Executive Summary

Capstone Group 5: Healthcare

Temesgen Fekadu
Jack Lynn
Hayden Muscha
Sam Wainright

Introduction

Diabetes in the United States has been increasing over the past several decades and has major impacts on healthcare spending, treatment outcomes, and quality of life. We are investigating US residents with diabetes and the communities that they occupy in order to narrate diabetes's impact on American society.

Questions

Group 5's research focuses on diabetes data with the goal of providing a holistic view of diabetes prevalence, incidence, and maintenance in the US. Our research will use aspects of data analysis and data engineering to address the following questions:

1. Which demographics are most likely to develop diabetes in the US?
2. What measurable bodily attributes contribute to the development of diabetes?
3. How are diabetes patients' blood glucose levels tracked in real time?
4. Which demographics are exhibiting higher spikes in blood glucose levels?
5. Looking at specific regions within the US, how do different lifestyles contribute to diabetes prevalence?
6. Does food scarcity impact diabetes incidence?

Model

Our first model will predict diabetes diagnoses based on readily available medical vitals, such as blood pressure, mineral levels, and body mass index.

Our second model will predict notification times for insulin shots based on real time blood glucose levels of Type-1 diabetes patients.

Datasets

Dataset Source	Significance
National Health and Nutrition Examination Survey from CDC via Kaggle	A CDC program of studies designed to assess the health and nutritional status of adults and children in the US
Aerobic Activity by State, from CDC	Shows data from each US state and overall national data for the percentage of total respondents who are aerobically active for at least 150 minutes per week
U.S. Chronic Disease Indicators: Diabetes from CDC	Shows the prevalence of diabetes in each state as well as the prevalence of diabetes by race and ethnicity
Incident of Diabetes in Adults from Rui-Ci Health Care	A Chinese dataset containing over 200,000 datapoints for unique patients regarding medical vitals and Type-2 diabetes diagnoses; often used for prediction of diabetes diagnoses based on other medical manifestations
Continuous Glucose Monitoring (CGM) Data from Jaeb Center for Health Research	Contains CGM data from a research study on patients with Type-1 diabetes; glucose readings are taken every 5 minutes over a 26-week period for each patient
Food Security in the US via the US Department of Agriculture	Shows food security issues on a state-by-state basis
Place of Birth by Educational Attainment in the United States via the US Census Bureau	Shows levels of educational attainment for each state in the US
Income in the Past 12 Months (In 2020 Inflation-Adjusted Dollars) via the U.S. Census Bureau	Shows income brackets for each state and the percentage of the total state population that resides within that bracket

References

Centers for Disease Control and Prevention. (2017, January 26). National Health and Nutrition Examination Survey. Retrieved August 5, 2022, from www.kaggle.com website: https://www.kaggle.com/datasets/cdc/national-health-and-nutrition-examination-survey

Centers for Disease Control and Prevention. (2021, June 3). National Center for Chronic Disease Prevention and Health Promotion, Division of Nutrition, Physical Activity, and Obesity. Data, Trend and Maps [online]. Retrieved August 5, 2022, from Cdc.gov website: https://www.cdc.gov/nccdphp/dnpao/data-trends-maps/index.html

Centers for Disease Control and Prevention. (2022, March 24). U.S. Chronic Disease Indicators: Diabetes | Chronic Disease and Health Promotion Data & Indicators. Retrieved August 5, 2022, from Socrata website: https://chronicdata.cdc.gov/Chronic-Disease-Indicators/U-S-Chronic-Disease-Indicators-Diabetes/f8ti-h92k/data

Chen, Y., Zhang, X.-P., Yuan, J., Cai, B., Wang, X.-L., Wu, X.-L., … Li, X.-Y. (2018, August 21). Data from: Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. Retrieved August 5, 2022, from datadryad.org website: https://doi.org/10.5061/dryad.ft8750v

JAEB Center for Health Research. (2019). A Randomized Clinical Trial to Assess the Efficacy and Safety of Continuous Glucose Monitoring in Youth < 8 with Type 1 Diabetes. Retrieved August 5, 2022, from public.jaeb.org website: https://public.jaeb.org/dataset/563

U.S. Department of Agriculture. (2021, September 8). Food Security in the U.S. USDA (2019). Retrieved from www.ers.usda.gov website: https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/interactive-charts-and-highlights/

United States Census Bureau. (2022a, March 17). American Community Survey: B06009: PLACE OF BIRTH BY EDUCATIONAL ATTAINMENT IN THE UNITED STATES. Retrieved August 5, 2022, from data.census.gov website: https://data.census.gov/cedsci/table?q=education%20by%20state&tid=ACSDT5Y2020.B06009

United States Census Bureau. (2022b, March 17). American Community Survey: S1901: INCOME IN THE PAST 12 MONTHS (IN 2020 INFLATION-ADJUSTED DOLLARS) U.S. Census Bureau (2020). Retrieved from data.census.gov website: https://data.census.gov/cedsci/table?q=S1901&tid=ACSST5Y2020.S1901