

# Analisi esplorativa del tasso di suicidi in diversi paesi tra il 2013 e il 2019

Tesina di Data Science for Health Systems

Paolo Speziali

*Dipartimento di Ingegneria  
Università degli Studi di Perugia  
Perugia, Italia  
paolo.speziali@studenti.unipg.it*

**Abstract**—Questo documento riporta i risultati dell’analisi esplorativa del dataset fornito dalla World Health Organization relativo al tasso di mortalità per suicidio standardizzato per età. L’obiettivo di questa analisi è quello di esaminare se, in differenti paesi tra il 2013 e il 2019, la distribuzione del tasso di suicidi sia significativamente differente o meno, aprendo la strada a potenziali ipotesi su fenomeni globali che possano averne influenzato la tendenza. Tramite l’analisi di tre gruppi distinti è possibile osservare come una distribuzione comune, in alcuni casi, emerga.

**Index Terms**—suicide, EDA, WHO, GDP

## I. INTRODUZIONE

Il suicidio è definito come l’atto di terminare intenzionalmente la propria vita [1]. Come accade da decenni, il suicidio rimane una delle cause più frequenti di morte nel mondo occidentale [2]. Come osserva il Dr. Peeter Värnik dell’*Estonian-Swedish Mental Health and Suicidology Institute* nel suo paper “Suicide in the World” [3]: “Il suicidio è un atto individuale ma, una volta che i dati sono aggregati a livello di paese, le variazioni da un anno all’altro sono piuttosto ridotte e di solito non ci sono grandi fluttuazioni. Si è tentati di affermare che il miglior predittore del tasso di suicidi nel breve periodo è il tasso di suicidi del passato. Tuttavia, nel lungo periodo possono verificarsi e si sono verificati grandi cambiamenti”. Nello studio che segue andremo ad osservare i cambiamenti nel tasso di suicidio a livello mondiale tra il 2013 e il 2019 e verificare se è possibile affermare che non esistono differenze significative nella distribuzione tassi di suicidio tra paesi differenti.

## II. DATASET

### A. Descrizione del Dataset

Il dataset utilizzato proviene dal portale web della World Health Organization (WHO) ed ha come titolo “Age-standardized suicide rates (per 100000 population)” [4]. Contiene dati riguardanti i tassi di suicidi in ogni paese dal 2000 al 2019 di soggetti maschili, femminili e di entrambi i sessi. Tali dati sono standardizzati prendendo in considerazione la distribuzione dell’età della popolazione del paese, è stato scelto questo dataset rispetto a “Crude suicide rates (per 100000

population)” in quanto la standardizzazione rispetto all’età dà luogo a statistiche più rappresentative quando l’evento preso in considerazione non è raro [5]. I dati sono stati acquisiti attraverso il processo di censimento della popolazione, registrazioni anagrafiche e l’analisi di certificati medici contenenti informazioni sulla causa dei decessi. Il dataset originale si compone di 10980 campioni e 34 feature, di cui molte con valori totalmente nulli o non importati ai fini della nostra analisi.

### B. Modellazione del Dataset

Prima di iniziare con l’analisi esplorativa il dataset è stato ridotto di dimensione andando a prendere unicamente le feature d’interesse qui di seguito riportate:

•

## III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads— $\text{\LaTeX}$  will do that for you.

### A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

### B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often

leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.

- Do not mix complete spellings and abbreviations of units: “Wb/m<sup>2</sup>” or “webers per square meter”, not “webers/m<sup>2</sup>”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use “cm<sup>3</sup>”, not “cc”).

### C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \quad (1)$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

### D. $\LaTeX$ -Specific Advice

Please use “soft” (e.g., `\eqref{Eq}`) cross references instead of “hard” references (e.g., (1)). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don’t use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in  $\LaTeX$  will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you’ve discovered a new method of counting.

$\BIBTeX$  does not work by magic. It doesn’t get the bibliographic data from thin air but from .bib files. If you use  $\BIBTeX$  to produce a bibliography you must send the .bib files.

$\LaTeX$  can’t read your mind. If you assign the same label to a subsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

$\LaTeX$  does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it’s supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won’t be any anyway) and it might stop a wanted equation number in the surrounding equation.

### E. Some Common Mistakes

- The word “data” is plural, not singular.
- The subscript for the permeability of vacuum  $\mu_0$ , and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
- Do not use the word “essentially” to mean “approximately” or “effectively”.
- In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
- Do not confuse “imply” and “infer”.
- The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the “et” in the Latin abbreviation “et al.”.
- The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [?].

### F. Authors and Affiliations

**The class file is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you

to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

H. Figures and Tables

a) *Positioning Figures and Tables:* Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

TABLE I  
TABLE TYPE STYLES

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy <sup>a</sup>		

<sup>a</sup>Sample of a Table footnote.



Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [?].

REFERENCES

[1] Shneidman ES. “The definition of suicide”. New York, NY and London: John Wiley and sons; 1985.

[2] Gvion Y, Apter A. “Suicide and suicidal behavior”. Public Health Reviews. 2012;34: epub ahead of print.

[3] Värnik, P. “Suicide in the world”. Int. J. Environ. Res. Public Health 2012, 9, 760-771. <https://doi.org/10.3390/ijerph9030760>

[4] World Health Organization. (2023). “Suicide rates”. Recuperato da <https://www.who.int/data/gho/data/themes/mental-health/suicide-rates>

[5] Statistics Canada. (2023). “Age-standardized rates”. Recuperato da <https://www.statcan.gc.ca/en/dai/btd/asr>