

Can Macro Data Predict Suicide in Adolescents?

Appendix

Anita Xia
Dana Hall School
Wellesley, MA 02482
anita.xia@danahall.org

1 Data Details

Data processing. The suicide and macro data cover most of the states in the U.S., but there are still some missing data, especially in suicide indicators because not all states participate in the survey and not all questions are asked in every survey year. We remove the states that have at least 1 missing suicide indicator or macro data for all years and get a dataset with data from 45 states: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KS, KY, LA, ME, MD, MA, MI, MS, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OH, OK, PA, RI, SC, SD, TN, TX, UT, VT, VA, WV, WI, WY. To align macro data with suicide data, we keep the data every other year and drop the rest. Then we turn GDP into GDP per Capita (USD) by dividing GDP by population to focus more on individual-level data. Finally, we normalize all macro data to range from 0 to 1.

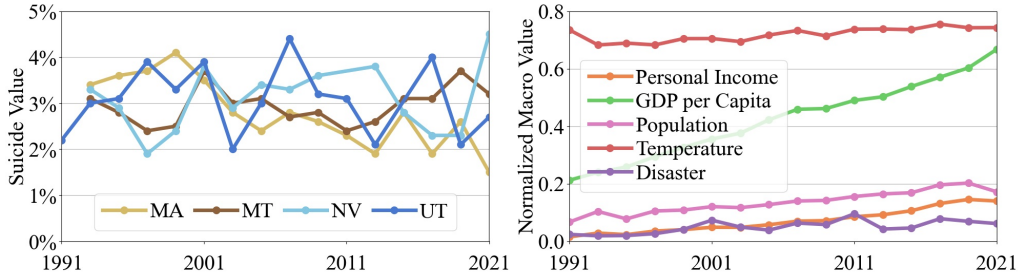


Figure 1: Data visualization. A. Suicide indicator IV of four states with the most comprehensive survey data; B. Processed macro data averaged across 45 states.

Table 1: Data statistics of suicide and processed macro data

Data	Min	Max	Avg.	Std.
Suicide Indicator I	10.40%	30.70%	17.83%	3.88%
Suicide Indicator II	8.10%	26.10%	14.31%	3.03%
Suicide Indicator III	3.60%	17.60%	8.87%	2.04%
Suicide Indicator IV	1.00%	9.70%	3.03%	1.01%
Personal Income	0.0017	1.0000	0.0654	0.0932
GDP per Capita	0.0990	1.0000	0.3720	0.1748
Population	0.0115	1.0000	0.1477	0.1595
Temperature	0.5007	1.0000	0.7144	0.1058
Disaster	0.0000	1.0000	0.0441	0.0748

Visualization of statistics. Figure 1 shows the Suicide indicator IV of four states with the most comprehensive survey data and processed macro data from 1991 to 2021 averaged across 45 states.

For Suicide data, we can see that suicide rate keeps changing and is different among states. For macro data, GDP per Capita keeps increasing over all 30 years. Personal Income and Population both increase in most years, but drop down from 2019 to 2021, which may result from the influence of Covid 19. Temperature and Disaster increase in fluctuations, and it may be due to global warming that we are experiencing more extreme weather patterns. Detailed statistics can be found in Table 1.

2 Limitations and Broader Impacts

In this paper, we employ mature machine learning techniques for early prediction of suicide in Adolescents. The limitations of our study might come from two aspects. The first one is that we employ the macro data, five features that we uncovered from the literature that have correlations with depression and suicide. There might have other types of macro data that can be involved to further enhance predictive accuracy. The second is that other advanced time-series models can produce better performance.

Predicting adolescent suicide holds significant promise for improving mental health outcomes by enabling early intervention and personalized support. By identifying at-risk individuals before symptoms escalate, predictive models can facilitate timely access to resources and interventions, potentially reducing the loss of lives. Our study not only supports better mental health outcomes but also promotes a shift towards proactive mental health care, reducing stigma and normalizing discussions around mental well-being among adolescents.

3 Code and Reproducibility

We provide our code, instructions, and implementation in an open-source repository: <https://github.com/straeCS/SuicideModel.git>. The experiments were conducted on a Windows 11 (Version 10.0.22631) laptop using Intel Core i7-8850U processor with 16 GB RAM.