**HEALTH EXPENDITURE AND SUICIDE RATES**

By DEVIKA JAGARLAMUDI

**INTRODUCTION:**

The statement highlights two important public health concerns: healthcare expenditure and suicide rates. Healthcare expenditure refers to the resources used to provide healthcare services, including medical personnel, facilities, equipment, drugs, and other associated costs. Suicide rates, on the other hand, are an indicator of mental health and well-being. Studies have shown a correlation between healthcare expenditure and suicide rates, with higher healthcare expenditure associated with lower suicide rates and vice versa. This suggests that investing in healthcare services and resources can have a positive impact on mental health and well-being, potentially leading to lower suicide rates.

There are several reasons why this correlation may exist. Firstly, increased healthcare expenditure can lead to better access to mental health services, including counselling, therapy, and medication. This can help individuals who are struggling with mental health concerns, such as depression or anxiety, to receive the support they need to manage their symptoms and improve their overall mental health. Secondly, higher healthcare expenditure can lead to better overall health outcomes, which can have a positive impact on mental health. For example, improved access to preventative care, such as vaccinations, can help to reduce the incidence of infectious diseases, which can cause stress and anxiety. Similarly, better management of chronic conditions, such as diabetes or heart disease, can improve overall health and reduce the risk of complications, which can also contribute to better mental health outcomes.

Policymakers and researchers must focus on this relationship between healthcare expenditure and suicide rates to promote better mental health and reduce suicide rates. This may involve increasing funding for mental health services, improving access to healthcare services, and addressing social factors that may contribute to poor mental health outcomes, such as poverty, unemployment, and social isolation. By investing in healthcare services and resources, policymakers can help to improve mental health outcomes and reduce the incidence of suicide.

**DATASET:**

The World Health Organization's Mortality Statistics and uses the International

Classification of Diseases, 10th Revision (ICD-10) codes to classify causes of death. The WHO Mortality Database is a comprehensive source of mortality data from member countries, including information on causes of death, age, sex, and other demographic characteristics. The ICD-10 is a widely used system for classifying diseases and health conditions, which allows for consistent and standardized recording and reporting of causes of death across different countries and regions. This dataset, obtained from Kaggle, contains information on mortality rates and various health indicators across different countries and regions, making it a valuable resource for researchers, policymakers, and other stakeholders interested in understanding trends and patterns in global health outcomes.

Link for the dataset: [https://www.kaggle.com/datasets/fernandoretamales/healthexpenditure-and-suiciderates](https://www.kaggle.com/datasets/fernandoretamales/health-expenditure-and-suicide-rates)

* Country - Country names
* Year- Years
* Population- Number of Population
* Deaths\_All\_Types- Total Deaths
* Deaths\_Suicides-Deaths by suicides according to ICD10
* HExp\_Pctage\_Y- Health Expenditure as % of GDP
* MHExp\_Pctage\_20-Mental Health Expenditure index in 2011
* Dep\_Num\_2015-Depression estimate index in 2015
* Suicide\_p100- Number of suicides per 100,000 population

**LANGUAGE & TOOLS:**

**Language:**

* Python: Python is a programming language that was created in 1991 and is now widely used in various fields such as data science, machine learning, web development, and scientific computing. Being an interpreted language, it is easy to learn and write code, making it a popular choice for beginners and experts alike.

**Tools:**

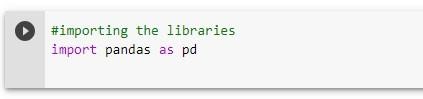
* Google Collab: Google Collaboratory, also known as Google Collab, is a free cloud-based platform provided by Google that enables users to write, run and share Python code through a web browser interface. It eliminates the need for installing any additional software or setting up a local machine. It works as a Jupyter notebook environment and provides a hassle-free way of collaborating on code.
* Tableau: Tableau is a software used for data visualization and business intelligence that enables users to easily connect, visualize, and share data. It is widely used by individuals, businesses, and organizations to analyse and visualize data in different formats, including spreadsheets, databases, and cloud services, in a user-friendly manner.

**Libraries:**

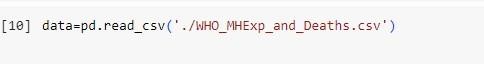
* NumPy: NumPy, or Numerical Python, is a Python library commonly used for scientific computing and data analysis. It offers a powerful N-dimensional array object, which is used to efficiently compute arrays and matrices, making it a popular choice for mathematical operations and data analysis.
* pandas: Pandas is a widely used open-source Python library that is utilized for data manipulation and analysis. It offers data structures that enable efficient storage and manipulation of large datasets, along with tools for cleaning, merging, and reshaping data. It is a versatile tool that is useful for a wide range of data-related tasks.

**DATA PREPROCESSING:**

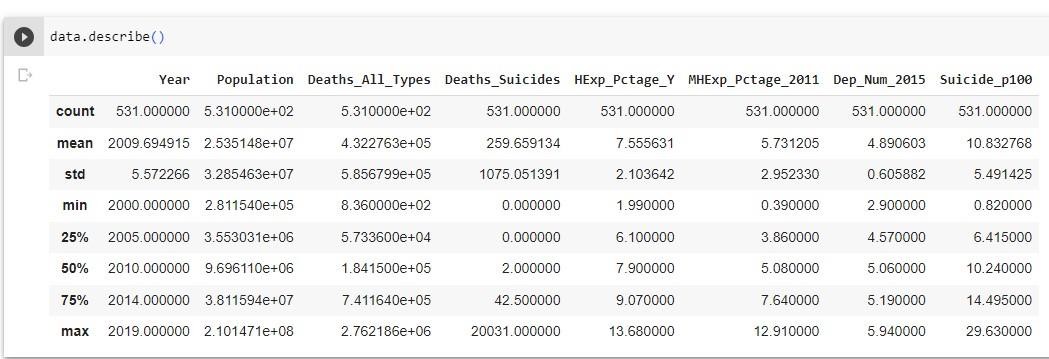
Step 1: In most programming languages, libraries are collections of pre-written code that developers can use to perform common tasks without having to write the code from scratch. Importing a library means that you are making the functions, classes, or other code contained within that library available to your program. When you import a library in your program, you are telling the compiler or interpreter to load that library into the program's memory so that the code contained within it can be executed. This allows you to use the functions and classes defined in the library in your own code. Importing the required libraries like pandas as shown below.



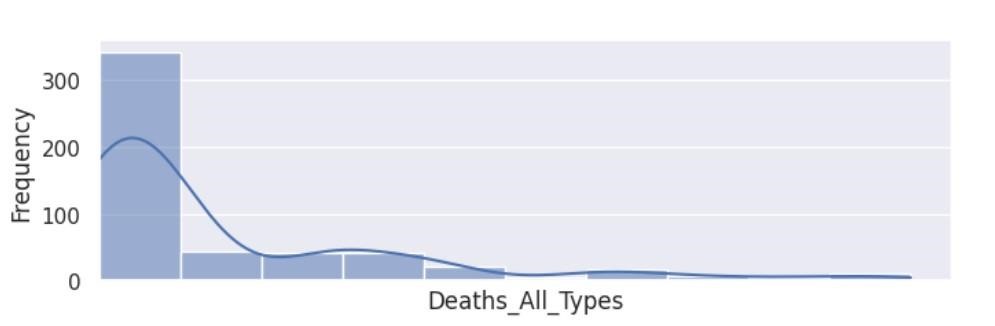
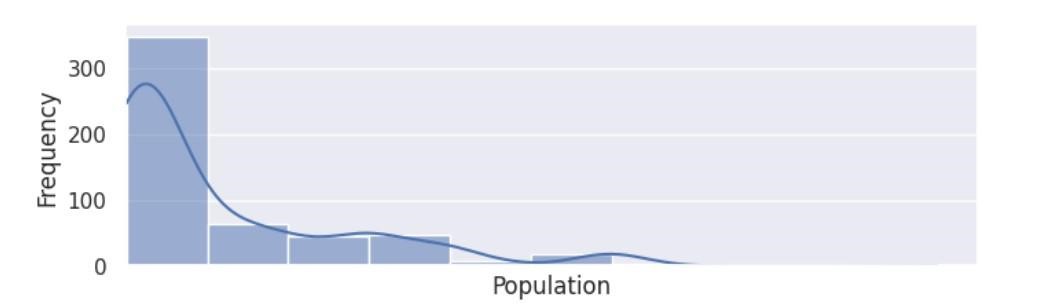
Step 2: read\_csv is a function provided by the popular Python library Pandas for reading data from a CSV (comma-separated values) file into a Pandas Data Frame. A data frame is a 2-dimensional tabular data structure commonly used for data analysis and manipulation. Importing the CSV file as shown below.

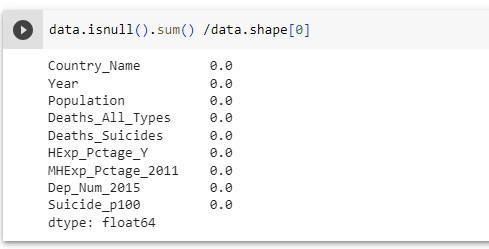


Step 3: Descriptive analysis is a statistical analysis technique used to describe and summarize a dataset's main features. This type of analysis is often used as a preliminary step in data analysis to gain a better understanding of the dataset and to identify any patterns or trends that may be present. Check for the descriptive analysis of the dataset.

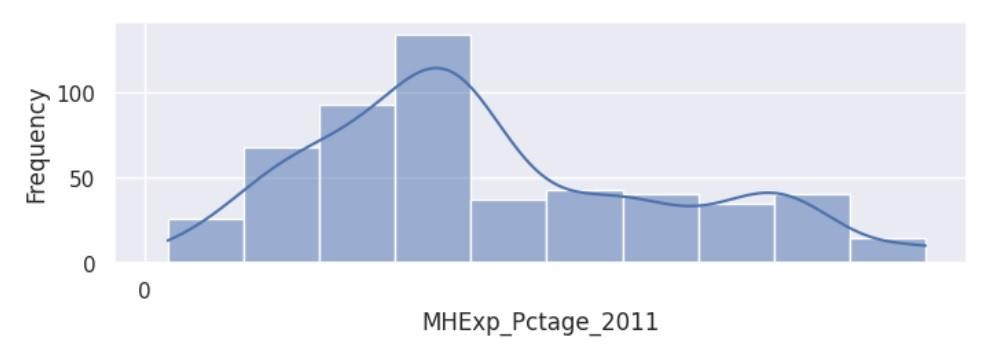


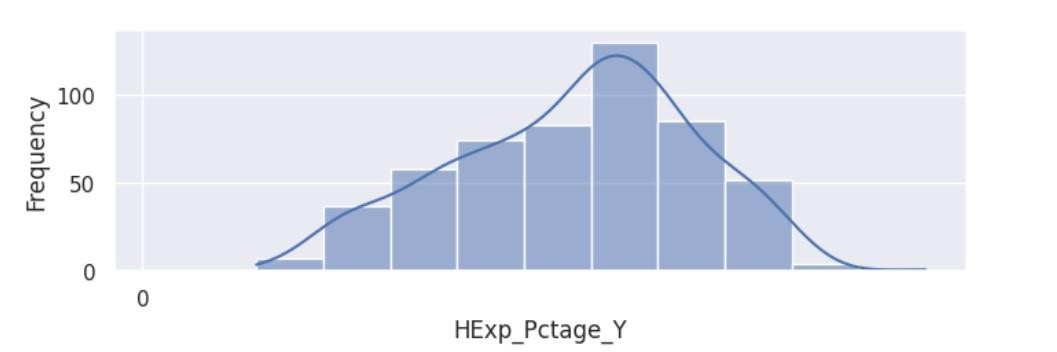
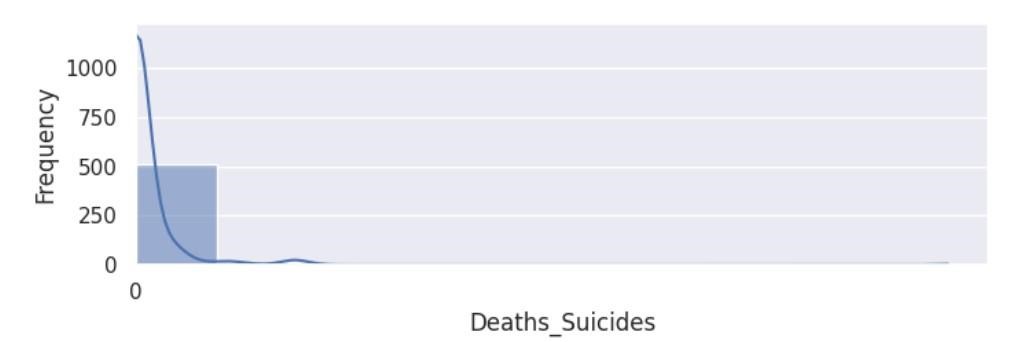
Step 4: Checking for null values is a major step in data analysis and data cleaning. Null values, also known as missing values, are values in a dataset that are undefined or unknown. They can occur for distinct reasons, such as data entry errors, data processing issues, or incomplete data. Check for the null values in the dataset as shown below.

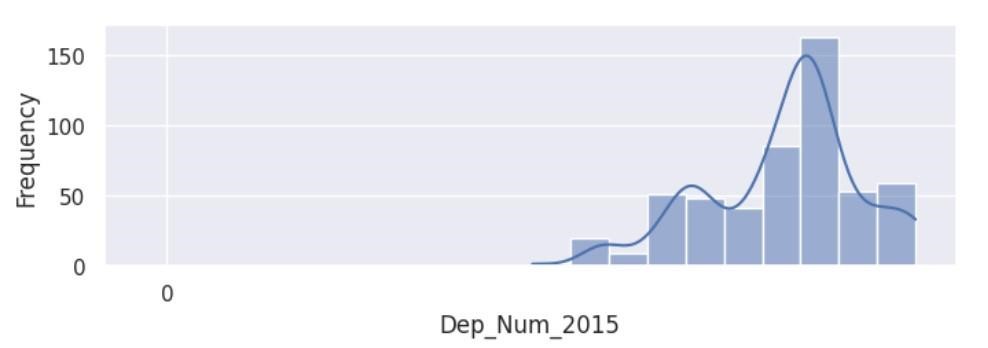




Step 5: Outliers are data points that are significantly different from other data points in a dataset. They can be caused by a variety of factors, such as measurement errors, data processing errors, or true anomalies in the data. Outliers can have a significant impact on data analysis and modeling, as they can skew the distribution of the data and affect statistical analyses or machine learning models. You can see the outliers below for the dataset.





Step 6: As there are no null values in the dataset, we are good to perform the visualization in Tableau.

**Graphs Used:**

I have chosen geo maps, tree maps, line chart, bar chart, Area chart to analyses the results.

**Geographical maps:** A geographical map in Tableau is a type of data visualization that displays data on a map, with the data points plotted at their corresponding latitude and longitude coordinates.

**Area chart:** An area chart is a type of data visualization in which data points are connected by a line and the area between the line and the x-axis is filled with colour.

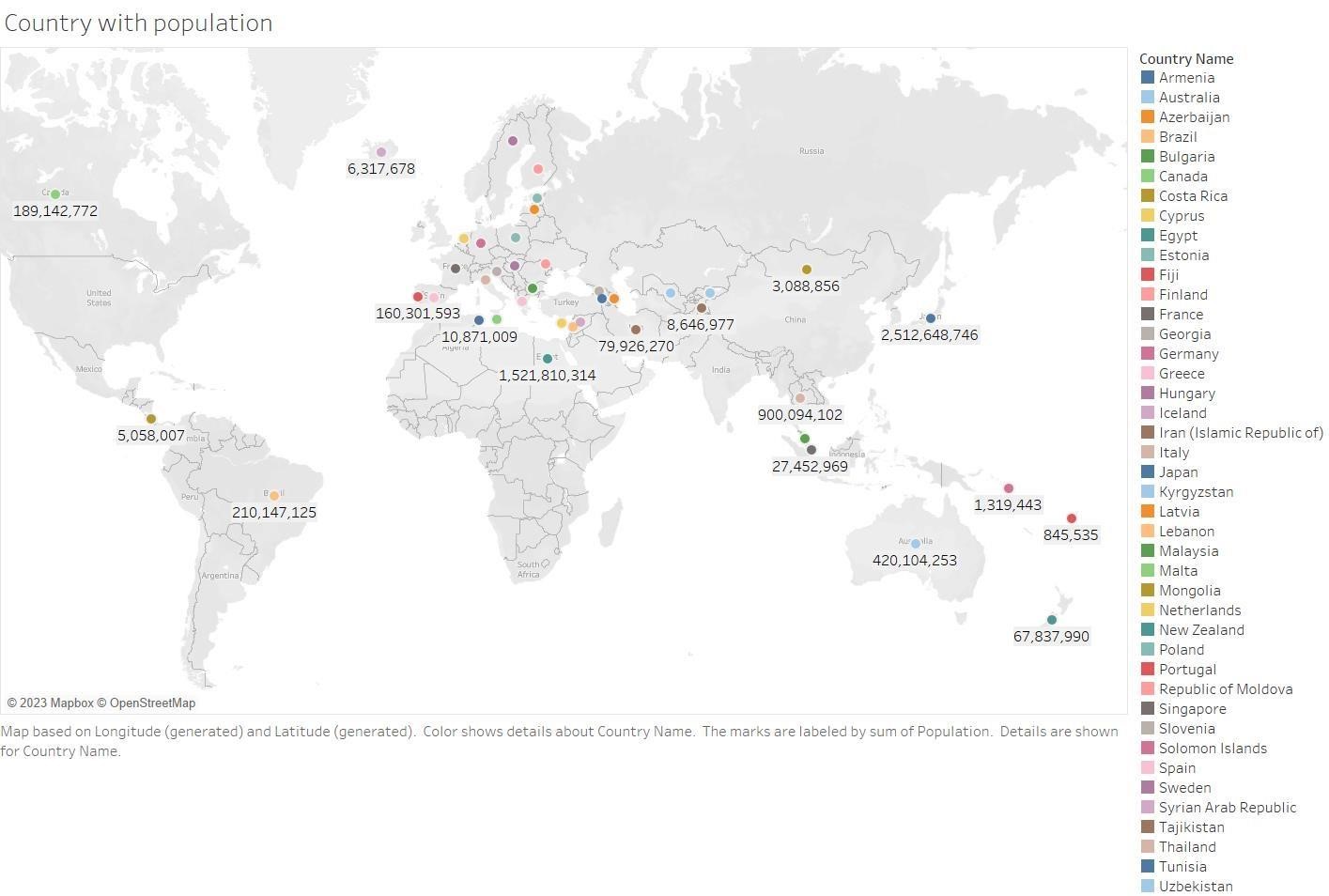
**Line chart:** A line chart is a type of data visualization that displays data as a series of points connected by a line.

**Bar chart:** A bar chart is a type of data visualization that displays data as rectangular bars, with the length of each bar proportional to the value of the data it represents.

**Tree map:** A tree map is a type of data visualization that displays hierarchical data using nested rectangles, with each rectangle representing a specific category or subcategory of the data.

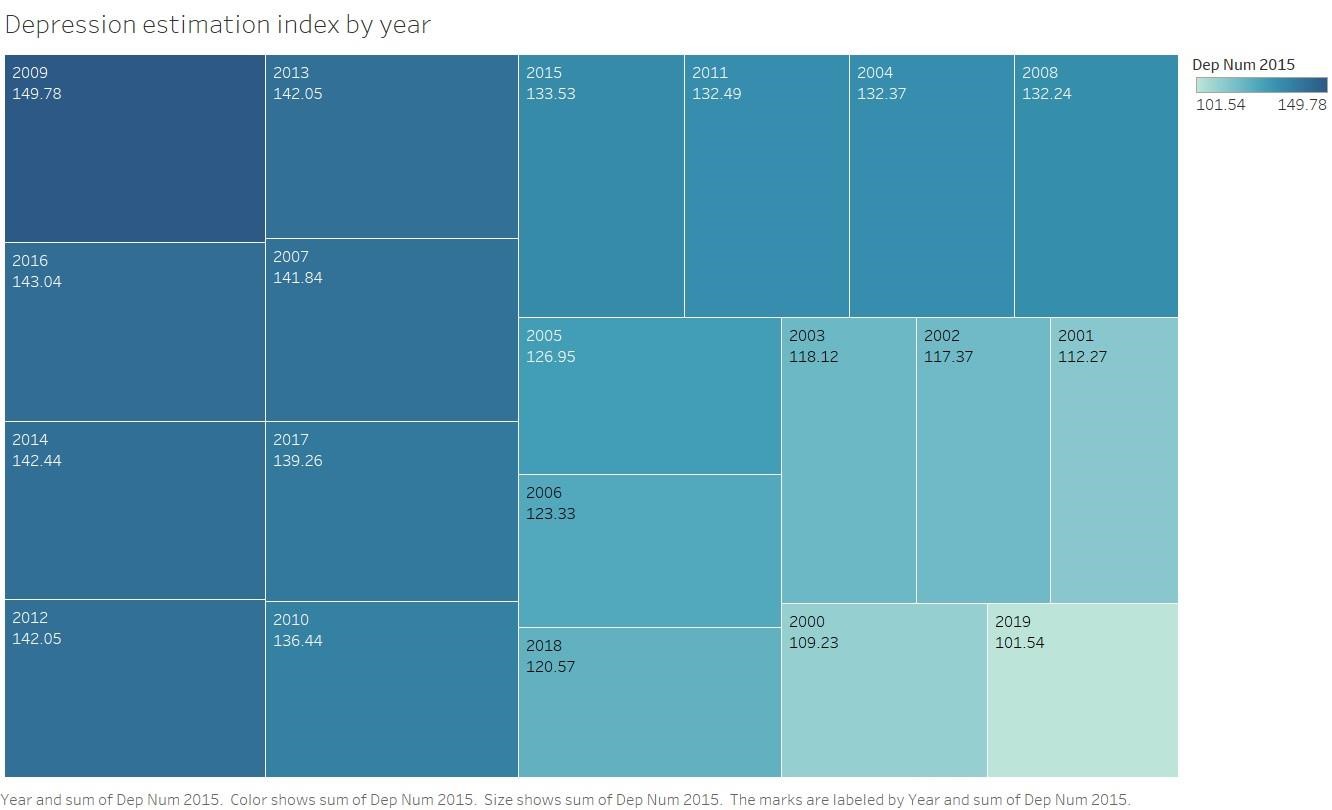
**Hypothesis:**

**1.Which country has the highest population?**



**Inference:** This graph shows you the population in the country **Result:** Brazil is having the highest population with 210147125 counts.

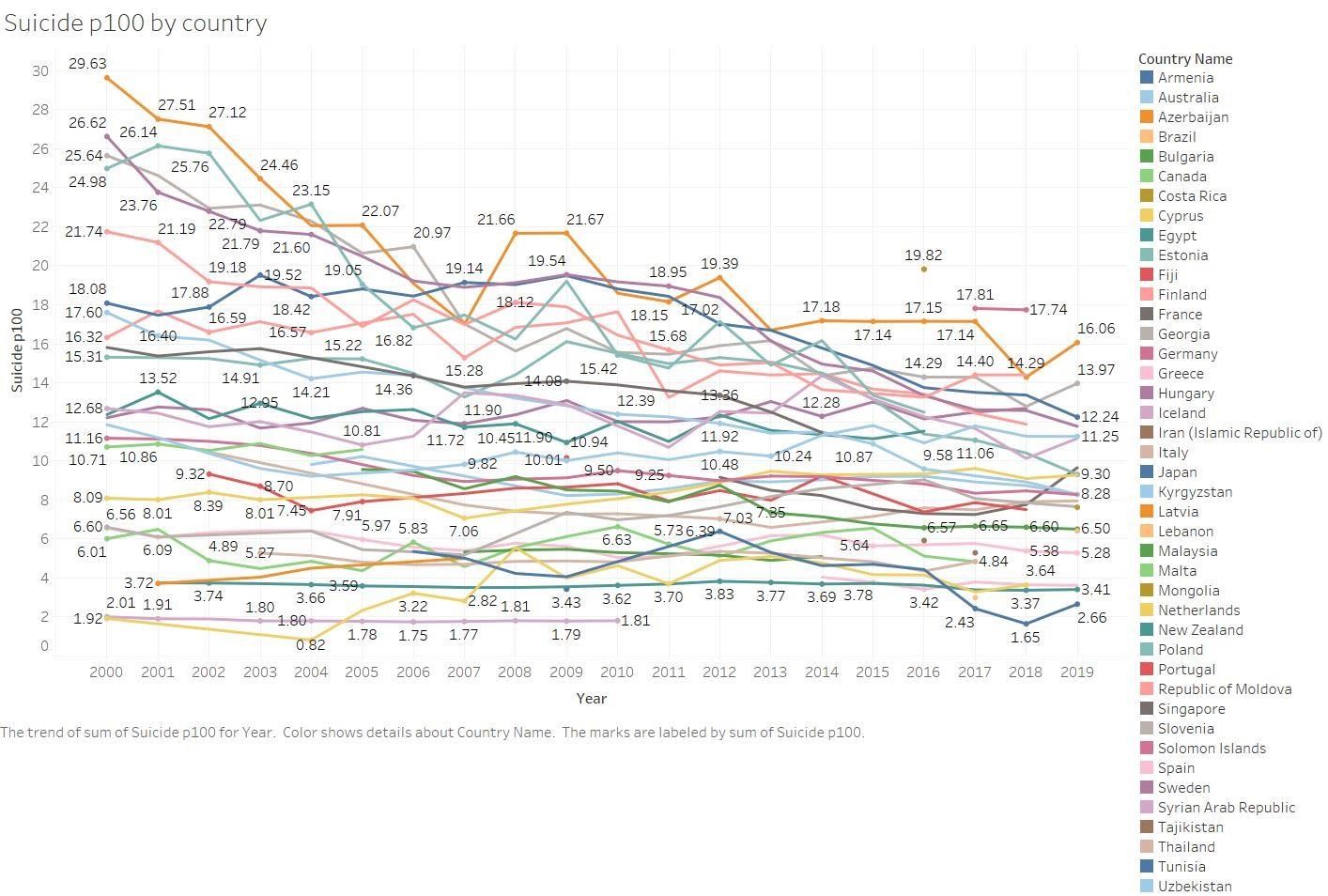
**2.In which year is the depression estimation index the highest and lowest?**



**Inference:** This line graph shows you the suicide in the different year in the different countries

**Results:** The highest depression estimates is in the year 2009 with 149.78 The lowest depression estimates are in the year 2019 with 101.54.

**3. In which year suicide is the least and highest and in which country it is?**



**Inference:**

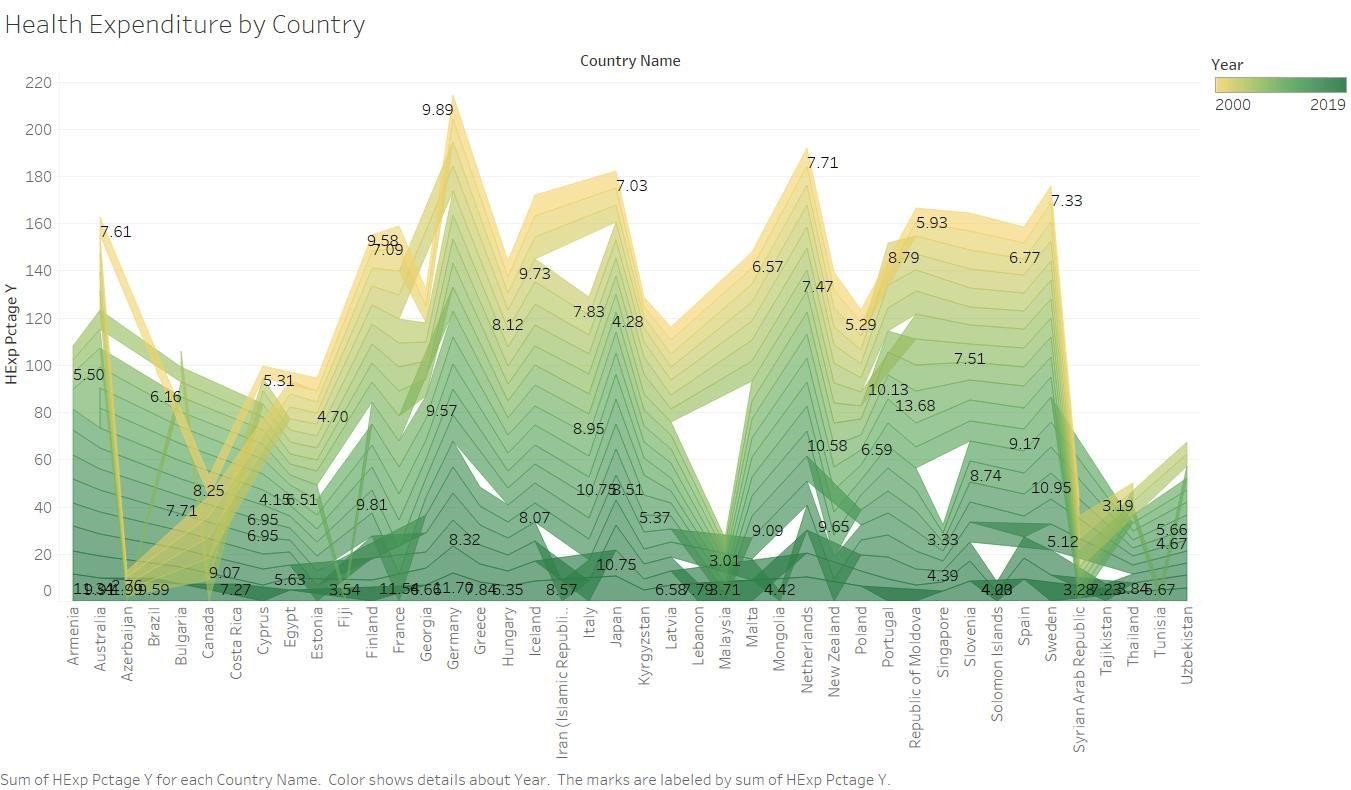
This graph shows you the depression estimation index in the different

years.

**Result:**

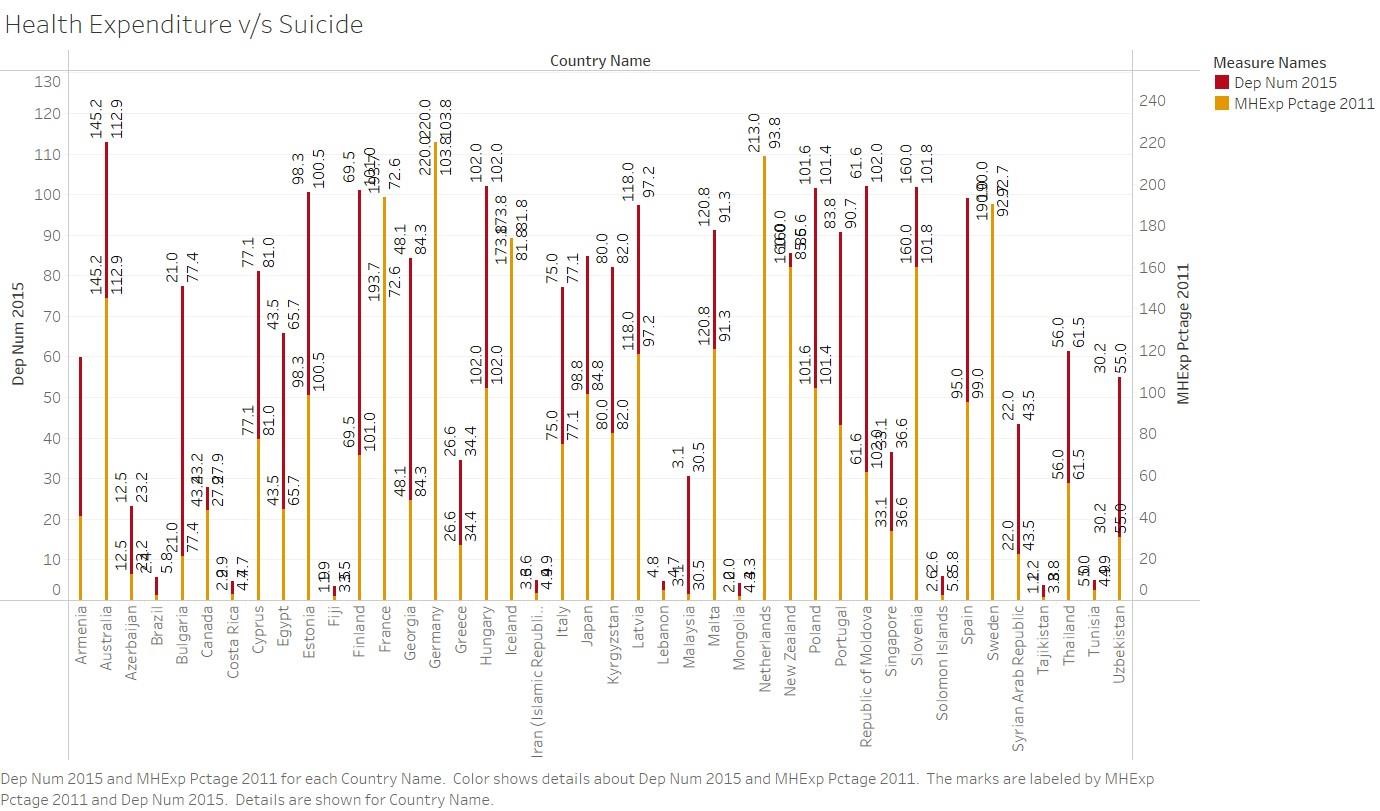
Cyprus is having the least suicide in the year 2004 and the highest suicide in Latvia in the year 2000.

**4.In which year is the health expenditure the**  **highest in the countries?**



**Inference:** This graph shows you the health expenditure in the different countries **Result:** Republic of Moldova has the highest health expenditure in the year 2009.

**5. In the country Thailand what is the Suicide and Health expenditure index?**



**Inference:** This graph shows you the health expenditure and suicide in the country **Result:** In Thailand, the suicide index and Health expenditure is 56.0 and 61.5 respectively.

**Conclusion:**

In summary, a visualization of health expenditure and suicide rate suggests a negative correlation between the two variables. Countries with higher health expenditure tend to have lower suicide rates, and vice versa. However, outliers exist due to unique socio-economic, cultural, or political factors.

Further analysis is needed to fully understand the complex relationship between these variables.

**References:**

* [https://ijret.org/volumes/2015v04/i06/IJRET20150406008.pdf- f](https://ijret.org/volumes/2015v04/i06/IJRET20150406008.pdf-)or data visualization we have referenced this document.
* https://public.tableau.com/en-us/s/ - for data visualization we have referenced this document.