INFSCI 2415: Information Visualization
Analysis of Global Life Expectancy

Vishruth Pradeep Reddy | vpr8@pitt.edu | 12.04.2023

GitHub Link: https://github.com/vishruthreddy18/Global-Life-Expectancy-Analysis
Dataset: https://www.kaggle.com/datasets/lashagoch/life-expectancy-who-updated/data

Box Plot (Main figure used in Midterm Report)

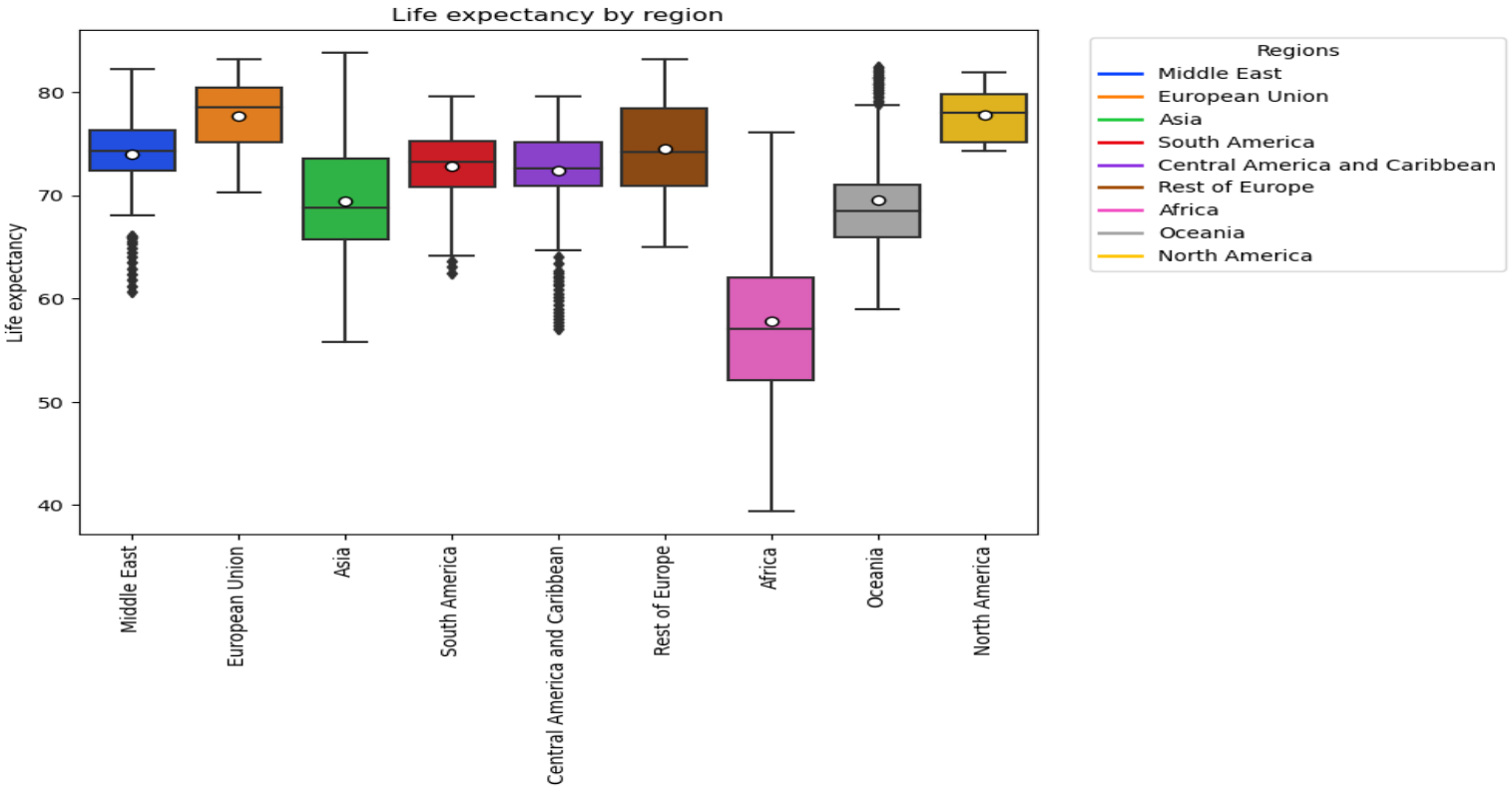


Figure Explained:
The above box plots are a graphical representation of my dataset's summary statistics. The above boxplot shows the life expectancy of the people in various continents. Middle East, South America, Central America, and Oceania have outliers. Africa has the widest and the lowest range of life expectancy compared to the other regions.
The legend on the side displays the color used to represent each region in the boxplots.

- Box: The central part of the plot is a rectangular box. The box represents the interquartile range (IQR), which includes the middle 50% of the data. The bottom and top of the box mark the first quartile (Q1) and third quartile (Q3), respectively. The width of the box shows the spread of the data within this middle 50%.
- Q1 (First Quartile): The 25th percentile of the data, i.e., the value below which 25% of the data falls.
- Q3 (Third Quartile): The 75th percentile of the data, i.e., the value below which 75% of the data falls.
- IQR (Interquartile Range): The range between Q1 and Q3, representing the middle 50% of the data.
- Median: A vertical line inside the box represents the median (Q2) of the data, which is the middle value when the data is sorted.
- Whiskers: Lines extending from the top and bottom of the box, indicate the range of the data.
- Whisker Range: The range between the whiskers, which typically covers the entire data range.
- Outliers: Any data points outside the whiskers are considered outliers and are usually plotted as individual points. Outliers are values that are significantly different from the rest of the data and can provide insights into data anomalies.

Key Findings:

- Life expectancy varies significantly by region.
    - The box plot shows a wide range in life expectancy, from around 50 years in Africa to over 80 years in the Middle East.
    - This suggests that factors like access to healthcare, sanitation, and education play a significant role in how long people live.
- There are outliers in all regions.
    - The whiskers on the box plot extend beyond the boxes in all regions, indicating that there are people who live much shorter or longer lives than the typical range.
    - This could be due to individual factors like genetics or lifestyle choices.
- The distribution of life expectancy is not symmetrical.
    - The boxes in the box plot are not all the same size.
    - For example, the box for Africa is much taller than the box for North America. This suggests that life expectancy is not normally distributed in all regions.
    - In some regions, there may be more people who live very short or very long lives, while in other regions, life expectancy may be more evenly distributed.

KDE (Kernel Density Estimate) Plot

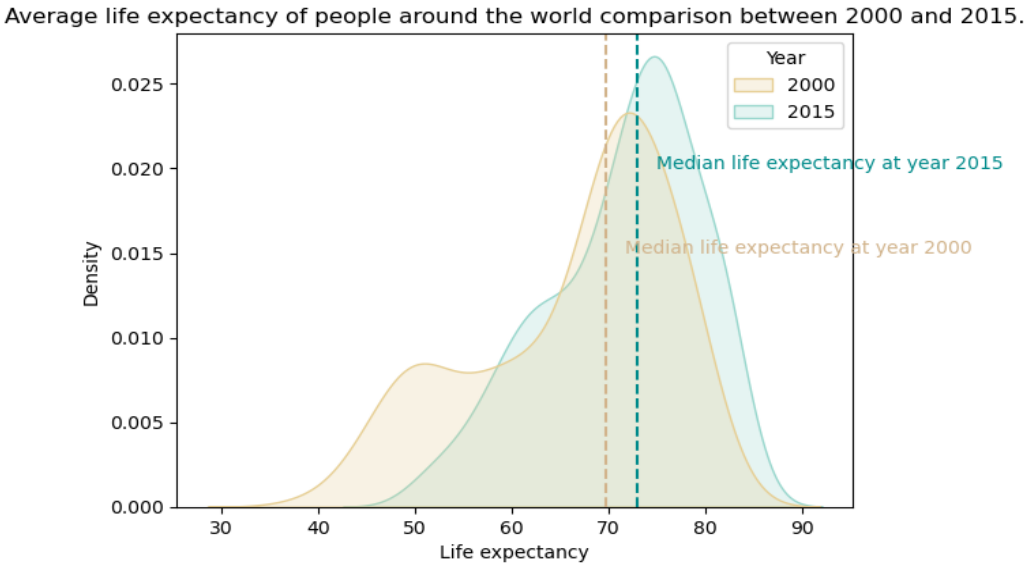


Figure Explained:

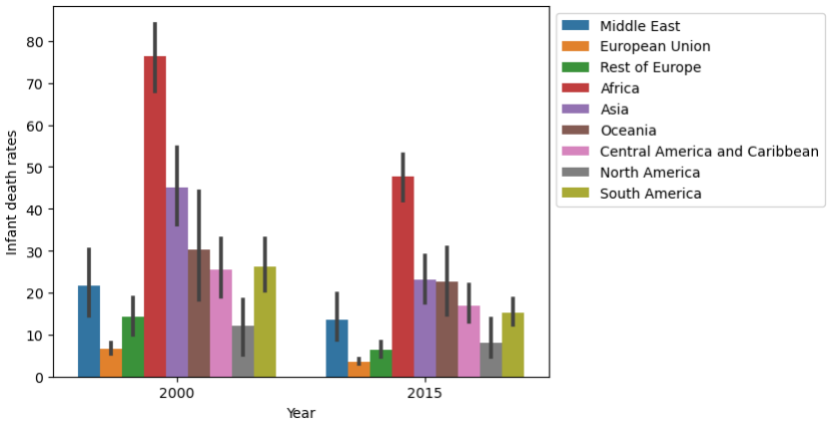
This shows the density vs life expectancy graph using the KDE plot. It shows the shift in the trend over time.

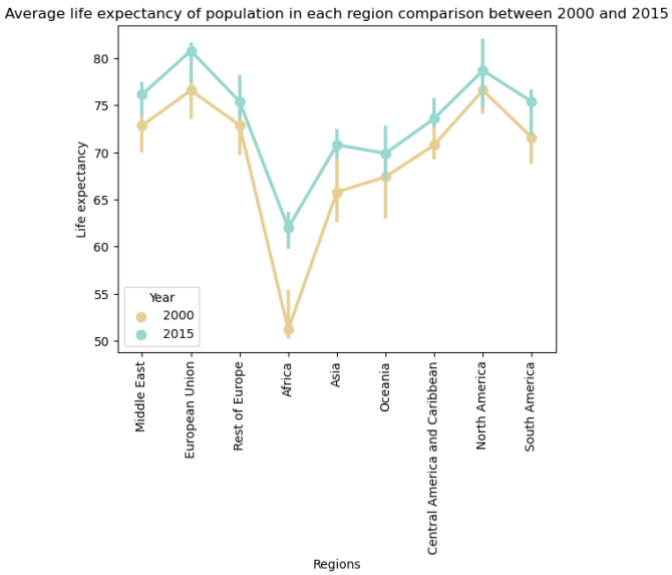
- Yellow color represents life expectancy density for the year 2000.
    - Yellow dotted line is the median for the year 2000.
- Green color represents life expectancy density for the year 2015.
    - Green dotted line is the median for the year 2015.

Key Findings:

- Global Life Expectancy Soared.
    - This refers to a noticeable rise in both the average and middle life expectancy worldwide during the mentioned time frame.
    - The increase could be attributed to various factors such as improvements in healthcare, better access to medical facilities, advancements in sanitation practices, and possibly other socio-economic developments.
    - These advancements likely contributed to longer and healthier lives for people around the world during this period.
- Increase in the Number of Individuals with Higher Life Expectancy.
    - This observation signifies that more people experienced longer life spans. The skewness reduced over time from 2000 to 2015. This indicates a shift towards a more balanced distribution of life expectancies across the population, with fewer extreme values.
    - The plot becoming narrower in 2015 suggests a convergence of life expectancies towards a central range, implying that more individuals were living closer to the average or median life expectancy compared to the earlier period.

Bar Plot and Point Plot





Figures Explained:

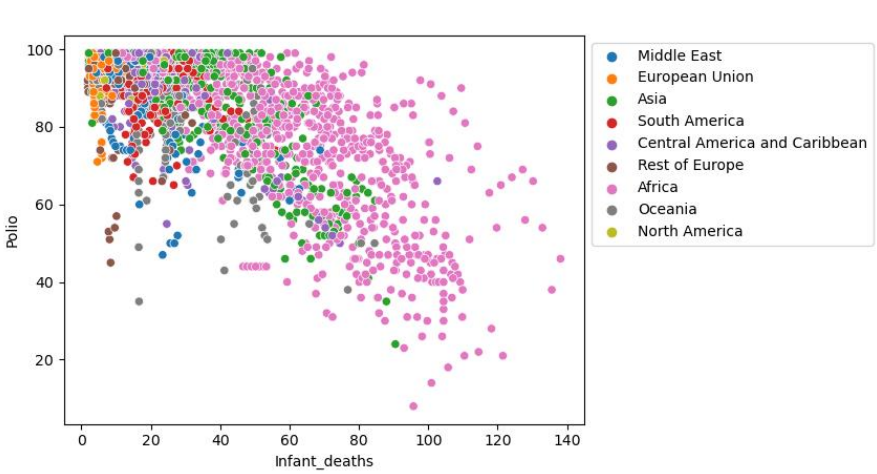
The first figure shows the bar graph of Infant death rate for a period of 2 years (2000 and 2015) over various regions. The second figure shows the point plots for Life Expectancy for various regions over 2 years of time.

- The colors in the bar graph represent the various regions as shown in the legend and the lengths represent the infant death rate.
- The colors in the point plots represent the time period (yellow: 2000 and green: 2015) as shown in the legend.

Key Findings:

- Infant Death Rate Decrease
    - Regardless of the region, there has been a consistent decrease in infant death rates from the year 2000 to 2015.
    - This suggests that efforts, such as improvements in healthcare, sanitation, and nutrition, have been successful in reducing the number of infant deaths during this time period.
- Increase in Life Expectancy
    - Over the same period (2000-2015), there has been an overall increase in life expectancy across regions.
    - This indicates that people, on average, are living longer lives. Factors contributing to this improvement may include advancements in medical technology, better access to healthcare, and overall improvements in living conditions.
- Inverse Proportionality between Life Expectancy and Infant Death Rate
    - There is an inverse relationship between life expectancy and infant death rate.
    - This correlation underscores the importance of addressing factors that contribute to both infant mortality and overall life expectancy.
- Regional Disparities
    - Notably, there are regional disparities in these trends. For instance, Africa is highlighted as having the highest infant death rates but the lowest life expectancy.
    - This suggests that, while progress has been made in reducing infant mortality, there are still challenges in improving overall life expectancy, possibly due to a combination of factors such as disease burden, healthcare infrastructure, and socio-economic conditions.

Scatter Plot



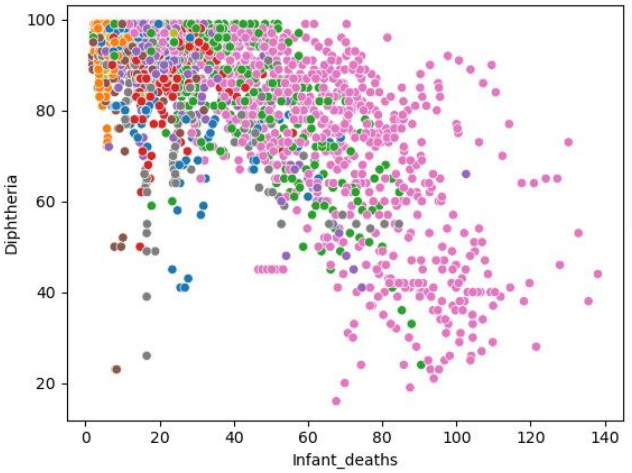


Figure Explained:

- The dots in the figure represent the distribution of % of vaccination coverage across the population vs the infant death rate.
- The colors in the figure represent the various regions across the globe as shown in the legend.

Key Findings:

- Disparities in Healthcare Distribution
    - The data reveals significant inequalities in the availability and accessibility of healthcare resources among various regions worldwide.
    - Africa and Asia exhibit lower vaccine coverage compared to the more robust coverage seen in the European Union (EU) and North America. This discrepancy suggests disparities in healthcare infrastructure, access to vaccines, and public health initiatives among these regions.
- Inverse Relationship between Vaccination Coverage and Infant Death Rates
    - A crucial finding is the inverse proportional relationship between vaccination coverage and infant death rates.
    - In regions where vaccination coverage is lower, such as Africa and Asia, there is a correspondingly higher rate of infant mortality.
    - This correlation underscores the importance of vaccination programs in preventing diseases that can be fatal to infants.
    - Conversely, regions with higher vaccination coverage, exemplified by the EU and North America, tend to experience lower infant death rates, indicating the protective impact of widespread immunization on child health.