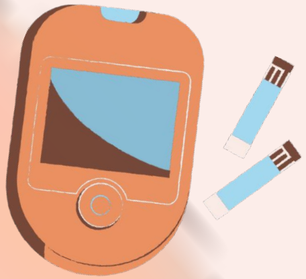
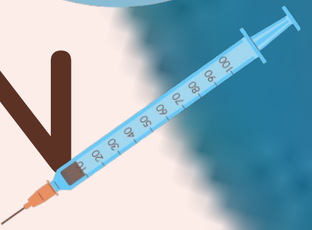
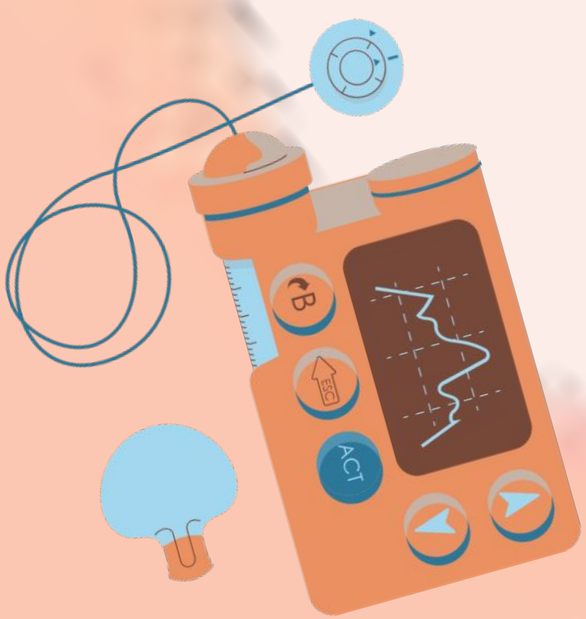
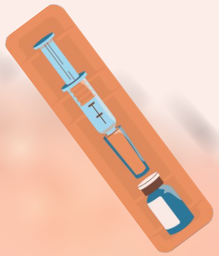


DIABETES PREDICTION ANALYSIS







TODAY'S AGENDA

- About dataset
- Visualization of data preprocessing
- Hypothesis Testing
- Key findings and insights
- Conclusion and Recommendations

WHY DIABETES MATTERS?

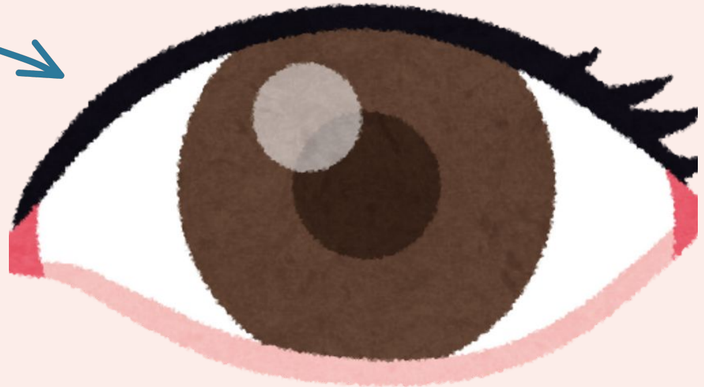
"Diabetes is a silent condition. Many people don’t know they have it until it causes serious problems like heart disease, kidney failure, or vision loss.
But what if we could find out earlier? Early detection can save lives by helping people take action before it’s too late.
Our project aims to use data to uncover patterns predict diabetes, and make early d possible."

Early Detection Saves Lives

No Symptoms

Diagnosis

Complications



ABOUT DATASET?

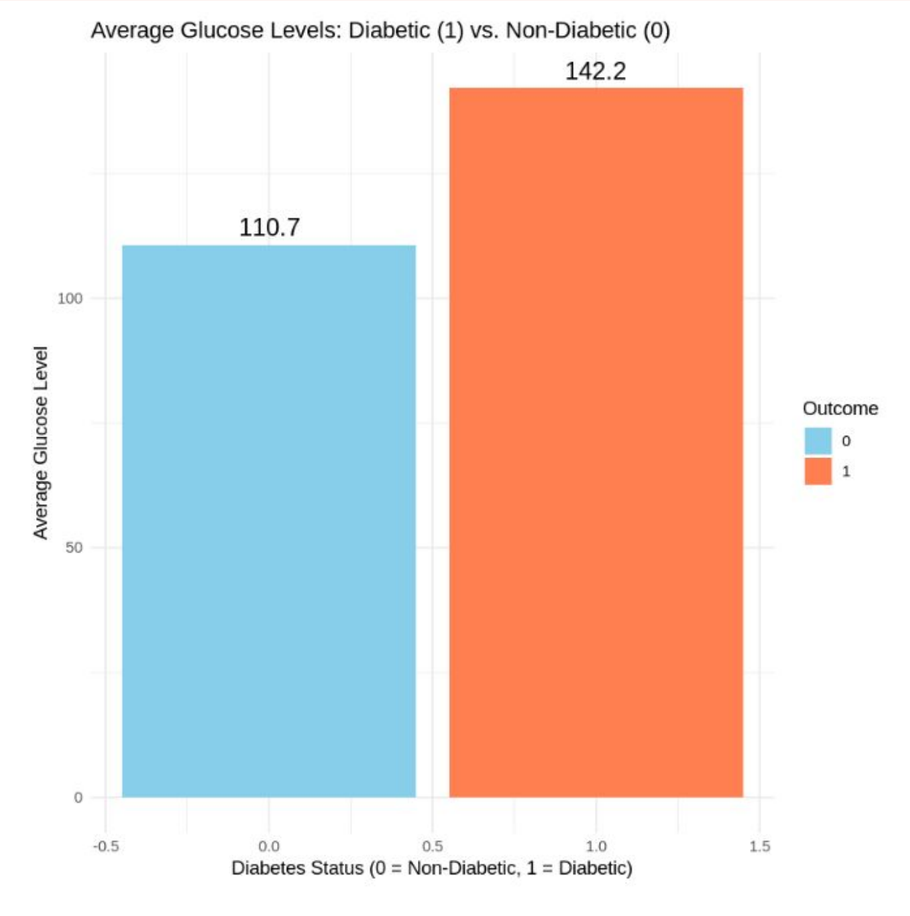
- Preprocessing data display data (num of columns, rows, duplicates)
- Handling Zero Values in the Dataset using (Mean, Median)
- using visualization to analyze data to get insights

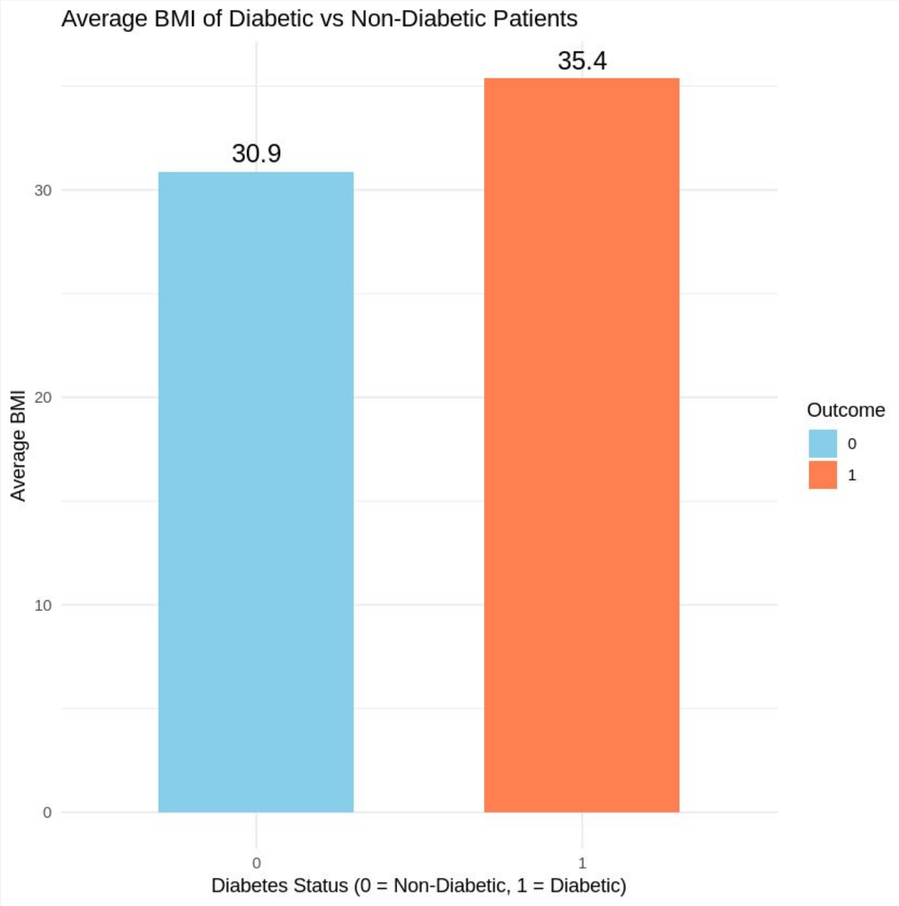
WHAT ARE THE VISUALIZATION AFTER PREPROCESSING?

WHAT IS DIABETES?

Write a short background about diabetes.

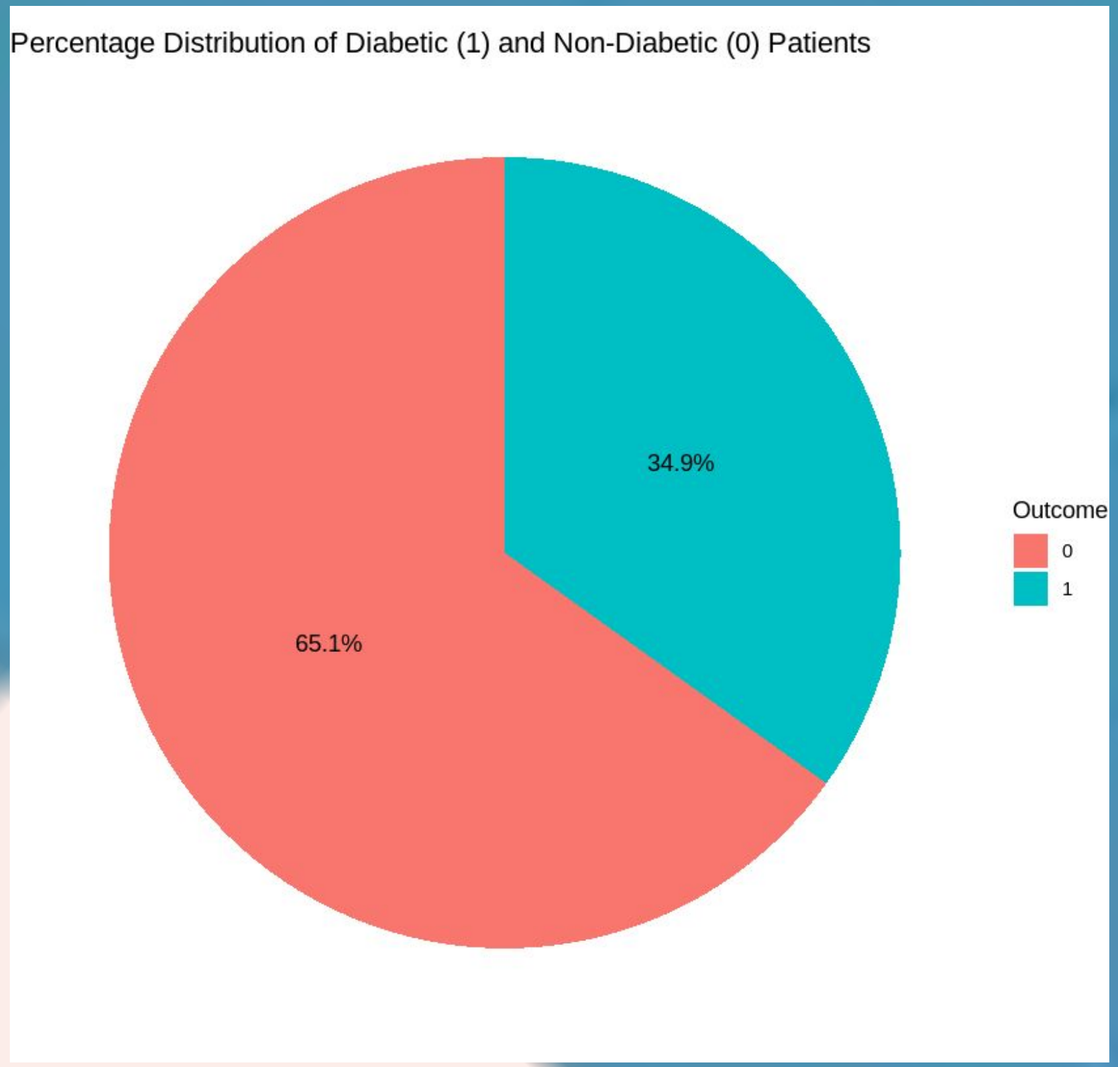
Using Bar Chart to visualize





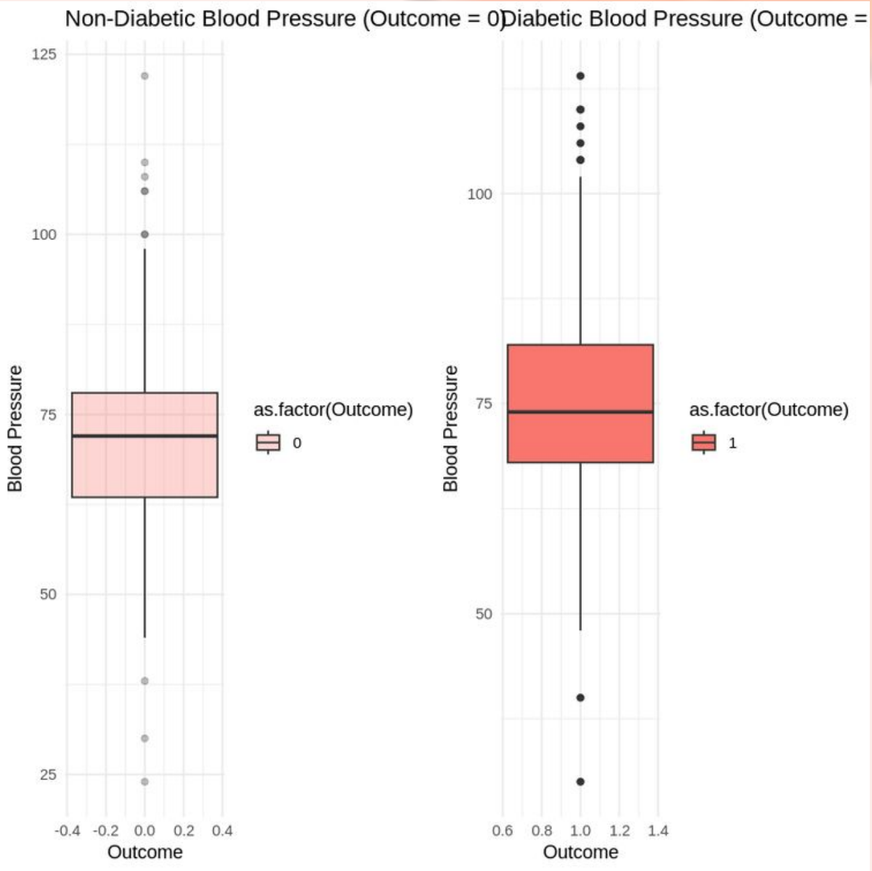
- average of Glucose level between Diabetic, Non-Diabetic patients

- average of BMI level between Diabetic, Non-Diabetic patients

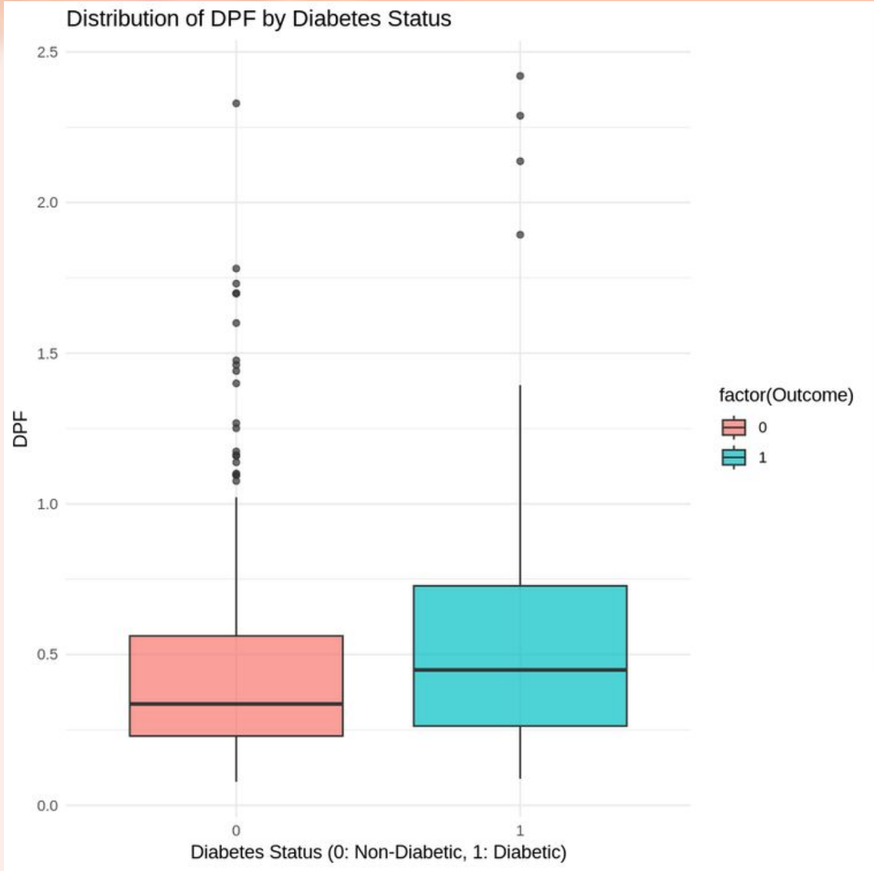


Using pie chart to represent the percentage among patients (Diabetic and Non-Diabetic)

Using Boxplot chart to represent median of

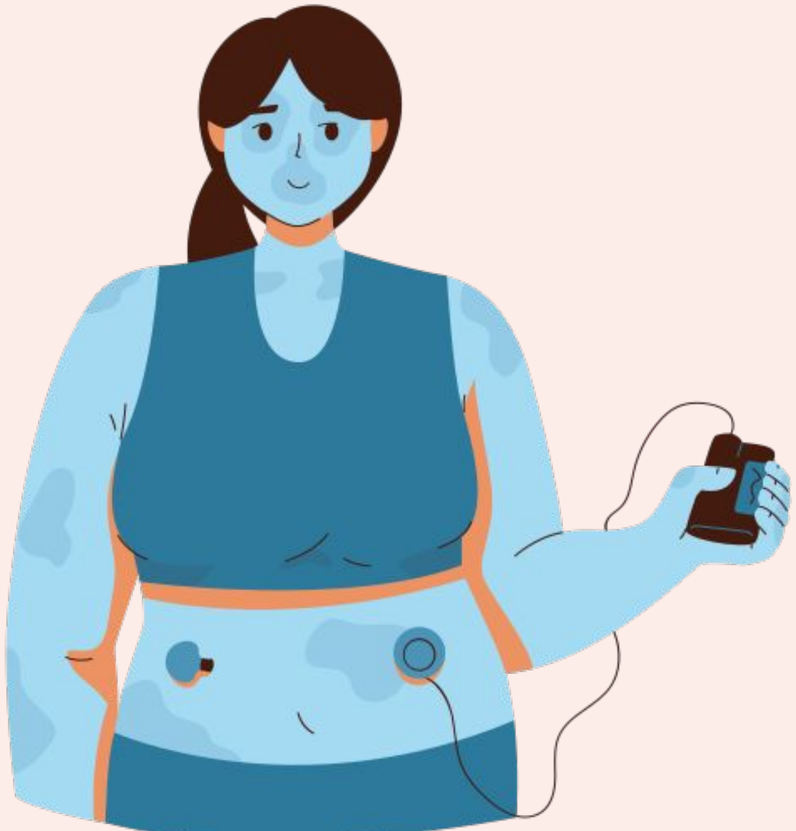


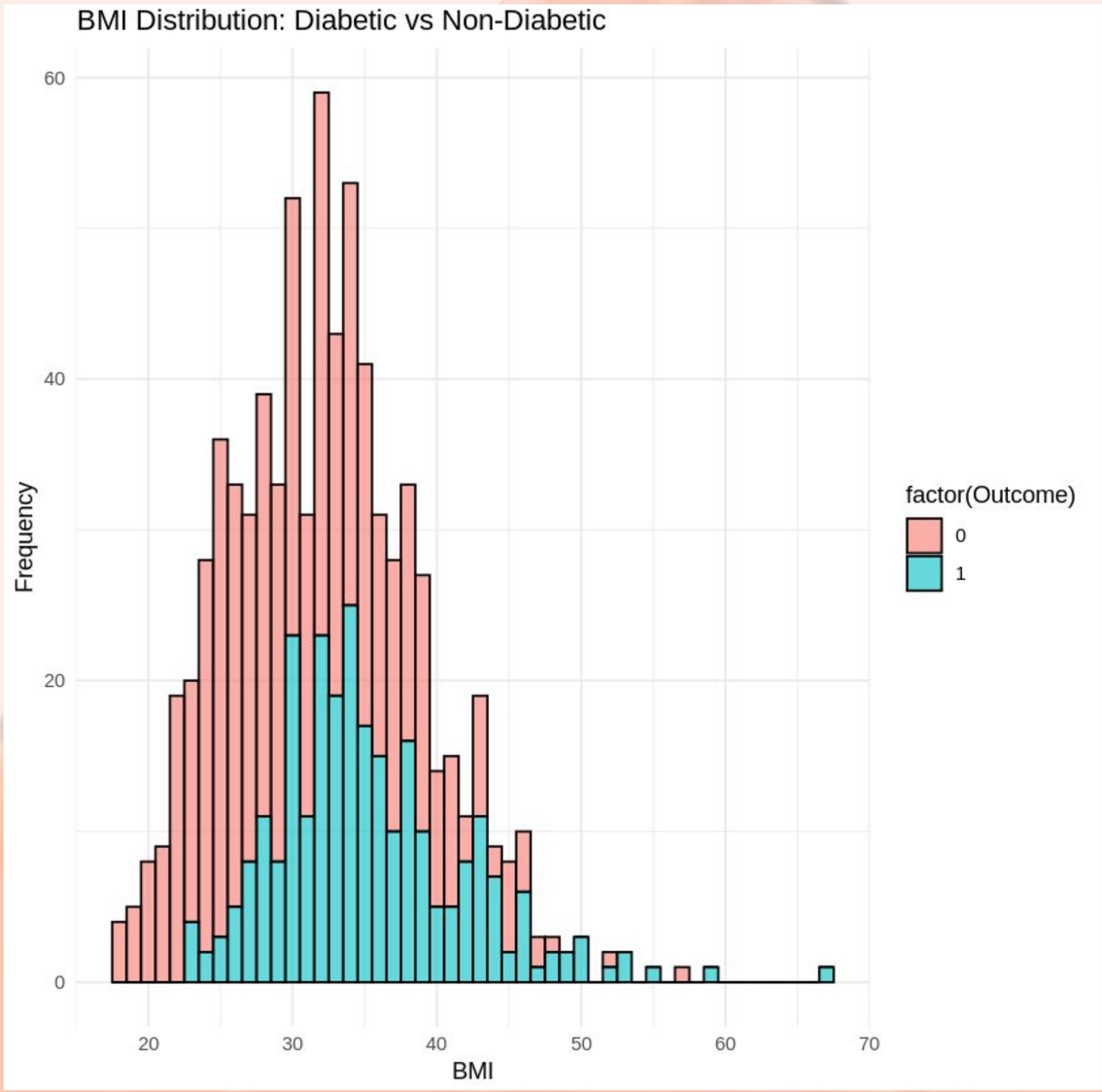
blood pressure among patients (Diabetic and Non-Diabetic)



DPF values for (diabetic and non-diabetic)

Using Histogram to visualize BMI Distribution





HYPOTHESIS TESTING

Claim 1: Significant difference in glucose levels

between diabetic and non-diabetic patients.

- H_0: $\mu_1 = \mu_2$ (Mean glucose level of diabetic = non-diabetic)
- H_a: $\mu_1 \neq \mu_2$ (Mean glucose level of diabetic ≠ non-diabetic)

Results: Reject H_0 → Significant difference confirmed.

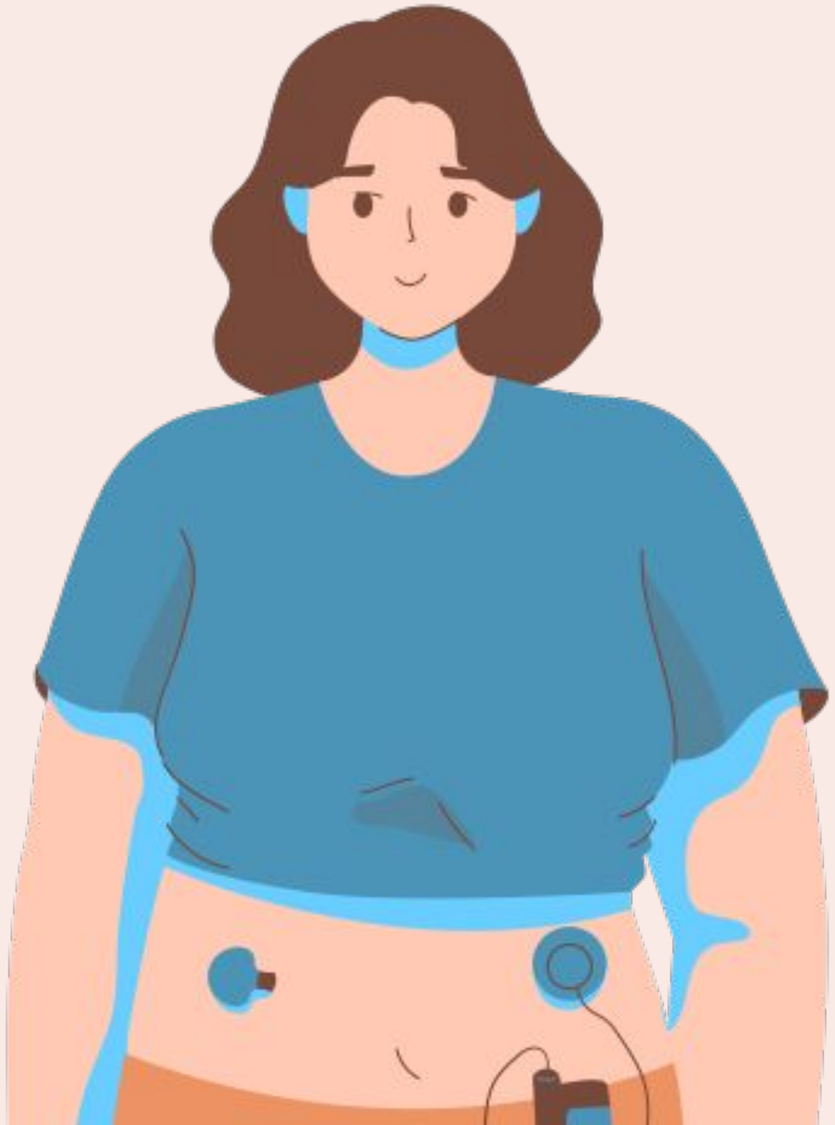
HYPOTHESIS TESTING

Claim2: Significant difference in BMI between diabetic and non-diabetic patients.

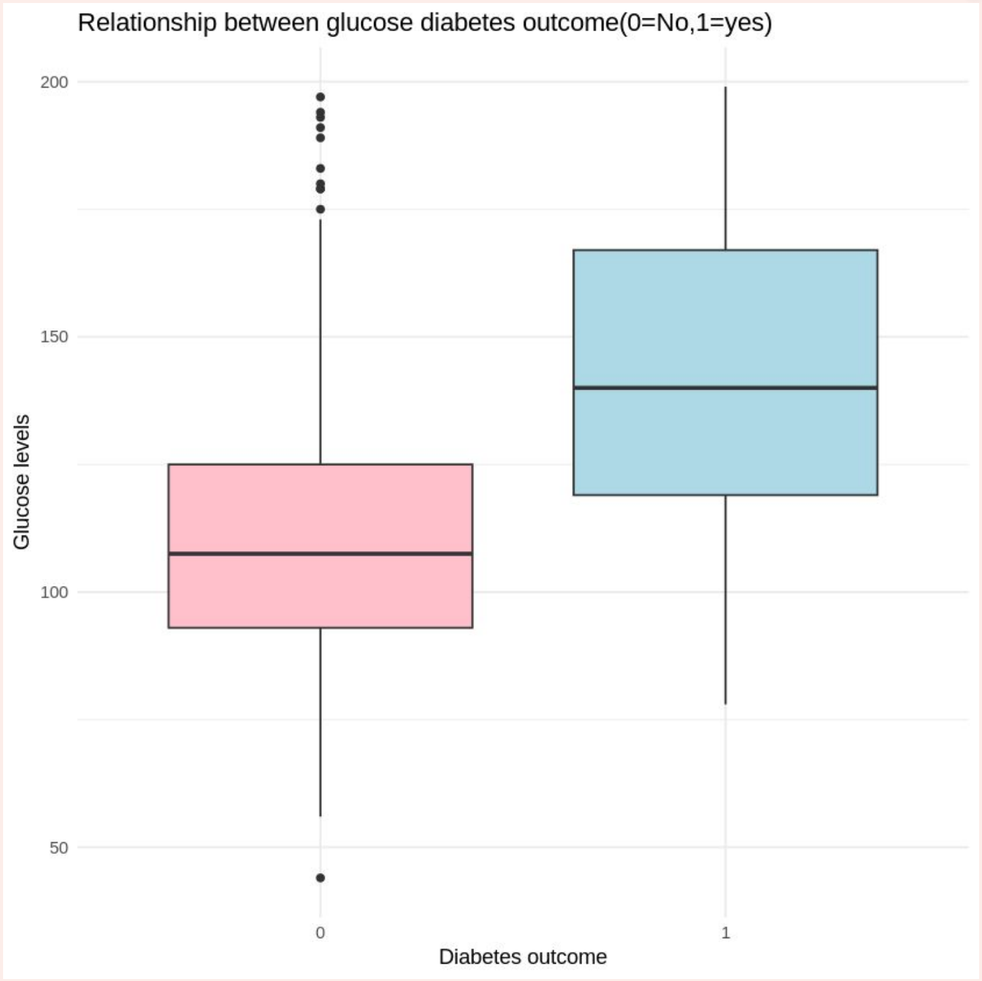
- H_0: $\mu_1 = \mu_2$ (Mean BMI of diabetic = non-diabetic)
- H_a: $\mu_1 \neq \mu_2$ (Mean BMI of diabetic ≠ non-diabetic)

Decision: Reject H_0 → Significant difference confirmed.

.



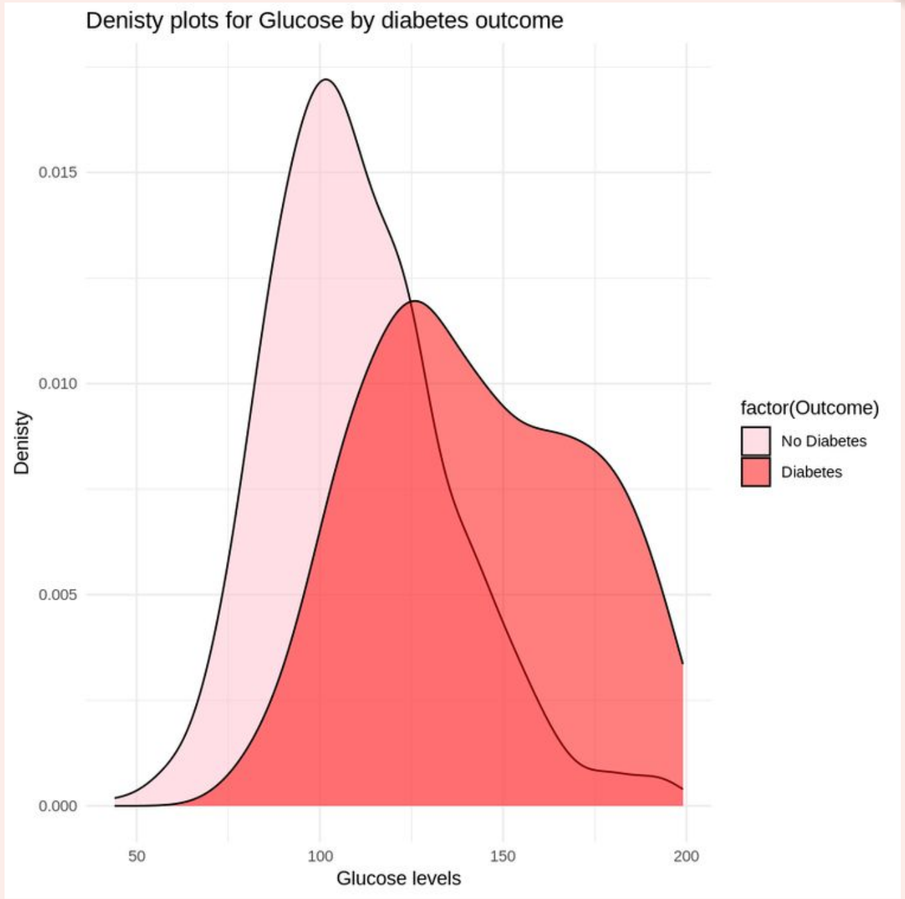
KEY FINDINGS AND INSIGHTS



Boxplot Results

- Diabetic patients had significantly higher glucose levels compared to non-diabetic patients.

- The Interquartile Range (IQR) for the diabetic group was notably wide, showing that most glucose measurements were very high among diabetics.

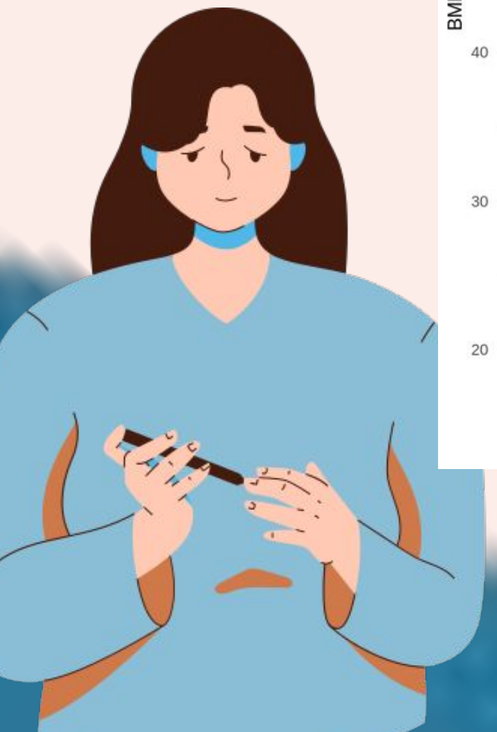


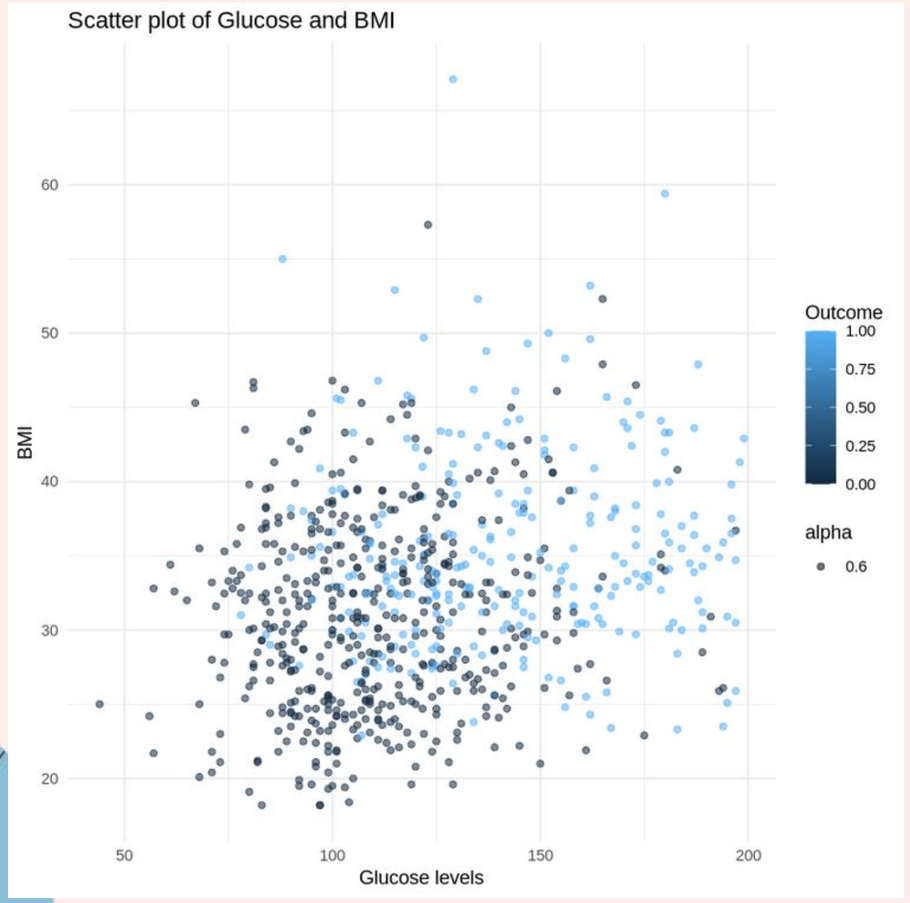
DENSITY PLOT FINDINGS

- The density curve for diabetic patients is concentrated at higher glucose levels than for non-diabetic patients.

- The overlap between the two distributions highlights that glucose level is a critical factor in distinguishing diabetics from non-diabetics.

- The correlation between glucose levels and diabetes outcome is 0.4666, indicating a moderate positive relationship between high glucose levels and the likelihood of having diabetes.

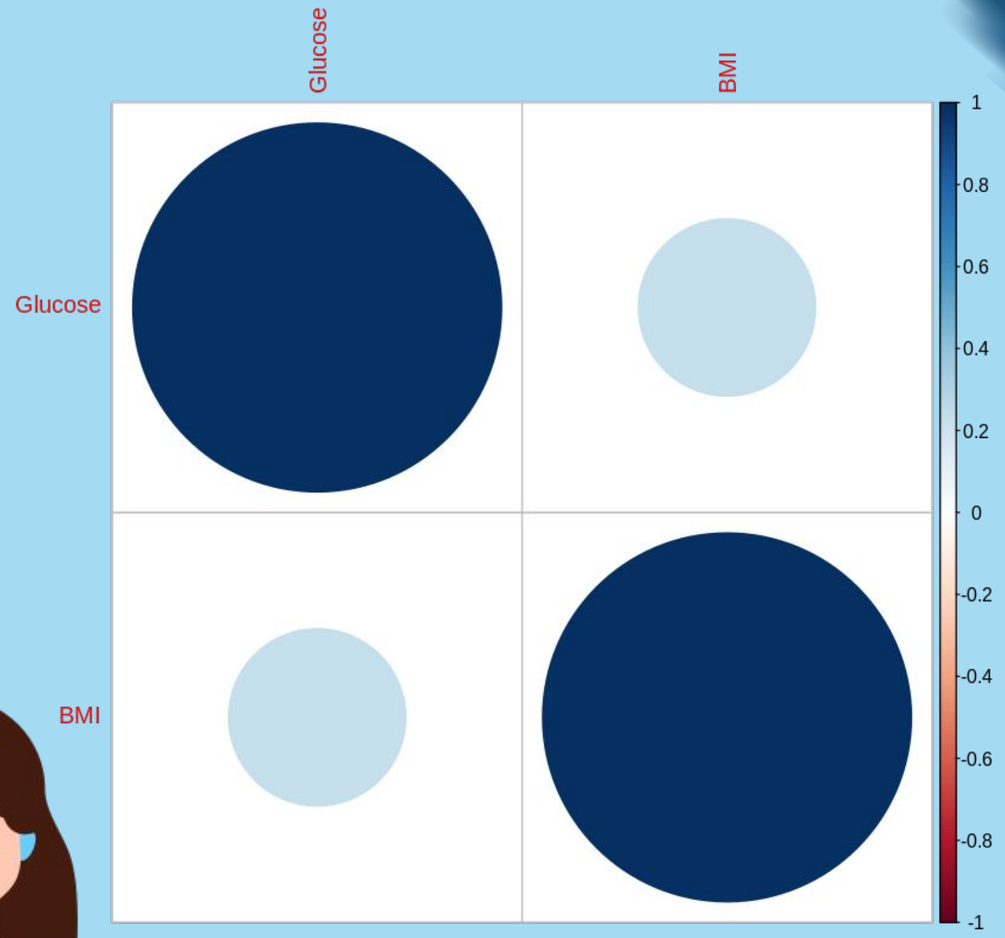




Scatter Plot Findings

- The scatter plot shows a positive relationship between glucose levels (x-axis) and weight (y-axis).

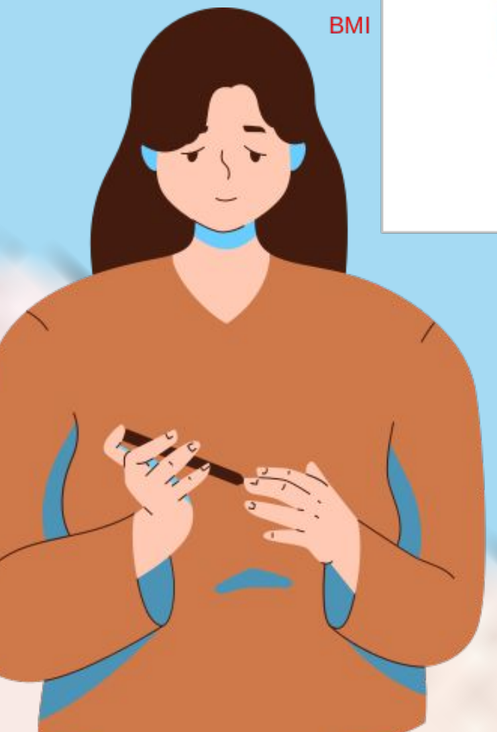
- As glucose levels increase, weight may also increase, but the correlation is weak.

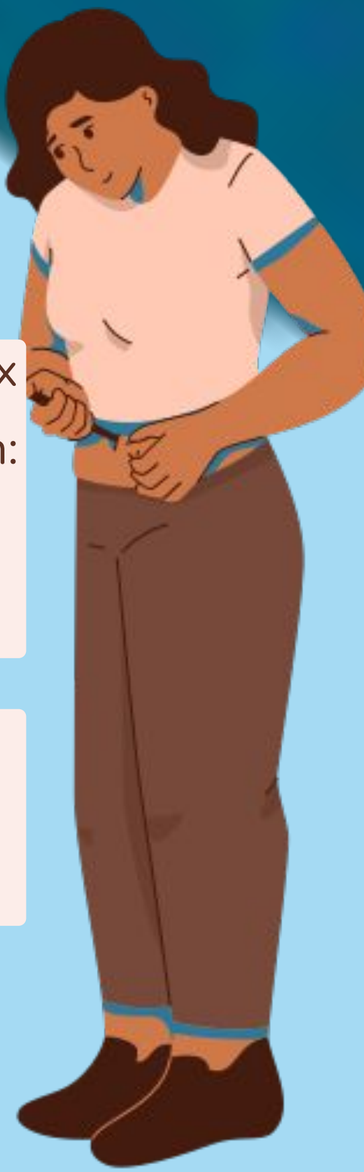


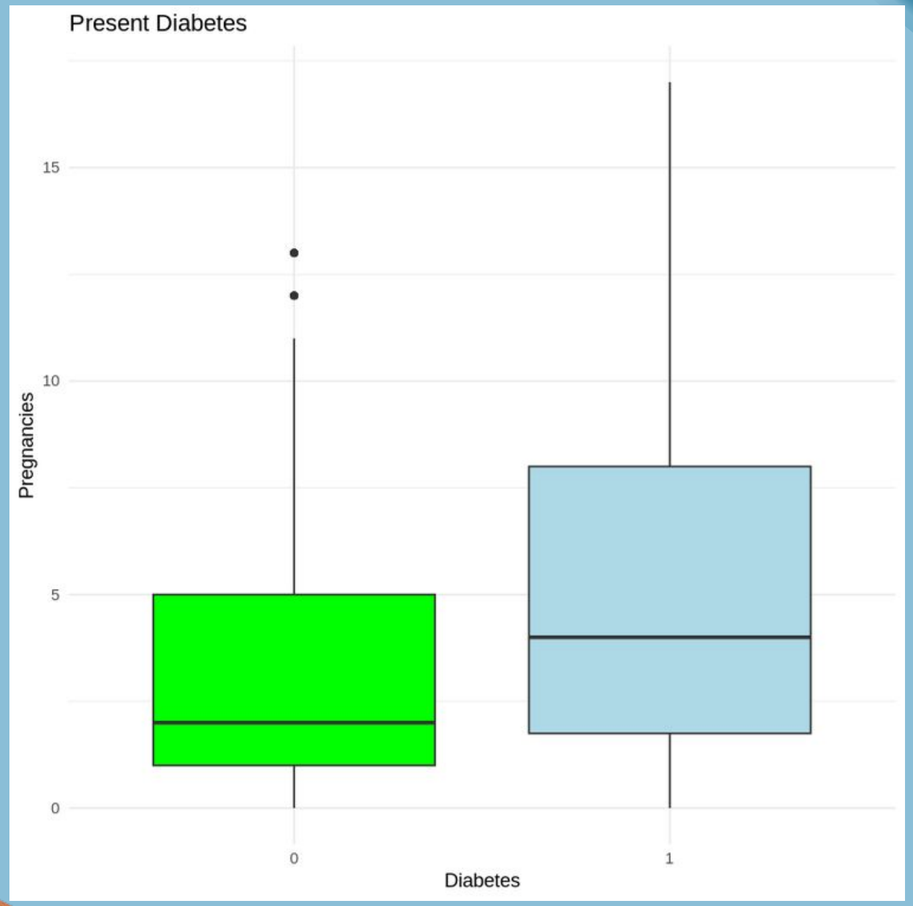
Correlation Matrix Findings

- The blue circles in the correlation matrix highlight a positive correlation between:
  Weight and BMI.
  Glucose levels and BMI.

- These relationships suggest an association but not a strong one.

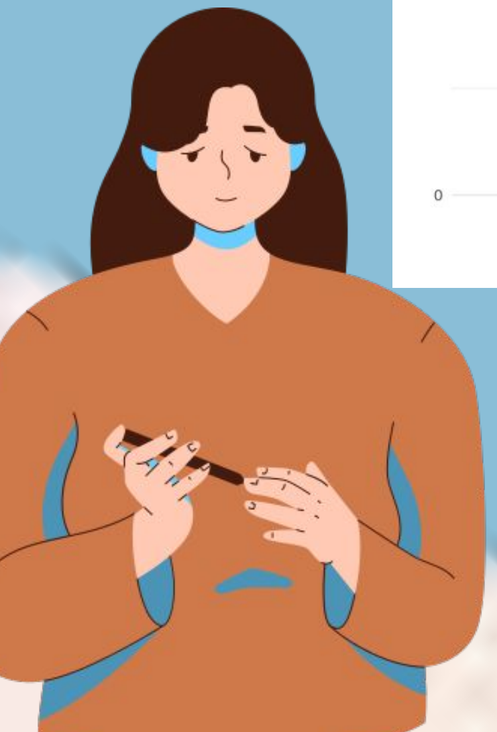


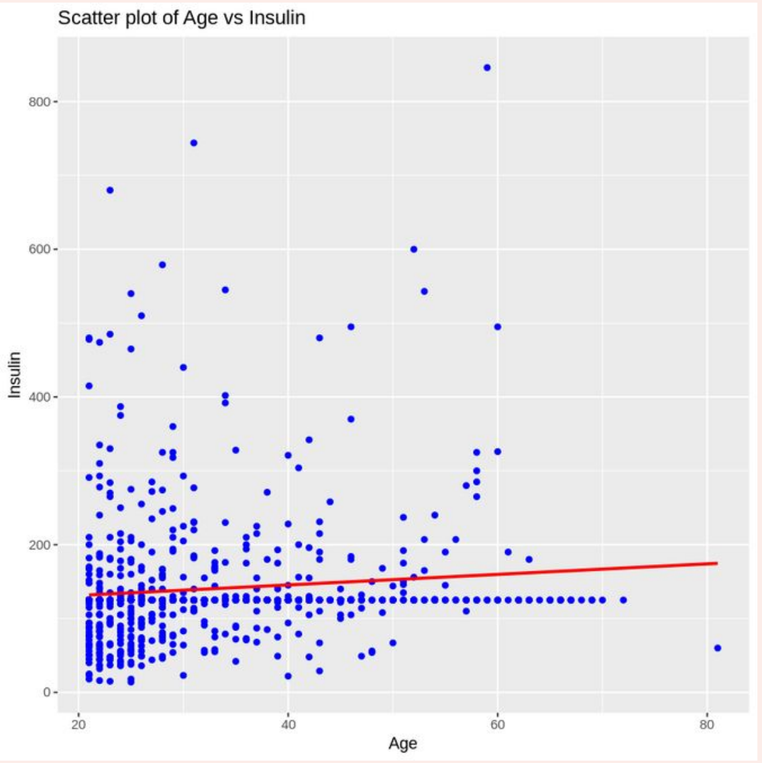




Boxplot Findings

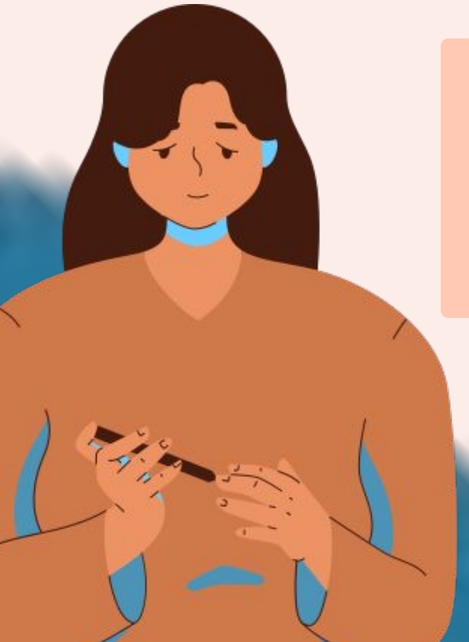
- Patients with diabetes tend to have a higher number of pregnancies.
- There is a greater disparity in the number of pregnancies among diabetic patients compared to non-diabetic patients





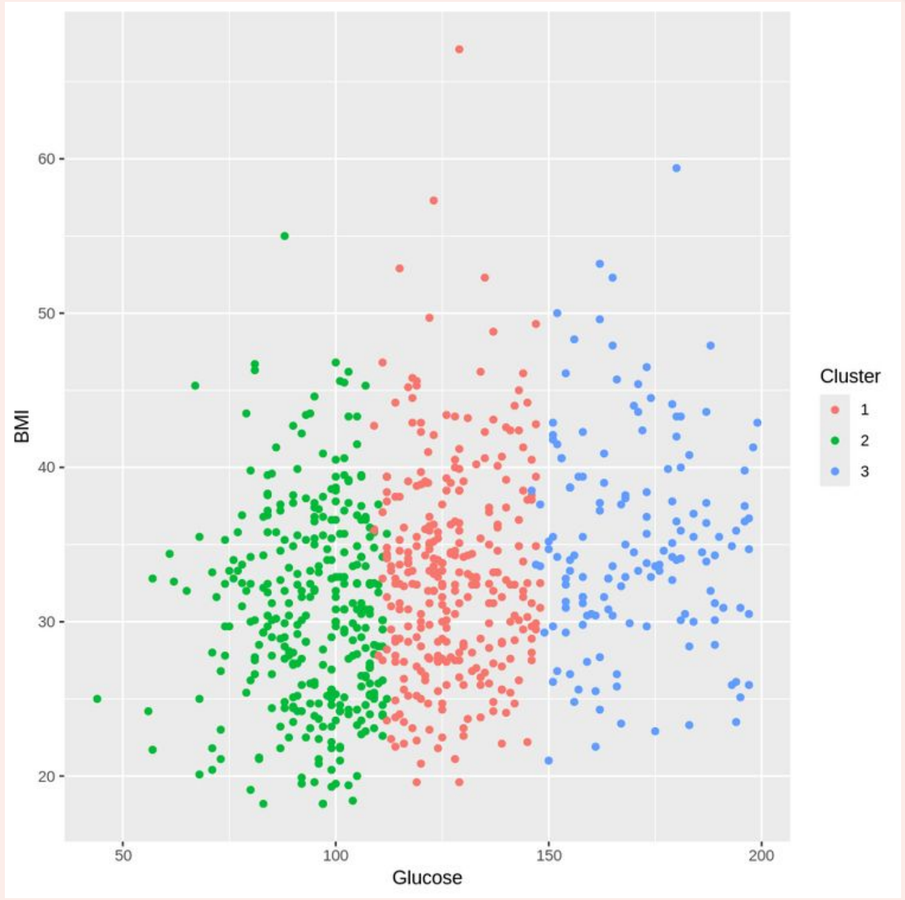
Boxplot Findings

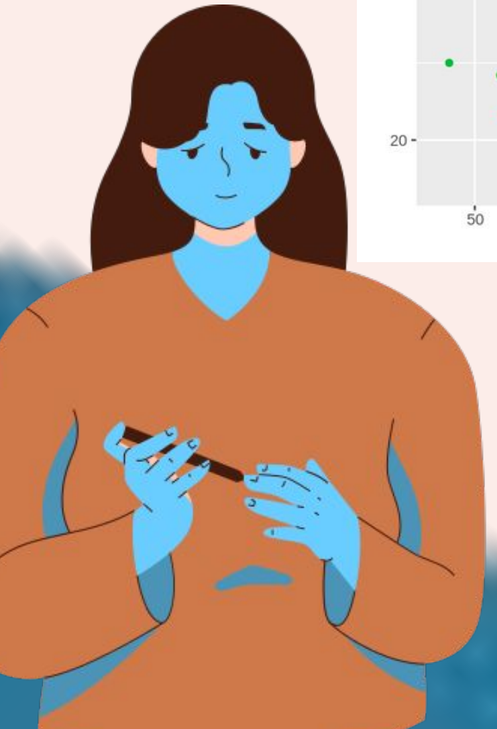
- Older patients might have higher glucose levels, but age is not a significant factor affecting insulin levels.



- A weak negative correlation exists between age and insulin levels (-0.0422), indicating minimal impact of age on insulin levels.

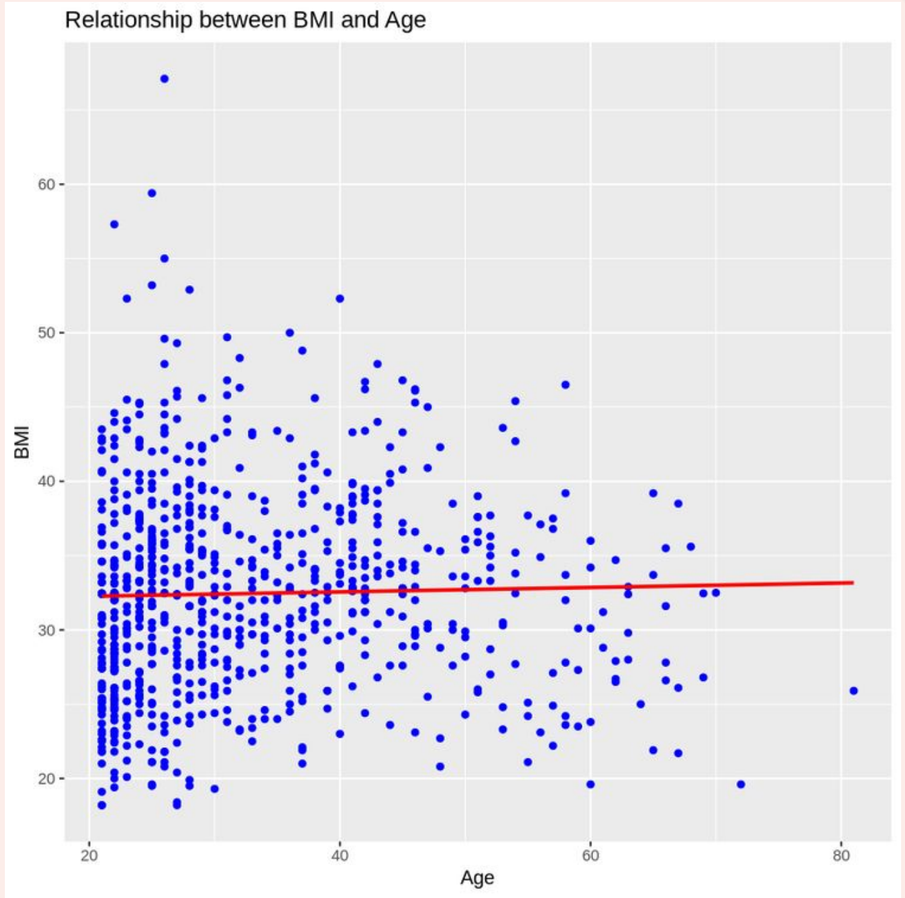
- A weak positive correlation exists between age and glucose levels (0.2635), suggesting glucose levels may slightly increase with age.





Cluster Analysis Findings

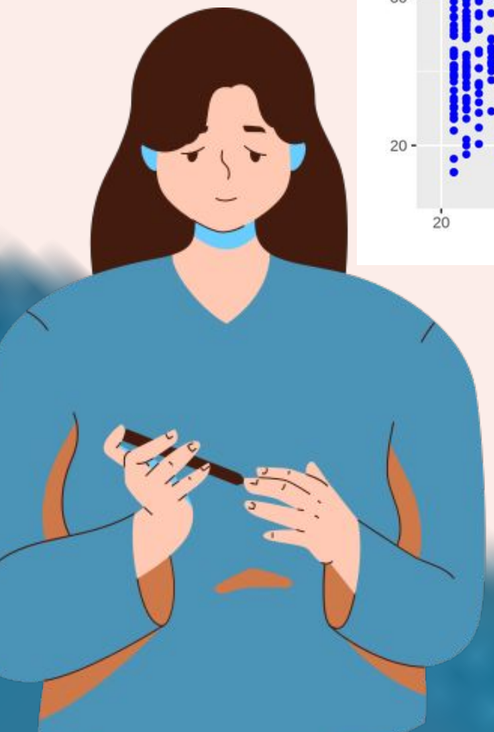
- Cluster analysis grouped patients into three categories based on glucose and BMI levels: low (low glucose and BMI), moderate (medium glucose and high BMI), and high risk (high glucose and BMI).
- High glucose and BMI levels strongly correlate with a higher risk of diabetes, emphasizing their importance in risk assessment.

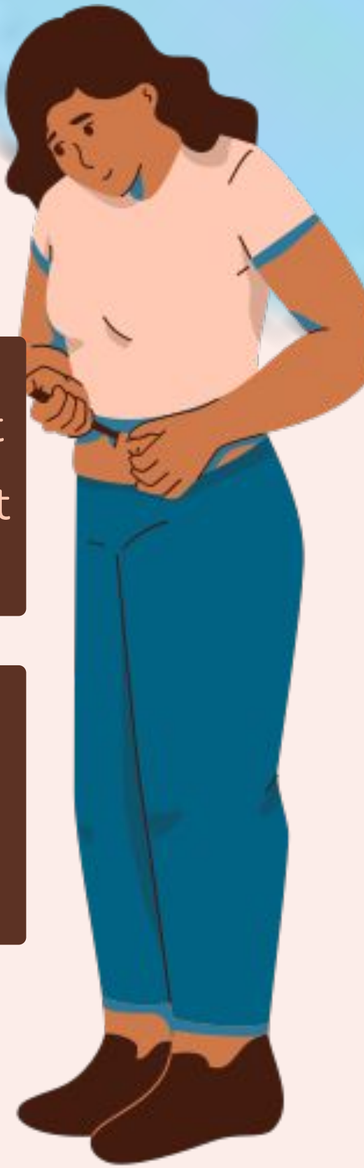


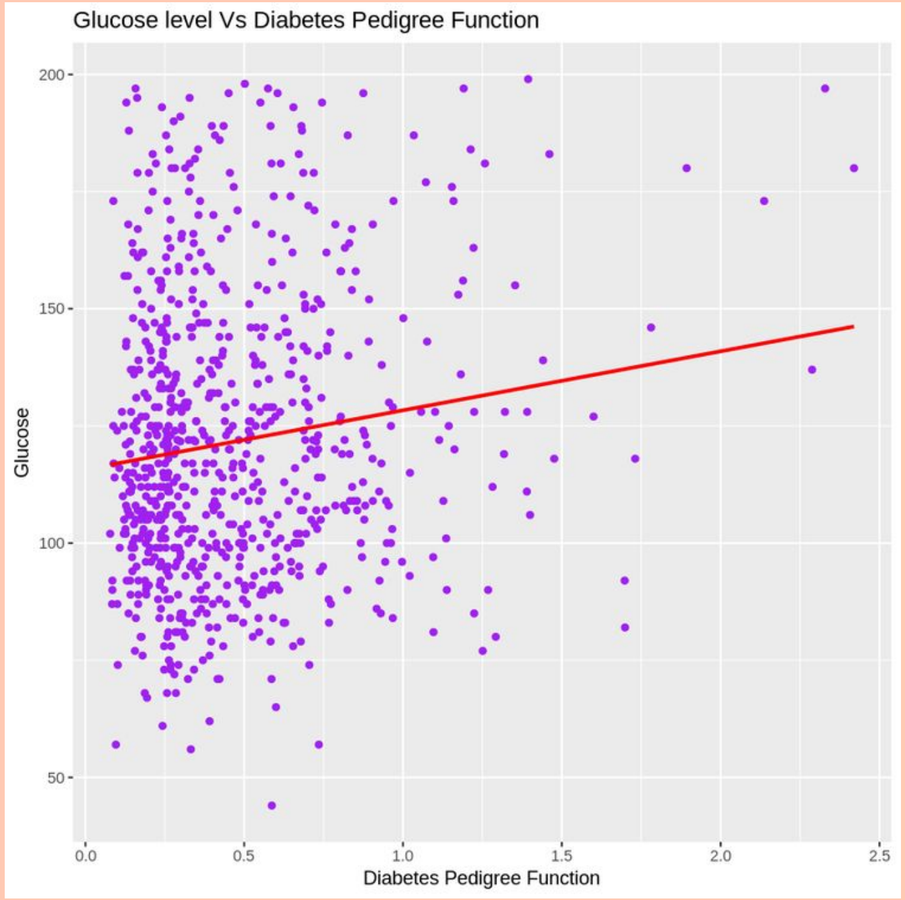
Age and BMI Relationship

- A weak positive correlation exists between age and BMI, suggesting that BMI may slightly increase with age, but the relationship is not strong.

- Age is not a primary factor driving BMI changes, as the regression line shows only a slight increase in BMI with age.

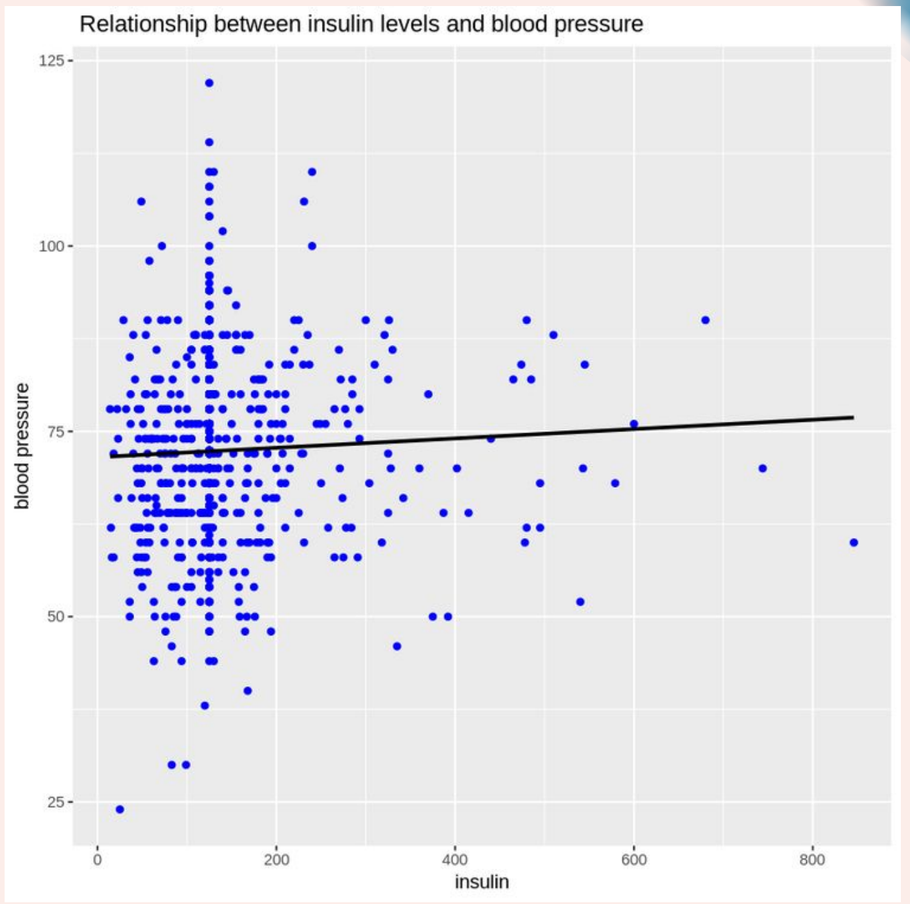






Family History and Glucose Levels

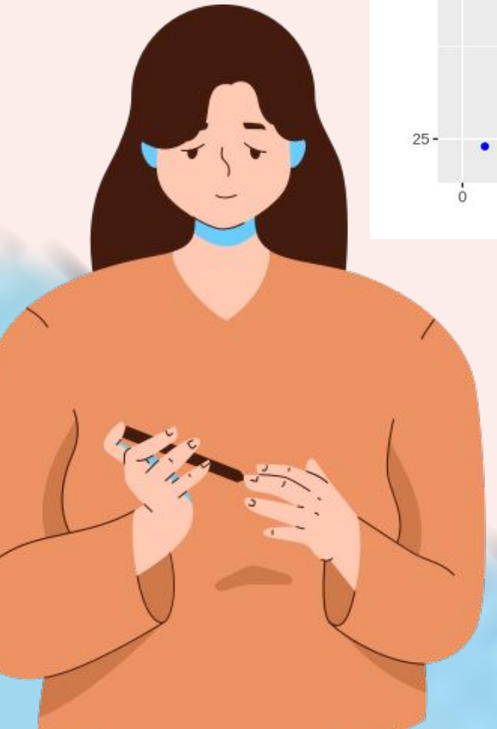
- Family history shows a slight positive influence on glucose levels, but the relationship is weak, as indicated by the wide spread around the regression line.
- While family history may affect glucose levels, other factors likely play a more significant role in influencing glucose levels.

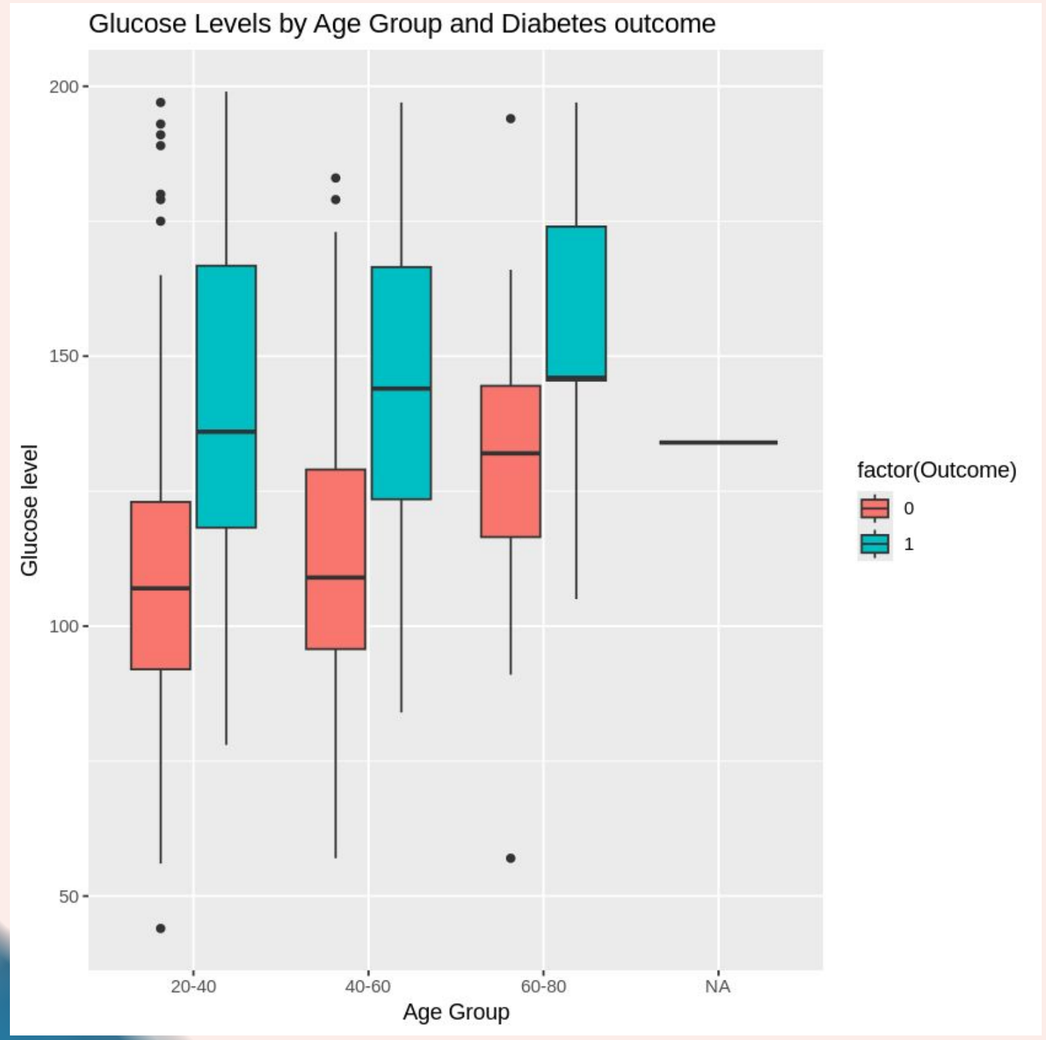


Insulin and Blood Pressure Relationship

- A slight positive correlation exists, indicating that blood pressure may increase slightly with higher insulin levels.

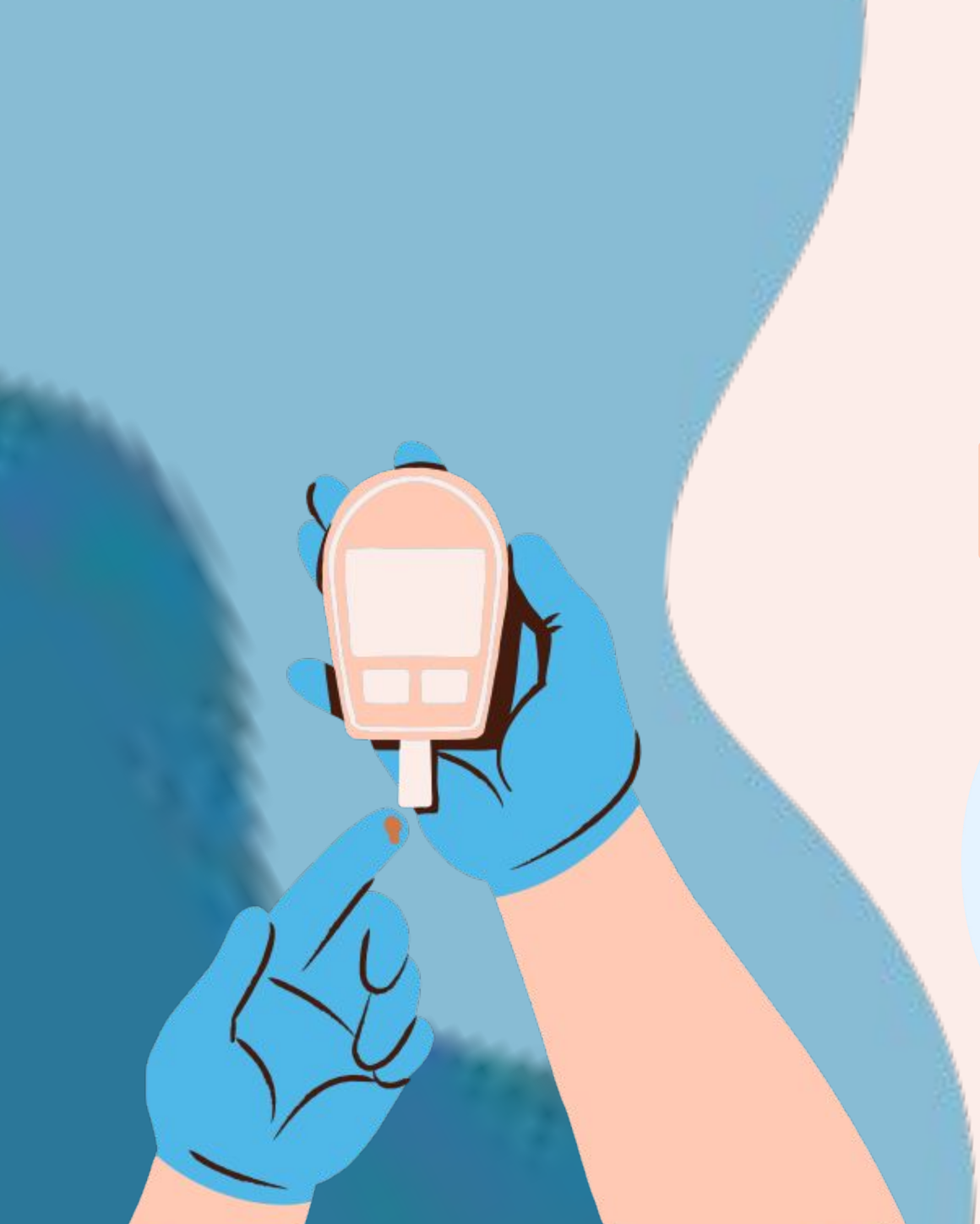
- While there is a connection, the relationship between insulin levels and blood pressure is minimal and not strongly indicative.





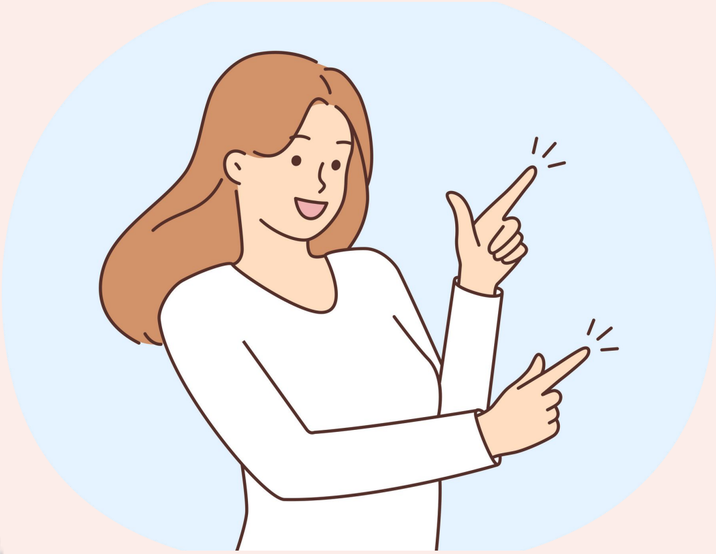
Age Groups and Glucose Levels

- Across all age groups (20-40, 40-60, 60-80), patients with diabetes consistently have higher glucose levels than non-diabetic patients, with the distribution becoming more widespread in older age groups
- High glucose levels are more prevalent with increasing age, particularly among diabetic individuals, highlighting the combined impact of age and diabetes on glucose levels.



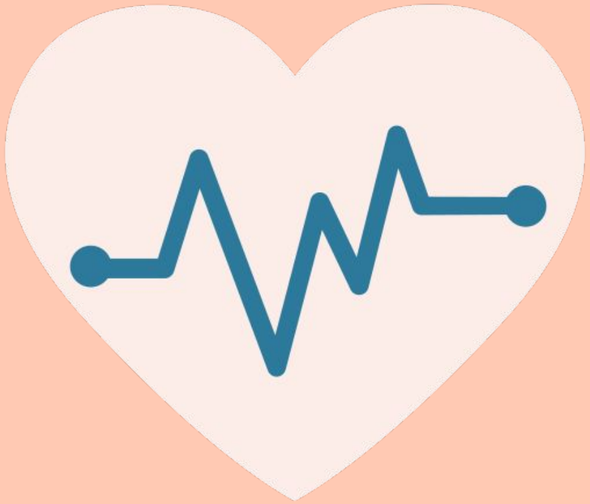
CONCLUSION AND RECOMMENDATIONS

Recommendations



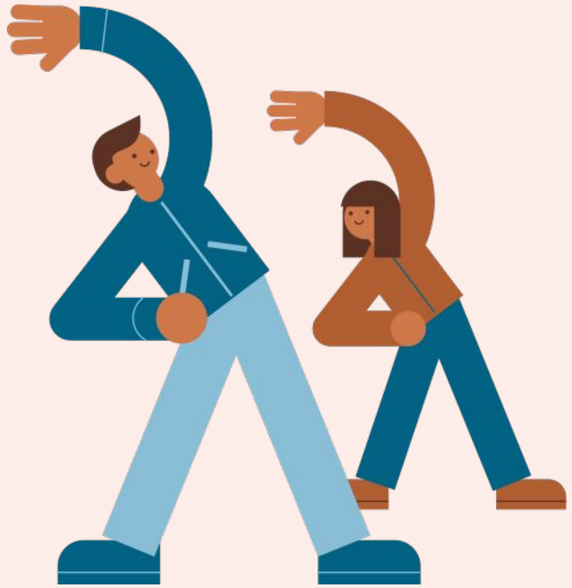
REGULAR HEALTH CHECKS

Check glucose and BMI regularly, especially for older people or those with a family history of diabetes.



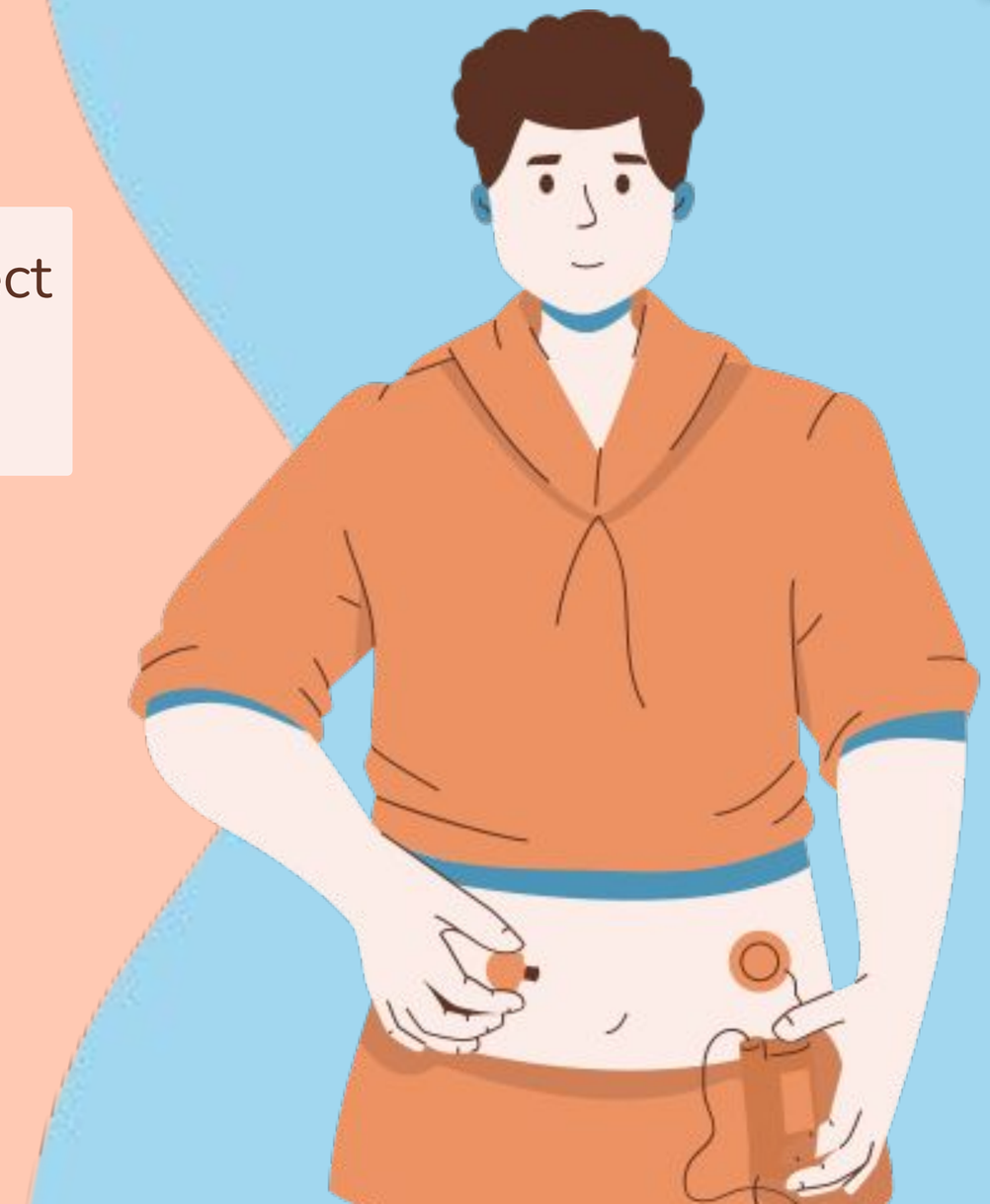
HEALTHY LIFESTYLE

Encourage exercise and healthy eating to manage weight and reduce the risk of diabetes.



PUBLIC AWARENESS

Teach people how age, glucose, and BMI affect the chances of getting diabetes.



CONCLUSION

We found through data analysis for the Diabetes Dataset

- Diabetic patients have significantly higher glucose levels than non-diabetic patients.
- BMI and age impact diabetes risk, but the relationships are not strong enough to rely on alone.
- Family history slightly affects glucose levels, but other factors play a more significant role.
- Glucose levels increase with age among diabetic patients, but age is not the main determining factor.

These findings highlight the importance of analyzing multiple factors to better assess and prevent diabetes risk.

TEAM MEMBERS

1. Nourhan Deif Sayed
2. Rahma Mourad
3. Rahma Ibrahim
4. Gehad Mahmoud
5. Habiba Moahmed