AN ANALYSIS OF FACTORS THAT INFLUENCE SUICIDE

Dan Zhao 4809602529

INTRODUCTION

In recent years, the suicide rate in America has been gradually increasing. In 2001, the suicide rate was 10.7%, while in 2013, it increased to 13%, with 41,149 deaths and over $44 billion in economic loss.²

Recent studies have found that there are two categories of factors that contribute to suicide. They are environmental factors and biological factors. Individuals prone to suicide are subject to a combination of both types of factors.² Environmental factors include social status (poverty, unemployment, disability, etc.),³ group characteristics that may face discrimination (race, gender, sexual orientation, etc.),³ abuse, emotional support from family or community and so on. Biological factors include family history, mental health status,⁴ and chemical reactions (drug use, alcohol abuse, etc.).⁵

In this project, I gathered data for most of the above factors and analyzed their possible influence on suicide rate in each county of California.
DATA SOURCES

The Shapefile of counties in California were obtained from the U.S. Census, 2010 TIGER/Line® Shapefilesvii.

Data for suicide, unemployment, poverty, race, gender, major depression, drug use, uninsured and disability were obtained from Community Health Status Indicators (CHSI) to Combat Obesity, Heart Disease, and Cancer Dataset of Centers for Disease Control and Prevention (CDC)viii.

To get data for gender, age, mean household income and marriage rate, I used American Community Survey (ACS) from American Fact Finder. I downloaded the AGE AND SEX 2009-2013 American Community Survey 5-Year Estimatesix and MARITAL STATUS 2009-2013 American Community Survey 5-Year Estimatesx.

CHALLENGES

I had difficulty in finding data for some important factors like mental health diseases, family history and so on. So I only analyzed the 11 factors that I could get access to the data.

Another challenge was that I could not use the Ordinary Least Squares tool in ArcMap. I generated integer Unique ID column as it required, but I still got Error 00748: field does not exist within table. I also tried changing the field names, but it still wouldn’t work. At last I gave up and used Stata instead.
RESULTS

Poverty

The picture shows California’s suicide rate and poverty rate for each county. Counties with dark green are the ones that have highest suicide rates. However, as we can see, the counties with higher poverty rate are not the ones with higher suicide rate, which indicates there may not be a positive relationship between the suicide rate and poverty.
Age

The pie charts show the breakdown of age groups in each county. According to the literature I read, I supposed teenagers and elderly people are more likely to commit suicide. In the picture, we can see in the counties with lower suicide rates, the age group of 19 to 64 seems slightly greater than in the counties with higher suicide rates. It means states with larger amount of teenagers and elderly people may result in higher suicide rates.
Some literature suggests that people who are White have higher suicide rates. In this picture, though we cannot see the ratio of Blacks and Native Americans clearly, we can compare White and Hispanic people in most counties. It seems counties with higher suicide rates do have more White people relative to other races, which shows that being White is associated with an increase in the possibility of suicide.
In most counties, there are no big differences between the population of males and females, so it’s difficult to see the impact of gender on suicide in this picture. We will use OLS regression to test it later. For now, we can see that a few counties that obviously have greater number of males do have higher rates of suicide (while one of them lacks suicide data), though not the highest. It seems male are more likely to commit suicide.
Marriage

Married people are usually supposed to be more responsible and less likely to commit suicide. However, in this picture, it is difficult to find a negative relationship between the marriage rate and the suicide rate.
This picture shows the unemployment rate and the suicide rate of each county in California. As we can see, counties with higher suicide rates are not those with higher unemployment rate. So we cannot find positive relationship between suicide and unemployment in this picture.
This picture shows the rate of people with major depression and the suicide rate in each county in California. And we can clearly see the counties with higher depression rates tend to have higher suicide rates, which indicates a positive relationship between major depression and suicide.
Drug Use

This picture is about reported recent drug use and suicide. And it indicates that counties with a higher drug use rate tend to have a higher suicide rate, which seems reasonable according to our common knowledge.
Some literature reflects that people with disability have a higher likelihood to commit suicide. However, I could not get the data for disabled people in each county, so I used the data for disabled people with Medicare instead. And we can see that counties with higher disabled Medicare rates also have higher suicide rates.
Uninsured

This picture shows rate of uninsured people and the suicide rate in each county in California. I supposed uninsured people might have higher suicide rate since they may not be able to get access to a health center when they get sick. However, the picture didn’t show an obvious relationship between the uninsured rate and the suicide rate.
It is believed that lack of sunshine can result in depression, thus may increase suicide rates. Generally speaking, California is a dry, low latitude state with enough sunshine, but there are still differences among each county. And the comparison of the drought and the suicide rate does show some interesting results. In the drought picture, we can see the northwest and southeast get more rain, while the middle part of the state is extremely dry. The suicide rate picture also shows an expected result: higher suicide rate in northwest and southeast, and lower in the middle. In addition, most counties with the highest suicide rates are in the north, which is high latitude with a shorter time of sunshine.
**OLS Regression**

Pictures could only show rough results about what might influence suicide rate, so I planned to use Ordinary Least Squares Regression to test the statistical results. Since I could not use OLS in ArcMap as I explained, I used Stata to do the regression. In the previous part, I only used data in California, which was quite a small sample. So in the OLS regression, I used data from all the counties in the United States to get a more reliable result. I dropped about 500 counties that have missing data. The result is as below:

```
. regress Suicide Poverty Age19_Under Age19_64 Age65_84 Black NativeAmerican Asian Hispanic male Meanhouseholdincome MarriageRate UnemployedRate MajorDepressionRate DrugUseRate DisabledRate UninsuredRate, vce(robust)
```

```
Linear regression                               Number of obs     =      2,610
F(16, 2593)       =      50.45
Prob > F          =     0.0000
R-squared         =     0.3553
Root MSE          =     4.2951

|               Robust   | Coef.   | Std. Err. | t     | P>|t| |  [95% Conf. Interval] |
|---------------------|---------|-----------|------|-----|-----------------------|
| Suicide             |         |           |      |     |                       |
| Poverty             | -.267502 | .0613707  | -4.36| 0.000| -.3878424 - .1471614  |
| Age19_Under         | 1.209599 | .2158378  | 5.60 | 0.000| .7863675  1.632831    |
| Age19_64            | 1.337216 | .190088   | 7.03 | 0.000| .964477   1.709956    |
| Age65_84            | 1.711659 | .2426636  | 7.05 | 0.000| 1.235825  2.187493    |
| Black               | -.053191 | .0105348  | -5.05| 0.000| -.073845  -.0325334   |
| NativeAmerican      | .2832262 | .0521274  | 5.43 | 0.000| .1810106  .3854418    |
| Asian               | -.0856678| .028143   | -3.04| 0.002| -.1408528 -.0304829   |
| Hispanic            | -.0202377| .0153567  | -1.32| 0.188| -.0503504 .0098749    |
| male                | 30.09976 | 5.682072  | 5.30 | 0.000| 18.9579  41.24162     |
| Meanhouseholdincome | -.0000386| .000109   | -3.54| 0.000| -.00006  -.0000172    |
| MarriageRate        | .0327221 | 2.425608  | 0.01 | 0.989| -4.723602 4.789046    |
```
The result shows that about 35% of the suicide rate can be explained by the variables I used. The factors of being under 85 years of age, Native American, male, married, having depression, disabled, and uninsured are positively related with suicide, while the marriage rate is not significant. And other variables not mentioned are negatively related with suicide, while being of the Hispanic race and drug use are insignificant. The result is similar with what I got from the previous pictures.

However, there are obvious limitations like omitted related variables and correlation between some variables (poverty, unemployed, uninsured, etc.).

**CONCLUSION&RECOMMENDATIONS:**

Given the limitations of my model, I could not determine with certainty which are the most influential factors of suicide in a given county in California or the US.

However, the variables for gender, major depression, and uninsured rate not only positively associated with suicide as expected, but also show high levels of significance and influence on the suicide rate. According to the high statistical significance of male,
major depression and uninsured rate, I recommend that County Departments of Health Services and Social Services consider males, people with depression and uninsured individuals to be more at risk for suicide. And these three factors seem to be correlated with each other. Males may prone to higher levels of pressure due to stress at work and “discrimination” when seeking help for mental stress, since they are usually supposed to be brave and strong. And it will be more difficult for males without insurance to get mental health care. Without proper health care, their mental stress might become more serious and finally fall into serious depression.

So I recommend we first focus on males. Perhaps create some outreach programs directed to males prone to suicide. Department of Health Services and communities should also notice that sometimes males might need more mental care than females and should focus on them more. Also, health care centers in counties could increase mental health treatments to both insured and uninsured people.

---

i Centers for Disease Control and Prevention (CDC)


vii U.S. Census, 2010 TIGER/Line® Shapefiles http://www.census.gov/cgi-bin/geo/shapefiles2010/main

viii Centers for Disease Control and Prevention, “Community Health Status Indicators (CHSI) to Combat Obesity, Heart Disease, and Cancer” (2010)

ix AGE AND SEX 2009-2013 American Community Survey 5-Year Estimates
http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_5YR_S0101&prodType=table

x ARITAL STATUS 2009-2013 American Community Survey 5-Year Estimates
http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_5YR_S1201&prodType=table