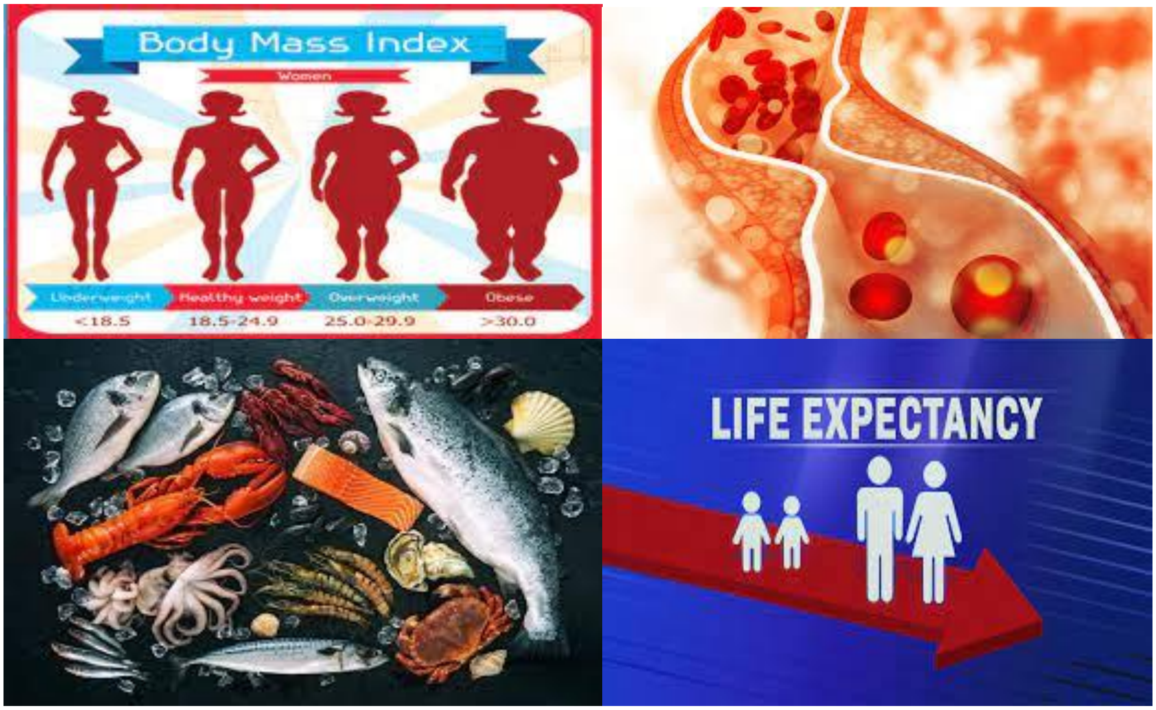


ANALYSIS OF FACTORS AFFECTING HEALTH FOR PUBLIC POLICY

An analysis of Fish consumption, Body Mass Index, and Fat levels in blood vs. Life Expectancy

Affiliation: Mathematics and Actuarial Science Department, Fundamentals of Data Science

Acknowledgment: This study was conducted under the supervision of Dr. Seif El Dawlatly.

Table of Contents

Introduction:

In recent years, healthcare costs have been rising globally given the COVID-19 pandemic and the ever-increasing population of older people, who require more medical attention, in western nations such as the United States and Japan (Tang, Li, & Hu, 2022). Yet simultaneously, recent events such as the Ukraine conflict, the COVID recession, and the 2008 economic crisis all signal the importance of economic stability and optimizing governmental expenditures to be able to prepare for future economic shocks. Hence, the purpose of this study is to examine what factors affect human health, with life expectancy as the measurement of health, in order to allow governments to optimize healthcare expenditures and prioritize specific factors that affect life expectancy the most, potentially enhancing citizen's health as a result. Furthermore, this study aims to spur future research on the topic of factors affecting life expectancy as well, to allow for more and more factors to be covered, hence for more insights to be provided to governments. Overall, in this study, the factors of fish consumption, Body Mass Index (BMI), and cholesterol levels in blood are analyzed with respect to life expectancy, in order to determine whether they impact life expectancy or not and subsequently recommendations were provided concerning public health policy.

Section 1: Data Summary

We have 8 main datasets:

- **BMI (Male and Female)**

  Body Mass Index of the both the male and female population calculated by dividing a person's weight in kg by their height in meters squared (kg/m^2)

- **Fat in blood/ Cholesterol level (Male and Female)**

  The mean total cholesterol of both the male and female population of each country, counted in mmol per L.

- **Life expectancy (Male and Female)**

  The average number of years a newborn child would live if current mortality patterns were to stay the same.

- **Fish and seafood consumption per capita (kg)**

  The average supply of fish and seafood across the population, measured in kilograms per person per year. Food supply is defined as food available for human consumption.

- **Total population for each country**

Each of these datasets provide the according data starting from 1980 until 2008. The averages in the above datasets were calculated as if each country has the same age composition as the world population and so is not extremely accurate given the sizeable sample. The datasets offer data for about 200 countries which we shrunk down to 40 using random sampling where we chose 8 main world regions and chose 2-6 countries from each one and calculated their average for each year which allowed for a variety and diversity of countries without the large sample size. The data was described using basic statistics such mean, median, standard deviation, and quartiles. In this model, life expectancy is assumed to be the main measure or indicator of good or bad health.

Section 2: Analyzing the relationship between fish consumption and life expectancy

In this comparison, the outcomes of dietary habits on health will be examined through discussing how the inclusion of seafood in people's nutrition plans has an effect on life expectancy. This is because generally a person with less vulnerability to illness is more likely to live longer, if we exclude other causes of death such as: accidents or murder. Concisely, this makes fish consumption a factor affecting health and life expectancy an indicator of health levels.

Our main objective from this report is to predict how some regions' medical costs are going to change in the future. Therefore, we have to analyze the datasets, detecting relationships between fish consumption and life expectancy. This is vital for drawing conclusions on each region's health over the period between 1980 and 2008.

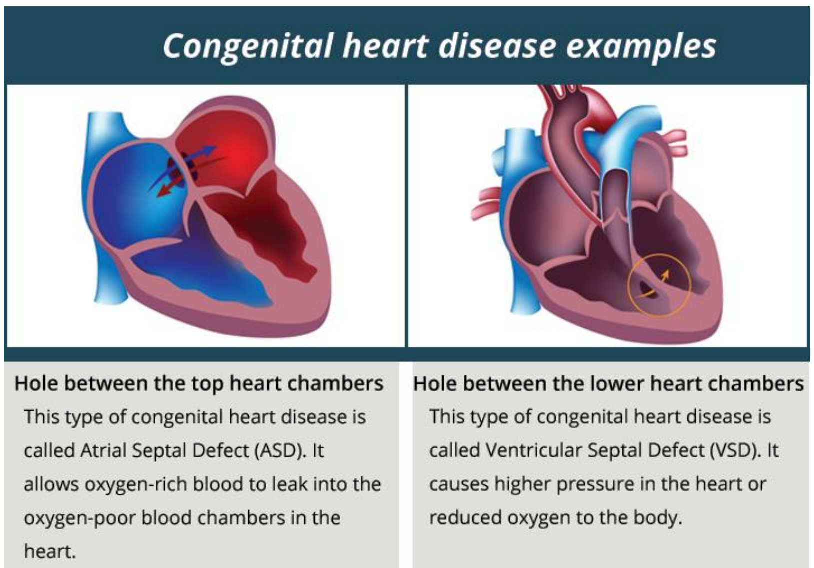
Before analyzing the datasets, retrieving key background information from other research papers is pivotal in order to evaluate the accuracy of our sample datasets in representing the whole population. This will be achieved through comparing the correlation between fish consumption and life expectancy we will deduce to the relationship between these two aspects in the other credible sources we consult.

According to The American Journal of Cardiology, there were "19 observational studies with 228,864 participants, representing a large and diverse population sample" conducted by Dr Paul Whelton and his colleagues about the effects of fish consumption on human health. When discussing the results of these studies, the researchers found that "consumption is associated with an approximately 20% reduction in the risk of fatal CHD1 and a 10% reduction in total CHD" (Whelton et al., 2004). Emphasizing on the significance of CHD's effects on life expectancy, the

[1] Congenital Heart Disease

peer-reviewed journal eClinicalMedicine produced an article that highlights how “CHD is the leading cause of deaths from non-communicable diseases (NCDs) in those under 20 years.” and that “Global CHD deaths in 2019 were 217,000” (Su et al., 2022). Scrutinizing these pieces of statistical evidence, it can be deduced that fish consumption proves to decrease the risks of suffering from a major cardiovascular illness, thus increasing life expectancy which portrays a positive relationship between the consumption of fish and health represented by life expectancy.

Figure 1: congenital heart disease example.

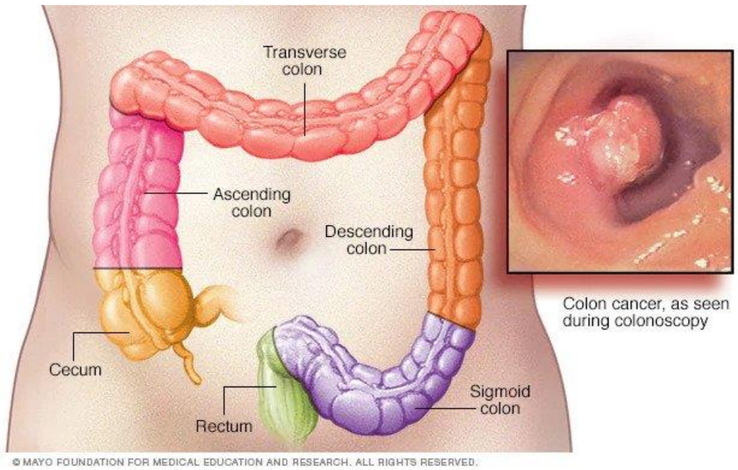


Additionally, another lethal illness is cancer which killed about “6 million” people in the year 2000 (Parkin, 2001). Relating cancer to fish consumption, “fish and fish oil are a rich source of n–3 fatty acids and the fat-soluble vitamins A and D. The n–3 fatty acids are important components of cell membranes (35) and appear to have anti-inflammatory effects and inhibit in vitro[2] the growth of colon, breast, and prostate cancers” (Fernandez et al., 1999). Therefore, this example clearly pinpoints how fish consumption works towards increasing health and life

[2] Trying to stop it from growing or working properly in a lab outside of a living body

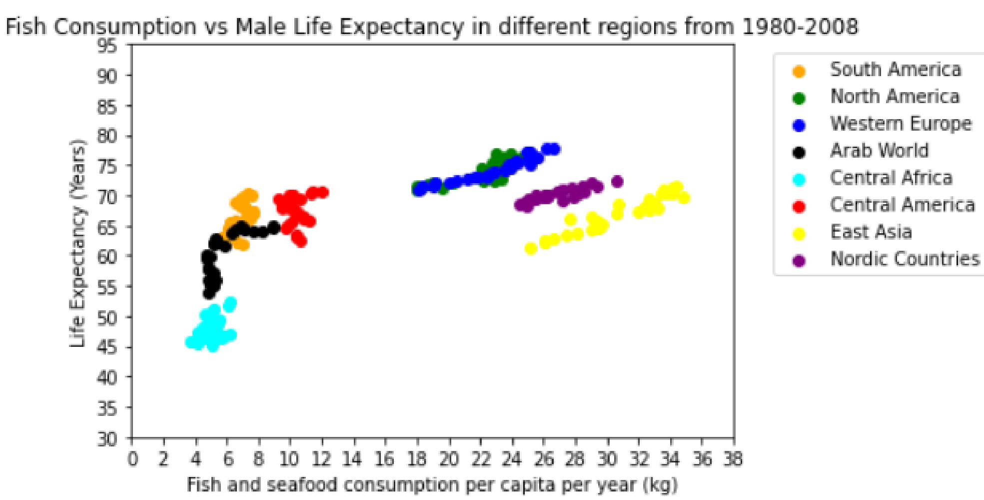
expectancy since all people of different ages are vulnerable to cancer. So, this reinforces the fact that fish consumption and life expectancy (and therefore health) are positively correlated.

Figure 2: Colon Cancer



Since covariance is a measure of correlation, we will calculate the covariance of fish consumption and life expectancy for males and females for each region in our datasets and evaluate the relationship accordingly.

Figure 1: Fish Consumption vs Life Expectancy of males in different regions from 1980 to 2008



- **Covariance calculations for males:**

Cov(x,y) = $\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$

Let x = Fish and seafood consumption per capita per year (kg) for a specific region

Let y = Life expectancy (years) per year for the same specific region for males

N = 29 (2008-1980= 28+1(including 1980) = 29)

South America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 6.789138)(y_i - 66.2977) \approx 1.002$

North America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 22.0069)(y_i - 73.94138) \approx 2.750$

Western Europe = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 22.82408)(y_i - 74.33908) \approx 5.085$

Arab World = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 5.952966)(y_i - 60.41034) \approx 3.681$

Central Africa = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 4.999272)(y_i - 48.1181) \approx 0.579$

Central America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 10.43236)(y_i - 67.43563) \approx 0.177$

East Asia = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 30.82966)(y_i - 67.03586) \approx 8.420$

Nordic Countries = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 26.84996)(y_i - 70.12701) \approx 1.558$

- **Correlation calculations for males:**

$$Correlation = \frac{Cov(x\,,y)}{\sigma_x * \sigma_y}$$

$Cov(x\,,y) = Covariance\ between\ x\ and\ y$

σ_x = standard deviation of fish consumption (kg)

σ_y= standard deviation of life expectancy of males (years)

South America $\approx \frac{1.002}{0.619 * 2.623} \approx 0.617$

North America $\approx \frac{2.750}{1.888 * 1.863} \approx 0.782$

Western Europe $\approx \frac{5.085}{2.574 * 2.177} \approx 0.907$

Arab World $\approx \frac{3.681}{1.366 * 3.875} \approx 0.695$

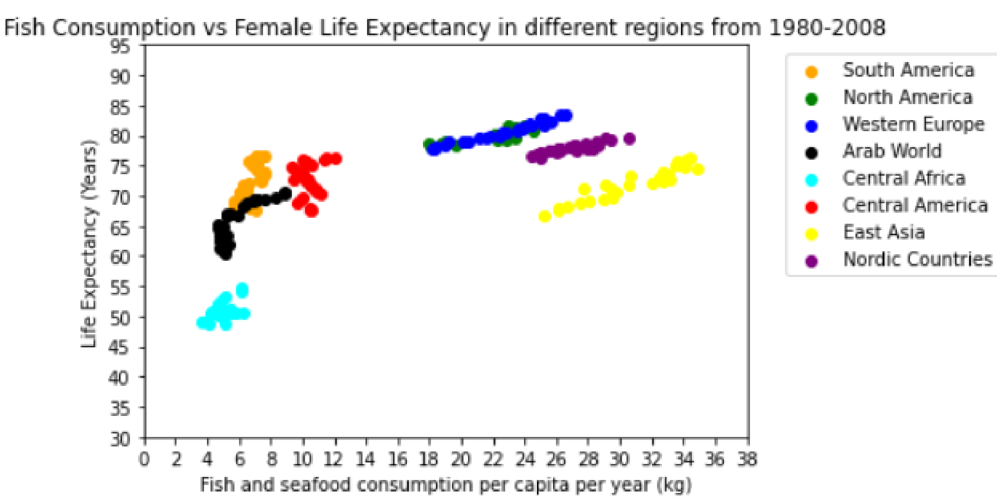
Central Africa $\approx \frac{0.579}{0.679 * 2.081} \approx 0.410$

Central America $\approx \frac{0.177}{0.638 * 2.353} \approx 0.118$

East Asia $\approx \frac{8.420}{2.969 * 3.122} \approx 0.908$

Nordic Countries $\approx \frac{1.558}{1.773 * 1.169} \approx 0.752$

Figure 4: Fish Consumption vs Life Expectancy of females in different regions from 1980 to 2008



- **Covariance calculations for females:**

Cov(x,y) = $\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$

Let x = Fish and seafood consumption per capita per year (kg) for a specific region

Let y = Life expectancy (years) per year for the same specific region for females

N = 29 (2008-1980= 28+1(including 1980) = 29)

South America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 6.789138)(y_i - 66.2977) \approx 1.156$

North America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 22.0069)(y_i - 73.94138) \approx 1.537$

Western Europe = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 22.82408)(y_i - 74.33908) \approx 4.300$

Arab World = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 5.952966)(y_i - 60.41034) \approx 3.525$

Central Africa = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 4.999272)(y_i - 48.1181) \approx 0.534$

Central America = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 10.43236)(y_i - 67.43563) \approx 0.200$

East Asia = $\frac{1}{29}\sum_{i=1}^{N}(x_i - 30.82966)(y_i - 67.03586) \approx 7.492$

- **Correlation calculations for females**

$$Correlation = \frac{Cov(x\,,y)}{\sigma_x * \sigma_y}$$

$Cov(x\,,y) = Covariance\ between\ x\ and\ y$

σ_x = standard deviation of fish consumption (kg)

σ_y= standard deviation of life expectancy of females (years)

South America $\approx \frac{1.156}{0.619 * 2.981} \approx 0.626$

North America $\approx \frac{1.537}{1.888 * 1.006} \approx 0.809$

Western Europe $\approx \frac{4.300}{2.574 * 1.773} \approx 0.812$

Arab World $\approx \frac{3.525}{1.366 * 3.447} \approx 0.749$

Central Africa $\approx \frac{0.534}{0.679 * 1.636} \approx 0.481$

Central America $\approx \frac{0.200}{0.638 * 2.819} \approx 0.111$

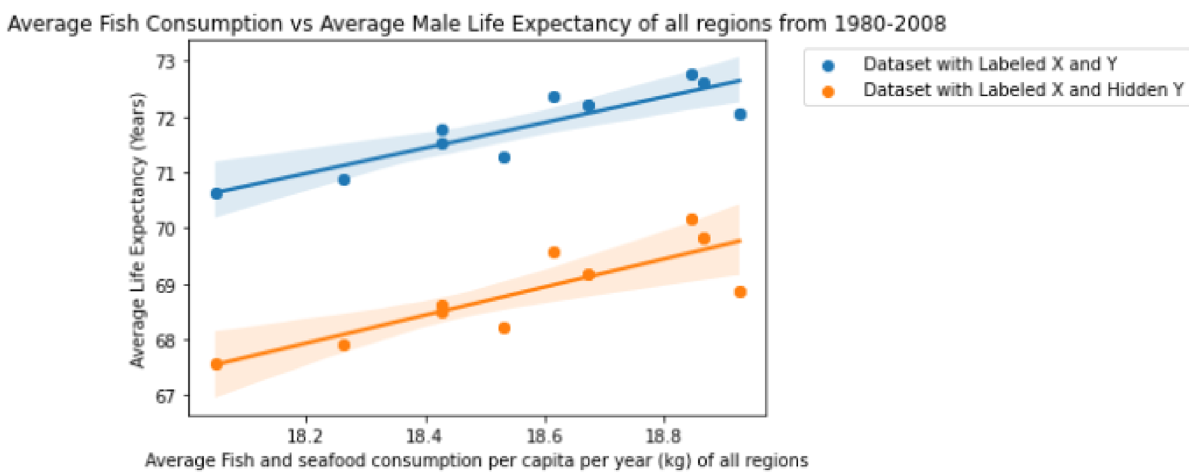
East Asia $\approx \frac{7.492}{2.969 * 2.791} \approx 0.904$

Nordic Countries $\approx \frac{1.284}{1.773 * 0.939} \approx 0.771$

Mainly, the above graphs and calculations aimed to portray and draw connections between fish consumption and life expectancy over the period between 1980-2008 in order to be able to deduce the relationship between fish consumption and health. Eventually, this will guide our research towards its main purpose of predicting the medical costs of different regions in the future. From the scatter plots, most of the regions shed light on the strong positive relationship between fish consumption and the life expectancy of both genders. However, this graphical representation had to be confirmed or refuted using quantitative calculations of conventional measures of the relationship between two variables. First, the covariance between fish consumption and life expectancy for each region and both genders was calculated to get an overview of the direction of the relationship between both variables. Since all values were positive, this indicated a positive relationship between fish consumption and life expectancy, thus complying with our background research.

Although covariance also gives a general idea about the strength of the relationship, finding the correlation coefficient seemed like a better option since the range of the variance is unbounded, unlike the correlation coefficient whose values (r) lie in the interval $-1 \leq r \leq 1$. Using the covariance and standard deviations of the variables (life expectancy once for males and once for females) to find the correlation, the results were remarkable. For North America, Western Europe, East Asia and the Nordic countries, their correlation coefficients (for females and males) fell in the range $0.7 \leq r \leq 0.9$ or higher indicating a strong positive relationship between fish consumption and life expectancy. Additionally, the Arab world is approximately in this range of a strong positive relationship. Concerning the moderate positive correlation range $0.4 \leq r \leq 0.7$ (Akoglu ,(2018)), South America and Central Africa fall in this range. On the other hand, Central America is the only region where the correlation coefficient is in the weak relationship range where r is close to 0.1. Relating to our original purpose from this research paper, the majority of regions' correlation coefficients pinpoint that fish consumption prevents major health dilemmas (where health is indicated by increased life expectancy), which foreshadows that fish consumption plays a role in maintaining healthy citizens and reducing the medical costs of the regions who consume the greatest seafood quantities such as Nordic countries and East Asia that have an average fish consumption of about 30 kilograms per capita per year. Thus, this encourages these regions to maintain or increase its fish consumption. This also acts as an incentive for the governments of regions with low fish consumption and life expectancy like Central Africa to perhaps reduce the price of local seafood in order to encourage citizens to increase their consumption, for the sake of benefiting citizens with healthy diets as well as reducing the region's total medical costs.

Figure 5: Linear regression



For this regression model, we utilized the 'sklearn' package in order to apply supervised learning on two new excel sheets we created from the existent dataset we had. The first one contained the average of all regions' fish consumption in one row, a cell for each year from 1990 to 2018, and another excel sheet having the average of all regions' male life expectancy in one row, a cell for each year 1990 to 2018. For the training phase, we gave the compiler average values of fish consumption and its corresponding average values of male life expectancy for the years 1999-2008. After completing this part, we started providing python with the test values of average fish consumption for all regions in the years 2009-2018 and asked it to predict the corresponding average male life expectancy values from 2009-2018, while hiding the actual average male life expectancy values for these years. The blue regression line and points represent the actual values x and y-values present in the data frames. As for the orange points and regression line, they represent the actual x-values plotted against the predicted y-values by the machine learning algorithm after providing it with x test values. Although the predicted regression line and points seems to deviate from the values of the actual regression line, the machine learning algorithm successfully recognized and plotted the strong positive relationship between average fish consumption and average life expectancy. Perhaps, this discrepancy is caused by the fact that the training data was for 10 years, which is a small sample to create an algorithm with high prediction accuracy. Additionally, the difference in values could also be influenced by the fact that we took into consideration 2-6 countries from each region which cannot yield completely accurate results for the whole region. Last but not least, plotting average values could have affected the accuracy of the model in the sense that average values do not represent the maxima and minima of the data, meaning that it only represents a sector of these regions' citizens not the whole population. Finally, this regression model relates to our data and main purpose by highlighting how there is a positive relationship between fish consumption and health (indicated through life expectancy), proving how governmental policies to increase demand on fish consumption would not only reduce total medical costs but would also improve their citizens' health and productivity over time.

Section 3: Analyzing the relationship between Body Mass Index (BMI) and life expectancy

Another factor which may influence one's life expectancy is one's weight, where weighing more and being obese is generally correlated with chronic diseases such as diabetes and heart disease (Walls & Backholer, 2012). In this analysis, the variable utilized as a measure of weight is the Body Mass Index (BMI), where the average BMI for both genders was calculated for each region per year, from 1980 to 2008. This was done given that 'weight' as a research variable, is not a standardized one, meaning that the health effects of weighing 150kg for a person that is 6'1 could be completely different for a person that weighs the same but whose height is 5'6. In terms of analytical methods used, the relationship between BMI and life expectancy was evaluated using mean BMI, and life expectancy data, scatter plots, and a regression analysis for each region's respective average BMI and life expectancy trends.

Figure 6A: Preliminary analysis using average BMI and life expectancy for males for the following years: 1980, 1990, 2000, 2008

Table A: Males' BMI Data

	1980	1990	2000	2008
SA	23.700000	24.333333	25.283333	26.000000
NA	25.450000	26.350000	27.350000	27.950000
WE	25.166667	25.533333	26.166667	26.783333
AW	23.760000	24.360000	25.160000	25.860000
CA	20.275000	20.550000	21.025000	21.600000
CAM	23.316667	23.966667	24.733333	25.533333
EA	21.360000	21.580000	22.180000	22.660000
NC	24.916667	25.133333	25.500000	26.250000

Table B: Males' Life Expectancy Data

	1980	1990	2000	2008
SA	62.366667	64.950000	68.150000	70.650000
NA	71.200000	73.000000	75.350000	77.050000
WE	71.200000	73.066667	75.816667	78.166667
AW	55.440000	58.060000	64.280000	65.720000
CA	46.750000	46.400000	48.875000	52.850000
CAM	63.366667	66.750000	69.250000	70.800000
EA	62.160000	65.720000	69.320000	71.820000
NC	68.483333	69.816667	70.533333	72.716667

An apparent trend in the above tables, is that both the BMI and the life expectancy averages have increased over the 28-year period from 1980 to 2008. Indicating that there is a potential positive correlation between, meaning that both increase together. Further insights include that the Western Europe (WE) and North America (NA) have both the highest BMI's and life expectancies. On the contrary, Central Africa (CA) has the lowest BMI and life expectancy over the time period. While the Arab World (AW) has the 2nd lowest life expectancy, even though it ranks amongst the higher regions in terms of BMI.

Figure 6B: Preliminary analysis using average BMI and life expectancy for females for the following years: 1980, 1990, 2000, 2008

Table C: Females' BMI Data

	1980	1990	2000	2008
SA	23.800000	24.683333	26.000000	26.850000
NA	24.650000	25.700000	26.800000	27.500000
WE	24.583333	24.833333	25.200000	25.433333
AW	24.620000	25.520000	26.560000	27.500000
CA	20.225000	20.800000	21.650000	22.350000
CAM	23.966667	25.050000	26.216667	27.283333
EA	21.080000	21.580000	22.140000	22.480000
NC	24.950000	25.166667	25.366667	25.816667

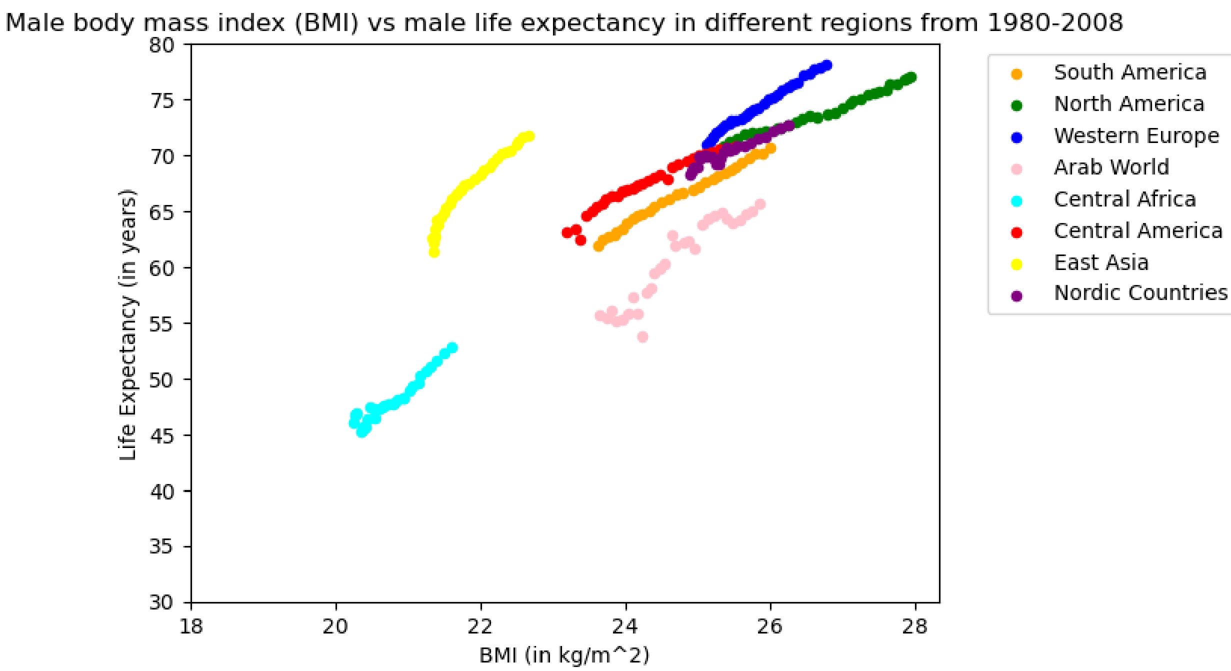
Table D: Females' Life Expectancy

	1980	1990	2000	2008
SA	68.050000	71.366667	75.000000	77.100000
NA	78.450000	79.700000	80.600000	81.800000
WE	77.900000	79.916667	81.983333	83.516667
AW	60.940000	63.700000	68.820000	70.780000
CA	50.600000	50.650000	50.925000	55.425000
CAM	68.466667	71.816667	74.483333	76.033333
EA	67.000000	70.600000	74.100000	76.960000
NC	76.600000	77.516667	78.216667	79.900000

Unlike the case with males, the Nordic countries, Russia, and Ukraine region (NC) has the highest BMI and the third highest life expectancy for females. As for the NA and WE regions, they still rank high in terms of BMI, and represent the top regions for life expectancy. Similar to the case with males however is the AW region, where they have a rather high average BMI, yet the second lowest life expectancy, and the CA region, which is the lowest in terms of both variables. Thus, the fact that NA and WE regions, which have the highest level of economic development in the world, have high BMI and life expectancy, indicates that there are possibly economic considerations in the relationship between BMI and life expectancy. Same applies for AW and CA regions, as they include some of the least economically-developed nations in the world, or nations that have been war-torn, hence, economically depleted such as Iraq.

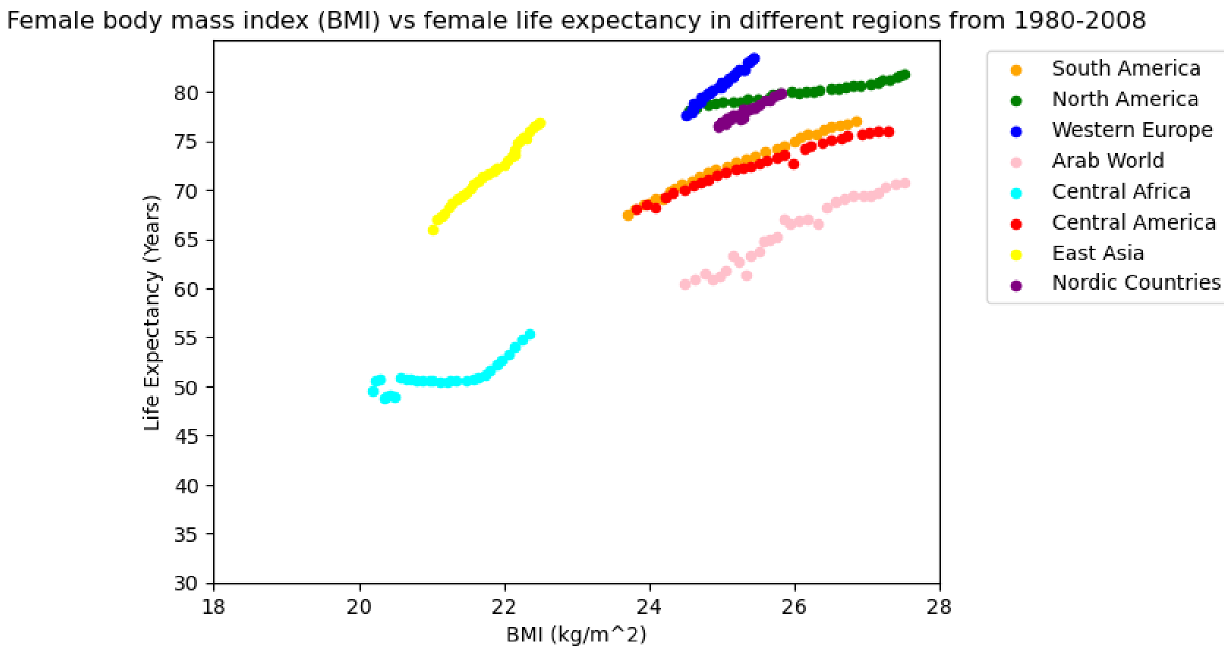
In order to validate if there is a positive correlation that exists between BMI and life expectancy, a scatter plot analysis, covariance analysis, and regression analysis were all conducted.

Figure 7A: Scatterplot depicting the relationship between male BMI and male life expectancy for all regions over the 28-year time period.



The above scatter plot shows that there is an apparent upward trend in the relationship between BMI and life expectancy over time for all regions, where both BMI and life expectancy have been increasing. Consequently, this supports the notion that a positive correlation exists between both variables. Furthermore, as was the case with the above tables, NA and WE represent the regions with the highest BMI's and life expectancies, while CA represents the lowest in both.

Figure 7B: Scatterplot depicting the relationship between female BMI and female life expectancy for all regions over the 28-year time period.



In terms of female's BMI and life expectancy, the same upward trend applies as was the case with male's data. Yet again, NA and WE represent the highest regions in terms of both BMI and life expectancy, with the addition of NC as one of the highest, and CA represents the lowest in both. AW region also exhibits the same behavior as the other regions, where BMI and life expectancy increase together, and albeit it has comparable BMI to NA region, it has a much lower life expectancy. This observation indicates that there are other factors, potentially economic and healthcare related as stated above that affect the relationship between BMI and life expectancy, potentially masking some of the effects of having a high BMI and being obese on life expectancy. In order to accurately test whether there is strong positive correlation between BMI and life expectancy, the relationship between both was analyzed using Ordinary Least Squares (OLS) regression.

Table 2: Coefficient of determination R^2 data by region, between BMI and life expectancy, for both genders.

Region	Male R^2	Female R^2
South America (SA)	0.993	0.991
North America (NA)	0.984	0.975
Western Europe (WE)	0.994	0.988
Arab World (AW)	0.880	0.957
Central Africa (CA)	0.923	0.691
Central America (CAM)	0.950	0.977
East Asia (EA)	0.948	0.990
Nordic Countries + Russia, Ukraine (NC)	0.902	0.953

From the data produced using the regression model, it is apparent that all R^2 values are greater than 0.69. Thus, there is in fact a strong positive correlation that exists between BMI and life expectancy. This can be further demonstrated in the following graphs representing the line of best fit for NA, and CA for both males and females as examples for the overall trend demonstrated for all regions.

Figure 8A: Male regression model graph for NA and CA regions

Figure A: NA Male

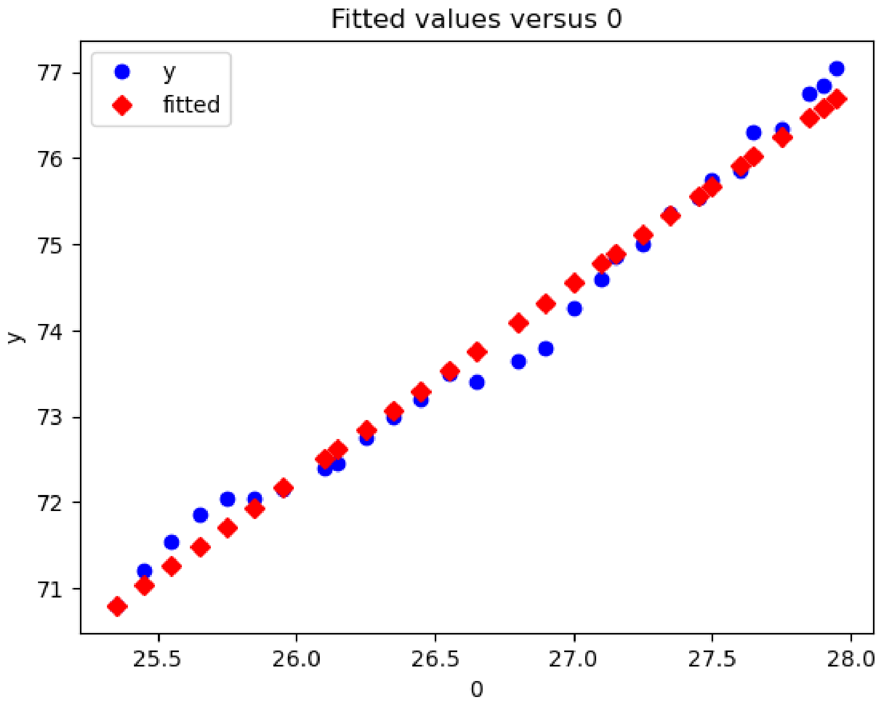


Figure B: CA Male

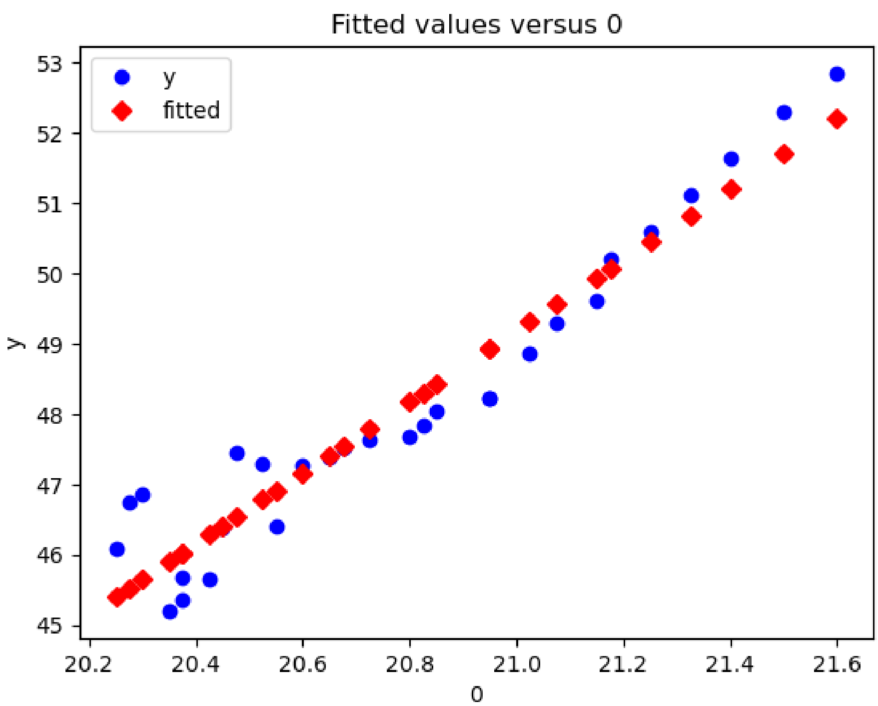


Figure 8B: Female regression model graph for NA and CA regions

Figure A: NA Female

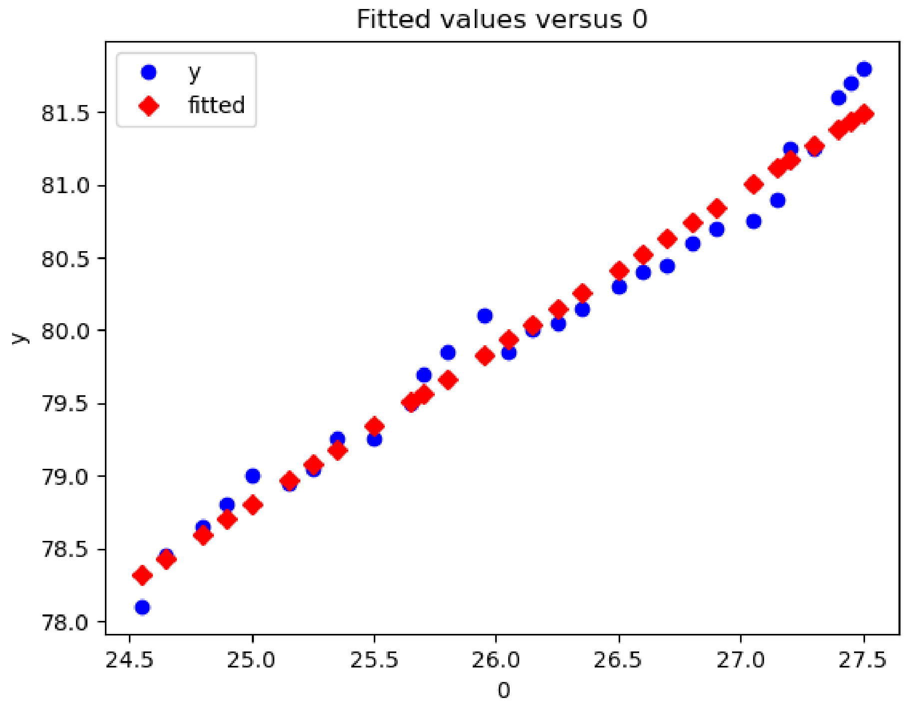
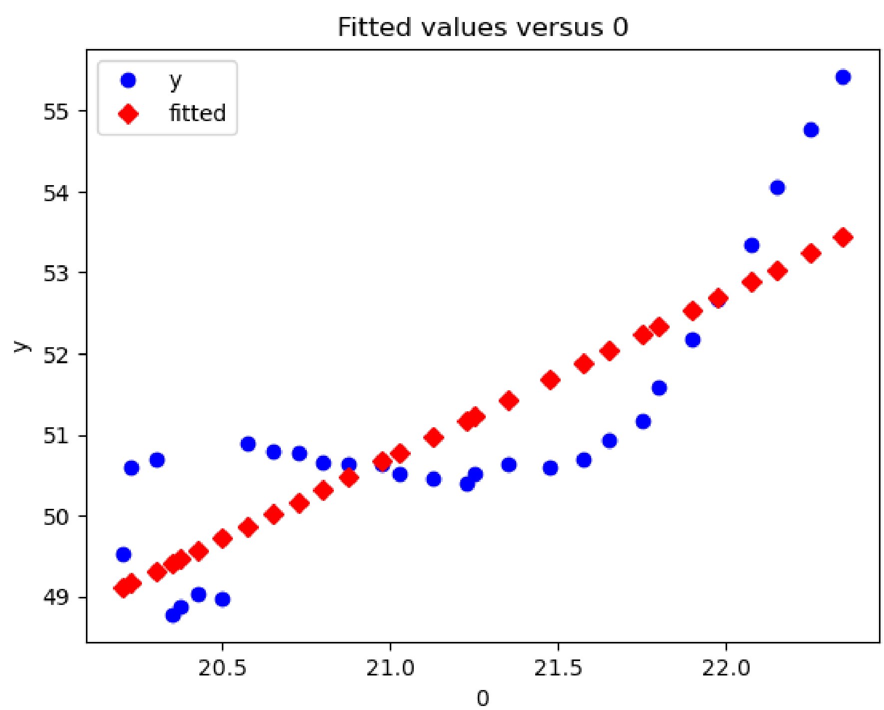


Figure B: CA Female



Furthermore, there is a multiplicity of factors that can mask the negative effects of obesity, as well as misconceptions concerning BMI in general, that are in part causing the above observations and the positive trend between BMI and life expectancy. These factors include healthcare quality, economic development level, and low BMI risk. At first glance, the established positive correlation seems unlikely and rather counter-intuitive, given that obesity and increases in weight are generally believed to cause chronic conditions such as cardiovascular diseases (Fekri & Khaloo, 2020), which in turn increase risk of death and can lower life expectancy. However, that is not the case given that firstly, healthcare quality has been increasing in recent decades, leading to higher quality medications for diseases resulting from obesity or weight gain such as heart disease and diabetes. Thus, the negative effects of obesity have been offset by advances in medicine (Walls & Backholer, 2012) to the point that the effects of obesity on life expectancy are no longer significant. Secondly, economic development level per region influenced that region's BMI and life expectancy, for example, NA and WE, which are the most economically developed, had the highest BMI and life expectancy in general. While, AW and CA, which are less developed, had lower life expectancy, even though AW had high BMI. This is the case, as in AW and CA, economic development is not sufficient for healthcare quality to improve to the extent where it can offset the effects of obesity, unlike NA and WE. Furthermore, an explanation behind why low BMI can be correlated to lower life expectancy, as is the case with CA, is that severely low BMI is associated with malnutrition and weakened immune systems, hence a higher risk of diseases and lower life expectancy overall (Walls & Backholer, 2012). This is the case with CA countries specifically given their lower economic development and history of war and famines (Gebre-Egziabher Kiros, 2001) (Bain & Awah, 2013).

An additional factor to consider that is potentially behind the positive correlation trend is high BMI for older individuals. Studies have shown that the acceptable BMI range actually increases with age, meaning that it is recommended for older people to have a higher BMI (Walls & Backholer, 2012) (U.S National Library of Medicine , n.d.). This is the case as having higher fat levels when older can aid in strengthening one's immunity against aging-related conditions such as osteoporosis; 'thinning of the bones' (U.S National Library of Medicine , n.d.). Finally, an additional factor to consider is that having a high BMI or being obese is not directly correlated with mortality, which would cause a reduction in life expectancy, rather is correlated with diseases which may or may not cause death.

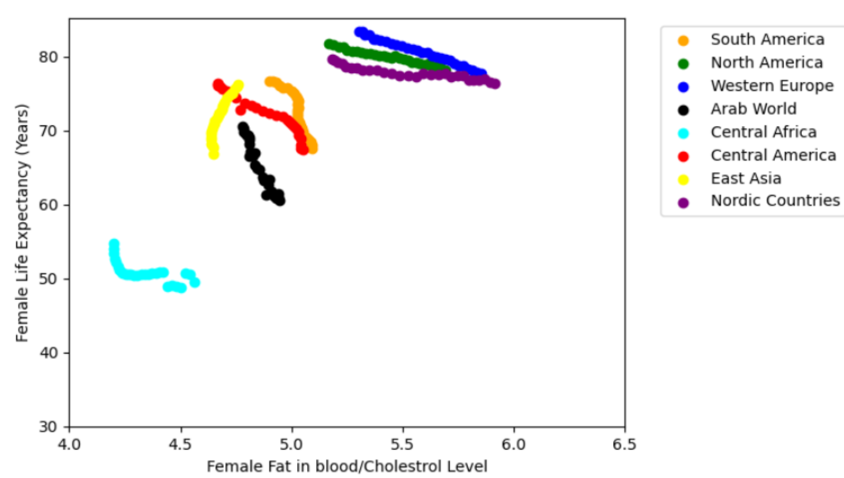
Overall, there is a positive relationship between BMI and life expectancy, where factors such as healthcare quality improvements, have led to the impacts of having a high BMI on life expectancy being counteracted. Furthermore, economic state of a region is an influential factor in determining the level of healthcare in that region, hence, the extent to which healthcare can offset the impacts of high BMI on life expectancy. Finally, it is important also to remember that low BMI is correlated with low life expectancy given that a low BMI indicates malnutrition and a weakened immunity, similarly, a higher BMI is actually healthy for older individuals, hence this is potentially part of the reason why a high BMI is correlated to high life expectancy.

Section 4: Life Expectancy vs Fat level Analysis

The relationship between life expectancy and the level of fat in blood, specifically cholesterol, is a crucial aspect of a healthy, long life. Cholesterol is a lipid molecule that plays a vital role in the body's physiological functions, but elevated levels of certain types of cholesterol, such as LDL cholesterol, can increase the risk of developing heart disease and other cardiovascular conditions which is ultimately a significant factor in life expectancy and long-term health.

Elevated levels of LDL cholesterol can contribute to the development of atherosclerosis, a condition characterized by the accumulation of fatty deposits in heart arteries. Over time, this can lead to reduced blood flow to vital organs, including the heart and brain, increasing the risk of heart attacks, strokes, and other cardiovascular, as well as neurological conditions. Several studies have consistently demonstrated a clear association between elevated cholesterol levels and increased mortality from cardiovascular disease. High levels of LDL (low-density lipoprotein) cholesterol have been identified as a major risk factor for the development of atherosclerosis (restricted blood flow) and subsequent cardiovascular events. On the other hand, maintaining healthy cholesterol levels, like having low LDL cholesterol and high HDL cholesterol, is associated with a reduced risk of cardiovascular disease and improved life expectancy.

Figure 9: female cholesterol level vs. female life expectancy from 1980 to 2008



- **Correlation and covariance calculations for female cholesterol vs. female life expectancy**

$$Correlation = \frac{Cov(x\,, y)}{\sigma_x * \sigma_y}$$

$Cov(x\,, y) = Covariance\ between\ x\ and\ y$

σ_x = standard deviation of fat level in blood (mmol/L)

σ_y= standard deviation of life expectancy of females (years)

South America $\approx \frac{-0.132}{0.044 * 2.981} \approx$ -1.006

North America $\approx \frac{-0.149}{0.161 * 1.006} \approx$ -0.920

Western Europe $\approx \frac{-0.298}{0.161 * 1.773} \approx$ -1.044

Arab World $\approx \frac{-0.177}{0.070 * 3.447} \approx$ -0.727

Central Africa $\approx \frac{-0.118}{0.074 * 1.636} \approx$ -0.975

Central America $\approx \frac{-0.381}{0.091 * 2.819} \approx$ -1.485

East Asia $\approx \frac{0.098}{0.030 * 2.791} \approx$ -1.170

Nordic Countries $\approx \frac{-0.197}{0.224 * 0.939} \approx$ -0.937

$$Covariance = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$$

μx = mean of fat level in male blood (mmol/L)

μy = mean of life expectancy of males (years)

South America ≈ 0.059714550264550484

North America ≈ -0.2761081349206347

Western Europe ≈ -0.32477263378353904

Arab World ≈ -0.24995359788359744

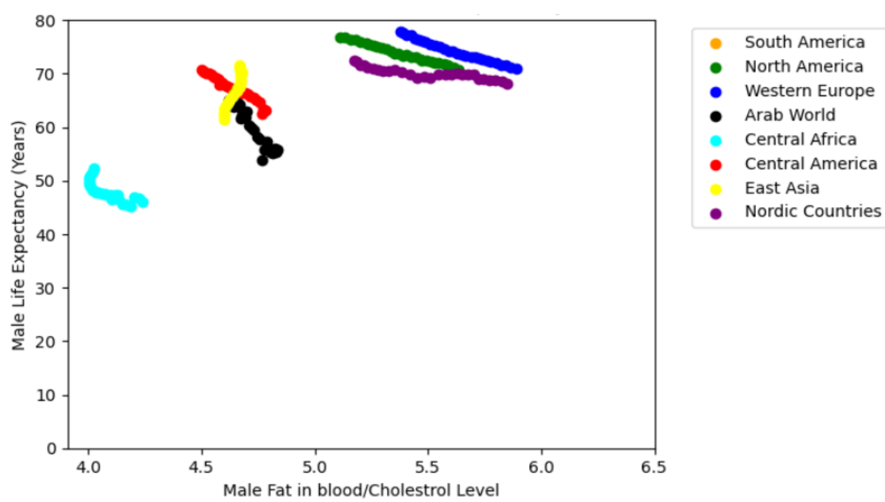
Central Africa ≈ -0.11256258267195768

Central America ≈ -0.20144429766222813

East Asia ≈ 0.08639518518518 56

Nordic Countries ≈ -0.21158098197602893

Figure 10: Male Cholesterol level vs. male life expectancy from 1980 to 2008



- **Correlation and covariance calculations for male cholesterol vs. male life expectancy**

$$Correlation = \frac{Cov(x, y)}{\sigma_x * \sigma_y}$$

$Cov(x, y) = Covariance\ between\ x\ and\ y$

σ_x = standard deviation of fat level in male blood (mmol/L)

σ_y= standard deviation of life expectancy of males (years)

South America ≈ $\frac{-0.060}{0.044 * 2.623}$ ≈ -0.519

North America ≈ $\frac{-0.276}{0.161 * 1.863}$ ≈ -0.920

Western Europe ≈ $\frac{-0.325}{0.161 * 2.177}$ ≈ -0.927

Arab World ≈ $\frac{-0.250}{0.070 * 3.875}$ ≈ -0.922

Central Africa ≈ $\frac{-0.113}{0.074 * 2.081}$ ≈ -0.734

Central America ≈ $\frac{-0.201}{0.091 * 2.353}$ ≈ - 0.939

East Asia ≈ $\frac{0.086}{0.030 * 3.122}$ ≈ 0.918

Nordic Countries ≈ $\frac{-0.212}{0.224 * 1.169}$ ≈ -0.809

$$Covariance = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$$

μx = mean of fat level in male blood (mmol/L)

μy = mean of life expectancy of males (years)

South America ≈ 0.059714550264550484

North America ≈ -0.2761081349206347

Western Europe ≈ -0.32477263378353904

Arab World ≈ -0.24995359788359744

Central Africa ≈ -0.11256258267195768

Central America ≈ -0.20144429766222813

East Asia ≈ 0.0863951851851856

Nordic Countries ≈ -0.21158098197602893

As mentioned above and deduced from the overall trends in the graph, there is a negative correlation between life expectancy and cholesterol levels in blood. As the fat in blood increases, a person's life expectancy decreases as seen for most of the chosen regions. A negative, decreasing slope indicates that high cholesterol levels are directly associated with decreased life expectancy. North America, western Europe, the Arab world, central Africa, central America, the Nordic region all had correlation coefficients between -0.7 and -1 which indicate a strong negative correlation. These values also had mainly negative covariance values where a strong relationship is also signified. However, the correlation coefficient values for east Asia were both positive with a noticeably increasing slope which is an outlier as it completely violates the overall trends for both the male and female curves. All of the correlation and covariance values are negative except that of East Asia. For both male and female scatter plots, their values are positive which indicate a strong increasing relationship which opposes most of the conducted research and studies. However, some research indicates that a higher fat level in blood results in a faster metabolism and, therefore, a better life expectancy. The exact measurements of cholesterol were not clear in the datasets nor was the method of data collection. Therefore, such elevated or increasing cholesterol levels may be HDL (high density lipoprotein) cholesterol which is healthy and which high levels of do not harm ultimately improving life expectancy instead of the opposite as shown in the trend.

According to a study conducted by Mayo Clinic, high levels of unhealthy cholesterol such as low-density lipoprotein, also known as hypercholesterolemia, reduce life expectancy by approximately 15-30 years. However, constant, intense treatment can manage and improve such odds and yield a much better outcome. Efforts to manage cholesterol levels and improve life expectancy involve actions on several fronts. Public health interventions focus on promoting healthy lifestyle choices, including a balanced diet low in saturated and trans fats, regular physical activity, smoking cessation, and weight management. Healthy lifestyle choices, such as avoiding tobacco products, excessive amounts of alcohol and remaining physically active are greatly encouraged. Dietary alterations and appropriate nutrition are also crucial like avoiding saturated fats present in foods, such as dairy products, meat, fast or junk food, and oily pastry. A vegan lifestyle is a viable, effective option for hypercholesterolemia patients. Eliminating trans fats such as vegetable oil in mass produced foods such as crackers, chips, or cookies is also essential as this decreases the amount of unsaturated fat in blood and decreases cholesterol. Medications, such as statins, are commonly prescribed to individuals with elevated cholesterol levels to reduce the risk of cardiovascular events and improve long-term outcomes as high cholesterol or LDL causes incredible strain on the heart as blood flow is restricted due to the unsaturated, indigestible fats in the blood. Health education and awareness campaigns also have a significant role in encouraging cholesterol management and reducing cardiovascular disease-related mortality. These initiatives aim to increase public understanding of the importance of cholesterol monitoring, encourage routine screening, and promote adherence to lifestyle modifications and prescribed medications.

In conclusion, the relationship between life expectancy and the level of fat in the blood, specifically cholesterol, is significant, with elevated levels of LDL cholesterol contributing to an increased risk of cardiovascular disease and mortality. Maintaining healthy cholesterol levels, along with adopting a healthy lifestyle and managing other risk factors, can improve life expectancy and reduce the burden of cardiovascular disease. Public health efforts focusing on cholesterol management, lifestyle modifications, and awareness campaigns are crucial in promoting better cardiovascular health and improving overall longevity.

Conclusion:

Overall, this study, utilizing the aforementioned data sets and a mixture of statistical methods such as regression, and covariance calculations, has established: firstly, a positive relationship between fish consumption and life expectancy, where fish is generally healthy to eat and can strengthen one's immunity against chronic illnesses. Secondly, a positive relationship between BMI and life expectancy, where it has been stated that due to overall improvements in medicine globally, the effects of obesity and having a high BMI have been masked. However, there are certain considerations concerning BMI, such as the fact that having a slightly higher BMI for older people is recommended as a healthy practice, which are potential causal factors for the relationship between BMI and life expectancy. Finally, a negative relationship has been established between life expectancy and cholesterol levels for both genders, where higher levels of cholesterol can cause serious health complications, consequently, they can impact life expectancy negatively.

Limitations:

1. Sampling means that the results of the analysis can never be 100% reflective of the entire population.
2. BMI data does not differentiate between those who have a higher BMI due to having a higher fat content, and those who have a higher BMI due to having higher muscle mass. Which limits the extent to which the analysis can be detailed and thorough.
3. Original per country data are from samples of the population of the respective country itself, indicating that there is an added layer of uncertainty in the data.

However, these limitations do not restrict nor limit the validity of the data overall, given the overall strong coefficients of determination, correlation, and covariance values generated within the data.

Recommendations:

1. We recommend that governments encourage the consumption of fish as much as possible, offering incentives to, and strengthening the fishing industry in their respective nations. This would not only improve overall health, but would also counteract the negative impacts of high cholesterol. this is the case as other protein sources such as beef and eggs are generally known to cause cholesterol to rise more rapidly in comparison to fish.
2. We recommend that governments encourage older people aged higher than 65, to have a slightly higher BMI of between 25 and 28, this would allow them to have better immunity to combat aging-related conditions such as weight loss. In addition, this could potentially even reduce the healthcare burden associated with older individuals, where if it is possible to maintain their health, then they would require less frequent check-ups and overall lower government expenditure.
3. We recommend that governments in less economically developed regions such as the Arab World and Central Africa to also procure higher quality medication related to obesity and obesity-caused conditions such as diabetes and heart conditions. This would allow these nations to offset the effects obesity has on life expectancy, increasing their citizen's overall life expectancy as is the case with the North America and Western Europe regions.
4. Finally, we recommend that governments engage in extensive anti-cholesterol health campaigns, encouraging citizens to adopt a healthier lifestyle and to avoid trans fats and fatty foods in general.

References:

Bain, L., & Awah, P. (2013). Malnutrition in Sub – Saharan Africa: burden, causes and prospects. *The Pan African Medical Journal*.

Fekri, N., & Khaloo, P. (2020). Association of body mass index with life expectancy with and without cardiovascular disease. *International Journal of Obesity*.

Gebre-Egziabher Kiros, D. P. (2001). War, famine and excess child mortality in Africa: the role of parental education. *International Journal of Epidemiology* .

Tang, B., Li, Z., & Hu, S. (2022). Economic Implications of Health Care Burden for Elderly Population. *Sage Journal*.

U.S National Library of Medicine . (n.d.). *Body mass index: MedlinePlus*. Retrieved from MedlinePlus Medical Encyclopedia: https://medlineplus.gov/ency/article/007196.htm#:~:text=In%20older%20adults%20it%20is,of%20the%20bones%20(osteoporosis).

Walls, H., & Backholer, K. (2012). Obesity and Trends in Life Expectancy. *Hindawi Journal of Obesity*.

Medical Knowledge Tream, Ada's. "Familial Hypercholesterolemia." *Ada*, 6 July 2022, ada.com/conditions/familial-hypercholesterolemia/.

Staff, Mayo Clinic. "High Cholesterol." *Mayo Clinic*, 11 Jan. 2023, www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/symptoms-causes/syc-20350800#:~:text=With%20high%20cholesterol%2C%20you%20can,a%20heart%20attack%20or%20stroke.

Staff, Mayo Clinic. "Can Lifestyle Changes Benefit Your Cholesterol?" *Mayo Clinic*, 2 Sept. 2022, www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/in-depth/reduce-cholesterol/art-20045935.

Taylor, William C. "Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity ..." *Annals of Internal Medicine*, 2007, www.acpjournals.org/doi/abs/10.7326/0003-4819-146-5-200703060-00004.