**Milestone 1 (Project Proposal)**

Health is the prerequisite for all human endeavors. Being healthy is something most of us take for granted, but are all internally aware is at risk. As of 2023, 92% of Americans had health insurance (public or private) (Keisler-Starkey, K. & Bunch, L. N.). This insurance comes at a substantial cost. In fact, the average health insurance premium for family coverage in 2024 was $25,572. An increase of 24% since 2019 (KFF). These costs sum to an insurance market with a volume of $318.4 billion dollars a year (Health Insurance - worldwide: Statista market forecast).

Health insurance, while it is a necessity, can also be therefore seen as a burden. It then begs the question, which we would like to investigate, *is there a practical way for individuals to reduce how much their insurance company spends on them, and thereby reduce for the individual how much they spend?*

To answer this question, we would like to analyze the Medical Insurance Cost Prediction dataset from Kaggle. This dataset includes information about 100,000 individuals including their demographics, socioeconomic status, health conditions, lifestyle factors, insurance plans, and medical expenditures. Specifically, we would like to analyze how lifestyle and habits impacts medical expenditures.

The lifestyle and habits factors we will analyze are bmi, whether the individual smokes, frequency of alcohol consumption, frequency of exercise, and hours of sleep. We would like to see how these things affect medical expenditures because there may be possibility for individuals to change these aspects of their life. And if this improves medical expenditures for the insurance company, it would seem natural for a lowering of the individual's health insurance premium. Thus benefitting the consumer, who works to live a healthier life and, in theory, would be able to reduce their premium. But also benefitting the company, as the amount they would then have to payout is reduced as well, and almost certainly reduced more than the amount of income lost from lowering premiums of those living healthy lifestyles. In this scenario, both parties would financially benefit. Since the company would benefit more, the onus would be on them to promote programs that would reduce a customer's premium if they live a healthy lifestyle.

This project should be feasible. I will work on this project alone. The aforementioned dataset is listed as having a CC0: Public Domain license, so it is completely legal to be used. The model predicting medical cost will intentionally leave off factors such as preexisting medical conditions and socioeconomic status, as I want to focus on things that people could, with oftentimes excruciating effort and discipline, change about their life. I will use multilinear regression. I believe that this is an appropriate model as it will allow for statistical inference on each of the marginal effects of each factor. It will also allow me to analyze whether some of these lifestyle choices do not have a statistically significant effect on how much the insurance company has to payout.

In order to complete this project by the due date, I have allotted 5 hours for working on this project each week. There is room for adjustment if this doesn't end up being enough. I will continue to use APA formatting for the rest of the project.

**Work Cited**


- *Health Insurance - worldwide: Statista market forecast*. (n. d.). Statista. https://www.statista.com/outlook/fmo/insurances/non-life-insurances/health-insurance/worldwide?srsltid=AfmBOorCm_d8dYNvj6Lg4gyHFabogYVBwQh880-epbVoX38UPX8tfnW7
- Keisler-Starkey, K., & Bunch, L. N. (2024, September 11). *Health insurance coverage in the United States: 2023*. Census.gov. https://www.census.gov/library/publications/2024/demo/p60-284.html#:~:text=In%202023%2C%20most%20people%2C%2092.0%20percent%20or,for%20some%20or%20all%20of%20the%20year.
- Kff. (2025, October 14). *2024 employer health benefits survey*. KFF. https://www.kff.org/health-costs/2024-employer-health-benefits-survey/#e3efa8b3-48d2-458b-a2f7-c4d5add1983b--h-section-1-cost-of-health-insurance