Suicidal Behaviors Among Adolescents

Grace Sigalla
Drexel University
gas83@drexel.edu

Kunal Chhabria
Drexel University
kc3575@drexel.edu

Nick Babcock
Drexel University
nb987@drexel.edu

Sreerekha Rajendran
Drexel University
sr3547@drexel.edu

***Abstract*—Suicide is one of the leading concerns of health around the world, especially adolescents. Although there has been progress in the field of suicide prevention, little is known about how to recognize teenagers who may be at risk of such thoughts. This project aims to analyze and understand several factors among teenagers that lead to such thoughts.**

We aim to do this using the dataset created by the GSHS-Global School-based Student Health Survey of adolescents based on Gender from 27 countries. It is divided into ten main modules that cover the world's top causes of morbidity and death in both children and adults. Some of the modules are alcohol use, dietary behaviors, drug use, tobacco use, etc.

Keywords: suicidal attempt, adolescents, alcohol, drugs, survey

I. INTRODUCTION

For children and adolescents all across the world, suicide is a serious public health issue. Between the 1960s and 1990s, the suicide rate in the US doubled for those aged 15 to 19 and tripled for those aged 10 to 14. Although it is not just due to more reporting, the causes of this trend are not obvious. Increased rates of drug and alcohol misuse, anxiety, social and familial instability are a few factors that have been proposed.

Between 1975 and 2015, the most recent year included in the many studies, there were around 100,000 teen suicide fatalities (15–19 years old). Three to four times as many boys as females commit suicide. Teenage boy suicide rates increased, reached their high in the late 1980s and early 1990s, then began to decline in the late 1990s and early 2000s. This might be linked to pediatricians treating depression more frequently. Unfortunately, the rate has been increasing once more since 2006. Teenage females' suicide rates followed a comparable but less striking pattern. Over the past 10–12 years, there has been a rise in the number of girls[3]. It's important to remember that the suicide rates across all age groups change over time. Comparing youth rates to every older age group, they continue to be significantly lower.

For this project, we are focusing on countries with lesser population and that has adolescents who are more exposed and prone to some of the factors mentioned above. The dataset created by Global School-based Student Health Survey of adolescents helped us concentrate on 26 such countries. The Global School-based Student Health Survey is intended to give reliable information on health behaviors and social support among students to: - Helping nations establish priorities, implement programs, and advocate for funding school health and youth health policies and programs. - Permit international organizations, nations, and others to compare the prevalence of protective factors and health behaviors across nations. - For use in evaluating school health and youth health promotion, identify changes in the prevalence of protective factors and healthy behaviors by nation.

The GSHS is a school-based survey that is mostly done among adolescents between the ages of 13 and 17 and uses a self-administered questionnaire that may be finished in one regular class hour as well as a common school-based methodology. Using the standardized core questions as a starting point, countries can then add country-specific questions to create their own country-specific questionnaire. The questions are translated into the students' suitable instructional language. To improve the survey's flow and the student's understanding, all of the questions have elements in common. No skip patterns are permitted in order to preserve student privacy[2].

II. DATASET DESCRIPTION

```
Data columns (total 17 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Country                           106 non-null    object
 1   Year                              106 non-null    int64
 2   Age Group                         106 non-null    object
 3   Sex                               106 non-null    object
 4   Currently_Drink_Alcohol           106 non-null    float64
 5   Really_Get_Drunk                  106 non-null    float64
 6   Overwieght                        106 non-null    float64
 7   Use_Marijuana                     106 non-null    float64
 8   Have_Understanding_Parents        106 non-null    float64
 9   Missed_classes_without_permssion  106 non-null    float64
 10  Had_sexual_relation               106 non-null    float64
 11  Smoke_cig_currently               104 non-null    float64
 12  Had_fights                        106 non-null    float64
 13  Bullied                           102 non-null    float64
 14  Got_Seriously_injured             106 non-null    float64
 15  No_close_friends                  106 non-null    float64
 16  Attempted_suicide                 106 non-null    float64
dtypes: float64(13), int64(1), object(3)
memory usage: 14.2+ KB
```

Fig. 1. Dataset Description

The dataset we decided to use is based off a culmination of surveys which were conducted among high school students between the years of 2010 and 2018. The Global School-based Student Health Survey (GSHS) is a school-based survey aimed

to collect data about young people's health behaviors and is funded by the World Health Organization (WHO) and the US Centers for Disease Control and Prevention (CDC). They take place in various countries from around the world and contain a plethora of questions regarding many different aspects of life such as diet, drug use, hygiene, and metal health.

While this data is publicly available in WHO's NCD Microdata Repository, the dataset we were able to obtain was provided from a source on Kaggle.com. We chose to use this dataset because it was a more convenient source in terms of gaining access to the actual data. To obtain the data in a workable form from WHO's website can be a complicated task whereas the dataset from Kaggle is simply downloaded.

The dataset has 17 columns which gives us information about which country it is, gender of the student, their age and percentages of students who are involved in multiple factors that could lead to suicidal thoughts including use of marijuana, consumption of alcohol, not having understanding parents, being bullied and being overweight.

A. Notable features in the dataset

For our analysis, we are looking at:
- 27 countries
- 9 years (2010 – 2018)
- 2 age groups: 13-15 and 16-17

III. RELATED WORK

Approximately 1 in 5 high school students reported seriously considering suicide in the previous year, according to the Youth Risk Behavior Survey from 2019. In addition, 1 in 11 high school students reported making a suicide attempt in the previous year, and about 2.5% sought medical assistance after attempting suicide. We are learning more about teen suicide thanks to research.

Although attempts at suicide are much more prevalent than suicide fatalities, any suicidal conduct is severe and can have long-lasting effects. The majority of suicidal teenagers do not truly wish to die. They are experiencing excruciating emotional anguish and want it to end. They can be unclear about how to go and whether there are any strategies to feel better.

All studies and research indicate that suicide is mainly avoidable. Early detection, prompt treatment, and support from communities, families, schools, and friends all play important roles.

IV. EXPLORING DATA

A. Data Cleaning

We had minor issues with respect to our data and hence did not have to spend much time in the data pre-processing section. The following was done to clean the data:

- Replaced null values in columns "smoke_cig_currently" and "bullied with their mean values.
- Labeled the two age groups "13-15" and "16-17" as 0 and 1, to make it easier when conducting the analysis.

```
df['Bullied'] = df['Bullied'].fillna(df['Bullied'].mean())
df['Smoke_cig_currently'] = df['Smoke_cig_currently'].fillna(df['Smoke_cig_currently'].mean())
```

Fig. 2. Replacing null values with mean

B. Exploratory Data Analysis

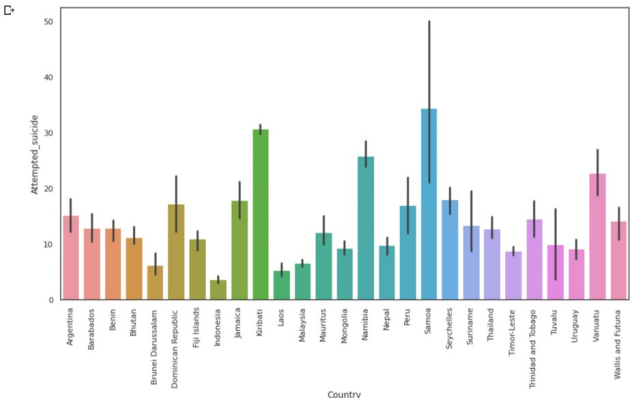

Fig. 3. Suicide Rates Comparison by Countries

We see that the 4 main countries with attempted suicide mean over 20 being Samoa with 34, followed by Kiribati with 31, then Namibia at Almost 26 and Vanuatu at almost 23. Then the lowest countries being Indonesia with 3.6, followed by Laos with 5.

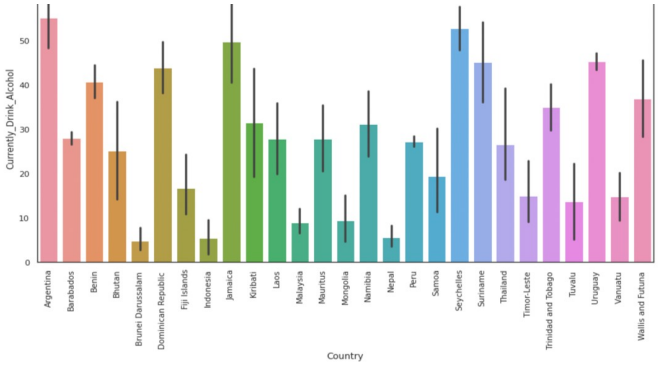

Fig. 4. Alcohol Consumption For Each Country

It is very evident from the graph that adolescents from Argentina, Seychelles and Jamaica indulged in alcohol more than any other country. However, from the previous graph, of the 4 countries with the highest attempted suicide rate, only Kiribati and Namibia seem to have students who consume over 30% alcohol.

And then we did some analysis to see how both the age groups are affected by certain factors over the years. We were

able to see different trends. Which year did students aged 16-17 were more involved in the use of marijuana, which year saw a lower usage of alcohol among students aged 13-15 and so on.

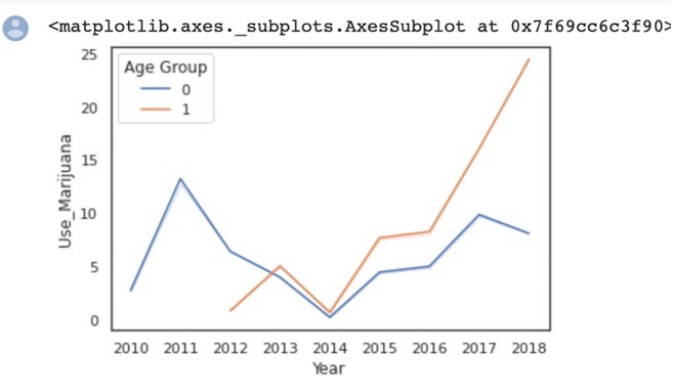

Fig. 5. Marijuana Consumption For Each Age Group

When we look at the use of marijuana among the adolescents, we can see a clear trend in teens of age group 16-17 indulging in marijuana from year 2012, which shoots up from 2016. As opposed to the age group of 13-15 which from 2010 had a higher increase in marijuana usage, which seemed to have reached a peak in 2011, before starting to decrease and then gain momentum from 2015 but at a slower rate compared to the age group of 16-17.

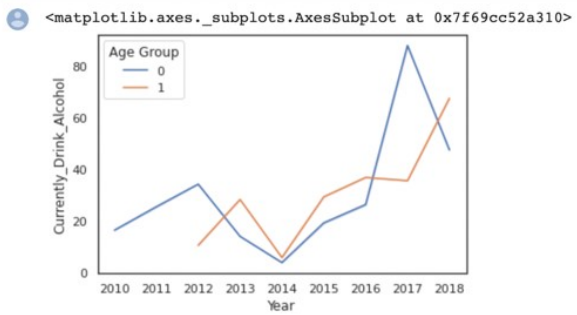

Fig. 6. Alcohol Consumption For Each Age Group

It was unexpected to see a higher proportion of adolescents in age group 13-15 indulging in the use of alcohol compared to adolescents aged 16-17.

Teens' use of alcohol, marijuana, and thoughts of or actual attempts at suicide pose serious, independent but interrelated public health issues. Studies showed that drinking alcohol or using marijuana within three months of being admitted to a mental health facility increased the likelihood that a suicidal attempt would be made on the day of admission as well as throughout the course of time. Using retrospective longitudinal data, researchers also discovered that marijuana use—but not alcohol use—increased a teen's likelihood of supporting suicidal thoughts both immediately and over time[3].

Even though researchers have published on this subject since the 1970s, there hasn't been much focus in the public health discourse on the effect of teenage marijuana use on the risk of developing depressive symptoms and mood disorders.

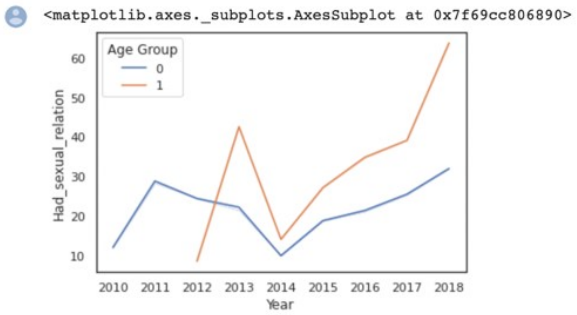

Fig. 7. Sexual Activity Compared For Each Age group

Age group 16-17 had a higher rate of involvement in sexual relations as the years progressed from 2012, and with more increments from 2014, which is really not as surprising for the age group.

We also conducted some bivariate analysis and diagnostic analysis to further strengthen our conclusions. Bivariate analysis - is one of the simplest forms of quantitative analysis. It involves the analysis of two variables, for the purpose of determining the empirical relationship between them. Bivariate analysis can be helpful in testing simple hypotheses of association. Diagnostic analytics - is a form of advanced analytics that examines data or content to answer the question, "Why did it happen?". Since we have numerical data that is highly affecting our class label AttemptedSuicide and our few categorical data, we will dive deeper into the Root Cause Analysis . Example, examining do overweight people get bullied more often based on country? OR Based on gender did they commit suicide if they had no friends?

Some of the conclusions we were able to make from the above mentioned analysis were:

- Those who had a higher rate of missed classes and got bullied a lot, had a higher chance of attempting suicide.
- Those adolescents who got bullied the most and didn't have close friends, had a higher chance of attempting suicide.
- If someone is overweight and gets bullied more, they have a higher chance of attempting suicide. However, it is also true if a person if bullied a lot but not very overweight, they could also attempt suicide.
- Those who got really drunk and got into fights, had a higher chance of attempting suicide. Also, those who got into

a lot more fights, had higher chances of attempting suicide.

- It is observed that those adolescents who got into more fights and didn't have understanding parents, had a higher chance of attempting suicide.

Our next aim was to isolate those surveys which had a higher attempted suicide rates. We focused on surveys with more than 20% attempted suicides. This helped us take a closer look on what factors affected the teenagers more and to narrow down factors to come to a solid conclusion. Below are some results which showed behaviors that higly contributed to suicidal attempts in teenagers.

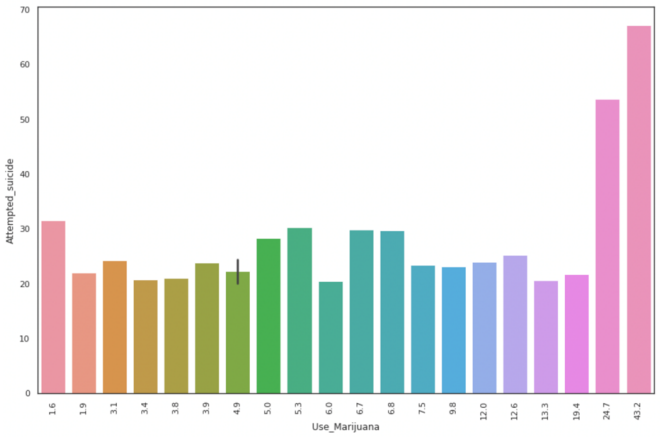

Fig. 8. Marijuana Usage

Those with suicide rate over 20% showed that for the survey that showed a higher percentage of marijuana usage (20% and above) had double the chances of attempted suicide.

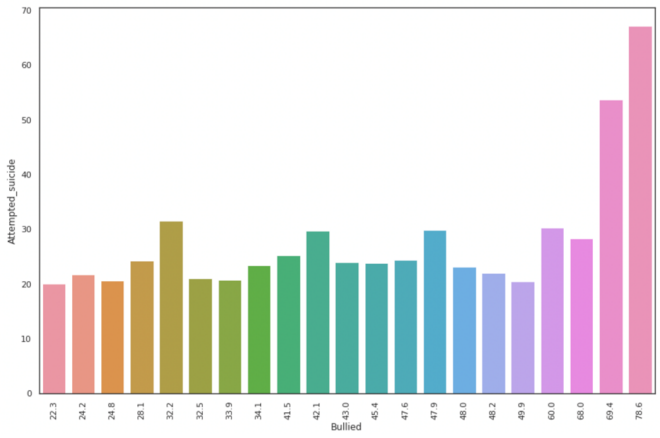

Fig. 9. Got Bullied

Those with a bullied rate of 70% and higher had attempted suicide rate by almost double compared to those with less.

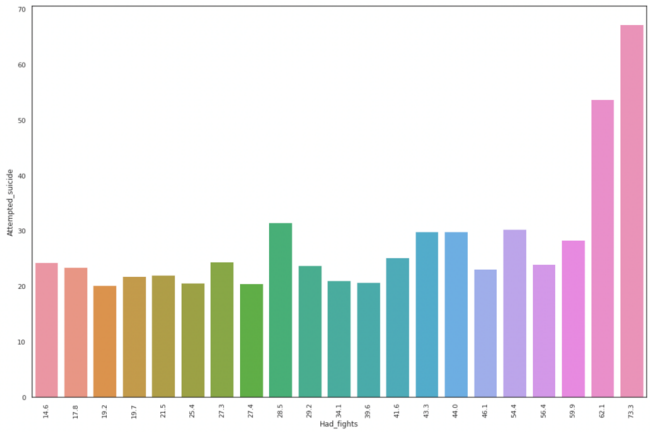

Fig. 10. Got into Fights

Those with fights rate of 60% and higher had attempted suicide rate more that 2 times that of surveys with less.

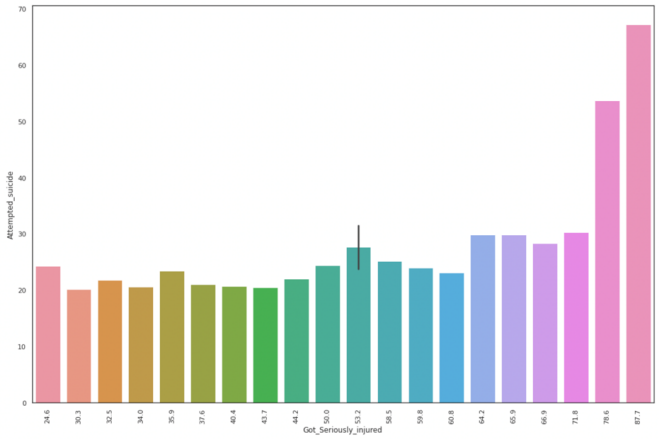

Fig. 11. Seriously Injured

We see a slight increase of those surveys which indicated serious injury rate of 64-75%, and then a sharp increase from 78% and above of injury, with a significant increase in attempted suicide rate by over 20%.

V. RESULTS

Based on the surveys analysis conducted, we determined that while there are more females overall in the surveys collected, there were in fact more males who attempted suicide compared to females for those surveys with a higher attempted suicide rate. Moreover, of the 2 ages groups, that of 13-15 reported the most suicide attempts. Countries with the highest attempted suicide rates are Samoa, Kiribati and Namibia.

Below are the behaviors that have the highest contribution to attempted suicide:

- Marijuana usage

- Getting into fights
- Being bullied
- Getting seriously injured

VI. CONCLUSIONS

Every suicide fatality in adolescence is tragic. An important objective is to be able to anticipate suicide attempts and maybe prevent them. Through this project, we have observed how several macro environmental elements affect adolescents. The application of advanced modeling approaches in the prediction of suicidal actions and thoughts might prove to be vital in addressing the many small elements that contribute to suicidal behavior. Machine learning algorithms have the potential to be an effective tool for reflecting or condensing the complex processes, which include a variety of components and numerous phases that lead to suicide.

VII. FUTURE WORK

The issue of accurately predicting the risk of suicide among adolescents and children within a manageable time frame is crucial yet difficult. Only very few studies have comprehensively considered the clinical risk factors available to produce risk scores for estimation of short- and long-term suicide risk for pediatric population. Based on the findings of their surveys, we hope to employ machine learning techniques to evaluate how well they predict suicidal behavior in children and adolescents. Logistic Regression, K-NN Classifier, Decision Tree Classifier, Random Forest, and Stacking are a few machine learning methods that might be employed.

VIII. APPENDIX

All the code for this project are uploaded to GitHub repository: https://github.com/GraceSigalla/DSCI-591-SNKG.git

REFERENCES

[1] CDC, "FastStats - Adolescent Health," cdc.gov, 2019. https://www.cdc.gov/nchs/fastats/adolescent-health.htm

[2] "Global school-based student health survey." https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-school-based-student-health-survey (accessed: Dec. 06, 2022).

[3] "Alcohol and marijuana use as daily predictors of suicide ideation .." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9262037/ (accessed: Dec. 06, 2022).