

PIMA INDIAN DIABETES REPORT

GENERAL INFORMATION

Names of columns

['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome']

Shape of dataset

(768, 9)

DESCRIPTIVE STATISTICS

Mean

('Pregnancies', 3.8450520833333335)
('Glucose', 120.89453125)
('BloodPressure', 69.10546875)
('SkinThickness', 20.536458333333332)
('Insulin', 79.79947916666667)
('BMI', 31.992578124999998)
('DiabetesPedigreeFunction', 0.47187630208333325)
('Age', 33.240885416666664)
('Outcome', 0.3489583333333333)

Mode

Pregnancies Glucose BloodPressure ... DiabetesPedigreeFunction Age Outcome 0 1.0 99 70.0 ...
0.254 22.0 0.0 1 NaN 100 NaN ... 0.258 NaN NaN [2 rows x 9 columns]

Median

('Pregnancies', 3.0)
('Glucose', 117.0)
('BloodPressure', 72.0)
('SkinThickness', 23.0)
('Insulin', 30.5)
('BMI', 32.0)
('DiabetesPedigreeFunction', 0.3725)
('Age', 29.0)
('Outcome', 0.0)

Variance

('Pregnancies', 11.354056320621465)
('Glucose', 1022.2483142519557)
('BloodPressure', 374.6472712271838)
('SkinThickness', 254.47324532811822)
('Insulin', 13281.180077955238)
('BMI', 62.15998395738266)
('DiabetesPedigreeFunction', 0.1097786378731394)
('Age', 138.30304589037377)

('Outcome', 0.22748261625380273)

Standard Deviation

('Pregnancies', 3.3695780626988694)
('Glucose', 31.97261819513622)
('BloodPressure', 19.355807170644777)
('SkinThickness', 15.952217567727637)
('Insulin', 115.24400235133817)
('BMI', 7.884160320375446)
('DiabetesPedigreeFunction', 0.3313285950127749)
('Age', 11.760231540678685)
('Outcome', 0.47695137724279896)

Univariate Analysis

---Skewness---

('Pregnancies', 0.9016739791518588)
('Glucose', 0.17375350179188992)
('BloodPressure', -1.8436079833551302)
('SkinThickness', 0.10937249648187608)
('Insulin', 2.272250858431574)
('BMI', -0.42898158845356543)
('DiabetesPedigreeFunction', 1.919911066307204)
('Age', 1.1295967011444805)
('Outcome', 0.635016643444986)

---Kurtosis---

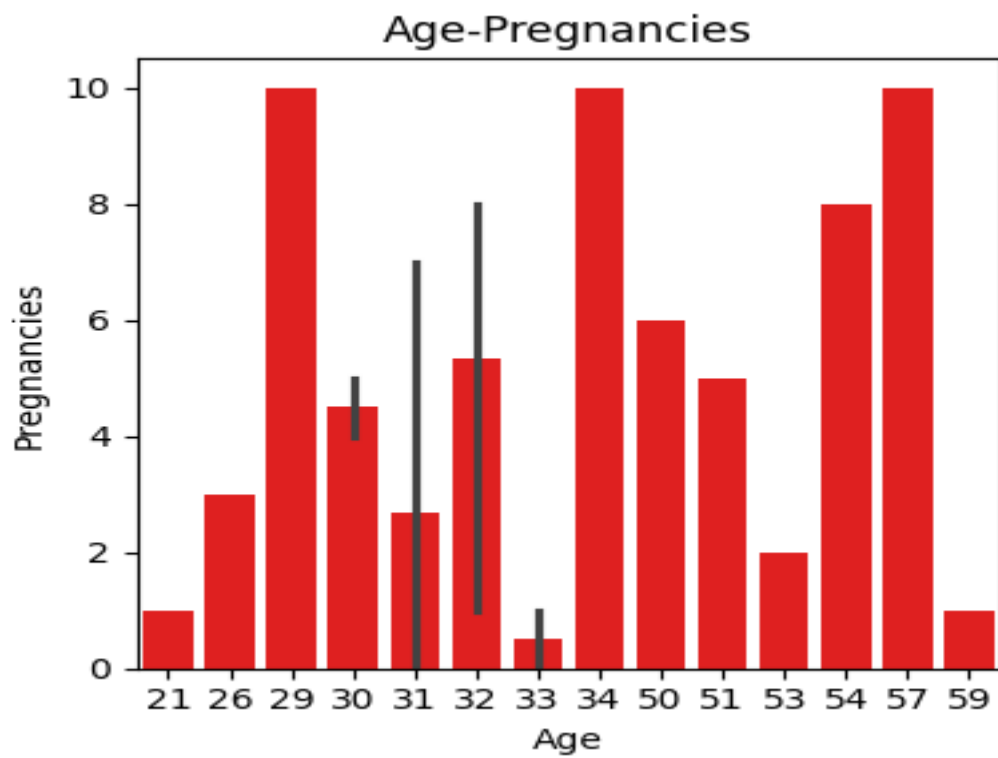
('Pregnancies', 0.15921977754746486)
('Glucose', 0.6407798203735053)
('BloodPressure', 5.180156560082496)
('SkinThickness', -0.520071866153013)
('Insulin', 7.2142595543487715)
('BMI', 3.290442900816981)
('DiabetesPedigreeFunction', 5.5949535279830584)
('Age', 0.6431588885398942)
('Outcome', -1.600929755156027)

Bivariate Analysis

OLS Regression Results =====
===== Dep. Variable: SkinThickness R-squared: 0.154 Model: OLS Adj.
R-squared: 0.153 Method: Least Squares F-statistic: 139.6 Date: Sat, 07 Jun 2025 Prob
(F-statistic): 1.05e-29 Time: 20:54:44 Log-Likelihood: -3152.0 No. Observations: 768 AIC: 6308. Df
Residuals: 766 BIC: 6317. Df Model: 1 Covariance Type: nonrobust =====
===== coef std err t P>|t| [0.025
0.975] -----
const -4.8753 2.215 -2.201
0.028 -9.224 -0.526 BMI 0.7943 0.067 11.814 0.000 0.662 0.926 =====
===== Omnibus: 9.542
Durbin-Watson: 1.938 Prob(Omnibus): 0.008 Jarque-Bera (JB): 9.612 Skew: -0.273 Prob(JB):
0.00818 Kurtosis: 3.045 Cond. No. 138. =====
===== Notes: [1] Standard Errors assume that the
covariance matrix of the errors is correctly specified.

DATA VISUALIZATION

Age-Pregnancies Chart:



Glucose-Insulin Chart:

