Predicting the health risk factors for pregnant patients

Objectives:

1. To identify relationship between different health factors like age, systolic blood pressure, diastolic blood pressure, blood sugar level, body temperature, and heart rate with pregnancy risk level.

2. To identify any correlation between health factors.

Data:

1.Data Collection:

- Kaggle website (Maternal Health Risk Data Set.csv file)

2.Data Preprocessing:

- Data cleaned and prepared for analysis

3.Analysis Techniques:

- Data visualization techniques like matplotlib, statistical methods, and simple linear regression.

Dataset had following information:

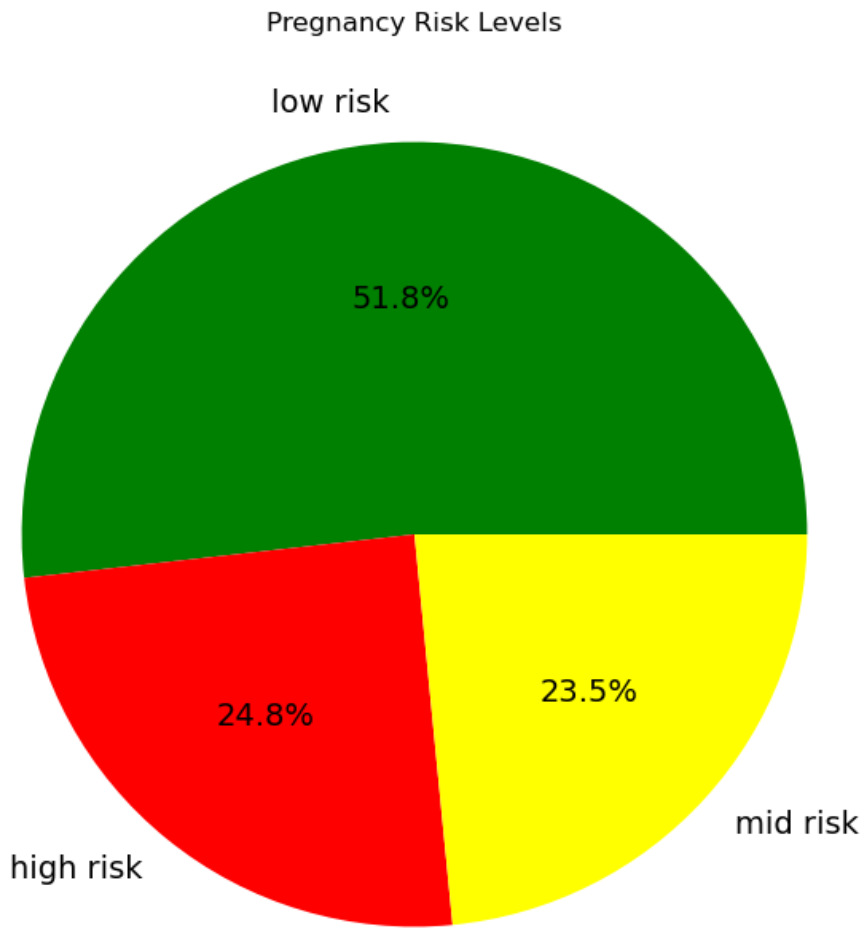
- **Age:** Age in years when a woman is pregnant.
- **SystolicBP**: Upper value of Blood Pressure in mmHg(Normal = < 120)
- **DiastolicBP**: Lower value of Blood Pressure in mmHg (Normal = < 80)
- **BS**: Blood glucose levels is in terms of a molar concentration, mmol/L. (Normal = < 7.8)
- **Heart Rate**: A normal resting heart rate in beats per minute.(Normal = 60 – 100)
- **Body Temperature**: A normal body temperature in Fahrenheit. (Normal = 97-99)
- **Risk Level**: Risk Intensity Level during pregnancy High risk, Mid risk and Low risk

Distribution of Risk Levels

The pie chart
51.8% - low risk

23.5% - mid risk

24.8% - high risk



Relationship between Risk level and Age

Group bar chart shows:

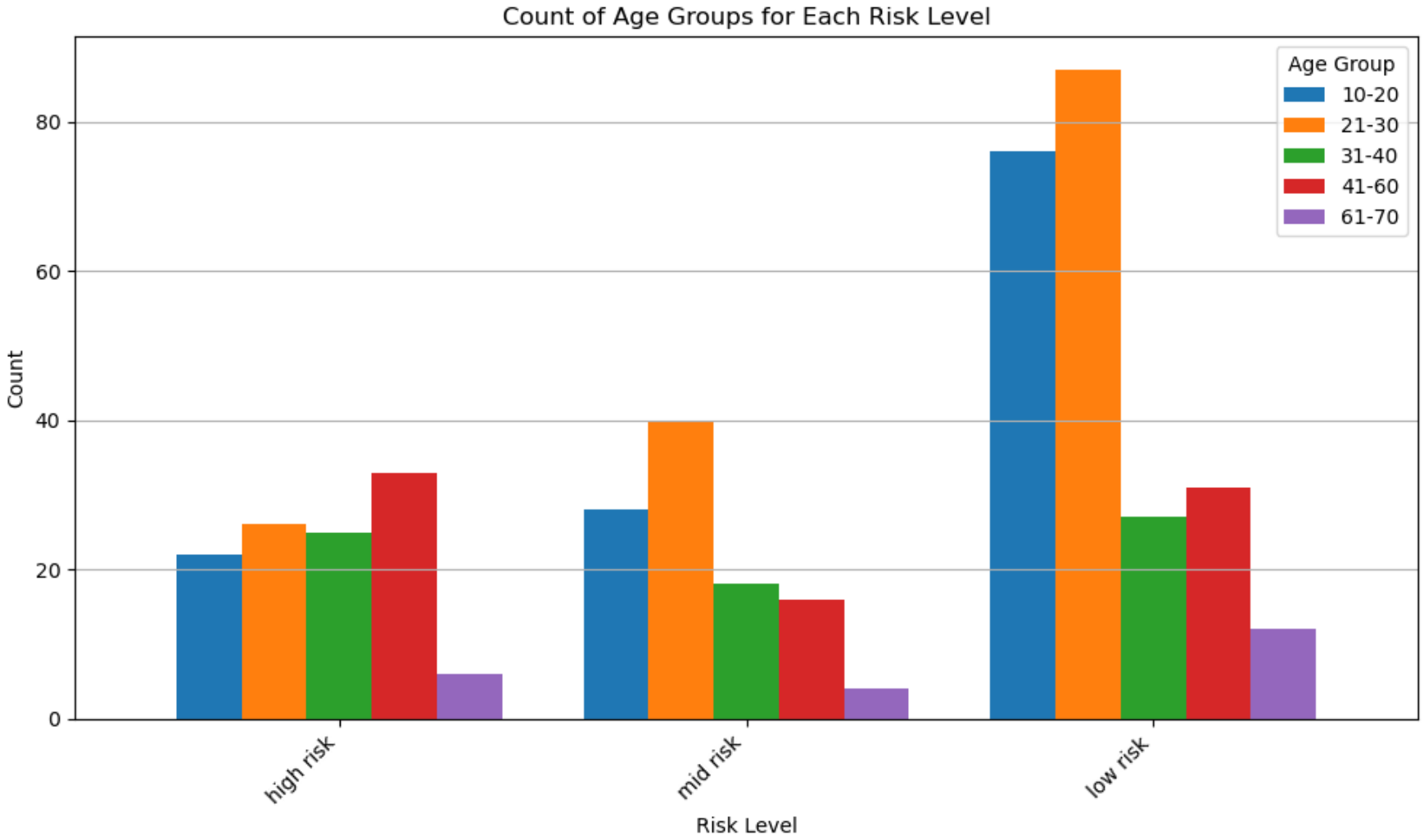
Low risk and mid risk women

- 10-20 years age group
- 21-30 years age group.

High risk pregnant women

- 31-40 years
- 41-60 years

As Age increases , risk level in pregnancy increases



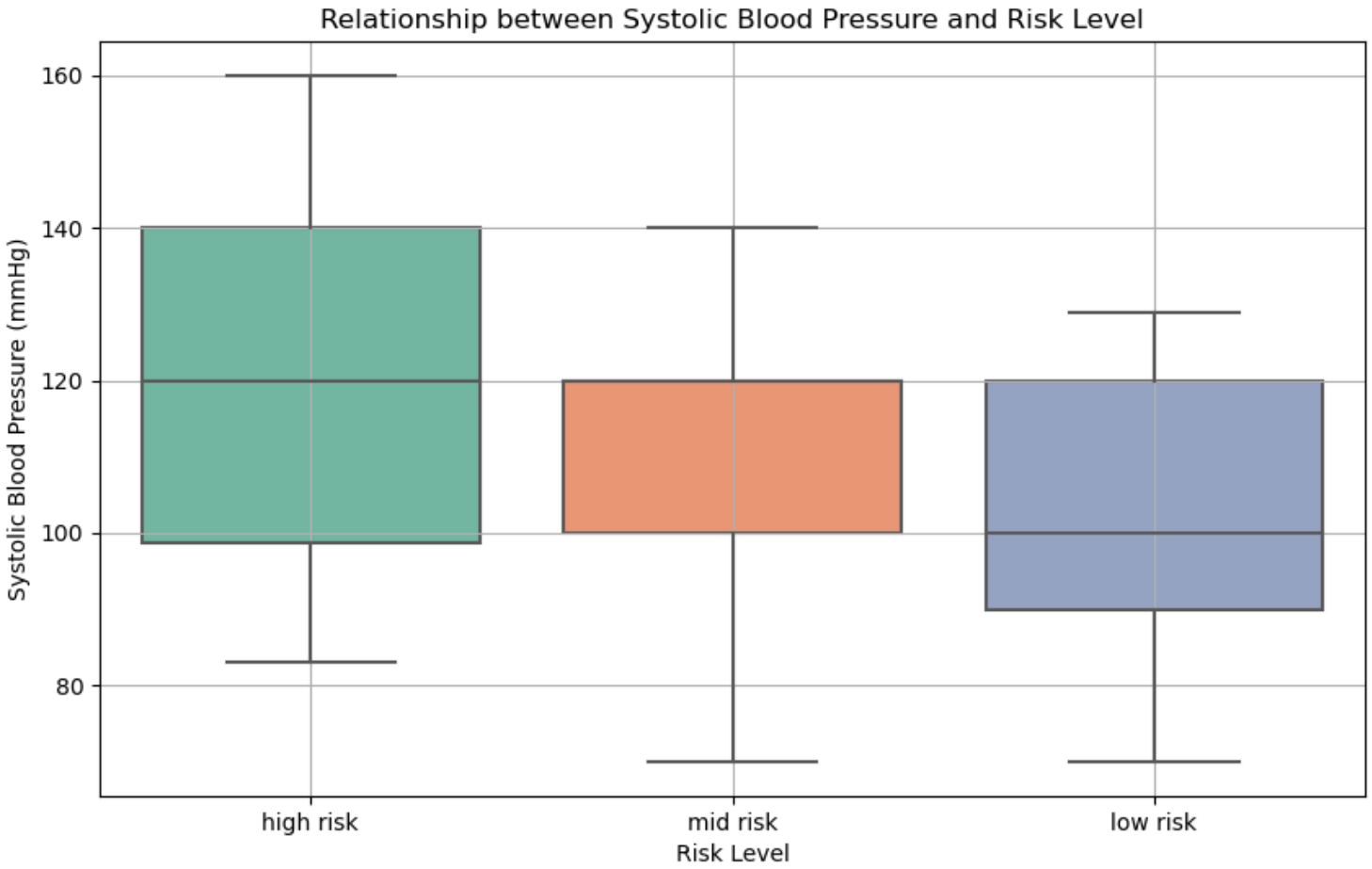
Relationship between Risk level and Systolic BP

Normal systolic blood pressure is <120

Boxplot shows:

- low risk females - 90-120mmHg
- Mid risk has BP -100-120mmHg
- High risk have - 100-140mmHg

As Systolic BP increase >120, risk level also increases.



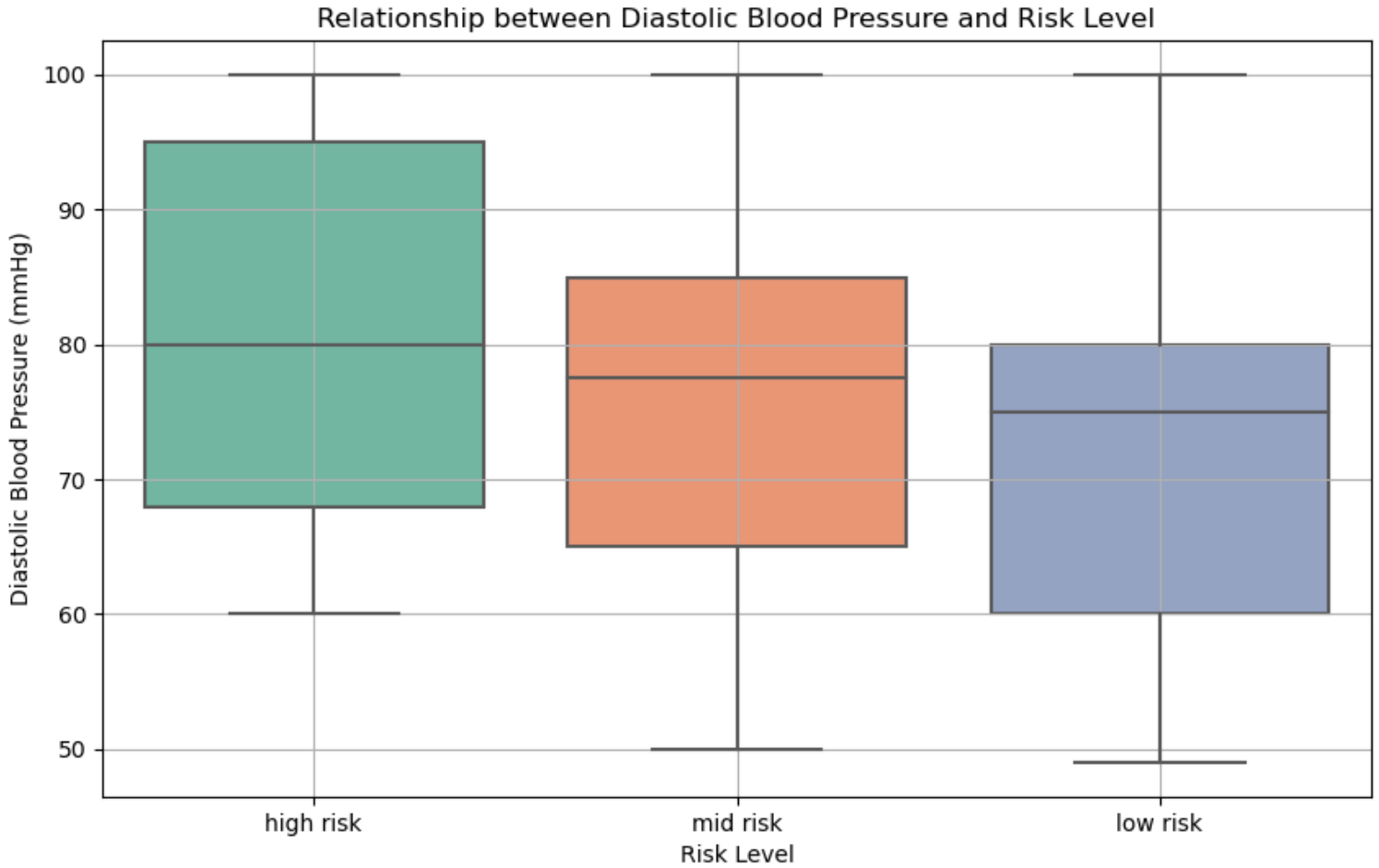
Relationship between Risk level and Diastolic BP

The normal diastolic blood pressure is <80 mm Hg.

Boxplot:

- Low risk - 60-80mmHg
- Mid risk - 65-85mmHg
- High risk - 68-95mmHg

As diastolic BP increase >80mmHg, risk level also increases.

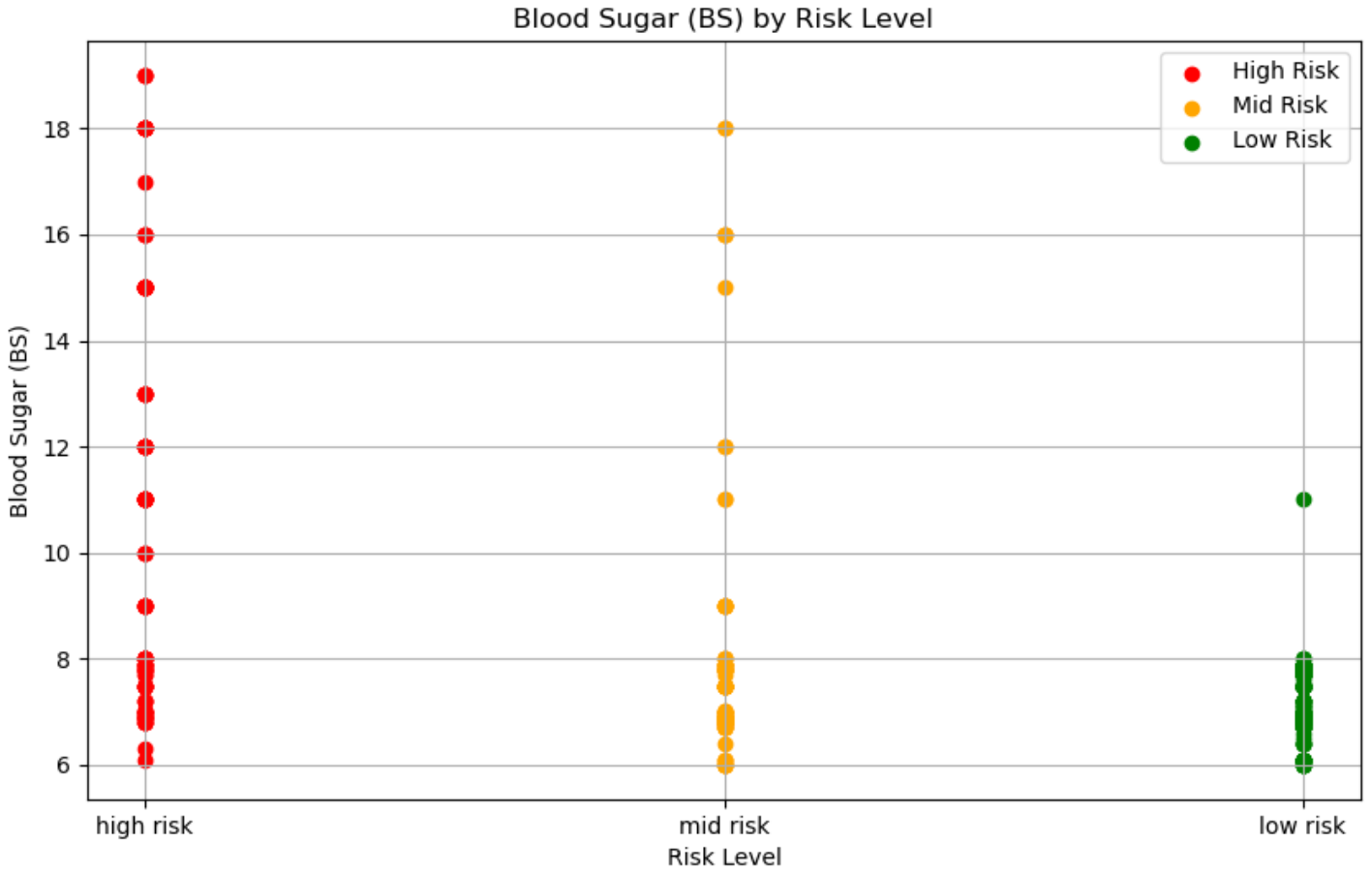


Relationship between Risk level and Blood Sugar

A blood sugar level less than (7.8 mmol/L) is normal.

Scatter plot:
Most of sugar level values are concentrated between 6 to 8 mmol/L for mid and low risk females.
High risk group shows values ranging from 6 to 19 mmol/L.

Blood sugar > 8mmol/L your risk level increases.

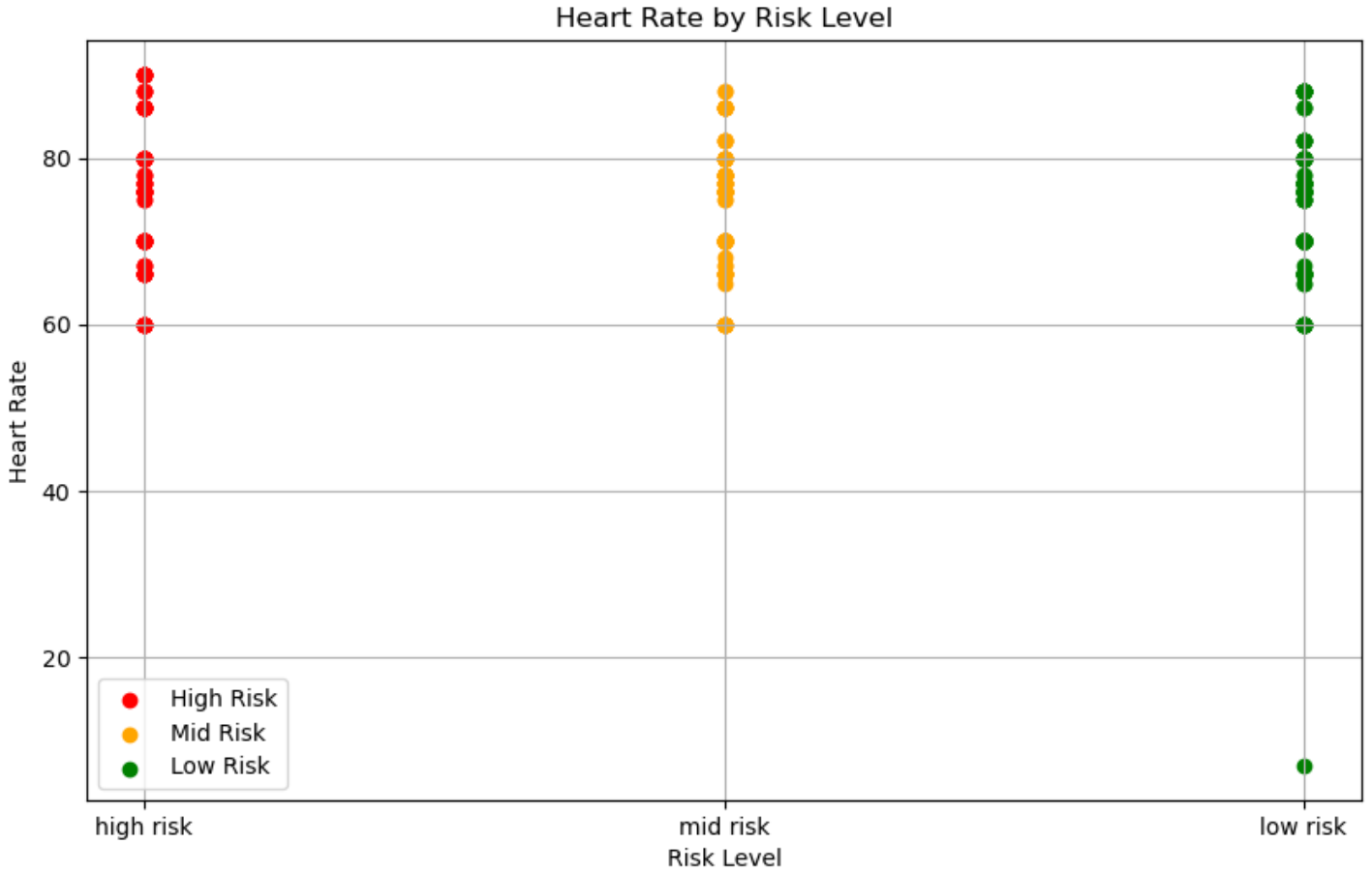


Relationship between Risk level and Heart Rate

A normal resting heart rate ranges from 60 to 100 beats per minute.

Heart rate for all risk levels falls between 60 to 90bpm, which is within normal range.

Heart rate does not relate with risk level .

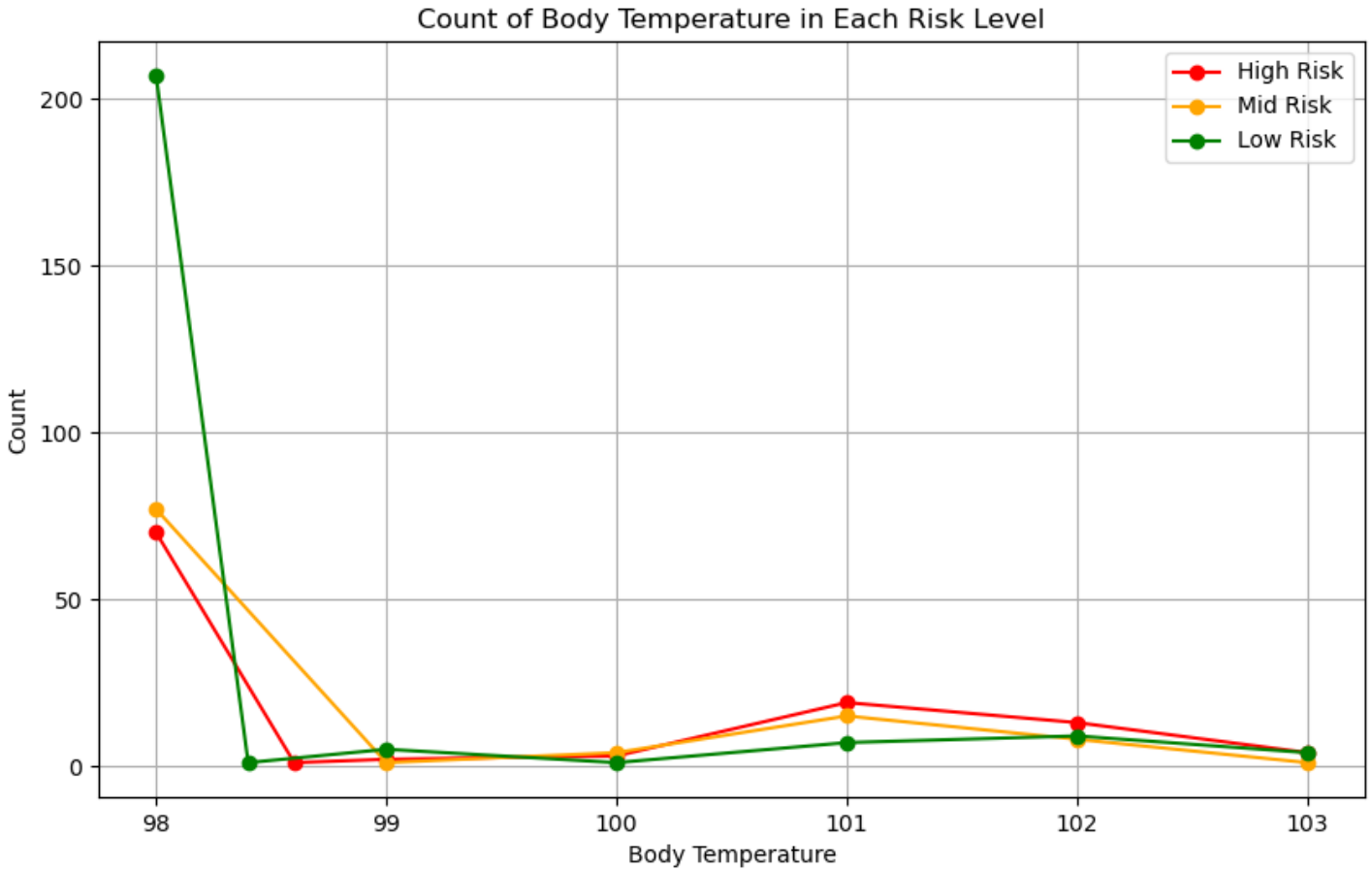


Relationship between Risk level and Body Temperature

Normal body temperature ranges between 97 F (36.1 C) and 99 F (37.2 C).

- All risk levels fall between normal body temperature group from 98 to 99F.

Body temperature does not relate with risk level.



Correlation coefficient

The Pearson correlation coefficient, often denoted as "r", is a measure of the linear relationship between two variables.

Interpretation:

1. If r = +1, it indicates a perfect positive linear relationship.

2. If r = -1, it indicates a perfect negative linear relationship.

3. If r = 0, it indicates no linear relationship between the variables.

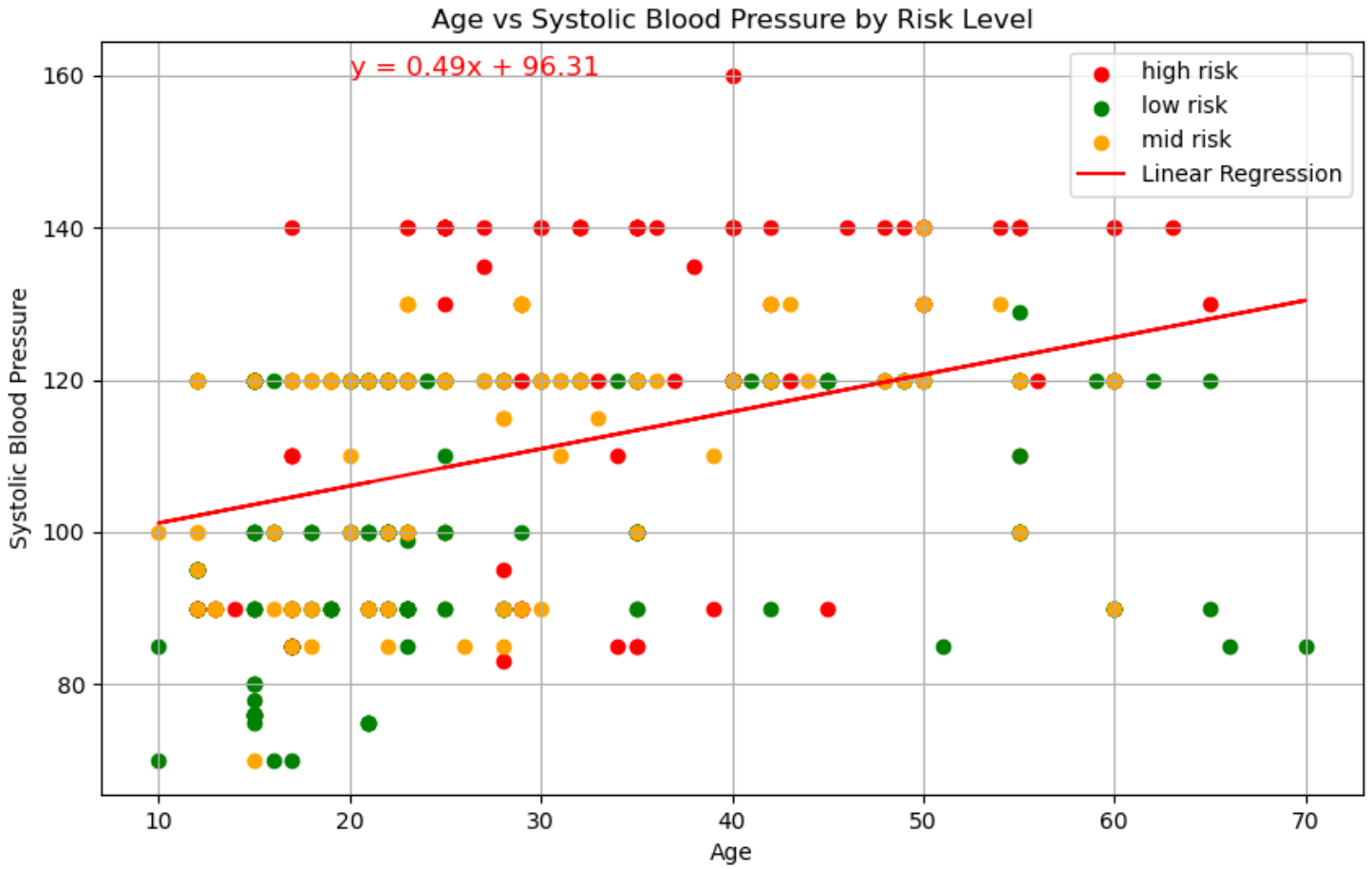
0.	-	0.19	very weak correlation
0.20	-	0.39	weak correlation
0.40	-	0.59	moderate correlation
0.60	-	0.79	strong correlation
0.80	-	1	very strong correlation

(BMJ /Publications/Statistics at square one/11. Correlation and regression)

Relationship between Age and Systolic BP

(r value) = 0.37
a weak positive linear relationship between the two variables.

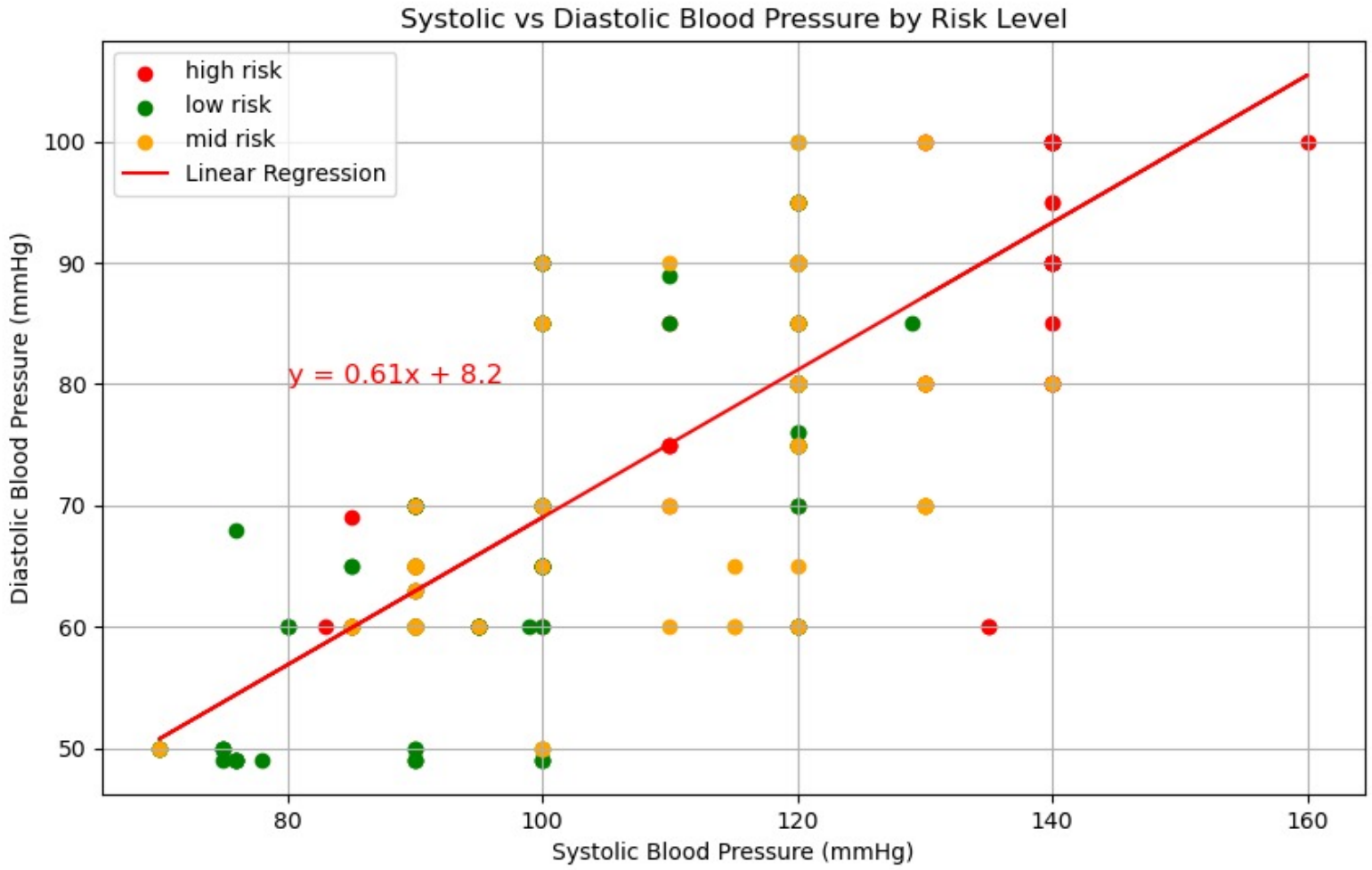
Systolic BP increases >120 the risk level increases.



Relationship between Diastolic BP and Systolic BP

(r value) = 0.79
a strong positive linear relationship between the two variables.

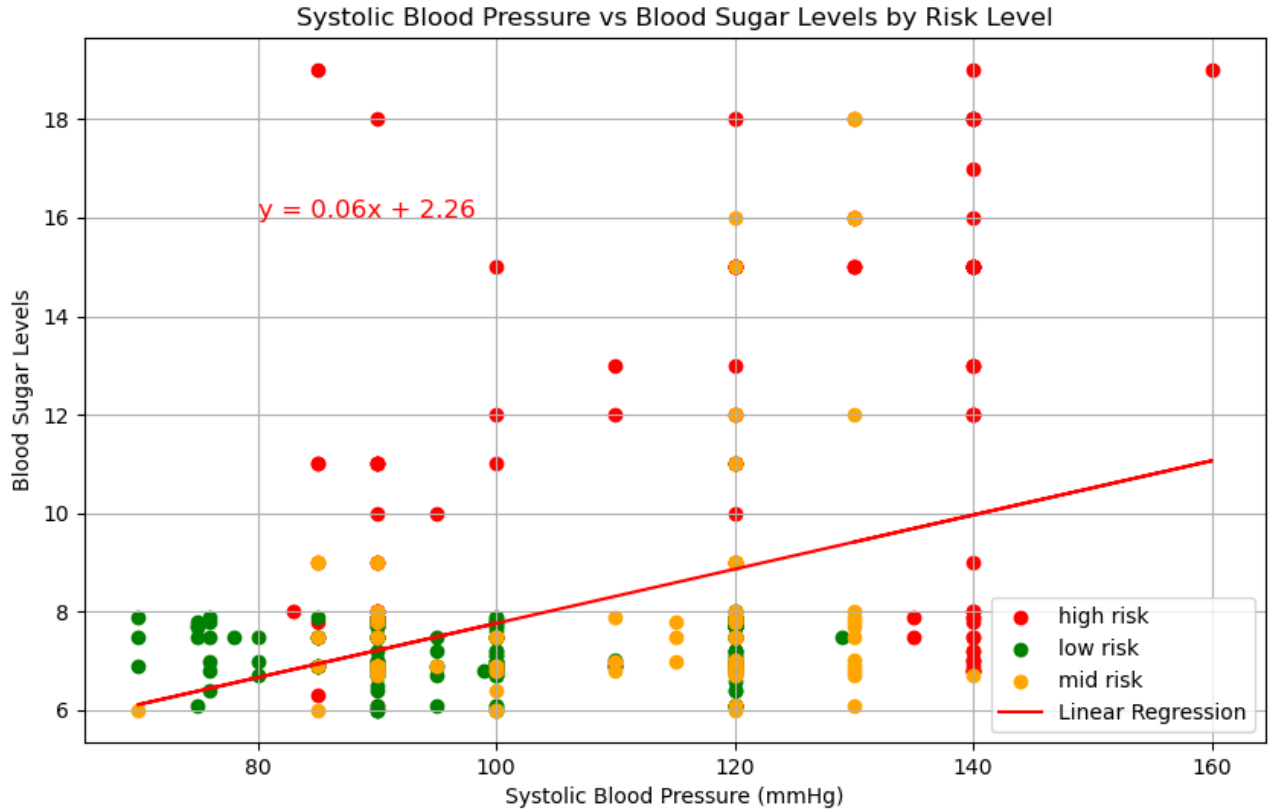
Diastolic BP >80
And
Systolic BP >120,
Risk level increases.



Relationship between Blood sugar and Systolic BP

(r value) = 0.34
a weak positive linear relationship between the two variables.

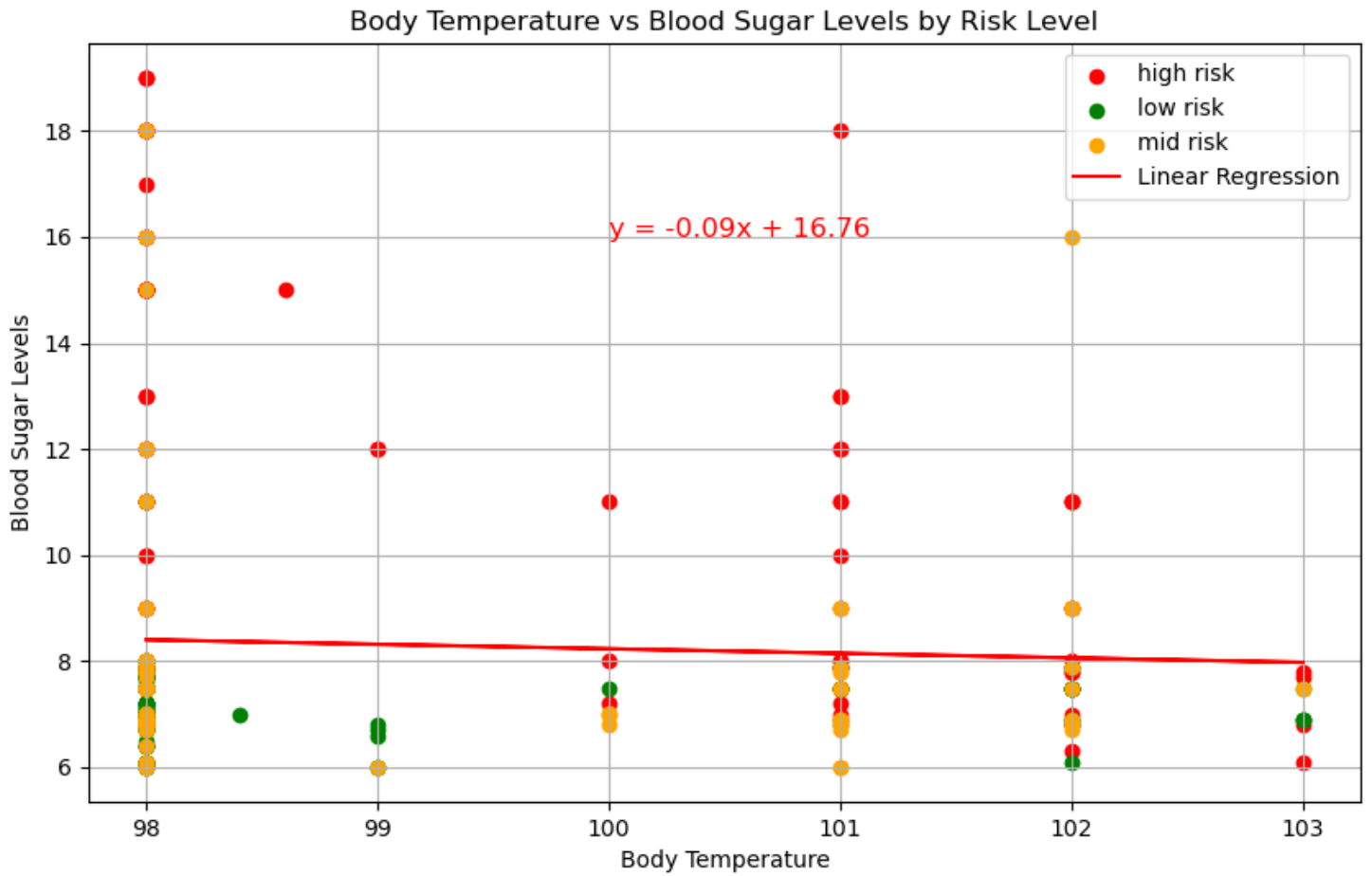
Systolic BP >120
Blood sugar >8
Risk level increases



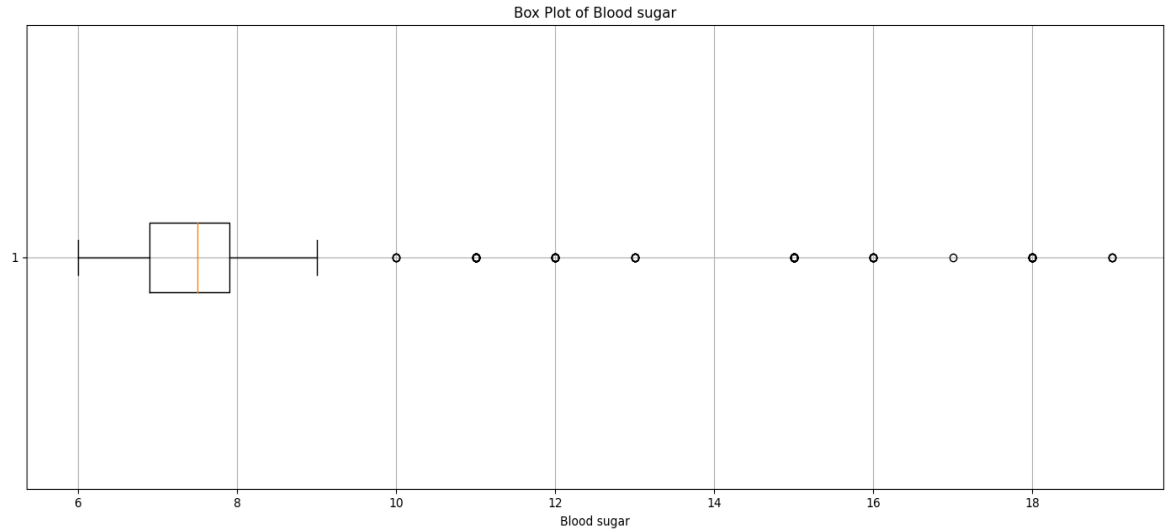
Relationship between Body temperature and Blood Sugar

(r value) = -0.04

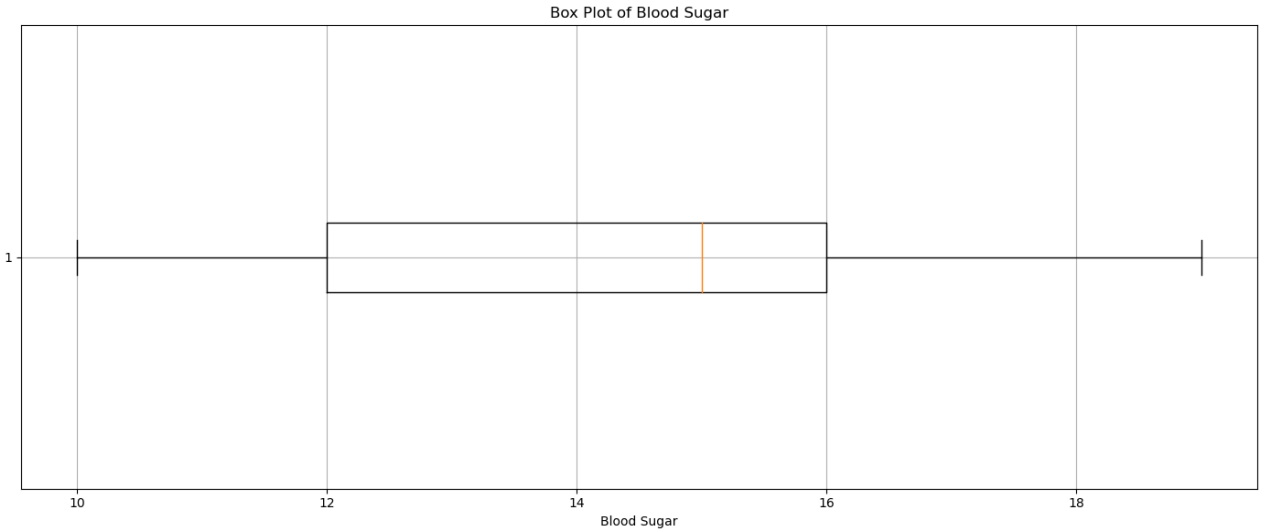
No linear relationship between blood sugar and body temperature.



Outliers for Blood sugar



Boxplot with outliers



Boxplot after removing outliers

Conclusion:

Health Factors	Pregnancy Risk level	Correlation
Increase in Age	Increase	Yes
Systolic BP > 120 mmHg	Increase	Yes
Diastolic BP > 80mmHg	Increase	Yes
Blood sugar > 8mmol/L	Increase	Yes
Heart rate	No effect	No
Body temperature	No effect	No

A strong + linear relationship between Systolic and Diastolic Blood Pressure
A weak + linear relationship between age and systolic blood pressure.
A weak + linear relationship between Systolic Blood Pressure and Blood Sugar Level
No relationship between blood sugar and body temperature