Health & Wealth: An interactive study on obesity and cardiovascular disease
Long Reflection

Alyssa Wu Zhang
UT EID: awz227
April 23, 2024

Historical Background and Interest

The creative project *Health & Wealth: An interactive study on obesity and cardiovascular disease* centers on an aesthetically-pleasing interactive data visualization dashboard about health, wealth, obesity, and cardiovascular disease around the world. Existing research indicate that obesity impacts higher-income countries more than lower-income countries (Talukdar et al., 2020), while cardiovascular disease impacts lower-income countries more than higher-income countries (National Academies Press, 2010). Furthermore, the scientific consensus is that obesity is a cause of cardiovascular disease (Powell-Wiley et al., 2021).

My project draws inspiration from statistian Hans Rosling's "200 Countries, 200 Years, 4 Minutes" presentation, which used an engaging animated bubble chart to illustrate the global trend towards improved health and wealth over time. This global progression, despite persistent health challenges such as obesity and cardiovascular disease, sparked my interest in crafting a compelling data visualization narrative centered around these specific health issues.

Thesis Statement

The goal of *Health & Wealth* is to investigate the statistical connections between obesity prevalence, cardiovascular disease prevalence, gross domestic product (GDP) per capita, and life expectancy across many countries. By employing statistical analysis and interactive visualization, the project creates a compelling narrative on the relationships between obesity, cardiovascular disease, wealth, and health. Understanding these relationships is important for informing policies aimed at decreasing obesity and cardiovascular disease prevalence across various income groups.

Data Sources

The data sources for *Health & Wealth* are Our World in Data (OWiD), the Global Burden of Disease (GBD) study conducted by the Institute for Health Metrics and Evaluation (IHME), the Global Health Observatory (GHO) maintained by the World Health Organization (WHO), and the World Bank (WB). These sources provide four distinct datasets: the estimated prevalence of obesity (in percent) among adults spanning from 1975 to 2016, the number of cases of cardiovascular disease per 100,000 people (in percent) spanning from 1990 to 2019, the GDP per capita (in current US dollars) spanning from 1960 to 2022, and the life expectancy at birth (in years) spanning from 1960 to 2021 across numerous countries. A merged version of the four datasets, carefully cleaned and wrangled with the NumPy and Pandas Python libraries, is the

dataset used for *Health & Wealth*. This dataset contains 4912 rows and 8 columns. The columns are year (spanning from 1990 to 2016), country (with 185 unique entries), country codes, continent, obesity prevalence, cardiovascular disease prevalence, GDP per capita, and life expectancy. Data for this project is reliable because OWiD has been cited in peer-reviewed journals, the IHME's GBD study is peer-reviewed, the WHO's GHO is a sourced by government agencies and universities for health statistics, and the WB is a leading international financial institution with a transparent data collection process.

Data Analysis

Data analysis for *Health & Wealth* involves calculating Pearson correlation coefficients among quantitative variables and identifying the strongest correlations. Table 1 presents the correlation matrix, revealing high correlations (above 0.5) between life expectancy and both obesity prevalence (0.569286) and GDP per capita (0.573077). The relationship between life expectancy and obesity prevalence shows a linear best fit, with similar R-squared values for linear, quadratic, and cubic fits. By contrast, the relationship between life expectancy and GDP per capita demonstrates a cubic best fit, which has the highest R-squared value. Refer to Graphs 1 and 2, made with the Seaborn Python library, for visual representations of these relationships.
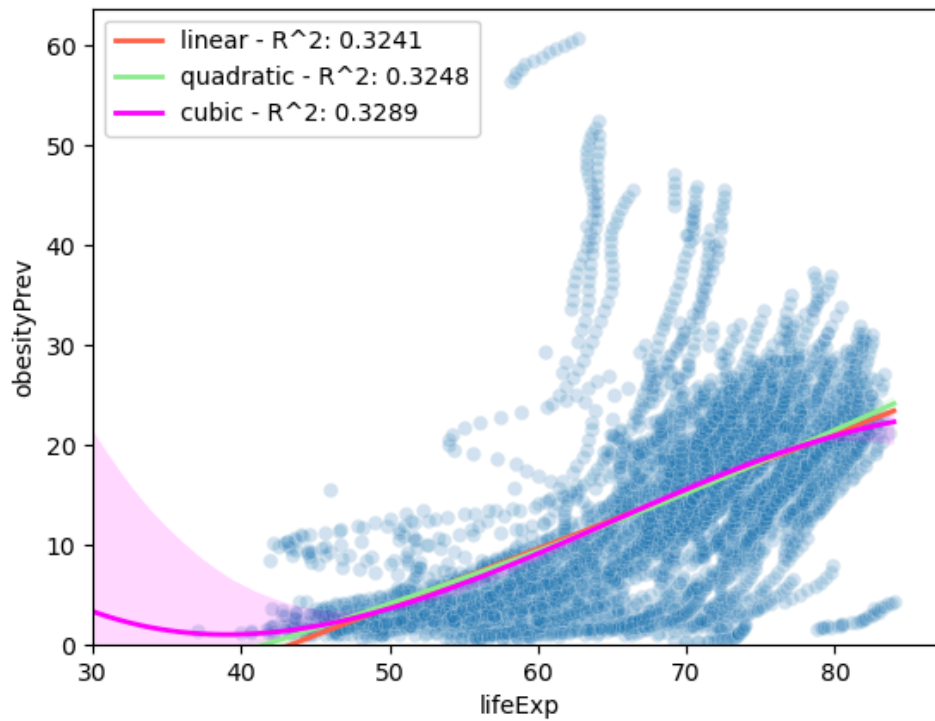
Table 1

Correlation Matrix

|	year	obesityPrev	cardioPrev	lifeExp	gdpPcap
year	1	0.309993	-0.04967	0.228291	0.212192
obesityPrev	0.309993	1	0.198029	0.569286	0.365642
cardioPrev	-0.04967	0.198029	1	-0.00363	-0.06066
lifeExp	0.228291	0.569286	-0.00363	1	0.573077
gdpPcap	0.212192	0.365642	-0.06066	0.573077	1

Graph 1

Obesity Prevalence vs. Life Expectancy Regression Plot

Graph 2

GDP per Capita vs. Life Expectancy Regression Plot
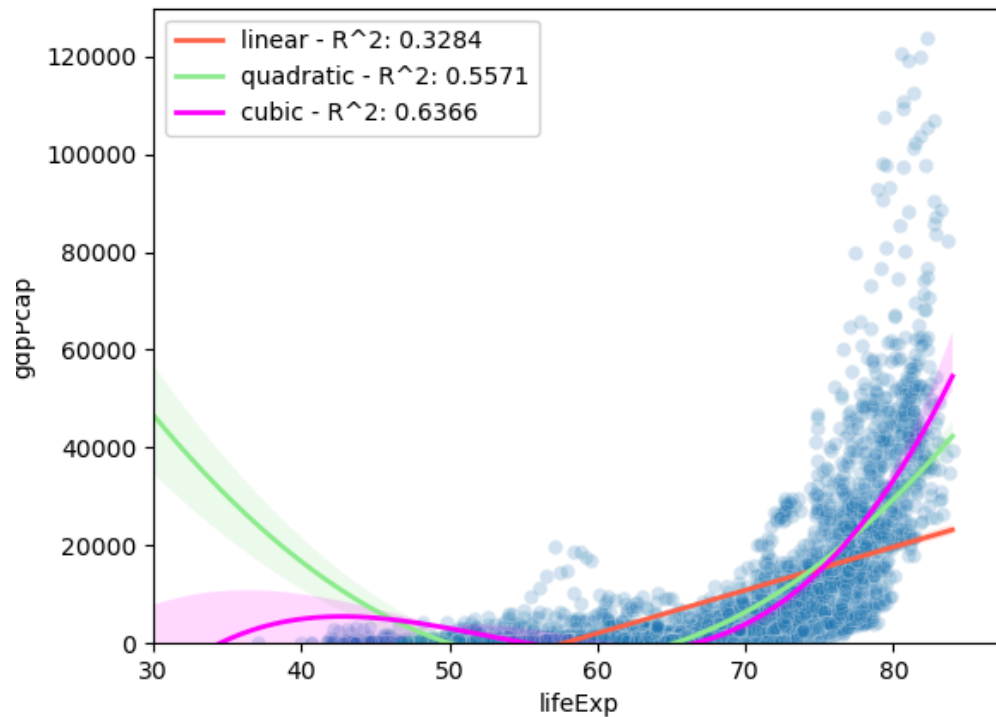
Data Visualizations

The focus of *Health & Wealth* is a dashboard with six interactive data visualizations crafted with the Plotly and Dash Python libraries. The visualizations intend to support five arguments, which are as follows:

I.   Countries have become healthier and wealthier over time.

II. Obesity prevalence has increased over time, while cardiovascular disease prevalence has remained largely stagnant.

III. Obesity and cardiovascular disease prevalence are positively correlated.

IV. There are global disparities in obesity and cardiovascular disease prevalence.

V.   Obesity impacts wealthier countries more, while cardiovascular disease does not.
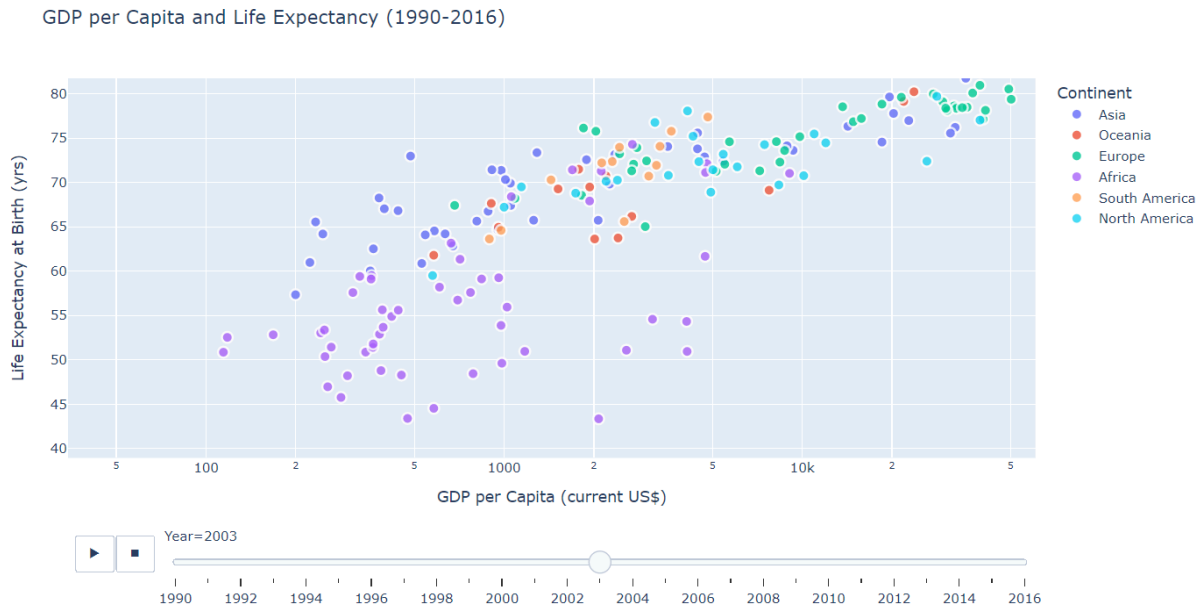
Argument I

The dashboard begins with an animated scatterplot showing the global state of GDP per capita and Life Expectancy from 1990 to 2016 (see Snapshots 1A to 1C). GDP per capita in current US dollars is on the x-axis and life expectancy at birth in years is on the y-axis. The bubbles represent 185 countries colored by continent. The countries in the bottom left are poor and sick, while the countries in the top right are rich and healthy. Interacting with the time slider shows a worldwide trend where, presumably due to advancements in healthcare, technological progress, and economic development, people have become healthier and wealthier over time.
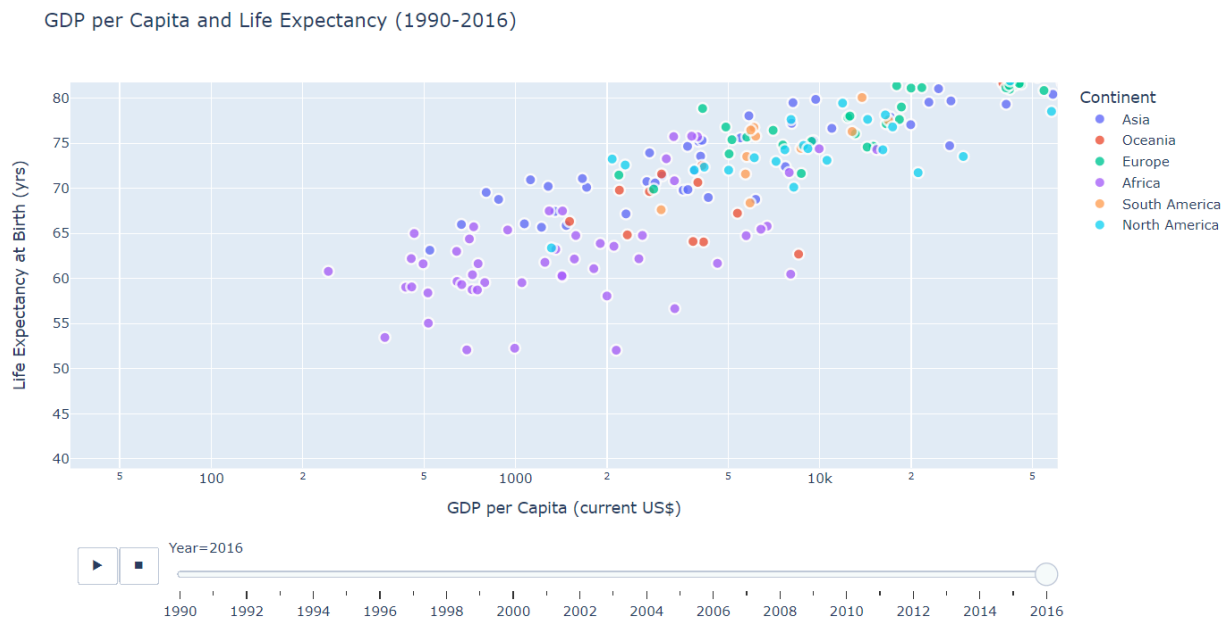
Snapshot 1A

Snapshot 1B

GDP per Capita and Life Expectancy (1990-2016)

Year=2003

Snapshot 1C

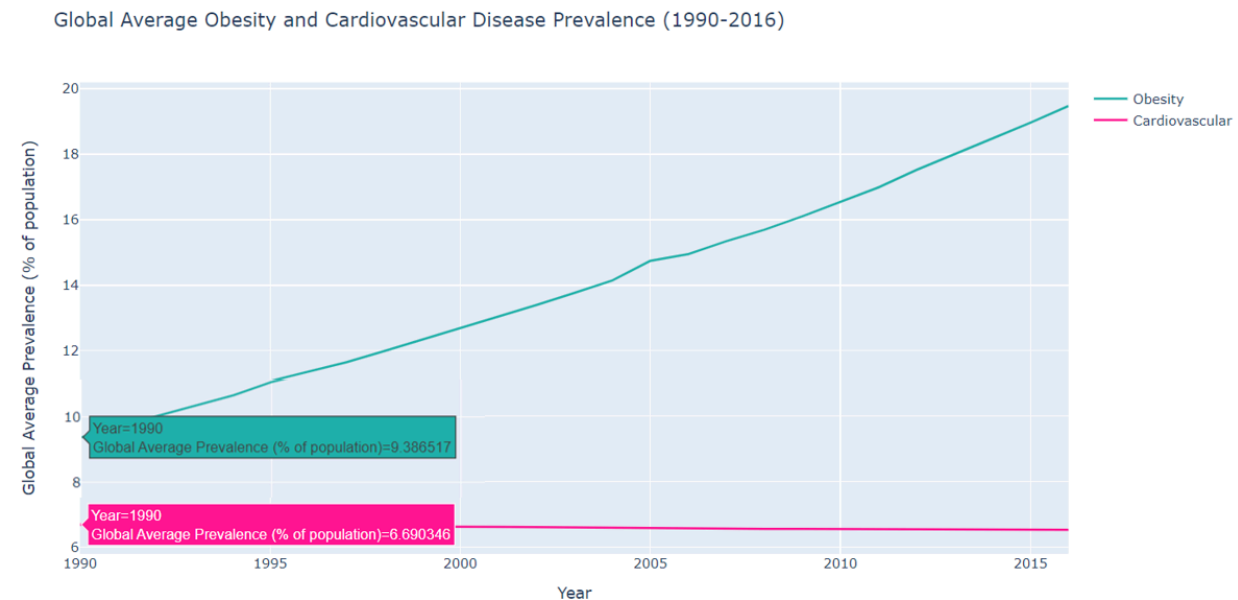### GDP per Capita and Life Expectancy (1990-2016)
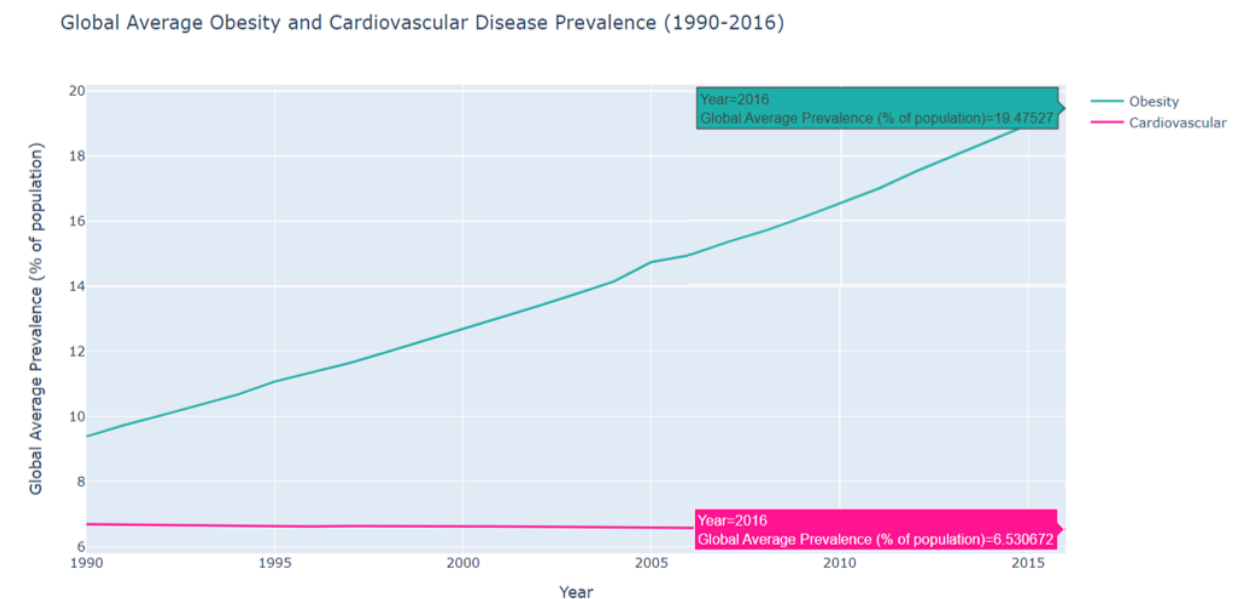

Year=2016
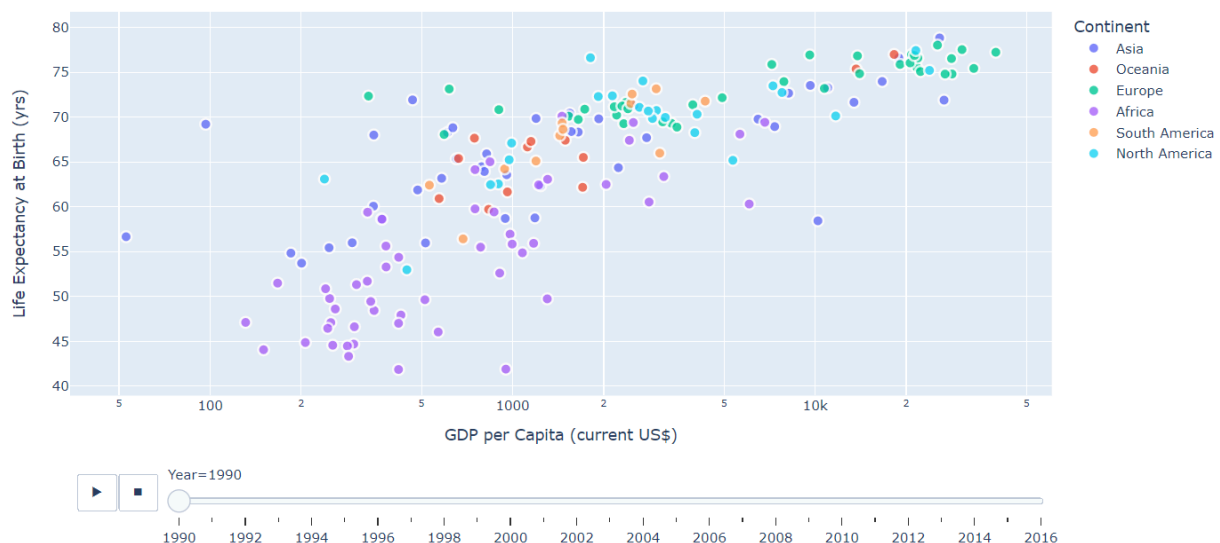
Argument II

Despite significant progress in healthcare and economic development, obesity and cardiovascular disease persist as formidable challenges. The dashboard shows this with an interactive line chart depicting the global average prevalence of obesity and cardiovascular disease from 1990 to 2016 (see Snapshots 2A and 2B). The x-axis is year and the y-axis is the percentage of population with either obesity or cardiovascular disease averaged across the 185

countries. The green line represents the percentage for obesity, and the pink line represents the percentage for cardiovascular disease. The green line shows a linear increase from 9.4% in 1990 to 19.5% in 2016, indicating that obesity has become more of a global issue in recent years. The pink line stagnates around 6.5 to 6.6% throughout the years, indicating that although cardiovascular disease has not become more of a problem over the years, it has not been significantly addressed.
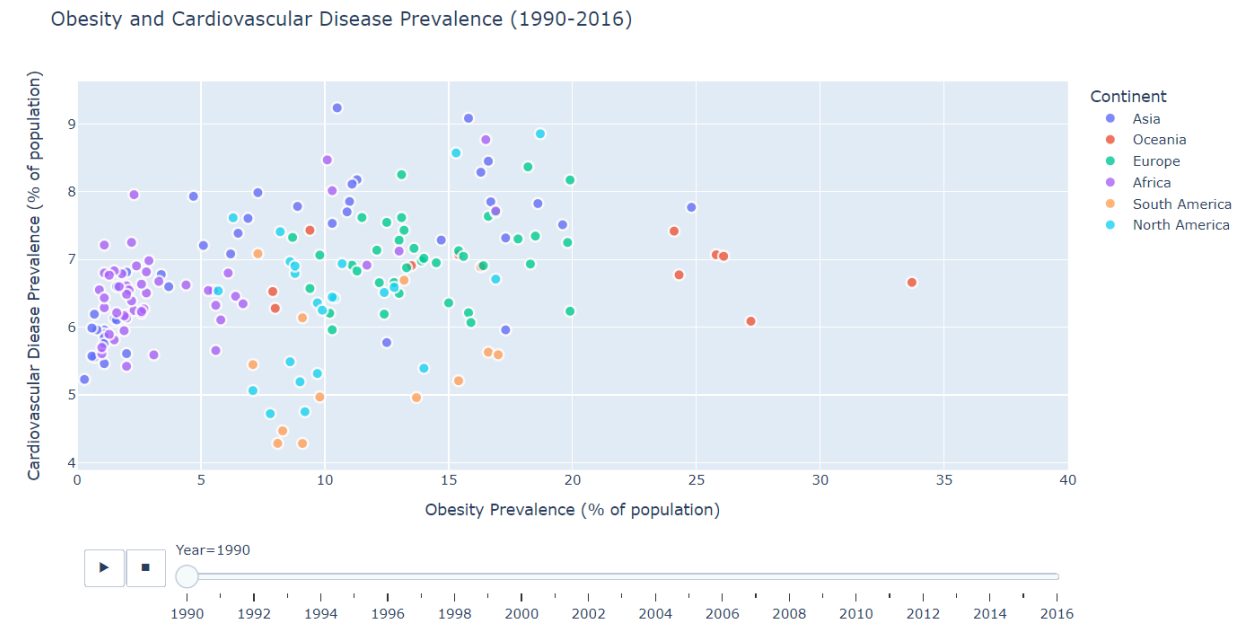
Snapshot 2A

Global Average Obesity and Cardiovascular Disease Prevalence (1990-2016)

Year=1990
Global Average Prevalence (% of population)=9.386517

Year=1990
Global Average Prevalence (% of population)=6.690346

— Obesity
— Cardiovascular

Snapshot 2B

Global Average Obesity and Cardiovascular Disease Prevalence (1990-2016)

Year=2016
Global Average Prevalence (% of population)=19.47527

Year=2016
Global Average Prevalence (% of population)=6.530672

— Obesity
— Cardiovascular

Argument III

Next, consider the relationship between obesity and cardiovascular disease with another animated scatterplot, this one showing obesity and cardiovascular disease prevalence from 1990 to 2016 (see Snapshots 3A to 3C). The percentage of population with obesity is on the x-axis, while the percentage of population with cardiovascular disease is on the y-axis. The bubbles once again represent 185 countries colored by continent. The countries in the bottom left are fitter with healthy hearts, while countries in the bottom right are fatter with unhealthy hearts. Interacting with the time slider shows that, as obesity prevalence increases, cardiovascular disease prevalence either increases or remain the same. The data suggests a weak positive correlation between obesity and cardiovascular disease. However, the scientific consensus by the American Heart Association is that obesity leads to cardiovascular disease (Powell-Wiley et al., 2021). Despite what the data visualization suggests, there is a causal relationship between obesity and cardiovascular disease.

Snapshot 3A

Snapshot 3B

Obesity and Cardiovascular Disease Prevalence (1990-2016)

Snapshot 3C

Obesity and Cardiovascular Disease Prevalence (1990-2016)


Argument IV.

Dive into the geographical distribution of obesity and cardiovascular disease from 1990 to 2016 with animated choropleth maps. For the map of obesity prevalence (see Snapshots 4-1A to 4-1C), the lightest color represents 0% obese while the darkest color represents 40% obese. Interacting with the time slider shows that nearly all countries have become more obese

overtime, especially in the Americas, the Middle East, Oceania, and Europe. The map of cardiovascular prevalence (see Snapshots 4-2A to 4-2C) has a smaller range, with the lightest color representing 4% of population having cardiovascular disease and the darkest color representing 9%. Interacting with the time slider reveals that, other than a slight lessening of cardiovascular disease in North America, there has not been considerable change in cardiovascular disease from 1990 to 2016. Generally, countries with higher prevalence of cardiovascular disease are in North America, the Middle East, and Southern/Eastern Europe.
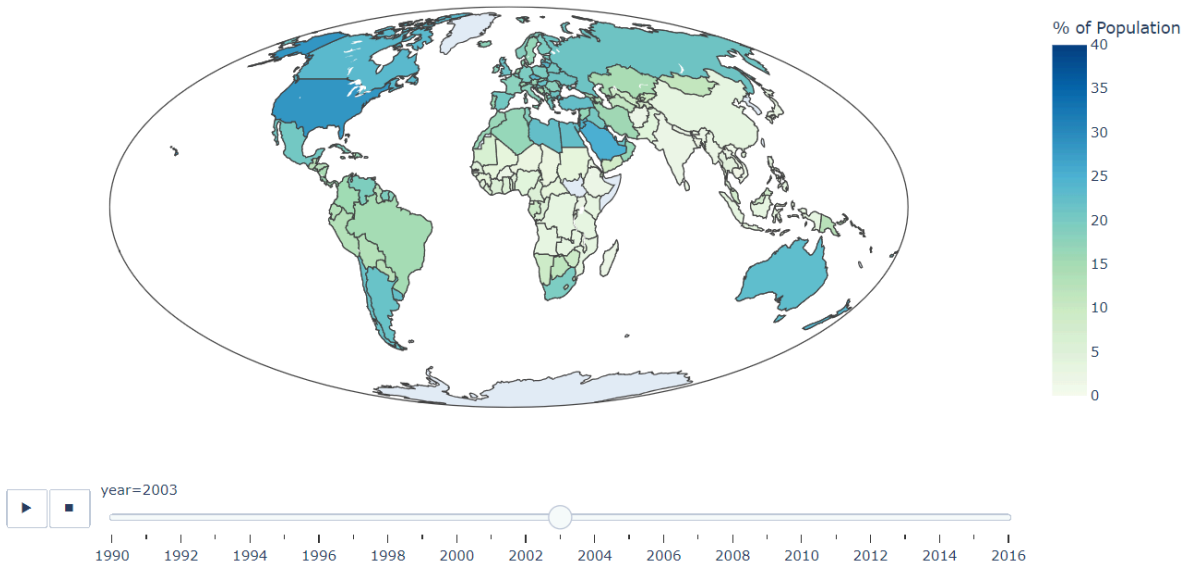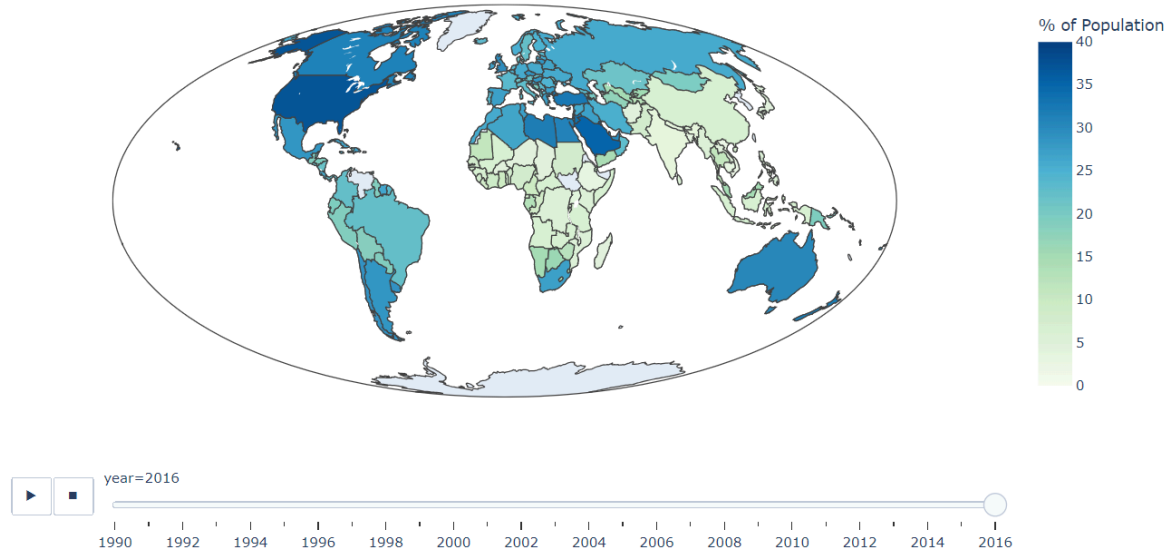
Snapshot 4-1A

Obesity Prevalence (1990-2016)
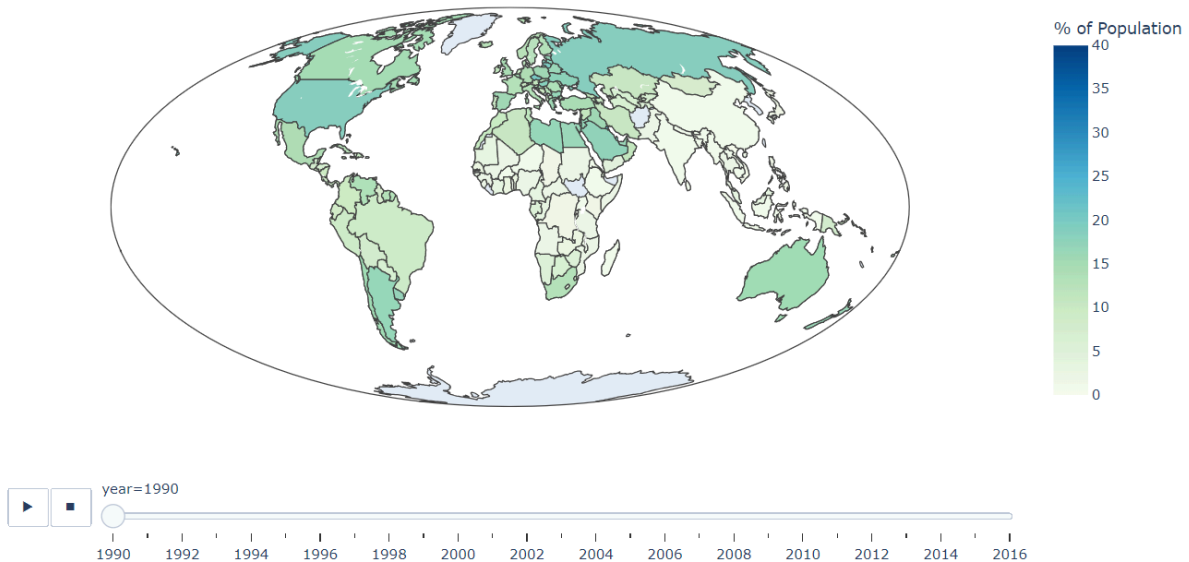
Snapshot 4-1B
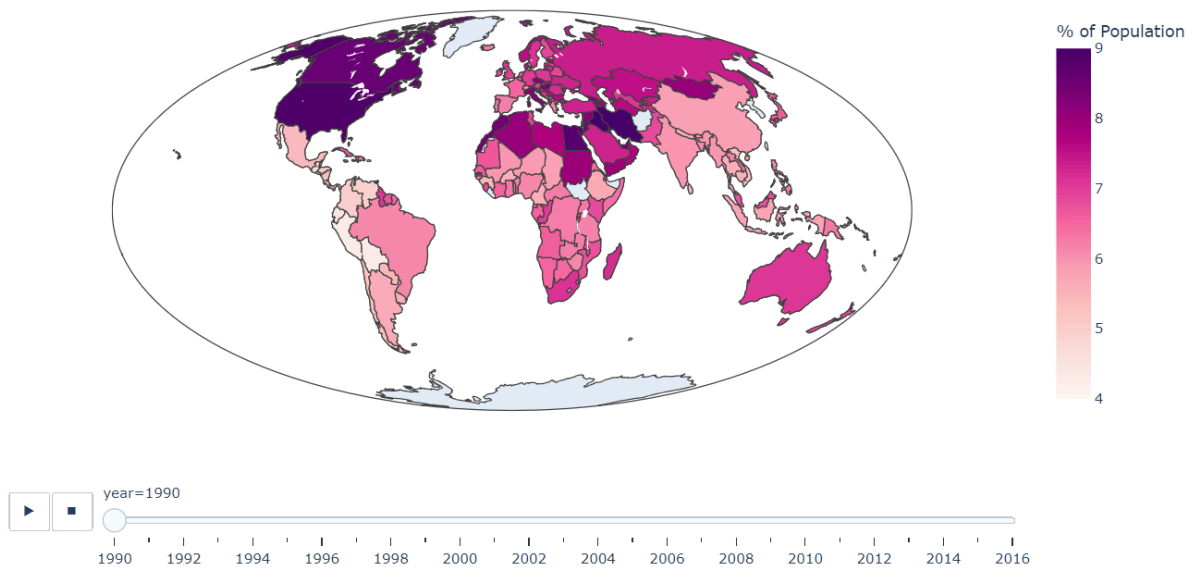
Obesity Prevalence (1990-2016)
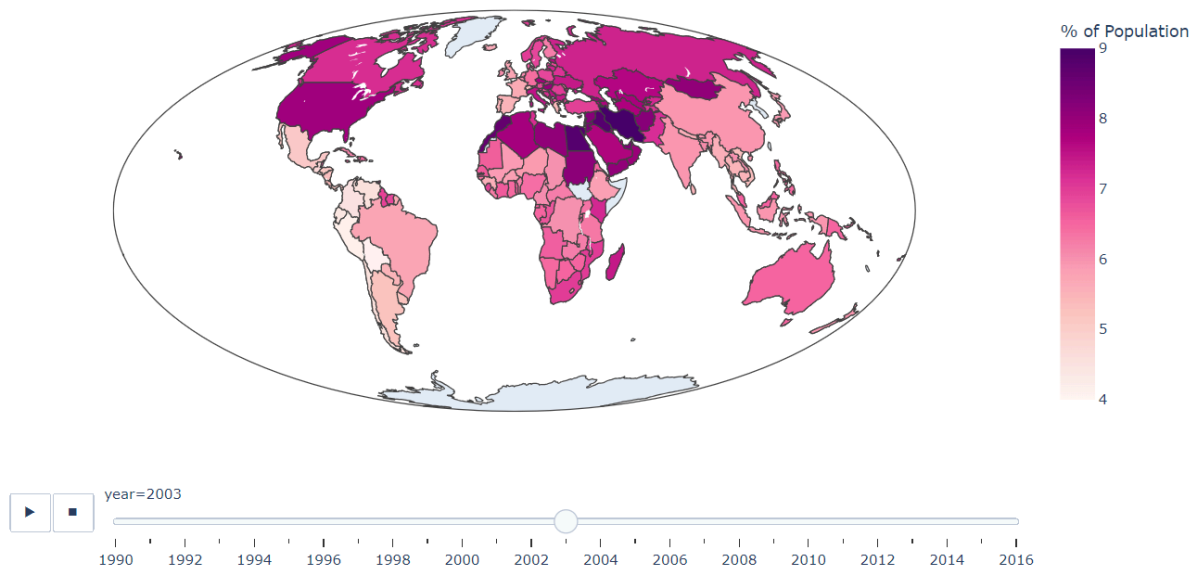
Snapshot 4-1C

Obesity Prevalence (1990-2016)

year=2016

1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016

Snapshot 4-2A

Cardiovascular Disease Prevalence (1990-2016)

year=1990

1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016
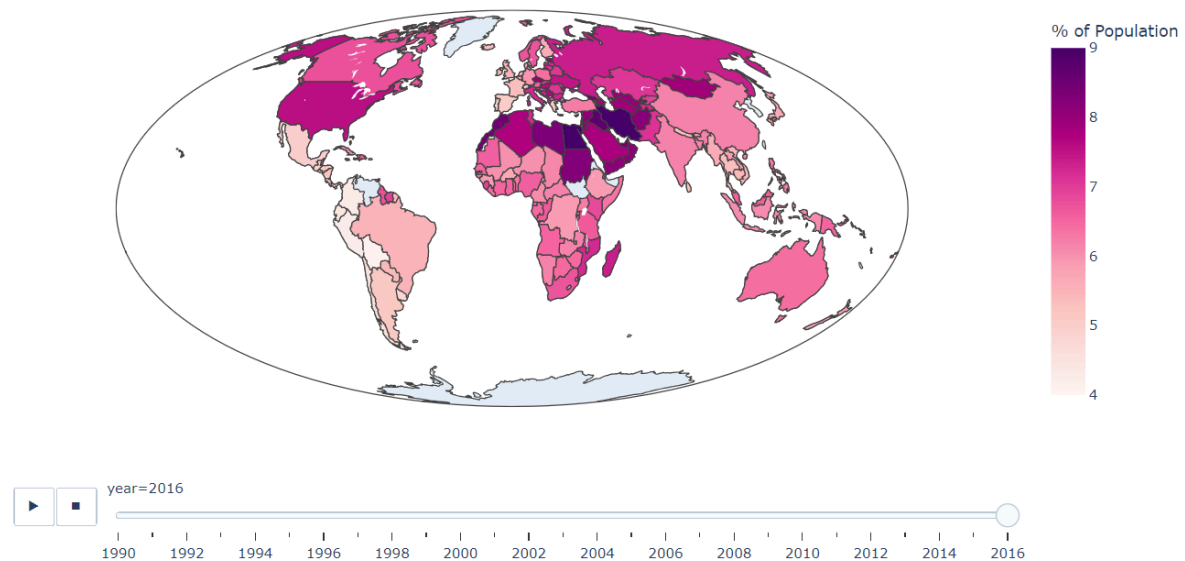
Snapshot 4-2B

Cardiovascular Disease Prevalence (1990-2016)

Snapshot 4-2C
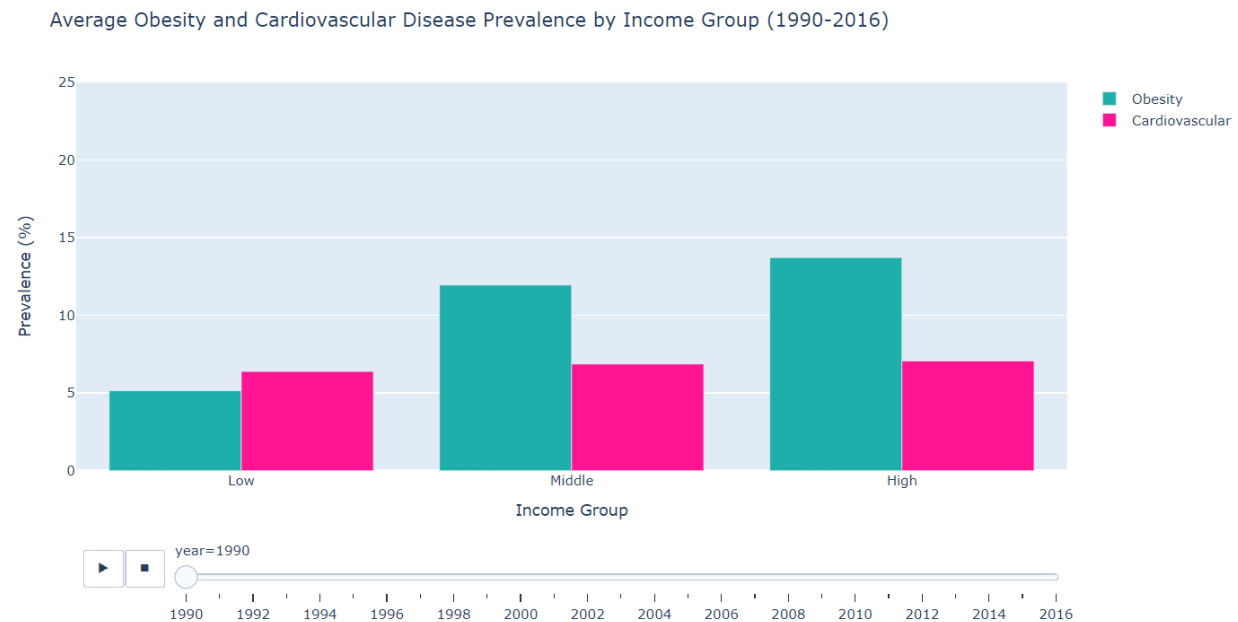
Cardiovascular Disease Prevalence (1990-2016)
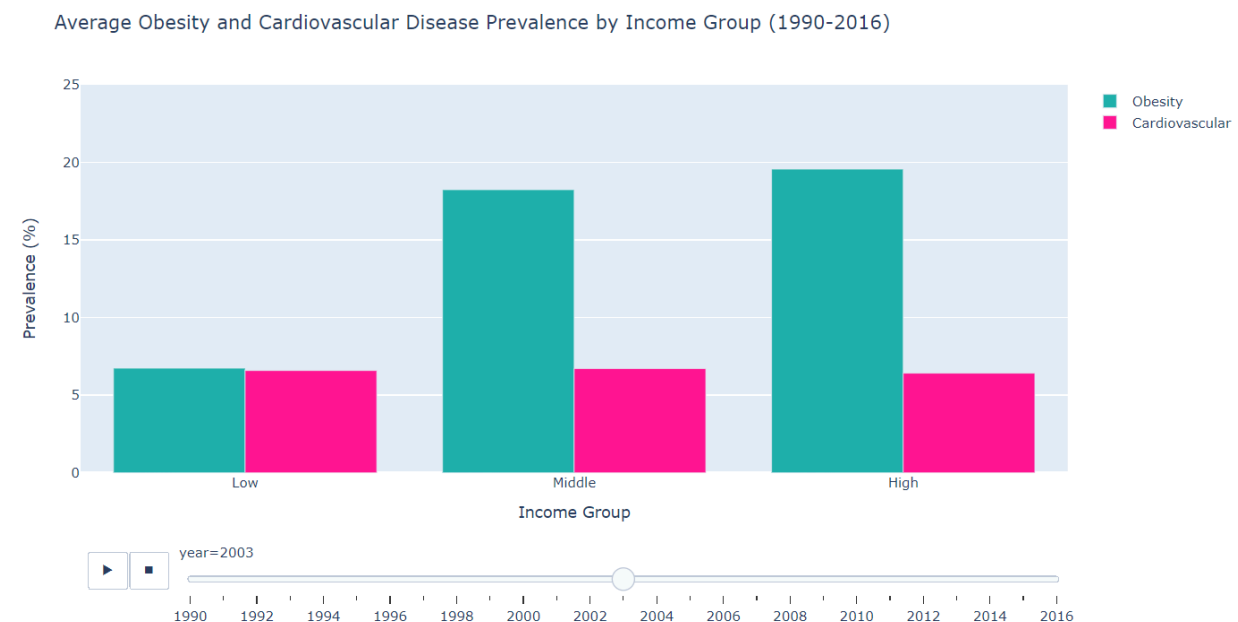

Argument V.

Uncover the role of wealth in shaping the burden of obesity and cardiovascular disease. Snapshots 5A to 5C show an animated side-by-side barchart comparing the average obesity and cardiovascular disease prevalence of low, middle, and high income countries from 1990 to 2016.

Countries were grouped by income based on GDP per capita z score. The y-axis represents the average percentage of the population with either obesity or cardiovascular disease for each income group. Green represents obesity, while pink cardiovascular disease. Interacting with the time slider shows that obesity impact healthier countries more, while cardiovascular disease does not. Although the data visualization implies that cardiovascular disease prevalence is unrelated to wealth, a National Academies Press publication indicates that cardiovascular disease impacts poorer countries more (2010).

Snapshot 5A

Average Obesity and Cardiovascular Disease Prevalence by Income Group (1990-2016)

year=1990

Snapshot 5B

Average Obesity and Cardiovascular Disease Prevalence by Income Group (1990-2016)

year=2003

Snapshot 5C

Average Obesity and Cardiovascular Disease Prevalence by Income Group (1990-2016)

Conclusion

   Health & Wealth was a journey of self-learning and skill development. Completing LinkedIn learning certificate and auditing Coursera courses for *Health & Wealth* equipped me with valuable self-learning skills, allowing me to perform data analysis with the Pandas, NumPy, Seaborn, and Plotly Python packages. *Health & Wealth* gave me an appreciation for data visualization, which I find to be an artistic and powerful tool for conveying complex information. Furthermore, the project applied statistical knowledge from my mathematics major to extract meaningful insights from data.

   If I had more time, I would have enhanced the interactivity of my *Health & Wealth* data visualization dashboard by incorporating linked brushing. This would have made the connections between different visualizations more obvious. Additionally, I would have made the dashboard publicly accessible so that a wider audience would be able to interact with my work. Lastly, I would have investigated the few discrepancies between my data visualizations and existing research conclusions. Overall, *Health & Wealth* has been a rewarding project filled with growth and exploration.

Acknowledgements

Many thanks to my faculty mentor, Computer Science Professor Shyamal Mitra, for meeting with me on a weekly basis for academic discussion and guidance. I am grateful to the UT Austin Digital Arts & Media Bridging Disciplines Program, through which I completed this project.

Literature Cited

Institute for Health Metrics and Evaluation, Global Burden of Disease Study (2019) via Our
World in Data. *Prevalence rate of cardiovascular diseases* [Data set].
https://ourworldindata.org/grapher/prevalence-rate-of-cardiovascular-disease

Powell-Wiley, T. M., Poirier, P., Burke, L. E., Després, J.-P., Gordon-Larsen, P., Lavie, C. J.,
Lear, S. A., Ndumele, C. E., Neeland, I. J., Sanders, P., St-Onge, M.-P. (2021). Obesity
and cardiovascular disease: A scientific statement from the american heart association.
Circulation, 143(21). https://doi.org/10.1161/CIR.0000000000000973

*Promoting cardiovascular health in the developing world: A critical challenge to achieve global
health.* (2010). National Academies Press. https://doi.org/10.17226/12815

World Bank. *GDP per capita (current US $)* [Data set].
https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

World Bank. *Life expectancy at birth, total (years)* [Data set].
https://data.worldbank.org/indicator/SP.DYN.LE00.IN

World Health Organization, Global Health Observatory (2022) via Our World in Data. *Obesity in
adults* [Data set]. https://ourworldindata.org/grapher/share-of-adults-defined-as-obese

Appendix

The computer code written for *Health & Wealth* is available in the following GitHub repository: https://github.com/alyssawzhang/health-wealth-interactive