

Project Description

Final Project SP by Vadym Kuzyak

1) Research Questions and Hypothesis

This project aims to answer the following research question: How did life expectancy and fertility change between 1964 - 2013 for different countries and continents.

My hypothesis is that in the majority of countries, the life expectancy increased, while the fertility decreased.

The focus of the project lies on the visualization.

A second (smaller) research question which will also be answered: Are there intervals (for example, certain decades) for which the change of life expectancy and fertility is different?

2) Data Sources and Brief Project description

The research done is based on the data from <https://www.gapminder.org/>. This project contains an extensive visualization of life expectancies and fertilities for different countries to answer the research question in a user-friendly way. The project aims to utilize the Bokeh library to make powerful visualizations. The focus is on the visualization, as the data already comes in a processed way.

3) Detailed project description

The project consists of a .jupyter-Notebook and a .csv-file with precomputed values. Running the .jupyter-notebook will produce two .html-files (already included in the GitHub), LifeExp_Fertilities.html and Hypothesis.html. They contain the results of the project: First, a map which shows the life expectancy and fertility for every country graphically on a map (with colors), with the user being able to

- 1) select the year
- 2) run an animation which shows them for all years

Second, it contains a map which aims to show the relationship between life expectancy and fertility. It colors a country green, if my hypothesis holds, grey if there is no significant change and red, if my hypothesis does not hold. The user is able to select any interval of years between 1964 and 2013, for example 1988-1995. This is helpful to answer the second research question. The length of the interval (e.g. 5 vs. 50 years) was weighted into the decision, whether a change was not significant – for smaller intervals, the threshold for non-significant changes is smaller than for larger intervals.

4) Main Research Question: Results

Overall, the hypothesis seems to hold for the majority of countries, as the majority of the countries is colored green (in the initial hypothesis map from 1964-2013, only this map is relevant for the question as the question is dealing with this fixed time interval). Furthermore is true for the majority of countries in general, but also the majority of countries in every region/continent. The Bokeh visualization made it possible to come to this conclusion very fast and was thus very helpful.

5) Smaller Research Question: Results

- There are definitely intervals for which my hypothesis does not hold. Most notably, for the 2000s (2000-2013) the number of green and red countries seems slightly even (with most countries

contradicting the hypothesis being in Europe and Asia). This is also for the 1970/80s (here, most African and -at that time East-Block-countries contradict the hypothesis). These are just two examples for significant outliers.

This shows while overall there was an increase of life expectancy and decrease of fertility between 1964 and 2013, this was not an constant development and for some decades a different development can be observed in the majority of countries.