DIABETES Analysis Cases in United States

Our Team

Hanna Sofea binti Mohd Zaid
2024793105

Aliyah Majdina binti Mudzaffar Alfian
2024905023

Nur Fatihah binti Sanusi
202474519

Muhammad Afif Ezani bin Mohd Nor Apendi
2024556405

Project Background

- Diabetes: A chronic health condition caused by body's failure to regulate blood glucose level
- Remains a significant public health concern in US
- The project is based on health dataset of patient records in the US

Problem Statement

- Variability of the data (mixed data types)
- Large amount of data
- Difficult to manually interpret relationships between predictors & diabetes outcomes

Data Context

- Diabetes Clinical Dataset from Kaggle
- 100k records of patient data
- Combining demographic information with important physiological parameters (17 attributes)
- 8-year period from 2015 to 2022 from 55 different distinct states in the US

Target Audience

- **Public healthcare analysts**

discover causes, relationships, and evidence

- **Healthcare practitioners**

analyzing healthcare data to support operational

- **Medical researchers**

disease patterns, distribution, and causes

Project Aims

- Perform thorough analysis of a large-scale diabetes clinical dataset
- Improve knowledge how different health indicators contribute to diabetes
- Convert complex clinical dataset into valuable insights

Research Objectives

1) To identify demographic and health distribution patterns among patients

2) To compare health risk indicators between diabetic and non-diabetic patients

3) To examine the distribution and trends of diabetes prevalence

Research Questions (Objective 1)

1) What is the distribution and trend of diabetes status from 2015 to 2022?

2) How are diabetic and non-diabetic patients distributed by gender, age group and race?

3) Does diabetes status differ by gender?

4) How are diabetes cases geographically distributed across US states?

5) What is the proportion of patients with hypertension or heart disease among age groups?

Research Questions (Objective 2)

1) How does average BMI differ between diabetic and non-diabetic petients across BMI categories?

2) How does HbA1c level and blood glucose level differ between diabetic and non-diabetic petients?

3) How does the prevalence of hypertension differ between diabetic and non-diabetic patients?

4) How does the prevalence of heart disease differ between diabetic and non-diabetic patients?

5) How does smoking history differ between diabetic and non-diabetic patients?

Research Questions (Objective 3)

1) How does diabetes prevalence vary by age groups across genders?

2) How has the distribution of diabetes cases across BMI categories changed from 2015 to 2022?

3) How has the number of diabetes cases changed from 2015 to 2022 across gender categories?

Derivates Attributes

Age Group

Senior, Adult, Teenager, Young Adult, Child

BMI Category

Underweight, Normal, Overweight, Obese

Race

African American, Asian, Hispanic, Caucasian, Other

Conclusion

- Hypertension and heart disease are more common in people with diabetes
- Diabetic patients exhibit higher BMI, particularly in overweight and obese categories
- HbA1c and blood glucose levels are significantly higher among diabetics
- Smoking history, especially former and current smokers, shows higher diabetes prevalence

THANK YOU

https://app.powerbi.com/links/YyMhXVOtX5?ctid=cdcbb0e2-9fea-4f54-8670-672707797ada&pbi_source=linkShare&bookmarkGuid=4cbfd7f2-a6d5-4129-903f-20edfaa79b28