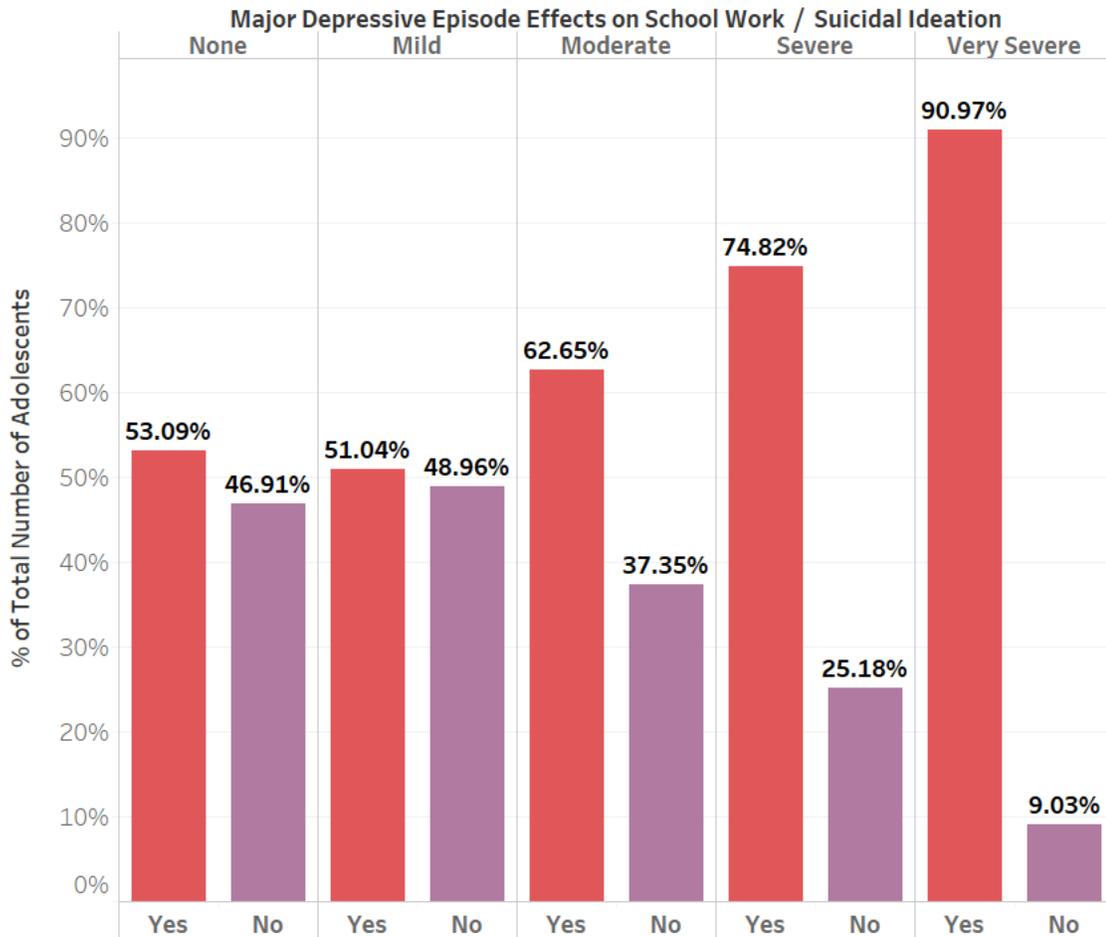


Figure 4.8. Functional impairment in Social Work vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

This is one of the crucial parts of the major depressive episode (MDE) which affects the schoolwork of an adolescents. Figure 4.8 shows that almost 60-70% of adolescents in the moderate to severe categories have suicidal thoughts. Almost more than 60% of the total adolescents suffers from major depressive episode (MDE) which impacts their schoolwork and triggers suicidal ideation

thoughts. It is quite alarming to see that more than 50% of the adolescents from this dataset that showed inability in doing their schoolwork and had suicidal thoughts.

Total % of Adolescents with Suicidal Ideation Distributed by MDE: Family Relationships

Family Relationships

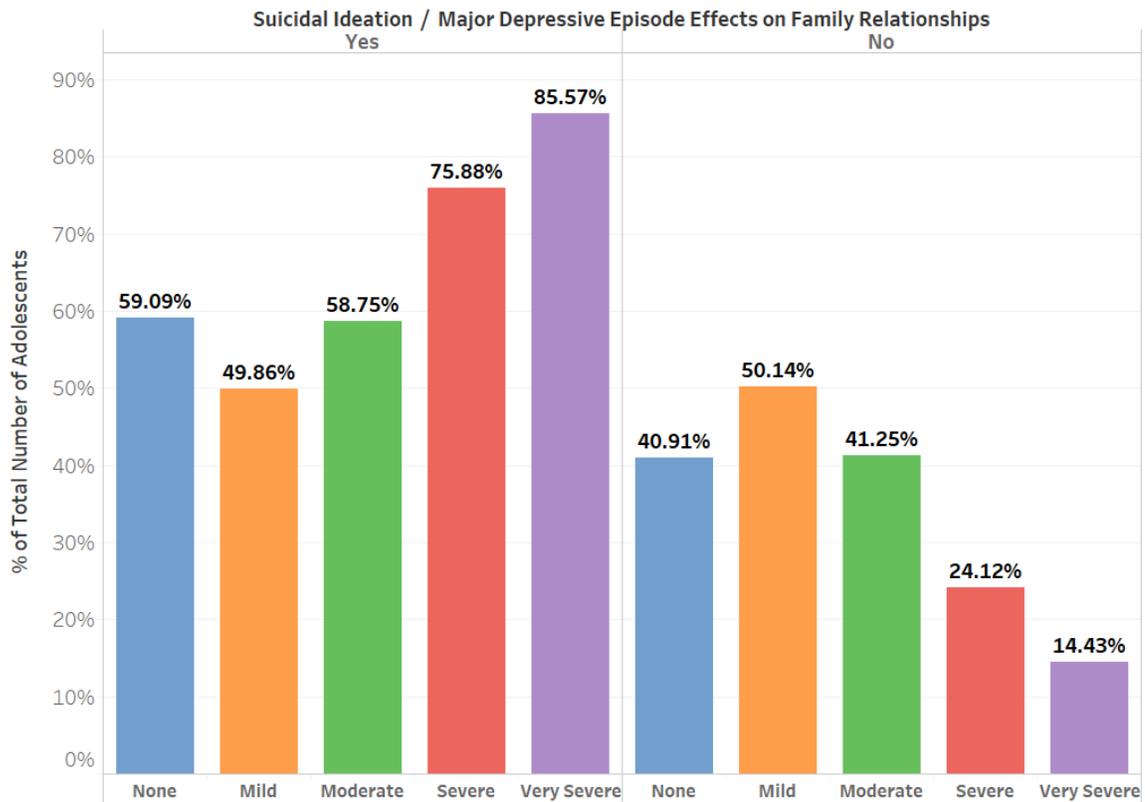


Figure 4.9. Functional impairment in Family Relationships vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

Figure 4.9 indicates an important role that functional family relationships plays on adolescents with major depressive episode (MDE). Sixty to sixty five percent (60- 85%) of the adolescents in each category who had moderate to very severe poor family relationships, during MDE also said they had suicidal ideation. In total almost more than 50% of adolescent's dysfunctional family who had MDE has suicidal ideation at the same time.

Total % of Adolescents with Suicidal Ideation Distributed by Alcohol Usage

Alcohol

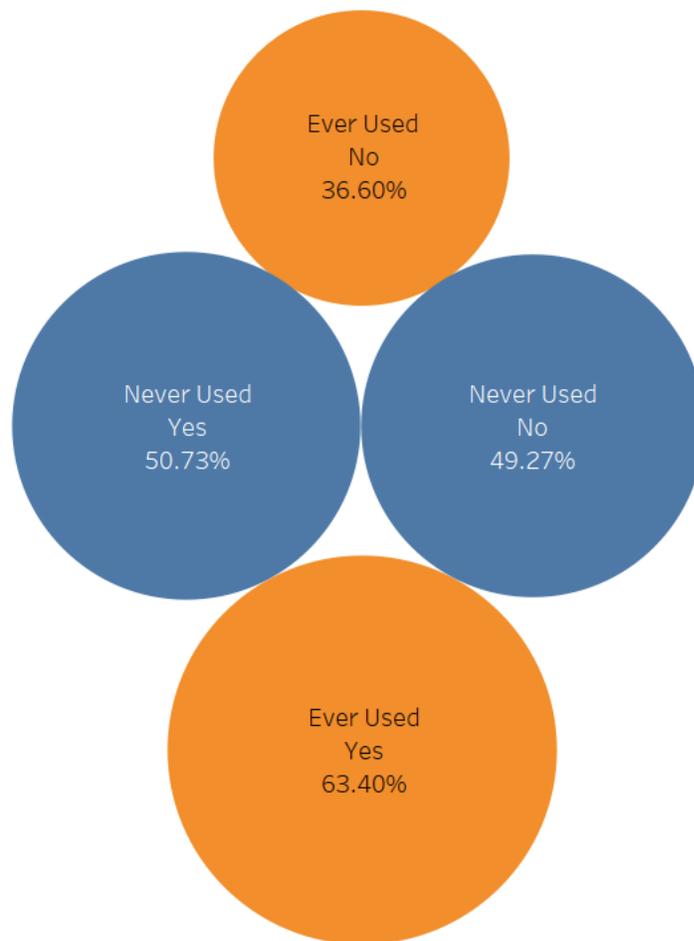


Figure 4.10. Alcohol Usage Ever vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 Suicidal Ideation

Total % of Adolescents with Suicidal Ideation Distributed by Race Gender

Race Gender

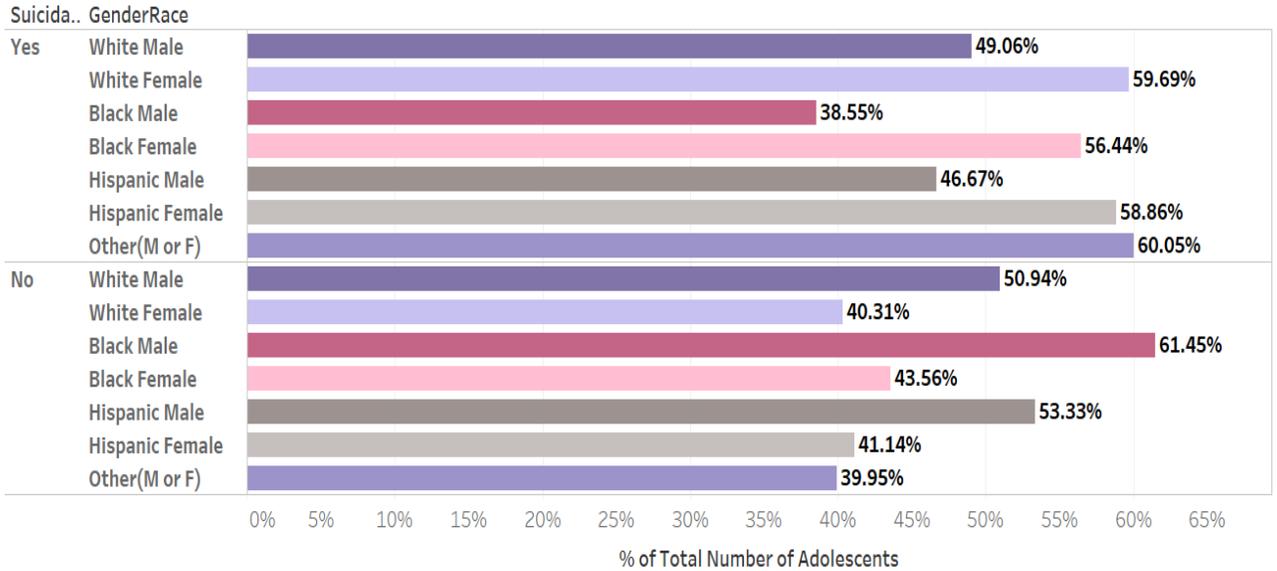


Figure 4.11. Race Sex vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

Total % of Adolescents with Suicidal Ideation Distributed by Race

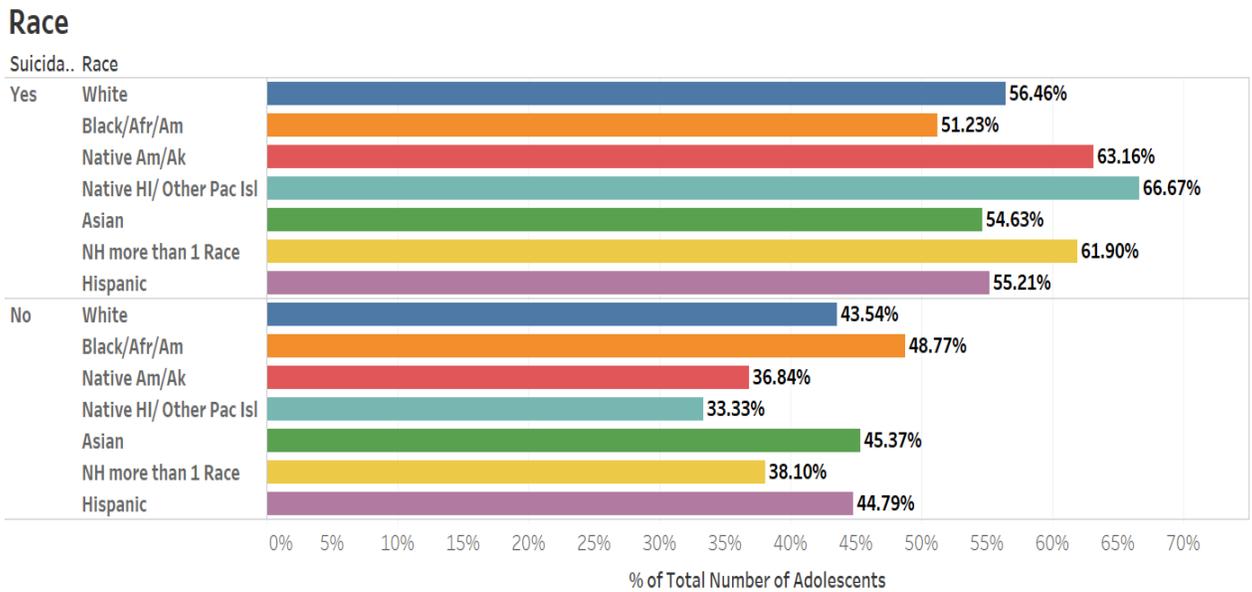


Figure 4.12. Race vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

Total % of Adolescents with Suicidal Ideation Distributed by Gender

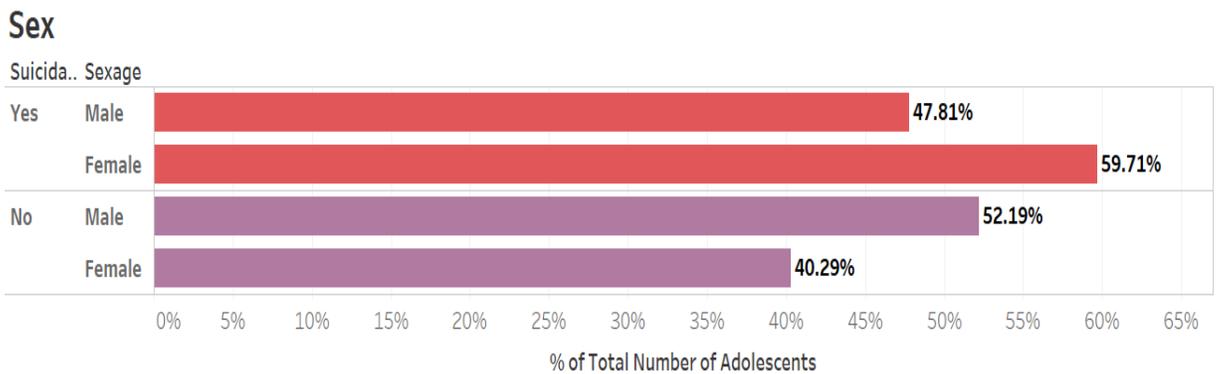


Figure 4.13. Sex vs Total percentage of U.S. civilian, non-institutionalized Adolescents population from year 2018 with Suicidal Ideation

From the Figure 4.11, 4.12, 4.13 the graph and area chart show the distribution of the sex, race and ethnicity of adolescents with respect to having suicidal ideation. From Figure 4.8, we can see the most impacted subgroup are White Females with 59% of the total adolescents in this category, then Hispanic female with 58% and then black female with 56% confirmed to have suicidal ideation. Similarly, Figure 4.9, we see that the most impacted race is the Native American with almost 63% of the total adolescent's population in this category. The analysis shows that almost 40% of the female adolescents from the overall dataset has suicidal ideation. However, the dataset does not have equal proportion of the females and males. In Figure 4.10 we looked into percentage across sex, 60% of the female population out of total confirmed to have suicidal ideation. On the other hand, almost 47% of the total male population confirmed to have suicidal thoughts. In general females are most likely to have suicidal ideation as compared to males. The female population is more at risk as compared to male for having thoughts about suicide.

CHAPTER FIVE: PROPOSED SOLUTION

From the above analysis, we see that more than 50% of adolescents in this dataset developed suicidal ideation their teenage years. As they grow from 12 years – 17 years, we saw the slight increase. We also saw the changes in the ideation thoughts with respect to four different domains of life, home chores, schoolwork, social life and family relationships.

We also saw that most of the adolescents approximately 95% reported major depressive episode in their early ages of life. It is important that teen's emotions, feelings, thoughts etc. be validated periodically rather being ignored or dismissed by the family, friends and community (The American Academy of Child and Adolescent Psychiatry (AACAP), 2019). Healthy relationships with family and friends can help the teens to cope up with emotional challenges in the life. In case, the teen is still struggling, they should be directed to receive professional help.

We also saw from the analysis that almost around 40% of the adolescents said they have used alcohol. The legal drinking age in United States is 21, and all these in the analysis are 17 years or below. Moreover, 27% of those also said 'Yes' for suicidal ideation, this could be the sign of negative peer pressure. It includes the pressure to use alcohol, drugs, or any other risky activities. As adolescents wants to fit in and be liked by their peers; so naturally, they may succumb to peer pressure. To be liked by others they might get engaged in the

risky activities like drugs, alcohol etc. There should be a constant reminder about the self-worth rather than validating it from others. Every now and then there should be comprehensive educational program conducted by school to make teen understand the risk involved with use of alcohol, drugs etc. (Rosalind Brannigan, Mathea Falco, Linda Dusenbury, and William B. Hansen,2004).

In the analysis we also see the trend of having suicidal ideation as the adolescent matures, as we saw a gradual rise in number of adolescents having suicidal ideation as we move from 12 years to 17 years. The reason could be the uncertainty of the future and succeeding in a career. As they get closer to the transition from high school to a college, they may feel lost and direction less (The American Academy of Child and Adolescent Psychiatry (AACAP), 2014). Schools may be able to provide mentoring programs in advance to help them identify their career path. The programs will help teens to identify their interest and talents (Woods, C. & Preciado, M. 2016).

In order to provide the resources to adolescents it's important to identify them first and proactively direct them to the resources. Even after so many awareness programs offered by schools, universities; resources for professional help and counselling should be made available. Suicide is still one of the major health concerns and adolescents still have suicidal ideation.

Given the findings of the exploratory data analysis performed in this project, we attempted to build a machine learning model which helps in predicting

the suicidal ideation among adolescents using the dataset provided by the NSDUH series 2018.

For the scope of the project we filtered the NSDUH 2018 dataset for just adolescents from aged 12-17 years. We had around N = 2876 specific adolescents in the dataset, however there were some missing information in the dataset which we removed to process our machine learning model. After data cleaning, we left with N= 1736, for our model preparation. Initially, the dataset has more than 2000 features but narrowing down to our scope (12-17 years adolescents) and cleaning the redundant information helped to lower the number features to 78.

Flowchart for Cleaning the Dataset

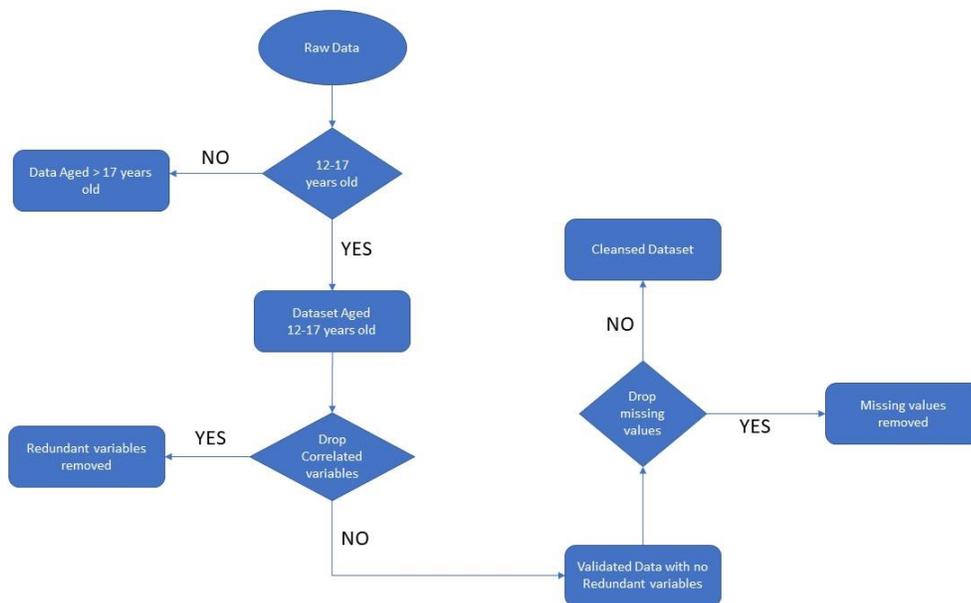


Figure 5.1. Flowchart used to identify features for a machine learning model to predict suicidal ideation from 2018 National Survey on Drug Use and Health Data.

Features for Machine Learning Model

Independent Variables

Independent variables involve the socio- demographic information with variables like sex, race, county, age, family income, overall health. Adolescents depression and their effects on different domains of life. During major depressive episode (MDE) the measurement of effects on social life, family relationships, schoolwork, home chores, visits to therapists or clinics etc. It also includes the substance use (alcohol, marijuana, pain reliever, inhalants, tobacco, cigarettes, pipes, cocaine, crack, heroin, hallucinogens, opioids).

Algorithm to Predict the Suicidal Ideation (SI)

The dataset was split into two: training sets and test set, with test size = 0.3 (Rinu Gour, 2019). The training sets was used to train the model to predict the target variable that is suicidal ideation. The target variable is categorical where 'YES' is encoded as 1 and 'NO' is encoded as 0. Three machine learning methods were trained used the training set: Random Forest (RF), Logistic Regression (LR) and Decision Tree (DT). The reason that we are using these models because of the target variable (Suicidal Ideation) which has two binary output (Kirasich, Kaitlin; Smith, Trace; and Sadler, Bivin, 2018). Since we have a target variable (Suicidal Ideation) which a form of supervised learning in which algorithm aims to classify an input to see in which category it belongs to. After testing the all these different machine learning algorithms, the comparison was done based on the accuracy score, confusion matrix, precision, f1-score. The

logistics regression model has the accuracy score within the range of 68-70%, random forest has the has the accuracy score within the range of 66-69% and decision tree has the accuracy within the range of 59-61%. Based on the accuracy score and the classification report we chose random forest algorithm to predict the suicidal ideation.

Random forest is one of the most common algorithms with most accurate results for prediction. It is the collection of the decision tress, while training the model you can specify the number of trees through n-estimators which is a parameter to the model. Moreover, you can define the depth of the decision tree through max_depth which is also a parameter to the model. The algorithm also has the randomness to the model as it searches for the best features or independent variable among the random set of the data.

Feature Importance & Selection for Prediction

Random forest also has a classic property to measure the importance of each feature in predicting the target variable (Pedregosa *et al* 2011). In our project we also implemented the feature importance after training the model, to see which feature (in our dataset we have 77) plays the at most important in predicting the suicidal ideation among adolescents. The column bar graph shows the feature importance as per the trained model to predict the suicidal ideation.

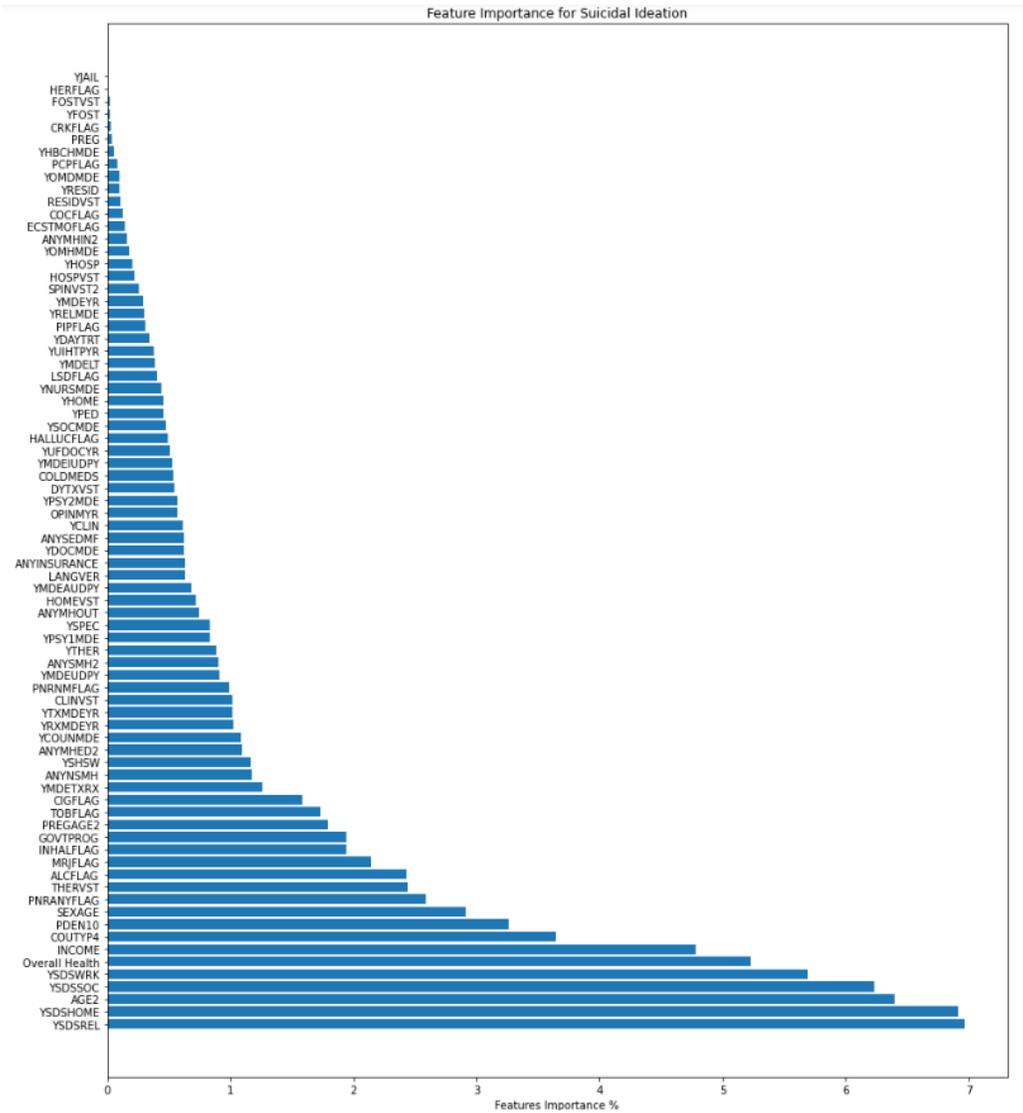


Figure: 5.2. Random forest feature importance for predicting suicidal ideation

For further tuning our random forest model to improve the accuracy and prediction its importance to select the best feature which best helps in predicting the suicidal ideation. The feature selection property helps to select only those features whose importance is greater than the mean of all the features importance. Using the feature selection method in our random forest model

helped us in identifying the 20 features out of 77 that is used in predicting the suicidal ideation among adolescents. However, accuracy score didn't improve but it remained the same even after reducing it 20 features

Application Interface

The project also attempted to provide the idea for a web application to predict the suicidal ideation with user inputs based on the 20 features that was selected using the feature selection method. The interface of the web application will have 20 questions which has options to select from the dropdown. Once the selection is done based on all 20 questions, the details is used by the trained random forest classifier to predict if the adolescents is at risk for suicidal ideation or not.

The idea behind the web application is to be used by healthcare professionals like Doctor, Psychiatrist, therapist etc. as a screening tool in general with all the patients that are aged from 12-17 years old. The solution also proposes that there should a routine checkup by the doctors for the adolescents. And as a part of the routine checkup without explicitly telling the adolescents, healthcare professional can use this web application to see if the adolescents are under the risk to have suicidal ideation. If the adolescents are found to be under risk for the suicidal ideation they should be directed to the resources as well as concerned authorities should be informed like parents.

The goal is to have early detection of the suicidal ideation in adolescents and proactively providing them the resources and professional help if needed.

The algorithm used by the web-application has accuracy that fluctuates between the 66-70%, therefore the web application sometime may lead to inaccurate results. Hence, it's important to cross-verify with other clinical methods, tests, and other factors as well which might influence the prediction.

Limitation

Data used for the analysis has the information related to substance use like alcohol, drugs etc., health, demographic, depression. In today's world analyzing adolescents could not be just limited to the aspects of health, depression, demographics. Since we are surrounded by technology and rapid growth of social media such as Snapchat, Facebook, Instagram, Twitter have a powerful influence on lives of the adolescents. Their popularity is being seen by the likes, followers, shares and comments a person gets on a social media platform. As a result, this could create a feeling of insecurity, jealousy, loneliness and mental pressure. Also, we also see the increase in cybercrime, cyberbullying and trolling which have a greater impact on the mental health of the adolescents. Therefore, it is really important to look into usage of social media aspects by adolescents. Including these factors to the analysis would have a different outcome and more accurate results for the prediction of the suicidal ideation (Roy, A., Nikolitch, K., McGinn, R., Jinah, S., Klement, W., & Kaminsky, Z. A.,2020).

CHAPTER SIX: FUTURE WORK AND CONCLUSION

The algorithm used for the prediction can be more improved in the accuracy score to predict more accurately. The current dataset is just from the one source, the validation of the random forest model should be done with other series of the dataset provided by the NSDUH. The dataset is available for various years in the website. In this project we have used the 2018 NSDUH series dataset. Therefore, validating our model with more data could help in boosting the predicting accuracy.

Human beings are complex creatures and every individual is different. It is therefore important to cross validate the predictive power of the random forest model with other machine learning models like logistics regression, decision tree, support vector machine etc. Currently, the web application gives the result on the basis of one algorithm, however it would be better to check the details using other models as well and then come up with the conclusion if the adolescents has suicidal ideation or not. More advancement has to be done at the design level of the application to be more specific as per healthcare professional.

Moreover, the data we are using is currently only from the U.S. which involves socio -demographics, substance use and adolescent's depression. However, since we are surrounded by the technology, social media plays a vital role in the adolescent's life. Hence, including the social media information will

help us to understand the emotions, sentiments, mood, and other factors as well which can further help in predicting the suicidal ideation. In the future we may use the data from other countries for the comparative studies.

To conclude, our research has shown that more than 50% of the adolescents suffer from suicidal ideation. Even after so much advancement and progress in technology, suicide is one of the leading health concerns in United States. As per data for the latest year 2019, the suicide rate among teenagers in the past year 17.2 percent of high school students seriously considered attempting suicide and 7.4% of them did attempted suicide (America's Health Rankings analysis of CDC WONDER Online Database, 2020). Additionally, due to the pandemic COVID 19, it might become worse. There is a likely chance of having impacts of the pandemic on people with mental illness or on the population in general. In the current situation when there is loss of employment and financial crisis, suicide rate could also surge as the mentioned reasons are well recognized risk factors for suicide (Gunnell, D et al. 2020).

According to our analysis from the NSDUH dataset 2018, we have seen that ages 15-17 years are more likely to have thoughts of suicidal ideation. We have also seen that the adolescents with moderate to severe inability in doing home chores, schoolwork, social life, and family relationships during depression are under risk to have suicidal ideation. We have further implemented the machine learning algorithm i.e. random forest to predict the suicidal ideation based on the features provided in the NSDUH dataset.

The project also attempted to implement the web application which will identify the adolescents that are at risk for having suicidal ideation. The motive behind the web application is early detection of adolescents who are at risk. This will further help in proactively providing the resources and professional help early, preventing any extreme step taken by the adolescents.

APPENDIX A:
PYTHON CODE FOR MACHINE LEARNING ASPECT OF STUDY

Python Code for Machine Learning Aspect of Study

```
import numpy as np # imports a fast numerical programming library
import scipy as sp #imports stats functions, amongst other things
import matplotlib as mpl # this actually imports matplotlib
import matplotlib.cm as cm #allows us easy access to colormaps
import matplotlib.pyplot as plt #sets up plotting under plt
import pandas as pd #lets us handle data as dataframes
import seaborn as sns #sets up styles and gives us more plotting options

df1 = pd.read_excel("NSDUH_DATA_JMP_Update.xlsx")
df1.head(5)

df1.columns = df1.iloc[0]
df1 = df1.drop(df1.index[0])

df1.head(5)
df1.shape
df1 = df1[df1.YOWRSTHK != 94 ]
df1 = df1[df1.YOWRSTHK != 97]
df1.shape

df1.info(all)
df1.describe()
df1.columns[df1.isnull().any()]
df = df1.drop(columns=['YOWRSATP', 'YOWRSPLN', 'YO_MDEA9'])
df = df.astype('category')
df.info(all)

df2 = df.drop(columns=[ 'NEWRACE2', 'PREGNANT', 'PREG2', 'SEXRACE' ])
df2.shape

Y = df2.YOWRSTHK.values
Y
cols = df2.shape[1]
X = df2.loc[:, df2.columns != 'YOWRSTHK', ]
X.columns;
X.shape
```

```

from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.metrics import r2_score, mean_squared_error
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.30,
random_state=99)

# Logistic Regression
logreg = LogisticRegression()
logreg.fit(X_train, Y_train)
Y_pred = logreg.predict(X_test)
acc_log = round(logreg.score(X_train, Y_train) * 100, 2)
acc_log

# Decision Tree
decision_tree = DecisionTreeClassifier()
decision_tree.fit(X_train, Y_train)
Y_pred = decision_tree.predict(X_test)
acc_decision_tree = round(decision_tree.score(X_test, Y_test) * 100, 2)
acc_decision_tree

# Random Forest
random_forest = RandomForestClassifier(n_estimators=100)
random_forest.fit(X_train, Y_train)
Y_pred = random_forest.predict(X_test)
random_forest.score(X_train, Y_train)
acc_random_forest = round(random_forest.score(X_test, Y_test) * 100, 2)
acc_random_forest

print("Scores of different Machine Learning Algorithms with Target variable
Suicide Symptoms")
models = pd.DataFrame({

```

```

'Model': [ 'Logistic Regression', 'Random Forest', 'Decision Tree'],
'Score': [ acc_log,acc_random_forest, acc_decision_tree]])
models.sort_values(by='Score', ascending=False)

```

```

from sklearn.metrics import classification_report, confusion_matrix
from sklearn.ensemble import RandomForestClassifier

```

```

x,y = df.loc[:,df.columns != 'YOWRSTHK'], df.loc[:, 'YOWRSTHK']
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.3,random_state =
1)
rf = RandomForestClassifier(random_state = 4)
rf.fit(x_train,y_train)
y_pred = rf.predict(x_test)
cm = confusion_matrix(y_test,y_pred)
print ('Confusion matrix: \n',cm)
print ('Classification report: \n',classification_report(y_test,y_pred))
y_test.value_counts()

```

```

feature_importance = pd.DataFrame(random_forest.feature_importances_,
                                columns=['importance'])
feature_importance['features']= X.columns
feature_importance.sort_values(by=['importance'],ascending=False,inplace=True
)
print ("Random Forest Feature Importance:\n", feature_importance)

```

```

#Feature_importance plot of random forest classification
plt.figure(figsize=(15, 18))
feature_importance_per = 100.0 * (feature_importance['importance'])
pos=feature_importance['features']
plt.barh(pos,feature_importance_per, align='center')
plt.yticks(pos)
plt.xlabel('Features Importance %')
plt.title('Feature Importance for Suicidal Ideation')
plt.show()

```

```

from sklearn.feature_selection import SelectFromModel

```

```
sel = SelectFromModel(RandomForestClassifier(n_estimators = 100))
sel.fit(X_train, y_train)
```

```
sel.get_support()
selected_feat= X_train.columns[(sel.get_support())]
len(selected_feat)
print(selected_feat)
```

REFERENCES

- A Historical Perspective on Suicide. (n.d.). Psychiatric Times. Retrieved June 26, 2020, from <https://www.psychiatrictimes.com/view/historical-perspective-suicide>
- Adolescence, C. on. (2000). Suicide and Suicide Attempts in Adolescents. *Pediatrics*, 105(4), 871–874. <https://doi.org/10.1542/peds.105.4.871>
- America's Health Rankings analysis of CDC WONDER Online Database, Underlying Cause of Death, Multiple Cause of Death files, United Health Foundation, AmericasHealthRankings.org, Accessed 2020
- Brent, D. A., Oquendo, M., Birmaher, B., Greenhill, L., Kolko, D., Stanley, B., Zelazny, J., Brodsky, B., Bridge, J., Ellis, S., Salazar, J. O., & Mann, J. J. (2002). Familial Pathways to Early-Onset Suicide Attempt: Risk for Suicidal Behavior in Offspring of Mood-Disordered Suicide Attempters. *Archives of General Psychiatry*, 59(9), 801–807. <https://doi.org/10.1001/archpsyc.59.9.801>
- Campbell, C., & Fahy, T. (2002). The role of the doctor when a patient commits suicide. *Psychiatric Bulletin*, 26(2), 44–49. <https://doi.org/10.1192/pb.26.2.44>
- Cox, B. J., Enns, M. W., & Clara, I. P. (2004). Psychological Dimensions Associated with Suicidal Ideation and Attempts in the National Comorbidity Survey. *Suicide and Life-Threatening Behavior*, 34(3), 209–219. <https://doi.org/10.1521/suli.34.3.209.42781>

Curtin, S. C., Warner, M., & Hedegaard, H. (2016). Increase in Suicide in the United States, 1999-2014. NCHS Data Brief, 241, 1–8.

Division (DCD), D. C. (2015, June 7). Does depression increase the risk for suicide? [Text]. HHS.Gov. <https://www.hhs.gov/answers/mental-health-and-substance-abuse/does-depression-increase-risk-of-suicide/index.html>

Ethem Alpaydin (2020). Introduction to Machine Learning (Fourth ed.). MIT. pp. xix, 1–3, 13–18. ISBN 978-0262043793.

Evans, E., Hawton, K., Rodham, K., Psychol, C., & Deeks, J. (2005). The Prevalence of Suicidal Phenomena in Adolescents: A Systematic Review of Population-Based Studies. *Suicide and Life-Threatening Behavior*, 35(3), 239–250. <https://doi.org/10.1521/suli.2005.35.3.239>

Evans, R., White, J., Turley, R., Slater, T., Morgan, H., Strange, H., & Scourfield, J. (2017). Comparison of suicidal ideation, suicide attempt and suicide in children and young people in care and non-care populations: Systematic review and meta-analysis of prevalence. *Children and Youth Services Review*, 82, 122–129. <https://doi.org/10.1016/j.childyouth.2017.09.020>

Fuller-Thomson, E., Hamelin, G. P., & Granger, S. J. R. (2013, January 30). Suicidal Ideation in a Population-Based Sample of Adolescents: Implications for Family Medicine Practice [Research Article]. *ISRN Family Medicine*; Hindawi. <https://doi.org/10.5402/2013/282378>

- GBD 2015 Mortality and Causes of Death Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* (London, England), 388(10053), 1459–1544. [https://doi.org/10.1016/S0140-6736\(16\)31012-1](https://doi.org/10.1016/S0140-6736(16)31012-1)
- Gunnell, D., Appleby, L., Arensman, E., Hawton, K., John, A., Kapur, N., Khan, M., O'Connor, R. C., Pirkis, J., Appleby, L., Arensman, E., Caine, E. D., Chan, L. F., Chang, S.-S., Chen, Y.-Y., Christensen, H., Dandona, R., Eddleston, M., Erlangsen, A., ... Yip, P. S. (2020). Suicide risk and prevention during the COVID-19 pandemic. *The Lancet Psychiatry*, 7(6), 468–471. [https://doi.org/10.1016/S2215-0366\(20\)30171-1](https://doi.org/10.1016/S2215-0366(20)30171-1)
- Hemenway, D., & Miller, M. (2002). Association of rates of household handgun ownership, lifetime major depression, and serious suicidal thoughts with rates of suicide across US census regions. *Injury Prevention*, 8(4), 313–316. <https://doi.org/10.1136/ip.8.4.313>
- History of suicide. (2020). In Wikipedia. https://en.wikipedia.org/w/index.php?title=History_of_suicide&oldid=959778684
- Im, Y., Oh, W.-O., & Suk, M. (2017). Risk Factors for Suicide Ideation Among Adolescents: Five-Year National Data Analysis. *Archives of Psychiatric Nursing*, 31(3), 282–286. <https://doi.org/10.1016/j.apnu.2017.01.001>

- Ivey-Stephenson, A. Z., Crosby, A. E., Jack, S. P. D., Haileyesus, T., & Kresnow-Sedacca, M.-J. (2017). Suicide Trends Among and Within Urbanization Levels by Sex, Race/Ethnicity, Age Group, and Mechanism of Death—United States, 2001-2015. *Morbidity and Mortality Weekly Report. Surveillance Summaries* (Washington, D.C.: 2002), 66(18), 1–16. <https://doi.org/10.15585/mmwr.ss6618a1>
- Kirasich, Kaitlin; Smith, Trace; and Sadler, Bivin (2018) "Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets," *SMU Data Science Review: Vol. 1: No. 3 , Article 9*. Available at: <https://scholar.smu.edu/datasciencereview/vol1/iss3/9>
- Medline ® Abstracts for References 6,7 of 'Suicidal ideation and behavior in children and adolescents: Evaluation and management'—UpToDate. (n.d.). Retrieved June 24, 2020, from <https://www.uptodate.com/contents/suicidal-ideation-and-behavior-in-children-and-adolescents-evaluation-and-management/abstract/6,7>
- NIMH » Suicide in America: Frequently Asked Questions. (n.d.). Retrieved June 26, 2020, from <https://www.nimh.nih.gov/health/publications/suicide-faq/index.shtml>
- NIMH » Suicide. (n.d.). Retrieved June 26, 2020, from <https://www.nimh.nih.gov/health/statistics/suicide.shtml>
- Oates, R. K. (2004). Sexual abuse and suicidal behavior. *Child Abuse & Neglect*, 28(5), 487–489. <https://doi.org/10.1016/j.chiabu.2004.04.002>

- Pedregosa et al., *JMLR* 12, pp. 2825-2830, 2011, From Scikit-learn: Machine Learning in Python,
- Rich, A. R., Kirkpatrick-Smith, J., Bonner, R. L., & Jans, F. (1992). Gender Differences in the Psychosocial Correlates of Suicidal Ideation Among Adolescents. *Suicide and Life-Threatening Behavior*, 22(3), 364–373.
<https://doi.org/10.1111/j.1943-278X.1992.tb00741.x>
- Rogers, M. L., Ringer, F. B., & Joiner, T. E. (2018). The association between suicidal ideation and lifetime suicide attempts is strongest at low levels of depression. *Psychiatry Research*, 270, 324–328.
<https://doi.org/10.1016/j.psychres.2018.09.061>
- Roy, A., Nikolitch, K., McGinn, R., Jinah, S., Klement, W., & Kaminsky, Z. A. (2020). A machine learning approach predicts future risk to suicidal ideation from social media data. *Npj Digital Medicine*, 3(1), 1–12.
<https://doi.org/10.1038/s41746-020-0287-6>
- Rudd, M. D., Berman, A. L., Joiner, T. E., Nock, M. K., Silverman, M. M., Mandrusiak, M., Van Orden, K., & Witte, T. (2006). Warning signs for suicide: Theory, research, and clinical applications. *Suicide & Life-Threatening Behavior*, 36(3), 255–262.
<https://doi.org/10.1521/suli.2006.36.3.255>
- Schwartz, J., Somaiya, M., Cosner, C., & Foster, A. (2018). Natural History of Suicidal Ideation: Insights from the Electronic Health Record. *Journal of*

Technology in Behavioral Science, 3(2), 58–62.

<https://doi.org/10.1007/s41347-017-0037-z>

Séguin, M., Lynch, J., Labelle, R., & Gagnon, A. (2004). Personal and Family Risk Factors for Adolescent Suicidal Ideation and Attempts. *Archives of Suicide Research*, 8(3), 227–238.

<https://doi.org/10.1080/13811110490444379>

Spirito, A., & Overholser, J. (2003). The suicidal child: Assessment and management of adolescents after a suicide attempt. *Child and Adolescent Psychiatric Clinics of North America*, 12(4), 649–665.

[https://doi.org/10.1016/S1056-4993\(03\)00034-8](https://doi.org/10.1016/S1056-4993(03)00034-8)

Standardized Screening for Suicidal Adolescents in Primary Care. (n.d.).

Retrieved June 24, 2020, from

https://www.researchgate.net/publication/43130818_Standardized_Screening_for_Suicidal_Adolescents_in_Primary_Care

Stewart, J. G., Shields, G. S., Esposito, E. C., Cosby, E. A., Allen, N. B., Slavich, G. M., & Auerbach, R. P. (2019). Life Stress and Suicide in Adolescents. *Journal of Abnormal Child Psychology*, 47(10), 1707–1722.

<https://doi.org/10.1007/s10802-019-00534-5>

Stewart, S. M., Kennard, B. D., Lee, P. W. H., Mayes, T., Hughes, C., & Emslie, G. (2005). Hopelessness and suicidal ideation among adolescents in two cultures. *Journal of Child Psychology and Psychiatry*, 46(4), 364–372.

<https://doi.org/10.1111/j.1469-7610.2004.00364.x>

Suicidal ideation. (2020). In Wikipedia.

https://en.wikipedia.org/w/index.php?title=Suicidal_ideation&oldid=962443

135

Suicide among teens and young adults reaches highest level since 2000. (2019,

June 18). PBS NewsHour. [https://www.pbs.org/newshour/nation/suicide-](https://www.pbs.org/newshour/nation/suicide-among-teens-and-young-adults-reaches-highest-level-since-2000)

[among-teens-and-young-adults-reaches-highest-level-since-2000](https://www.pbs.org/newshour/nation/suicide-among-teens-and-young-adults-reaches-highest-level-since-2000)

Suicide in antiquity. (2020). In Wikipedia.

[https://en.wikipedia.org/w/index.php?title=Suicide_in_antiquity&oldid=960](https://en.wikipedia.org/w/index.php?title=Suicide_in_antiquity&oldid=960791595)

[791595](https://en.wikipedia.org/w/index.php?title=Suicide_in_antiquity&oldid=960791595)

Suicide in the United States. (2020). In Wikipedia.

[https://en.wikipedia.org/w/index.php?title=Suicide_in_the_United_States&](https://en.wikipedia.org/w/index.php?title=Suicide_in_the_United_States&oldid=963963320)

[oldid=963963320](https://en.wikipedia.org/w/index.php?title=Suicide_in_the_United_States&oldid=963963320)

Suicide Trends Among and Within Urbanization Levels by Sex, Race/Ethnicity,

Age Group, and Mechanism of Death—United States, 2001–2015 |

Semantic Scholar. (n.d.). Retrieved June 24, 2020, from

[https://www.semanticscholar.org/paper/Suicide-Trends-Among-and-](https://www.semanticscholar.org/paper/Suicide-Trends-Among-and-Within-Urbanization-Levels-Ivey-Stephenson-Crosby/5ade1902e9dbc0a0e3f3a8cfe39022090ea541ad#paper-header)

[Within-Urbanization-Levels-Ivey-Stephenson-](https://www.semanticscholar.org/paper/Suicide-Trends-Among-and-Within-Urbanization-Levels-Ivey-Stephenson-Crosby/5ade1902e9dbc0a0e3f3a8cfe39022090ea541ad#paper-header)

[Crosby/5ade1902e9dbc0a0e3f3a8cfe39022090ea541ad#paper-header](https://www.semanticscholar.org/paper/Suicide-Trends-Among-and-Within-Urbanization-Levels-Ivey-Stephenson-Crosby/5ade1902e9dbc0a0e3f3a8cfe39022090ea541ad#paper-header)

WHO | Suicide data. (n.d.). WHO; World Health Organization. Retrieved June 26,

2020, from

http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/

Wintersteen, M. B., Diamond, G. S., & Fein, J. A. (2007). Screening for suicide risk in the pediatric emergency and acute care setting. *Current Opinion in Pediatrics*, 19(4), 398–404.

<https://doi.org/10.1097/MOP.0b013e328220e997>

Wong, M. M., & Brower, K. J. (2012). The prospective relationship between sleep problems and suicidal behavior in the National Longitudinal Study of Adolescent Health. *Journal of Psychiatric Research*, 46(7), 953–959.

<https://doi.org/10.1016/j.jpsychires.2012.04.008>