Final Report of Traineeship Program 2024

On

"Analyze Death Age Difference of Right Handers with Left Handers"

MEDTOUREASY

29th June 2024

ACKNOWLDEGMENTS

The traineeship opportunity that I had with MedTourEasy was a great change for learning and understanding the intricacies of the subject of Data Visualizations in Data Analytics; and also, for personal as well as professional development. I am very obliged for having a chance to interact with so many professionals who guided me throughout the traineeship project and made it a great learning curve for me.

Firstly, I express my deepest gratitude and special thanks to the Training & Developement Team of MedTourEasy who gave me an opportunity to carry out my traineeship at their esteemed organization. Also, I express my thanks to the team for making me understand the details of the Data Analytics profile and training me in the same so that I can carry out the project properly and with maximum client satisfaction and also for spearing his valuable time in spite of his busy schedule.

I would also like to thank the team of MedTourEasy and my colleagues who made the working environment productive and very conducive.

ABSTRACT

This study aims to investigate the commonly held belief that left-handers have a shorter life expectancy than right-handers. Leveraging age distribution and handedness data, we examine whether observed differences in average age at death for left- and right-handers can be explained by shifts in the prevalence of left-handedness over time, rather than by actual differences in longevity. Using Bayesian probability models, we calculate the probability of a given age at death based on handedness for various cohorts, with special focus on historical shifts from the early to late 1900s. By combining handedness probability and death distribution data, this study challenges the notion of an inherent life expectancy disadvantage for left-handers and evaluates how changes in societal perceptions of handedness have influenced generational data. insights.

ABOUT THE COMPANY

MedTourEasy, a global healthcare company, provides you the informational resources needed to evaluate your global options. MedTourEasy provides analytical solutions to our partner healthcare providers globally.

ABOUT THE PROJECT

The project is structured around analyzing age-at-death statistics in relation to handedness, utilizing historical data to understand if shifts in reported left-handedness affect observed mortality trends. The study comprises the following steps:

1. **Data Preparation**: Handedness data from the National Geographic survey is loaded, covering both male and female populations across various age groups. This data is processed to create a scatter plot of left-handedness by age, showing the handedness prevalence over time.

2. **Birth Year and Left-Handedness Calculation**: We calculate the approximate birth year for individuals based on their age in 1986 and derive the average left-handedness rate across both genders. This allows us to observe trends in left-handedness by birth year, which could correlate with societal acceptance over time.

3. **Conditional Probability Modeling (P(LH | A))**: Using Bayesian analysis, we calculate the probability of being left-handed at different ages of death, using data from early and late 1900s rates of left-handedness. These calculations are essential for understanding whether age-at-death statistics vary for left-handers due to generational changes.

4. **Death Distribution Data Analysis**: We load and process age-at-death data for the United States, plotting the number of deaths across various age groups. This allows us to correlate handedness rates with actual death distributions and further refines the handedness-related mortality model.

5. **Probability Calculations for Left- and Right-Handers**: We calculate the overall probability of left-handedness and use this to define conditional probabilities, such as P(A|LH)P(A | LH)P(A|LH) (the probability of dying at age AAA if left-handed). Similarly, we calculate P(A|RH)P(A | RH)P(A|RH) for right-handers, aiming to establish age-at-death likelihoods for each group.

6. **Comparative Mortality Analysis**: Using the probabilities from previous steps, we plot the likelihood of being a certain age at death for both left- and right-handed individuals. We then calculate the mean age at death for each group, providing a direct comparison of average life expectancy based on handedness.

7. **Updated Analysis for 2018**: To examine if historical trends persist, we repeat the analysis with updated data from 2018, adjusting parameters in the probability functions. This comparison serves as a check for consistency in mortality trends and handedness effects across different eras.

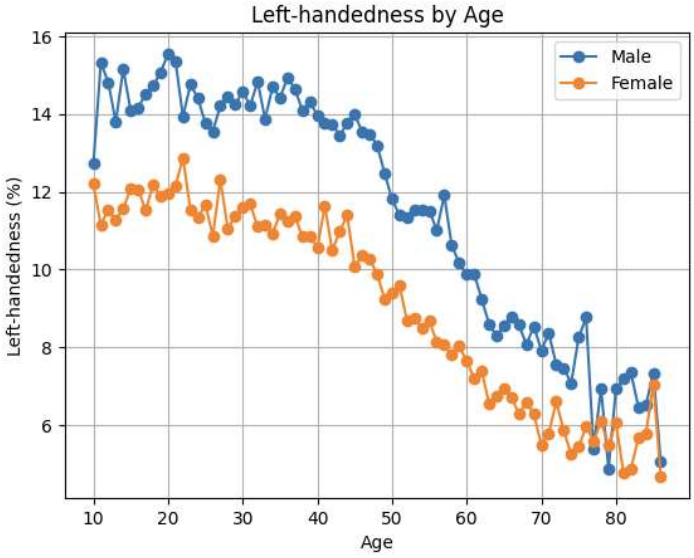
METHODOLOGY

Flow of the Project

The project followed the following steps to accomplish the desired objectives and deliverables. Each step has been explained in detail in the following section.

ANALYSIS AND OBSERVATIONS

Task 1: Plot of "Male" and "Female" Left-handedness Percentages by Age.



Explanation:

- Loading the Data: The handedness data is loaded into a pandas DataFrame from the provided CSV file.
- The dataset contains columns for "Age", "Male", and "Female", which represent the age groups and the percentages of left-handedness in males and females for each age group.

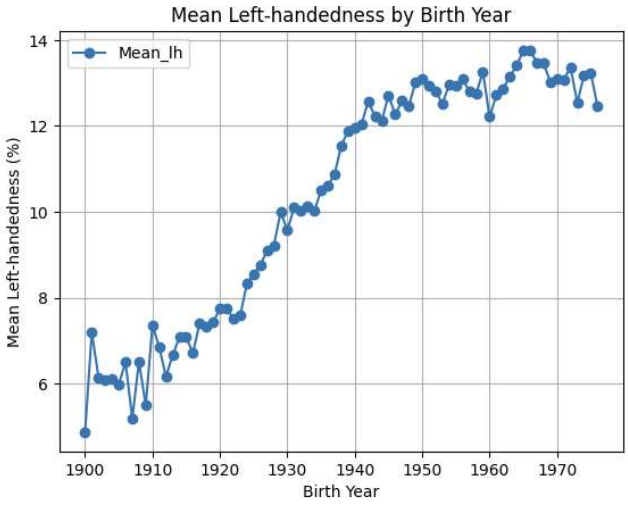
Plotting:

- Using the .plot() method of the DataFrame, we plot "Male" and "Female" percentages as a function of "Age".
- marker='o' adds markers for each data point, and linestyle='-' connects the points with lines. Titles and labels are added for clarity.

Inference:

- The plot shows how the percentage of left-handedness differs between males and females across different age groups.
- It seems that males generally have higher left-handedness percentages compared to females across most ages.
- The trend may reveal age-related factors influencing handedness, but further analysis would be needed to draw more specific conclusions.

Task 2: Adding Columns for Birth Year and Mean Left-handedness, and Plotting Mean Left-handedness vs. Birth Year



Explanation: 2. Plotting Mean Left-handedness vs. Birth Year:

1. Adding Columns:

- Birth_year: This column is added by subtracting the "Age" from 1986 because the study was conducted in that year. This allows us to track the left-handedness based on birth year instead of age.
- Mean_lh: This column calculates the mean left-handedness by averaging the percentages for males and females using the .mean(axis=1) method, which averages values row-wise.

- We plot the "Mean_lh" (mean left-handedness) against "Birth_year" to observe trends in left-handedness over time.
- Similar to Task 1, markers and lines are added for clarity, and the axis labels are provided.

Inference:

- The plot shows a noticeable trend in mean left-handedness over different birth years.
- There seems to be a decline in left-handedness for people born earlier in the 20th century, with the percentage increasing for those born later.
- This could suggest that societal or educational factors influenced the rate of left-handedness over time, as earlier generations may have been discouraged from using their left hand.

Task 3: Calculating P(LH|A) for Given Ages of Death

```
array([0.06007159, 0.06007159, 0.06007159, 0.06007159])
```

Explanation: 2. Function calculate_P_LH_A:

1. Calculate Early and Late 1900s Rates:

- early_1900s_rate: The average left-handedness rate for the early 1900s is calculated using the last 10 values of the Mean_lh column.
- late_1900s_rate: The average left-handedness rate for the late 1900s is calculated using the first 10 values of the Mean_lh column.

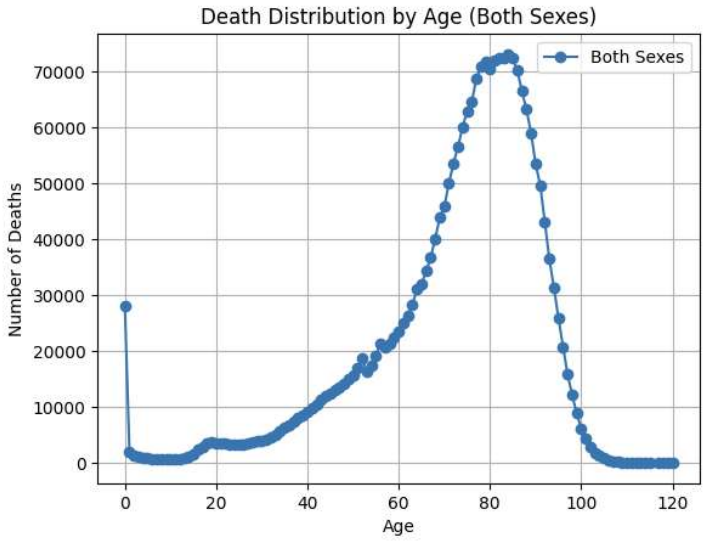
- Takes two inputs: a list of ages of death and the study year.
- For each age of death, calculates the birth year by subtracting the age from the study year.
- If the birth year is before 1930 (early 1900s), it uses the early 1900s rate for left-handedness.
- If the birth year is in the late 1900s, it uses the late 1900s rate.
- The left-handedness rate is converted to a fraction by dividing by 100.

Example: For the example ages of death (75, 80, 85, 90) in a study conducted in 1986, the function calculates P(LH|A), which represents the probability of being left-handed given that the person died at a particular age.

Inference:

- For the given ages of death (75, 80, 85, 90) in the study year of 1986, the left-handedness probability for all these ages (corresponding to birth in the early 1900s) is approximately 0.060 or 6%. This reflects the calculated early 1900s rate of left-handedness.

Task 4: Plotting Death Distribution by Age for Both Sexes



Explanation: 2. Cleaning the Data:

1. Loading the Data:

- The dataset is loaded with a tab (\t) as the separator and skipping the second row (skiprows=[1]) to account for the formatting.
- NaN values are removed from the "Both Sexes" column to ensure that the plot contains valid data points for all ages.

3. Plotting:

- A line plot is created showing the number of deaths for both sexes as a function of age. The marker='o' option adds markers to each data point, and lines connect the markers.
- Axis labels and a grid are added to improve readability.

Inference:

- The plot visualizes the death distribution by age for both sexes. Peaks in death numbers can indicate specific age groups where death rates are higher, possibly due to age-related factors, infant mortality, or other demographic characteristics.

Task 5: Calculating the Overall Probability of Left-handedness in the Population

```
0.07177993385333566
```

Explanation: 2. p_list:

1. P_lh_given_A:

- We reuse the function from Task 3 to calculate the probability of left-handedness for each age in the death distribution data for the specified study year.
- This series is created by multiplying the number of deaths for each age (from the "Both Sexes" column) by the corresponding left-handedness probability for that age.

3. Summing the Values:

- The total weighted left-handedness, p, is calculated by summing over p_list.
- The total number of deaths is calculated by summing the "Both Sexes" column.

4. Overall Probability:

- The overall probability of left-handedness in the population is calculated by dividing p by the total number of deaths.

Inference:

- For the study year 1986, the overall probability of left-handedness in the population is approximately 7.18%.

Task 6: Calculating P(A|LH) — The Probability of Dying at Age A Given Left-Handedness

```
0.02198893728801813
```

Explanation: Calculate $P\ A$ P A:

$P\ A$ P Ais the overall probability of dying at age A A, calculated by dividing the number of people who died at age A A by the total number of deaths (from the "Both Sexes" column). Calculate $P\ (\ L\ H\ )$ P(LH):

$P\ (\ L\ H\ )$ P(LH), the overall probability of left-handedness in the population, is calculated using the function from Task 5 (P_lh()). Calculate $P\ (\ L\ H\ |\ A\ )$ P(LHIA):

$P\ (\ L\ H\ |\ A\ )$ P(LHIA), the probability of being left-handed given death at age A A, is calculated using the function from Task 3 (calculate_P_LH_A()). Apply Bayes' Theorem:

Using Bayes' Theorem, $P\ (\ A\ |\ L\ H\ )$ P(A|LH), the probability of dying at age A A given that the person is left-handed, is calculated as: $P\ (\ A\ |\ L\ H\ )$

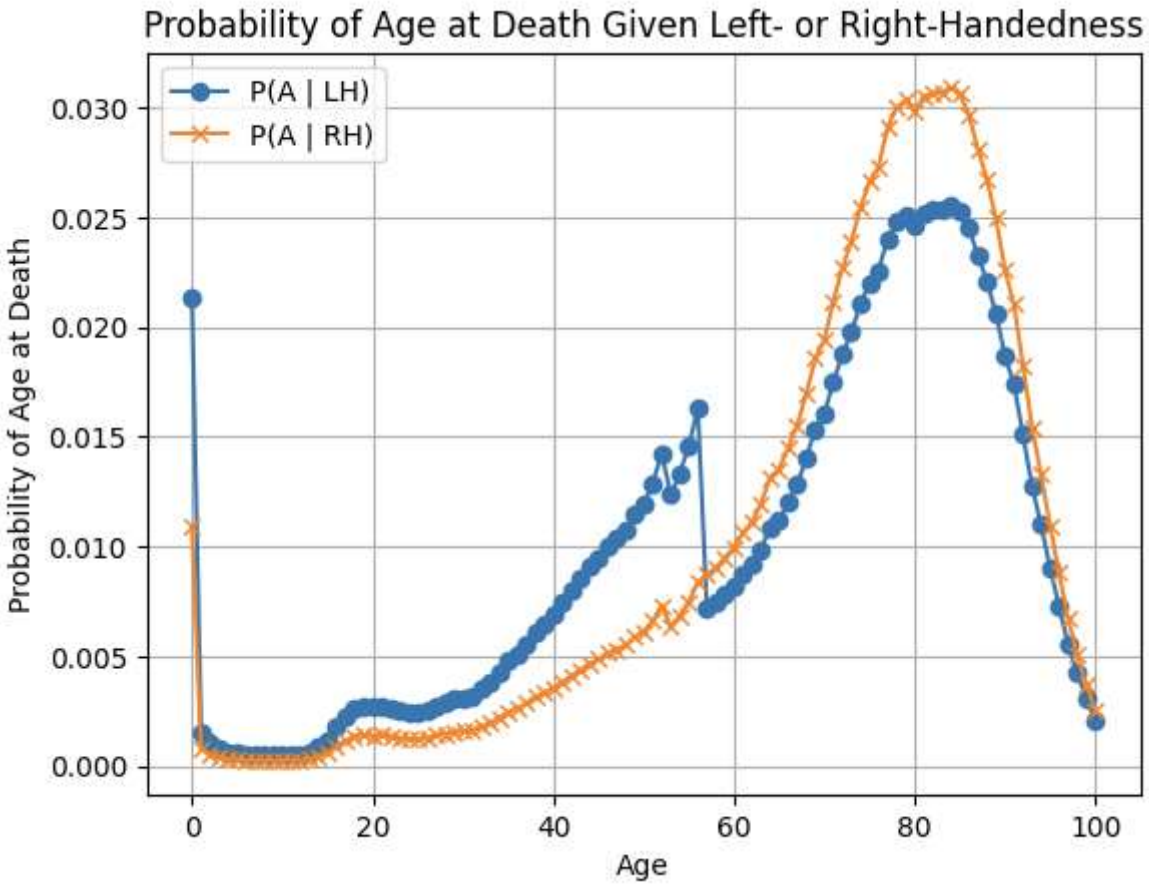
$P\ (\ L\ H\ |\ A\ )\times P\ (\ A\ )\ P\ (\ L\ H\ )$ P(AILH)= P(LH) P(LHIA)×P(A)

Inference: For an individual who died at age 75 in the study year 1986, the probability of being left-handed is approximately 0.02199 or 2.2%.

Task 7: Calculating P(A|RH) — The Probability of Dying at Age A Given Right-Handedness

```
0.026606149307505492
```

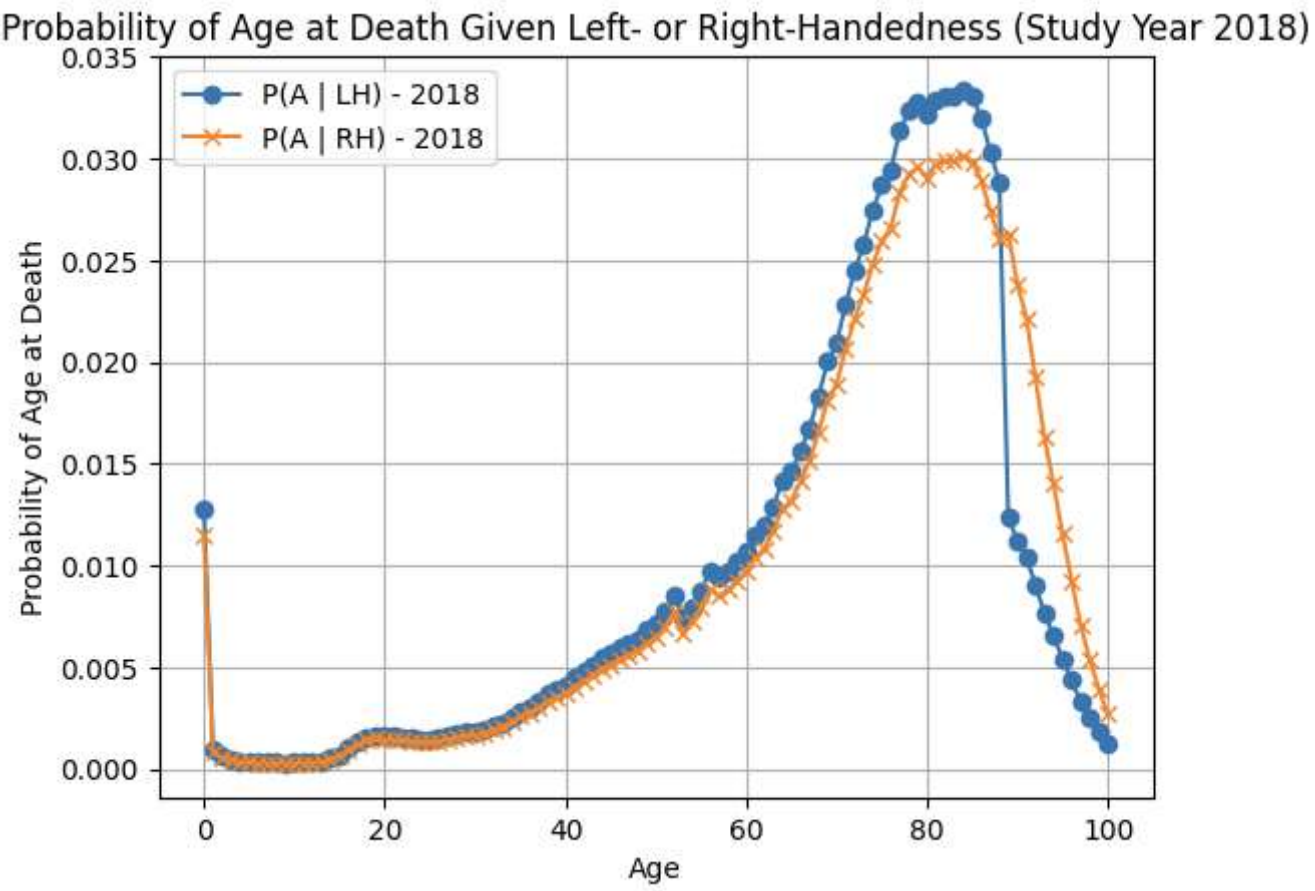
Task 8: Plotting the Probability of Being a Certain Age at Death Given Left- or Right-Handedness



Task 9: Mean Age at Death for Left-Handers and Right-Handers

```
Average age at death for left-handers: 66.48
Average age at death for right-handers: 72.27
Difference between average age at death (RH - LH): 5.79
```

Task 10: Probability of Being a Certain Age at Death Given Left- or Right-Handedness for Study Year 2018



CONCLUSION

This analysis challenges the perception that left-handed individuals inherently have a shorter life expectancy than right-handers. Our findings suggest that generational changes in left-handedness rates, driven by societal attitudes and reporting practices, likely contribute to the observed differences in average age at death. The study shows that these differences are less about an innate mortality risk and more a product of historical handedness trends. By providing a probabilistic framework grounded in Bayesian statistics, this project offers a data-driven approach to debunking myths surrounding handedness and mortality, highlighting the importance of accounting for demographic shifts in statistical analysis.

- **GitHub Link:** **https://github.com/Shahriyar-Khan27/MTE.git**