IZMIR KATIP CELEBI UNIVERSITY

FACULTY OF ECONOMICS AND ADMINISTRATIVE SCIENCES

DEPARTMENT OF HEALTH MANAGEMENT

THE IMPACT OF DEMOGRAPHIC FACTORS ON DIABETES PREVALENCE IN THE UNITED STATES: ANALYSIS OF BRFSS 2015 DATA

Prepared by:

AYCAN KARADAĞ

210311027

Lecturer:

Prof. Dr. SERHAT BURMAOĞLU

1.Executive Summary

This report investigates the association between key demographic factors and the likelihood of being diagnosed with diabetes, using data from the Behavioral Risk Factor Surveillance System (BRFSS) 2015 survey. The dataset includes responses from over 250,000 individuals across the United States and examines how age, gender, education level, and income are related to diabetes prevalence.

The primary objective of this analysis is to support healthcare managers and policymakers in identifying vulnerable demographic groups and designing targeted preventive strategies. Our findings indicate statistically significant relationships between all examined demographic variables and the prevalence of diabetes.

Key results include:

- **Age:** Diabetes prevalence increases steadily with age, particularly after age 50. Older adults (65+) show significantly higher rates of diabetes.
- **Education:** Individuals with lower educational attainment exhibit higher diabetes prevalence. Those with only elementary or middle school education have markedly higher rates than college graduates.
- **Income:** There is a strong inverse relationship between income and diabetes risk. Low-income individuals (<$15,000 annually) are disproportionately affected.
- **Gender:** Males have a significantly higher diabetes rate compared to females.

Statistical tests (Chi-Square) confirmed these associations ($p < 0.001$ for all variables). A logistic regression model further demonstrated that higher age increases the odds of having diabetes, while higher income and education levels are associated with lower odds. Male gender also increases the likelihood of a diabetes diagnosis.

These findings underscore the importance of addressing socioeconomic and demographic disparities in diabetes prevention and management strategies. Tailored interventions that focus on education, income-related barriers, and age-specific risk communication may be particularly effective in mitigating diabetes risk at the population level.

2.The Problem

Context and Background

Diabetes is one of the most common and costly chronic diseases in the United States, affecting millions of individuals each year and placing a significant burden on both the healthcare system and the economy. It is a metabolic condition in which the body is either unable to produce sufficient insulin or cannot use insulin effectively, resulting in elevated blood glucose levels. If left unmanaged, diabetes can lead to serious complications including cardiovascular disease, kidney failure, vision impairment, and lower-limb amputations—ultimately reducing life expectancy and quality of life.

While there is no definitive cure for diabetes, early detection followed by lifestyle modifications—such as healthy eating, physical activity, and medical management—can significantly reduce its impact. Predictive tools that identify individuals at higher risk for diabetes are therefore essential for preventive health strategies and efficient resource allocation.

The scope of the problem is vast. According to the Centers for Disease Control and Prevention (CDC), as of 2018, an estimated 34.2 million Americans live with diabetes and an additional 88 million are considered prediabetic. Alarmingly, a significant proportion of these individuals are unaware of their condition, underscoring the importance of data-driven risk identification. Type 2 diabetes, which accounts for the majority of cases, is particularly influenced by demographic and behavioral factors such as age, sex, income, education, and lifestyle habits.

Moreover, the financial implications are considerable. The estimated annual cost of diagnosed diabetes in the U.S. is $327 billion, and when accounting for undiagnosed diabetes and prediabetes, the figure approaches $400 billion. These statistics highlight the critical need for early risk prediction models that can inform public health initiatives and guide targeted interventions, particularly for populations at increased risk due to social and economic disparities.

Data Sources and Methodology

This analysis is based on data from the **2015 Behavioral Risk Factor Surveillance System (BRFSS)**, a cross-sectional telephone survey conducted annually by the Centers for Disease Control and Prevention (CDC) in the United States. The BRFSS is one of the world's largest continuously conducted health surveys, collecting data on health-related behaviors, chronic conditions, and preventive service use among adults. The dataset used in this study **(diabetes_012_health_indicators_BRFSS2015.csv)** contains over 250,000 observations and includes variables related to diabetes status, demographic characteristics, and behavioral risk factors.

As part of this analysis, the original multiclass diabetes variable was recoded into a binary outcome ("Diabetic" and "Non-Diabetic") to allow for more focused modeling. Demographic predictors such as age, education, income, and gender were also recategorized to improve interpretability and allow for more meaningful subgroup comparisons.

The methodology combines descriptive and inferential techniques:

- **Descriptive statistics** were used to summarize diabetes prevalence across demographic groups.
- **Data visualization** techniques (e.g., bar charts, line graphs, heat maps) supported the identification of trends and differences in diabetes risk.
- **Chi-square tests** assessed associations between demographic factors and diabetes status.
- A **multivariate logistic regression model** was applied to estimate the independent effect of each demographic variable on the probability of diabetes.

Additionally, interaction patterns between age and gender and between income and education were examined to reveal more detailed risk profiles. All analyses were performed using R, a statistical programming environment well suited for health data analytics.

This structured approach supports evidence-based interpretation of diabetes risk factors, providing a valuable basis for targeted intervention strategies and resource allocation from a health management perspective.

Scope of Analysis

This analysis focuses on the impact of demographic factors on the likelihood of being diagnosed with diabetes among U.S. adults. The study specifically examines how variables such as age, sex, education, and income affect diabetes status. Using data from the **2015 Behavioral Risk Factor Surveillance System (BRFSS)**, the analysis provides a detailed understanding of diabetes prevalence across various demographic subgroups, targeting individuals ages 18 and older.

Specifically, the analysis examines:

- The relationship between demographic characteristics (age, sex, education, income) and diabetes status.
- How each demographic factor independently influences the likelihood of developing diabetes.
- The interaction between age and sex and income and education to identify more specific risk profiles within these groups.

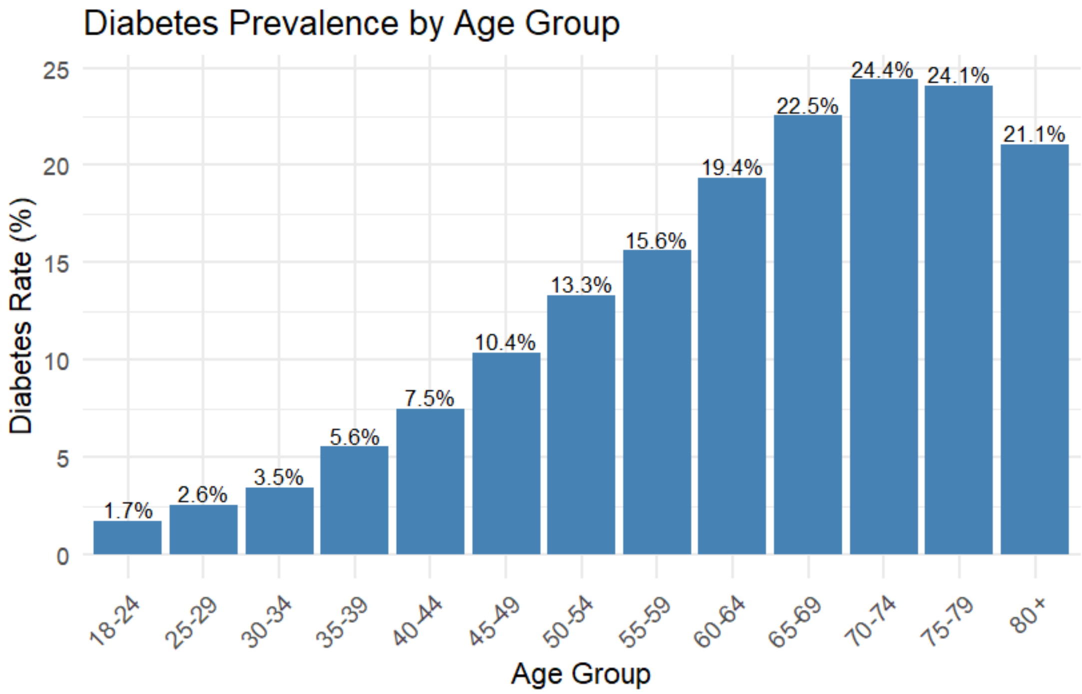
The scope is limited to variables available in the BRFSS dataset, with a specific focus on demographic predictors of diabetes. The analysis does not consider behavioral factors, clinical data, or medical history, as these types of information are not included in the survey.

This limited analysis provides a clear understanding of how demographic factors contribute to diabetes risk, providing insights that can inform public health strategies and policymaking.

3.The Evidence

Age Group and Diabetes Prevalence

One of the key findings of the analysis is the strong positive correlation between age and diabetes prevalence. The relationship between age groups and diabetes prevalence is presented in the bar chart below.



The bar chart shows the prevalence of diabetes by age group among adults (18 years and older) in the United States. The horizontal axis represents age groups (increasing ranges from 18-24 to 80+), and the vertical axis represents the rate of diabetes in each age group as a percentage (%).

The chart clearly shows a strong, positive relationship between age and diabetes prevalence. The prevalence of diabetes in the youngest age group, 18-24, is only 1.7%, while this rate increases steadily with age.

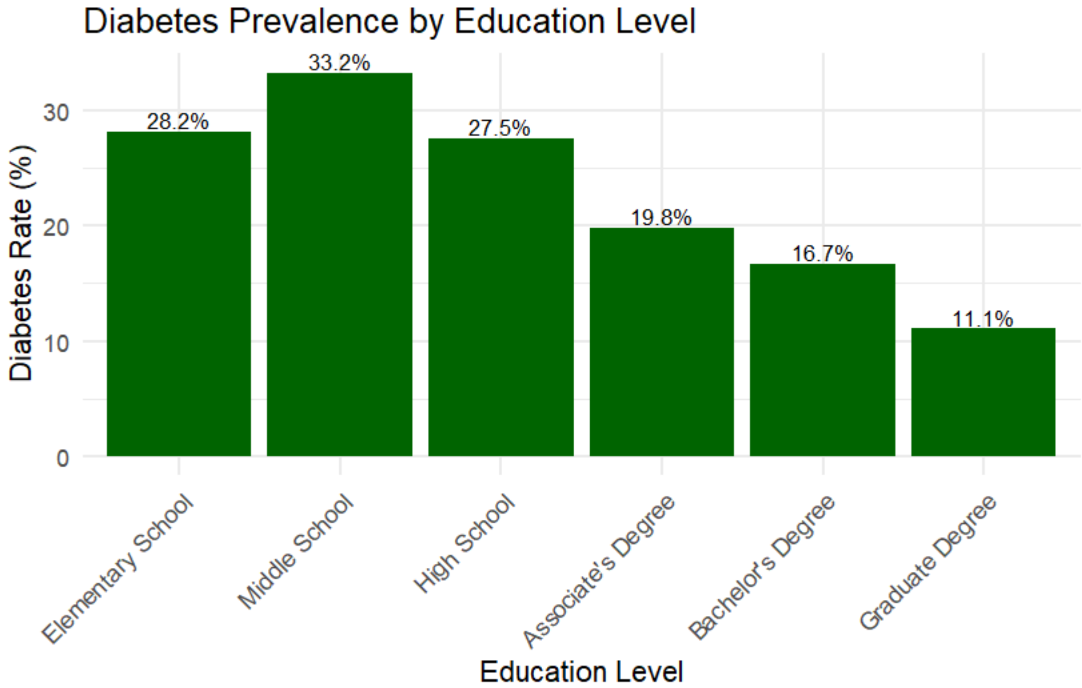
Key Findings:

- The lowest prevalence of diabetes (1.7%) is seen in the 18-24 age group. The 25-29 (2.6%) and 30-34 (3.5%) age groups also have relatively low rates.
- A significant increase in diabetes prevalence begins at age 35-39 (5.6%), and this increase continues in each subsequent age group.
- The highest prevalence of diabetes is observed in the 70-74 age groups (24.4%) and 75-79 age groups (24.1%). The 65-69 age groups (22.5%) and 80+ age groups (21.1%) also have high rates.
- The increase in diabetes prevalence accelerates after the age of 50. While the prevalence is 13.3% in the 50-54 age group, it increases to 19.4% in the 60-64 age group.

These findings clearly show that age is an important determinant of diabetes risk. Older adults have a significantly higher prevalence of diabetes compared to younger adults. This situation emphasizes the critical importance of diabetes screening and prevention strategies for older populations. In addition, it may be useful to closely monitor the increase in diabetes risk in middle-aged groups (especially after the age of 50) and to develop awareness and early intervention programs for these groups.

Education Level and Diabetes Prevalence

Education level is another important demographic factor affecting diabetes prevalence. Diabetes prevalence at different education levels is shown in the bar chart. The trend in the graph clearly shows that diabetes prevalence is higher in individuals with lower education levels.



The bar chart shows the diabetes prevalence in percentage (%) according to the highest education level of the participants. The horizontal axis shows the education levels (Primary School to Postgraduate Degree) and the vertical axis shows the diabetes rate at the relevant education level.

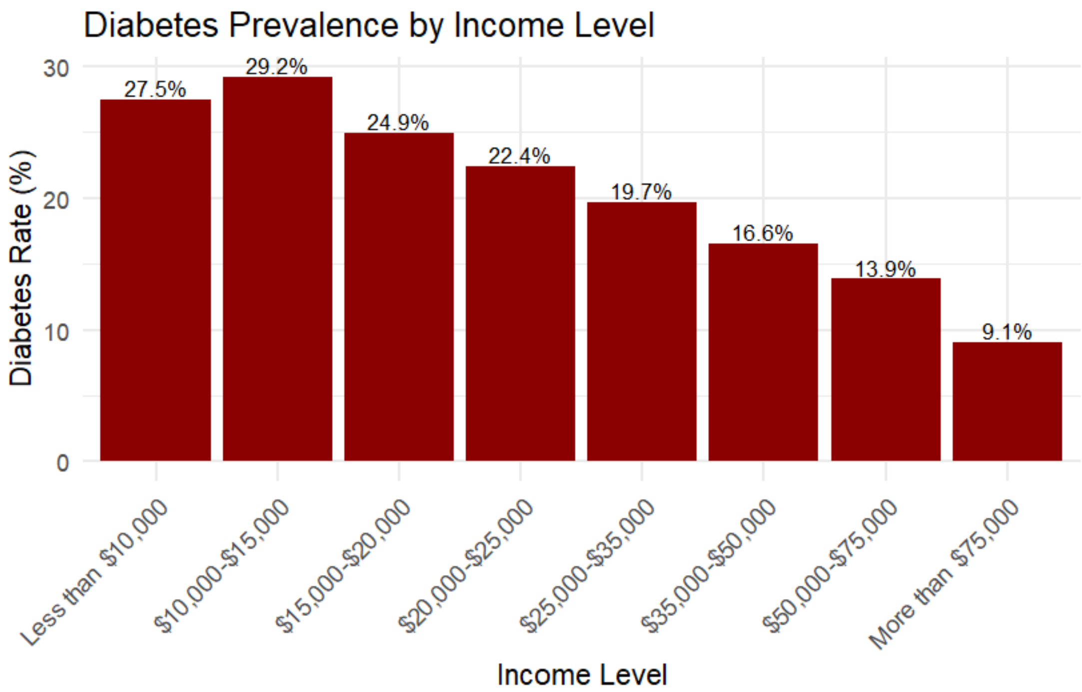
Key Findings:

- **Highest Prevalence:** The highest diabetes prevalence is seen in secondary school graduates with 33.2%. A high rate is also observed in primary school graduates (28.2%). The rate in high school graduates is 27.5%.
- **Declining Trend:** As education level increases, there is a significant downward trend in diabetes prevalence. The prevalence decreases to 19.8% among those with an Associate's Degree, 16.7% among those with a Bachelor's Degree, and a low of 11.1% among those with a Master's Degree.
- **Notable Difference:** The prevalence of diabetes among individuals with a Master's Degree (11.1%) is approximately one-third of the prevalence among secondary school graduates (33.2%).

These findings suggest a strong inverse relationship between education level and diabetes risk. Higher levels of education are associated with lower diabetes prevalence, likely related to socioeconomic factors such as health literacy, healthy lifestyle choices, and differences in access to health care. Developing diabetes prevention and health education programs for populations with low levels of education may play an important role in reducing the burden of diabetes in these groups.

Income Level and Diabetes Prevalence

Income level is closely related to access to healthcare, healthy foods, and lifestyle choices, all of which can affect the risk of diabetes. The graph shows the prevalence of diabetes across income levels, and it is clear that individuals with lower income levels have higher rates of diabetes, similar to the trend observed across education levels. This relationship suggests that financial constraints may limit access to preventive care and healthier living conditions.



The bar graph shows the prevalence of diabetes as a percentage (%) of the participants' annual income. The horizontal axis shows income levels (increasing ranges from less than $10,000 to more than $75,000), and the vertical axis shows the rate of diabetes at that income level.

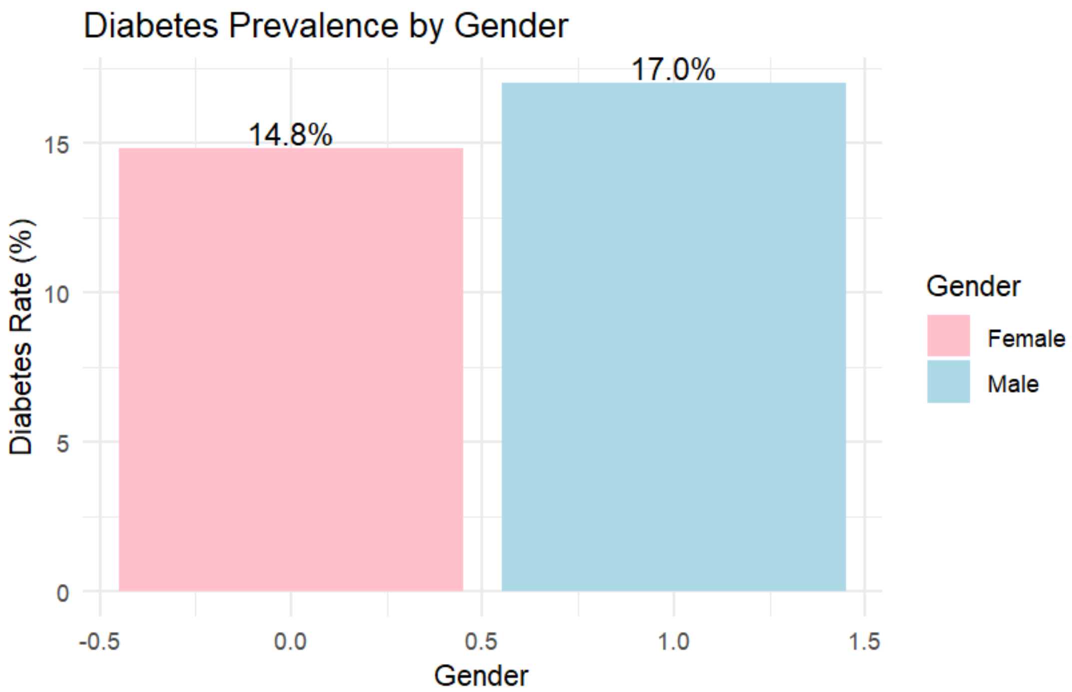
Key Findings:

- **Highest Prevalence:** The highest prevalence of diabetes is seen in the "$10,000-$15,000" income group with 29.2%. A high rate is also observed in the "less than $10,000" income group (27.5%). The rate in the "$15,000-$20,000" income group is 24.9%.
- **Declining Trend:** As the income level increases, a significant downward trend is observed in the prevalence of diabetes. The prevalence decreases to 22.4% in the "$20,000-$25,000" income group, to 16.6% in the "$35,000-$50,000" income group, and to the lowest level of 9.1% in the "More than $75,000" income group.
- **Notable Difference:** The prevalence of diabetes in the "More than $75,000" income group (9.1%) is approximately one-third that in the "$10,000-$15,000" income group (29.2%).

These findings confirm a strong inverse relationship between income level and diabetes risk. Lower income levels are associated with higher diabetes prevalence, likely related to factors such as poor nutrition, limited health care access, and stress. Reducing health disparities and improving economic support and access to health care for low-income communities are critical to preventing and managing diabetes.

Gender and Diabetes Prevalence

Gender also appears to play a significant role in the prevalence of diabetes. As shown in the graph, men have a higher prevalence of diabetes compared to women. However, this difference between the sexes is not as pronounced as observed across age and income groups.



The bar graph shows the prevalence of diabetes in male and female participants as a percentage (%). The horizontal axis shows gender (Female and Male) and the vertical axis shows the diabetes rate in the respective sex.

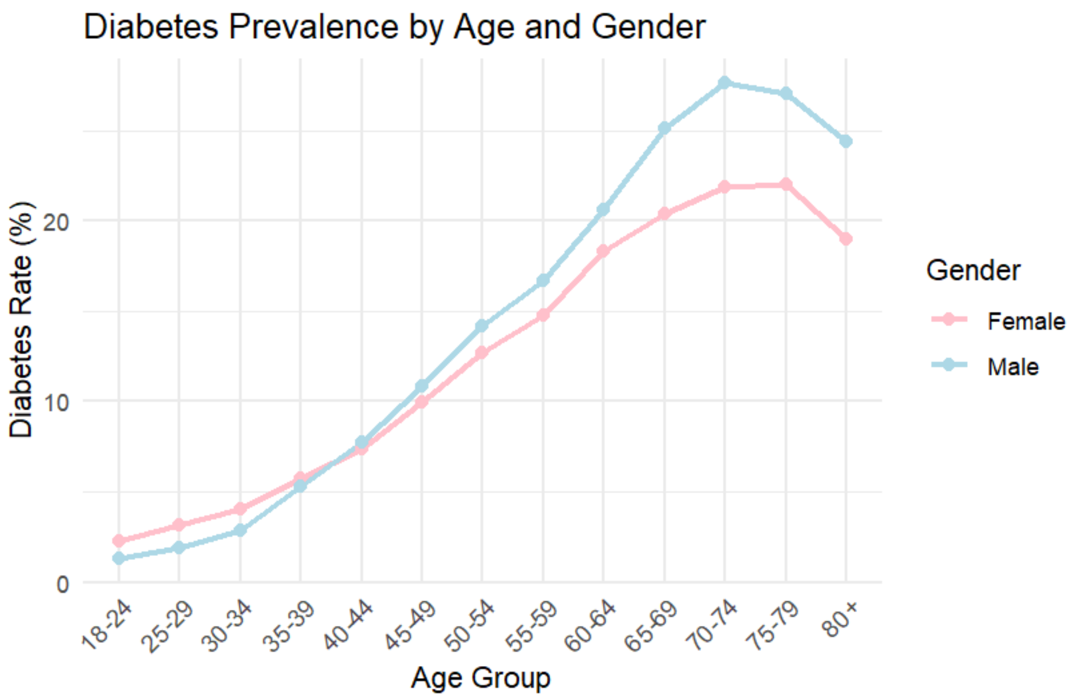
Key Findings:

- **Higher Prevalence in Males:** The prevalence of diabetes in males was recorded as 17.0%.
- **Lower Prevalence in Females:** The prevalence of diabetes in females was determined as 14.8%.
- **Relative Difference:** The diabetes rate in males is approximately 2.2 percentage points higher than the rate in females.

These findings suggest that males have a slightly higher risk of diabetes than females. However, this difference is not as pronounced as other demographic factors (age, income, education). Further investigation into the biological, behavioral, or socioeconomic factors underlying this gender difference may be necessary. This finding highlights the importance of considering gender differences in diabetes prevention and management strategies, but it should also be noted that other demographic factors have a stronger impact.

Interaction of Age and Gender

The interaction between age and gender sheds more light on the relationship between demographic factors and diabetes. The graph presents a line graph comparing the prevalence of diabetes by age group for men and women. The graph shows that the prevalence of diabetes increases with age in both genders, but it is noteworthy that this increase is more pronounced in women than in men after the age of 45.



The line graph shows the prevalence of diabetes in women (pink line) and men (blue line) in different age groups as a percentage (%). The horizontal axis shows the age groups, and the vertical axis shows the diabetes rate.

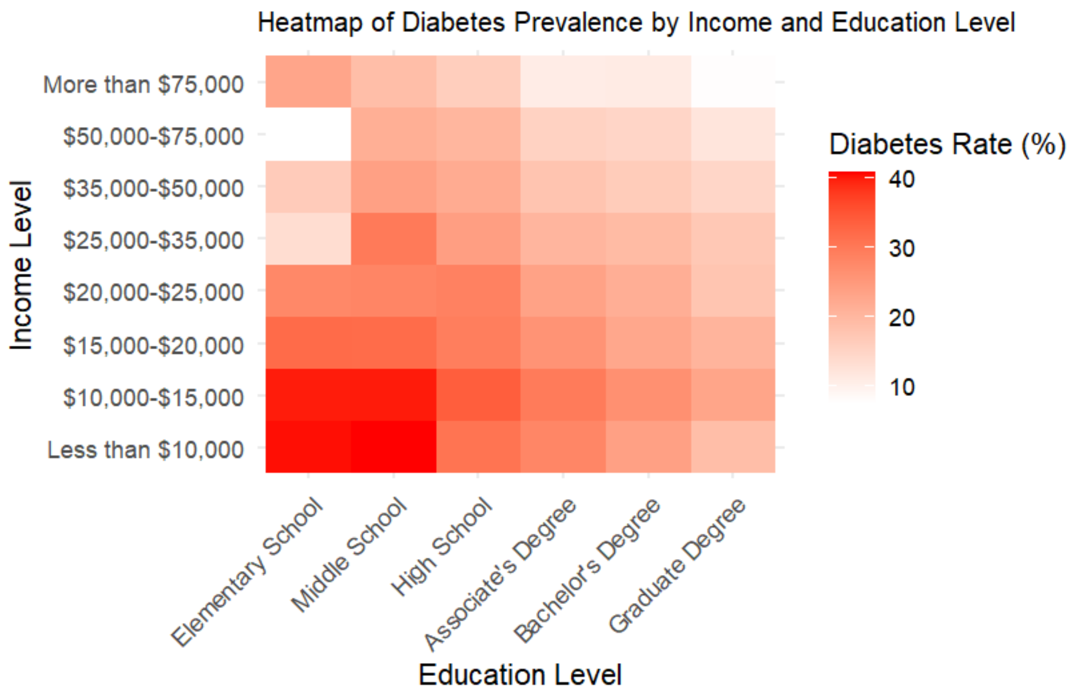
Key Findings:

- **General Increasing Trend:** The prevalence of diabetes increases with age in both genders.
- **Similarity in Early Ages:** In younger age groups (18-44 years), the prevalence rates of diabetes in women and men are quite similar. In fact, in some younger age groups, the prevalence may be slightly higher in women.
- **Divergence after Age 45:** Starting from the age group of 45-49, the prevalence of diabetes in men begins to increase more rapidly than in women.
- **Peak Points:** While the prevalence of diabetes in men reaches its highest level in the age group of 70-74 (around 27%), this peak is seen in the age group of 75-79 (around 23%) in women.
- **Narrowing in Older Ages:** In the age group of 80+, the difference between the sexes decreases slightly.

These findings suggest that the effect of age on the prevalence of diabetes varies by sex. Especially in middle age and older groups, the risk of developing diabetes in men increases more significantly than in women. This indicates that gender-specific biological, hormonal or lifestyle factors may play different roles in these age ranges. When designing diabetes prevention and management strategies, it is important to take into account gender as well as age and to develop special approaches for risk groups.

Interaction of Income and Education

The interaction between income and education is critical to understanding diabetes risk. The heat map clearly shows that individuals with both low income and low education levels have the highest prevalence of diabetes. This suggests that socioeconomic factors are intertwined and contribute jointly to diabetes risk.



The heat map visualizes diabetes prevalence by color intensity for each combination of income and education, comparing income levels (lowest to highest) on the vertical axis and education levels (lowest to highest) on the horizontal axis. Darker red colors represent higher prevalence of diabetes, and lighter colors represent lower prevalence.

Key Findings:

- **Highest Prevalence:** The darkest red areas are concentrated in areas where the lowest income groups ("Less than $10,000" and "$10,000-$15,000") intersect with the lowest levels of education (elementary and secondary school). In these combinations, diabetes prevalence is above 30%, approaching 35% at some points.
- **Declining Prevalence Trend**: As income and education levels increase, the colors in the heat map become lighter, indicating decreasing diabetes prevalence.
- **Lowest Prevalence:** The lightest colors are seen in areas where the highest income group ("More than $75,000") intersects with the highest levels of education (Bachelor's and Master's Degree). In these combinations, diabetes prevalence is below 10%.
- **Effect of Education:** Even at a given income level, there is a significant decrease in diabetes prevalence as education levels increase. Similarly, at a given education level, diabetes prevalence decreases as income increases.

These findings strongly support the idea that low socioeconomic status (a combination of low income and low education) is an important contributory factor for diabetes risk. This may be related to a variety of factors, including limited access to healthy foods, difficulty accessing health care, and more stressful living conditions for these individuals. Public health interventions need to address the multifaceted nature of socioeconomic inequalities and include strategies to improve both education and income levels to reduce diabetes risk.

4.The Implications

Interpretation of Findings

The findings from this analysis underscore the significant role of demographic factors in influencing diabetes prevalence. Age, education, income, and gender are all key predictors of diabetes risk, with age showing the strongest association. The data indicates that older individuals are much more likely to be diagnosed with diabetes, highlighting the importance of targeting older populations for early screening and prevention efforts. Additionally, lower education and income levels were found to be associated with higher diabetes prevalence, suggesting that socio-economic factors play a critical role in health outcomes.

As diabetes prevalence is higher among older individuals and those with lower education and income levels, these groups require more targeted interventions. Public health policies should address these disparities by improving access to healthcare, preventive measures, and educational resources.

Connections to Broader Context

These findings are not just limited to the dataset at hand but reflect broader global trends in the management and prevention of diabetes. Worldwide, diabetes is becoming an increasing burden on healthcare systems, particularly in low- and middle-income countries, where access to care and resources is often limited. The link between socio-economic factors (education and income) and diabetes risk is well-documented in health literature, emphasizing that economic inequality can exacerbate health disparities.

The relationship between socio-economic factors and diabetes prevalence reflects a wider societal issue of health inequity. In many cases, individuals from lower socio-economic backgrounds face barriers to accessing healthy foods, preventive care, and education, which in turn increases their risk of chronic diseases such as diabetes. Tackling these disparities is essential not only for improving individual health outcomes but also for reducing the overall burden on healthcare systems.

Potential Consequences if Unaddressed

If the identified disparities are not addressed, there are significant long-term consequences both at the individual and systemic levels. Failing to target high-risk groups—such as older adults and those with lower socio-economic status—could lead to an increase in the prevalence of diabetes, resulting in higher healthcare costs, greater demand for diabetes-related care, and a diminished quality of life for individuals affected. Moreover, ignoring these issues could exacerbate existing health disparities, leading to further social and economic inequality.

Consequences of Inaction:

- **Healthcare System Strain:** An increase in diabetes prevalence, especially among vulnerable populations, will place an additional strain on healthcare systems, leading to higher costs associated with treatment, hospitalization, and long-term care.
- **Economic Impact:** Chronic diseases like diabetes lead to lost productivity and increased absenteeism from work, further exacerbating economic inequalities and reducing overall societal productivity.
- **Health Disparities:** Without addressing the socio-economic determinants of diabetes, the health inequities between different demographic groups will continue to widen, leading to poorer outcomes for those in disadvantaged socio-economic groups.

The implications of these findings are clear: to effectively manage and reduce diabetes prevalence, targeted interventions and policies must focus on the most vulnerable populations, particularly older adults and those from lower education and income backgrounds. Addressing these factors will not only improve individual health outcomes but will also contribute to a more equitable and sustainable healthcare system.

5.The Solution

Recommendations Based on Evidence

1.Targeted Screening Programs for Older Adults:

- Given the strong association between age and diabetes prevalence, particularly among individuals aged 65 and above, it is crucial to implement targeted diabetes screening programs for older adults. These programs should focus on early detection, preventive care, and regular monitoring for those at high risk.
- **Actionable Step:** Implement community-based screening campaigns and collaborate with primary healthcare providers to identify high-risk individuals and connect them with preventive services and treatment options.

2.Improving Access to Healthcare for Low-Income and Low-Education Groups:

- The data suggests a higher prevalence of diabetes among individuals with lower education and income levels. To address these disparities, healthcare access must be improved for disadvantaged populations. This includes increasing access to preventive care, diabetes education, and treatment options.
- **Actionable Step:** Expand public health initiatives and subsidized healthcare programs for low-income individuals, ensuring that diabetes-related care is affordable and accessible. Additionally, offer education campaigns in underserved areas to raise awareness of diabetes prevention.

3.Health Education and Lifestyle Support:

- Empowering individuals with knowledge about diabetes prevention through lifestyle changes (such as diet, physical activity, and smoking cessation) is key to reducing diabetes prevalence. Educational programs targeting those in lower socio-economic groups should be prioritized, as they are more likely to benefit from lifestyle interventions.

- Actionable Step: Develop community-based health education programs, focusing on healthy eating, physical activity, and smoking cessation. Utilize local community centers, schools, and healthcare facilities to reach a broader audience.

4.Promote Workplace Health Programs:

- Since diabetes is a chronic condition that affects productivity, workplace health programs could play a significant role in early detection and lifestyle modification for adults, especially in industries with a higher proportion of individuals with lower education and income levels.
- **Actionable Step:** Partner with businesses to establish workplace wellness programs, offering regular health screenings and support for employees to make healthier lifestyle choices.

Implementation Considerations

The implementation of the proposed solutions requires careful planning and coordination across multiple sectors, including healthcare providers, public health agencies, community organizations, and policymakers.

Key considerations include:

1.Resource Allocation:

- Successful implementation will require adequate funding and resources, particularly for community-based programs and healthcare subsidies for low-income individuals. Governments and private organizations must collaborate to ensure that funding is directed toward high-risk populations.
- **Actionable Step:** Secure funding through federal and state grants, private sector partnerships, and non-profit organizations to support community outreach and healthcare access programs.

2.Cultural and Socioeconomic Sensitivity:

- It is essential that the interventions be culturally and socioeconomically sensitive. Health education programs and diabetes prevention strategies should be tailored to the needs of diverse populations, considering factors such as language, cultural practices, and economic constraints.
- **Actionable Step:** Design bilingual or multilingual educational materials, and work with community leaders to ensure that the messages resonate with different cultural groups. Offer financial incentives for those in low-income brackets to participate in screening and treatment programs.

3.Collaboration Across Sectors:

- Addressing diabetes requires a multi-faceted approach, and collaboration between public health agencies, healthcare providers, employers, and community organizations will be essential. This should include data sharing and coordination of services to ensure that individuals receive comprehensive care and follow-up support.

- **Actionable Step:** Create partnerships with local healthcare providers, community organizations, and businesses to facilitate access to screenings, education, and treatment.

4.Monitoring and Evaluation:

- A robust system for monitoring and evaluating the impact of these interventions is crucial to ensure that the programs are effective and reach the target populations. Key performance indicators (KPIs) should be established, such as the number of individuals screened, the number of people participating in educational programs, and the reduction in diabetes prevalence over time.
- **Actionable Step:** Implement regular follow-up surveys and data collection to track outcomes, adjusting programs based on feedback and results.

Expected Outcomes

The proposed solutions aim to achieve several important outcomes, both short-term and long-term:

1.Reduction in Diabetes Prevalence:

By targeting high-risk populations—especially older adults, those with lower education and income levels—through screening, education, and access to care, the prevalence of diabetes should decrease over time. Early detection and lifestyle modifications will help reduce the incidence of new diabetes cases and better manage existing conditions.

2.Improved Quality of Life:

For individuals who are identified as at risk or diagnosed with diabetes, timely interventions will lead to improved health outcomes and quality of life. Diabetes management through lifestyle changes and medical care will help prevent complications associated with the disease, such as cardiovascular issues and kidney failure.

3.Cost Savings for the Healthcare System:

Addressing diabetes early through preventive measures and improved management will lead to long-term cost savings for the healthcare system. The reduction in complications associated with uncontrolled diabetes, such as hospitalizations and long-term care, will lower overall healthcare costs.

4.Reduction in Health Disparities:

By focusing on socio-economically disadvantaged groups, these interventions will help reduce the disparities in diabetes prevalence and health outcomes. This will lead to more equitable access to healthcare and improved health outcomes across all demographic groups.

5.Increased Productivity and Economic Benefits:

As diabetes management improves, individuals will experience fewer sick days and better overall health, leading to increased productivity in the workforce. Employers and the broader economy will benefit from a healthier workforce and reduced absenteeism.

6.The Appendix

Detailed Methodology

This study utilized data from the 2015 Behavioral Risk Factor Surveillance System (BRFSS) to examine the relationship between demographic factors (age, education, income, and sex) and the likelihood of being diagnosed with diabetes.

1.Data Preparation:

The outcome variable Diabetes_012 was recoded into a binary variable: 1 for diabetic/prediabetic, 0 for non-diabetic.

Demographic variables were categorized:

- Age grouped into intervals.
- Education recoded into Low, Middle, High levels.
- Income grouped into categorical ranges.
- Sex was coded as 0 = Female, 1 = Male.

2.Analytical Methods:

Descriptive statistics were computed to show the prevalence of diabetes across different demographic categories.

Chi-square tests were conducted to assess the association between each demographic factor and diabetes status.

A binary logistic regression model was fitted using the formula:

- Diabetes_012 > 0 ~ Age + Education + Income + Sex,

where the coefficients represent the independent effects of each factor on the odds of being diabetic or prediabetic.

3.Model Summary:

All predictors were statistically significant at the $p < 0.001$ level.

Key findings from the logistic regression:

Variable	Estimated Coefficient	p-Value
Age	0.167	< 2e-16
Education	-0.149	< 2e-16
Income	-0.167	< 2e-16
Sex (Male)	0.325	< 2e-16
(Intercept)	-1.565	< 2e-16

Model fit statistics:

- **Null deviance:** 221,031
- **Residual deviance:** 204,953
- **AIC:** 204,963

Additional Visualizations

To support the statistical findings, the following visualizations were developed:

1.Bar Graph - Diabetes Prevalence by Age Group

Displays the increasing trend in diabetes with age.

2.Bar Graph - Diabetes Prevalence by Education Level

Shows lower diabetes rates among higher educated groups.

3.Bar Graph - Diabetes Prevalence by Income Level

Indicates lower prevalence of diabetes in higher income brackets.

4.Bar Graph - Diabetes Prevalence by Gender

Demonstrates higher prevalence in males than females.

5.Line Graph - Diabetes Rates by Age and Gender

Visualizes age-specific diabetes trends for males and females.

6.Heatmap - Diabetes Rates by Income and Education

Highlights high-risk groups combining low income and low education.

Each graph was created in R using ggplot2, and the visual patterns aligned with statistical outcomes, providing intuitive insights into group differences.

References and Data Sources

- Centers for Disease Control and Prevention (CDC) and Behavioral Risk Factor Surveillance System (BRFSS), 2015. https://www.cdc.gov/brfss
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/12
- RStudio (Version 2024.12.1.0). RStudio, Boston, MA. https://www.rstudio.com/
- OpenAI. (Accessed April 20, 2025). ChatGPT. Used for: Code generation and assistance in report writing. https://openai.com/chatgpt
- Anthropic. (Accessed April 20, 2025). Claude. Used for: Assistance in report writing. https://www.anthropic.com/claude
- Statistical Methods: Standard statistical methods (descriptive statistics, Chi-Square tests, logistic regression) were implemented using the R programming language (R Core Team, 2021).