Analyse Health and Demographic Data to Identify Common Traits Leading to Heart Disease

Capstone Project

Prepared By	Jyoti Prakash Das
Program	Business Analytics
Mentor-Led Internship	**upGrad**
Industry Partner	Practo
Date of Submission	January 20, 2026

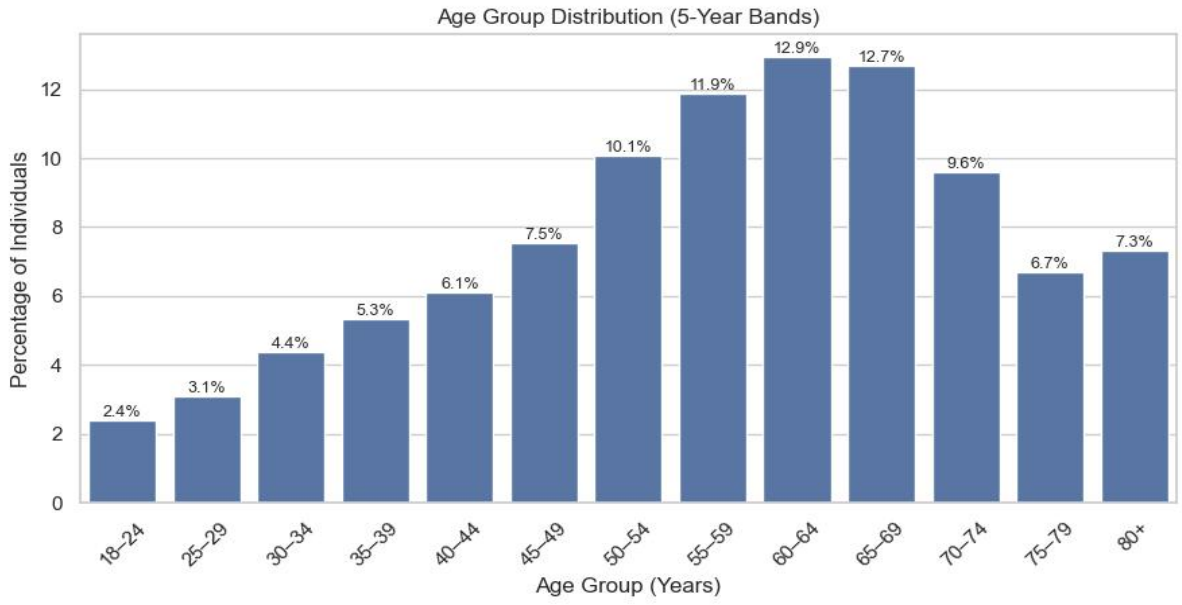
Heart Disease Insights & Executive Summary

Heart disease prevalence rises sharply with **age** and is higher in **men** than women. Older adults (especially 55+) and males show far greater heart disease rates. Lifestyle and health status are critical: individuals reporting *poor overall health* are far more likely to have heart disease than those in excellent health. Physical limitations also strongly predict disease – for example, about 23% of people with serious walking or climbing difficulties have heart disease versus only ~7% of those without such limitations. In contrast, cardiovascular risk is lower among those with higher socioeconomic status: people with higher education or income have **lower** heart disease rates. Clinical risk factors like high blood pressure, high cholesterol, and **obesity** further elevate risk (general health had the strongest correlation of 0.25 in our analysis). Habitual smoking showed a modest increase in heart disease, whereas regular exercise and moderate alcohol intake were somewhat protective (heavy drinkers had only slightly lower disease rates, a finding likely confounded by other factors). Overall, heart disease prevalence is most strongly associated with older age, male sex, poor health status, mobility limitations, hypertension, high cholesterol, and lower socioeconomic status.

Key Risk Factors and Their Impact

Demographics (Age & Gender)

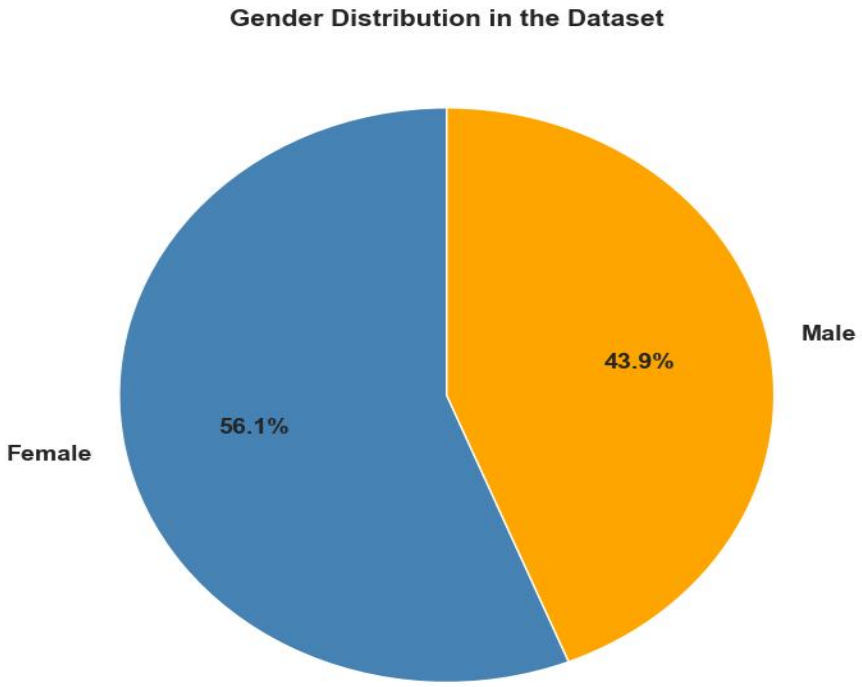
The BRFSS 2015 sample is skewed toward middle-aged and older adults. The largest age groups are 60–64 (12.9%) and 65–69 (12.7%), while younger adults (18–24) are underrepresented.



Cardiovascular risk rises with age, so the focus on older adults makes this dataset useful for studying heart disease (but less generalizable to youth). Men and women are both represented, though females (56.1%) slightly outnumber males (43.9%). Since heart disease risk differs by sex and age, subsequent analyses and interventions should account for this composition.

- **Age Trend:** *Prevalence of heart disease climbs sharply with advancing age, reflecting the strong effect of aging on cardiovascular risk.*

- **Gender:** The female majority in the sample should be noted, as biological and behavioral differences mean men and women have different heart disease patterns.



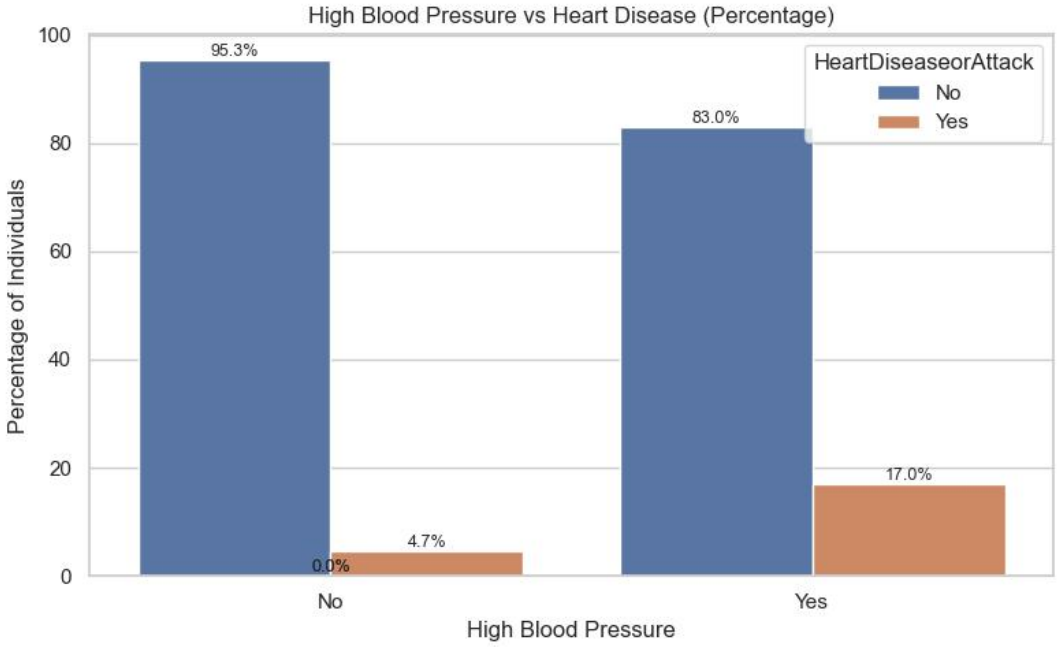
- **Implications:** Screening and prevention programs may focus on older populations, and analyses must control for age and sex to avoid confounding.

Clinical Conditions

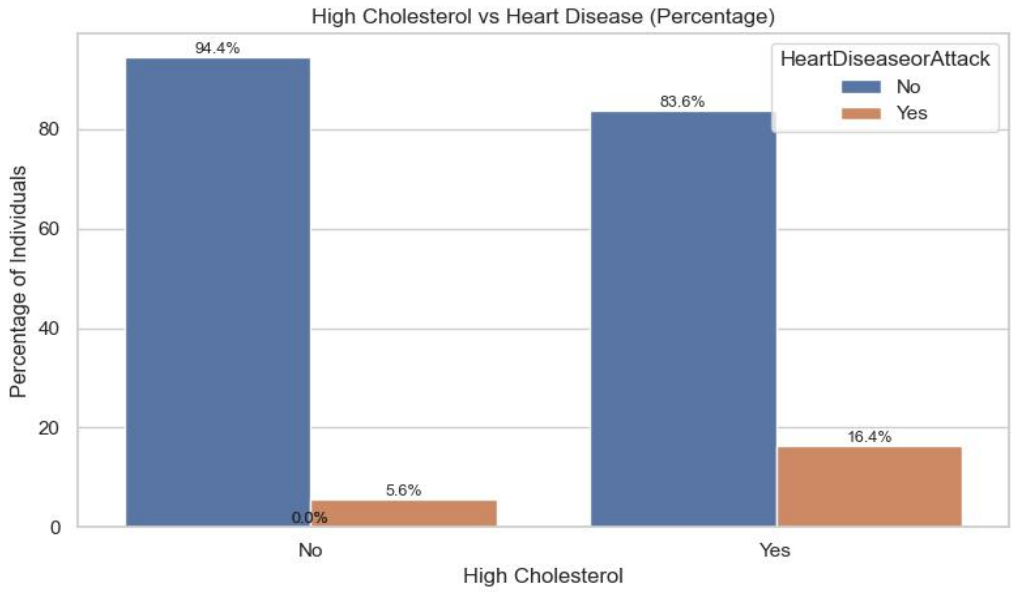
Several health conditions strongly co-occur with heart disease in the data:

- **Hypertension (High Blood Pressure):** Individuals with high blood pressure have a *substantially* higher rate of heart disease than those without hypertension. Over 43% of the sample report hypertension, underscoring its prevalence. Hypertension's

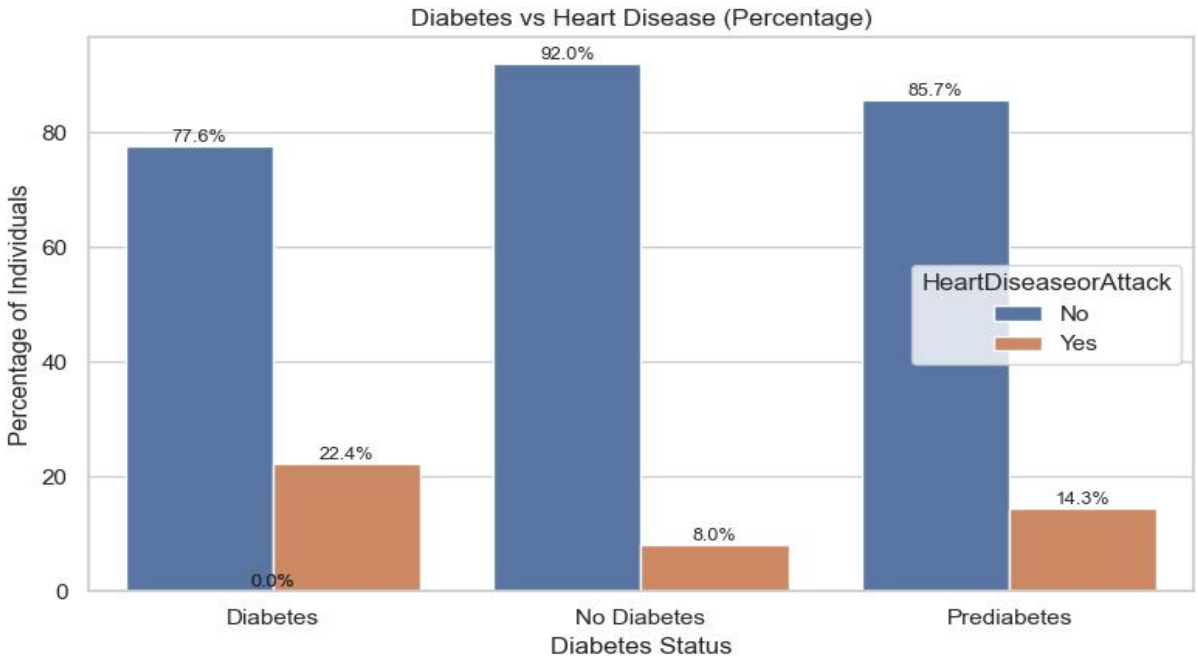
association with heart disease is well-established and is clearly observed in the EDA.



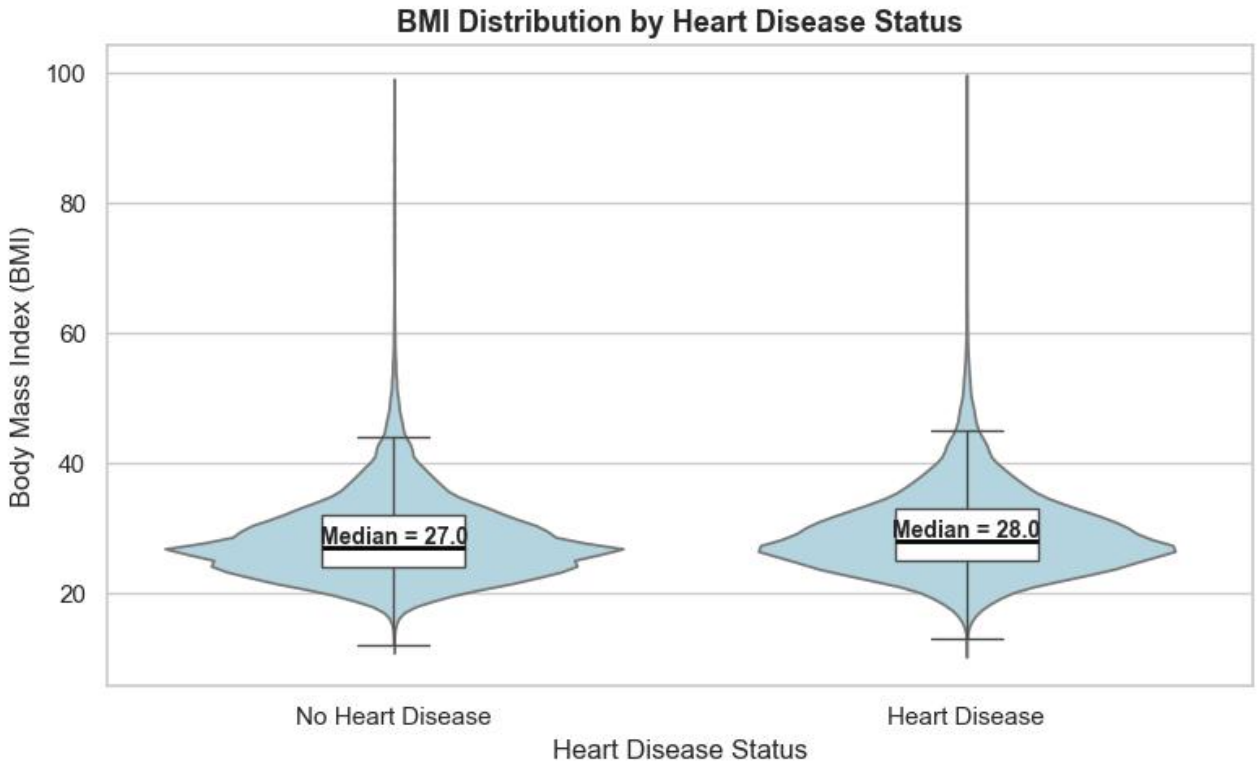
- **High Cholesterol**: Similarly, those reporting high cholesterol show a higher heart disease prevalence. About 42% of respondents have high cholesterol. This supports known cardiovascular risk factors and suggests aggressive lipid management is important.



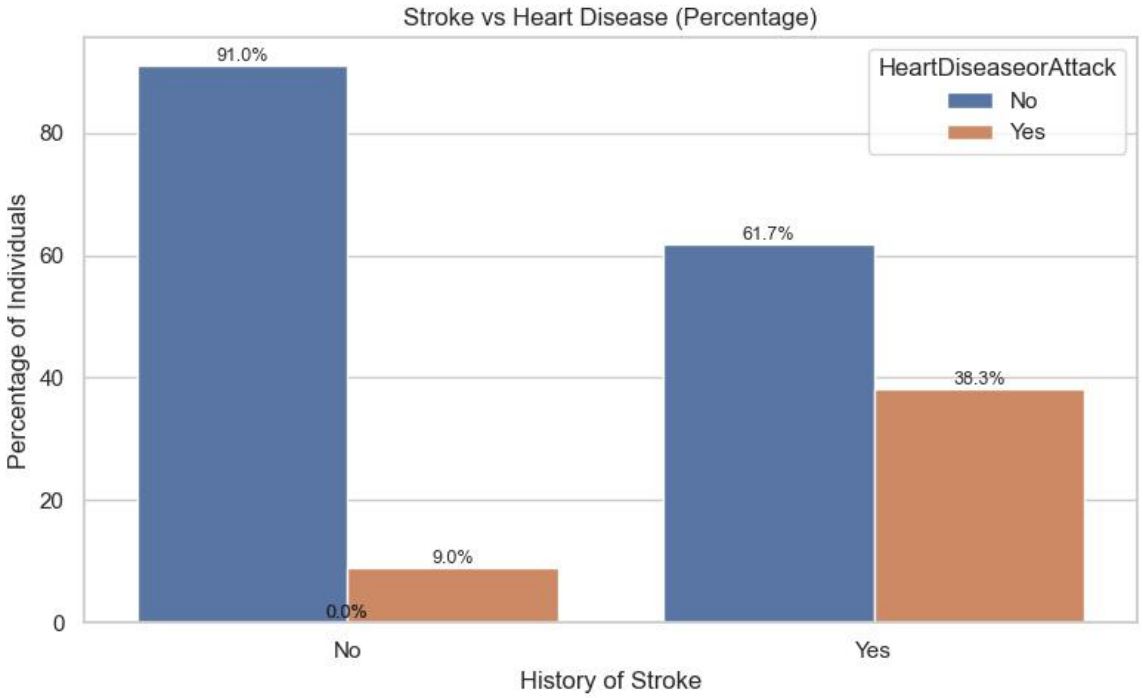
- **Diabetes:** Heart disease cases rise sharply with diabetes status. Respondents with diabetes (and even prediabetes) have much higher heart disease rates than those without diabetes. In other words, poor blood sugar control strongly predicts cardiovascular risk.



- **Obesity (BMI):** Heart disease patients tend to have higher BMI. The distribution of BMI is shifted upward in the heart disease group, indicating overweight/obesity are more common in those with cardiovascular issues. This aligns with BMI's links to hypertension and diabetes.



- ***History of Stroke:*** *People who have had a stroke show a dramatically higher prevalence of heart disease compared to those without a stroke history. Stroke and heart disease share many risk factors, so stroke history is a red flag for cardiovascular risk.*

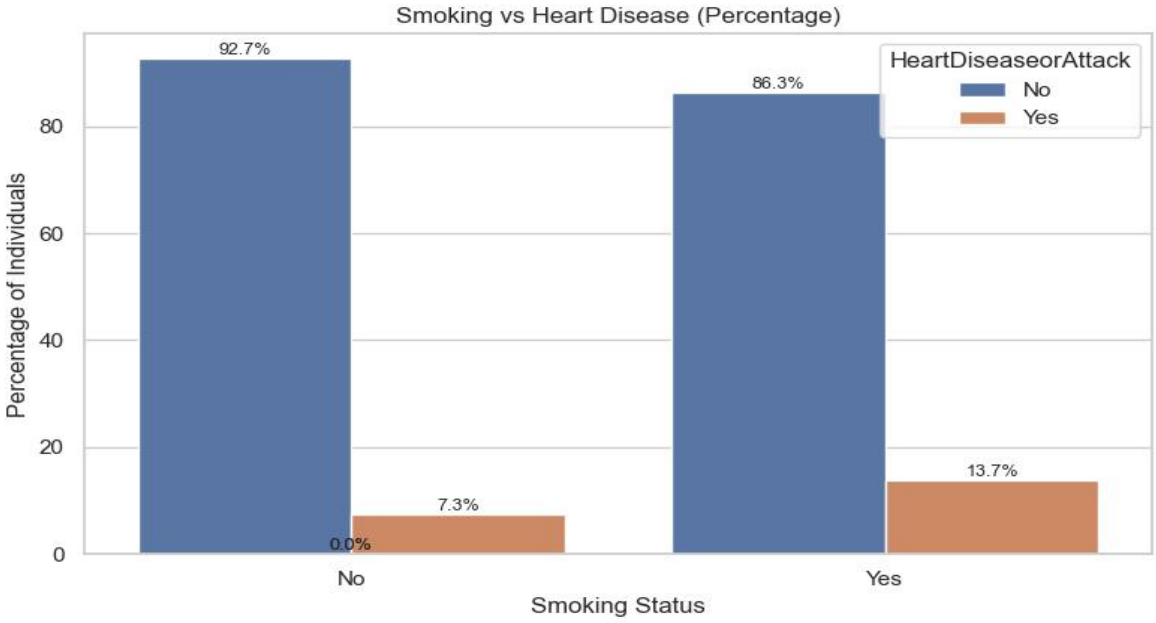


These findings reflect well-known clinical risk factors: nearly 9–10% of respondents report a prior heart attack or coronary heart disease, against a backdrop where 43% have high blood pressure, 42% high cholesterol, and 44% are smokers . The analysis confirms that managing these conditions should be a top priority.

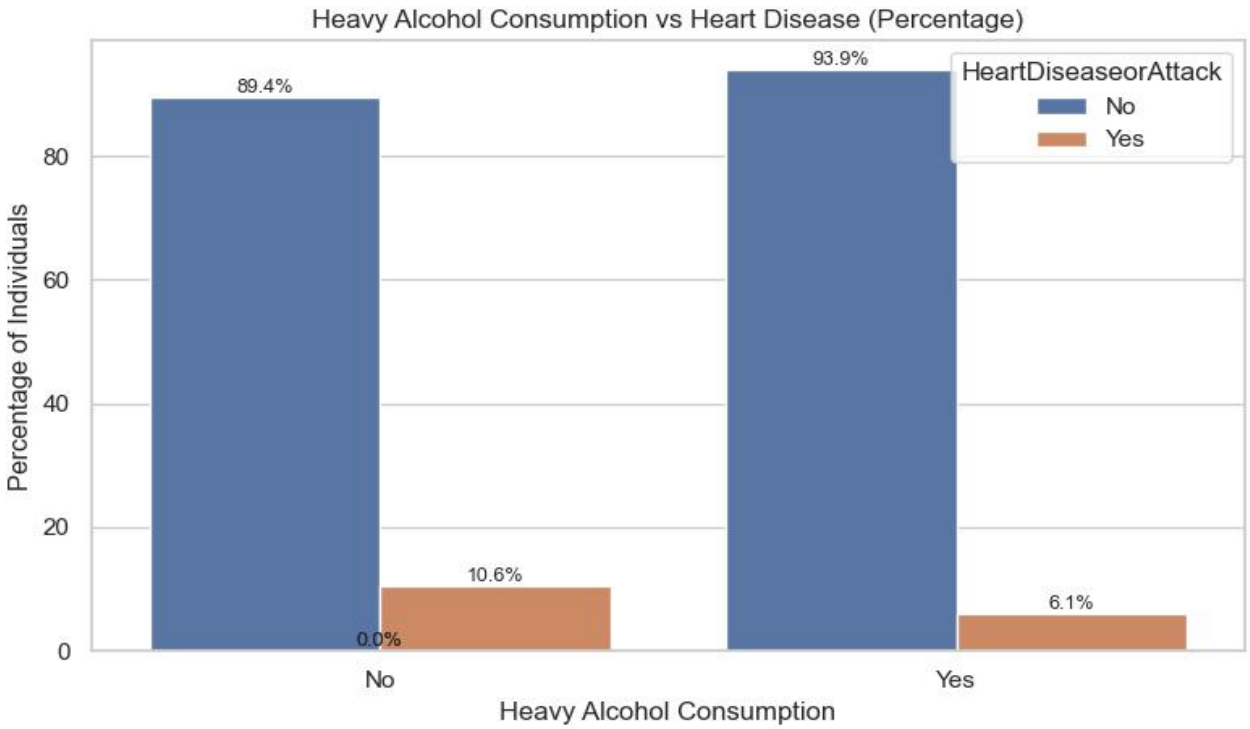
Lifestyle Factors

Modifiable behaviors also align with heart disease patterns:

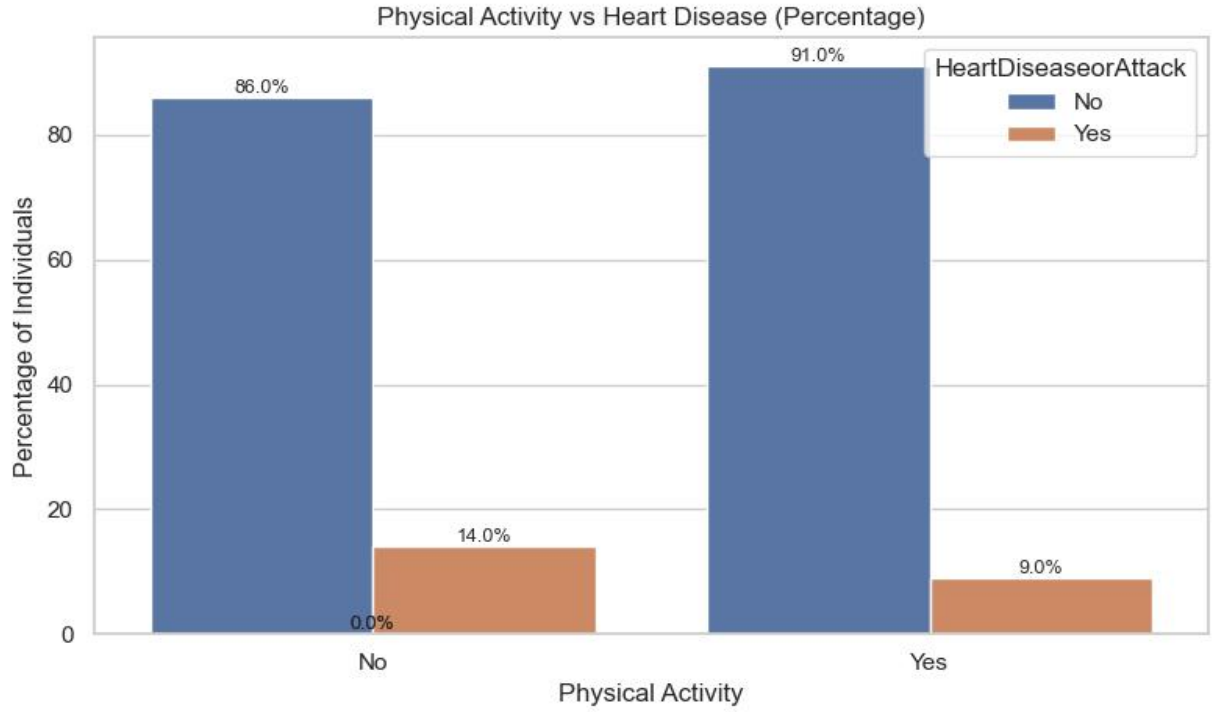
- **Smoking:** Current smokers have a higher heart disease rate than nonsmokers. This confirms that smoking is a significant modifiable risk factor for cardiovascular disease.



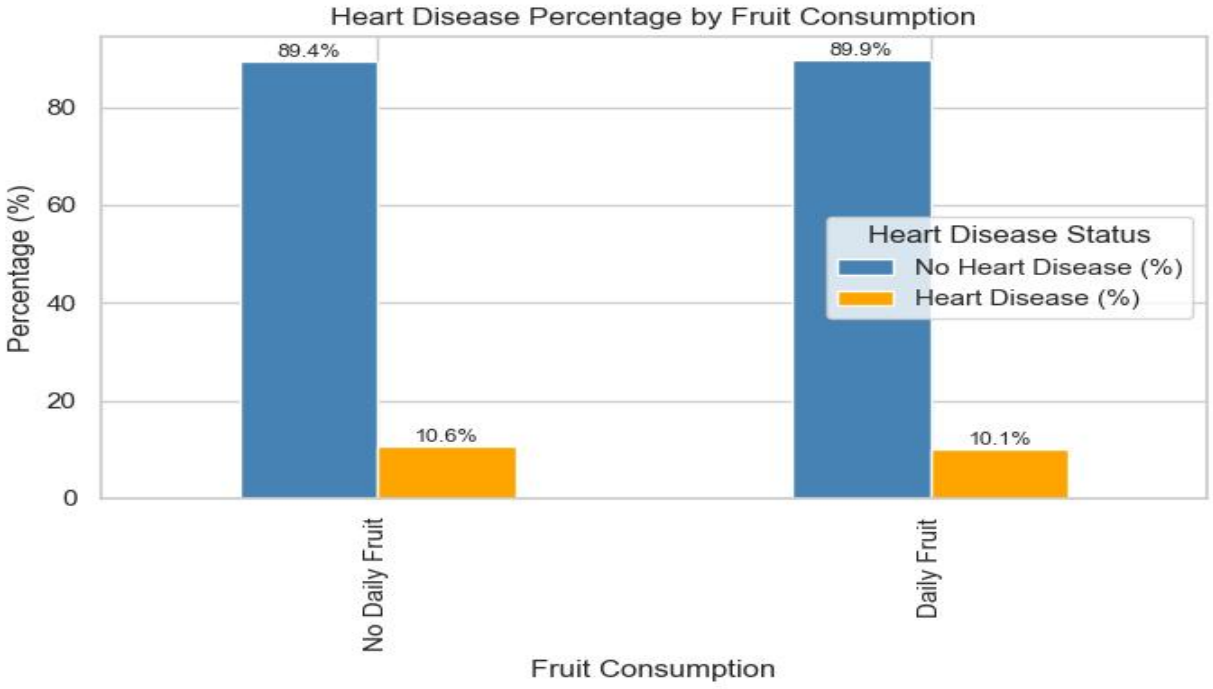
- **Alcohol:** Surprisingly, heavy drinkers in this survey show lower heart disease prevalence. This counterintuitive result is likely an artifact: heavy drinkers tend to be younger in this sample, and younger adults have naturally lower heart disease rates. In other words, age is confounding the apparent effect. This finding should be interpreted cautiously and does not contradict medical evidence that excessive alcohol increases risk.

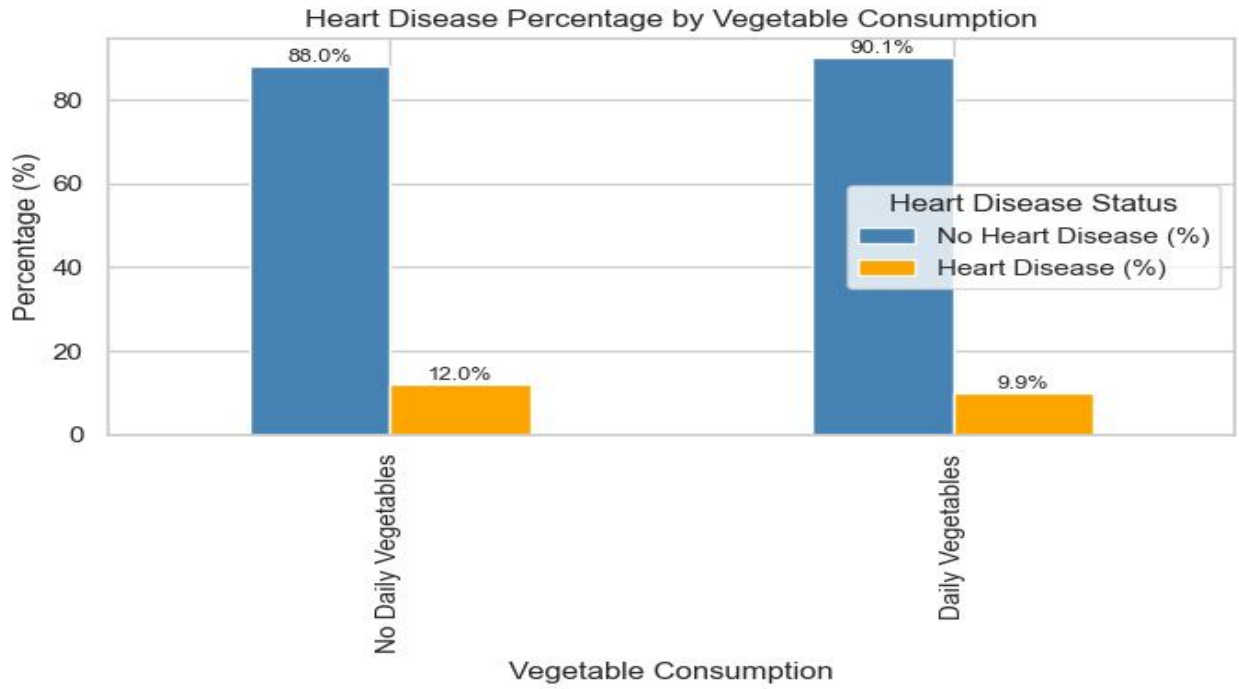


- ***Physical Activity:*** *Regular exercise is clearly protective. Only about 9% of physically active respondents have heart disease, versus ~14% of those reporting no leisure-time physical activity. This suggests that even simple activity (e.g. any exercise in the last month) is associated with substantially lower cardiovascular prevalence.*



- **Diet:** The survey's diet measures show minimal direct effects. Daily fruit consumption made almost no difference in heart disease prevalence. There is a hint that daily vegetable eaters have a slightly lower heart disease rate (~10% vs ~12%), but the difference is small. (These variables only capture whether fruits/vegetables are eaten daily, not overall diet quality, so interpretations are limited.)

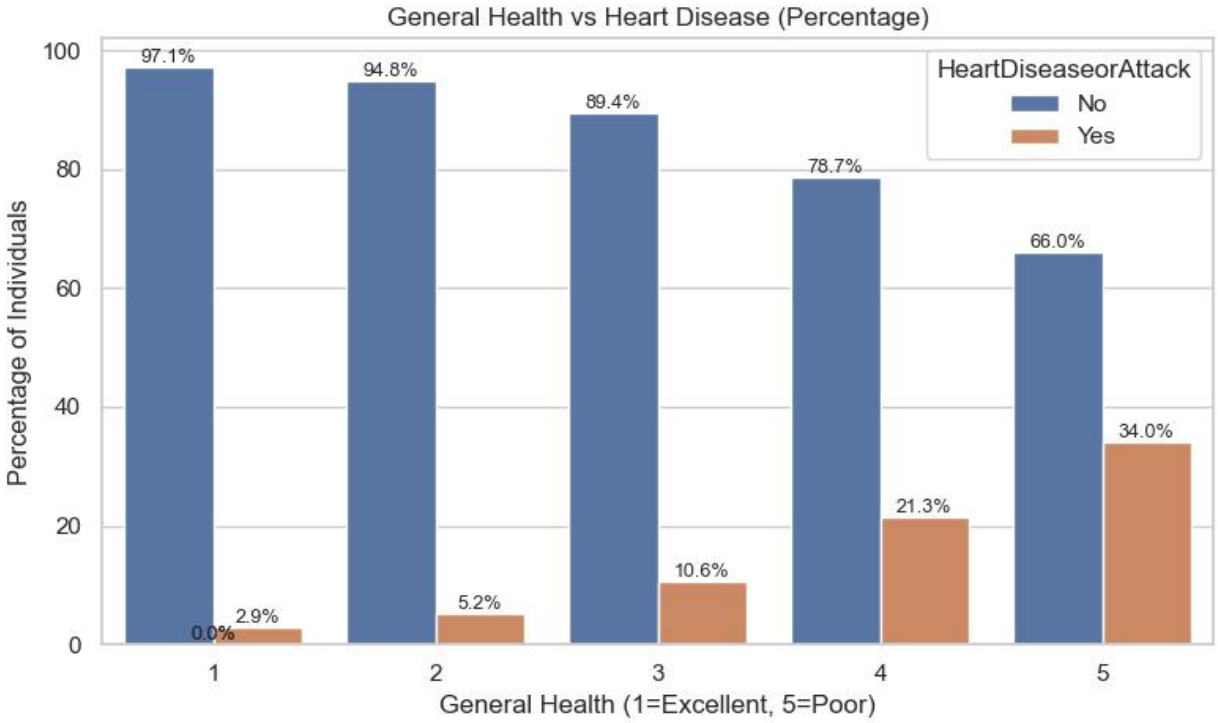




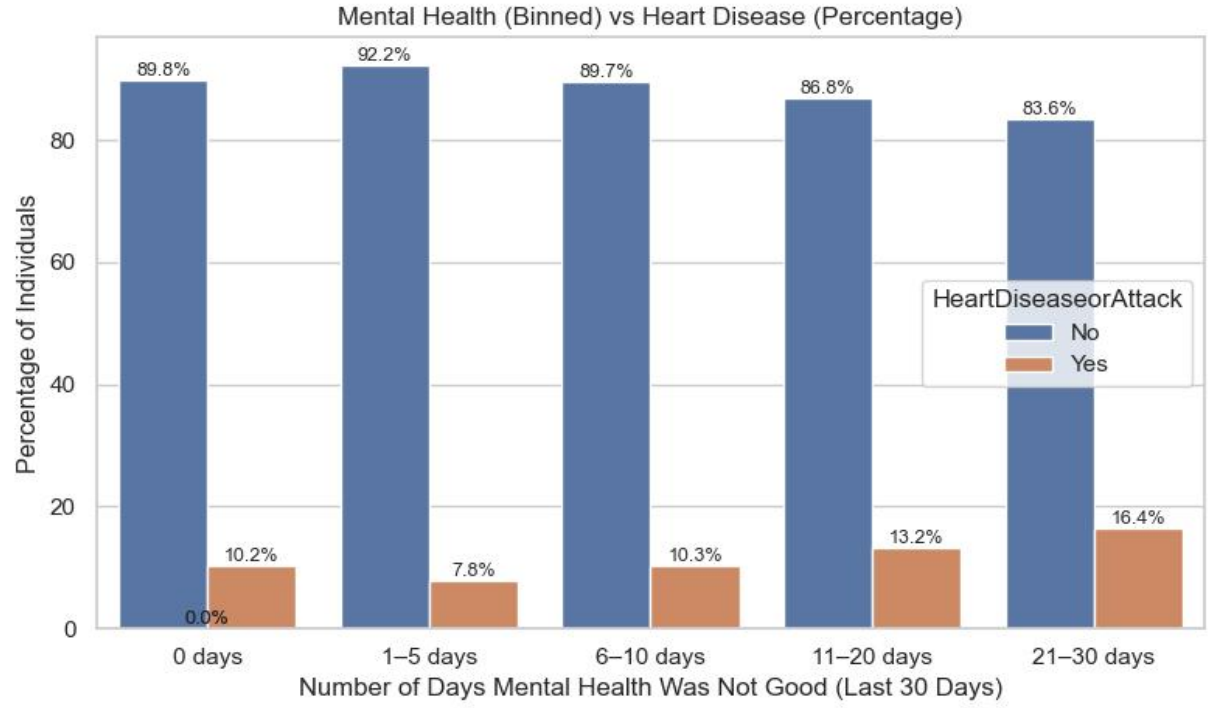
Self-Rated Health

Self-reported health status emerged as a **key indicator** of heart disease risk. Heart disease prevalence climbs sharply as general health rating worsens. Individuals who rate their health as "poor" have more than ten times the heart disease prevalence of those reporting "excellent" health. This striking gradient means that self-rated health effectively captures underlying risk: people who feel unhealthy are much more likely to have cardiovascular disease.

- **General Health:** Poor or fair health is a powerful flag for heart disease; this single survey question strongly correlates with objective heart disease prevalence.



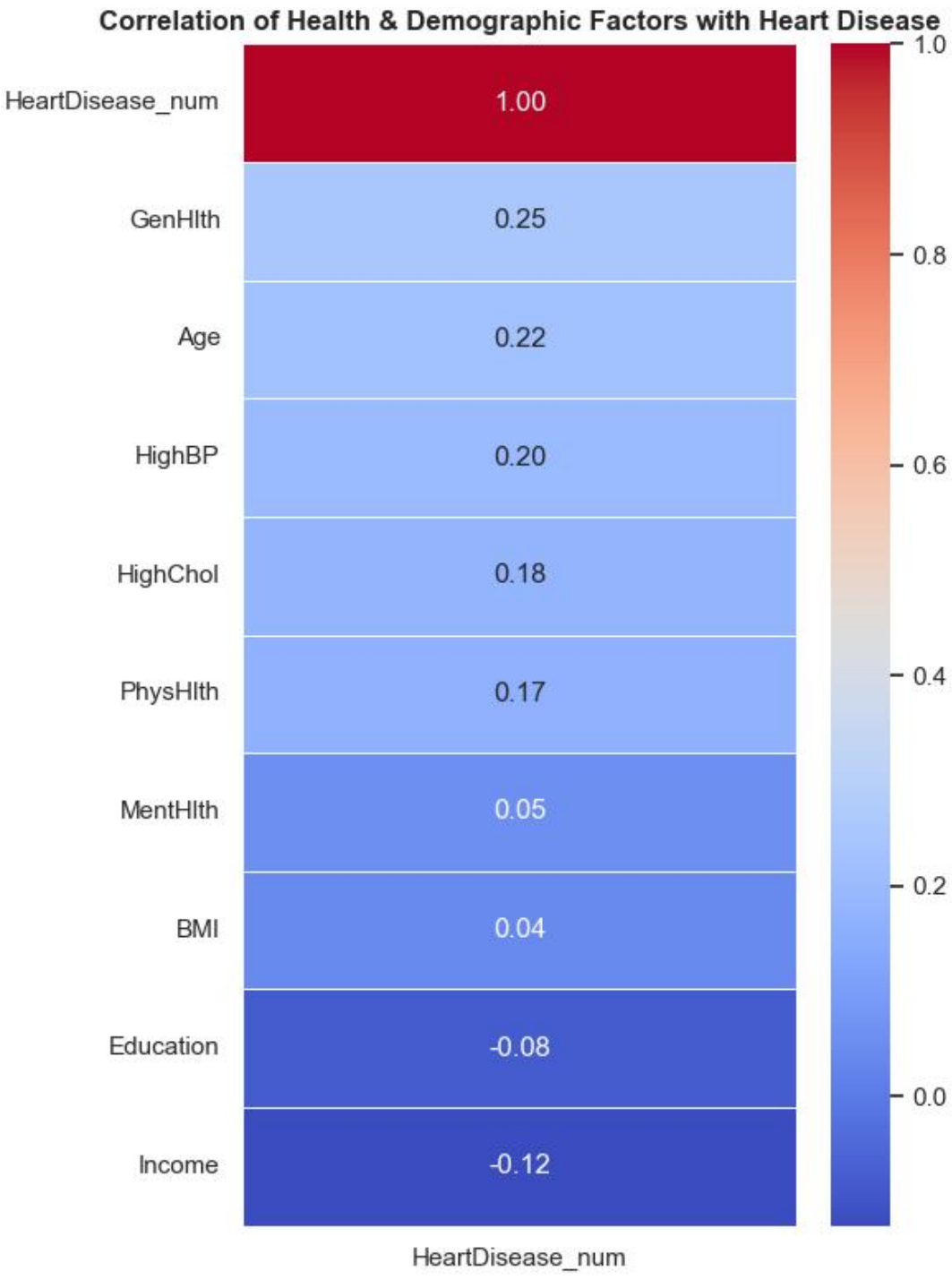
- **Mental/Physical Health:** Similarly, more days of poor mental or physical health are associated with higher heart disease rates. Those with frequent bad health days report more heart problems, reflecting the burden of chronic illness.



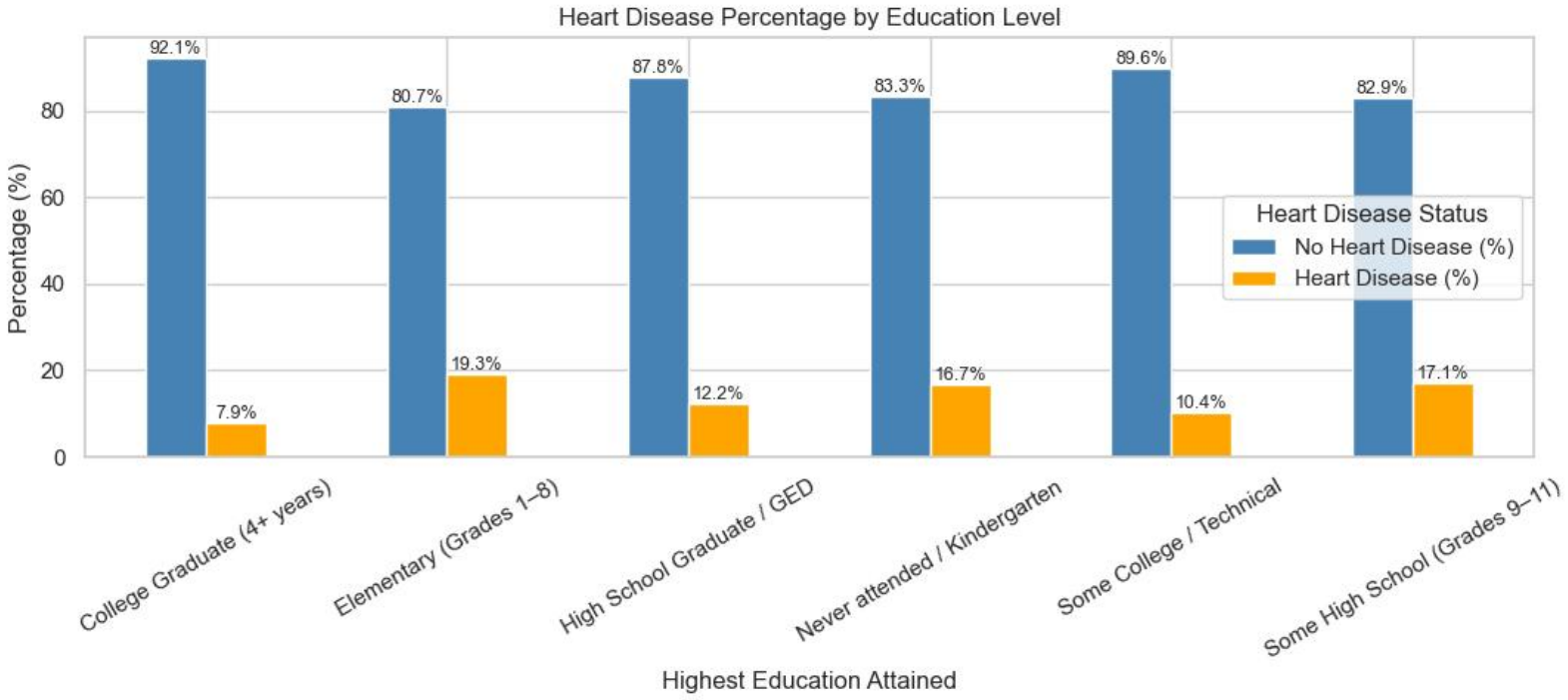
Public health leaders can leverage these insights: routine collection of self-rated health can help identify high-risk individuals for further evaluation, and improvements in overall health status will likely reduce heart disease burden.

Socioeconomic Indicators

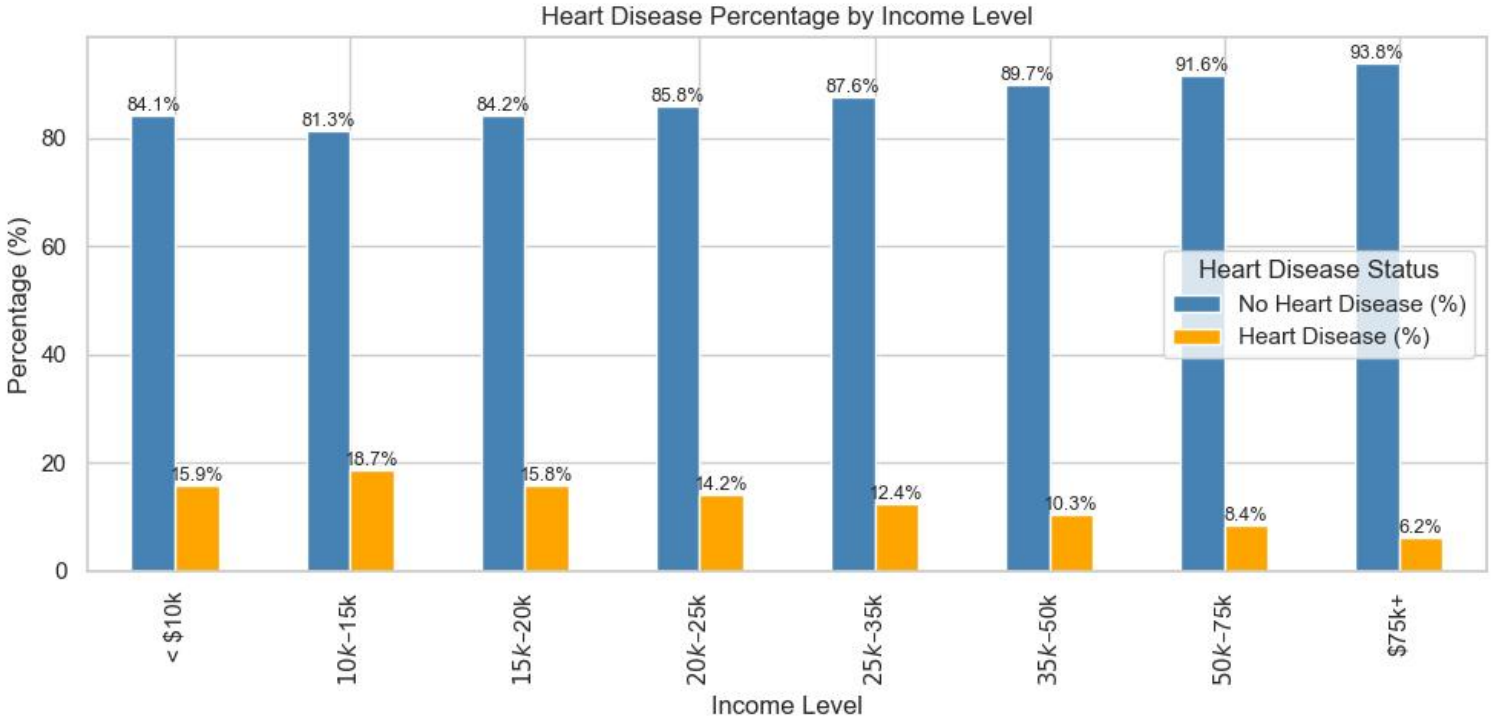
Social determinants have clear effects:



- **Education:** There is a pronounced inverse relationship between education level and heart disease. People with only elementary education have the highest heart disease rate (~19.3%), while college graduates have the lowest (~7.9%). Those with some college or a high school degree fall in between. This suggests that lower education (a proxy for health literacy, occupation, etc.) is linked to higher cardiovascular risk.



- **Income:** Heart disease prevalence falls steadily as income rises. About 16–19% of people in the lowest income brackets (<$15k) report heart disease, compared to only ~6% in the highest income group (>$75k). Middle-income groups (~$25k–$50k) show intermediate prevalence (~10–12%). This strong gradient reflects better access to healthcare, nutrition, and living conditions among higher-income groups.



These findings underscore that heart disease risk is not just about biology – social and economic context matter too. Lower-income, less-educated populations bear a disproportionate burden, likely due to factors like stress, access to care, and health behaviors. Public health strategies should explicitly address these disparities.

Strategic Recommendations

Based on these insights, we recommend action on multiple fronts:

- **Target High-Risk Clinical Groups:** Strengthen hypertension and cholesterol control programs, especially among older adults. Expand screening for diabetes and coronary issues in patients with those conditions. In practice, this could mean routine cardiovascular risk assessments in clinics that treat hypertensive or diabetic patients.
- **Promote Healthy Lifestyles:** Invest in community campaigns for smoking cessation and physical activity. The data show clear benefits: for example, physically active respondents had far

lower heart disease prevalence. Public policies could include subsidized exercise programs and stricter tobacco control measures.

- **Use Self-Rated Health as a Tool:** Incorporate general health questions into primary care and surveys. Individuals who self-report poor health should receive prioritized evaluation for cardiovascular risk. For instance, outreach programs could use self-rated health screening to identify and refer at-risk individuals.
- **Address Socioeconomic Barriers:** Develop interventions for low-income and less-educated communities. This might include mobile health clinics in underserved areas, education about heart disease prevention, and ensuring affordable medications. The inverse income and education gradients imply that social support (e.g. Medicaid expansion, health education) can yield cardiovascular benefits.
- **Coordinate Care for Comorbidities:** Given the overlap with stroke and mental health issues, integrate cardiovascular prevention into broader chronic-disease management programs. For example, stroke survivors should automatically get heart disease risk evaluation due to their dramatically higher prevalence.

In summary, the EDA confirms that **non-modifiable factors (age, sex)** intersect with **modifiable ones (BP, smoking, activity)** and **social factors (education, income)** to shape heart disease risk. Strategic efforts that combine medical management (e.g. blood pressure control) with public health measures (e.g. anti-smoking and active living campaigns) and that explicitly reach disadvantaged groups will be most effective in reducing heart disease burden.

Sources: Key findings are drawn directly from the 2015 BRFSS exploratory analysis report. These cited observations support the insights and recommendations above. Each placeholder chart represents data visualizations to accompany these points.