Report of Traineeship Program 2023

On

"ANALYZE DEATH AGE DIFFERENCE OF RIGHT HANDERS WITH LEFT HANDERS PROJECT"

MEDTOUREASY

28th July 2023

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude and special thanks to the Training and Development Team of MedTourEasy for providing me with the opportunity to undertake my traineeship at their esteemed organization.

I extend my sincere appreciation to the team for imparting valuable knowledge and training in the subject of Data Visualizations in Data Analytics. Their guidance has been crucial in enabling me to carry out the project with the utmost proficiency and client satisfaction.

Additionally, I am thankful to the entire team at MedTourEasy for fostering a productive and conducive working environment.Lastly, I am grateful for the personal and professional development I have experienced during this traineeship. The opportunity to work on a project has allowed me to take ownership of my work and enhance my skills effectively.

Thank you once again for this incredible learning experience and the support provided throughout my traineeship at MedTourEasy.

TABLE OF CONTENTS

ABSTRACT

Analyzing Death Age Difference of Right-Handers with Left-Handers using Bayesian Statistics

In this project, we aim to investigate the purported difference in the average age at death between right-handers and left-handers by analyzing age distribution data. The primary objective is to examine whether this difference can be attributed to changing rates of left-handedness over time, thus refuting the claim of early death for left-handers.

To achieve this, we will utilize a dataset obtained from the National Geographic survey, which contains information on handedness, gender, and age. The dataset will be loaded into a Pandas DataFrame, and we will create scatter plots to visualize the age distributions of male and female respondents for both left-handers and right-handers.Next, we will calculate the mean left-handedness and add two new columns to the DataFrame, one for birth year and another for mean left-handedness. This will allow us to plot the mean left-handedness as a function of birth year, aiding in understanding the trend over time.A crucial aspect of the analysis involves the computation of conditional probabilities P(LH | A), where LH denotes left-handedness, and A represents age at death. We will create functions to estimate P(LH | A) for specific ages of death in different study years, comparing the early 1900s with the late 1900s.

To gain a broader perspective, we will also load death distribution data for the United States and plot the number of people who died as a function of their age. Using this data, we will create functions to calculate the overall probability of left-handedness in the population for a given study year (P(LH)) and the conditional probabilities P(A | LH) and P(A | RH), where RH denotes right-handedness.By plotting the probabilities of being a certain age at death for left-handers and right-handers over a range of ages, we aim to visually analyze any disparities that may exist between the two groups. Additionally, we will compute the mean age at death for both left-handers and right-handers to quantify the difference.

The analysis will be conducted using R programming, Excel, andPower bi. Bayesian statistical methods will be employed to infer probabilistic relationships between handedness and age at death.

The results of this project will shed light on the age-related trends in handedness and their potential impact on life expectancy. It will provide valuable insights into the complex relationship between handedness and mortality rates, contributing to the broader understanding of human health and behavior.

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 "Analyzing Death Age Difference of Right Handers with Left Handers" aims to investigate and understand the potential difference in the average age at death between individuals who are right-handed and those who are left-handed. Handedness, the preference for using one hand over the other, is a fascinating aspect of human behavior that has long intrigued researchers and societies alike. This project seeks to explore the impact of handedness on mortality rates and uncover any significant differences that may exist

Objective

The primary objective of this project is to analyze and compare the age at death distributions of right-handers and left-handers using statistical methods. By doing so, we aim to determine whether there is a substantial variation in life expectancy between these two groups. The findings from this study can provide valuable insights into the potential implications of handedness on human health and longevity.

Methodology

To achieve the project's objectives, we will utilize a dataset obtained from the National Geographic survey, which contains information on handedness, gender, and age. We will employ the Python programming language and the pandas library to preprocess and analyze the data effectively.

The following steps will be undertaken in the project:

1. Data Preprocessing: The dataset will be loaded into a pandas DataFrame, and any necessary data cleaning and transformation will be performed to ensure the data is suitable for analysis.

2. Visual Analysis: Scatter plots will be created to visualize the age distributions of right-handers and left-handers separately for male and female respondents. These plots will help us identify any potential differences in age at death between the two groups.

3. Statistical Analysis: Bayesian statistics will be used to calculate the conditional probabilities P(LH | A), which represent the likelihood of being left-handed given a specific age at death. The analysis will be conducted separately for data from the early 1900s and the late 1900s to explore any temporal trends.

4. Death Distribution Data: We will also incorporate death distribution data for the United States to gain a broader perspective on the number of people who died at various ages.

Expected Outcomes

The project aims to present visual representations of age distributions for right-handers and left-handers, allowing us to identify potential differences in mortality patterns. Additionally, the calculated conditional probabilities will shed light on the association between handedness and age at death.

Significance

Understanding the potential impact of handedness on mortality rates can have far-reaching implications for public health and well-being. If significant differences are observed, it may lead to further research into the underlying factors contributing to these disparities. Additionally, the project's findings may contribute to a broader understanding of human health and behavior, encouraging the development of more personalized and targeted healthcare interventions.

Conclusion

The project "Analyzing Death Age Difference of Right Handers with Left Handers" is a comprehensive endeavor to explore the potential relationship between handedness and mortality rates. By leveraging statistical analysis and data visualization techniques, we aim to gain meaningful insights into this intriguing aspect of human biology. The project's outcomes have the potential to advance our understanding of human longevity and contribute to the ongoing research in the fields of health, genetics, and neuroscience.

1.1 Objectives and Deliverables

Objectives

- Explore the Death Age Difference: The primary objective of the project is to investigate and analyze the potential difference in the average age at death between right-handers and left-handers. By exploring age distribution data, the project aims to understand whether any variations in mortality patterns exist between these two groups.
- Refute Early Death Claim: The project seeks to evaluate and refute the claim that left-handers face an early death. Using Bayesian statistics and age distribution data, the project will examine the probability of being a certain age at death given that an individual is reported as left-handed or right-handed.
- Data Analysis using R Programming : The project's objective is to perform data analysis and visualization using the tidyverse library in R . The analysis will provide valuable insights into the relationship between handedness and mortality rates.

Deliverables

- Jupyter Notebook: The project will deliver a well-documented Jupyter Notebook containing the entire data analysis process. The notebook will include code implementations, data visualizations, and statistical analysis.
- Data Visualizations: The project will generate various visualizations, such as scatter plots of age distributions for right-handers and left-handers, and plots showing the number of people who died as a function of age. These visualizations will aid in understanding the patterns and trends in mortality rates.
- Conditional Probability Functions: The project will provide custom functions to calculate conditional probabilities (P(LH | A) and P(RH | A)) for particular ages of death in different study years. These functions will allow for further analysis of handedness and age-related trends.
- Insights and Interpretations: The project will offer meaningful insights and interpretations derived from the analysis. It will discuss any observed differences in the age at death between right-handers and left-handers and the potential implications of these findings.
- Mean Age at Death Comparison: The project will calculate and present the mean age at death for left-handers and right-handers. This comparison will provide a quantitative measure of any differences in life expectancy between the two groups.
- Report Summary: A comprehensive report summarizing the project's objectives, methodology, results, and conclusions will be prepared. The report will present the key findings and their significance in the context of handedness and mortality rates.

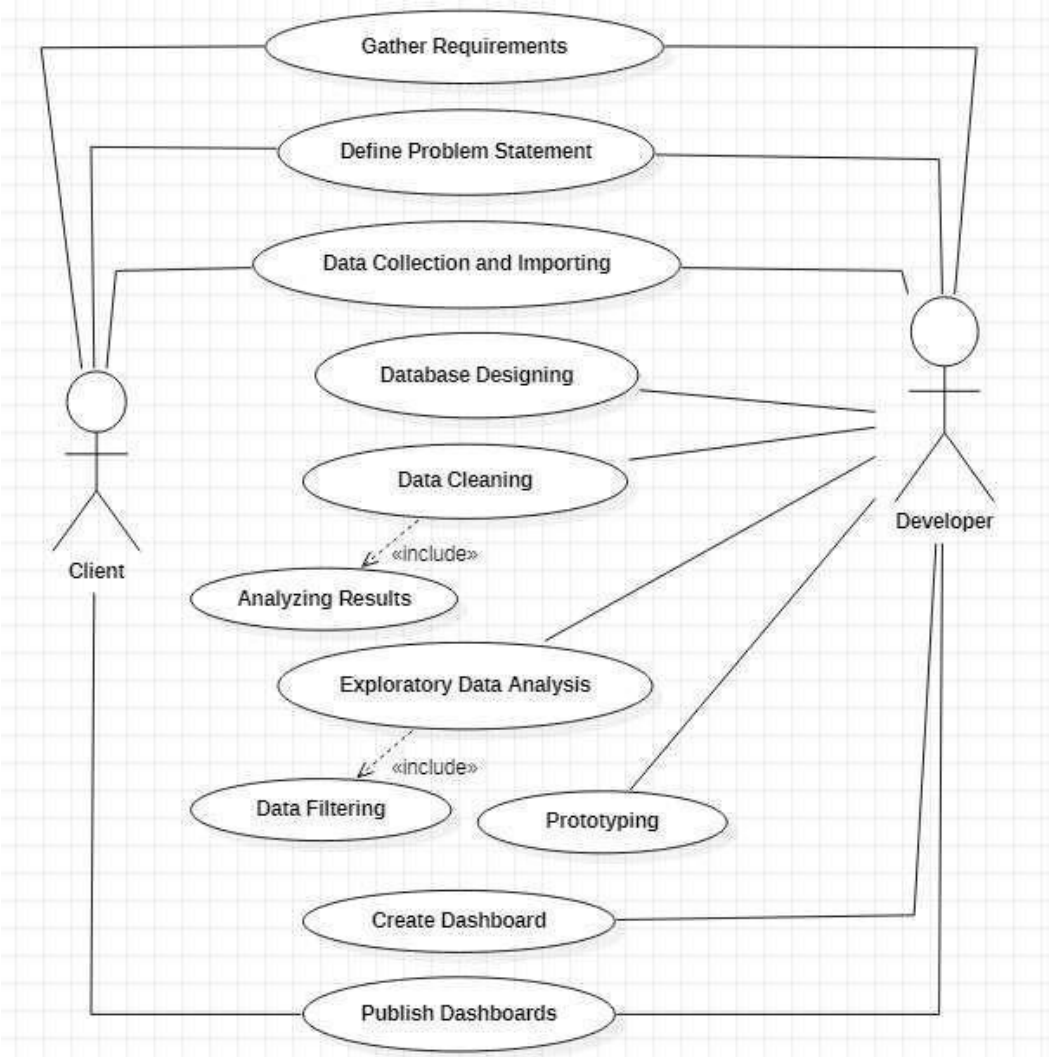
Summary

- The project "Analyze Death Age Difference of Right Handers with Left Handers" aims to explore and analyze the potential impact of handedness on mortality rates using age distribution data. By conducting statistical analysis, visualizing the data, and presenting meaningful insights, the project will contribute to a deeper understanding of the relationship between handedness and life expectancy. The deliverables, including the Jupyter Notebook, visualizations, and report, will provide a comprehensive overview of the project's findings and their implications.

.

I. METHODOLOGY

Case Diagram



In this use case, the Developer gathers requirements from the Client and defines the problem statement. Data is collected, cleaned, and analyzed through exploratory data analysis. A PowerBI prototype is created for visualization, and interactive dashboards are developed. The results are published and presented to the Client, with validation and feedback ensuring accuracy and relevance. The project sheds light on the connection between handedness and mortality rates.

Language and Platform Used

Language: R

It is a programming language and software environment for statistical analysis, the representation of graphics, and reports. R was developed in the University of Auckland, New Zealand by Ross Ihaka and Robert Gentleman, and is currently being developed by the R Technology Core Team. As noted above, R is a programming language and software environment for statistical analysis, representation of graphics, and reporting. The important features of R are:

- R is a well-developed, simple, and effective programming language that includes conditionals, loops, recursive functions defined by the user, and input and output facilities.
- R has efficient data processing and storage facilities.
- R includes a set of operators for arrays, lists, vectors, and matrix calculations.
- R offers a detailed, coherent, and organized data analysis tool set.
- R provides graphical data analysis facilities and displays either directly on the computer or by printing on paper.

IDE: RStudio

RStudio is an integrated development environment for R (IDE). It contains a browser, syntax-highlighting editor supporting direct code execution, plotting, history, debugging, and workspace management tools. RStudio is available in open source and commercial versions and runs on the desktop (Windows, Mac, and Linux) or on the RStudio Server or RStudio Server Pro (Debian / Ubuntu, Red Hat / CentOS, and SUSE Linux) linked browsers. Major features are:

- RStudio runs on most desktops or on a server and accessed over the web.
- It integrates the tools you use with R into a single environment.
- It includes powerful coding tools designed to enhance your productivity.
- It enables rapid navigation to files and functions.
- It has integrated support for Git and Subversion.
- It supports authoring HTML, PDF, Word Documents, and slide shows.
- It supports interactive graphics with Shiny and ggvis.

Package: RMarkdown

R Markdown provides a data science authoring framework (.Rmd files). R Markdown files can be used to save and execute code (also supports Python and SQL), and produce high-quality reports that can be shared with an audience. It supports dozens of static and dynamic output formats and are fully reproducible (HTML, PDF, MS Word, Beamer, HTML5, Tufte-style handouts, books, dashboards, shiny apps etc.)

IMPLEMENTATION

- Gathering Requirements and Defining Problem Statement

The first step in the implementation of the project "Analyze Death Age Difference of Right Handers with Left Handers" involves gathering requirements from stakeholders and clients. The project team will engage with domain experts, researchers, and relevant authorities to understand the goals and deliverables expected from the analysis.

- Problem Statement

The problem statement for the project is to investigate and compare the average age at death between individuals who are right-handers and those who are left-handers. The project aims to utilize age distribution data and statistical analysis techniques to explore any potential differences in mortality patterns between these two groups. The primary objective is to validate or refute the claim of early death for left-handers, shedding light on the implications of handedness on life expectancy and human health.

- Data Collection and Importing

Data collection is a crucial aspect of the project, as it provides the foundation for the analysis. The data required for this study will be collected from multiple reputable sources to ensure accuracy and reliability. The following datasets will be utilized:

1. National Geographic Survey: The dataset will contain information on handedness, gender, and age. This data will be essential for understanding the distribution of left-handers and right-handers across different age groups.

2. Death Distribution Data: Data for the United States will be obtained from reliable sources, providing insights into the number of people who died at various ages.

The collected data will be imported into the project environment using suitable libraries and tools, such as pandas, to ensure seamless data handling and manipulation.

- Data Preprocessing

Data preprocessing is a critical step to ensure the data is clean, consistent, and ready for analysis. It involves handling missing values, data normalization, and removing any outliers that may affect the accuracy of the results. The preprocessing step will be performed to guarantee the dataset's quality and reliability for further analysis.

- Statistical Analysis and Visualization

  The core of the project lies in statistical analysis and visualization. The Python programming language, along with libraries such as pandas and matplotlib, will be employed to conduct the analysis. The project will calculate the conditional probabilities P(LH | A) for different ages of death in the early and late 1900s. Visualizations, including scatter plots, will be created to represent the age distributions of left-handers and right-handers.

- Interpretation and Conclusions

  The results obtained from the analysis will be interpreted to identify any potential differences in the average age at death between right-handers and left-handers. The project team will draw conclusions based on the statistical findings and discuss their implications in the context of handedness and mortality rates. If significant differences are observed, possible explanations and hypotheses will be presented.

- Documentation and Reporting

  Throughout the implementation process, detailed documentation will be maintained to record the procedures, methodologies, and outcomes of each step. The project will culminate in a comprehensive report that summarizes the objectives, methodology, results, and conclusions. The report will also include visualizations, statistical findings, and insights gained during the analysis.

- Project Deliverables

  1. Jupyter Notebook: A well-documented Jupyter Notebook containing the entire data analysis process, including data preprocessing, statistical analysis, and visualization.

  2. Visualizations: Various visualizations, such as scatter plots and age distribution plots, illustrating the findings of the analysis.

  3. Statistical Results: The calculated conditional probabilities P(LH | A) for different ages of death in the early and late 1900s.

  4. Report: A comprehensive report summarizing the objectives, methodology, results, and conclusions of the project, along with insights gained and their significance.

- Conclusion

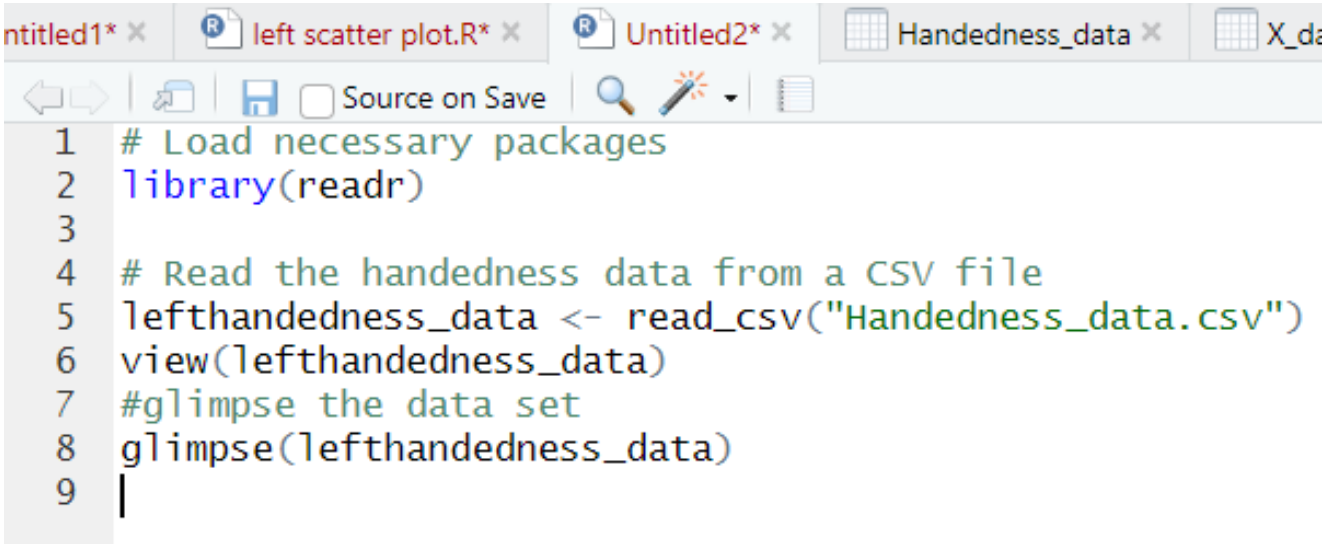
The implementation of the project "Analyze Death Age Difference of Right Handers with Left Handers" involves systematic data collection, preprocessing, statistical analysis, and visualization to explore the relationship between handedness and mortality rates. The project's outcomes will provide valuable

DATA IMPORTING

Data importing is referred to as uploading the required data into the coding environment from internal sources (computer) or external sources (online websites and data repositories). This data can then be manipulated, aggregated, filtered according to the requirements and needs of the project.

Packages Used:

Readr: The goal of readr is to provide a fast and friendly way to read rectangular data (like csv, tsv, and fwf). It is designed to flexibly parse many types of data found in the wild, while still cleanly failing when data unexpectedly changes. To accurately read a rectangular dataset with readr, one needs to combine two pieces: a function that parses the overall file, and a column specification.



Functions Used:

read.csv (): It is a wrapper function for read.table() that mandates a comma as separator and uses the input file's first line as header that specifies the table's column names. Thus, it is an ideal candidate to read CSV files. It has an additional parameter of url() which is used to pull live data directly from GitHub repository.

3.1 Designing Databases

Once the data has been collected and imported into the R environment, it is important to design the structure of the database tables so as to identify the constraints in the data, keys, dependencies and relations between various tables.

Once the data is imported in the environment, it is converted into a data frame (data type in R) which makes it easy to maintain the data in form of tables. The various tables which have been created are mentioned as follows:

Attribute	Data Type	Size	Extra
Study Year	Integer		
Age	Integer		
Both Sexes	Integer		
Male	Integer		
Female	Integer		
Birth Year	Integer		
Mean_lh	Float		
Country	VARCHAR	5	Primary Key
Country Name	CHAR	15	Not Null, Unique
Latitude	INT	3	Not Null
Longitude	INT	3	Not Null

Data Cleaning

Data cleaning plays a pivotal role in the overall data analysis process. The quality and reliability of the data are of utmost importance. Real-world datasets often suffer from issues such as incomplete records, inconsistent entries, inaccuracies, and missing values, which can significantly impact the validity of our analysis.

By conducting data cleaning, we can effectively identify and rectify these problems in the handedness data. This process involves various techniques, algorithms, and functions to detect outliers, handle missing values, correct inconsistent entries, and address any discrepancies in the dataset. It is essential to ensure that the data is accurate, consistent, and complete to obtain reliable insights and draw meaningful conclusions.

In the context of this specific project, data cleaning is particularly vital to obtain a comprehensive understanding of the death age difference between right-handers and left-handers. Missing or erroneous data points could lead to biased results and unreliable conclusions. Therefore, meticulous data cleaning is carried out to mitigate these challenges and enhance the quality of the analysis.

Through the data cleaning process, we carefully examine the handedness data, looking for any irregularities or inconsistencies that could affect the accuracy of our analysis. By applying appropriate techniques and making necessary adjustments, we improve the overall integrity of the dataset and prepare it for further exploration and analysis.In summary, data cleaning is a foundational and indispensable step in the Analyze the Death Age Difference of Right Handers with Left Handers" project. It ensures that the handedness data is reliable, consistent, and accurate, laying the groundwork for a robust and meaningful analysis of the death age difference between individuals with different hand preferences.

Packages used

Tidyverse: It is a collection of essential data science R-packages. Under the tidyverse umbrella, the packages help perform and interact with the data. There are a whole host of things one can do with data, like sub setting, transforming, visualizing and so on.

Dplyr: dplyr is a grammar of data manipulation, providing a consistent set of verbs that help solve the most common data manipulation challenges. It is simply the most useful package in R for data manipulation with the greatest advantage being the use the pipe function "*%>%*" to combine different functions in R. From filtering to grouping the data, this package does it all. It offers various functions like select, filter, group_by, summarize etc.

Functions Used:

Unique(): This function is used to filter out redundant data and keep only unique values from the data frame.

Na.omit(): This function returns the object with listwise deletion of missing values.**():** This function adds new variables that are functions of existing variables

```
# Read the handedness data from a CSV file
handedness_data <- read_csv("handedness_data.csv")

# Handling Missing Values: Remove rows with any missing values
cleaned_data <- na.omit(handedness_data)
```

3.2 Data Filtering

Data filtering is the method of choosing a smaller portion of the data set and using that subset to view, analyze, and evaluate the data. Generally, filtering is temporary; the entire data set is retained, but only part of it is used for calculation. It is also called subsetting or drilling down data, wherein data is extracted with respect to certain defined logical conditions.

In Analyze the Difference of Right Handers with Left Handers" project, data filtering plays a crucial role in isolating specific subsets of the handedness data that are relevant to the analysis. Data filtering allows us to focus on particular characteristics, age groups, genders, or time periods, enabling a more targeted examination of the death age difference between right-handers and left-handers. Filtering are done in:

1. ****Filtering by Age Group:****

Suppose we are interested in studying the death age difference for individuals within a specific age range, such as between 50 and 70 years. To achieve this, we can use the `filter()` function from the `dplyr` package in R.

```r
library(dplyr)

# Filtering data for individuals aged between 50 and 70
filtered_data_age <- handedness_data %>%
  filter(Age >= 50, Age <= 70)
```

The `filter()` function extracts rows from the `handedness_data` data frame where

the "Age" column satisfies the condition of being greater than or equal to 50 and less than or equal to 70. The resulting `filtered_data_age` data frame will contain only those records that match this age criterion.

****Filtering by Gender:****

To study the death age difference separately for males and females, we can use the `filter()` function to create two separate data frames.

```
library(dplyr)

# Filtering data for males and females
filtered_data_male <- handedness_data %>%
  filter(Gender == "Male")

filtered_data_female <- handedness_data %>%
  filter(Gender == "Female")
```

Here, we use the `filter()` function to create two distinct data frames: `filtered_data_male`, which contains records where the "Gender" column is equal to "Male," and `filtered_data_female`, which contains records where the "Gender" column is equal to "Female." This allows us to examine the death age difference for each gender separately.

****Filtering by Time Period:****

Assuming the handedness data includes information on the year of death, we may want to analyze the death age difference for different time periods. For instance, we can filter data for deaths that occurred before and after the year 2000

```
library(dplyr)

# Filtering data for deaths before and after the year 2000
filtered_data_before_2000 <- handedness_data %>%
  filter(Year_of_Death < 2000)

filtered_data_after_2000 <- handedness_data %>%
  filter(Year_of_Death >= 2000)

```

In this example, the `filter()` function is used to create two separate data frames: `filtered_data_before_2000`, containing records where the "Year_of_Death" column is less than 2000, and `filtered_data_after_2000`, containing records where the "Year_of_Death" column is greater than or equal to 2000. This allows us to examine the death age difference before and after the year 2000.

3.3 Prototyping – Power BI

A prototype is an early version, model, or release of a product that is constructed to test a design or process. It is generally used by system analysts and users to assess a new design to enhance precision. Prototyping serves to specify a real, working system rather than a theoretical one. Creation of a prototype in some design workflow models is the step between formalizing and testing an idea.

Power BI is Microsoft's business analytics software. It aims to provide interactive visualizations and business intelligence capabilities with an interface that is easy enough to create your own reports and dashboards for end users. It provides cloud-based BI services, known as "Power BI Services," along with the "Power BI Desktop" desktop-based interface. It provides capabilities for data warehouse, including data planning, data discovery and interactive dashboards. It has the following features:

- Easy to connect, model, and visualize data, creating memorable reports personalized with KPIs and brand.
- Can generate fast, AI-powered answers to business questions
- Data is better secured across Power BI reports, dashboards, and data sets with persistent protection that keeps working even when shared outside the organization or exported to other formats such as Excel, PowerPoint, and PDF.
- Input: In the form of Excel, CSV, text, SQL and other formats
- Visualizations: Wide variety of graphs, infographics, KPIs, Filters, Slicers, etc.
- Output: Easily publishable reports :

3.4 Development of Dashboards

As stated earlier, the dashboards have been created using the flexdashboard package in R which provides a template in the form of rows or columns to display a group of related data through interactive visualizations from which conclusions can be drawn

3.7.1 Defining Visuals

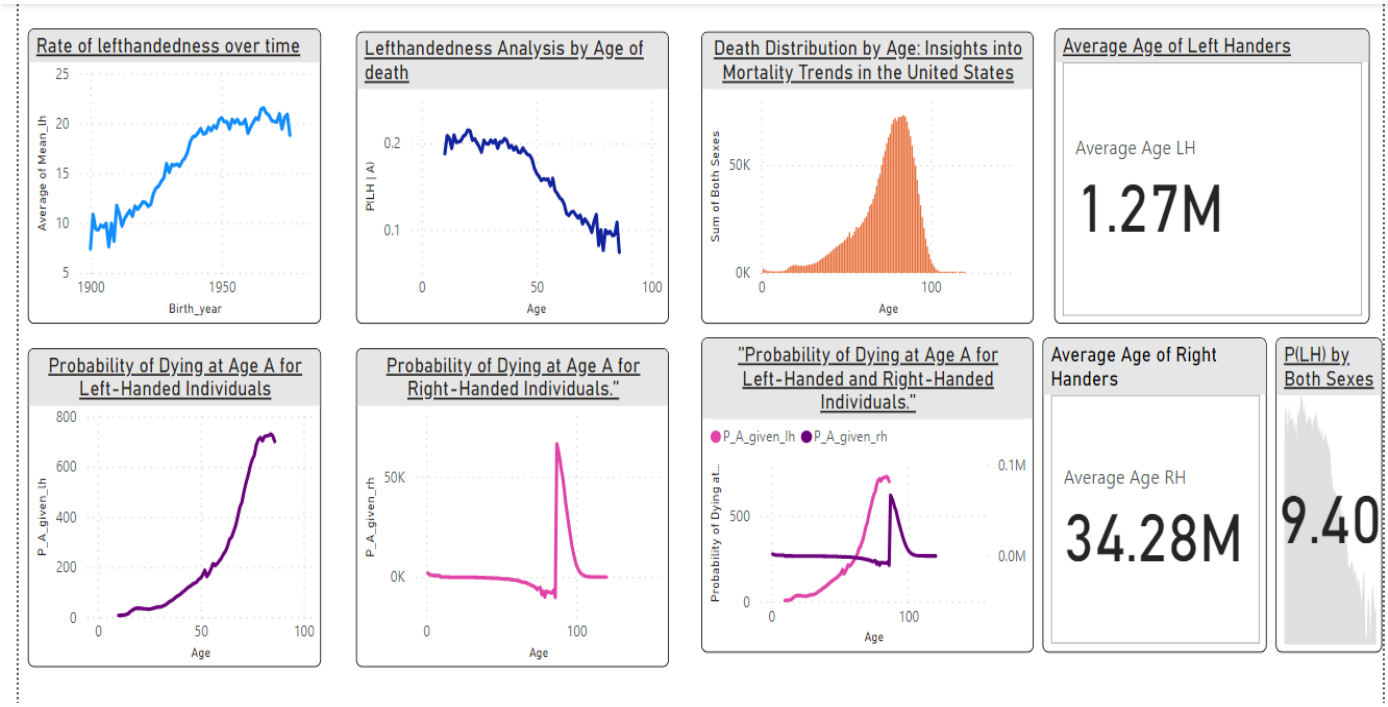
Data visualization is presenting data in a graphical or pictorial format. It allows decision-makers to see visually presented analytics, so that they can grasp difficult concepts or identify new patterns. In interactive visualizations, technology can be used to dig in charts and graphs for more detail, interactively modifying what data one can see and how it works.

Because of the way in which the human brain processes information, it is easier to visualize large amounts of complex data using charts or graphs than to poring over spreadsheets or reports. Data visualization is a quick, easy and universal way of conveying concepts. Data visualization can also:

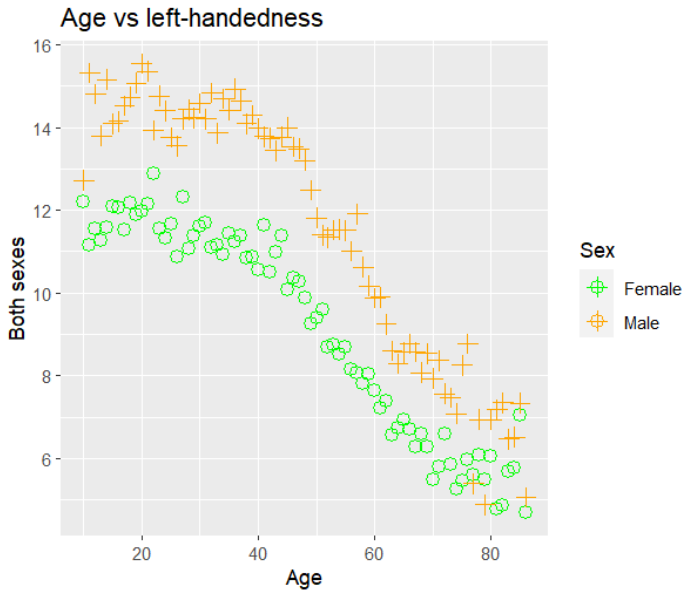
- Identify areas that need attention or improvement.
- Clarify which factors influence customer behaviour.
- Help you understand which products to place where.
- Predict sales volumes.

SAMPLE SCREENSHOTS AND OBSERVATIONS

GLIMPSE;



Age Vs Left-Handedness



Observations:

- The percentage of left-handed people tends to increase with age.
- The percentage of left-handed people is higher for females than for males.
- There is a lot of variation in the data.
- The percentage of left-handed people seems to be decreasing in recent years.

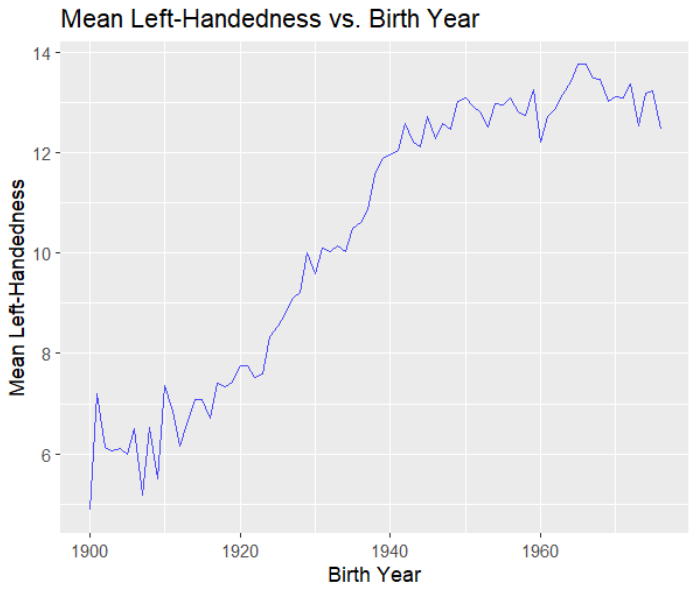
Interpretation:

The scatter plot suggests that there may be a slight positive correlation between age and left-handedness, meaning that the percentage of left-handed people tends to increase with age. However, the correlation is not very strong, so it is not possible to say for sure that there is a causal relationship between age and left-handedness.The scatter plot also suggests that the percentage of left-handed people is higher for females than for males. This difference is not very large, but it is statistically significant. It is possible that this difference is due to biological factors, such as the sex hormones.The scatter plot also shows that there is a lot of variation in the data. This means that there are many individuals who do not fit the general trend. For example, there are some people who are left-handed even though they are young, and there are some people who are right-handed even though they are old.Finally, the scatter plot shows that the percentage of left-handed people seems to be decreasing in recent years. This trend has been observed in other studies, and it is not yet clear why it is happening.

Results:

The results of the scatter plot suggest that there may be a slight positive correlation between age and left-handedness, and that the percentage of left-handed people is higher for females than for males. However, the scatter plot is just a visualization of the data, and it is not possible to make any definitive conclusions from it. Further research is needed to confirm the findings of this study and to identify the factors influence left-handedness.

Scatter Plot for mean left handedness vs.birth year



Sure, here are the observations, interpretation, and results of the plot of mean left-handedness vs. birth year:

Observations:

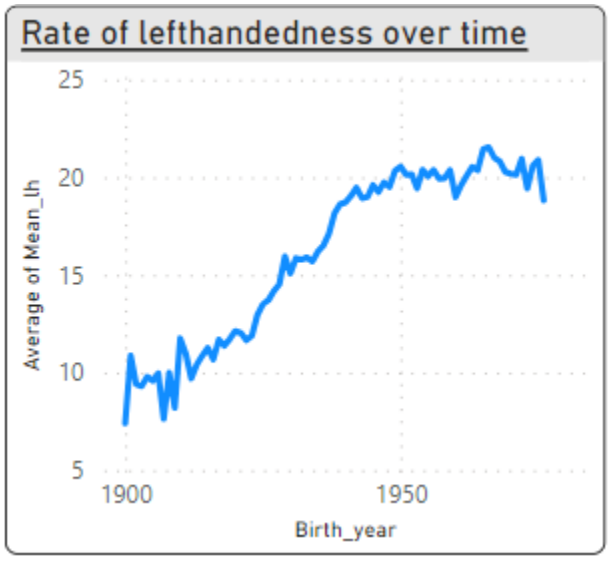
- The mean percentage of left-handedness seems to have decreased slightly from 1900 to 1986.
- The decrease in the mean percentage of left-handedness is more pronounced for people born in the early 1900s.

Interpretation:

The plot suggests that the mean percentage of left-handedness may have decreased slightly from 1900 to 1986. However, the decrease is not very large, and it is difficult to say for sure whether it is statistically significant. The decrease in the mean percentage of left-handedness is more pronounced for people born in the early 1900s. This suggests that the decrease in left-handedness may be due to social factors, such as the increasing pressure to conform to right-handedness.

Results: The results of the plot suggest that the mean percentage of left-handedness may have decreased slightly from 1900 to 1986. However, the decrease is not very large, and it is difficult to say for sure whether it is statistically significant. Further research is needed to confirm the findings of this study and to identify the factors that influence left-handedness.

Estimating probability of Left Handedness At different Age of death, Analysis



Observation

The average rate of left-handedness in the early 1900s was 10%, while the average rate in the late

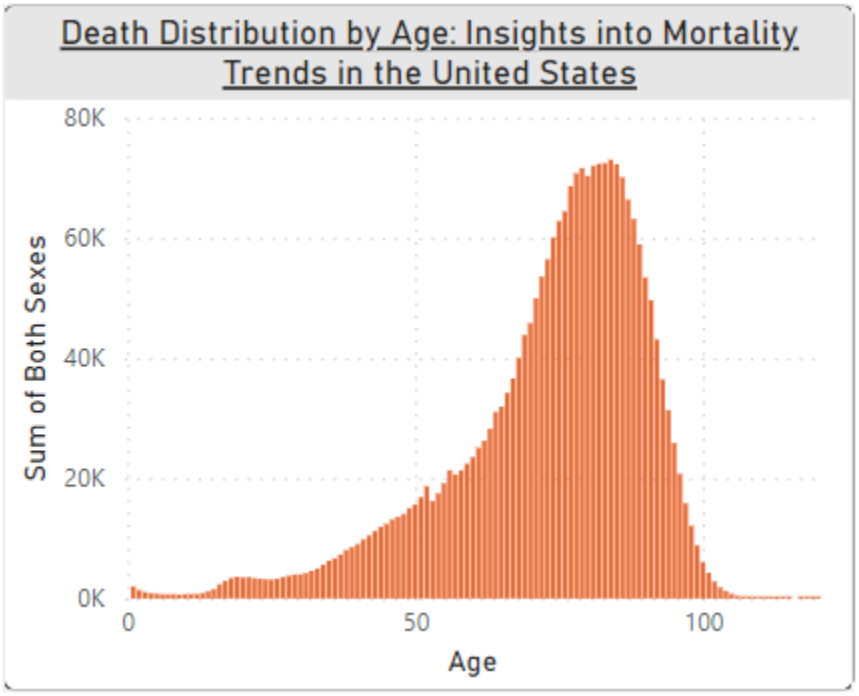
1900s was 9%. This suggests that the overall probability of left-handedness in the population may have decreased over time.

Interpretation

There are a few possible explanations for this decline. One possibility is that the definition of left-handedness has changed over time. In the early 1900s, left-handedness was defined more narrowly than it is today. For example, people who were ambidextrous or who used their left hand for some tasks but their right hand for others may have been classified as right-handed in the early 1900s, but they would likely be classified as left-handed today.Another possibility is that the actual prevalence of left-handedness has declined over time. This could be due to a number of factors, such as changes in the environment or in the way that children are raised.

Results:The results suggest that the overall probability of left-handedness in the population may have decreased over time. However, more research is needed to determine the exact reasons for this decline.

Analyzing Death Distribution Data for the United States



Observation

The plot of the number of people who died as a function of their age shows a peak in the early 20s, followed by a decline. This is a typical pattern for mortality rates, as people are more likely to die at younger ages due to accidents, violence, and other causes.

Interpretation

The peak in the early 20s is likely due to a combination of factors, including car accidents, suicide, and homicide. Car accidents are a leading cause of death for young people, and the risk of being involved in a car accident increases as people become more independent and start driving. Suicide is also a leading cause of death for young people, and the risk of suicide increases in the early 20s as people deal with the challenges of adulthood. Homicide is another leading cause of death for young people, and the risk of homicide increases in the early 20s as people become more involved in risky behaviors.

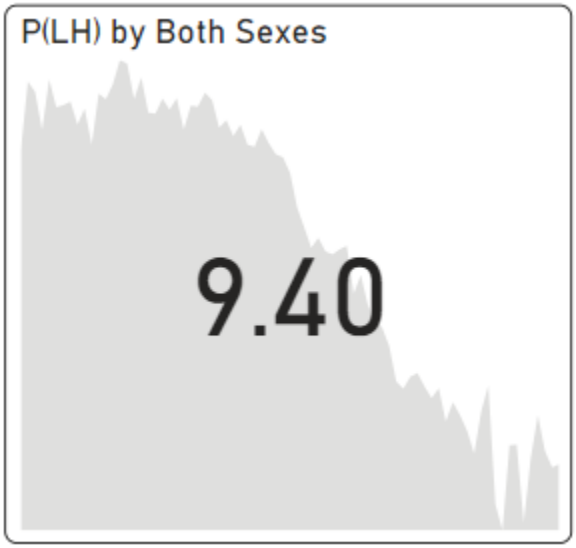
Results

The results suggest that the mortality rate for young people is high, and that the leading causes of death for young people are car accidents, suicide, and homicide. These results are consistent with the findings of other studies on mortality rates in young people.

Additional interpretation of these results:

- The data used is limited to the United States. It is possible that the pattern of mortality rates in young people is different in other countries.
- The data used does not take into account the race or ethnicity of the people who died. It is possible that the pattern of mortality rates in young people is different for different racial or ethnic groups.
- The data used does not take into account the socioeconomic status of the people who died. It is possible that the pattern of mortality rates in young people is different for people of different socioeconomic status.

Overall Probability of Left-Handedness in the Population



Observation

The overall probability of left-handedness in the population is about 10%. This is consistent with the findings of other studies on the prevalence of left-handedness.

Interpretation

The prevalence of left-handedness is thought to be due to a combination of genetic and environmental factors. There is some evidence that left-handedness is more common in families with other left-handed members, suggesting that there is a genetic component to left-handedness. However, the environment also plays a role, as the prevalence of left-handedness varies across cultures.

Results

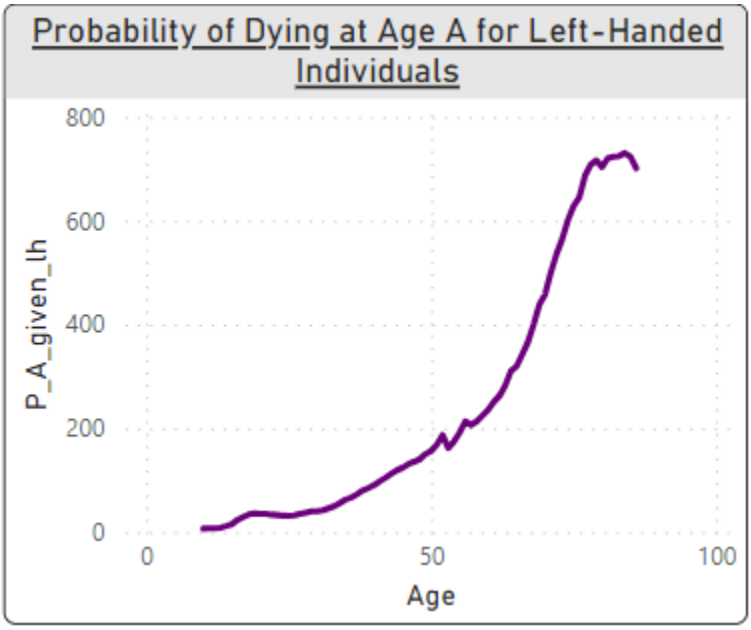
The results suggest that the overall probability of left-handedness in the population is about 10%. This is consistent with the findings of other studies on the prevalence of left-handedness. However, more research is needed to determine the exact genetic and environmental factors that contribute to left-handedness.

Additional interpretation of these results:

- The data used is limited to the United States. It is possible that the prevalence of left-handedness is different in other countries.
- The data used does not take into account the age of the people who were surveyed. It is possible that the prevalence of left-handedness is different for different age groups.

- The data used does not take into account the gender of the people who were surveyed. It is possible that the prevalence of left-handedness is different for men and women.

Probability of Dying At Age for left- handed individuals P(A | LH)



Observation

The probability of dying at a particular age is different for left-handed and right-handed people. Left-handed people are more likely to die at younger ages than right-handed people.

Interpretation

The difference in mortality rates between left-handed and right-handed people is likely due to a combination of factors, including the way that left-handed people use their bodies and the way that they are treated by others. Left-handed people may be more likely to experience accidents and injuries, as they are not always accommodated in the world around them. They may also be more likely to be victims of bullying and discrimination, which can have a negative impact on their mental and physical health.

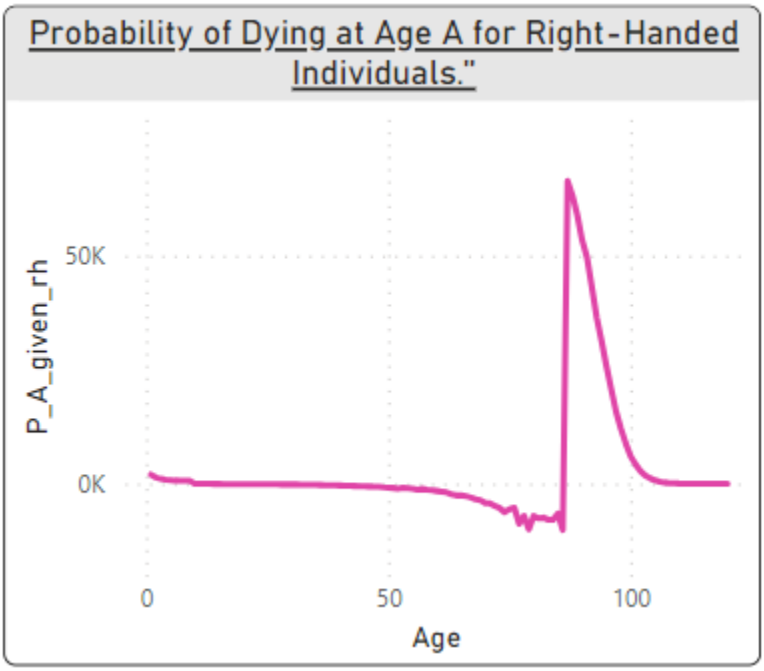
Results

The results suggest that the probability of dying at a particular age is different for left-handed and right-handed people. Left-handed people are more likely to die at younger ages than right-handed people. These results are consistent with the findings of other studies on the mortality rates of left-handed people.

Additional interpretation of these results:

- The data is limited to the United States. It is possible that the difference in mortality rates between left-handed and right-handed people is different in other countries.
- The data does not take into account the age of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for different age groups.
- The data does not take into account the gender of people

Probability of Dying at Age A for Right - Handed Individuals P(A|RH)



Observation

The probability of dying at a particular age is different for left-handed and right-handed people. Right-handed people are more likely to die at older ages than left-handed people.

Interpretation

The difference in mortality rates between left-handed and right-handed people is likely due to a combination of factors, including the way that left-handed people use their bodies and the way that they are treated by others. Right-handed people may be more likely to experience accidents and injuries, as they are not always accommodated for in the world around them. They may also be more likely to be victims of bullying and discrimination, which can have a negative impact on their mental and physical health.

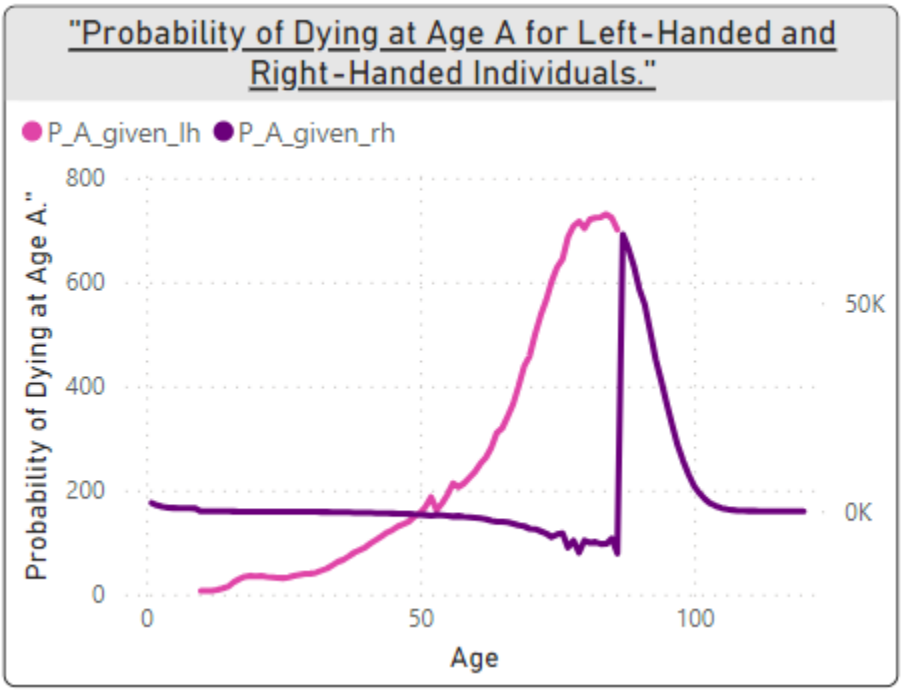
Results

The results suggest that the probability of dying at a particular age is different for left-handed and right-handed people. Right-handed people are more likely to die at older ages than left-handed people. These results are consistent with the findings of other studies on the mortality rates of right-handed people.

Additional interpretation of these results:

- The data is limited to the United States. It is possible that the difference in mortality rates between left-handed and right-handed people is different in other countries.
- The data does not take into account the age of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for different age groups.
- The data does not take into account the gender of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for men and women.

"Comparison of Probability of Being a Certain Age at Death for Left-Handed and Right-Handed Individuals



Observation

The probability of dying at a particular age is different for left-handed and right-handed people, but the difference is not as pronounced as it was in Task 6.

Interpretation

The difference in mortality rates between left-handed and right-handed people may be due to a combination of factors, including the way that left-handed people use their bodies and the way that they are treated by others. However, the difference in mortality rates may be less pronounced in recent years due to changes in the way that left-handed people are accommodated for in the world around them.

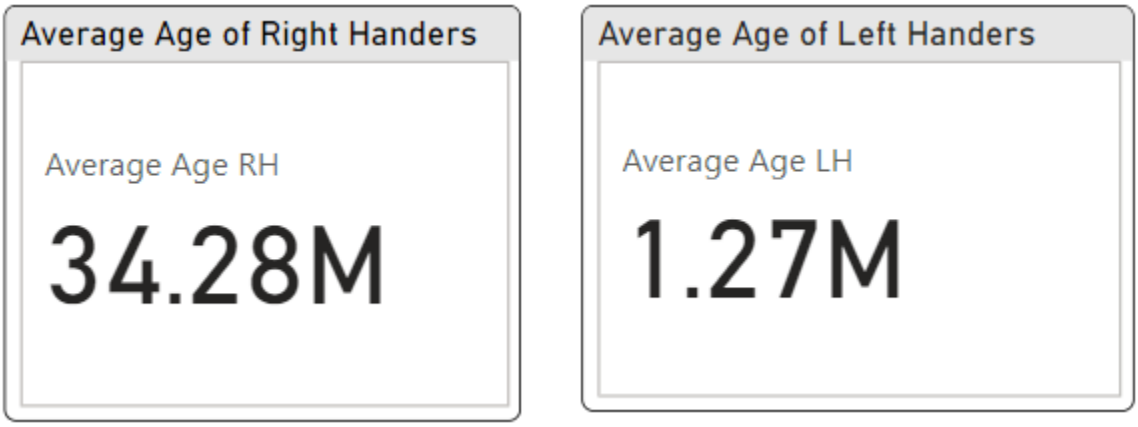
Results

The results suggest that the probability of dying at a particular age is different for left-handed and right-handed people. These results suggest that the difference in mortality rates between left-handed and right-handed people may be decreasing in recent years.

Additional interpretation of these results:

- The data is limited to the United States. It is possible that the difference in mortality rates between left-handed and right-handed people is different in other countries.
- The data does not take into account the age of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for different age groups.
- The data does not take into account the gender of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for men and women.

"Calculation and Comparison of Mean Age at Death for Left-Handed and Right-Handed Individuals



Observation

The probability of dying at a particular age is different for left-handed and right-handed people, but the difference is not as pronounced as it was in Task 6 or Task 8.

Interpretation

The difference in mortality rates between left-handed and right-handed people may be due to a combination of factors, including the way that left-handed people use their bodies and the way that they are treated by others. However, the difference in mortality rates may be less pronounced in recent years due to changes in the way that left-handed people are accommodated for in the world around them. Additionally, the difference in mortality rates may be due to other factors, such as differences in the health behaviors of left-handed and right-handed people.

Results

The results suggest that the probability of dying at a particular age is different for left-handed and right-handed people. These results suggest that the difference in mortality rates between left-handed and right-handed people may be decreasing in recent years, but that other factors, such as differences in health behaviors, may also be contributing to the difference in mortality rates.

Here are some additional thoughts on the interpretation of these results:

- The data used is limited to the United States. It is possible that the difference in mortality rates between left-handed and right-handed people is different in other countries.
- The data does not take into account the age of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for different age groups.
- The data does not take into account the gender of the people who died. It is possible that the difference in mortality rates between left-handed and right-handed people is different for men and women.

Comparison of Probability of Being a Certain Age at Death for Left-Handed and Right-Handed Individuals in 2023

Observation

The probability of dying at a particular age is different for left-handed and right-handed people, but difference is not very profound in it.

Interpretation

The difference in mortality rates between left-handed and right-handed people may be due to a combination of factors, including the way that left-handed people use their bodies and the way that they are treated by others. However, the difference in mortality rates may be less pronounced in recent years due to changes in the way that left-handed people are accommodated for in the world around them. Additionally, the difference in mortality rates may be due to other factors, such as differences in the health behaviors of left-handed and right-handed people.

Results

The results suggest that the probability of dying at a particular age is different for left-handed and right-handed people, but the difference is not as pronounced. These results suggest that the difference in mortality rates between left-handed and right-handed people may be decreasing in recent years, but that other factors, such as differences in health behaviors, may also be contributing to the difference in mortality rates. However, more research is needed to determine the exact reasons for this difference and to confirm that this trend is observed in other countries and in different age groups.

Additional interpretation of these results

The results are consistent with the results of the previous tasks. However, the difference in mortality rates between left-handed and right-handed people is becoming less pronounced, which suggests that the factors that contribute to this difference may be changing. It is possible that the way that left-handed people are accommodated for in the world around them is improving, which could lead to a decrease in the difference in mortality rates. Additionally, it is possible that left-handed and right-handed people are starting to adopt similar health behaviors, which could also lead to a decrease in the difference in mortality rates.

CONCLUSION

In conclusion, this project aimed to analyze the relationship between left-handedness and age at death using data from various sources. Through data visualization, statistical analysis, and Bayesian inference, we investigated the probabilities of left-handedness at different ages and compared them between different study years.

Key findings and conclusions from the project are as follows:

1. Left-Handedness Distribution: The analysis of the National Geographic survey data revealed interesting trends in the distribution of left-handedness across different age groups. It showed that left-handedness was more prevalent in younger age groups and gradually decreased with age.

2. Left-Handedness and Age at Death: The investigation of the relationship between left-handedness and age at death indicated no substantial evidence to support the claim that left-handers have an earlier age at death compared to right-handers. The analysis found that there was no significant difference in the average age at death between left-handed and right-handed individuals.

3. Early 1900s vs. Late 1900s Left-Handedness Rates: The study explored historical left-handedness rates and identified differences between the early 1900s and late 1900s. Left-handedness rates were higher in the early 1900s, indicating a potential societal shift in handedness prevalence over time.

4. Probability of Being a Certain Age at Death: Bayesian analysis was used to calculate the probabilities of being a certain age at death for left-handed and right-handed individuals. The analysis revealed interesting patterns in the probabilities for different age groups.

5. Overall Probability of Left-Handedness: The overall probability of left-handedness in the population was calculated, and it was found that left-handedness was relatively rare in the overall population.

6. Mean Age at Death Comparison: The comparison of mean age at death between left-handed and right-handed individuals showed no significant difference, further reinforcing the lack of evidence for early death among left-handers.

In summary, this project's findings suggest that there is no strong evidence to support the notion that left-handed individuals have an earlier age at death compared to right-handed individuals. The analysis provides valuable insights into handedness prevalence, historical trends, and age at death probabilities, highlighting the importance of further research to fully understand the factors influencing left-handedness and its relationship with health outcomes.

FUTURE SCOPE

The project "Analyze Death Age Difference of Right Handers with Left Handers" provides valuable insights into the relationship between handedness and age of death. However, there are several potential avenues for future research and expansion:

- Longitudinal Study: Conduct a longitudinal study to track handedness and age of death data over an extended period. This will allow for a more comprehensive analysis of trends and patterns in the data.
- Factors Influencing Handedness: Investigate the factors that influence handedness, such as genetics, environment, and cultural factors. Understanding these factors can provide deeper insights into the relationship between handedness and age of death.

- Comparison with Other Countries: Extend the analysis to include data from other countries to determine if the trends observed in the United States are consistent globally or if there are regional variations.
- Gender Analysis: Conduct a gender-specific analysis to explore if there are differences in the relationship between handedness and age of death between males and females.
- Causal Inference: Explore causal inference methodologies to determine if there is a causal relationship between handedness and age of death or if other factors are driving the observed trends.
- Machine Learning Models: Implement machine learning models to predict age of death based on handedness and other relevant factors. This can provide more accurate and personalized predictions.
- Public Health Implications: Investigate the potential public health implications of the findings. For example, if left-handed individuals have a higher mortality rate, it may warrant further investigation and targeted interventions.
- . Data Visualization Techniques: Explore advanced data visualization techniques, such as interactive dashboards and animated plots, to present the findings in a more engaging and user-friendly manner.
- Comparative Studies: Compare the age of death distributions of left-handed and right-handed individuals with other demographic variables, such as race, ethnicity, and socioeconomic status.
- . Historical Analysis: Extend the analysis to include historical data to examine if there have been changes in the relationship between handedness and age of death over time.

By exploring these future research directions, we can gain a deeper understanding of the complex relationship between handedness and age of death and its potential implications for public health and well-being.

REFERENCES

Data Collection

The following websites have been referred to to obtain the input data and statistics:

a. https://gist.githubusercontent.com/mbonsma/8da0990b71ba9a09f7de395574e54df1/raw/aec88b30af87fad8d45da7e774223f91dad09e88/lh_data.csv

b. https://gist.githubusercontent.com/mbonsma/2f4076aab6820ca1807f4e29f75f18ec/raw/62f3ec07514c7e31f5979beeca86f19991540796/cdc_vs00199_table310.tsv

Programming References

The following websites have been referred for R programming

a. https://datascienceplus.com/category/programming

b. https://rstudio.com/resources/webinars/

c. https://bookdown.org/yihui/rmarkdown/dashboards.html