The changing relation between GDP per capita and life expectancy from 1961 to 2017

- ## Graph Objective

  The graph objective is finding the changing relation between GDP and life expectancy from 1961 to 2017.

  By observing the data, I found that the changes between each year's data are slight. So, I set the time interval to 5 years to analyze the changing relationship directly. On the other hand, for fairness, I only selected data from different countries.

  In general, this graph studies the relationship between life expectancy and GDP per capita every five years from 1961, based on the annual data of each country.

- ## Data Management

  The data are sourced from the World Bank Open Data website, for GDP and for life expectancy. .

  The raw data include data specified by region, country and organization. It is only fair to do the comparison between same jurisdiction-level. I chose data specified by country, the data are sufficient for visual mining, and the GDP situation is diverse.

- ## Data Validation

  Based on the graph goal, we analyzed the data from 1961 to 2017, using 5 years as the interval to mine changes. I deleted all the data that is not recorded by country, not match the year.

  The data to be processed should contain 3 features, year, GDP and life expectancy. I deleted countries with no GDP, life expectancy records, and countries with too few records, less than 3 values. Then I inner joined all the data together by country code. Then I found that some records have GDP without same year life expectancy, these data should also be deleted.

  Because we do not need to the relationship between countries, the cleaned table has only three features, year, GDP and life expectancy.

  Next, I used visual mining to analyze the cleaned data.

- ## Visual Mining

By simply observing the cleaned data, I found that the data have high density, and more attention needs to be paid to local characteristics. So, I chosen the LoWeSS smoother as visual mining tool.

I confirmed the bandwidth parameter by plotting the residual plot using Stata. I compared the residual plot of bandwidth = 0.4 0.6 0.8, and finally choose 0.8 as the bandwidth. And I also smooth the residual with bandwidth 0.8.

- ## Visual Implantations and retinal variables

Because the residual plot and changing relation both use the LoWess method for smoothing, I use yellow color to represent the fitted line and blue scatter to present the original data. Because the LoWess method is used, the value of bandwidth is marked in the subtitle.

I placed the residual plot and changing relation plot vertically in the dashboard, and I set the range of the GDP axis to same range for easy observation. For the residual plot, the graph target does not care about the residual value, so I delete the grid and keep only the residual axis zero line.

The graph objective focuses on the changing relationship with time, so I use motion to show the relationship between GDP and life every five years.

- Insight and Conclusion

  Here is the dashboard of final solution. Readers can observe the plot by selecting the year or automatic playback.

  We can observe that as GDP rises, life expectancy is rising, but after GDP rises to a higher level, life expectancy remains stable. Because high-income people have the ability to invest more in health. However, capital investment is not enough. The level of medical care and technological development also limit the life expectancy. On the other hand, the gap between rich and poor in the country also limits the average life expectancy.

  With the passage of time, the upper limit of average life expectancy continues to increase. The life expectancy was below 80 years before the 1990s, and the average life expectancy in high-income countries after the 1990s could reach 80-90 years. This shows that with the change of time, human medical level, technological level has improved, and the life expectancy has also improved overall.