In [2]:
```
import pandas as pd
```

In [3]:
```
import numpy as np
```

In [4]:
```
import random
```

In [5]:
```
import matplotlib.pyplot as plt
```

In [6]:
```
import statistics
```

In [8]:
```
import seaborn as sb
```

In [27]:
```
from sklearn.model_selection import train_test_split
```

In [33]:
```
from sklearn.linear_model import LinearRegression
```

In [34]:
```
df=pd.read_csv("BostonHousing.csv")
```

In [35]:
```
df.head(50).isnull().sum()
```

Out[35]:
```
crim       0
zn         0
indus      0
chas       0
nox        0
rm         0
age        0
dis        0
rad        0
tax        0
ptratio    0
b          0
lstat      0
medv       0
dtype: int64
```

In [36]:
```
df
```

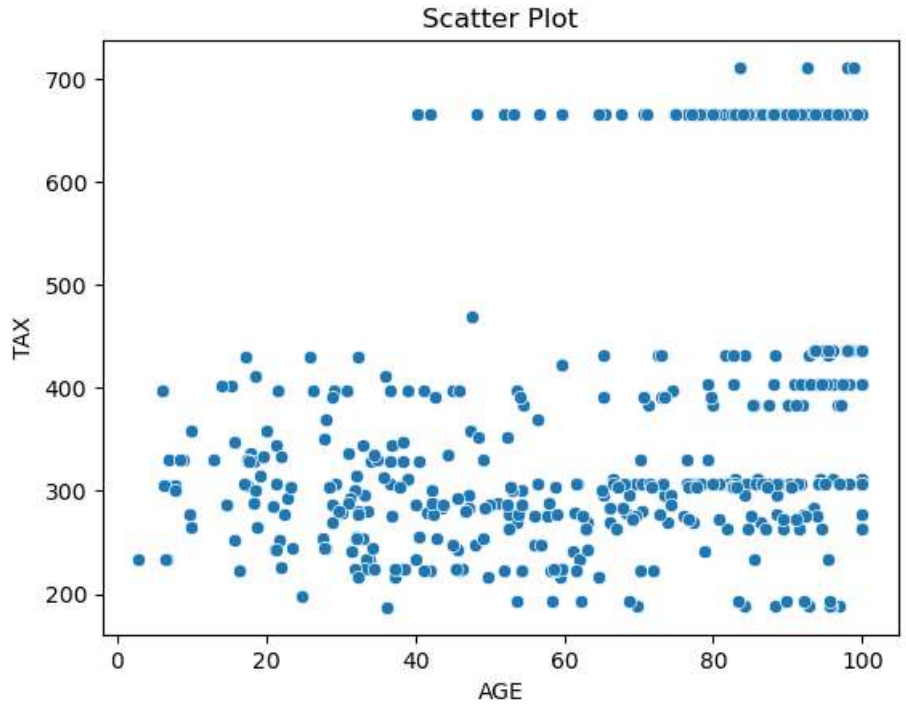
Out[36]:

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
0	0.00632	18.0	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	396.90	5.33	36.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
501	0.06263	0.0	11.93	0	0.573	6.593	69.1	2.4786	1	273	21.0	391.99	9.67	22.4
502	0.04527	0.0	11.93	0	0.573	6.120	76.7	2.2875	1	273	21.0	396.90	9.08	20.6
503	0.06076	0.0	11.93	0	0.573	6.976	91.0	2.1675	1	273	21.0	396.90	5.64	23.9
504	0.10959	0.0	11.93	0	0.573	6.794	89.3	2.3889	1	273	21.0	393.45	6.48	22.0
505	0.04741	0.0	11.93	0	0.573	6.030	80.8	2.5050	1	273	21.0	396.90	7.88	11.9

506 rows × 14 columns

In [46]:
```
x=df["age"]
y=df["tax"]
sb.scatterplot(x,y,data=df)
plt.title("Scatter Plot")
plt.xlabel("AGE")
plt.ylabel("TAX")
```

Out[46]:
```
Text(0, 0.5, 'TAX')
```



In [47]:
```
x=df[["age"]]
y=df["tax"]
```

In [48]:
```
x_test,x_train,y_test,y_train=train_test_split(x,y,test_size=0.25)
```

In [49]:
```
regr=LinearRegression()
```

In [50]:
```
regr.fit(x_train,y_train)
```

Out[50]:
```
LinearRegression()
```

In [51]:
```
y_pred=regr.predict(x_test)
```

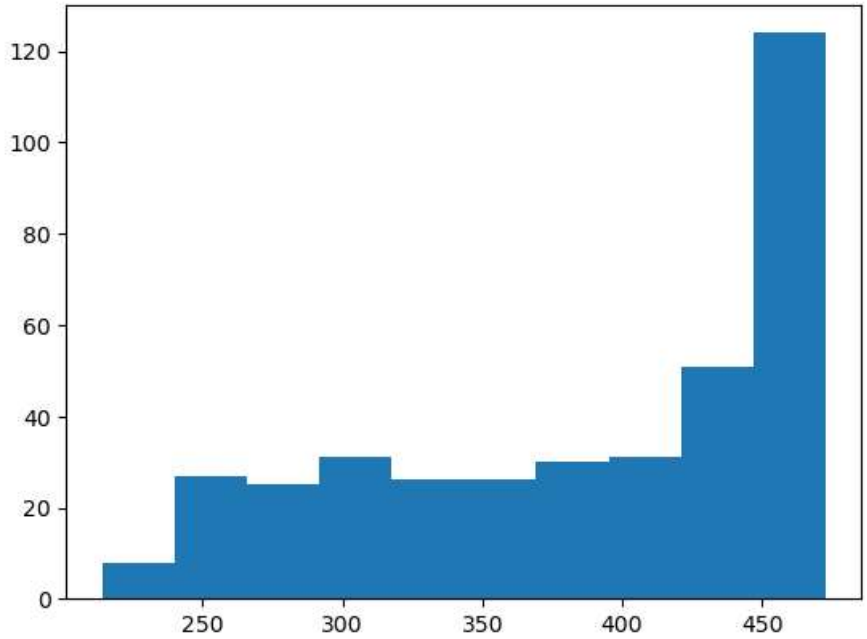
In [52]:
```
y_pred
```

Out[52]:
```
array([472.62722516, 440.4156027 , 263.65099683, 283.084455  ,
       370.13569916, 370.13569916, 369.86948741, 456.92073156,
       361.08449946, 440.4156027 , 303.84897196, 319.02304204,
       249.80798552, 295.06398402, 410.06746254, 440.4156027 ,
       472.62722516, 314.23123044, 465.173296  , 467.83541355,
       453.72619049, 418.85245048, 424.70910911, 279.88991393,
       444.14256728, 430.03334422, 414.32685063, 394.36096894,
       430.56576774, 467.03677829, 294.79777226, 443.34393202,
       433.76030881, 472.62722516, 251.67146782, 338.72271197,
       455.58967279, 472.62722516, 472.62722516, 268.44280843,
       386.90703978, 460.11527263, 330.7363593 , 463.57602546,
       380.25174588, 348.04012343, 370.93433443, 322.75000662,
       344.5793706 , 472.62722516, 452.39513172, 449.73301416,
       308.90699532, 472.62722516, 297.45988982, 413.26200361,
       362.68177   , 430.29955598, 464.64087248, 461.97875493,
       335.5281709 , 318.75683029, 327.80802998, 319.82167731,
       459.31663737, 431.09819125, 463.04360195, 457.98557859,
       455.58967279, 464.64087248, 445.47362606, 428.70228545,
       401.01626284, 307.8421483 , 472.62722516, 391.43263963,
       328.07424174, 404.47701566, 336.85922968, 298.25852509,
       292.1356547 , 403.14595689, 382.38143993, 461.97875493,
       472.62722516, 472.62722516, 420.18350926, 458.25179034,
       356.82511137, 391.69885138, 404.21080391, 280.15612569,
       301.98548967, 320.35410082, 255.66464415, 334.72953564,
       442.81150851, 243.41890338, 223.98544521, 382.64765169,
       462.51117844, 355.49405259, 472.62722516, 472.62722516,
       383.44628695, 465.43950775, 472.62722516, 349.63739396,
       301.45306616, 356.55889961, 432.96167354, 455.05724927,
       449.20059065, 458.5180021 , 332.59984159, 448.93437889,
       386.37461627, 295.59640753, 467.83541355, 222.38817467,
       410.06746254, 281.22097271, 255.3984324 , 385.30976924,
       263.38478507, 468.10162531, 472.62722516, 434.55894407,
       228.77725681, 421.51456804, 284.14930203, 350.70224099,
       291.33701943, 339.25513549, 253.00252659, 261.78751454,
       307.04351303, 427.10501491, 360.2858642 , 346.17664114,
       253.00252659, 472.62722516, 292.1356547 , 472.62722516,
       460.38148439, 450.53164942, 263.38478507, 429.23470896,
       268.70902019, 440.68181446, 450.79786118, 285.48036081,
       291.33701943, 462.24496668, 463.57602546, 435.6237911 ,
       423.91047384, 472.62722516, 233.03664491, 459.31663737,
       472.62722516, 436.68863812, 411.13230956, 453.19376698,
       472.62722516, 460.9139079 , 328.3404535 , 380.51795764,
       286.81141959, 387.9718868 , 379.98553413, 364.81146405,
       404.74322742, 223.98544521, 456.38830805, 256.46327942,
       284.41551378, 253.80116186, 346.44285289, 283.35066676,
       439.08454392, 432.96167354, 232.50422139, 472.62722516,
       465.43950775, 443.61014377, 467.83541355, 445.73983782,
       319.2892538 , 457.71936683, 446.80468484, 472.62722516,
       472.62722516, 303.84897196, 447.86953187, 373.59645199,
       294.26534875, 305.71245425, 472.62722516, 430.29955598,
       442.01287324, 440.68181446, 365.34388756, 410.59988605,
       326.4769712 , 391.96506314, 422.84562682, 428.16986193,
       416.45654468, 450.79786118, 255.3984324 , 462.51117844,
       401.54868635, 342.1834648 , 227.17998628, 337.12544144,
       304.38139547, 399.4189923 , 346.9752764 , 470.76374287,
       415.92412117, 327.27560647, 288.40869012, 264.71584385,
       310.2380541 , 290.27217241, 289.20732539, 424.1766856 ,
       304.64760723, 469.43268409, 461.18011966, 405.54186269,
```

```
       427.90365018, 442.01287324, 472.62722516, 467.5692018 ,
       344.31315885, 369.07085214, 426.04016789, 400.48383933,
       472.62722516, 321.95137136, 350.16981747, 469.43268409,
       411.39852132, 460.38148439, 312.90017166, 452.92755523,
       264.71584385, 465.70571951, 329.67151227, 440.4156027 ,
       472.62722516, 469.16647233, 419.11866224, 460.38148439,
       299.32337211, 448.66816713, 406.0742862 , 365.34388756,
       253.53495011, 472.62722516, 371.99918145, 458.5180021 ,
       464.10844897, 467.03677829, 328.3404535 , 379.98553413,
       393.82854543, 469.69889584, 470.23131936, 397.28929826,
       472.62722516, 349.10497045, 298.25852509, 362.94798175,
       442.27908499, 472.62722516, 296.3950428 , 282.01960798,
       414.85927414, 463.84223722, 418.58623872, 294.5315605 ,
       453.99240225, 214.13561024, 283.35066676, 472.62722516,
       337.65786495, 472.62722516, 292.40186646, 458.78421386,
       461.97875493, 259.92403225, 472.62722516, 467.5692018 ,
       460.9139079 , 440.68181446, 472.62722516, 371.46675794,
       462.51117844, 265.78069087, 435.09136759, 291.60323119,
       464.90708424, 450.26543767, 248.21071499, 437.75348514,
       264.4496321 , 463.84223722, 374.1288755 , 465.43950775,
       275.09810233, 316.89334799, 399.95141582, 241.02299758,
       413.52821536, 464.64087248, 303.84897196, 378.65447535,
       431.63061476, 456.92073156, 468.36783707, 433.76030881,
       448.66816713, 448.13574362, 461.44633141, 347.24148816,
       469.69889584, 452.92755523, 374.39508726, 410.59988605,
       431.364403  , 289.20732539, 267.37796141, 313.96501868,
       449.99922591, 324.61348891, 417.25517994, 465.97193126,
       276.42916111, 289.7397489 , 436.15621461, 394.36096894,
       472.62722516, 379.71932237, 349.10497045, 313.69880692,
       362.68177   , 444.94120255, 472.62722516, 326.21075945,
       393.29612192, 389.56915734, 472.62722516, 423.64426208,
       419.11866224, 426.83880315, 454.79103752, 262.5861498 ,
       444.6749908 , 465.70571951, 462.24496668, 365.0776758 ,
       395.42581597, 255.66464415, 351.2346645 , 263.38478507,
       460.9139079 , 451.06407294, 245.54859743, 441.74666148,
       291.60323119, 252.20389133, 464.37466073, 308.37457181,
       467.03677829, 466.50435477, 426.30637964])
```

In [56]:
```
plt.hist(y_pred)
```

Out[56]:
```
(array([  8.,  27.,  25.,  31.,  26.,  26.,  30.,  31.,  51., 124.]),
 array([214.13561024, 239.98477173, 265.83393323, 291.68309472,
        317.53225621, 343.3814177 , 369.23057919, 395.07974068,
        420.92890218, 446.77806367, 472.62722516]),
 <BarContainer object of 10 artists>)
```



In [59]:
```
y_test.head(10)
```

Out[59]:
```
383    666
433    666
190    398
495    391
235    307
269    223
36     279
30     307
305    222
438    666
Name: tax, dtype: int64
```

In [65]:
```
x_test.head(10)
```

Out[65]:

	age
383	100.0
433	87.9
190	21.5
495	28.8
235	61.5
269	61.5
36	61.4
30	94.1
305	58.1
438	87.9

In [60]:
```
y_pred[0:10]
```

Out[60]:
```
array([472.62722516, 440.4156027 , 263.65099683, 283.084455  ,
       370.13569916, 370.13569916, 369.86948741, 456.92073156,
       361.08449946, 440.4156027 ])
```

In [61]:
```
train_score=regr.score(x_train,y_train)
```

In [62]:
```
test_score=regr.score(x_test,y_test)
```

In [63]:
```
train_score
```

Out[63]:
```
0.2146353945961217
```

In [64]:
```
test_score
```

Out[64]:
```
0.24412538742091794
```

In [ ]: