

```
In [148...]
```

```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.datasets import load_diabetes
df=load_diabetes(as_frame=True).frame
sns.set_theme(style="whitegrid")
```

```
In [149...]
```

```
df
```

```
Out[149...]
```

| | age | sex | bmi | bp | s1 | s2 | s3 | s4 | s5 | s6 | target |
|-----|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------|
| 0 | 0.038076 | 0.050680 | 0.061696 | 0.021872 | -0.044223 | -0.034821 | -0.043401 | -0.002592 | 0.019907 | -0.017646 | 151.0 |
| 1 | -0.001882 | -0.044642 | -0.051474 | -0.026328 | -0.008449 | -0.019163 | 0.074412 | -0.039493 | -0.068332 | -0.092204 | 75.0 |
| 2 | 0.085299 | 0.050680 | 0.044451 | -0.005670 | -0.045599 | -0.034194 | -0.032356 | -0.002592 | 0.002861 | -0.025930 | 141.0 |
| 3 | -0.089063 | -0.044642 | -0.011595 | -0.036656 | 0.012191 | 0.024991 | -0.036038 | 0.034309 | 0.022688 | -0.009362 | 206.0 |
| 4 | 0.005383 | -0.044642 | -0.036385 | 0.021872 | 0.003935 | 0.015596 | 0.008142 | -0.002592 | -0.031988 | -0.046641 | 135.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 437 | 0.041708 | 0.050680 | 0.019662 | 0.059744 | -0.005697 | -0.002566 | -0.028674 | -0.002592 | 0.031193 | 0.007207 | 178.0 |
| 438 | -0.005515 | 0.050680 | -0.015906 | -0.067642 | 0.049341 | 0.079165 | -0.028674 | 0.034309 | -0.018114 | 0.044485 | 104.0 |
| 439 | 0.041708 | 0.050680 | -0.015906 | 0.017293 | -0.037344 | -0.013840 | -0.024993 | -0.011080 | -0.046883 | 0.015491 | 132.0 |
| 440 | -0.045472 | -0.044642 | 0.039062 | 0.001215 | 0.016318 | 0.015283 | -0.028674 | 0.026560 | 0.044529 | -0.025930 | 220.0 |
| 441 | -0.045472 | -0.044642 | -0.073030 | -0.081413 | 0.083740 | 0.027809 | 0.173816 | -0.039493 | -0.004222 | 0.003064 | 57.0 |

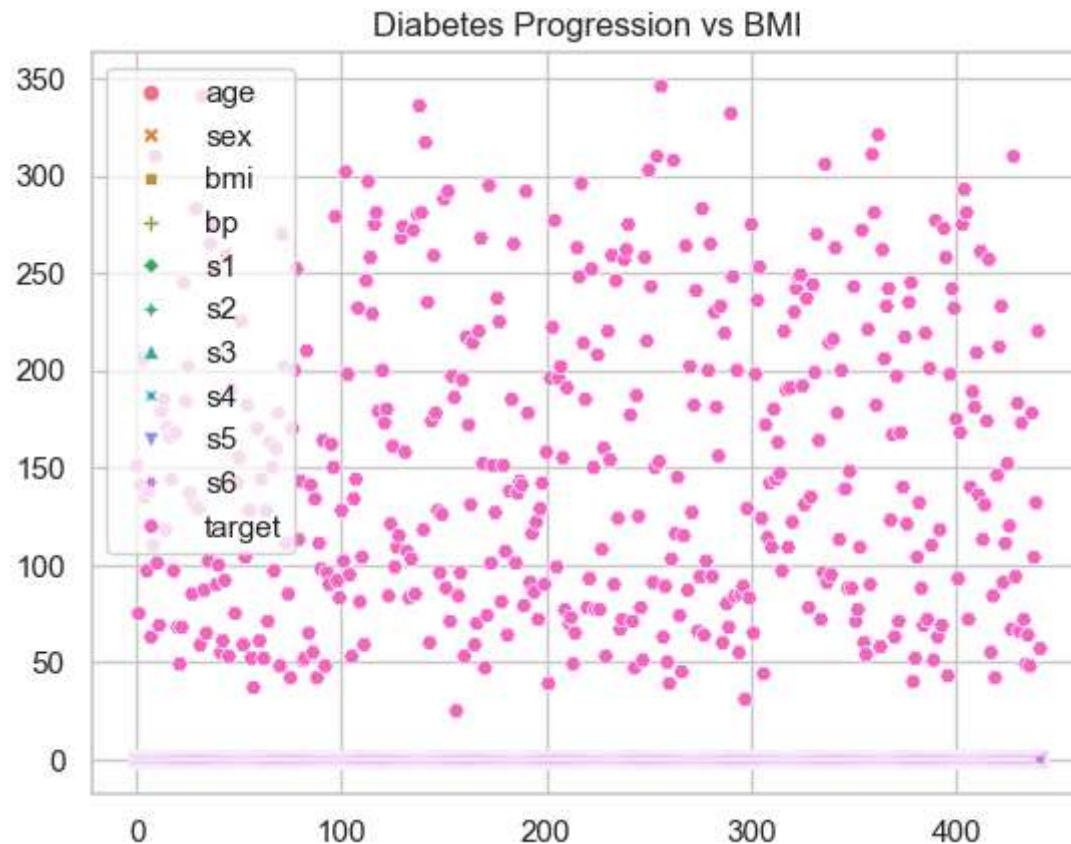
442 rows × 11 columns

```
In [150...]
```

```
sns.scatterplot(data=df)
plt.title("Diabetes Progression vs BMI")
```

```
Out[150...]
```

Text(0.5, 1.0, 'Diabetes Progression vs BMI')



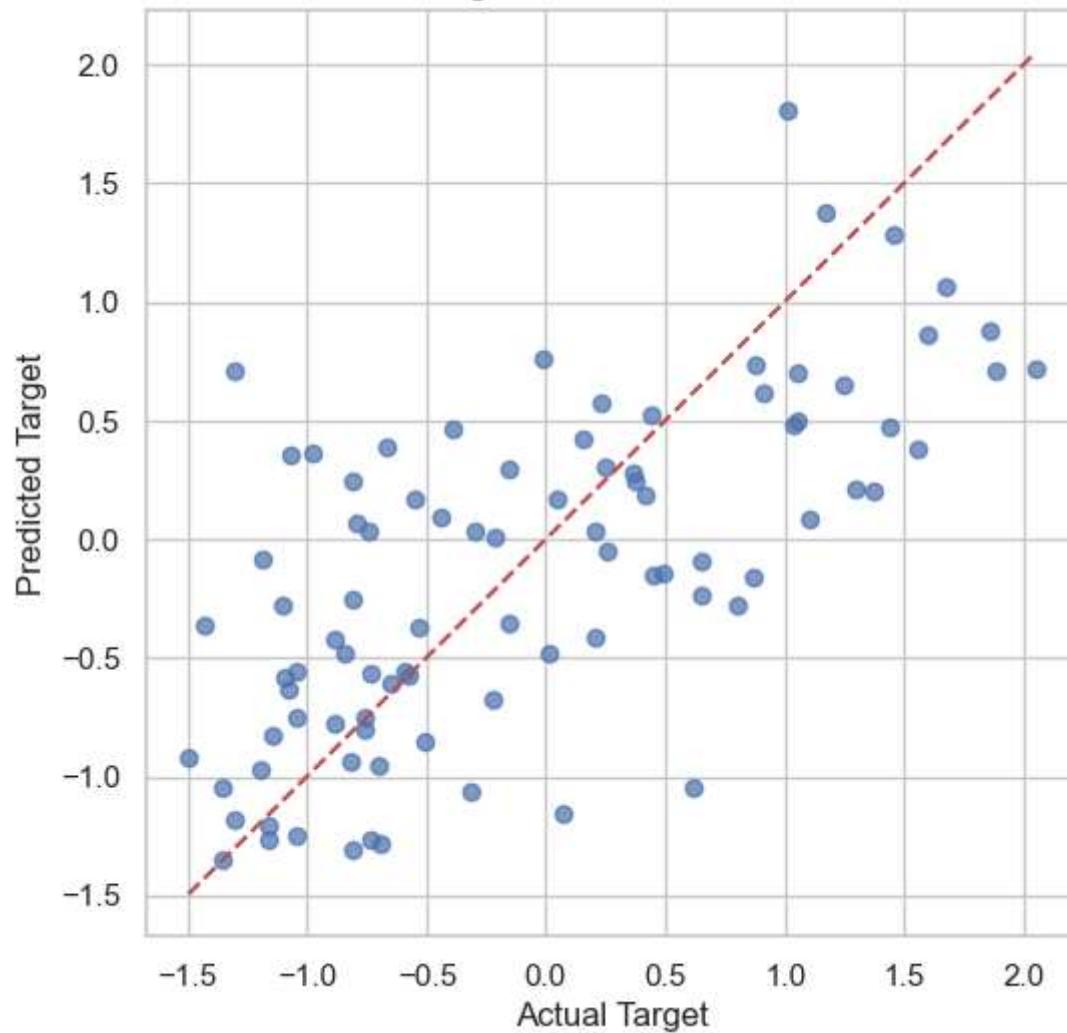
```
In [151]:  
x = df.drop(columns=["target"])  
y = df["target"]
```

```
In [152]:  
from sklearn.model_selection import train_test_split  
from sklearn.linear_model import LinearRegression  
#standardize the features  
from sklearn.preprocessing import StandardScaler  
scaler = StandardScaler()  
X = scaler.fit_transform(X)  
y= scaler.fit_transform(y.values.reshape(-1, 1)).flatten()  
# Split the data into train and test sets  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)  
  
# Fit a Linear regression model  
lr = LinearRegression()
```

```
lr.fit(X_train, y_train)

# Predict on the test set
y_pred_lr = lr.predict(X_test)
y_pred_train = lr.predict(X_train)
plt.figure(figsize=(6,6))
plt.scatter(y_test, y_pred_lr, alpha=0.7)
plt.xlabel("Actual Target")
plt.ylabel("Predicted Target")
plt.title("Linear Regression: Actual vs Predicted")
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], 'r--')
plt.show()
```

Linear Regression: Actual vs Predicted



In [153]:

x_train

```
Out[153]: array([[ 1.48782782,  1.06548848,  0.25474221, ...,  0.72130245,
   0.57529619, -0.02265729],
 [-0.19230663,  1.06548848, -0.37972943, ...,  1.49710409,
  0.00572853,  0.41275676],
 [ 0.11317236, -0.93853666,  1.04783175, ..., -0.05449919,
  0.35816289, -0.28390573],
 ...,
 [ 0.6477606 , -0.93853666, -0.42504883, ..., -0.83030083,
 -0.22922772, -0.02265729],
 [-0.26867637, -0.93853666, -0.49302793, ..., -0.05449919,
 -0.80856925, -0.80640259],
 [-1.94881082, -0.93853666,  0.59463772, ..., -0.83030083,
 -0.10810835, -0.02265729]])
```

```
In [154]: X_test
```

```
Out[154... array([[ 0.95323959, -0.93853666, -0.13047271, -0.33635919,  2.62836772,
   2.63213934,  0.40338351,  0.72130245,  0.68185057, -0.1097401 ],
 [ 1.9460463 , -0.93853666,  0.77591533,  0.45984057, -0.52475728,
 -0.35021759,  0.0163735 , -0.83030083, -0.4733829 , -0.45807135],
 [ 1.33508832,  1.06548848, -0.08515331, -0.2639774 ,  2.16552368,
  1.02574841,  1.17740353, -0.05449919,  1.77633271, -0.37098854],
 [ 2.02241605, -0.93853666,  1.09315115,  1.66644512,  1.15305235,
  0.76898921, -1.60906855,  2.9711272 ,  2.07395673,  1.28358487],
 [ 0.26591186,  1.06548848, -0.42504883, -0.04683201,  0.80591932,
  1.11791838, -0.1384305 ,  0.72130245, -0.10810835, -0.19682291],
 [ 0.18954211, -0.93853666, -0.51568763, -0.55350458,  2.07874042,
  1.98036597,  1.48701154, -0.05449919, -0.44981062,  0.15150833],
 [-0.19230663,  1.06548848,  3.58571828,  0.31507698,  0.63235281,
  0.70973708, -0.44803851,  0.72130245,  0.70753094,  0.6740052 ],
 [-0.49778562,  1.06548848,  0.95719294,  0.45984057,  2.31016244,
  1.86844529,  0.0163735 ,  0.72130245,  1.55977434,  1.28358487],
 [-1.94881082,  1.06548848, -1.89792941, -1.20494075, -0.52475728,
 -0.63989464, -0.1384305 , -0.05449919,  0.50572921,  0.06442552],
 [-1.26148309,  1.06548848,  0.32272131, -0.40874099,  0.77699157,
  1.01258127,  0.40338351, -0.05449919, -0.64643795, -0.02265729],
 [-0.87963435, -0.93853666, -0.69696524, -0.48112279,  0.97948584,
  0.87432632,  1.17740353, -0.51998017, -0.5456329 , -0.80640259],
 [ 0.11317236, -0.93853666, -1.21813837, -0.48112279, -1.4215176 ,
 -1.43650654, -1.14465653, -0.05449919,  0.90185853, -1.76431351],
 [-1.87244107, -0.93853666, -1.28611748, -0.55350458, -1.16116783,
 -1.14682949,  0.86779553, -1.60610247, -1.97491811, -1.15473383],
 [ 0.41865135,  1.06548848,  0.30006161,  1.32842213,  0.31414754,
  0.4266436 , -0.98985253,  0.72130245,  0.98100774,  1.89316455],
 [-0.26867637,  1.06548848, -1.17281897, -0.04683201, -0.58261278,
 -0.61356037,  0.40338351, -0.83030083, -0.3585878 ,  0.93525363],
 [-0.72689486,  1.06548848,  0.118784 , -0.1192138 , -1.53722861,
 -1.31800229, -0.1384305 , -0.83030083, -0.9549857 ,  0.6740052 ],
 [ 0.87686984,  1.06548848,  1.50102577,  0.17031339,  0.80591932,
  0.33447363, -0.37063651,  0.72130245,  1.54329291,  1.80608174],
 [ 1.33508832,  1.06548848,  1.86358099,  1.47318573,  0.42985854,
  0.78873991, -1.06725453,  1.49710409,  0.61592484,  1.5448333 ],
 [-0.65052511,  1.06548848,  0.59463772,  1.47318573, -2.66541095,
 -2.24628558, -1.14465653, -1.0087352 , -0.64643795,  0.32567395],
 [ 1.02960933,  1.06548848,  1.22910936,  1.47318573,  0.28521978,
  0.43322717, -0.44803851,  0.72130245,  0.46260918,  0.58692238],
 [-1.49059233, -0.93853666,  0.82123474, -0.69826818, -0.26440751,
 -0.72548105, -0.52544052, -0.05449919,  1.42408999, -0.28390573],
```

$[1.02960933, -0.93853666, -0.90090256, -1.13255896, 0.95055809,$
 $1.05208269, 0.71299152, -0.05449919, -0.5456329, -1.32889946],$
 $[-0.87963435, -0.93853666, -1.03686076, -0.77064997, -0.1486965,$
 $-0.47530541, 1.79661955, -0.83030083, -1.39787631, 0.15150833],$
 $[1.25871858, 1.06548848, 0.48133922, 1.03889495, 0.34307529,$
 $0.24888722, -0.29323451, -0.05449919, 0.83133333, 0.41275676],$
 $[-1.26148309, -0.93853666, 0.93453324, -0.40874099, -0.206552,$
 $-0.15929408, 0.48078551, -0.83030083, -0.57035505, -0.19682291],$
 $[0.87686984, 1.06548848, 0.30006161, 0.89413135, -0.64046829,$
 $-0.02762269, -0.91245053, -0.05449919, -0.69894856, 0.32567395],$
 $[0.6477606, 1.06548848, 1.25176906, 1.18365854, -0.46690178,$
 $0.02504586, -0.68024452, -0.05449919, -0.52129404, -0.37098854],$
 $[0.80050009, 1.06548848, -0.37972943, 1.40080393, -1.07438457,$
 $-0.35021759, -1.60906855, 0.72130245, -0.25011698, -0.28390573],$
 $[-1.18511334, -0.93853666, -1.55803389, -1.06017716, -0.52475728,$
 $-0.98882382, 1.95142356, -1.60610247, -1.28614752, -0.98056821],$
 $[0.34228161, 1.06548848, -0.44770853, -0.1915956, 0.71913607,$
 $1.0059977, 0.0163735, -0.05449919, -0.27138952, 0.49983957],$
 $[0.57139085, -0.93853666, 1.9542198, -1.10867296, 0.16950877,$
 $0.8348249, -0.60284252, 0.44201386, -1.01669526, 0.41275676],$
 $[-2.17792006, -0.93853666, -0.78760405, -0.55350458, 0.05379777,$
 $0.42006003, 0.24857951, -0.05449919, -1.43658851, -0.54515416],$
 $[-0.03956713, 1.06548848, 0.30006161, -1.56684974, 0.05379777,$
 $0.13038297, -0.29323451, -0.05449919, 0.40358265, 0.06442552],$
 $[0.57139085, 1.06548848, -0.13047271, 0.60460417, -0.35119077,$
 $-0.03420626, -1.22205854, 0.72130245, 0.61592484, 0.6740052],$
 $[2.02241605, -0.93853666, 0.84389444, -1.20494075, 0.95055809,$
 $1.27592405, -0.44803851, 0.76009253, 0.26387376, 0.49983957],$
 $[-1.64333183, 1.06548848, 1.63698398, 1.11127674, 1.64482414,$
 $1.35492688, 0.55818752, -0.05449919, 0.85509726, -0.19682291],$
 $[-0.19230663, 1.06548848, -0.65164584, -0.55350458, -0.23547976,$
 $-0.02103912, -0.44803851, -0.05449919, 0.13048914, 0.58692238],$
 $[-0.03956713, 1.06548848, -0.69696524, -0.384855, 0.66128056,$
 $0.90066059, -0.29323451, 0.41873981, 0.2150044, 0.58692238],$
 $[1.41145807, 1.06548848, -0.65164584, 0.09793159, 0.5166418,$
 $0.7492385, -0.60284252, 0.72130245, 0.49135587, 1.71899893],$
 $[0.41865135, -0.93853666, -1.21813837, -1.20494075, -0.03298549,$
 $-0.26463119, 1.56441355, -0.83030083, -1.28614752, -1.59014789],$
 $[-0.03956713, -0.93853666, -1.46739509, -0.2639774, -0.00405774,$
 $-0.19221193, 1.48701154, -0.83030083, -1.32275163, 0.84817082],$
 $[1.41145807, 1.06548848, -0.62898614, 1.20754453, -0.00405774,$
 $-0.33046688, 1.56441355, -1.06304132, -0.80856925, 0.15150833],$

```

[ 0.87686984,  1.06548848, -0.47036823,  0.60460417, -1.39258985,
 -0.94932241, -1.29946054, -0.05449919,  0.06015558, -1.15473383],
[ 2.32789504,  1.06548848, -0.69696524, -0.48112279, -0.090841 ,
  0.4266436 , -1.29946054,  1.49710409,  0.32730811,  0.93525363],
[ 0.87686984, -0.93853666, -0.94622196,  0.72548177,  0.92163033,
 -0.33046688,  0.79039352, -0.30275571,  1.88997794,  0.15150833],
[-0.34504612, -0.93853666, -0.74228465, -0.55350458,  0.69020832,
  0.3608079 ,  2.10622756, -0.83030083, -1.47606729, -1.6772307 ],
[-2.02518057, -0.93853666, -0.76494435, -1.56684974, -0.8140348 ,
 -0.58064252,  0.32598151, -0.83030083, -1.55770788, -0.02265729],
[ 0.6477606 , -0.93853666, -1.05952046, -0.04683201, -0.92974581,
 -1.89077283,  2.49323757, -1.60610247, -0.38081857,  0.06442552],
[-0.42141587, -0.93853666, -1.7846309 , -0.55350458, -0.7561793 ,
 -0.71889748,  0.86779553, -1.08631537, -1.73191279, -0.98056821],
[-1.26148309, -0.93853666,  0.02814519, -0.62588638, -0.1486965 ,
 -0.4555547 ,  0.24857951, -0.05449919,  0.66881874, -1.15473383],
[ 0.11317236, -0.93853666,  1.22910936, -0.91541357, -1.53722861,
 -1.52209294,  0.40338351, -1.60610247, -1.08070454, -0.54515416],
[-2.02518057, -0.93853666, -1.46739509, -1.42208614, -0.40904627,
 -0.22512977,  0.32598151, -0.83030083, -0.98564884, -1.6772307 ],
[ 0.57139085,  1.06548848,  0.36804071, -0.69826818, -0.1486965 ,
  0.96649629, -1.37686254,  1.49710409, -2.02742872, -1.24181665],
[ 0.41865135, -0.93853666, -0.85558315, -0.33635919, -0.17762425,
 -0.3699683 ,  1.10000153, -0.83030083, -0.64643795,  0.06442552],
[-1.1087436 ,  1.06548848, -0.37972943,  1.69033112,  1.87624616,
  2.26345946, -0.83504852,  2.27290573,  0.75812511, -0.8934854 ],
[-0.57415536, -0.93853666,  1.36506757, -0.04683201, -0.52475728,
 -0.36338473,  0.48078551, -0.83030083, -1.28614752, -1.32889946],
[-0.49778562, -0.93853666, -0.96888166, -0.69826818,  0.69020832,
  0.76240564,  0.79039352, -0.05449919, -0.69894856,  0.23859114],
[ 0.80050009,  1.06548848,  1.29708846,  0.45984057, -0.92974581,
 -0.73206462, -0.91245053, -0.05449919,  0.41853093, -0.37098854],
[-0.57415536, -0.93853666, -0.37972943, -0.84303177, -0.06191324,
 -0.23829691,  0.79039352, -0.83030083, -0.18802414, -1.15473383],
[-0.80326461, -0.93853666, -1.15015927, -1.63923153, -0.69832379,
 -1.81835357,  2.95764959, -1.60610247, -0.40362428, -0.1097401 ],
[-0.49778562, -0.93853666,  0.63995713, -0.1192138 ,  1.7316074 ,
  1.93428099, -0.37063651,  1.49710409,  0.69469075,  0.06442552],
[-0.87963435, -0.93853666,  0.86655414, -0.55350458, -0.66939604,
 -0.63989464, -0.75764652,  0.06187106,  0.70753094, -0.37098854],
[-1.26148309,  1.06548848,  1.13847055, -0.40874099, -1.04545682,
 -1.02832524,  0.48078551, -0.83030083, -0.92470586, -0.1097401 ],

```

```

[ 0.34228161,  1.06548848, -0.94622196,  1.32842213,  0.22736428,
 -0.00787198,  1.33220754, -0.83030083, -0.64643795,  0.76108801],
[-0.87963435, -0.93853666, -1.37675628, -0.84303177, -0.11976875,
  0.30155578, -0.91245053,  0.72130245,  0.14773715, -0.28390573],
[ 1.56419757, -0.93853666,  0.39070041,  1.32842213,  1.29769111,
  0.90066059,  0.17117751, -0.05449919,  1.22018017, -1.24181665],
[-1.1087436 ,  1.06548848, -0.24377122,  1.18365854,  1.1819801 ,
  1.53268325, -0.83504852,  1.49710409,  0.64256343, -0.1097401 ],
[ 1.10597908,  1.06548848, -0.51568763,  1.18365854, -0.1486965 ,
 -0.10662552, -0.44803851, -0.05449919,  0.56168942, -0.80640259],
[-2.25428981, -0.93853666, -0.71962495, -1.42208614, -1.33473434,
 -1.48259153,  0.17117751, -0.83030083, -0.012861 , -1.6772307 ],
[-1.41422259,  1.06548848, -0.26643092, -0.84303177, -0.32226301,
  0.09746513, -1.22205854,  0.72130245,  0.40358265, -0.71931978],
[-1.56696208,  1.06548848, -0.37972943,  0.17031339, -0.40904627,
 -0.5213904 , -1.37686254,  0.72130245,  1.41527434, -0.37098854],
[-0.03956713,  1.06548848,  0.63995713,  1.11127674,  0.83484708,
  1.19033765, -0.83504852,  1.49710409,  0.53390096,  0.58692238],
[ 1.25871858, -0.93853666, -0.44770853,  1.83509471,  0.95055809,
  0.6636521 , -0.98985253,  1.49710409,  1.66345406,  2.85107547],
[-1.26148309,  1.06548848, -0.22111152, -0.3124732 , -1.04545682,
 -0.49505612, -1.22205854,  0.33340163, -0.2085301 , -0.71931978],
[ 1.41145807, -0.93853666, -1.28611748, -0.84303177, -0.55368503,
 -0.51480683,  0.71299152, -0.83030083, -1.18055136, -1.24181665],
[ 0.18954211,  1.06548848, -0.83292345,  0.60460417,  0.80591932,
  1.54585039, -1.53166655,  2.27290573,  0.32730811, -0.98056821],
[-0.57415536,  1.06548848,  1.27442876,  1.03889495,  1.7894629 ,
  1.81577674, -0.0610285 ,  0.72130245,  0.79492086,  1.02233644],
[-0.9560041 , -0.93853666,  0.82123474,  0.02554979,  0.34307529,
  0.32130649, -0.60284252,  0.55838411,  0.93616291, -0.54515416],
[ 0.95323959,  1.06548848,  0.41336012,  0.82174956,  0.42985854,
  0.54514785,  0.17117751, -0.05449919, -0.06939615,  0.41275676],
[ 0.26591186, -0.93853666, -0.42504883, -0.33635919,  0.25629203,
  0.44639431, -1.60906855,  2.27290573,  1.25889237, -0.45807135],
[-1.79607132, -0.93853666, -0.08515331, -0.1915956 , -0.06191324,
  0.16330082,  0.48078551, -0.83030083, -1.28614752, -0.28390573],
[-1.18511334, -0.93853666, -0.24377122, -0.69826818, -0.98760132,
 -1.00199096,  0.0937755 , -0.83030083, -0.16770981, -1.85139632],
[-1.03237385, -0.93853666, -1.35409658, -2.14590411, -0.06191324,
 -0.32388331,  1.33220754, -0.99321917, -0.69894856, -1.15473383],
[-0.57415536, -0.93853666, -1.26345777, -0.62588638,  0.97948584,
  0.42006003,  2.57063958, -0.83030083, -1.08070454, -0.19682291],

```

```
[ 0.03680262, -0.93853666, -1.37675628, -0.1192138 , -0.1486965 ,  
-0.40946972,  0.86779553, -0.83030083, -0.06939615,  0.15150833],  
[ 0.26591186, -0.93853666, -0.53834734, -0.84303177, -0.64046829,  
-0.94932241,  1.64181555, -1.60610247, -1.5165043 ,  0.23859114],  
[-0.57415536, -0.93853666, -1.33143688, -1.06017716, -1.88436164,  
-2.19361702,  1.10000153, -1.60610247, -1.18055136, -1.41598227],  
[-0.49778562, -0.93853666, -1.46739509, -1.34970435, -1.24795109,  
-1.06124309,  0.40338351, -0.83030083, -1.87392142, -1.06765102],  
[-1.33785284, -0.93853666,  0.75325563, -0.48112279, -0.64046829,  
-0.39630258, -0.1384305 , -0.05449919, -0.5456329 , -1.15473383]])
```

In [155...]: y_train

```
Out[155]: array([-1.05621783e-01, -2.77055191e-02,  1.66048019e+00, -3.52356618e-01,
 -1.20943552e+00, -1.13151925e+00,  1.67346624e+00,  1.62152206e+00,
 1.93057228e-01, -8.06868156e-01, -1.04061695e+00,  3.35903711e-01,
 -8.32840244e-01,  1.53061975e+00, -6.64021672e-01,  5.82638546e-01,
 -7.15965848e-01, -1.28735178e+00, -1.05360299e+00,  1.42673140e+00,
 -1.30033783e+00, -6.51035629e-01,  1.80071184e-01,  4.78750195e-01,
 2.70973492e-01,  8.81317557e-01,  6.99512942e-01, -7.15965848e-01,
 -1.19644947e+00, -1.18346343e+00,  1.16701052e+00, -4.04300794e-01,
 -3.13398486e-01, -6.25063541e-01,  1.46568953e+00, -2.61454310e-01,
 1.69943833e+00,  2.83959536e-01, -3.00412442e-01,  1.36180118e+00,
 -1.96524090e-01, -1.15749134e+00, -7.67910024e-01,  1.03715008e+00,
 7.25485030e-01,  1.41374536e+00,  3.48889755e-01,  1.37478723e+00,
 1.42673140e+00, -1.31332387e+00,  1.10208030e+00, -1.05360299e+00,
 -1.70552003e-01,  1.50464767e+00, -1.07957508e+00,  2.14096382e+00,
 1.25791283e+00,  2.42386567e-02,  5.17708327e-01, -4.69231013e-01,
 -9.23742551e-01, -3.91314750e-01,  1.38777327e+00,  5.04722283e-01,
 1.81631272e+00, -1.26137969e+00, -5.86105409e-01,  7.51457118e-01,
 -7.93882112e-01,  1.30985701e+00, -8.71798376e-01,  1.29687096e+00,
 -1.20943552e+00, -9.62700683e-01,  6.21596678e-01, -9.62700683e-01,
 1.20596866e+00,  2.96945580e-01, -1.43019827e+00, -3.26384530e-01,
 -1.28735178e+00, -7.54923980e-01, -6.25063541e-01,  6.08610634e-01,
 1.46568953e+00,  1.67346624e+00,  1.24492679e+00,  1.36180118e+00,
 8.16387337e-01,  1.95915920e+00,  2.32015360e-01, -1.20943552e+00,
 1.62152206e+00,  7.38471074e-01, -1.83538046e-01,  5.95624590e-01,
 -3.65342662e-01, -7.28951892e-01,  1.76436855e+00,  9.46247777e-01,
 1.46568953e+00, -6.64021672e-01, -1.26137969e+00,  5.95624590e-01,
 -1.31332387e+00,  1.29687096e+00, -1.14450530e+00,  8.81317557e-01,
 -2.74440354e-01,  7.77429206e-01, -1.31593871e-01, -6.38049585e-01,
 3.72247006e-02, -4.04300794e-01, -8.58812332e-01, -5.34161233e-01,
 -1.13151925e+00, -2.74440354e-01, -1.31332387e+00, -3.13398486e-01,
 -1.44579915e-01, -1.35228200e+00, -5.60133321e-01,  3.35903711e-01,
 -8.32840244e-01, -8.84784420e-01,  8.29373381e-01, -2.77055191e-02,
 -1.19644947e+00, -7.28951892e-01,  4.91736239e-01, -1.01464486e+00,
 1.64749415e+00,  3.87847887e-01,  1.02154920e-01,  1.20596866e+00,
 1.60853602e+00,  2.83959536e-01,  3.61875799e-01, -2.77055191e-02,
 5.69652502e-01, -1.83538046e-01, -7.15965848e-01,  1.21895470e+00,
 2.19290800e+00,  2.02408942e+00, -5.60133321e-01, -1.07957508e+00,
 3.87847887e-01,  1.37478723e+00,  1.15140964e-01,  3.35903711e-01,
 8.03401293e-01, -1.39124013e+00, -2.77055191e-02,  1.02154920e-01,
 -1.26137969e+00,  5.82638546e-01,  4.26806019e-01,  1.50464767e+00,
 2.05006151e+00, -3.78328706e-01, -1.09256112e+00, -1.04061695e+00,
 4.26806019e-01, -1.05621783e-01, -6.66636509e-02,  2.06043272e-01,
```

3.35903711e-01, 1.21895470e+00, -1.47194752e-02, -3.26384530e-01,
-8.97770464e-01, 2.33575448e+00, -1.73343121e-03, -5.60133321e-01,
-8.06868156e-01, -1.11853321e+00, 8.03401293e-01, -8.71798376e-01,
-3.00412442e-01, -8.19854200e-01, 1.38777327e+00, 9.98191953e-01,
6.21596678e-01, -9.75686727e-01, -1.27436574e+00, -1.57304475e+00,
-5.60133321e-01, 6.99512942e-01, -1.05621783e-01, -4.43258925e-01,
-8.97770464e-01, 1.16701052e+00, 1.38777327e+00, -1.04061695e+00,
1.41113052e-01, 3.74861843e-01, -1.44579915e-01, -1.05360299e+00,
-1.96524090e-01, 5.56666458e-01, 3.48889755e-01, -6.51035629e-01,
-2.74440354e-01, -1.36526805e+00, 1.07610822e+00, -9.75686727e-01,
5.95624590e-01, -7.67910024e-01, 1.28127008e-01, 9.46247777e-01,
1.59554997e+00, 4.00833931e-01, 1.99811734e+00, -9.23742551e-01,
-1.26137969e+00, -7.96496949e-02, 5.69652502e-01, 1.01117800e+00,
2.05006151e+00, -1.45617035e+00, -2.22496178e-01, 2.51755909e+00,
-1.41721222e+00, -3.13398486e-01, -9.75686727e-01, 1.07610822e+00,
-1.33929596e+00, -1.01464448e+00, -7.80896068e-01, -8.84784420e-01,
1.43971745e+00, -1.05621783e-01, -1.31593871e-01, 2.45262887e+00,
-4.82217057e-01, 7.61828325e-02, 1.56957789e+00, -8.71798376e-01,
-8.32840244e-01, 8.81317557e-01, -1.46915640e+00, -9.36728595e-01,
2.57987448e-01, 8.42359425e-01, 2.38769865e+00, -1.30033783e+00,
1.55659184e+00, -4.82217057e-01, -5.47147277e-01, -2.74440354e-01,
-1.05360299e+00, 1.59554997e+00, -4.43258925e-01, -1.65096101e+00,
-6.77007716e-01, 1.67346624e+00, 8.94303601e-01, 1.24492679e+00,
6.21596678e-01, -2.61454310e-01, -7.93882112e-01, -1.10554717e+00,
6.47568766e-01, -1.02763090e+00, -8.71798376e-01, 1.59554997e+00,
1.17999657e+00, -1.11853321e+00, 1.82929877e+00, 1.08909426e+00,
1.17999657e+00, -8.45826288e-01, -1.46915640e+00, 8.42359425e-01,
-7.80896068e-01, 1.86825690e+00, 1.81631272e+00, -1.31593871e-01,
-1.32630991e+00, -1.28735178e+00, -6.25063541e-01, -1.00165882e+00,
-4.17286837e-01, -1.31593871e-01, -1.18607827e-01, -6.89993760e-01,
-1.13151925e+00, -4.69231013e-01, 1.05013613e+00, 1.54099096e-01,
-7.41937936e-01, -1.20943552e+00, -1.70552003e-01, -9.26357388e-02,
3.22917667e-01, 4.26806019e-01, -7.15965848e-01, -1.43019827e+00,
6.34582722e-01, 1.15402448e+00, -1.06658903e+00, -9.62700683e-01,
-1.33929596e+00, -6.38049585e-01, -1.40422618e+00, -5.34161233e-01,
5.04722283e-01, -1.36526805e+00, 3.87847887e-01, -1.22242156e+00,
3.72247006e-02, -1.47194752e-02, -9.49714639e-01, -6.25063541e-01,
-1.18607827e-01, -1.73343121e-03, 2.32015360e-01, -1.00165882e+00,
6.21596678e-01, -3.65342662e-01, -7.93882112e-01, -1.33929596e+00,
1.41113052e-01, -1.28735178e+00, 1.69943833e+00, 3.35903711e-01,
8.68331513e-01, 6.21596678e-01, -5.08189145e-01, -5.08189145e-01,
-1.15749134e+00, -4.95203101e-01, -3.39370574e-01, 1.58256393e+00,

```
-8.32840244e-01,  2.06304756e+00, -8.97770464e-01, -1.05360299e+00,
-2.35482222e-01,  1.19298261e+00, -1.13151925e+00,  2.70973492e-01,
-1.23540761e+00, -1.09256112e+00, -1.44579915e-01,  1.53061975e+00,
-2.35482222e-01,  6.47568766e-01, -5.36776070e-02, -1.14450530e+00,
 1.94617316e+00])
```

In [156... y_test

```
Out[156... array([ 0.86833151, -1.06658903,  0.64756877,  1.011178 , -0.53416123,
 -0.88478442,  1.16701052,  1.55659184, -0.75492398, -0.72895189,
 -0.75492398,  1.29687096, -0.68999376,  1.88124294, -0.22249618,
 -1.10554717,  1.85527085,  1.45270349,  0.23201536,  1.59554997,
  2.05006151, -1.1445053 , -0.31339849,  1.03715008, -0.30041244,
 -0.44325893,  1.43971745, -0.97568673, -1.352282 , -0.58610541,
 -0.15756596, -0.50818915, -0.80686816,  0.1540991 ,  0.3618758 ,
  1.05013613, -1.43019827, -0.88478442,  0.25798745, -1.15749134,
 -1.352282 , -0.57311936,  0.05021074,  0.20604327, -0.80686816,
 -1.30033783,  0.62159668, -0.84582629, -0.80686816,  1.37478723,
 -0.20951013,  0.07618283, -1.07957508, -1.04061695,  0.2450014 ,
 -0.74193794, -1.04061695, -0.01471948,  0.20604327, -1.19644947,
 -0.39131475, -1.30033783,  0.45277811, -0.65103563,  0.80340129,
  1.24492679,  0.37486184, -0.54714728, -0.15756596,  0.64756877,
 -0.66402167,  0.90728965,  1.67346624, -1.18346343, -0.8198542 ,
 -0.79388211,  0.43979206,  0.88131756,  1.1020803 ,  1.05013613,
 -1.09256112,  0.49173624, -0.72895189, -1.04061695,  0.01125261,
 -0.7029798 , -1.49512849, -1.15749134,  0.41381998])
```

In [157... y_pred_lr

```
Out[157... array([-0.16344139,  0.35560625, -0.23497894,  1.80874224, -0.36807416,
       -0.77865797,  1.37780504,  0.3792423 , -0.80395781, -0.56488935,
      -0.75312337,  0.21169043, -1.28079657,  0.70770475, -0.67532916,
     -0.27877028,  0.8752234 ,  1.28103219,  0.5744422 ,  0.86281398,
      0.71705062, -0.82656267, -1.06096791,  0.47821962,  0.03575489,
     0.09386594,  0.46982402,  0.36695257, -1.35240582, -0.56046396,
     0.29407536, -0.85408455, -0.24901872,  0.420809 ,  0.2817806 ,
     0.49639284, -0.35994543, -0.42181758, -0.054305 , -1.20873359,
    -1.04550855, -0.57723793,  0.17297632,  0.03735144,  0.24562462,
   -1.17752084, -1.04494486, -0.48264141, -1.30579514,  0.20053713,
     0.00505723, -1.15806682, -0.6315901 , -0.55743852,  0.30527203,
     0.03206875, -0.74950823,  0.76109395, -0.41395435, -0.96768922,
     0.46488162,  0.70591993, -0.14936305, -0.60432768, -0.27828 ,
     0.64997567,  0.24669472,  0.16597133, -0.35593134, -0.09509908,
     0.38780055,  0.61398291,  1.06590589, -0.08021257, -0.93845531,
     0.067994 ,  0.52737157,  0.73714832,  0.08380857,  0.6997981 ,
    -0.57987579, -0.14541116, -1.26370039, -1.2493581 , -0.48206382,
   -0.95028808, -0.91645903, -1.2694314 ,  0.18337421])
```

```
In [158... y_pred_train
```

```
Out[158]: array([ 4.22910067e-01, -2.42844344e-03,  1.13107578e+00, -6.11835725e-01,
   -1.13109088e+00, -1.27387582e+00,  5.92352480e-01,  1.27884785e+00,
   4.21351811e-01, -6.16464920e-01, -1.14349051e+00,  5.43600399e-01,
  -6.39044371e-01,  1.85835949e+00, -7.36471262e-01, -1.76032260e-02,
  -6.29002244e-01, -3.19966269e-01, -3.73408403e-01,  2.58709693e-01,
  3.65106665e-01, -5.19653585e-01,  7.24741774e-01,  6.70782952e-01,
  7.78733696e-01,  5.82773909e-01,  1.16847939e-01, -6.21153778e-02,
  -2.48571805e-01, -5.17561000e-01,  1.86873631e-01,  2.35397894e-01,
  -7.58112680e-01, -1.41953756e+00,  1.75727307e-01, -3.55363195e-01,
  4.17752534e-01,  2.72129555e-01, -5.53577304e-01,  6.18520680e-01,
  -7.76470986e-01, -4.42804268e-01, -8.78808320e-01,  1.07507722e+00,
  1.07042762e+00,  1.09817545e+00,  1.03010703e-01,  1.80205235e+00,
  9.42392441e-02, -1.00978601e+00,  9.96286169e-01, -8.66581867e-01,
  2.30773173e-01,  9.22406990e-01, -7.34976418e-01,  1.02108914e+00,
  7.53215299e-01,  7.35104237e-02,  7.61179023e-01, -2.12604562e-01,
  -8.54045251e-02,  2.28593958e-01,  1.18117007e+00, -3.89309395e-01,
  6.43643824e-01, -8.87995923e-01,  3.57931840e-01, -1.19739991e-02,
  -8.00229912e-01, -3.81688360e-01, -7.88023670e-02, -1.70083765e-03,
  -9.94349126e-01, -1.23179779e+00, -4.78795118e-02, -4.98138641e-01,
  1.79782336e-01,  5.01941298e-01, -9.20879615e-01, -4.00846139e-01,
  -7.54110073e-01,  5.78502805e-02, -1.06012032e+00, -4.26708801e-01,
  6.91887604e-01,  1.35398101e+00,  8.17705150e-01,  9.98570669e-01,
  1.16784050e+00,  1.46572804e+00, -2.85914337e-01, -4.12794580e-01,
  4.39585667e-01,  3.54149548e-03, -1.04344257e+00,  4.01607557e-01,
  -1.79884576e-01, -9.11644612e-01,  8.04980222e-01,  9.41128005e-01,
  6.28562228e-01, -7.55627518e-01, -6.62979831e-02,  1.04119234e-01,
  -3.76166505e-01,  1.12498286e+00, -4.96355972e-01,  1.44306977e+00,
  1.43465327e-01,  5.34461874e-01, -1.09677098e-01, -2.81902039e-01,
  9.38655044e-01,  7.11278370e-01, -1.73120996e-02, -6.49011680e-01,
  -3.60001176e-01,  6.25518183e-01, -9.75720884e-01,  9.50903707e-01,
  4.14594510e-01,  5.09884086e-01, -1.47956206e-01, -3.57577438e-01,
  -6.13889572e-01,  2.91620433e-01,  2.52564667e-01,  7.53191524e-01,
  -4.75090695e-01, -8.58270087e-01,  2.17138914e-01, -3.48893797e-01,
  7.79789734e-01, -1.05624977e-02, -4.55902826e-01,  1.25638233e+00,
  3.99309632e-02,  3.90433294e-01,  9.05680254e-01,  6.57185180e-01,
  7.69664829e-02,  2.80154566e-01, -3.53418045e-01,  1.01492796e+00,
  1.14924134e+00,  1.49379140e+00,  6.85418102e-01, -6.17996646e-01,
  -4.14117012e-01,  1.00057796e+00,  4.21838788e-01,  4.81898276e-01,
  -4.52176278e-01, -1.56832187e+00, -3.99127845e-01, -4.78455318e-01,
  -1.43554834e+00,  4.93477452e-01, -1.26256573e-01,  7.37873456e-01,
  1.37187193e+00,  5.14343043e-01,  4.39081716e-01, -9.46178176e-01,
  8.45218751e-02, -4.35648532e-01,  3.03138024e-01, -1.43956281e-01,
```

1.84455953e-01, -5.62922411e-02, 1.39187214e-01, 2.18432244e-01,
-3.20307371e-01, 9.35640652e-01, -5.54888701e-01, 1.33979224e-01,
-6.18494619e-02, 3.10052595e-01, -7.25365948e-02, -1.29178498e+00,
5.45331726e-01, -3.97471322e-01, 5.82499110e-02, 4.41326596e-01,
7.61763570e-02, -1.02465004e+00, -7.61286147e-01, -6.99105823e-01,
-5.86867160e-01, 7.71691979e-02, 1.10842519e-01, -7.65391281e-01,
-1.07011068e+00, 1.70565375e+00, 9.93896454e-01, -1.33885611e+00,
4.95226873e-01, -8.80791760e-01, 2.25733745e-01, -8.64387517e-01,
5.73935212e-01, 1.06568176e+00, -4.86309988e-01, -9.03177726e-01,
-2.94530111e-01, -1.34621176e+00, 1.83718562e-01, -1.15727426e+00,
6.84351546e-01, -1.82575808e-01, -4.16944539e-01, 3.04060745e-01,
1.14956765e+00, -5.05775025e-01, 1.24008165e+00, -6.87761147e-01,
-1.06666774e+00, -1.43871511e-01, 1.05306575e-01, -1.87133307e-01,
1.29314278e+00, -1.61974828e-01, -2.94797057e-01, 1.54330767e+00,
-1.17416431e+00, 1.36973994e-01, -1.02122640e+00, 2.87472275e-01,
-7.54991129e-01, -8.34353060e-01, -3.93749000e-01, -7.41586630e-01,
1.46648960e+00, 3.97235178e-01, -5.84187267e-01, 1.61417892e+00,
-2.34570589e-01, -7.86593012e-01, 1.37361522e+00, 3.38218054e-01,
-5.12748690e-01, 2.62041467e-02, -1.20511233e+00, -9.18853191e-01,
-1.19056341e-01, 3.67531912e-01, 1.39583081e+00, -1.19139903e+00,
1.15699142e+00, -8.34698545e-01, 1.09113633e-01, 7.97462823e-02,
-4.56431454e-01, 8.35866518e-01, -6.20435525e-01, -7.77701771e-02,
-4.86777331e-02, 1.86430182e+00, 6.26488106e-01, 1.00274352e+00,
-4.38927555e-01, 1.19099708e+00, 3.42789118e-01, 4.16028141e-01,
7.23536711e-01, -2.24851239e-02, -7.74904433e-02, 1.24881021e+00,
1.58608128e+00, -3.31844520e-01, 6.44481377e-01, 1.00750724e+00,
1.20711223e+00, -8.67386056e-01, -1.13603610e+00, 9.27227638e-01,
-1.04347732e+00, 9.03092404e-01, 6.02533706e-01, -2.96327493e-01,
-1.26654795e-01, -4.56197934e-01, 6.81283846e-02, -1.07390393e+00,
9.85065078e-02, 5.09162194e-01, 2.80196597e-01, 9.79041548e-01,
-9.29561481e-01, -1.44688720e+00, 1.29161257e+00, 4.17043830e-01,
-1.79933363e-01, -3.02136400e-01, -3.30133120e-01, -3.16723869e-01,
-4.46036518e-01, 1.46017496e-01, -6.20291819e-01, -9.99010703e-01,
-7.51183889e-01, 4.74409513e-01, -1.17901105e+00, 4.72051405e-01,
-2.97904319e-01, -1.92424963e-01, -3.22395298e-01, 2.65475829e-01,
5.37197340e-01, -6.18440406e-01, -2.28016595e-01, 1.65411659e-01,
9.98342045e-02, 1.84804010e-01, -4.72560549e-01, -7.37160163e-01,
-1.10806197e+00, 7.87378373e-01, -6.06379275e-01, -1.00583633e+00,
-2.07509473e-01, 4.06170436e-01, 3.87368411e-01, -8.98350788e-01,
8.35660059e-01, -9.56720175e-01, 3.72063159e-01, -3.25721789e-01,
-1.01237368e-01, 3.26774195e-01, -4.71307180e-01, 6.22494465e-02,
-6.45849453e-01, -7.95321927e-01, 3.41733892e-01, 1.19230117e+00,

```
-1.05758840e-01, 3.21046804e-01, -1.79579196e-01, -5.69422591e-01,
-4.39265709e-02, 4.24549076e-01, -6.37480338e-01, 6.74472977e-01,
-1.41099083e+00, -4.74327477e-01, -1.99349011e-02, 1.07368266e+00,
-1.07937859e+00, 5.43466294e-01, -2.49380800e-01, -4.27856905e-01,
-1.39715162e-01])
```

```
In [159... df.fillna = df.fillna(df.mean())
# Check for missing values
missing_values = df.isnull().sum()
print("Missing values in each column:\n", missing_values)
```

Missing values in each column:

```
age      0
sex      0
bmi      0
bp       0
s1       0
s2       0
s3       0
s4       0
s5       0
s6       0
target    0
dtype: int64
```

```
In [160... df["target"].describe()
```

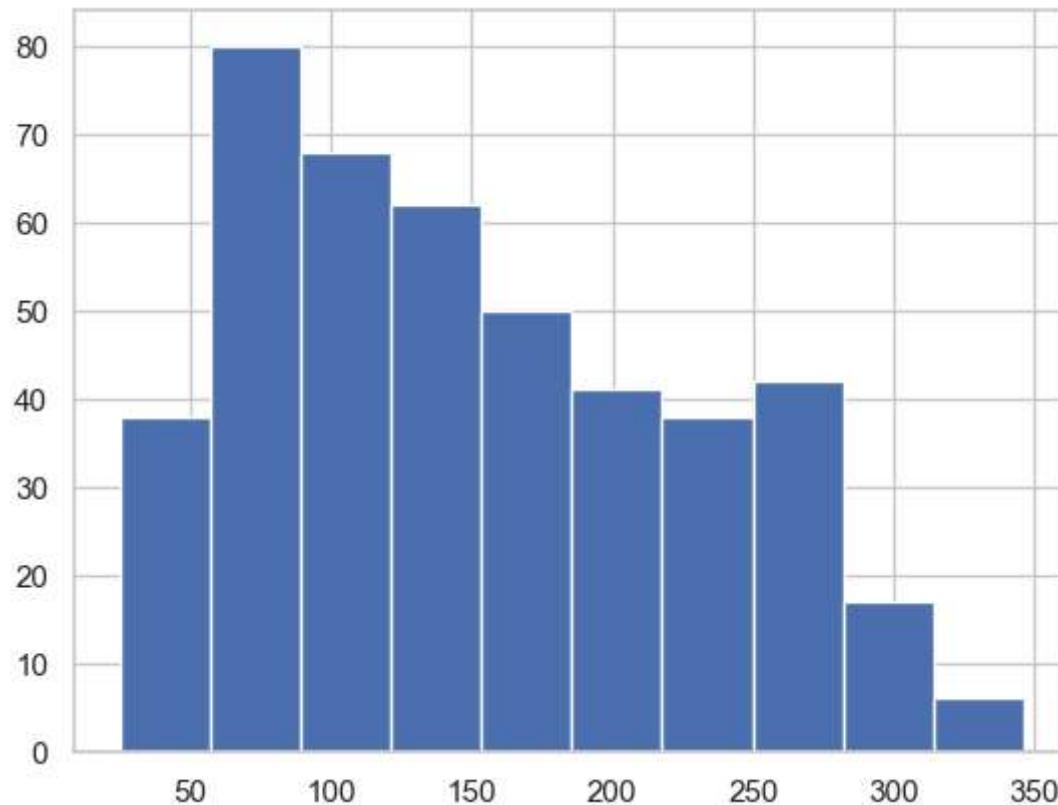
```
Out[160... count    442.000000
mean     152.133484
std      77.093005
min      25.000000
25%     87.000000
50%     140.500000
75%     211.500000
max     346.000000
Name: target, dtype: float64
```

```
In [161... df.nunique()
```

```
Out[161... age      58  
       sex      2  
       bmi     163  
       bp     100  
       s1     141  
       s2     302  
       s3      63  
       s4      66  
       s5    184  
       s6      56  
       target   214  
       dtype: int64
```

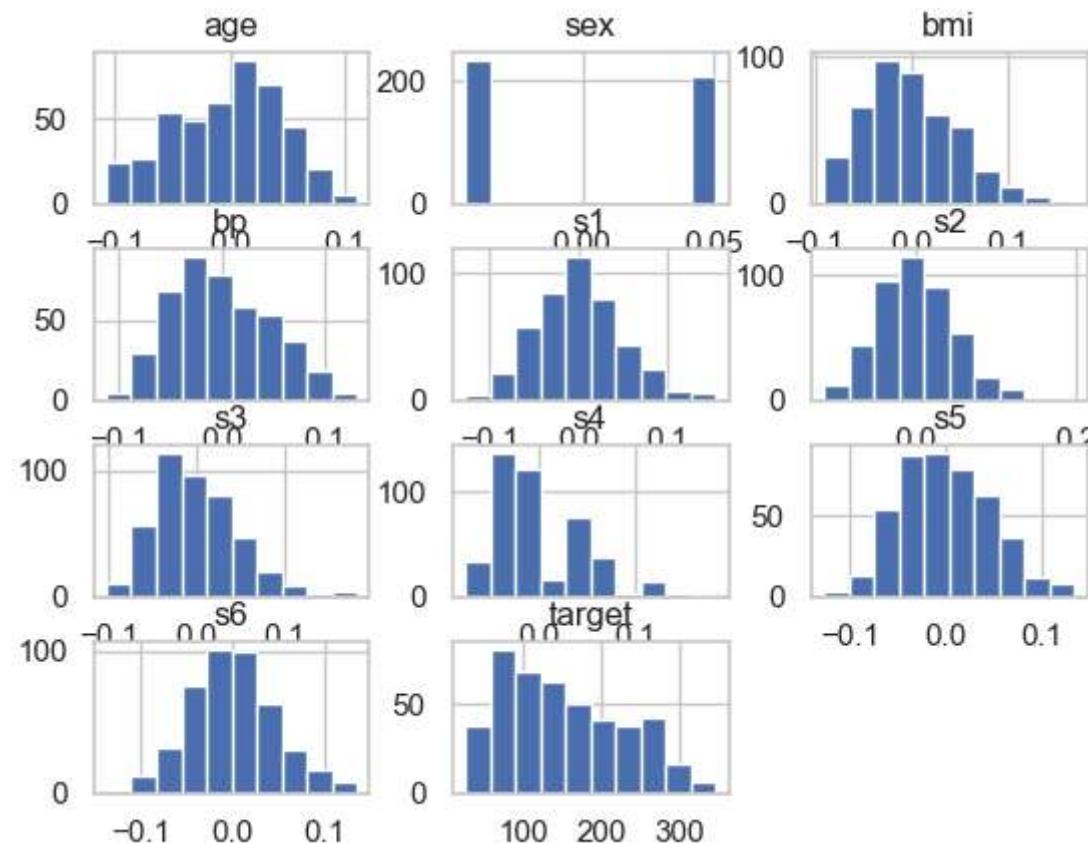
```
In [162... df['target'].hist()
```

```
Out[162... <Axes: >
```



```
In [163... df.hist()
```

```
Out[163... array([[[<Axes: title={'center': 'age'}>, <Axes: title={'center': 'sex'}>,
   <Axes: title={'center': 'bmi'}>],
  [<Axes: title={'center': 'bp'}>, <Axes: title={'center': 's1'}>,
   <Axes: title={'center': 's2'}>],
  [<Axes: title={'center': 's3'}>, <Axes: title={'center': 's4'}>,
   <Axes: title={'center': 's5'}>],
  [<Axes: title={'center': 's6'}>,
   <Axes: title={'center': 'target'}>], dtype=object)
```



```
In [164... df.corr()
```

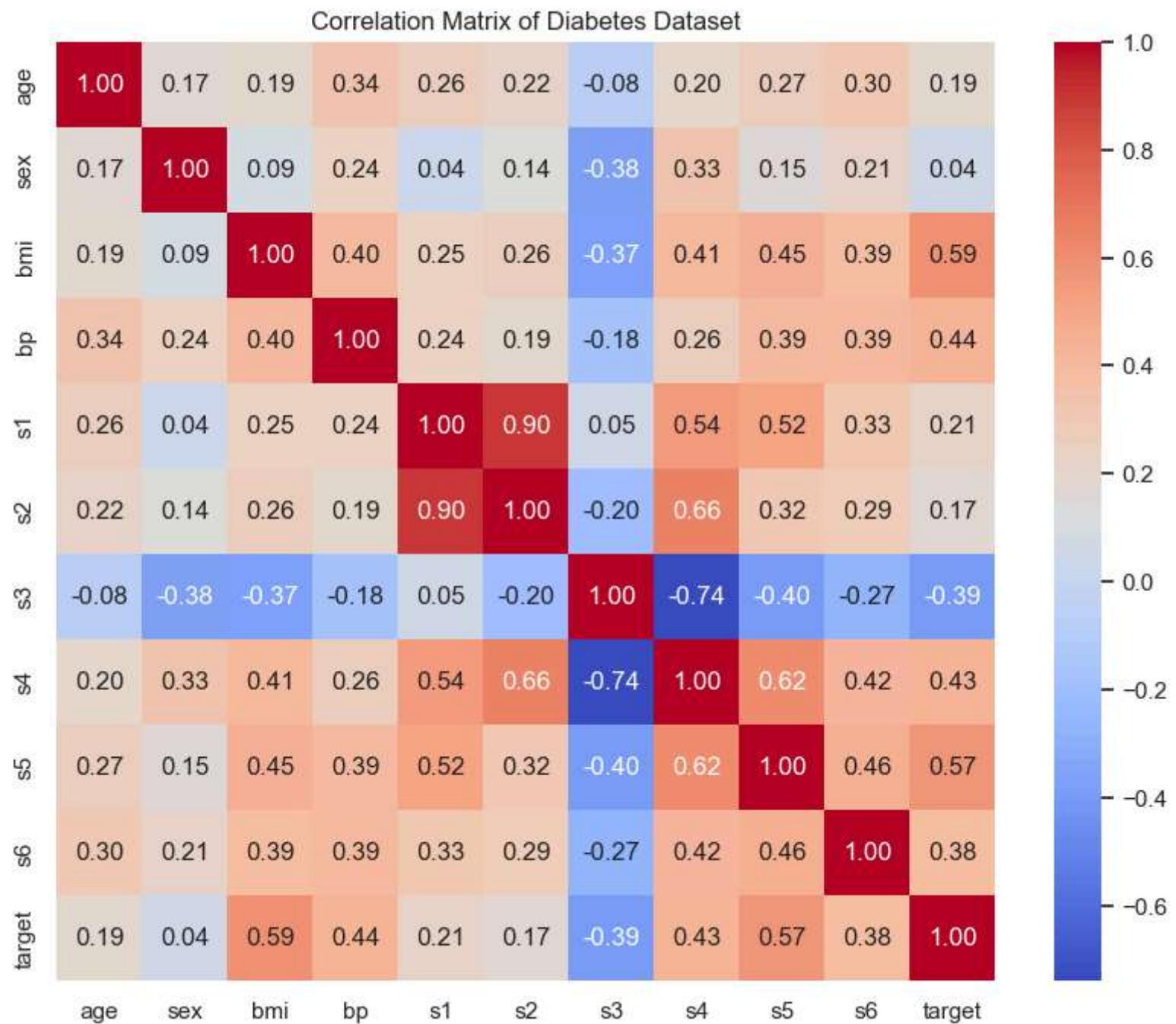
Out[164...]

| | age | sex | bmi | bp | s1 | s2 | s3 | s4 | s5 | s6 | target |
|---------------|------------|------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------------|
| age | 1.000000 | 0.173737 | 0.185085 | 0.335428 | 0.260061 | 0.219243 | -0.075181 | 0.203841 | 0.270774 | 0.301731 | 0.187889 |
| sex | 0.173737 | 1.000000 | 0.088161 | 0.241010 | 0.035277 | 0.142637 | -0.379090 | 0.332115 | 0.149916 | 0.208133 | 0.043062 |
| bmi | 0.185085 | 0.088161 | 1.000000 | 0.395411 | 0.249777 | 0.261170 | -0.366811 | 0.413807 | 0.446157 | 0.388680 | 0.586450 |
| bp | 0.335428 | 0.241010 | 0.395411 | 1.000000 | 0.242464 | 0.185548 | -0.178762 | 0.257650 | 0.393480 | 0.390430 | 0.441482 |
| s1 | 0.260061 | 0.035277 | 0.249777 | 0.242464 | 1.000000 | 0.896663 | 0.051519 | 0.542207 | 0.515503 | 0.325717 | 0.212022 |
| s2 | 0.219243 | 0.142637 | 0.261170 | 0.185548 | 0.896663 | 1.000000 | -0.196455 | 0.659817 | 0.318357 | 0.290600 | 0.174054 |
| s3 | -0.075181 | -0.379090 | -0.366811 | -0.178762 | 0.051519 | -0.196455 | 1.000000 | -0.738493 | -0.398577 | -0.273697 | -0.394789 |
| s4 | 0.203841 | 0.332115 | 0.413807 | 0.257650 | 0.542207 | 0.659817 | -0.738493 | 1.000000 | 0.617859 | 0.417212 | 0.430453 |
| s5 | 0.270774 | 0.149916 | 0.446157 | 0.393480 | 0.515503 | 0.318357 | -0.398577 | 0.617859 | 1.000000 | 0.464669 | 0.565883 |
| s6 | 0.301731 | 0.208133 | 0.388680 | 0.390430 | 0.325717 | 0.290600 | -0.273697 | 0.417212 | 0.464669 | 1.000000 | 0.382483 |
| target | 0.187889 | 0.043062 | 0.586450 | 0.441482 | 0.212022 | 0.174054 | -0.394789 | 0.430453 | 0.565883 | 0.382483 | 1.000000 |

In [165...]

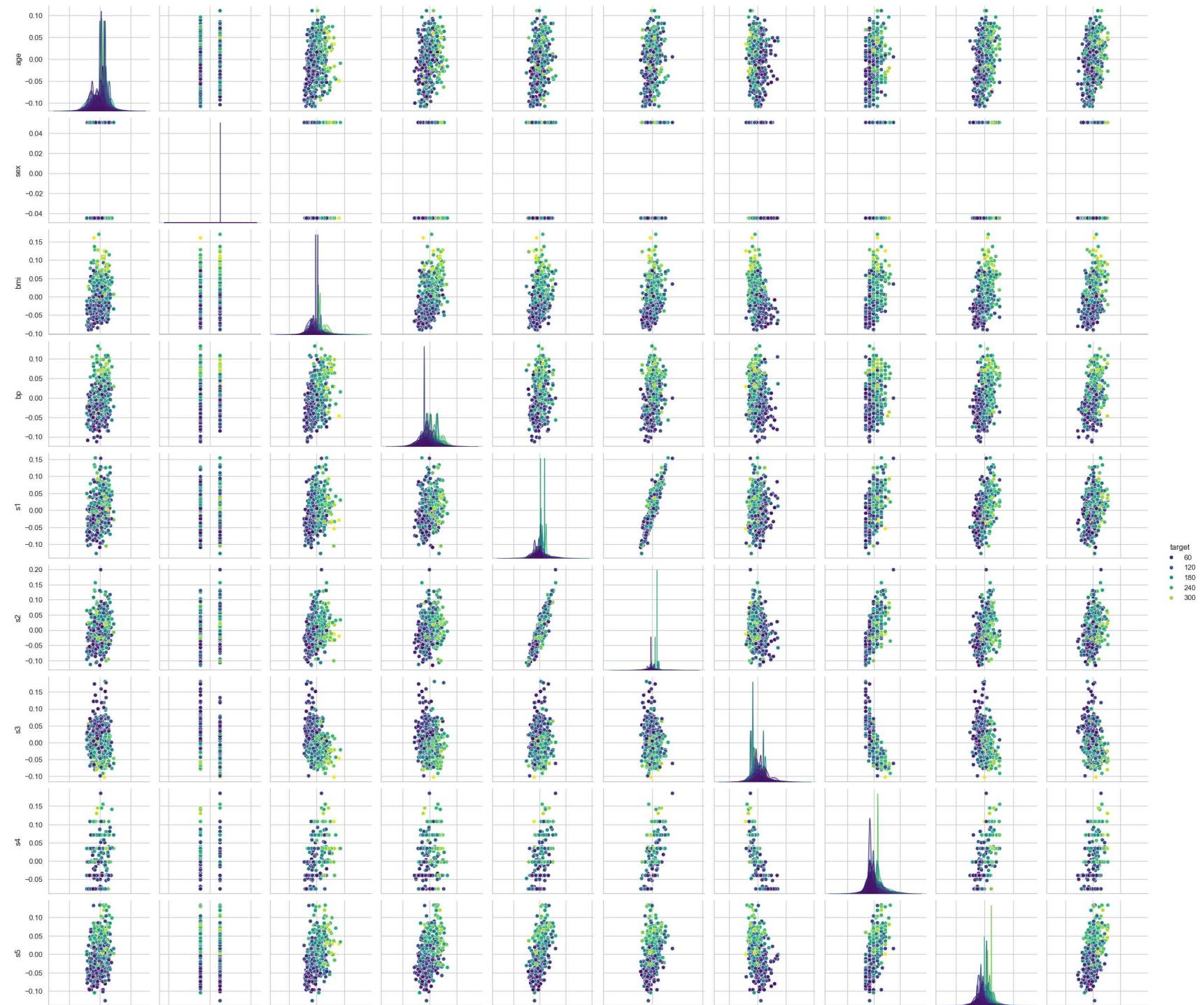
```
plt.figure(figsize=(10, 8))
sns.heatmap(df.corr(), annot=True, fmt=".2f", cmap="coolwarm", square=True)
plt.title("Correlation Matrix of Diabetes Dataset")
```

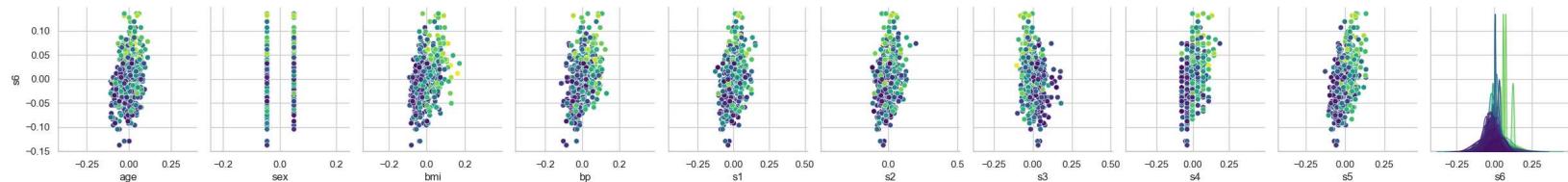
Out[165...]: Text(0.5, 1.0, 'Correlation Matrix of Diabetes Dataset')



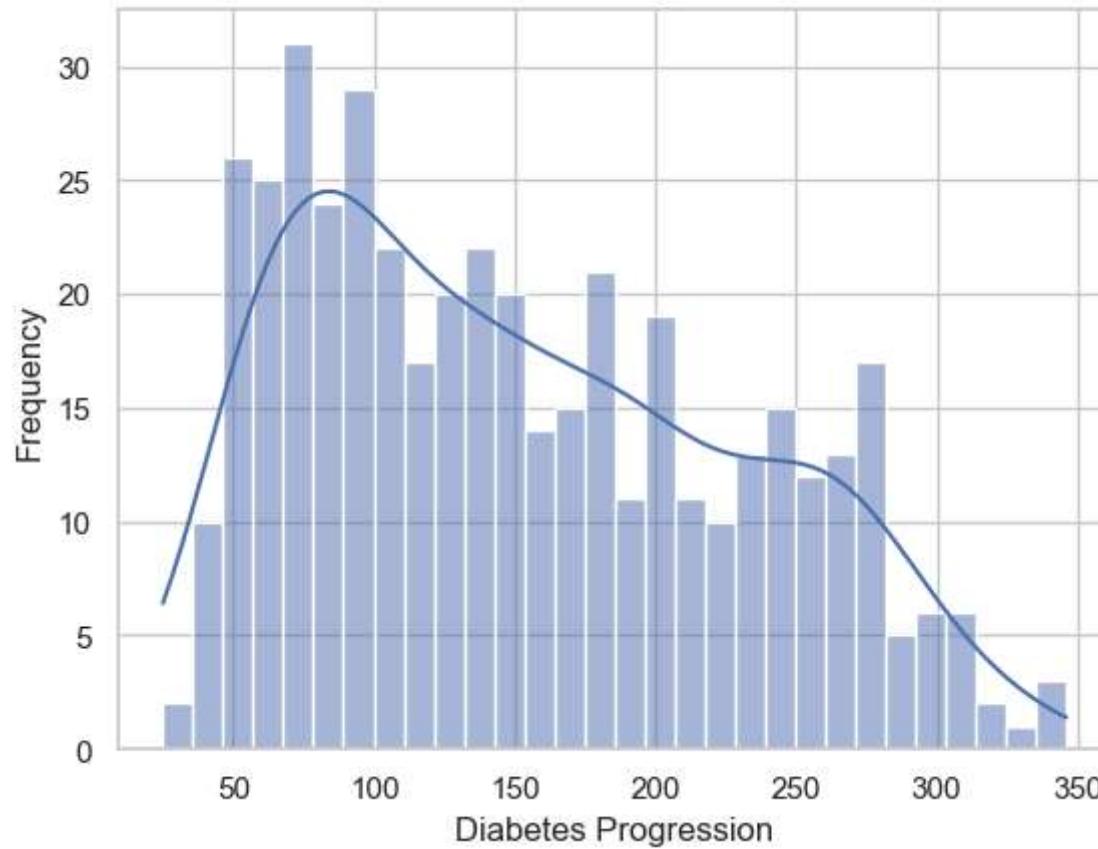
In [166...]

```
sns.pairplot(df, diag_kind='kde', markers='o', hue='target', palette='viridis')
# Show the pairplot
plt.show()
# Visualize the distribution of the target variable
sns.histplot(df['target'], kde=True, bins=30)
# Show the distribution plot plt.title("Distribution of Target Variable")
plt.title("Distribution of Target Variable")
plt.xlabel("Diabetes Progression")
plt.ylabel("Frequency")
plt.show()
```





Distribution of Target Variable

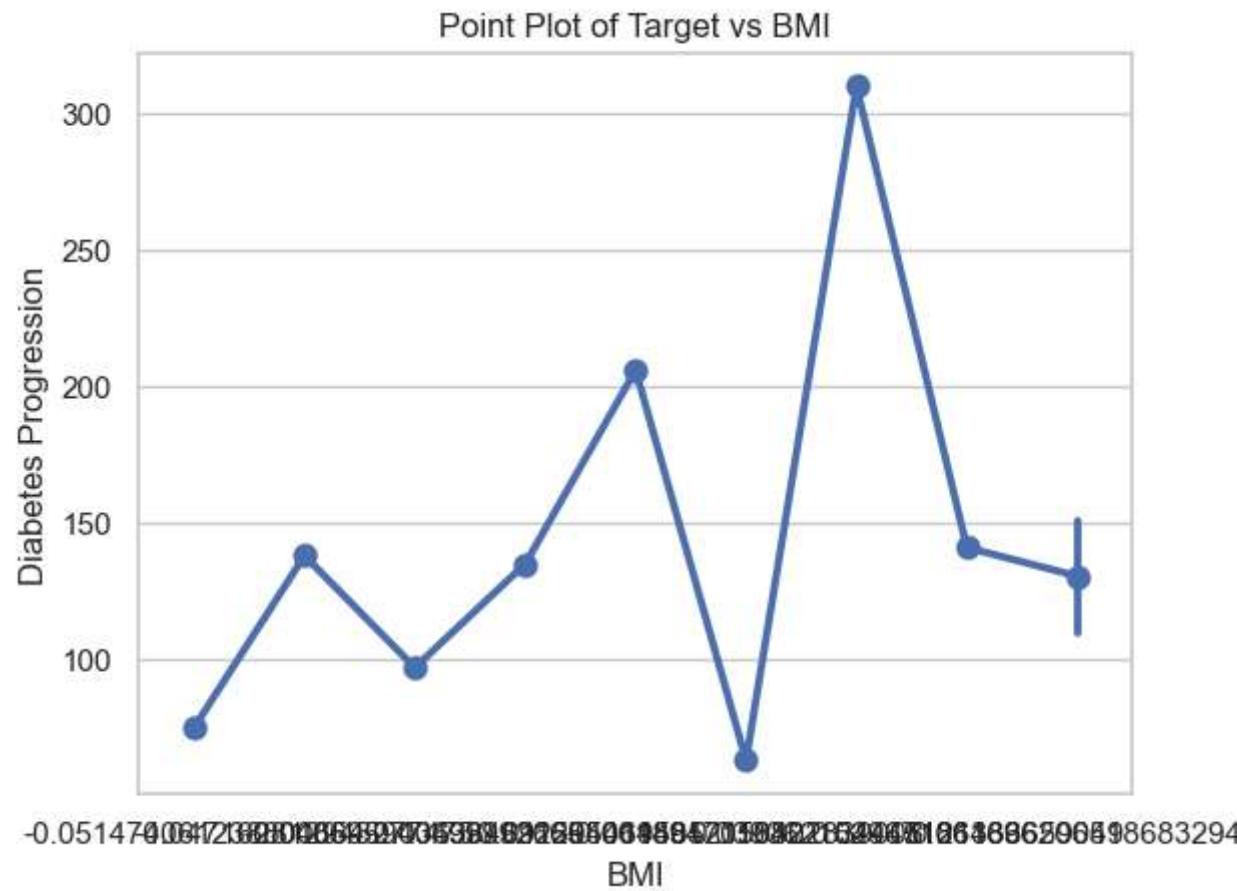


In [167...]

```

sns.pointplot(data=df.head(10), x='bmi', y='target')
# Show the point plot
plt.title("Point Plot of Target vs BMI")
plt.xlabel("BMI")
plt.ylabel("Diabetes Progression")
plt.show()

```



```
In [168]: df["target"].describe()
```

```
Out[168]: count    442.000000
mean     152.133484
std      77.093005
min      25.000000
25%     87.000000
50%    140.500000
75%    211.500000
max    346.000000
Name: target, dtype: float64
```

```
In [169]: df.skew()
```

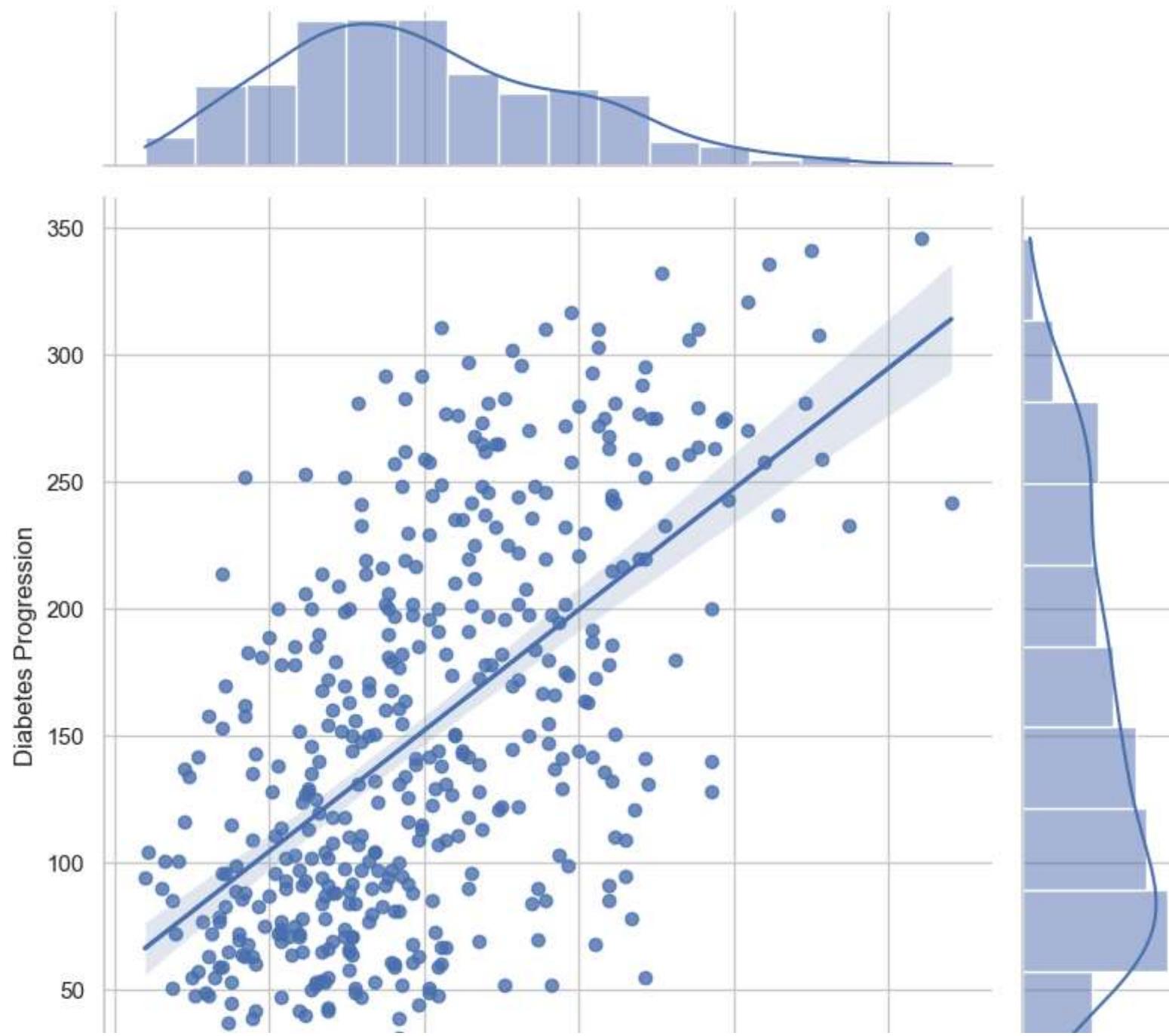
```
Out[169...]: age      -0.231382
              sex       0.127385
              bmi       0.598148
              bp        0.290658
              s1        0.378108
              s2        0.436592
              s3        0.799255
              s4        0.735374
              s5        0.291754
              s6        0.207917
              target     0.440563
              dtype: float64
```

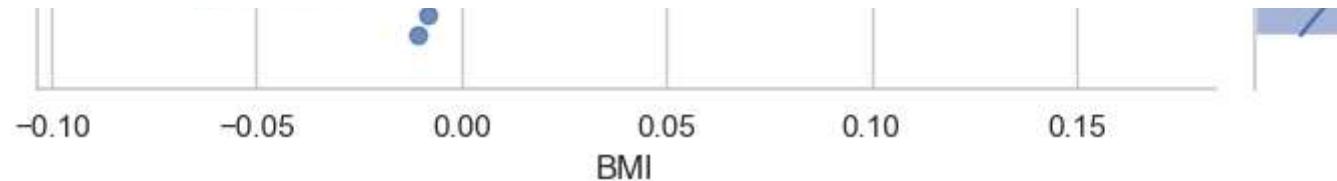
```
In [170...]: df.std()
```

```
Out[170...]: age      0.047619
              sex       0.047619
              bmi       0.047619
              bp        0.047619
              s1        0.047619
              s2        0.047619
              s3        0.047619
              s4        0.047619
              s5        0.047619
              s6        0.047619
              target    77.093005
              dtype: float64
```

```
In [171...]: sns.jointplot(data=df, x='bmi', y='target', kind='reg', height=8)
plt.suptitle("Joint Plot of Target vs BMI", y=1.02)
plt.xlabel("BMI")
plt.ylabel("Diabetes Progression")
plt.show()
```

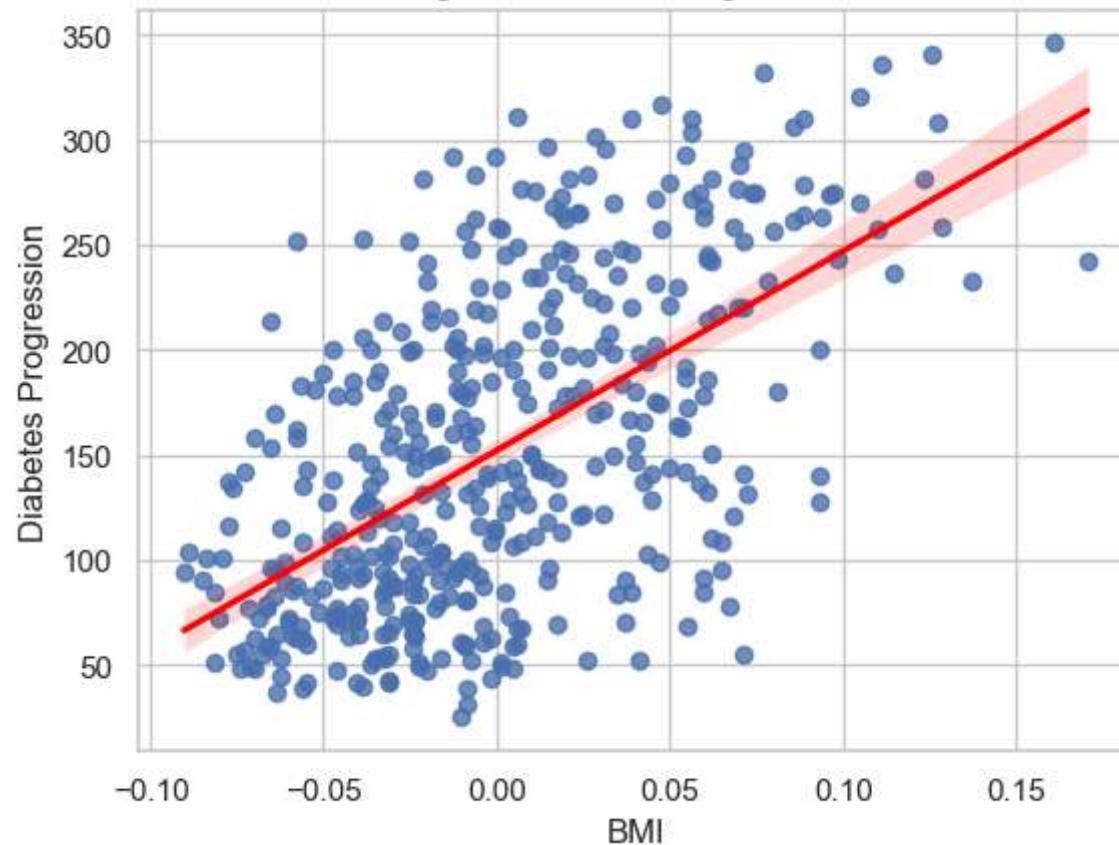
Joint Plot of Target vs BMI

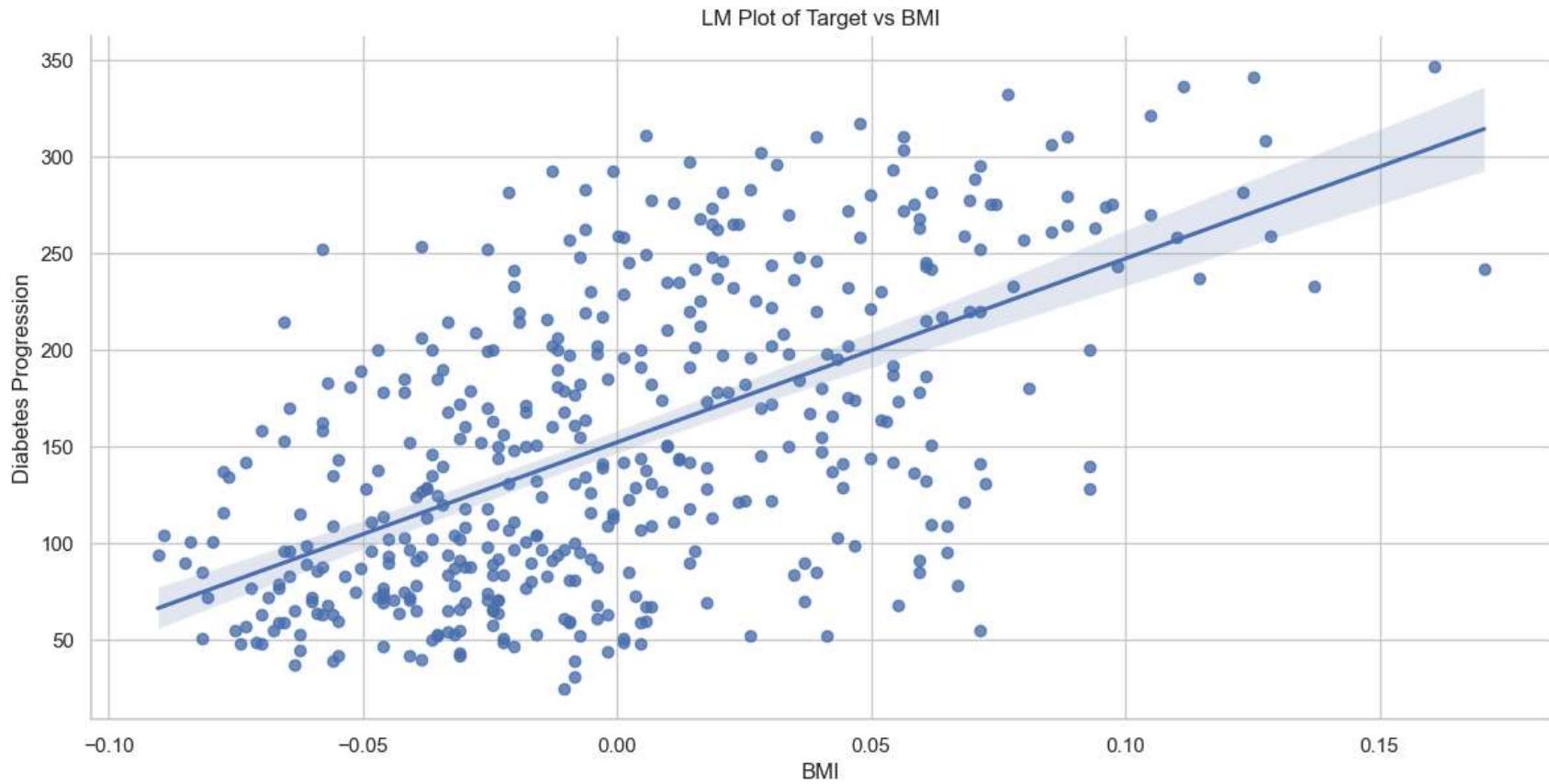




```
In [172]:  
sns.regplot(data=df, x='bmi', y='target', line_kws={'color': 'red'})  
plt.title("Regression Plot of Target vs BMI")  
plt.xlabel("BMI")  
plt.ylabel("Diabetes Progression")  
plt.show()  
sns.lmplot(data=df, x='bmi', y='target', aspect=2, height=6)  
plt.title("LM Plot of Target vs BMI")  
plt.xlabel("BMI")  
plt.ylabel("Diabetes Progression")  
plt.show()
```

Regression Plot of Target vs BMI



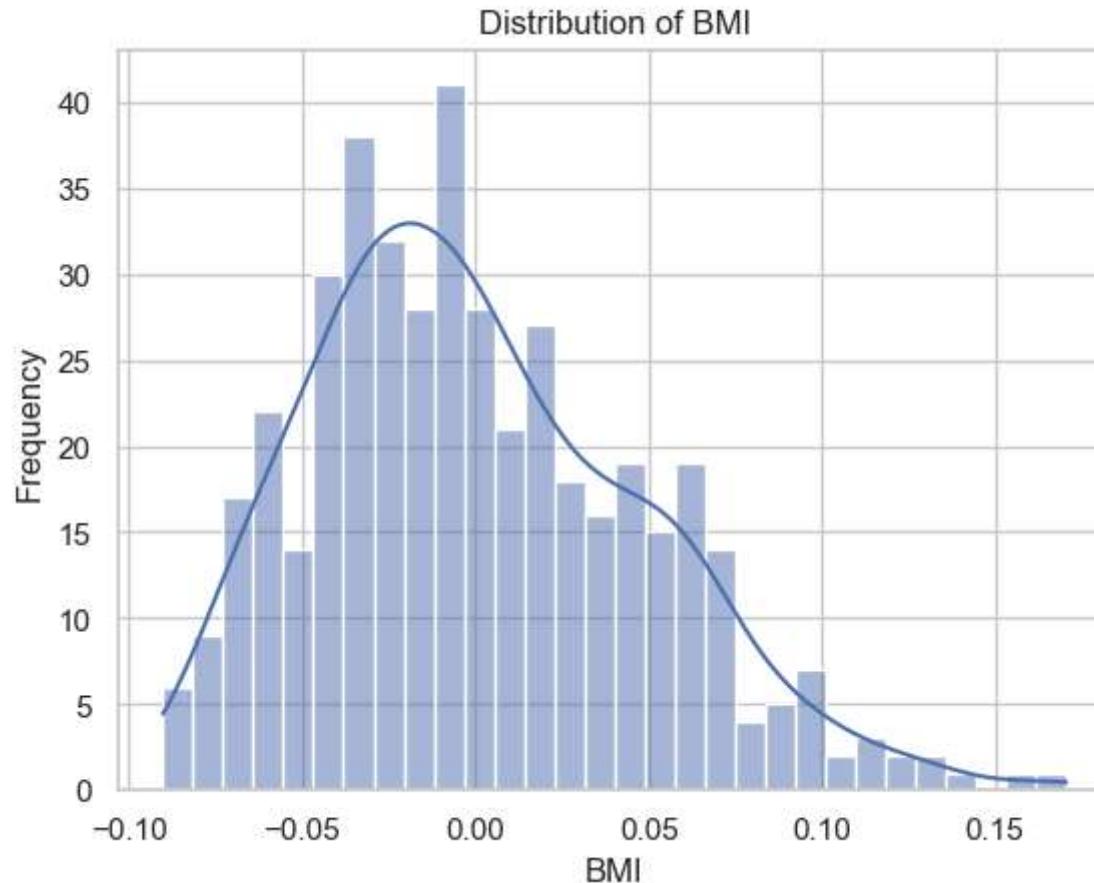


In [173]:

```
from scipy import stats
# Calculate the Pearson correlation coefficient
pearson_corr, p_value = stats.pearsonr(df['bmi'], df['target'])
print(f"Pearson correlation coefficient: {pearson_corr:.2f}, p-value: {p_value:.2e}")
# Calculate the Spearman correlation coefficient
spearman_corr, p_value_spearman = stats.spearmanr(df['bmi'], df['target'])
print(f"Spearman correlation coefficient: {spearman_corr:.2f}, p-value: {p_value_spearman:.2e}")
# Calculate the Kendall correlation coefficient
kendall_corr, p_value_kendall = stats.kendalltau(df['bmi'], df['target'])
print(f"Kendall correlation coefficient: {kendall_corr:.2f}, p-value: {p_value_kendall:.2e}")
# Visualize the distribution of BMI
sns.histplot(df['bmi'], kde=True, bins=30)
plt.title("Distribution of BMI")
plt.xlabel("BMI")
```

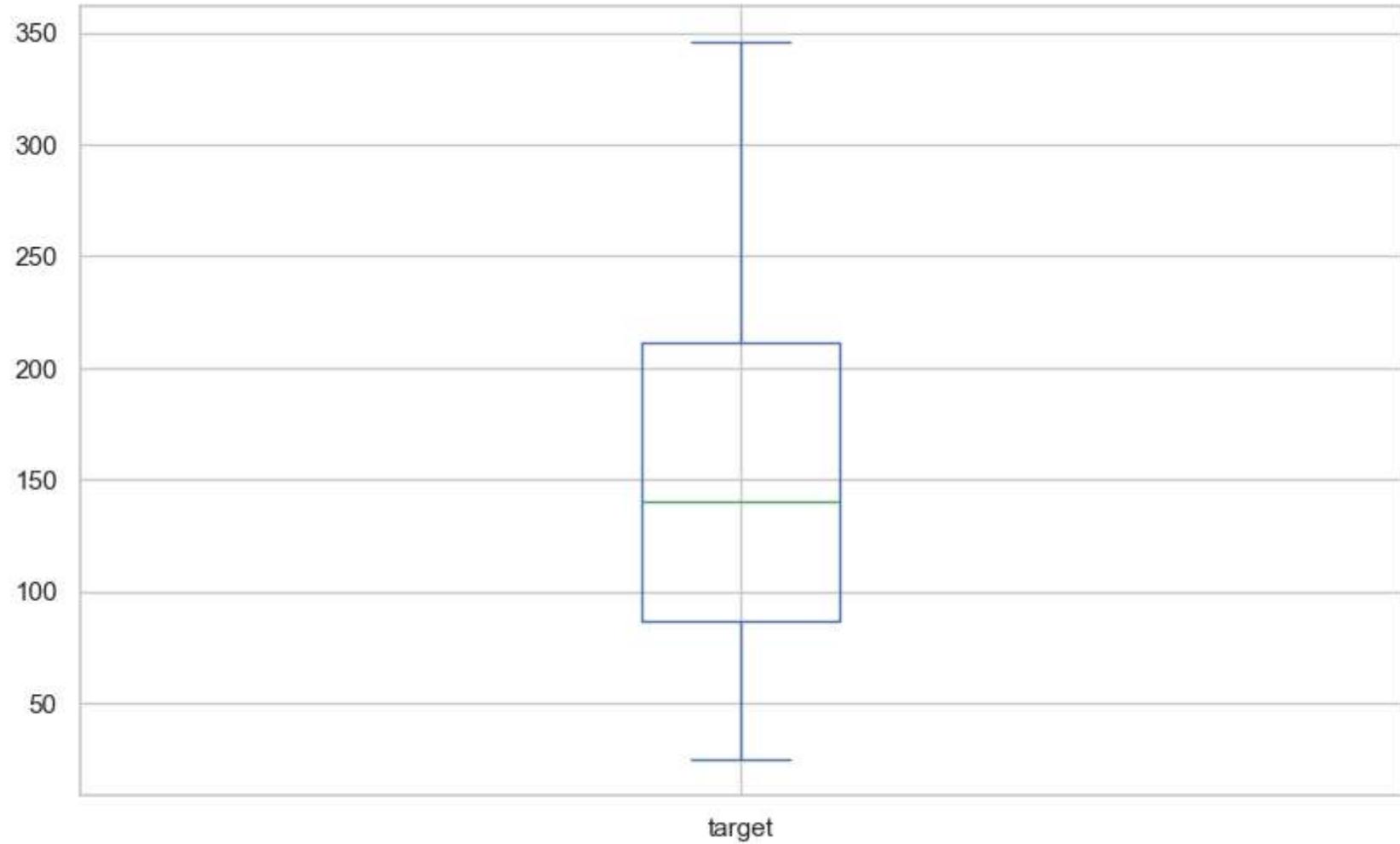
```
plt.ylabel("Frequency")
plt.show()
```

Pearson correlation coefficient: 0.59, p-value: 3.47e-42
Spearman correlation coefficient: 0.56, p-value: 4.57e-38
Kendall correlation coefficient: 0.39, p-value: 2.24e-34



In [174]: df['target'].plot(kind='box', figsize=(10, 6))

Out[174]: <Axes: >



In [175]:

y_pred_lr

```
Out[175... array([-0.16344139,  0.35560625, -0.23497894,  1.80874224, -0.36807416,
   -0.77865797,  1.37780504,  0.3792423 , -0.80395781, -0.56488935,
   -0.75312337,  0.21169043, -1.28079657,  0.70770475, -0.67532916,
   -0.27877028,  0.8752234 ,  1.28103219,  0.5744422 ,  0.86281398,
   0.71705062, -0.82656267, -1.06096791,  0.47821962,  0.03575489,
   0.09386594,  0.46982402,  0.36695257, -1.35240582, -0.56046396,
   0.29407536, -0.85408455, -0.24901872,  0.420809 ,  0.2817806 ,
   0.49639284, -0.35994543, -0.42181758, -0.054305 , -1.20873359,
  -1.04550855, -0.57723793,  0.17297632,  0.03735144,  0.24562462,
  -1.17752084, -1.04494486, -0.48264141, -1.30579514,  0.20053713,
  0.00505723, -1.15806682, -0.6315901 , -0.55743852,  0.30527203,
  0.03206875, -0.74950823,  0.76109395, -0.41395435, -0.96768922,
  0.46488162,  0.70591993, -0.14936305, -0.60432768, -0.27828 ,
  0.64997567,  0.24669472,  0.16597133, -0.35593134, -0.09509908,
  0.38780055,  0.61398291,  1.06590589, -0.08021257, -0.93845531,
  0.067994 ,  0.52737157,  0.73714832,  0.08380857,  0.6997981 ,
  -0.57987579, -0.14541116, -1.26370039, -1.2493581 , -0.48206382,
  -0.95028808, -0.91645903, -1.2694314 ,  0.18337421])
```

```
In [179... (y_test - y_pred_lr).std()
```

```
Out[179... np.float64(0.697494831405442)
```

```
In [177... from sklearn.model_selection import cross_val_score

# Perform 5-fold cross-validation for the linear regression model
cv_scores = cross_val_score(lr, X, y, cv=5, scoring='neg_mean_squared_error')
cv_mse = -cv_scores
print("Cross-validated MSE scores:", cv_mse)
print("Mean cross-validated MSE:", cv_mse.mean())
```

```
Cross-validated MSE scores: [0.46879889 0.5107749  0.54599501 0.507387   0.49077052]
Mean cross-validated MSE: 0.5047452627670396
```

```
In [178... from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score, median_absolute_error, mean_absolute_p

mse = mean_squared_error(y_test, y_pred_lr)
rmse = np.sqrt(mean_squared_error(y_test, y_pred_lr)) # Correct RMSE
mae = mean_absolute_error(y_test, y_pred_lr)
medae = median_absolute_error(y_test, y_pred_lr)
mape = mean_absolute_percentage_error(y_test, y_pred_lr)
r2 = r2_score(y_test, y_pred_lr)
```

```
print(f"Mean Squared Error: {mse:.2f}")
print(f"Root Mean Squared Error: {rmse:.2f}")
print(f"Mean Absolute Error: {mae:.2f}")
print(f"Median Absolute Error: {medae:.2f}")
print(f"Mean Absolute Percentage Error: {mape:.2f}")
print(f"R-squared: {r2:.2f}")
```

Mean Squared Error: 0.49
Root Mean Squared Error: 0.70
Mean Absolute Error: 0.56
Median Absolute Error: 0.49
Mean Absolute Percentage Error: 2.04
R-squared: 0.45