

Final Project Step 1

Alan Donahue

8/1/2021

Introduction

The topic that I have decided to choose is about military suicides. This has been a hot topic in the news for a few years now because of movements like Mission 22. The news has also reported how military suicide has been on the rise since 2016. I want to focus on both suicide incidents that happen to those those who are now veterans.

The military would be interested in the topic because they are trying to figure out a way to prevent military suicides from rising. Family and friends of military personnel and military personnel themselves would also be interested in the topic because it directly affects their lives.

This is a data science problem because there is a massive amount of information that is being collected from military suicides. This data can be used to help find a trend for the most vulnerable individuals. The trend can be used to identify those folks how are most vulnerable to committing suicide so that the proper folks reach out to them to help prevent suicide.

Research Questions

- 1) Does the trend in military suicides follow the same trend as the U.S. national suicide trend? If it doesn't, what is the reason for not having the same trend?
- 2) Does the current data suggest that those veterans who seek help are less likely to commit suicide?
- 3) Is there a trend in the data to show when a veteran is more vulnerable to the idea of committing suicide?
- 4) Does the location of a veteran within the U.S. have an affect on the rate of suicide?
- 5) Is there a difference in how veterans commit suicide versus the general population of the U.S.?

Approach

I plan to address the problem statement by finding a trend in the data sets to determine when a veteran is most vulnerable to the idea of suicide. The goal is to use a regression model to find the trend/predictors.

How your approach addresses (fully or partially) the problem

By discovering the predictors/trend of when a veteran is most vulnerable to the idea of suicide, the federal government can ensure they are allocating the proper funds to those veterans. Having the proper funds and help to the most vulnerable veterans will help reduce the suicide rate.

Data

Veteran Suicides By State of Death and by Sex, 2005-2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 5 Peculiarities of data: Some of the number of suicides column has ranges or says <10 instead of an exact number.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

Veteran and General Adult Population (Veteran and Non-Veteran) Suicides by State of Death and Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 7 Peculiarities of data: some of the number of suicides has ranges or says <10 instead of an exact number. Blank cells are represented by —.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

Veteran and Adult General Population (Veteran and Non-Veteran) Suicides by State of Death and Method, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 7 Peculiarities of data: Some of the numbers in the Excel say they are formatted as text instead of the rest which are as a number.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

Veteran Suicides by U.S. Territory of Death, Age Group, and Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 5 Peculiarities of data: Some of the numbers in the Excel say they are formatted as text instead of the rest which are as a number. Some numbers have a * by them and the blanks are represented by —.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Veteran Deaths by Suicide: Crude and Adjusted Suicide Rates by Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 14 Peculiarities of data: No issues.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Veteran Deaths by Suicide: Crude Suicide Rates by Sex and Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 12 Peculiarities of data: Blank cells are represented by “.”.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Non-Veteran Deaths by Suicide: Crude and Adjusted Suicide Rates by Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 14 Peculiarities of data: No issues.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Non-Veteran Deaths by Suicide: Crude Suicide Rates by Sex and Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 12 Peculiarities of data: Blank cells are represented by “.”.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Recent Veteran VHA User Deaths by Suicide: Crude and Adjusted Suicide Rates by Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 14 Peculiarities of data: No issues

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Recent Veteran VHA User Deaths by Suicide: Crude Suicide Rates by Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 5 Peculiarities of data: No issues

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Non-Recent Veteran VHA User Deaths by Suicide: Crude and Adjusted Suicide Rates by Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 14 Peculiarities of data: No issues.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All Non-Recent Veteran VHA User Deaths by Suicide: Crude Suicide Rates by Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 5 Peculiarities of data: No issues

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All U.S. Adult Population Deaths by Suicide: Crude and Adjusted Suicide Rates by Sex, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 14 Peculiarities of data: No issues

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

All U.S. Adult Population Deaths by Suicide: Crude Suicide Rates by Sex and Age Group, 2005–2018

Original Purpose: To track veteran suicides from 2005-2018 to be included in the annual report by the U.S. Department of Veteran Affairs When was it collected: 2005-2018 for 2020 release How many Variables: 12 Peculiarities of data: Blank cells are represented by “.”.

Mental Health. (2020). Retrieved from U.S. Department of Veteran Affairs: [link](#)

Underlying Cause of Death, 1999-2019

Original Purpose: To collect data on suicide rates in the U.S. When was it collected: 1999-2019, release for 2020 How many Variables: 5 Peculiarities of data: No issues.

Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death" "1999-2019 on CDC WONDER Online Database, released in 2020. Data are from the Multiple Cause of Death Files, 1999-2019, as" "compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed" "at link on Aug 1, 2021 5:59:36 PM"

Required Packages

- ggplot2
- car
- pastecs
- boot
- Hmisc
- readxl
- qqplotr
- plyr
- dplyr
- purrr
- stringr

Plots and Tables Needed

- Histogram
- Error bar
- Scatterplot
- Boxplots
- Line graph

Questions for Future Steps

At this point, I believe I have more than enough data to answer my topic. As far as technical concerns, I might not know how to conduct a more complex transforming of the data. I believe I can learn that by experiencing and messing around with the data.