What is the correlation between the human development index (HDI) and the mortality rate due to coronary artery disease (CAD) per 100,000 people in countries with an HDI above 0.75 for the years 2000 and 2019?

Introduction

As an aspiring biologist and data scientist, I have had a passion for understanding and addressing global health challenges leading me to delve into research for diverse illnesses and diseases. From the previous COVID-19 pandemic to prevalent issues like HIV and diabetes, my drive to understand and contribute to global healthcare has become greater. This is why I pursued an internship at the Fred Hutchinson Cancer Center, where I had the privilege of working alongside dedicated medical researchers and I gained valuable insights into the important of data-driven research and the need to explore diseases and healthcare from a broader perspective. These experiences sparked my curiosity about Coronary Artery Disease (CAD) and its impact on different populations all over the world.

Coronary Artery Disease or C.A.D is a common form of heart disease that is caused by atherosclerosis, which is when the coronary arteries have a buildup of plaque. As the plaque accumulates, the human body counterattacks by delivering monocytes or white blood cells to the cholesterol in the plaque. However, these initial white cells transform into foam cells which cause inflammation in the arteries. Foam cells also known as lipid-laden macrophages, are cells that contain cholesterol and can accumulate along the arterial walls. In some cases, these foam cells can break open and cause platelets in the arteries to stick to the injury which can form blood clots. Blood clots in the arteries can limit or prevent blood flow from occurring which can lead to less nutrients and oxygen getting delivered to the heart. This can cause heart attacks, angina, nausea, pain or discomfort, and weakness in the arms and shoulders.

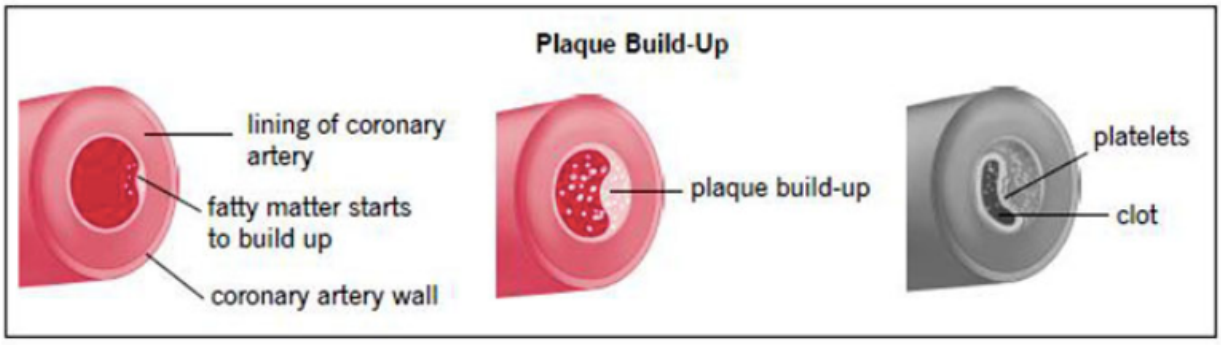


Figure 1: Plaque Build-up Demonstration

According to the World Health Organization and the CDC, cardiovascular diseases are currently the leading cause of death with almost 17 million people dying globally every year. Doctors can perform electrocardiograms, echocardiograms, chest x-rays, cardiac catheterization, and coronary angiograms to diagnose different heart diseases and can also use interventional procedures such as a balloon angioplasty, or surgical methods such as a coronary artery bypass graft to repair and treat the plaque buildup in the coronary arteries. Unfortunately, these treatments tend to be very expensive and can cause countries and individuals a large amount of money. For example, the United States spent close to 363 billion dollars from 2016 to 2017 just on heart diseases.

To understand further how socioeconomic factors can impact the prevalence and outcome of CAD I decided to start an investigation that explores the correlation between the Human Development Index (HDI) and Coronary Artery Disease (CAD) in countries with an HDI above 0.75, specifically focusing on the years 2000 to 2019 to see how the disease and healthcare has been changed over time. The Human Development Index or HDI was created to capture the standard of life in each country by combining a countries health, economic, and education conditions into a number from 0 to 1. Typically, a high HDI value shows that there is a high standard of living in that country and vice versa for a lower HDI. An HDI value of 0.75 is considered a reasonable indicator of a developed country.

Ultimately, this investigation will help investigate how to effectively combat CAD in high HDI countries and will hopefully show whether or not HDI has a positive or negative impact on CAD.

Hypothesis

According to the World Health Organization, an alarming 75% of deaths caused by heart diseases occur in countries with lower levels of development. The HDI value was developed as a way to measure overall human development, including factors such as education, life expectancy, and income. This suggests that a lower HDI value means a country is less developed whereas a higher HDI value means a more developed country. According to the United Nations Development Index, an HDI value of 0.75 is considered a developed country. Drawing from this evidence, it is reasonable to hypothesize that a developed country would have a lower mortality rate caused by CAD.

Null Hypothesis: There will be no correlation between countries with an HDI value above 0.75 and the mortality rate of coronary artery disease.

Alternative Hypothesis: There will be a negative correlation between countries with an HDI value above 0.75 and the mortality rate of coronary artery disease.

Materials

- Microsoft Excel Program
  - The MS Excel Program was used throughout my investigation to create and organize my data tables, calculate all values, and plot and process my data.
- World Health Organization
  - The World Health Organization or WHO has a database filled with the mortality rate due to coronary artery disease. This data base was selected since it is reliable and has well organized information.
  - (https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death)
- HDI Database
  - The Human Development Reports created by the United Nations Development Programme has a data center for HDI rankings that was used to get all countries

with an HDI above 0.85. This data base was selected since it is reliable and has recent data from which couldn't be found on other sites.

- (https://hdr.undp.org/en/countries)

Methodology and Trial Investigation

The following trial investigation was conducted to ensure that there is a correlation between the two variables, HDI and mortality rate per 100,000 for CAD. To select these countries a random number generator was used. The following countries were selected:

Countries: Greece, Belarus, the United States, Thailand, and Italy.

First, a raw data table for my trial investigation is created that compiles the countries names, HDI values, and CAD rates for the years 2000 and 2019.

Table 1: The number of HDI Values, CAD Mortality rate per 100,000, and randomly selected countries

Country	Human Development Index		Estimated CAD Mortality Rate per 100,000	
	2000	2019	2000	2019
Greece	0.810	0.888	256.13	201.18
Belarus	0.712	0.823	536.07	562.67
United States	0.891	0.926	214.42	153.47
Thailand	0.653	0.777	55.520	73.710
Italy	0.841	0.892	208.63	160.26

3 significant figures will be used for HDI values since it provides an accurate representation of the index. For the CAD mortality rate, the rate is recorded with 5 significant figures to ensure the values go till the 100ths place for accuracy.

After Table 1, an average of the years 2000 and 2019 were taken for both the Human Development Index and the estimated CAD mortality rate per 100,000 people. To calculate the averages the values in 2000 and 2019 were added and divided by 2.

Sample Calculations: Greece's HDI value average of 2000 and 2019

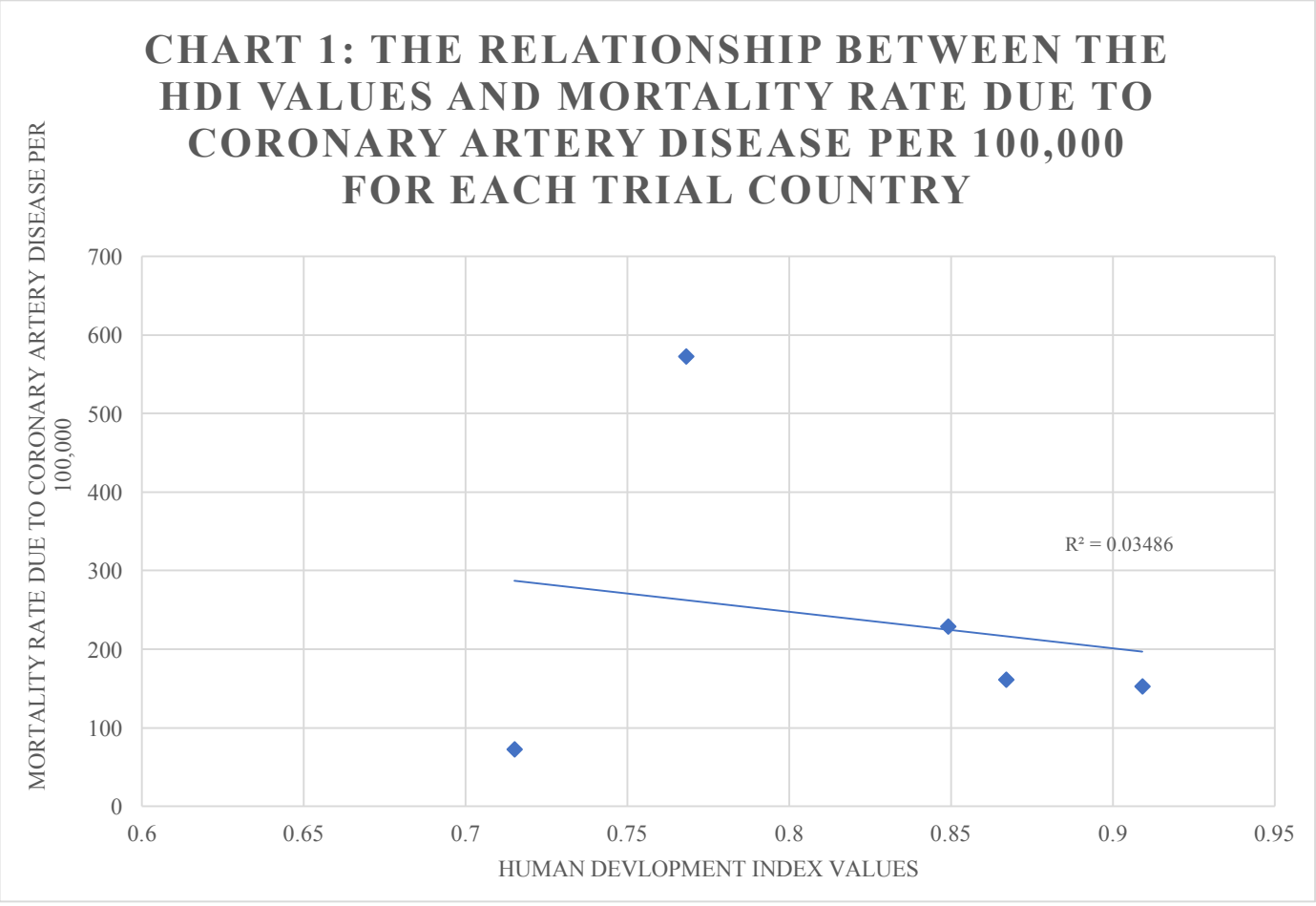
$$Average\ HDI\ value = \frac{HDI\ value\ for\ 2000 + HDI\ value\ for\ 2019}{2}$$

$$= \frac{0.810 + 0.888}{2} = 0.849$$

This sample calculation above was repeated with all values to create Table 2.

Table 2: Processed Data table showing the average mortality rate per 100,000 population and the average HDI		
Country	**Average HDI Value**	**Average CAD Mortality Rate per 100,000**
Greece	0.849	228.66
Belarus	0.768	572.32
United States	0.909	152.63
Thailand	0.715	72.310
Italy	0.867	161.34

After this, Table 2 was plotted on a scatterplot to see the correlation between the average HDI values and the average CAD mortality rate per country. This data was then plotted and created Chart 1.



This scatterplot for the trial countries shows a negative correlation. This correlation indicates that the higher HDI value a country has the lower mortality rate it can have. Chart 1 also indicates a very low R^2 value. The R^2 value is a type of statistical measure that shows the amount of variation between dependent variables. A number below 0.5 means that the data doesn't fit the regression table well and since the R^2 value for the sample trial is 0.03486 it indicates that the data provided doesn't fit the regression table well.

However, this preliminary experiment only contains the data of 5 countries which is not a large enough sample size to determine if the hypothesis is valid. Therefore, the investigation will follow the exact method as outlined above but expand with a larger sample size.

Investigation and Results

For the main investigation, the years 2000 and 2019 will be used as the question outlines. Using the database for the HDI values, I first pulled all countries with an HDI value above 0.75 for the year 2000 and 2019. After this, I removed countries that did not have an HDI value above 0.75 for both years such as Russia, countries that did not have reliable data such as Hong Kong, and countries that did not have a population above 7 million people such as Trinidad and Tobago. Therefore, a total of 23 countries consistently had an HDI above 7.5, have a population size above 7 million since 2000, and have reliable data.

From this list, I followed the steps of the trial investigation and created a raw data table that includes the country names, HDI values, and CAD rates, then a processed data table, and a chart showing the correlation.

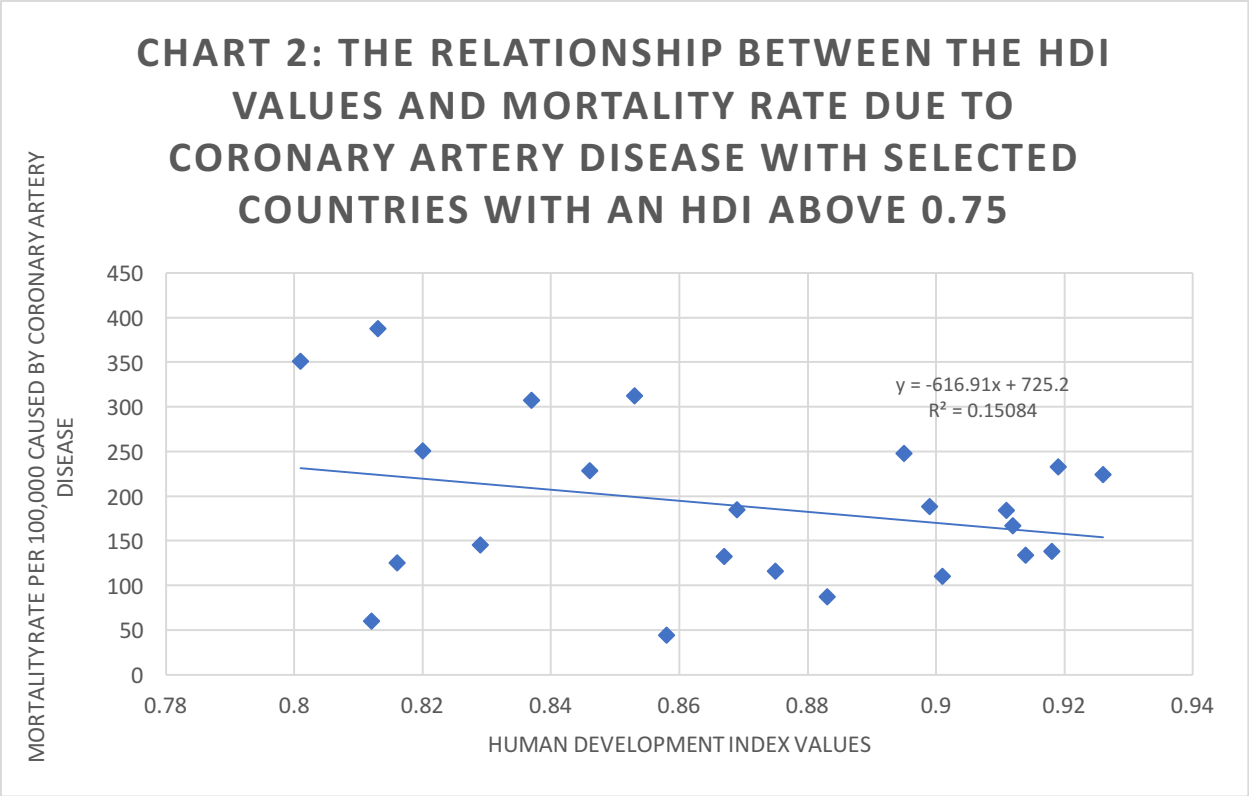
Table 3: The number of HDI Values, CAD Mortality rate per 100,000, and Countries

Country	HDI Values		CAD Mortality Rate per 100,000	
	2000	2019	2000	2019
Switzerland	0.887	0.962	217.04	124.70
Australia	0.896	0.941	151.97	85.510
Sweden	0.904	0.947	296.97	151.36
Germany	0.889	0.948	289.07	176.11
Netherlands	0.893	0.943	173.46	102.93
Belgium	0.887	0.936	200.90	132.87
Canada	0.890	0.937	160.49	107.12
United Kingdom	0.862	0.935	231.84	145.38
Japan	0.877	0.924	91.060	129.06
United States	0.891	0.930	214.42	153.47
Israel	0.844	0.921	114.95	60.130
Austria	0.871	0.919	298.41	196.96
United Arab Emirates	0.796	0.920	45.300	42.62
Spain	0.825	0.908	154.21	110.67

France	0.844	0.905	134.24	97.050
Italy	0.841	0.897	208.63	160.26
Czech Republic	0.808	0.897	328.32	295.92
Greece	0.810	0.889	256.13	201.18
Poland	0.793	0.881	303.85	310.55
Portugal	0.791	0.867	163.90	126.25
Chile	0.763	0.861	63.920	56.24
Hungary	0.773	0.853	395.97	379.63
Argentina	0.779	0.852	140.92	108.85

Table 4: Processed Data table showing the average mortality rate per 100,000 population and the average HDI of Countries from years 2000 and 2019

Country	**Average HDI Value**	**Average CAD Mortality Rate per 100,000**
Switzerland	0.9245	170.87
Australia	0.919	118.74
Sweden	0.926	224.17
Germany	0.919	232.59
Netherlands	0.918	138.20
Belgium	0.912	166.89
Canada	0.914	133.81
United Kingdom	0.899	188.61
Japan	0.901	110.06
United States	0.911	183.95
Israel	0.883	87.540
Austria	0.895	247.69
United Arab Emirates	0.858	43.960
Spain	0.867	132.44
France	0.875	115.65
Italy	0.869	184.45
Czech Republic	0.853	312.12
Greece	0.846	228.66
Poland	0.837	307.20
Portugal	0.829	145.08
Chile	0.812	60.080
Hungary	0.813	387.80
Argentina	0.816	124.89



This scatterplot shows that there is a negative correlation between the mortality rate for coronary artery disease and the selected countries with an HDI of above 0.75. However, the trend line has a weak negative correlation. This is because all the countries selected were very close to each other in their HDI values. This shows that countries with an HDI of above 0.75 all have a similar mortality rate since they are all well developed. This can also be indicated by the R^2 value since it is 0.07 which shows there wasn't much of a difference between the HDI value of each country.

Pearson Correlation Test

After this, a Pearson correlation test was conducted since the R^2 was low. This test is important to determine if there is a linear relationship between the mortality rate of coronary artery disease and the HDI values for the countries. This test also determines the strength of the relationship between the two variables.

Table 5: Testing Pearson's Correlation Coefficient with results from Table 4

Country	N (# of countries)	Average HDI value (Y)	Average CAD rate (X)	HDI *Rate XY	HDI^2 Y^2	$Rate^2$ X^2
Switzerland	1	0.9245	104.22	96.351	0.855	10861.81
Australia	2	0.919	99.94	91.845	0.845	9988.00
Sweden	3	0.926	126.5	117.139	0.857	16002.25
Germany	4	0.919	203.51	187.026	0.845	41416.32

Netherlands	5	0.918	156.02	143.226	0.843	24342.24
Belgium	6	0.912	87.073	79.411	0.832	7581.71
Canada	7	0.914	103.35	94.462	0.835	10681.22
United Kingdom	8	0.899	94.476	84.934	0.808	8925.71
Japan	9	0.901	179	161.279	0.812	32041.00
United States	10	0.911	84.48	76.961	0.830	7136.87
Israel	11	0.883	122.88	108.503	0.780	15099.49
Austria	12	0.895	135.23	121.031	0.801	18287.15
United Arab Emirates	13	0.858	109.89	94.286	0.736	12075.81
Spain	14	0.867	108.37	93.957	0.752	11744.06
France	15	0.875	150.89	132.029	0.766	22767.79
Italy	16	0.869	206.55	179.492	0.755	42662.90
Czech Republic	17	0.853	60.94	51.982	0.728	3713.68
Greece	18	0.846	130.55	110.445	0.716	17043.30
Poland	19	0.837	178.02	149.003	0.701	31691.12
Portugal	20	0.829	106.57	88.347	0.687	11357.16
Chile	21	0.812	112.67	91.488	0.659	12694.53
Hungary	22	0.813	97.86	79.560	0.661	9576.58
Argentina	23	0.816	194.13	158.410	0.666	37686.46
SUM		21.8175	3554.599	3077.708	19.082	601342.57

Then using the SUM values and using the Pearson's correlation coefficient formula the results can be found:

$$r = \frac{n(\sum \text{xy}) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

$$r = \frac{23(3077.708) - (3554.599)(21.8175)}{\sqrt{(23(601342.57) - (3554.599)^2)(23(19.082) - (21.8175)^2)}}$$

$$r = -0.318$$

The r value indicates that there is a reasonably strong negative correlation between the variables Average HDI Value (X) and the Average CAD Rate (Y). An r value that is close to 0.5 can be established as somewhat strong.

Therefore, the correlation coefficient or r_{calc} for the 23 countries is -0.318.

Based off of the r_{calc} value, the critical value or r_{crit} can be found using the degrees of freedom. The calculations for the degrees of freedom are below:

$$DF\ (degrees\ of\ freedom) = N - 2$$
$$= 23 - 2$$
$$= 20$$

Since the degrees of freedom is 20, the r_{crit} value is 0.317.

Since the absolute value of the r_{calc} value is greater than the r_{crit} value the null hypothesis can be rejected, and the alternate hypothesis that states that there is a negative correlation between the HDI values and mortality rate of CAD can be accepted.

Conclusion

As shown above, the initial hypothesis: There will be a negative correlation between countries with an HDI value above 0.75 and the mortality rate of coronary artery disease, was accepted for the question: What is the correlation between the human development index (HDI) and the mortality rate due to coronary artery disease (CAD) per 100,000 people in countries with an HDI above 0.75 for the years 2000 and 2019?

Since coronary artery disease occurs from a plaque buildup in the artery walls, a possible biological explanation for this could be that since HDI includes a countries socioeconomic status that people who live in places with lower HDI's might not be able to afford heart healthy food that don't create plaque such as vegetables, whole grains, and more. Another explanation for this could be that since coronary artery disease is very expensive to treat, people who live in places with a lower HDI might not be able to afford or have proper healthcare.

A study conducted by William M. Schultz, located on the AHA journals and on the national library of medicine created by the National Institute of Health, helps prove this hypothesis as well. In this study, Schultz and his colleagues explored the connection between Socioeconomic status and cardiovascular diseases to conclude that "Socioeconomic status has a significant impact on their cardiovascular health". The two variables used in this study are similar to HDI and CAD since Socioeconomic status is another form of a population's standard of life and cardiovascular health takes coronary artery disease into consideration.

Evaluation and Improvements

After conducting the investigation and accepting the alternative hypothesis, I noticed some strengths and limitations to this investigation. This investigation was strong since only reliable sources and data were used through this experiment. Having reliable data is necessary to understand and create proper conclusions from an experiment. This investigation was also strong since it conducted correlation tests to ensure and specify the type of correlation that occurred between HDI and the mortality rate of CAD per 100,000 people.

However, this investigation did have weaknesses as well. This investigation only pulled two years of data: 2000 and 2019. Having more years, or more spread apart years, could ensure that there is always a negative correlation between HDI and the mortality rate and not just a correlation that has occurred in the past 20 years. Another limitation to this experiment is that only 23 countries were used. If there was more time, I would have used even more countries to further my evidence as to why there is a negative correlation between my variables.

If there was more time for this investigation, I could have investigated the different factors for each country such as GDP per capita, climate change, or different health policies for each country. I hope to conduct a follow up experiment where I would test each countries HDI for that specific year and create charts to represent the correlation per year. This would allow me to discover whether HDI is consistently correlated with CAD every year or if this has been changing over time.

Works Cited

CDC. “Heart Disease Facts.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 7 Feb. 2022, https://www.cdc.gov/heartdisease/facts.htm.

Cleveland Clinic. “Coronary Artery Disease: Causes, Symptoms, Diagnosis & Treatments.” *Cleveland Clinic*, 21 June 2021, https://my.clevelandclinic.org/health/diseases/16898-coronary-artery-disease#:~:text=Coronary%20artery%20disease%20is%20the,discomfort%20and%20shortness%20of%20breath.

Kopecky , Stephen. “Coronary Artery Disease.” *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 5 June 2020, https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613.

Rainger, Ed. “Do Platelets Recruit White Blood Cells into Fatty Plaques in Atherosclerosis?” *Welcome to the British Heart Foundation*, 1 Sept. 2017, https://www.bhf.org.uk/research-projects/do-platelets-exacerbate-the-atherogenic-process-by-regulating-the-recruitment-differentiation-and-inflammatory-function-of-monocytes#:~:text=As%20part%20of%20plaque%20formation,blood%2C%20which%20form%20blood%20clots.

RTI International. “Cardiovascular Disease Costs Will Exceed $1 Trillion by 2035.” *ScienceDaily*, ScienceDaily, 14 Feb. 2017, https://www.sciencedaily.com/releases/2017/02/170214162750.htm.

Team, The Investopedia. “What Is the Human Development Index (HDI)?” *Investopedia*, Investopedia, 25 Feb. 2022, https://www.investopedia.com/terms/h/human-development-index-hdi.asp.

United Nations Development Programme. “Human Development Reports.” | *Human Development Reports*, 2000, https://hdr.undp.org/en/countries.

World Health Organization. “Cardiovascular Diseases (Cvds).” *World Health Organization*, World Health Organization, 11 June 2021, https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

World Health Organization. “Cardiovascular Diseases.” *World Health Organization*, World Health Organization, 10 Mar. 2022, https://www.euro.who.int/en/health-topics/noncommunicable-diseases/cardiovascular-diseases/cardiovascular-diseases2.

World Health Organization. “Global Health Estimates: Leading Causes of Death.” *World Health Organization*, World Health Organization, 2000, https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death.