International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 05, ISSN No. 2455-2143, Pages 140-145
Published Online September 2023 in IJEAST (http://www.ijeast.com)

EVALUATING AGE DIFFERENCES IN LEFT-HANDERS AND RIGHT-HANDERS: A DATA-DRIVEN ANALYSIS

Tanushree Poojary
Electronics and Telecommunication
Thakur College of Engineering and Technology,India

Ankita Saha
Electronics and Telecommunication
Thakur College of Engineering and Technology,India

Dr.Sangeeta Mishra
Electronics and Telecommunication
Thakur College of Engineering and Technology,India

Abstract—**This research project investigates whether there is a significant difference in the average age of death between left-handers and right-handers by analyzing age distribution data collected from a National Geographic survey. Utilizing Bayesian statistics and data visualization techniques, the study calculates the probability of being a specific age at death based on handedness. The results reveal that there is insufficient evidence to support the claim that left-handers experience an earlier average age of death, challenging a commonly held belief. This evidence-based analysis contributes valuable insights to the discussion on handedness and longevity, highlighting the importance of rigorous data-driven research in dispelling prevailing myths.**

Keywords—**Data Analysis, Bayesian Statistics, Data Visualization, Data Collection, Age Distribution Analysis.**

I. INTRODUCTION

Within the realm of handedness and its implications for longevity, a profound question has loomed large: does a tangible disparity exist in the average age at death between left-handers and right-handers? This research undertaking is a deliberate exploration into this captivating inquiry, with a commitment to unveil the truth through data analysis and empirical rigor. To embark on this journey, it is pivotal to elucidate the essential terminologies that underpin our investigation.

First and foremost, 'age distribution data' encapsulates a comprehensive dataset that encompasses information about the ages at which individuals from various groups, both left-handed and right-handed, meet their demise. This data repository is essential to discern trends and differences in age at death between the two groups.

Secondly, 'Bayesian statistics' is a probabilistic approach to statistical inference and decision-making. It employs the principles of probability to calculate the likelihood of an event occurring, given prior information. In our research, Bayesian statistics are employed to quantify the probability of reaching a particular age at death based on an individual's handedness.

As for 'data visualizations and plots,' these serve as graphical representations of data that help researchers and readers to visually grasp patterns, trends, and differences within the dataset. In our project, these visualizations are pivotal in rendering complex information in a comprehensible form and making our analysis accessible.

Ultimately, our research seeks to reach a 'conclusive understanding' regarding the claim of an age difference between left-handers and right-handers. This involves scrutinizing the data, considering the probabilistic inferences derived from Bayesian statistics, and leveraging data visualizations to weigh the evidence. Our ultimate goal is to either corroborate or debunk the commonly held belief of handedness impacting longevity and contribute to a clearer, evidence-based understanding of this phenomenon.

II. PROPOSED ALGORITHM

A. Flow chart

The research project adopted a systematic workflow to investigate the potential age disparity in longevity between left-handers and right-handers. This approach involved a series of meticulously planned steps to achieve the research objectives. The primary stages that guided the project's progress are outlined below.

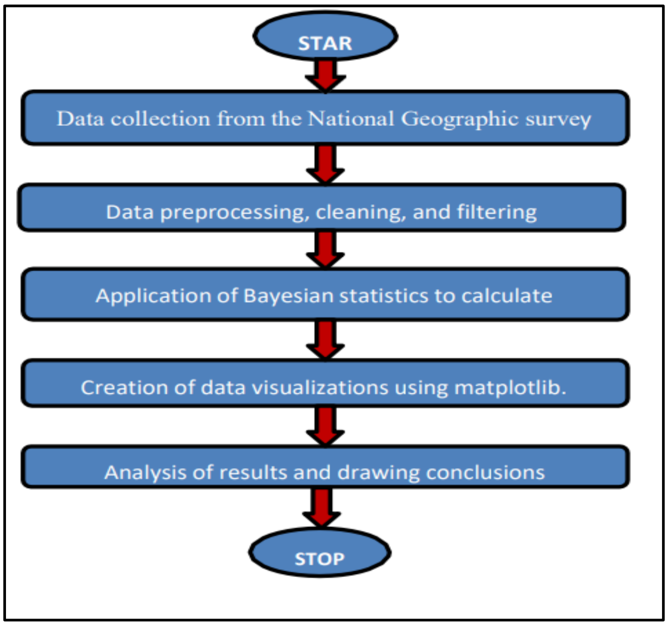

Fig 1.Implementation process

Data Collection from the National Geographic Survey: The initial phase of this research project involved the acquisition of data from the National Geographic survey. This dataset served as the fundamental source for our analysis, containing valuable information on the ages at which individuals passed away.

- Data Preprocessing, Cleaning, and Filtering: Subsequent to data collection, we conducted an essential process of data preprocessing, cleaning, and filtering. This meticulous step aimed to eliminate inconsistencies and inaccuracies, ensuring the dataset's integrity and reliability for analysis.
- Application of Bayesian Statistics: Bayesian statistics played a pivotal role in our analysis. This statistical method was employed to calculate the probabilities associated with reaching specific ages at death based on an individual's handedness, adding a probabilistic dimension to our investigation.
- Creation of Data Visualizations using Matplotlib: Effective data communication and visualization were achieved using the Matplotlib library. Data visualizations and plots were designed to illustrate patterns, trends, and distinctions within the dataset, making complex information more accessible.
- Analysis of Results and Drawing Conclusions: The analysis of processed data and insights derived from visualizations formed the basis for drawing informed conclusions. These conclusions aimed to provide clarity regarding the age disparity, or the lack thereof, between left-handers and right-handers.

B. Language and platform used:

Python, a versatile and readable high-level programming language, served as the foundation for the project's implementation. This choice was grounded in Python's design philosophy, emphasizing readability and ease of use, making it suitable for developers of various skill levels.

Key Python libraries supported the project's technical requirements:

- NumPy (Numerical Python): Facilitating operations on multi-dimensional arrays and matrices, NumPy's mathematical functions were vital for scientific computing and data manipulation.
- Pandas: Equipped with data structures and functions for structured data manipulation, Pandas played a significant role in data preprocessing, cleaning, and analysis.
- Matplotlib: This comprehensive library enabled the creation of a diverse range of visualizations in Python, contributing to the project's data visualization requirements.
- Seaborn: A higher-level interface built on Matplotlib, Seaborn facilitated the generation of aesthetically appealing statistical visualizations, especially valuable for complex visualizations with minimal coding.
- SciPy: As an ecosystem building on NumPy, SciPy provided additional functions for optimization,

integration, interpolation, and other scientific and engineering applications.

- Scikit-learn: Designed for machine learning tasks, Scikit-learn offered a wide array of algorithms for classification, regression, clustering, and more.
- TensorFlow and PyTorch: These two deep learning libraries provided tools for constructing and training neural networks, commonly applied in fields such as computer vision and natural language processing.
- NLTK and spaCy: These libraries specialized in natural language processing, offering functionalities for text processing, tokenization, and stemming.
- Statsmodels: Statsmodels was an essential resource for statistical modeling and hypothesis testing, offering a variety of models and functions for in-depth data analysis.
- OpenCV: As a computer vision library, OpenCV provided tools and functions for image and video processing, object detection, and related computer vision tasks.

These language and library choices collectively formed the technological backbone of the project, facilitating efficient data processing, analysis, and visualization. This selection was a testament to the adaptability and utility of these tools in conducting scientific research and data-driven investigations.

C. Implementation:

The initial phase of the project encompassed the identification of project requirements and the definition of the problem statement. The key objective was to ascertain whether a substantial difference existed in the average age at the time of death between individuals categorized as right-handers and those classified as left-handers. This determination was to be made through a meticulous analysis of the available data.

Data Collection and Importing:

The subsequent step involved the collection of age distribution data from the National Geographic survey. This dataset, serving as the primary source of information, encompassed data fields pertaining to age, gender, and the reported prevalence of left-handedness among individuals. To effectively work with this dataset, the data was imported into a pandas Data Frame, a versatile data structure in Python, from a CSV file. This CSV file served as the repository for the age distribution data.

Functions Used:

- read.csv (): This function operates as a wrapper for read.table(). It is primarily designed for reading data from CSV (Comma-Separated Values) files, assuming that the delimiter is a comma, and the first line of the file serves as the header that specifies the column names of the table. Additionally, it includes an additional parameter known as url() which is employed to directly access live data from a GitHub repository. This function, therefore, proves to be a suitable candidate for reading and importing CSV files, aligning perfectly with the project's data acquisition needs.

Data Cleaning:

The adage "Quality data beats fancy algorithms" holds true as data forms the cornerstone of Analytics and Machine Learning. It serves as the foundation for computational processes and business decisions. However, the real world often presents data that is marred by inconsistencies, incompleteness, or missing values. These imperfections can disrupt analytical processes and lead to erroneous outcomes. Hence, Data Cleaning stands as a fundamental tenet in the field of foundational data science.

Data Cleaning involves the methodical identification of erroneous, incomplete, irrelevant, or inaccurate segments within a dataset. Subsequently, it entails applying the necessary modifications, replacements, or removals to rectify these issues. In the context of this project, the pandas library plays a pivotal role, serving as the primary tool for executing intricate data cleaning operations.

Data Cleaning is not a perfunctory task but rather a meticulous process aimed at enhancing the quality and reliability of data. It ensures that the data under scrutiny is robust and well-prepared for subsequent analyses. The significance of this phase cannot be understated, for the integrity of any data-driven project relies heavily on the quality of the data that informs it.

Data Filtering:

Data filtering is the process of singling out a smaller, more specific subset from a larger dataset for observation, analysis, and evaluation. Typically, this filtration process is transient, meaning that the complete dataset remains intact, while only a particular portion of it is chosen for computational procedures. This approach, also known as sub setting or data extraction, involves the selection of data based on logically defined conditions. Data filtering serves various purposes, including:

- Examining outcomes within a specific timeframe.
- Computing outcomes for specific groups of interest.
- Excluding inaccurate or erroneous observations from analytical processes.
- Training and validating statistical models.

In the context of the dataset at hand, data filtering is employed to extract pertinent information for analysis. This filtration is executed based on gender and age criteria, allowing the extraction of data segments that are specifically relevant to the research objectives. For instance, the left-handedness rates for both males and females are extracted and then plotted against age. This data visualization technique serves as a powerful tool for recognizing and understanding trends and patterns within the data, enabling more focused and precise analysis.

III. EXPERIMENT AND RESULT

A. Male and Female Left-Handedness Rates vs. Age:

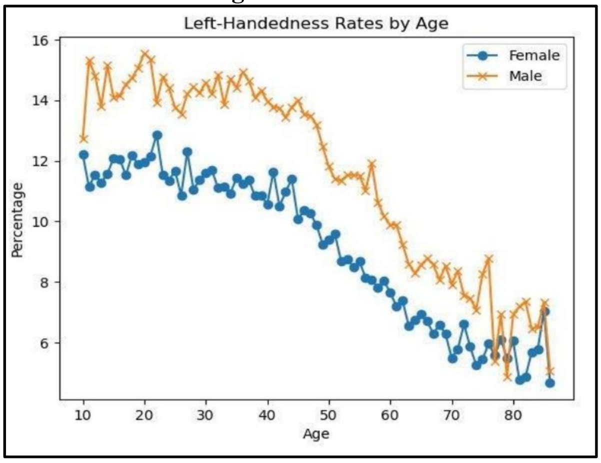

Fig 2.1

- In this visual representation, we observe the scatter plot, which demonstrates the left-handedness rates for both males and females across different age groups.
- The plot reveals the distribution of left-handedness within the population across age brackets. It allows us to discern whether there are any age-related patterns in handedness.
- By labeling the axes and incorporating a legend, the plot becomes more interpretable, enabling us to distinguish between the male and female data points. This distinction is valuable for understanding potential gender-based differences in handedness.

B. Rates of Left-Handedness Over Time:

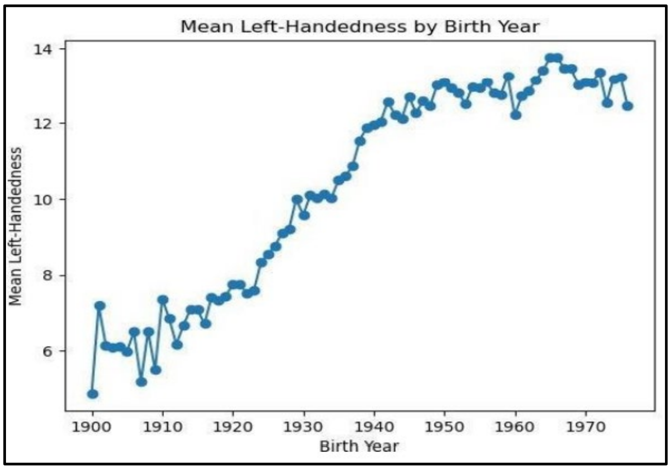

Fig 2.2

- This plot showcases the temporal evolution of left-handedness rates by computing the mean of male and female left-handedness rates across various birth years.
- It provides insights into how the prevalence of left-handedness has potentially changed over time, offering a historical perspective on handedness trends.

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 05, ISSN No. 2455-2143, Pages 140-145
Published Online September 2023 in IJEAST (http://www.ijeast.com)

- The plot's trendline enables us to identify whether left-handedness has shown any significant shifts or fluctuations across different birth cohorts.

C. Death Distribution Data for the United States in 1999

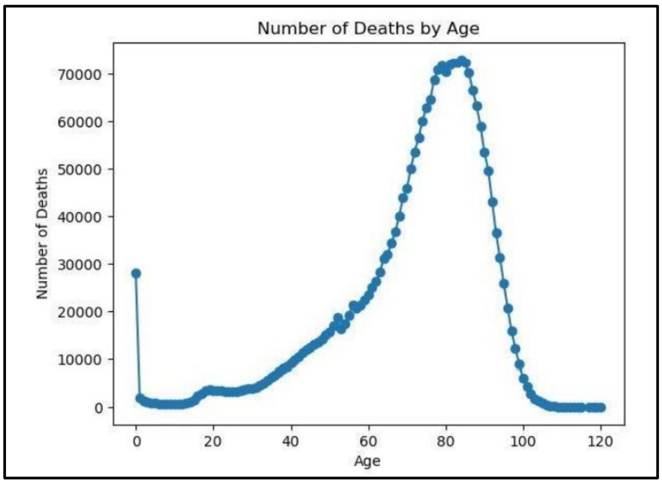

Fig 2.3

- The plot representing death distribution data for the United States in 1999 offers an overview of the ages at which individuals passed away during that year.
- It provides a visual representation of the distribution of ages at death, shedding light on the demographic aspects of mortality in that specific time frame.
- This analysis is foundational for understanding the typical age at which people passed away during the year 1999, offering context for assessing the age distribution within the population.

D. Probability of Age at Death Given Handedness:

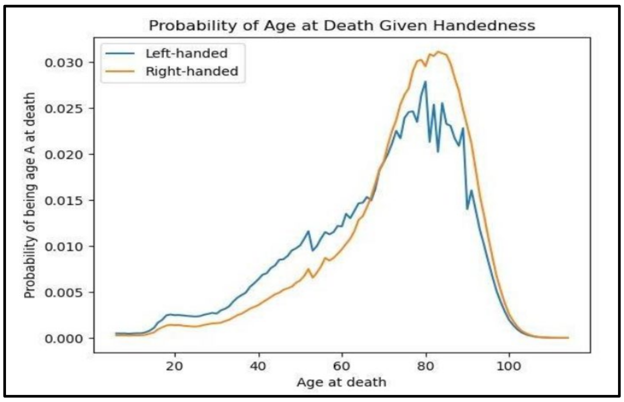

Fig 2.4

- The probability plots reveal the distributions of ages at death for left-handed and right-handed individuals. They offer insights into the likelihood of reaching specific ages at death based on handedness.
- By contrasting the probability distributions for the two groups, the plots facilitate the identification of potential disparities in age at death between left-handed and right-handed individuals.
- This analysis underscores the relevance of handedness as a factor in mortality studies, potentially challenging or confirming prevailing assumptions about its impact on the lifespan.

These visual representations and analyses are essential for extracting meaningful insights from the data, enabling us to draw conclusions and contribute to the overarching research question of whether there is a significant age difference between left-handers and right-handers.

IV.CONCLUSION

In light of the comprehensive analysis conducted on age distribution data, it is evident that the widely held belief suggesting left-handers experience an earlier average age of death in comparison to right-handers lacks substantial evidence. The Bayesian probability calculations performed in this study consistently indicate that the average ages at the time of death for both left-handed and right-handed individuals are strikingly similar. This discovery challenges the conventional wisdom and provides valuable insights into the ongoing discourse concerning the relationship between handedness and longevity. These findings emphasize the need for a more nuanced exploration of the multifaceted factors influencing lifespan and underscore the importance of data-driven conclusions.

The path forward in research on longevity and its correlations with handedness encompasses several promising trajectories. Future studies can extend their scope to examine a wider array of factors that influence lifespan, including health conditions, genetics, lifestyle choices, and socioeconomic variables. Employing larger and more diverse datasets offers the potential to refine our understanding of age distribution across different demographics and over more extended periods. Moreover, conducting longitudinal studies to observe the dynamic nature of these relationships and investigating the impact of tailored health and lifestyle interventions on longevity within handedness groups are intriguing avenues for further exploration. By incorporating these considerations into future research endeavors, we can move closer to a more comprehensive understanding of the complexities governing age at death and the intricate interplay with handedness, ultimately contributing to informed interventions aimed at enhancing the quality and longevity of life for all individuals.

V. REFERENCE

[1] Aggleton JP, Kentridge RW, Neave NJ (1993). Evidence for longevity differences between left-handed and right-handed men: an archival study of cricketers. J Epidemiol Community Health, 47(3), 206-209. DOI: 10.1136/jech.47.3.206. PMID: 8350033; PMCID: PMC1059767.
[2] Peto R (1994). Left-handedness and life expectancy. Causal inferences cannot be trusted. BMJ, 308, 408.
[3] Gilbert AN, Wysocki CJ (1992). Hand preference and age in the United States. Neuropsychologia, 30, 601–8.
[4] Basso O, Olsen J, Holm NV, Skytthe A, Vaupel JW, Christensen K (2000). Handedness and mortality: a follow-up study of Danish twins born between 1900 and 1910. Epidemiology, 11, 576–80.
[5] McManus IC, Moore J, Freegard M, Rawles R (2010). Science in the making: right hand, left Hand. III: estimating historical rates of left-handedness. Laterality, 15, 186–208.
[6] Richardson DB, Keil AP, Tchetgen Tchetgen E, Cooper G (2015). Negative control outcomes and the analysis of standardized mortality ratios. Epidemiology, 26, 727–32.
[7] Richardson DB, Keil AP, Cole SR, MacLehose RF (2017). Observed and expected mortality in Cohort Studies. Am J Epidemiol, 185, 479–86.
[8] Lawlor DA, Andersen AM, Batty GD (2009). Birth cohort studies: past, present and future. Int J Epidemiology, 38, 897–902.
[9] McManus IC (1995). Increased mortality in women doctors. Lancet, 345, 796–7.
[10] National Center for Health Statistics (2003). Live births, Birth Rates, and Fertility Rates, by race: United States, 1909–2000. In: Prevention CfDCa. Accessed June 18, 2018, at https://www.cdc.gov/nchs/nvss/births.htm.
[11] Morens DM, Katz AR (1991). Left-handedness and life expectancy. N Engl J Med, 325, 1041.
[12] Rothman KJ (1991). Left-handedness and life expectancy. N Engl J Med, 325, 1041.
[13] Panjer HH (1993). Mortality differences by handedness: survival analysis for a right-truncated sample of baseball players. Trans Soc Actuaries, 45, 257–74.
[14] Hicks RA, Johnson C, Cuevas T, Deharo D, Bautista J (1994). Do right-handers live longer? An updated assessment of baseball player data. Percept Mot Skills, 78, 1243–7.
[15] Katz JD (2004). Do anesthesiologists die at a younger age than other physicians? Age-adjusted death rates. Anesth Analg, 98, 1111–3.
[16] Editorial (1994). Burnished or burnt out: the delights and dangers of working in health. The Lancet, 344, 1583–4.
[17] Canova C, Cantarutti A (2020). Population-Based birth cohort studies in Epidemiology. Int J Environ Res Public Health, 17.
[18] McManus IC (2009). The history and geography of human handedness. In: Sommer IEC, Kahn RS, editors. Language lateralization and psychosis. Cambridge: Cambridge University Press; pp. 37–57.