Higher Diploma in Data Analytics (HDSDA)
Data Visualisation

Student Name: Siobhan Purcell

Student Number: 18195342

Continuous Assignment 2

Project Submission Date: August 7th 2019

Lecturer: Seán Carroll

Table of Contents

Contents

Table of Contents i

Table of Figures iii

1. Introduction 1

  1.1. Project Outline 1

2. Narrative report 2

  2.1. Global evolution of life expectancy 2

  2.2. Life expectancy is a misleading summary of survival 3

  2.3. Exploring the relationship between global life expectancy and fertility rates 6

  2.4. Health and socioeconomic factors of life expectancy 8

  2.5. Can Increased health care expenditure and schooling improve global life expectancy? 9

    2.5.1. Health care and child mortality 9

    2.5.2. Educational attainment and child mortality 10

  2.6. How can we improve life expectancy across the world? 10

  2.7. Conclusion Infographic 12

3. Interactive visualisation of data 13

  3.1. Introduction 13

  3.2. Tableau Storyboard: A Visual History of Life Expectancy 13

4. Technical Report 14

  4.1. Introduction 14

  4.2. Data pre-processing and visualisation methodology 14

    4.2.1. Figure 1: Global map view of life Expectancy in 1800, 1950 and 2018 14

    4.2.2. Figure 2: Time series of life expectancy by continental region 15

    4.2.3. Figure 3: Deaths globally by age group, 2017 15

    4.2.4. Figure 4: Child mortality rate by countries (1800-2016) 16

4.2.5. Figure 5: Global comparison of child mortality rates in 1970 and 2016 ... 17

4.2.6. Figure 6: Bubble chart comparing children (left) and adult (right) mortality between developing and developed countries in 2015. 18

4.2.7. Figure 7: Charting changes in life expectancy and fertility rate by regions from 1800-2018 19

4.2.8. Figure 8: Pearson's Correlation matrix between life expectancy and different health-associated variables 20

4.2.9. Figure 9: Bubble chart of child mortality rates vs. total health expenditure per capita by country and region, 2014... 21

4.2.10. Figure 10: Bubble chart of child mortality rates vs. mean years of schooling, 1950-2010 by country and region 22

4.2.11. Figure 11: Comparative bar charts of several death associated risk factors between developing and developed countries 23

4.2.12. Figure 12: Conclusion Infographic 23

5. References... 26

Table of Figures

Figure 1: Global map view of life Expectancy in 1800, 1950 and 2018 2

Figure 2:Time series of life expectancy by continental region 3

Figure 3: Deaths globally by age group, 2017 3

Figure 4: Child mortality rate by countries (1800-2016)... 4

Figure 5: Global comparison of child mortality rates in 1970 and 2016 5

Figure 6: Bubble chart comparing children (left) and adult (right) mortality between developing and developed countries in 2015. 5

Figure 7: Charting changes in life expectancy and fertility rate by regions from 1800-2018 7

Figure 8: Pearson's Correlation matrix between life expectancy and different health-associated variables... 8

Figure 9: Bubble chart of child mortality rates vs. total health expenditure per capita by country and region, 2014 9

Figure 10:Bubble chart of child mortality rates vs. mean years of schooling, 1950-2010 by country and region 10

Figure 11: Comparative bar charts of several death associated risk factors between developing and developed countries 11

Figure 12: Conclusion Infographic... 12

A Visual History on Global Life Expectancy

Student Name: Siobhan Purcell

Student Number: 18195342

1. Introduction

This project aims to address the subject of global life expectancy – is it improving, and if so, is this improvement consistent across countries? Which countries have fared the best or worst historically and where do they stand today? What are the factors that influence life expectancy such that we might utilise this knowledge to improve it?

To address these questions, the project utilises data from the World Health Organisation and a myriad of other open-source health related databases.

In an aim to contextualise the discussion of global life expectancy, a variety of other demographic factors and socio-economic factors were retrieved. All visualisations were aided by Tableau, ggplot2 in R Studio as well as graphic design tools using Canva, Powerpoint and Adobe Photoshop.

1.1. Project Outline

This project will be divided into three sections and their outputs will output from this project will proceed as follows:

1) **Narrative Report:** an ordered compilation of insights derived from the exploratory and visual analysis, a discussion into their implications, followed by a summarization of key findings through the illustration of a concluding infographic.
2) **Interactive Data Visualisation:** a dashboard published to Tableau Server, which enables the user to engage with the visualisations that furnish the insights for the Narrative Report.
3) **Visualisation Methodology Report:** this section outlines the methods for retrieving, processing and visualising the data, as well as an explanation on for the visual or design choices for conveying the data..

2. Narrative report

2.1. Global evolution of life expectancy

There has been a rapid and dramatic surge in life expectancy over the last few centuries.

This global heatmap compares life expectancy by country across three centuries and starkly visualises the progress that has been made globally since 1800 (Figure 1)

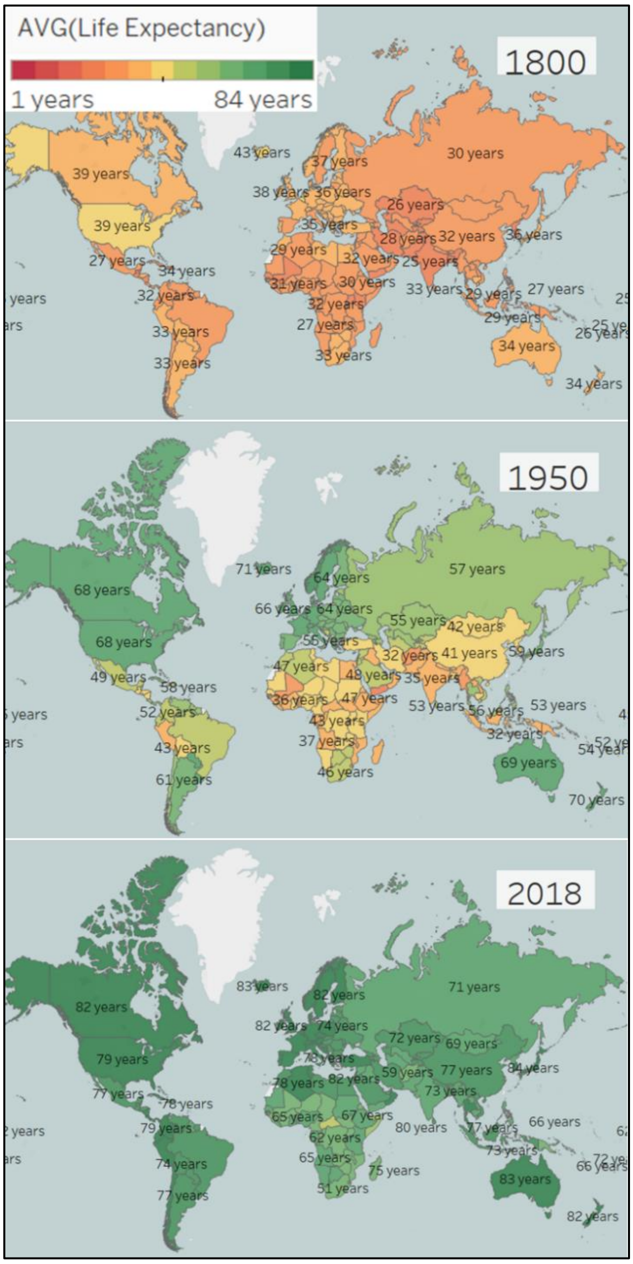


Figure 1: Global map view of life Expectancy in 1800, 1950 and 2018

During the 1800s, historical estimates place life expectancy at about 40 years for all regions of the world. In the following century, whilst global improvements are evident, there is a clear disparity in longevity across the world. Post-war Europe, US and Australia saw economic growth and associated improvements to healthcare provision. In contrast the struggle for the newly independent ex-colonies throughout Africa, and Asia to establish their infrastructure is evident in the lower life-expectancy.

Today, the picture is a lot more positive, with the global average standing at 72. However, there still exists inequalities between countries, with life expectancy ranging from 51-67 years in several Sub-Saharan countries while the average in Japan exceeds 84.

The time series chart below provides a better understanding as to when this transition occurs (Figure 2). It shows that Life expectancy in each region of the world stayed fairly stable for most of history but has doubled since then. It also shows how the onset of increased life expectancy occurred at different timepoints for various regions; Oceania began to see increases in life expectancy around 1870, while Africa didn't begin to see increases until around 1920.

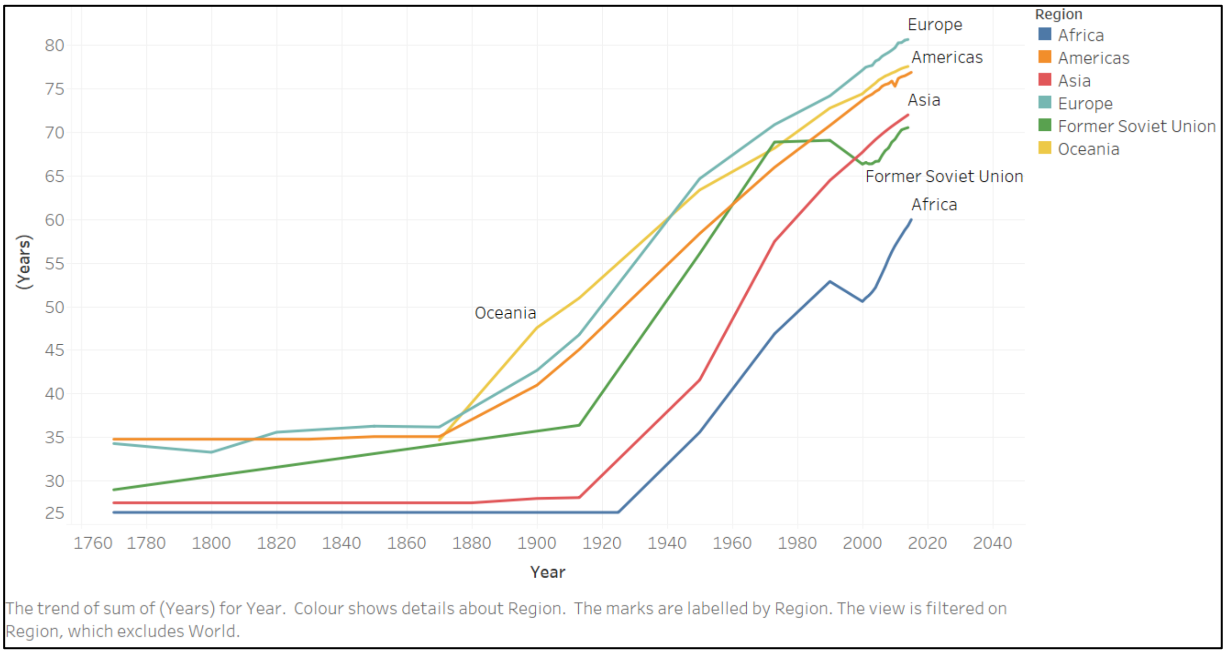


Figure 2:Time series of life expectancy by continental region

2.2. Life expectancy is a misleading summary of survival

However, today's global life expectancy can be a misleading measure of survival and prosperity. The current global average life expectancy of 71.5 years is heavily negatively skewed to a lower number because child mortality factors into this average (Figure 3). As shown by the histogram here, an estimated 5.4 million children under age 5 died in 2017. According UNICEF reports,

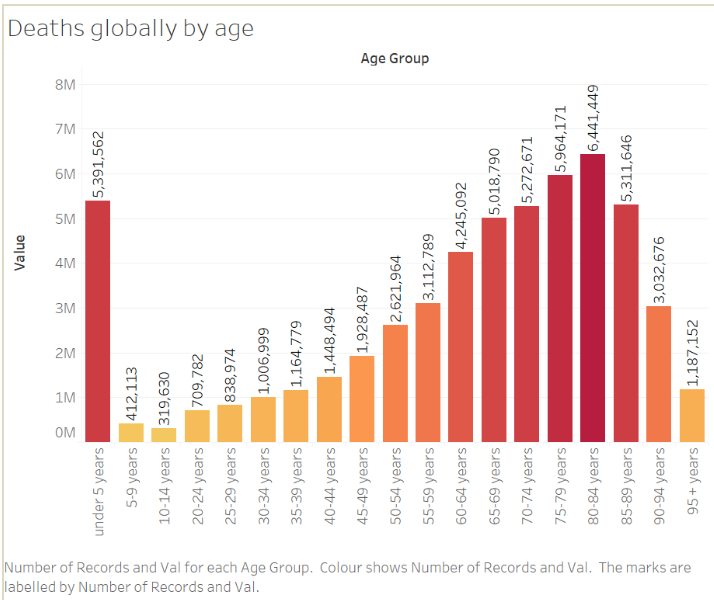


Figure 3: Deaths globally by age group, 2017

approximately 50% of these children deaths occur in sub-Saharan countries Africa (UNICEF, 2017).

However, in line with global life expectancy trends observed before, the survival prospects for this *Under 5 Age Group* have shown remarkable improvements over the years. This is seen below in a time series plotting the decline of child mortality over the years from 1800 - 2016 (Figure 4). As seen below, child mortality first showed signs of improvement in countries that became industrialised early, including the UK, US, Sweden, France and Canada and appears correlated with the economic progress of a nation.

As we can see, this decline happened very fast for more modern countries such as South Korea and Iran. In developing countries such as Ghana and India, the health of children is quickly improving – but child mortality is still much higher than in developed countries.

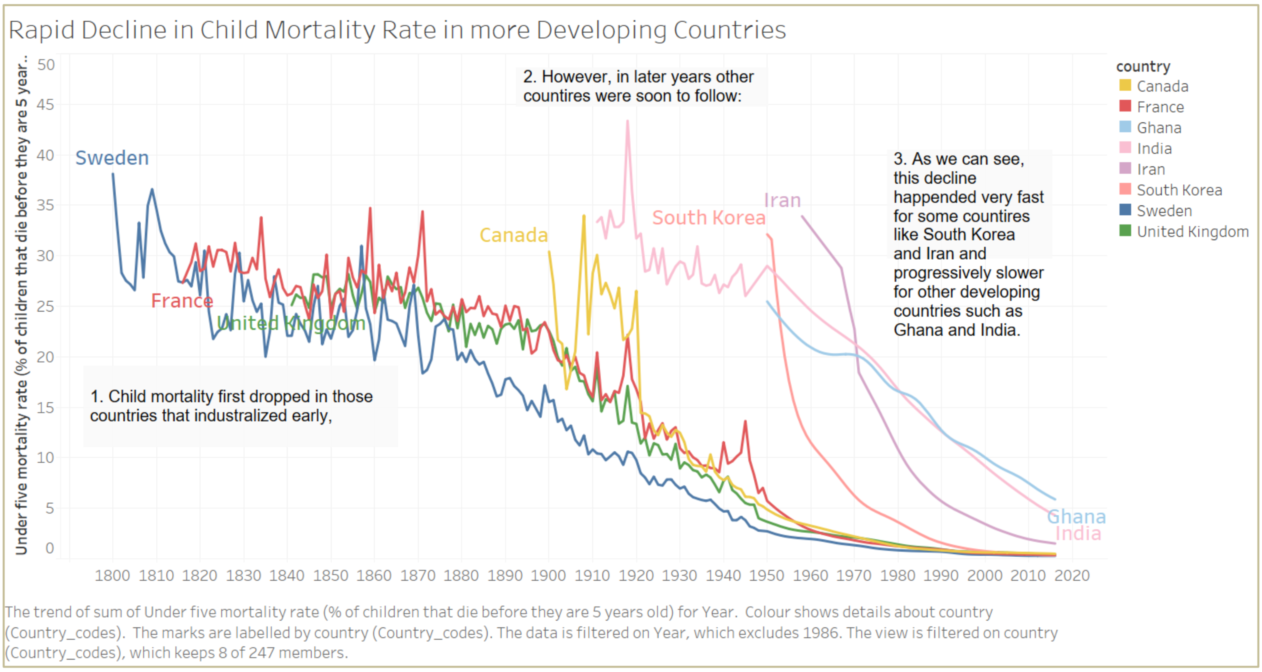


Figure 4: Child mortality rate by countries (1800-2016)

To show this dramatic decrease from a global perspective, two side-by side world map views of child mortality estimates were compared to each other at different timepoints (1970 and 2018) as shown in below in Figure 5. Here we can see that Sub-Saharan countries made the most significant improvements over these years.

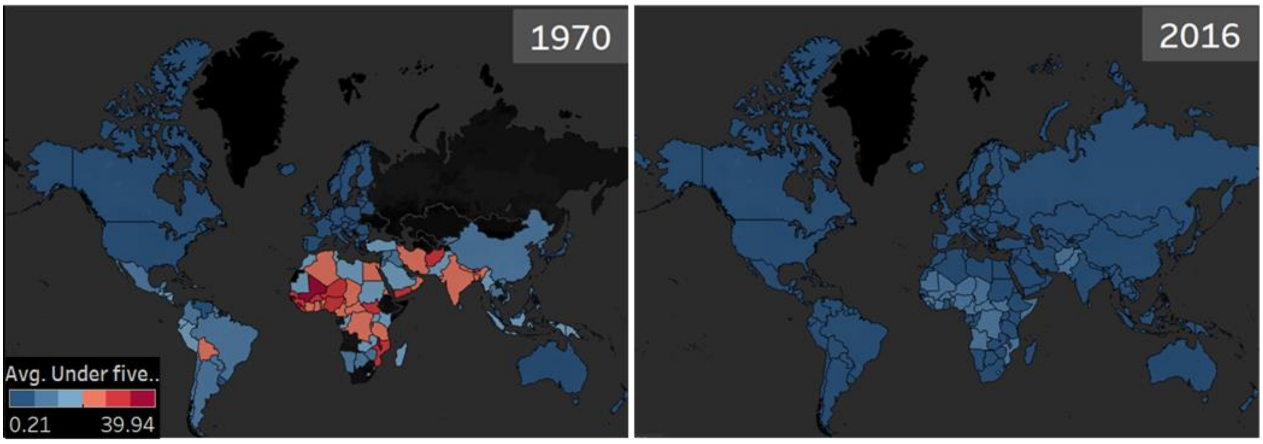


Figure 5: Global comparison of child mortality rates in 1970 and 2016

However, according to life expectancy data captured by the World Health Organisation (WHO), we see that the differences in child mortality rates between developing and developed countries are grossly larger than the differences in adult mortality rates (Figure 6). Therefore, the global burden of child deaths is a call for urgent and concerted action to further improve the survival rates of the children worldwide. By focusing our efforts in reducing infant mortality within developing countries could potentially make considerable contributions towards increasing the in the global life expectancy metric.

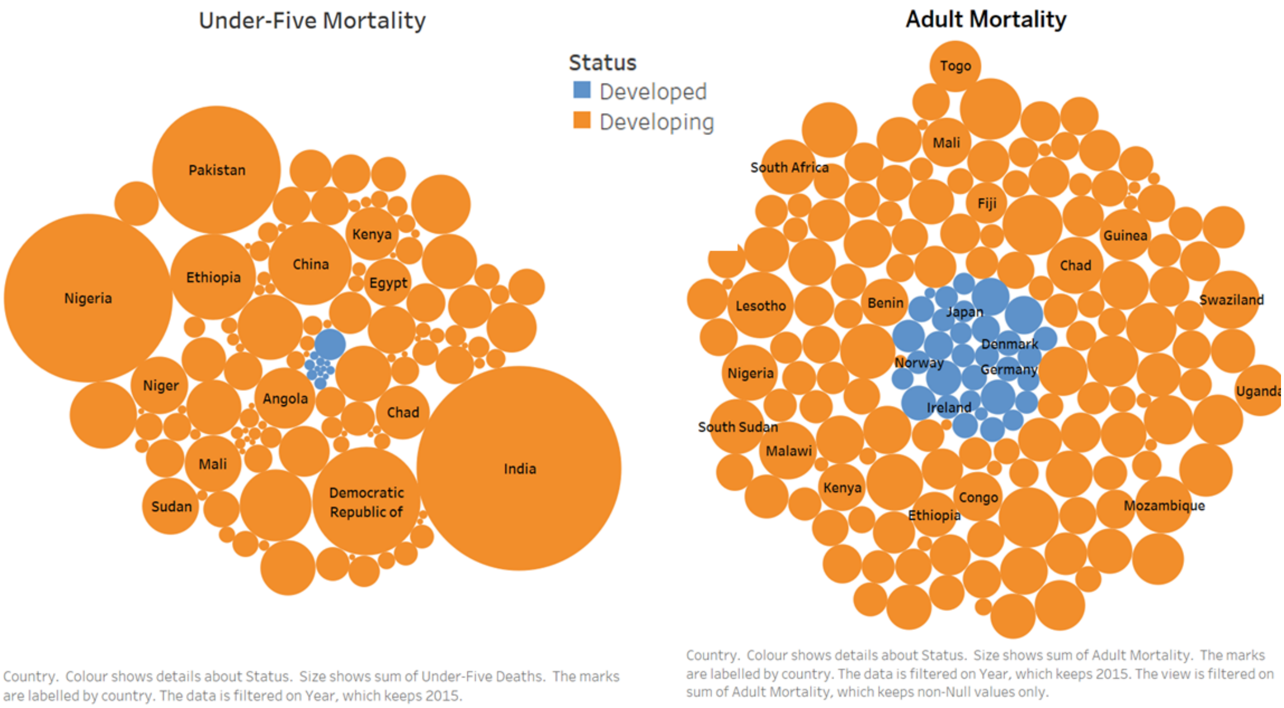


Figure 6: Bubble chart comparing children (left) and adult (right) mortality between developing and developed countries in 2015.

2.3. Exploring the relationship between global life expectancy and fertility rates

Perhaps, one of the reasons why child mortality rates is quite high in developing counties is primarily due to their high fertility rates. As shown here at four different timepoints, the following visualization uses the data on fertility – the number of children born per woman – over time and combines it with information on life expectancy over time (1800-2018). These charts are a powerful representation showing how these two aspects changed over the course of the 20th century- as life expectancy increases, fertility goes down (Figure 7). As previously illustrated, life expectancy is higher in developed countries and from this graphic we see that many of these countries today have far less children than those from the developing world. This link would suggest that as child's chances for survival increases, a parent's demand for the number of children decreases.

From a pessimistic view, one could argue a decrease in child mortality is a burden to the world as it may result in the overpopulation of the planet. However, the visual chart above strongly disproves this cynical justification. It provides clear evidence that as fertility rates decreases the temporary population growth comes to an end. Thus, in order to reduce the world's potential over-population in the future but increase the global average life expectancy, we must work to increasing child survival in developing countries.

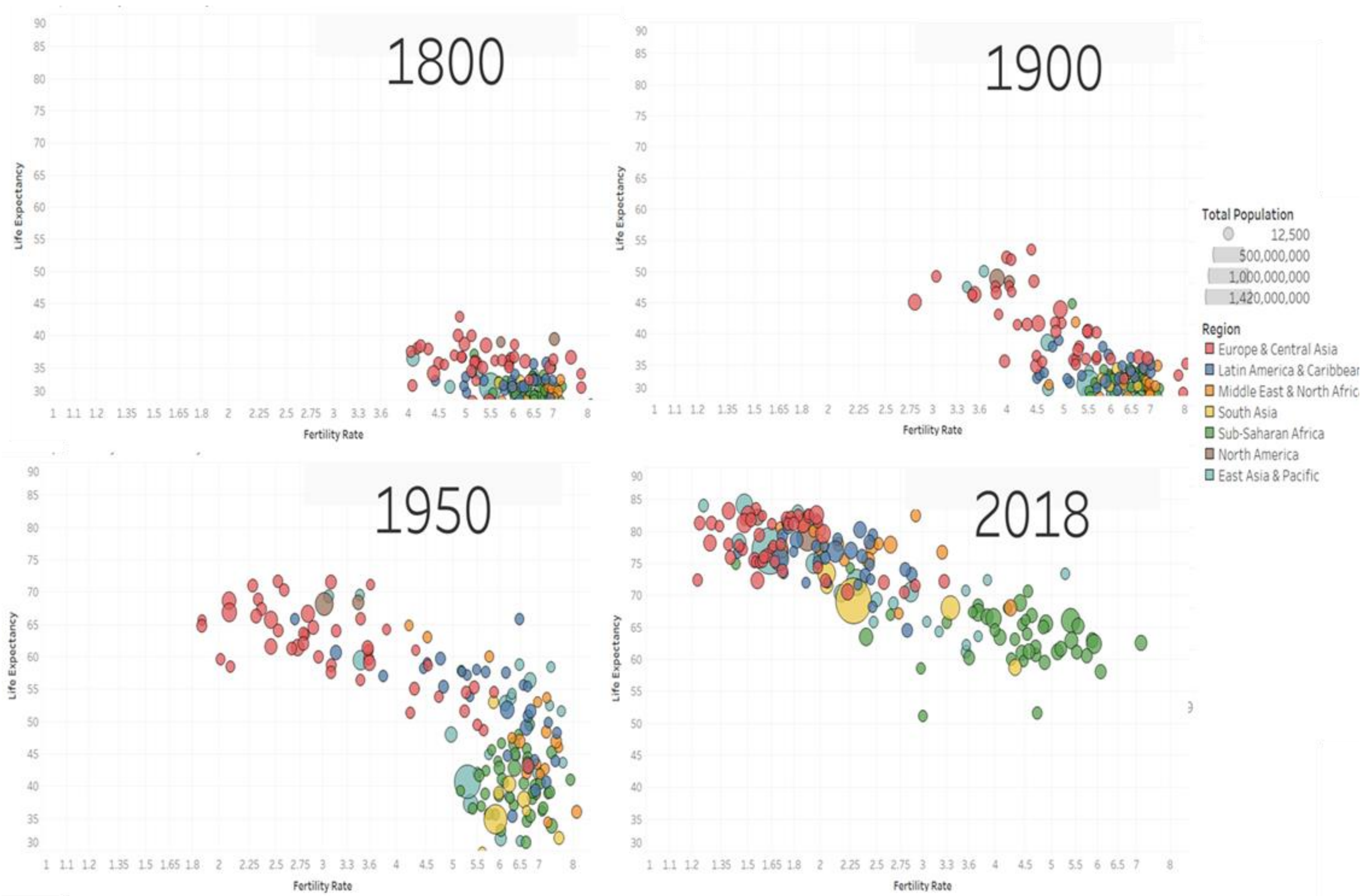


Figure 7: Charting changes in life expectancy and fertility rate by regions from 1800-2018

2.4. Health and socioeconomic factors of life expectancy

In order to identify potential determinants of life expectancy, a correlation matrix was generated in R studio from the WHO life expectancy data, to explore the potential relationships between this target feature and other independent health- and social-associated variables (Figure 8) . Strikingly, from the heatmap below, *life_expectancy* shows a strong positive correlation with *Income_Comp_Of_Resources* (0.72) as and *Schooling* (0.73) whilst, expectedly, shows a strong negative relationship with *Adult Mortality* (-0.7). *Life_expectancy* also shows a moderate negative correlation with HIV*/AIDS* (-0.59) and positive correlation with *BMI* (0.54) Naturally, we see *Adult Mortality* shares a similar level but opposite correlation with *AIDS/HIV* (-0.55). Upon closer inspection, *infant.deaths* and *under.five.deaths* share moderate positive correlations with *Measles* (0.53 and 0.52, respectively) and *population* (0.67 and 0.66, respectively). The latter relationship supports the insights found from last two previous visualisations (Figure 6 and Figure 7) in that rates of children death is related to population size of countries. Most of these countries from the previous graphics were shown to be developing and having high fertility rates.

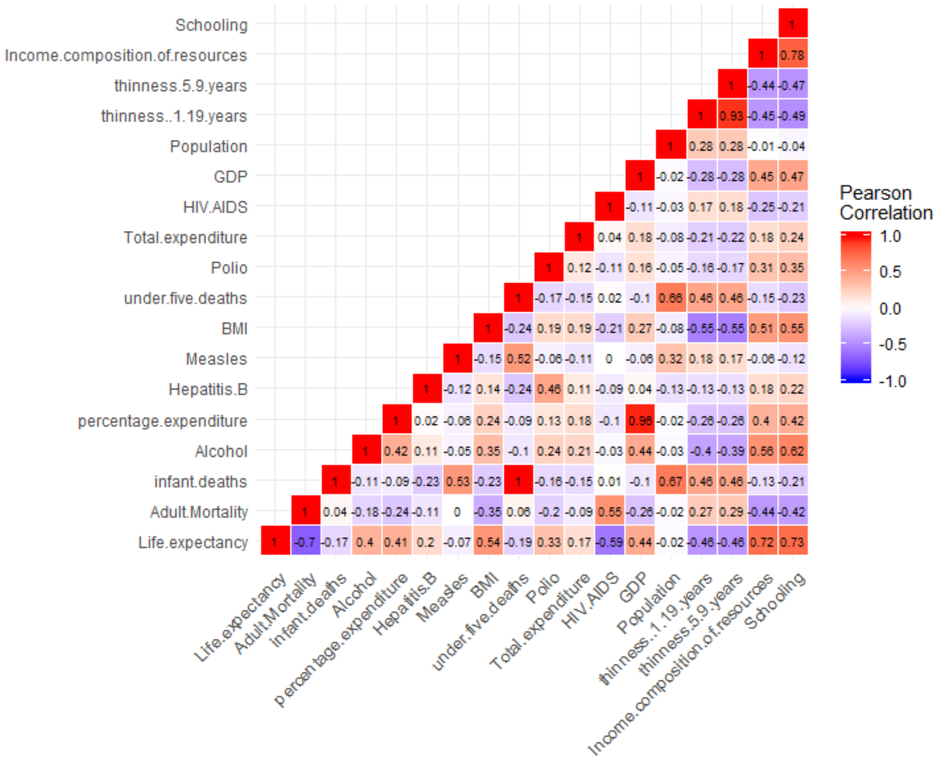


Figure 8: Pearson's Correlation matrix between life expectancy and different health-associated variables

2.5. Can Increased health care expenditure and schooling improve global life expectancy?

Following the insights which had generated from the previous heatmap, data on child mortality percentages from the World Bank was used as further supporting evidence of the correlations between these rates with health and education.

2.5.1. Health care and child mortality

The first graphic below compares a large representative sample of countries of varying socioeconomic status by their levels of child mortality and expenditure amount on health care (Figure 9). These shares of deaths dramatically decrease as a country invests more in disease control and medical facilities. Most of the lowest rates are within countries from either Europe or North America.

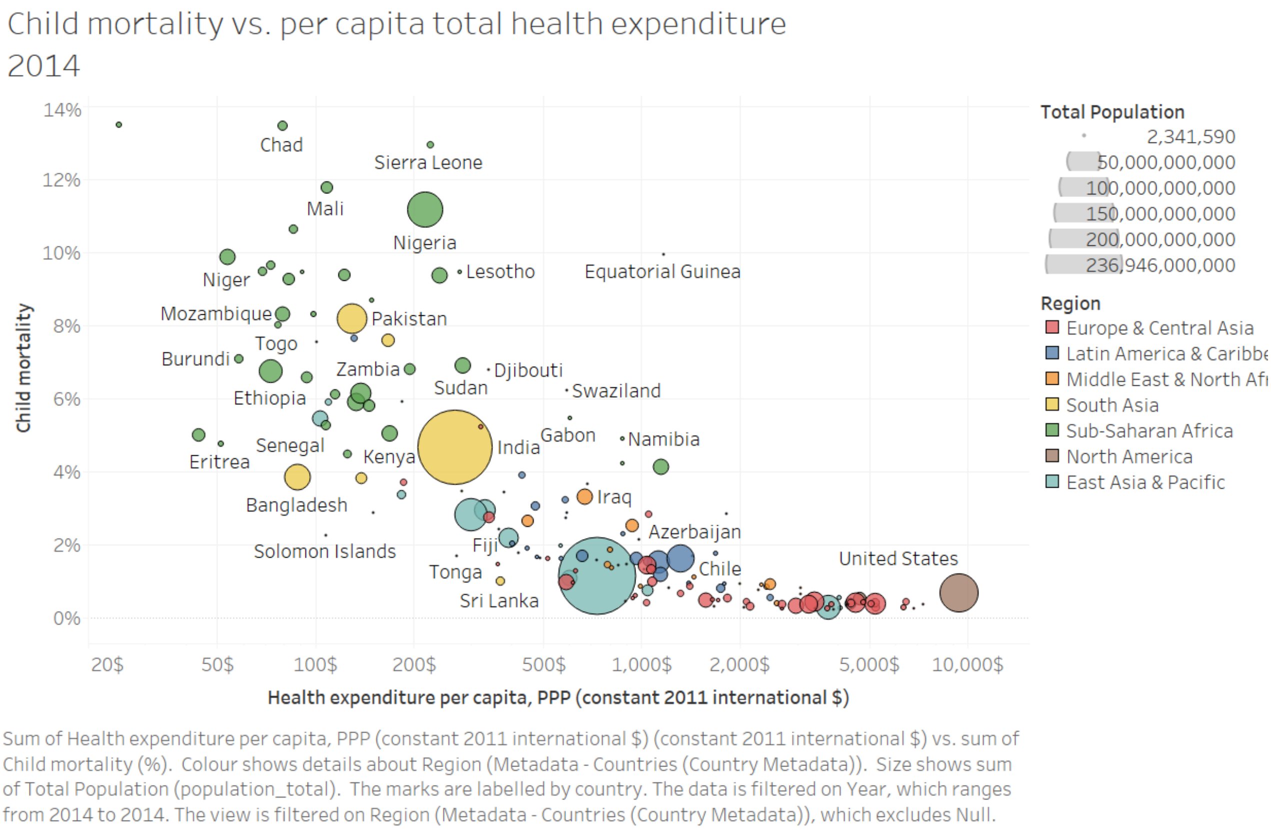


Figure 9: Bubble chart of child mortality rates vs. total health expenditure per capita by country and region, 2014

2.5.2. Educational attainment and child mortality

We can also see a clear relationship exists between education improvements and child mortality. The second chart here shows the mean years of schooling for those aged 15 and older, versus, the share of children who die before reaching the age of five (Figure 10) . Overall, we see that as the average years of education increase in all countries, child mortality declines.

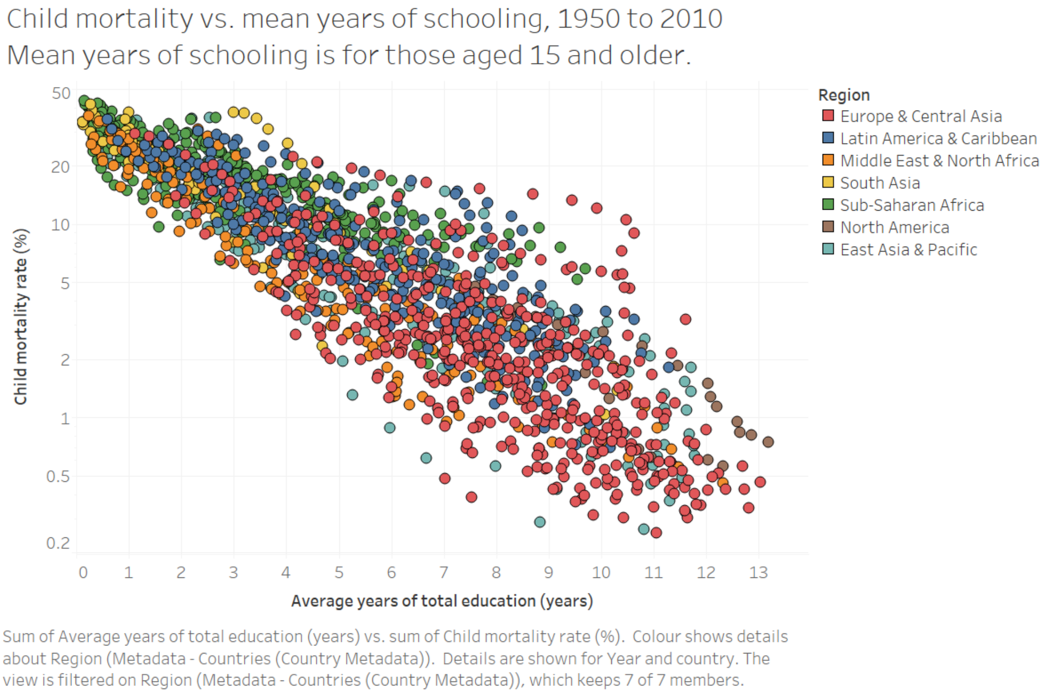
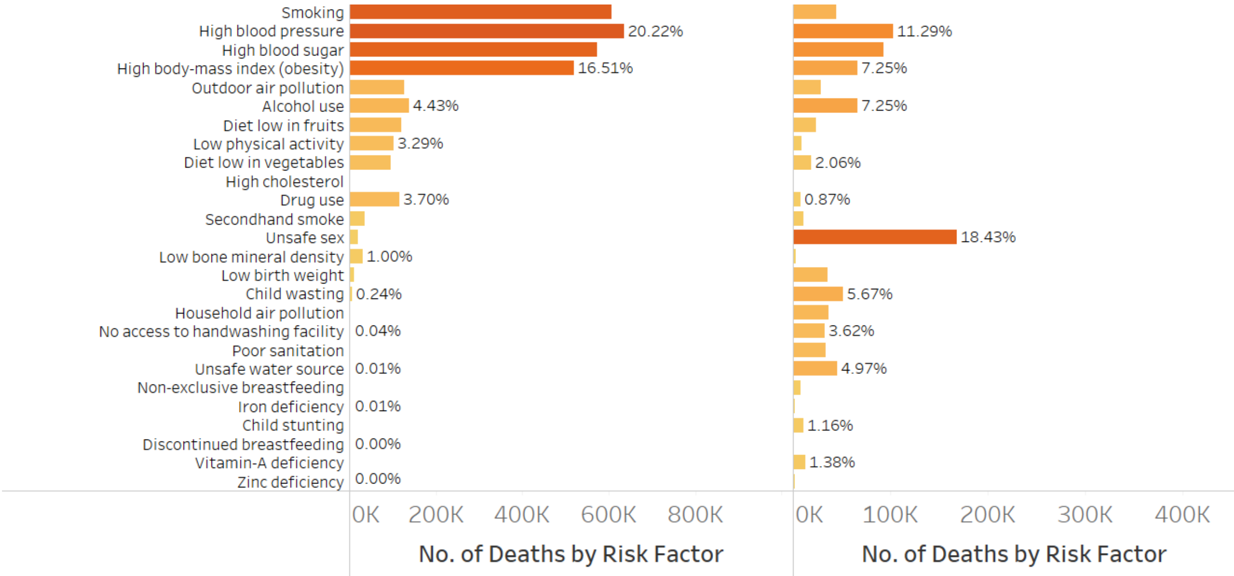


Figure 10:Bubble chart of child mortality rates vs. mean years of schooling, 1950-2010 by country and region

2.6. How can we improve life expectancy across the world?

As a final visual analysis, data collected from the Global Burden of Disease was used to show that the contribution of specific risk factors to mortality varies significantly by country. This data represents the numbers of deaths in which each risk factor was implicated as a contributing cause across a number of selected developed (France, Sweden, UK, US) and developing countries (Ghana, Togo, South Africa, Niger Zambia, Liberia) (Figure 11).

In the chart here, we see the number of deaths attributed to specific risk factors in 2017. This data is measured across all age groups and both sexes. We see a stark difference in the number deaths by risk factors between countries of different developmental status. The dominant risk factors for deaths in developed are notably related to those related to dietary and activity lifestyle factors (including blood pressure, physical activity, body-mass index, blood sugar, and dietary intake); smoking. These would support our previous finding of high BMI being associated with life expectancy in the previous Pearson's correlation matrix.



Sum of No. of Deaths by Risk Factor for each Risk Factors for Death broken down by Countries on page 2017. Colour shows % of Total No. of Deaths by Risk Factor. The marks are labelled by % of Total No. of Deaths by Risk Factor. The data is filtered on Year, country and country (copy). The Year filter ranges from 1990 to 2017. The country filter keeps 10 of 231 members. The country (copy) filter keeps 231 of 231 members. The view is filtered on Countries, which keeps 223 of 223 members.

Figure 11: Comparative bar charts of several death associated risk factors between developing and developed countries

For most high-income countries, the dominant risk factors are those related to dietary and activity lifestyle factors, including blood pressure (20.22%), physical activity, high body-mass index (19.30%), high blood sugar (18.21%), as well as smoking (19.30%) and alcohol (4.43%) intake. These would support our previous finding of high BMI being associated with life expectancy in the previous Pearson's correlation matrix.

We see that in developing low income countries, other top risk factors include child wasting (5.6%), household air pollution (3.6%), unsafe water source, poor sanitation (3.1%), and no access to handwashing facilities (3.2%). Furthermore, in developed countries, where HIV/AIDS is a major health burden such as particularly South Africa and Kenya, unsafe sex is the top risk factor (18.34%).

2.7. Conclusion Infographic

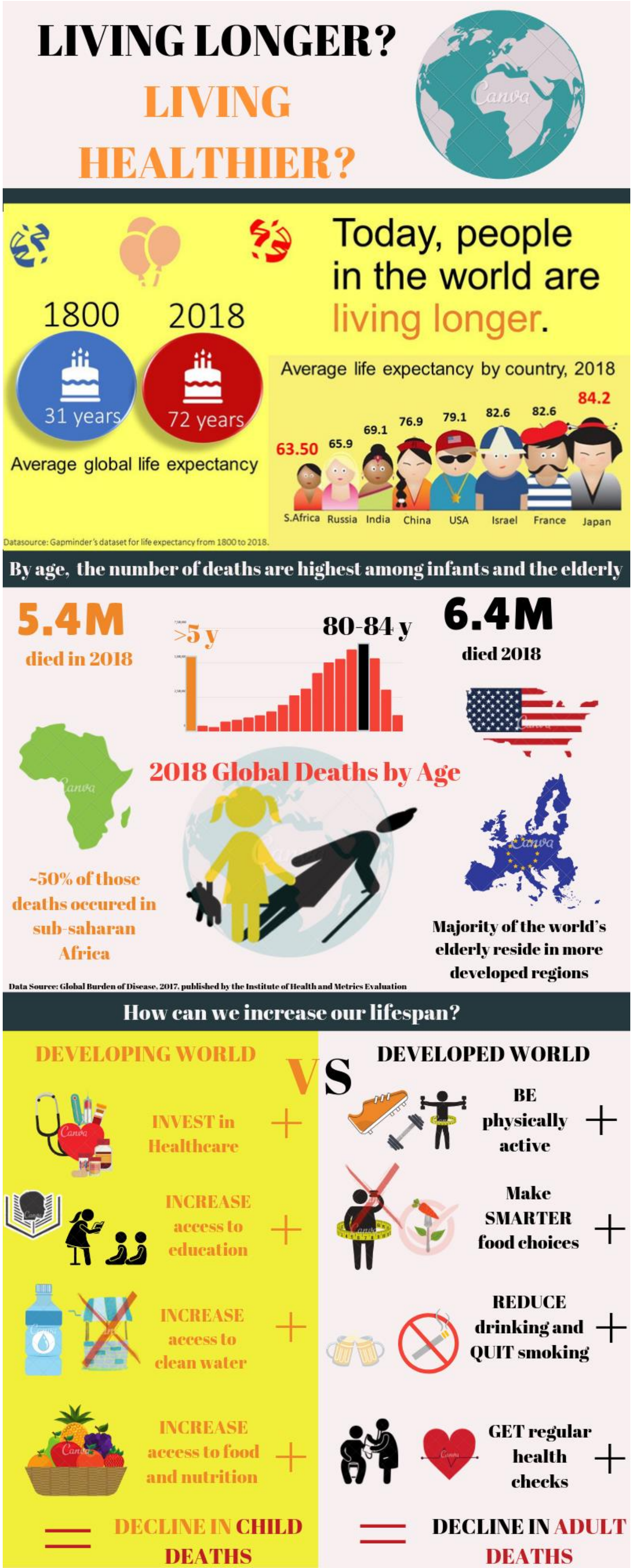


Figure 12: Conclusion Infographic

The infographic in Figure 12 provides an illustrative summary of the key findings from this analysis:

1. Firstly, life expectancy has made a dramatic turn over the last two centuries whereby people are living longer lives than they were before. By rank, we see that Japan is the country with the highest life expectancy whereas people in South Africa have the lowest.

2. Secondly, the global average life expectancy is being negatively skewed by the tragic deaths of children. Despite the global progress in reducing child mortality over the past few decades, an estimated 5.4 million children under age 5 died in 2017—roughly half of those deaths occurred in sub-Saharan Africa. By contrast we see that those aged between years 80-84 are most at risk of death, most of whom are from developed regions of the world such as the United States and Europe.

3. As a final take home message, this report provides recommendations on how we can improve global life expectancy. We can increase the survival of children from developing world by providing them with better healthcare, education, sanitary water and nutrition. Although not as urgent a call, the infographic provides tips on how to increase longevity in those dwelling in high income countries. In the developed world, we should promote a healthy living lifestyle by encouraging more physical activity and consumption of clean foods.

3. Interactive visualisation of data

3.1. Introduction

The majority of figures used in the narrative report and infographic were rendered using Tableau. These were packaged using Tableau's dashboarding capabilities. Following editing, these dashboards were further assembled using the Story feature. This is a progression of visualisations that are arranged in such a way as to convey a narrative. This was utilised to facilitate the storytelling component of the project by taking the user along a narrative, with appropriate context and details to support the reporting and infographics.

3.2. Tableau Storyboard: A Visual History of Life Expectancy

Features of this interactive dashboard include interactive charts with filtering capabilities on year and country as well as animations. In this way users can explore or delve deeper into the underlying datasets to further enhance their understanding. In addition, this facilitates a sense of engagement and ownership of the insights, as users are invited to interact and study these visualisations further.

To access this Story please follow this link: https://public.tableau.com/profile/siobhan2585#!/vizhome/AVisualHistoryofGlobalLifeExpectancy/FinalReportStory?publish=yes

4. Technical Report

4.1. Introduction

This section outlines, where relevant the data retrieval, processing and visualisation steps taken to render the visualisations that inform the Narrative Report and Infographic. It will also discuss and justify the rationale behind any aesthetic or design choices that were undertaken to convey key insights. It is structured to discuss each figure as a section in turn.

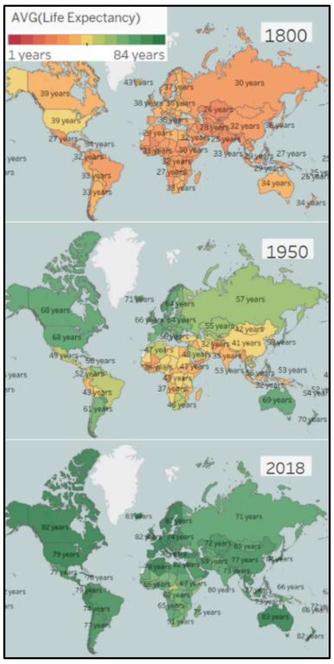
4.2. Data pre-processing and visualisation methodology

4.2.1. Figure 1: Global map view of life Expectancy in 1800, 1950 and 2018

Data Sources

Data was retrieved from Gapminder; values for life expectancy were available from the year 1800 onwards for a discreet set of countries. https://www.gapminder.org/data/documentation/gd004/ [Accessed and downloaded 01/07/2019 at 11:00]

Data pre-processing and visualisation



Life expectancy data was plotted on three global map views in Tableau for two purposes. The first aim was to analyse the data geographically so that life expectancy could be compared across all countries. The second reason was to show its global history over the last two centuries between the years 1700-2018. Choosing this method provides a dramatic overview of the massive change in life expectancy over the years worldwide and how its rate of change has varied by location.

Life expectancy was added as a label for each region to facilitate ease of comprehension, together with a red-green heat mapping. A red-yellow-green colour scale of 13 consecutive hues was used convey the range from low to high expectancy. Green is commonly associated with positivity, growth and progression, a fitting contrast to the typically negative connotations of Red. Interactive features include a *Play* option, allowing the user to see the transformation of life expectancy over years.

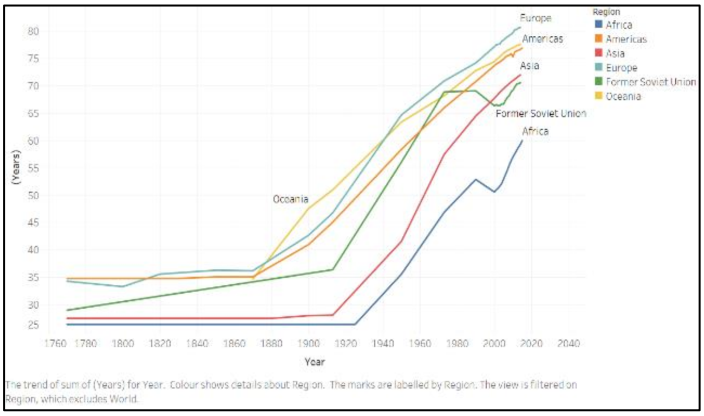
4.2.2. Figure 2:Time series of life expectancy by continental region

Data Sources

Data until 1990 are from: James C. Riley (2005) – Estimates of Regional and Global Life Expectancy, 1800–2001. Issue Population and Development Review. Population and Development Review. Volume 31, Issue 3, pages 537–543, September 2005). Data from 2000 and later are from the WHO and World Bank. Actual data was assembled by Max Rosner and downloaded from Our World In Data (OWID) https://ourworldindata.org/life-expectancy [Accessed and downloaded 01/07/2019 at 11:30]

Data pre-processing and visualisation

Data was already processed to a format that was compatible with Tableau visualisation. Here, a time series charts was generated to allow users to see the change in life expectancy over years by region. In doing so, key moments where a sudden transition in life expectancy occurred for each country is easily understood. Sharp and bold colours were used to easily distinguish between several representative regions.



The trend of sum of (Years) for Year. Colour shows details about Region. The marks are labelled by Region. The view is filtered on Region, which excludes World.

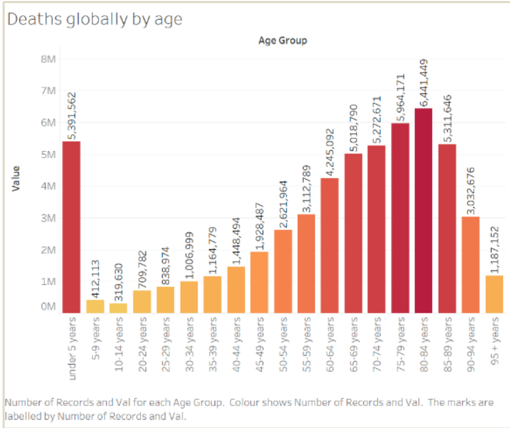
4.2.3. Figure 3: Deaths globally by age group, 2017

Data Sources

Data for all causes of death for all sexes and different age groups was selected for and downloaded from the Global Burden of Disease study, 2017 http://ghdx.healthdata.org/gbd-results-tool. [Accessed and downloaded 01/07/2019 at 11:42]

Data pre-processing and visualisation

After reading the file, columns containing redundant information were removed and the correct age group was relabelled for each row in the dataset.



Number of Records and Val for each Age Group. Colour shows Number of Records and Val. The marks are labelled by Number of Records and Val.

A histogram was used to demonstrate the non-normal distribution of deaths by age. This graphic effectively captures the extreme differences in the number of deaths across different ages group by highlighting those under five years of age as extreme negative outliers in the data. A sequential orange to red colour scale was used to differentiate high from number of deaths and to better highlight the under five years age group as an outlier in the distribution. Orange was selected to represent the lower range of age values, to highlight a sense of warning yet less urgent matter, whilst red was selected to invoke a sense of an alarm for the viewer.

4.2.4. Figure 4: Child mortality rate by countries (1800-2016)

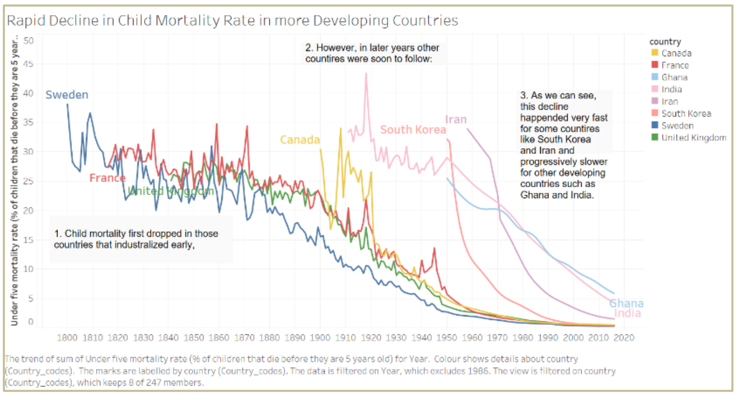
Data Sources

Data was retrieved from Gapminder; values for child mortality were available from the year 1800 onwards for a discreet set of countries.

https://www.gapminder.org/data/documentation/gd005/ [Accessed and downloaded 01/07/2019 at 11:47]

Data pre-processing and visualisation

Data was already processed to a format that was compatible with Tableau visualisation. A time series chart was used compare the differences in declining child mortality rates across the years 1800-2016 and their timepoints of change. Countries from different cultures and socioeconomic backgrounds were deliberately selected for comparison in order to highlight the major discrepancies between developing and developed regions. Bold and contrasting colours were used to differentiate between the representative countries.

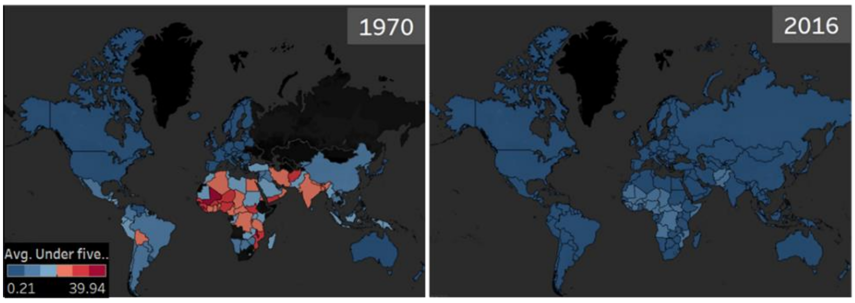


4.2.5. Figure 5: Global comparison of child mortality rates in 1970 and 2016

Data Sources

Data was retrieved from Gapminder; values for child mortality were available from the year 1800 onwards for a discreet set of countries.
https://www.gapminder.org/data/documentation/gd005/ [Accessed and downloaded 01/07/2019 at 11:47]

Data pre-processing and visualisation



Data was already processed to a format that was compatible with Tableau visualisation. Two side-by-side world map views of child mortality rates were created in Tableau to compare their best/worst estimates at different timepoints. Both this juxtaposition and colour choices were effective in showing its dramatic decrease over the last number of years. A red-to-blue heatmap colour scheme was used as a scale for declining child mortality rates. A child's death would typically arouse an overwhelming sense of injustice, senseless suffering and unfulfilled dreams. Thus, to create a dramatic and ominous mood, the map of the world was presented with a black background layer. This visualisation was also made interactive by allowing the user to press the play option and see the decline of child mortality over years on Tableau Desktop. This option is not available on at a server level.

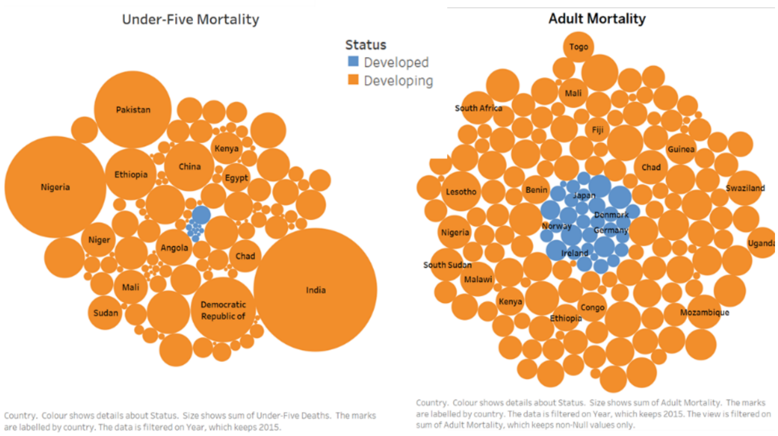
4.2.6. Figure 6: Bubble chart comparing children (left) and adult (right) mortality between developing and developed countries in 2015.

Data Sources

Data was originally published by the Global Health Observatory (GHO) data repository under World Health Organization (WHO). Actual data was retrieved from the Kaggle data sharing platform https://www.kaggle.com/kumarajarshi/life-expectancy-who/data [Accessed and downloaded 02/07/2019 at 11:58].

Data pre-processing and visualisation

For this visualisation, the data was already processed to a format that was compatible with Tableau Two packed bubble charts were generated in Tableau to depict the number of children and adult deaths by country; these were further delineated by their developmental status using colour.



Each bubble represents a country, with size indicating the magnitude of death of children or adults. When depicted in this way, the bubbles starkly highlight the greater discrepancy in child mortality between developing and developed countries. This effectively shows that infant survival is a far more serious issue in low income countries.

Colour was chosen to support the jarring nature of this insight. Contrasting colours were chosen to depict developmental status. Blue was chosen to represent developed countries as it is typically associated with stability and comfort. Orange was chosen for its conspicuousness, association with urgency and contrast with blue. This visualisation was also made interactive through the incorporation of a dynamic filter for *Year* and *Country.*

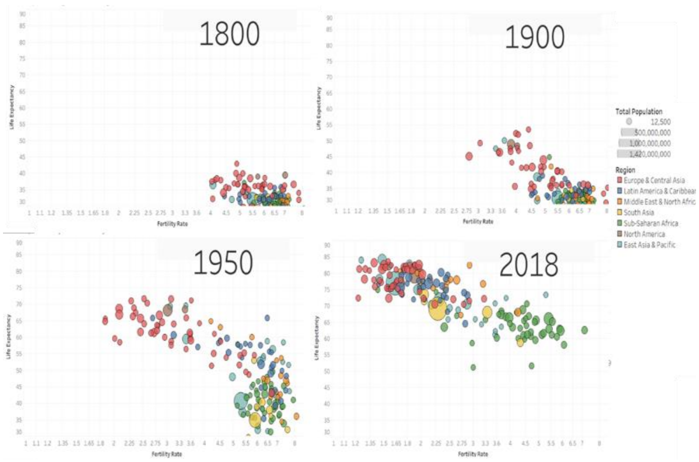
4.2.7. Figure 7: Charting changes in life expectancy and fertility rate by regions from 1800-2018

Data Sources

Datasets for "Children per woman(total fertility)", "Population.total" and "Life expectancy (years)" and "Regions of the World" were retrieved from Gapminder https://www.gapminder.org/data/?source=post_page. [Accessed and downloaded 01/07/2019 at 17:00]

Data pre-processing and visualisation

Columns in all data files were rearranged and renamed as appropriate in Tableau before editing their relationships to each other in order to merge their data sources together for visualisation.



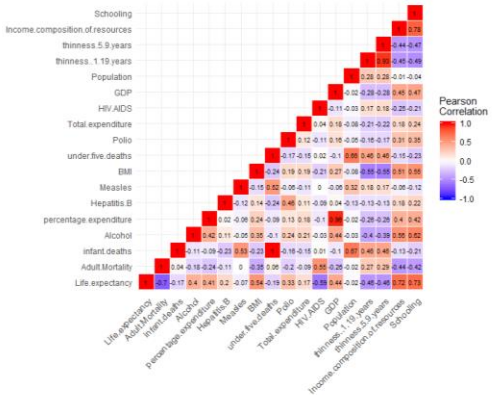
This animated graphic was inspired by Hans Rosling's representation of life expectancy data against fertility rates using data from Gapminder. Here, each bubble is a country. The size of the bubble represents the country's population, and the colour of the bubble its geographical region. The x-axis shows fertility rates and the y axis shows life expectancy at birth. This visualisation was also made interactive by allowing the user to press play and see an animated version of this chart on Tableau Desktop. This option is not available on at a Server level. To save the viewer from exploring the interactive visualisation, snapshots of the animation were taken at four different timepoints across the three centuries and assembled together in Powerpoint for the report.

4.2.8. Figure 8: Pearson's Correlation matrix between life expectancy and different health-associated variables

Data Sources

Data was originally published by the Global Health Observatory (GHO) data repository under World Health Organization (WHO). Actual data was retrieved from the Kaggle data sharing platform https://www.kaggle.com/kumarajarshi/life-expectancy-who/data [Accessed and downloaded 02/07/2019 at 11:58].

Data pre-processing and visualisation



For this visualisation, the dataset was read into R studio. An initial exploratory analysis of this large data was conducted in order to gain a preliminary assessment of all possible features within the dataset. The data was derived from the Global Health Observatory (GHO) data repository under World Health Organization (WHO) which monitors information on all developing and developed countries (193) related to life expectancy, health, social, mortality and economic factors.

The data contains 2938 observations and 22 variables. For a full comprehensive review of how the dataset was explored, please refer to the original R script. Upon initial inspection of the tabular dataset, a number of missing values were observed in the data. The distribution pattern of all total missing values was then visualised using the VIM package which indicated that most of the missing data was for population Hepatitis B and GDP. As finding missing values for this data was difficult to trace from other source, it was decided to omit all NA values from the from the final model dataset for the purpose of the correlation matrix. After removing these values, the results of the final dataset contained 1649 observations and 22 variables.

A Pearson' correlation matrix was used in the visualisation exploration to provide a wealth of information to the viewer of the multiple associations between life expectancy and health/social/economic factors. These relationships were chosen to be displayed in such an ordered manner to facilitate the process of recognising meaningful patterns. In this way, pairs of variables that have the highest correlation can be easily identified. This was further aided by the colour choices for this graphic. A simple sequential red to blue colour scale of 5 consecutive hues was used differentiate high from low values. White was used as a neutral colour to act as a reference value in the middle of the data range.

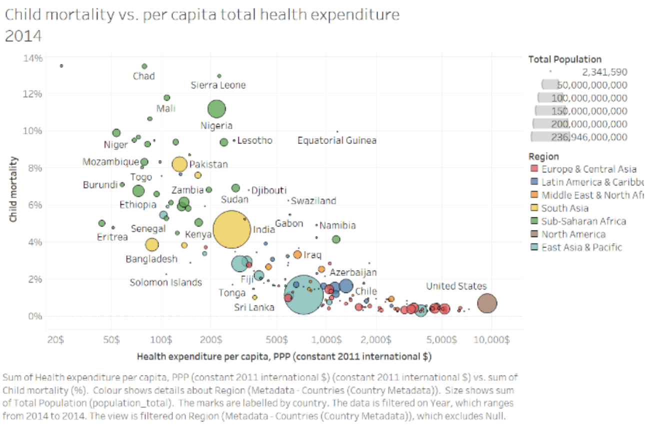
4.2.9. Figure 9: Bubble chart of child mortality rates vs. total health expenditure per capita by country and region, 2014

Data Sources

Data on Total Health Expenditure Per Capita (Constant 2011 Internation $) and Mortality Rate, Under 5 (Per1000 Live Births) are publish by the World Bank –World Development Indicators . http://data.worldbank.org/data-catalog/world-development-indicators Actual data was assembled by Max Rosner and downloaded from Our World In Data (OWID) https://ourworldindata.org/child-mortality [Accessed and downloaded 01/07/2019 at 19:00]

Data pre-processing and visualisation

Data was already processed to a format that was compatible with Tableau visualisation. Following from the results of the correlation matrix, bubble scatterplots were used to focus on key correlated pairs of variables using data from other sources for validation of the associated between disease and child mortality. The size of the bubble was used to represent population size. This was selected so that the viewer could effectively see the change in population size over



time as total healthcare expenditure care increased and child mortality rates decreases. Different contrasting colour for chosen for each world region so that they could easily distinguishable. This visualisation was also made interactive by allowing the user to press the play option and see a motion animation of the increase of life expectancy and decrease of fertility rates across world region over years on Tableau Desktop. This option is not available on at a server level.

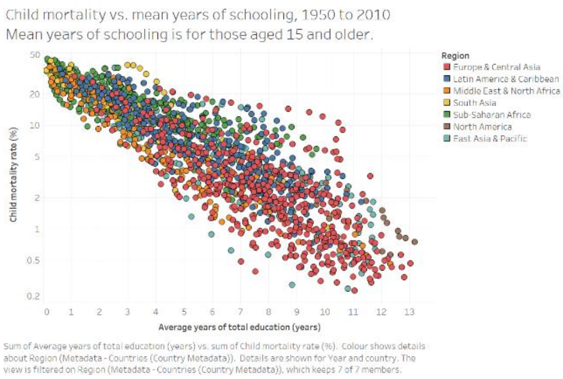
4.2.10. Figure 10:Bubble chart of child mortality rates vs. mean years of schooling, 1950-2010 by country and region

Data Sources

Data on Educational Attainment (Average Years of Total Education) was originally published by Barro, Robert and Jong-Wha (2010), A New Data Set of Educational Attainment in the World, 1950-2010. Journal of Development Economics, vol 104, pp.184-198. http://www.barrolee.com/. Data on Child Mortality Rates were originally published by United Nations, Department of Economic and Social Affairs, Population Division (2017). Actual data was retrieved was assembled by Max Rosner and downloaded from Our World In Data (OWID) https://ourworldindata.org/child-mortality [Accessed and downloaded 01/07/2019 at 17:00]

Data pre-processing and visualisation

Data was already processed to a format that was compatible with Tableau visualisation. Another scatter plot was employed to visual the correlation between child mortality rates and educational attainment of countries over the years, 1950-2010. Unlike the previous scatter plot in Figure 9, all datapoints were included in these visualisations, with countries being coloured by their region.

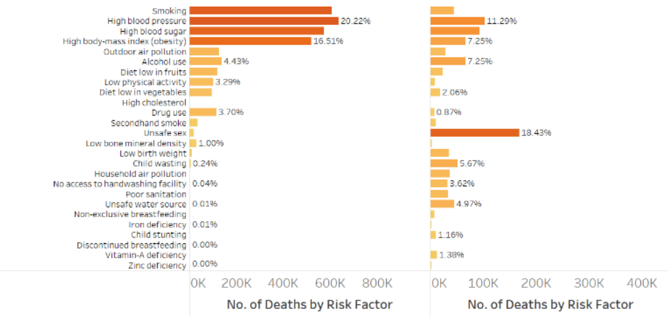


4.2.11. Figure 11: Comparative bar charts of several death associated risk factors between developing and developed countries

Data Sources

Multiple data sources were assembled and prepared by Max Rosner from OWID. A full description of these sources and where the original file was downloaded can be found at . https://ourworldindata.org/causes-of-death [Accessed and downloaded 01/07/2019 at 17:00]

Data pre-processing and visualisation



Sum of No. of Deaths by Risk Factor for each Risk Factors for Death broken down by Countries on page 2017. Colour shows % of Total No. of Deaths by Risk Factor. The marks are labelled by % of Total No. of Deaths by Risk Factor. The data is filtered on Year, country and country (copy). The Year filter ranges from 1990 to 2017. The country filter keeps 10 of 231 members. The country (copy) filter keeps 231 of 231 members. The view is filtered on Countries, which keeps 223 of 223 members.

Data was already processed to a format that was compatible with Tableau visualisation. Parallel bar charts of death associated risk factors were used to compare the differences between developing and developed countries. An orange to red colour scale was used to differentiate between low and high-risk values. Orange was chosen to represent the lower range values as it often implies a warning notice and anticipation of risk. Red was selected to represent the high-risk values as it is associated with danger and severe warning signals. This visualisation was also made interactive by allowing the user to press play and see an animated version of this chart on Tableau Desktop

4.2.12. Figure 12: Conclusion Infographic

Data Sources

The final infographic was created on Canva's online graphic design tools. All additional icons were sourced on Google images and Adobe photoshop image editing software was used to modify their background to fit into the final infographic.

Data sources referenced from Figure 1, 3, 9, 10 and 11 were used to generate the final conclusion infographic.

All icons that were used in this infographic were obtained following websites and modified using Photoshop:

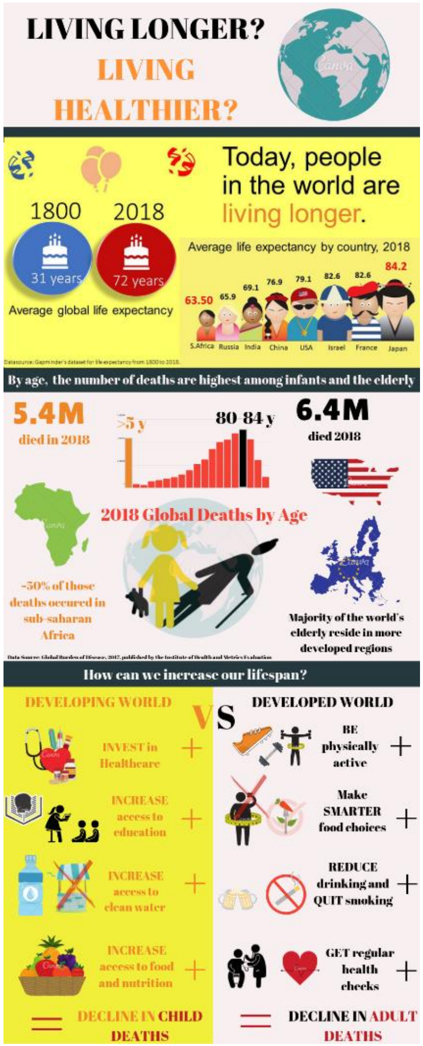
http://www.clipartpanda.com/clipart_images/life-expectancy-infographic-14313108[Accessed and downloaded 04/07/2019 at 18:00]

https://www.health.org.uk/infographic/poverty-and-health[Accessed and downloaded 04/07/2019 at 17:00]

Data pre-processing and visualisation

The purpose of this infographic was to visually summarize t3 key take-aways from this analysis, namely:

1. Global life expectancy has improved, but this has progressed unequally across countries.
2. Child mortality plays a key factor in overall life expectancy. Furthermore, this is an issue still prevalent within the Developing World.
3. Recommendations for actions that could be taken to improve life expectancy, one that considers the development status of a country.



Several key features were utilised in this infographic to inform the viewer and drive the pivotal points across, that is, colour choice, font, iconography and size.

In terms of colour, a bright yellow colour was selected for the background in the first section, to convey the positive mood that longevity would suggest. The colours orange and black were chosen to signify the youth vs. elderly respectively, and this is consistent where relevant across each subsequent view. Furthermore, orange was used to raise a sense of alarm and urgency in the viewer fitting to the subject matter of infant death. whilst black is an appropriate contrast that is suitably sombre.

In the second part of the infographic, the central image of an orange girl and her elderly shadow conveys the theme of this view - two age groups that were that are most at risk of death each year. This colour representation is again present in the fonts that describe each group, keeping consistent with this colour selection. Furthermore, in the histogram for child mortality, red was chosen to echo the negative connotations of the subject matter, death.

In the third section, yellow and beige were chosen to support the subject of this section – ending the infographic on a positive, affirmative note, with clear recommendations on how to improve life expectancy by age group according to a country's developmental status. Font colour was chosen to be consistent with choices for these two groups.

Font and colour were frequently leveraged together to both emphasise key facts. as well as further reinforce the dichotomy between age groups. Capitalisation of recommended actions and jarring facts was used as a technique to reinforce urgency.

Iconographic were chosen as a form of visual language to embody subject matter where space for words is scarce. These were intended to be simple elements that would be recognised and understood almost immediately. For instance, in the first pictorial bar chart, countries are conveyed using stereotypical representations of their people, with size signifying the magnitude of the data point. In the second view, maps are utilised to allow the user to easily recognise the region in question. Icons were employed to reinforce the recommendation section by synthesising each suggestion into a succinct visual.

5. References

United Nations Inter-agency Group for Child Mortality Estimation, Levels and Trends in Child Mortality: Report 2017.