CASE STUDY REPORT

Terro’s Real Estate Agency

1 summary statistics

#CRIME\_RATE

|  |  |
| --- | --- |
| *CRIME\_RATE* |  |
|  |  |
| Mean | 4.871976285 |
| Standard Error | 0.129860152 |
| Median | 4.82 |
| Mode | 3.43 |
| Standard Deviation | 2.921131892 |
| Sample Variance | 8.533011532 |
| Kurtosis | -1.189122464 |
| Skewness | 0.021728079 |
| Range | 9.95 |
| Minimum | 0.04 |
| Maximum | 9.99 |
| Sum | 2465.22 |
| Count | 506 |

1. The average crime rate per capital in town is 4.87 %.
2. The frequent crime rate is 4.82 %
3. The minimum rate is 0.04%
4. The maximum crime rate 9.99%

#AGE

|  |  |
| --- | --- |
| *AGE* |  |
|  |  |
| Mean | 68.57490119 |
| Standard Error | 1.251369525 |
| Median | 77.5 |
| Mode | 100 |
| Standard Deviation | 28.14886141 |
| Sample Variance | 792.3583985 |
| Kurtosis | -0.967715594 |
| Skewness | -0.59896264 |
| Range | 97.1 |
| Minimum | 2.9 |
| Maximum | 100 |
| Sum | 34698.9 |
| Count | 506 |

The proportion of house build prior to 1940

1)The average age of house 68.5

2)We can find 77.5 years of house more in the town

3The minimum age are up to 2.9

4)The maximum age of house is 100

#INDUS

|  |  |
| --- | --- |
| *INDUS* |  |
|  |  |
| Mean | 11.13678 |
| Standard Error | 0.30498 |
| Median | 9.69 |
| Mode | 18.1 |
| Standard Deviation | 6.860353 |
| Sample Variance | 47.06444 |
| Kurtosis | -1.23354 |
| Skewness | 0.295022 |
| Range | 27.28 |
| Minimum | 0.46 |
| Maximum | 27.74 |
| Sum | 5635.21 |
| Count | 506 |

The proportion of non rental business acres per town

1)The average proportion is 11.13

2)We can see 18.1 of more number non rental business acres per town

3The minimum is 0.46

4)The maximum age of house is 27.4

#NOX

|  |  |
| --- | --- |
| *NOX* |  |
|  |  |
| Mean | 0.554695 |
| Standard Error | 0.005151 |
| Median | 0.538 |
| Mode | 0.538 |
| Standard Deviation | 0.115878 |
| Sample Variance | 0.013428 |
| Kurtosis | -0.06467 |
| Skewness | 0.729308 |
| Range | 0.486 |
| Minimum | 0.385 |
| Maximum | 0.871 |
| Sum | 280.6757 |
| Count | 506 |

Nitric oxide concentration

1)The average concentration is 0.55

2)We can see 0.538 most frequent nitric oxide concentration

3The minimum concentration is 0.385

1. The maximum concentration is 0.871

#DISTANCE

|  |  |
| --- | --- |
| *DISTANCE* |  |
|  |  |
| Mean | 9.549407 |
| Standard Error | 0.387085 |
| Median | 5 |
| Mode | 24 |
| Standard Deviation | 8.707259 |
| Sample Variance | 75.81637 |
| Kurtosis | -0.86723 |
| Skewness | 1.004815 |
| Range | 23 |
| Minimum | 1 |
| Maximum | 24 |
| Sum | 4832 |
| Count | 506 |

Distance from highway (miles)

Average distance is 9.5

Minimum is 1

Maximum is 24

#TAX

|  |  |
| --- | --- |
| *TAX* |  |
|  |  |
| Mean | 408.2372 |
| Standard Error | 7.492389 |
| Median | 330 |
| Mode | 666 |
| Standard Deviation | 168.5371 |
| Sample Variance | 28404.76 |
| Kurtosis | -1.14241 |
| Skewness | 0.669956 |
| Range | 524 |
| Minimum | 187 |
| Maximum | 711 |
| Sum | 206568 |
| Count | 506 |

Full value property tax rate in 1000$

Average tax is 408.2

Minimum is 187

Maximum is 711

#PTRATIO

|  |  |
| --- | --- |
| *PTRATIO* |  |
|  |  |
| Mean | 18.45553 |
| Standard Error | 0.096244 |
| Median | 19.05 |
| Mode | 20.2 |
| Standard Deviation | 2.164946 |
| Sample Variance | 4.686989 |
| Kurtosis | -0.28509 |
| Skewness | -0.80232 |
| Range | 9.4 |
| Minimum | 12.6 |
| Maximum | 22 |
| Sum | 9338.5 |
| Count | 506 |

Pupil teacher ratio in the town

Average tax is 18.4

Minimum ratio is 12.6

Maximum ratio is 22

#AVG\_ROOM

|  |  |
| --- | --- |
| *AVG\_ROOM* |  |
|  |  |
| Mean | 6.284634 |
| Standard Error | 0.031235 |
| Median | 6.2085 |
| Mode | 5.713 |
| Standard Deviation | 0.702617 |
| Sample Variance | 0.493671 |
| Kurtosis | 1.8915 |
| Skewness | 0.403612 |
| Range | 5.219 |
| Minimum | 3.561 |
| Maximum | 8.78 |
| Sum | 3180.025 |
| Count | 506 |

Avg number of room in the house

Mean is 6.28

Minimum avg is 3.5

Maximum avg rooms 8.7

#LSTAT

|  |  |
| --- | --- |
| *LSTAT* |  |
|  |  |
| Mean | 12.65306 |
| Standard Error | 0.317459 |
| Median | 11.36 |
| Mode | 8.05 |
| Standard Deviation | 7.141062 |
| Sample Variance | 50.99476 |
| Kurtosis | 0.49324 |
| Skewness | 0.90646 |
| Range | 36.24 |
| Minimum | 1.73 |
| Maximum | 37.97 |
| Sum | 6402.45 |
| Count | 506 |

% of lower status of population

Mean is 12.6

Minimum population is 1.73

Maximum is 37.9

AVG\_PRICE

|  |  |
| --- | --- |
| *AVG\_PRICE* |  |
|  |  |
| Mean | 22.53281 |
| Standard Error | 0.408861 |
| Median | 21.2 |
| Mode | 50 |
| Standard Deviation | 9.197104 |
| Sample Variance | 84.58672 |
| Kurtosis | 1.495197 |
| Skewness | 1.108098 |
| Range | 45 |
| Minimum | 5 |
| Maximum | 50 |
| Sum | 11401.6 |
| Count | 506 |

12 bins / 11401.6

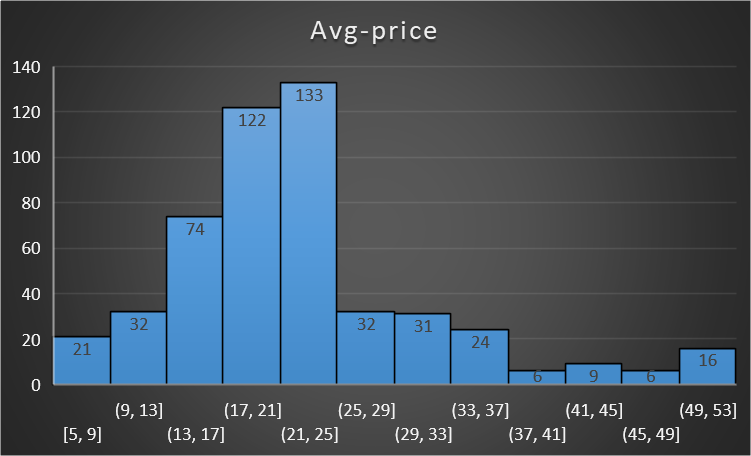
Average value of the house in $1000

Mean value 22.5

Minimum value is 5

Maximum is 50

2



We can observe more number of avg price come under 21 to 25

Bin width 4

Total observations 506

Total bins 12

Mode 133

It is an right tail skewness graph

3

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *CRIME\_RATE* | *AGE* | *INDUS* | *NOX* | *DISTANCE* | *TAX* | *PTRATIO* | *AVG\_ROOM* | *LSTAT* | *AVG\_PRICE* |
| CRIME\_RATE | 8.516148 |  |  |  |  |  |  |  |  |  |
| AGE | 0.562915 | 790.7925 |  |  |  |  |  |  |  |  |
| INDUS | -0.11022 | 124.2678 | 46.97143 |  |  |  |  |  |  |  |
| NOX | 0.000625 | 2.381212 | 0.605874 | 0.013401 |  |  |  |  |  |  |
| DISTANCE | -0.22986 | 111.55 | 35.47971 | 0.61571 | 75.66653 |  |  |  |  |  |
| TAX | -8.22932 | 2397.942 | 831.7133 | 13.0205 | 1333.117 | 28348.62 |  |  |  |  |
| PTRATIO | 0.068169 | 15.90543 | 5.680855 | 0.047304 | 8.743402 | 167.8208 | 4.677726 |  |  |  |
| AVG\_ROOM | 0.056118 | -4.74254 | -1.88423 | -0.02455 | -1.28128 | -34.5151 | -0.53969 | 0.492695 |  |  |
| LSTAT | -0.88268 | 120.8384 | 29.52181 | 0.48798 | 30.32539 | 653.4206 | 5.7713 | -3.07365 | 50.89398 |  |
| AVG\_PRICE | 1.162012 | -97.3962 | -30.4605 | -0.45451 | -30.5008 | -724.82 | -10.0907 | 4.484566 | -48.3518 | 84.41956 |

The variance of crime rate is 8.516148

The covariance of crime rate and age is 0.562915 indicating magnitude as 1 and direction is positive.

The covariance of crime rate and INDUS is -0.11022 indicating magnitude as 0 and direction is negative.

Similarly we can observe the covariance relationship in the above matrix.

4

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *CRIME\_RATE* | *AGE* | *INDUS* | *NOX* | *DISTANCE* | *TAX* | *PTRATIO* | *AVG\_ROOM* | *LSTAT* | *AVG\_PRICE* |
| CRIME\_RATE | 1 |  |  |  |  |  |  |  |  |  |
| AGE | 0.006859 | 1 |  |  |  |  |  |  |  |  |
| INDUS | -0.00551 | 0.644779 | 1 |  |  |  |  |  |  |  |
| NOX | 0.001851 | 0.73147 | 0.763651 | 1 |  |  |  |  |  |  |
| DISTANCE | -0.00906 | 0.456022 | 0.595129 | 0.611441 | 1 |  |  |  |  |  |
| TAX | -0.01675 | 0.506456 | 0.72076 | 0.668023 | 0.910228 | 1 |  |  |  |  |
| PTRATIO | 0.010801 | 0.261515 | 0.383248 | 0.188933 | 0.464741 | 0.460853 | 1 |  |  |  |
| AVG\_ROOM | 0.027396 | -0.24026 | -0.39168 | -0.30219 | -0.20985 | -0.29205 | -0.3555 | 1 |  |  |
| LSTAT | -0.0424 | 0.602339 | 0.6038 | 0.590879 | 0.488676 | 0.543993 | 0.374044 | -0.61381 | 1 |  |
| AVG\_PRICE | 0.043338 | -0.37695 | -0.48373 | -0.42732 | -0.38163 | -0.46854 | -0.50779 | 0.69536 | -0.73766 | 1 |

 The top 3 positively correlated pairs are

Distance|tax 0.91022

Indus|Nox 0.76365

Ages|Nox 0.73147

The top 3 negatively correlated pairs are

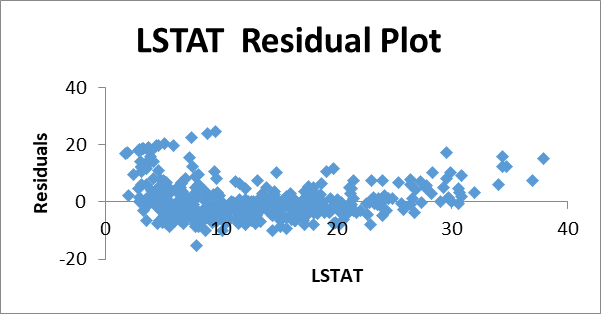
Tax|AVG room -0.29204

Tax|AVG price -0.46853

PTRatio|AVG\_Room -0.35550

5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.737663 |  |  |  |  |  |  |  |
| R Square | 0.544146 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.543242 |  |  |  |  |  |  |  |
| Standard Error | 6.21576 |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  | 0.05 |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 23243.91 | 23243.91 | 601.6178711 | 5.0811E-88 |  |  |  |
| Residual | 504 | 19472.38 | 38.63568 |  |  |  |  |  |
| Total | 505 | 42716.3 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 34.55384 | 0.562627 | 61.41515 | 3.7431E-236 | 33.44845704 | 35.65922 | 33.44846 | 35.65922 |
| LSTAT | -0.95005 | 0.038733 | -24.5279 | 5.0811E-88 | -1.0261482 | -0.87395 | -1.02615 | -0.87395 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| RESIDUAL OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Observation* | *Predicted AVG\_PRICE* | *Residuals* |  |  |  |  |  |  |
| 1 | 29.8226 | -5.8226 |  |  |  |  |  |  |
| 2 | 25.87039 | -4.27039 |  |  |  |  |  |  |
| 3 | 30.72514 | 3.974858 |  |  |  |  |  |  |
| 4 | 31.7607 | 1.639304 |  |  |  |  |  |  |
| 5 | 29.49008 | 6.709922 |  |  |  |  |  |  |
| 6 | 29.60408 | -0.90408 |  |  |  |  |  |  |
| 7 | 22.74473 | 0.155273 |  |  |  |  |  |  |
| 8 | 16.3604 | 10.7396 |  |  |  |  |  |  |
| 9 | 6.118864 | 10.38114 |  |  |  |  |  |  |
| 10 | 18.308 | 0.592003 |  |  |  |  |  |  |
| 11 | 15.12533 | -0.12533 |  |  |  |  |  |  |
| 12 | 21.94669 | -3.04669 |  |  |  |  |  |  |
| 13 | 19.62857 | 2.071434 |  |  |  |  |  |  |
| 14 | 26.70643 | -6.30643 |  |  |  |  |  |  |
| 15 | 24.80633 | -6.60633 |  |  |  |  |  |  |
| 16 | 26.50692 | -6.60692 |  |  |  |  |  |  |



a)

The Regression Summary Output in terms of variance is r square i.e 0.544146 , the coefficient value is -0.95005, Intercept is 34.55384 and the Residual plot is zero variance.

b)YES LSTAT variable significant for the analysis based on your model because 5.0811E-88 is less than 0.05.

6)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |  |
| Multiple R | 0.799100498 |  |  |  |  |  |  |  |  |
| R Square | 0.638561606 |  |  |  |  |  |  |  |  |
| Adjusted R Square | 0.637124475 |  |  |  |  |  |  |  |  |
| Standard Error | 5.540257367 |  |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |  |
| Regression | 2 | 27276.99 | 13638.49 | 444.330892 | 7E-112 |  |  |  |  |
| Residual | 503 | 15439.31 | 30.69445 |  |  |  |  |  |  |
| Total | 505 | 42716.3 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |  |
| Intercept | -1.358272812 | 3.172828 | -0.4281 | 0.66876494 | -7.5919 | 4.875355 | -7.5919 | 4.875355 |  |
| AVG\_ROOM | 5.094787984 | 0.444466 | 11.46273 | 3.4723E-27 | 4.22155 | 5.968026 | 4.22155 | 5.968026 |  |
| LSTAT | -0.642358334 | 0.043731 | -14.6887 | 6.6694E-41 | -0.72828 | -0.55644 | -0.72828 | -0.55644 |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| RESIDUAL OUTPUT | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Observation* | *Predicted AVG\_PRICE* | *Residuals* |  |  |  |  |  |  |  |
| 1 | 28.94101368 | -4.94101 |  |  |  |  |  |  |  |
| 2 | 25.48420566 | -3.88421 |  |  |  |  |  |  |  |
| 3 | 32.65907477 | 2.040925 |  |  |  |  |  |  |  |
| 4 | 32.40652 | 0.99348 |  |  |  |  |  |  |  |
| 5 | 31.63040699 | 4.569593 |  |  |  |  |  |  |  |
| 6 | 28.05452701 | 0.645473 |  |  |  |  |  |  |  |
| 7 | 21.28707846 | 1.612922 |  |  |  |  |  |  |  |
| 8 | 17.78559653 | 9.314403 |  |  |  |  |  |  |  |
| 9 | 8.104693384 | 8.395307 |  |  |  |  |  |  |  |
| 10 | 18.24650673 | 0.653493 |  |  |  |  |  |  |  |

X1 =7

X2 =20

Y=M1\*X1+M2\*X2

22.8163492

The company is overcharging.

 b. Yes.

The performance of this model is better than the previous model there is increase of 0.0938826 % i.e 0.63712447 in terms of adjusted R-square.

7

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |
| Multiple R | 0.832979 |  |  |  |  |  |  |
| R Square | 0.693854 |  |  |  |  |  |  |
| Adjusted R Square | 0.688299 |  |  |  |  |  |  |
| Standard Error | 5.134764 |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |
| Regression | 9 | 29638.86 | 3293.207 | 124.9045 | 1.9E-121 |  |  |
| Residual | 496 | 13077.43 | 26.3658 |  |  |  |  |
| Total | 505 | 42716.3 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Intercept | 29.24132 | 4.817126 | 6.070283 | 2.54E-09 | 19.77683 | 38.7058 | 19.77683 |
| CRIME\_RATE | 0.048725 | 0.078419 | 0.621346 | 0.534657 | -0.10535 | 0.202799 | -0.10535 |
| AGE | 0.032771 | 0.013098 | 2.501997 | 0.01267 | 0.007037 | 0.058505 | 0.007037 |
| INDUS | 0.130551 | 0.063117 | 2.068392 | 0.039121 | 0.006541 | 0.254562 | 0.006541 |
| NOX | -10.3212 | 3.894036 | -2.65051 | 0.008294 | -17.972 | -2.67034 | -17.972 |
| DISTANCE | 0.261094 | 0.067947 | 3.842603 | 0.000138 | 0.127594 | 0.394593 | 0.127594 |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* |
| TAX | -0.0144 | 0.003905 | -3.68774 | 0.000251 | -0.02207 | -0.00673 | -0.02207 |
| PTRATIO | -1.07431 | 0.133602 | -8.0411 | 6.59E-15 | -1.3368 | -0.81181 | -1.3368 |
| AVG\_ROOM | 4.125409 | 0.442759 | 9.317505 | 3.89287E-19 | 3.255495 | 4.995324 | 3.255495 |
| LSTAT | -0.60349 | 0.053081 | -11.3691 | 8.91E-27 | -0.70778 | -0.49919 | -0.70778 |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| RESIDUAL OUTPUT | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| *Observation* | *Predicted AVG\_PRICE* | *Residuals* |  |  |  |  |  |
| 1 | 30.11536 | -6.11536 |  |  |  |  |  |
| 2 | 27.00714 | -5.40714 |  |  |  |  |  |
| 3 | 32.83291 | 1.867087 |  |  |  |  |  |
| 4 | 31.20703 | 2.192966 |  |  |  |  |  |
| 5 | 30.59473 | 5.605271 |  |  |  |  |  |
| 6 | 28.07645 | 0.623553 |  |  |  |  |  |
| 7 | 25.29985 | -2.39985 |  |  |  |  |  |
| 8 | 22.54671 | 4.553287 |  |  |  |  |  |
| 9 | 14.17584 | 2.32416 |  |  |  |  |  |
| 10 | 22.67662 | -3.77662 |  |  |  |  |  |

The value Adjusted R Square is 0.693854

The coefficient value are

|  |  |
| --- | --- |
|  | *Coefficients* |
| Intercept | 29.24132 |
| CRIME\_RATE | 0.048725 |
| AGE | 0.032771 |
| INDUS | 0.130551 |
| NOX | -10.3212 |
| DISTANCE | 0.261094 |
| TAX | -0.0144 |
| PTRATIO | -1.07431 |
| AVG\_ROOM | 4.125409 |
| LSTAT | -0.60349 |
|  |  |

The intercept value is 29.24

And overall significant value is 1.93275554549125E-121

8

we found that crime rate is most insignificant

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.832836 |  |  |  |  |  |  |  |
| R Square | 0.693615 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.688684 |  |  |  |  |  |  |  |
| Standard Error | 5.131591 |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 8 | 29628.68 | 3703.585 | 140.643 | 1.9E-122 |  |  |  |
| Residual | 497 | 13087.61 | 26.33323 |  |  |  |  |  |
| Total | 505 | 42716.3 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 29.42847 | 4.804729 | 6.124898 | 1.85E-09 | 19.98839 | 38.86856 | 19.98839 | 38.86856 |
| AGE | 0.032935 | 0.013087 | 2.516606 | 0.012163 | 0.007222 | 0.058648 | 0.007222 | 0.058648 |
| INDUS | 0.13071 | 0.063078 | 2.072202 | 0.038762 | 0.006778 | 0.254642 | 0.006778 | 0.254642 |
| NOX | -10.2727 | 3.890849 | -2.64022 | 0.008546 | -17.9172 | -2.62816 | -17.9172 | -2.62816 |
| DISTANCE | 0.261506 | 0.067902 | 3.851242 | 0.000133 | 0.128096 | 0.394916 | 0.128096 | 0.394916 |
| TAX | -0.01445 | 0.003902 | -3.70395 | 0.000236 | -0.02212 | -0.00679 | -0.02212 | -0.00679 |
| PTRATIO | -1.0717 | 0.133454 | -8.03053 | 7.08E-15 | -1.33391 | -0.8095 | -1.33391 | -0.8095 |
| AVG\_ROOM | 4.125469 | 0.442485 | 9.3234 | 3.69E-19 | 3.256096 | 4.994842 | 3.256096 | 4.994842 |
| LSTAT | -0.60516 | 0.05298 | -11.4224 | 5.42E-27 | -0.70925 | -0.50107 | -0.70925 | -0.50107 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| RESIDUAL OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Observation* | *Predicted AVG\_PRICE* | *Residuals* |  |  |  |  |  |  |
| 1 | 30.04889 | -6.04889 |  |  |  |  |  |  |
| 2 | 27.04098 | -5.44098 |  |  |  |  |  |  |
| 3 | 32.69896 | 2.001035 |  |  |  |  |  |  |
| 4 | 31.14307 | 2.256931 |  |  |  |  |  |  |
| 5 | 30.58809 | 5.611913 |  |  |  |  |  |  |
| 6 | 27.85095 | 0.849047 |  |  |  |  |  |  |
| 7 | 25.0709 | -2.1709 |  |  |  |  |  |  |
| 8 | 22.63588 | 4.464117 |  |  |  |  |  |  |
| 9 | 14.00883 | 2.491167 |  |  |  |  |  |  |
| 10 | 22.84744 | -3.94744 |  |  |  |  |  |  |

1. Interpret the output of this model.

The over performance of the model is 68.87%

1. Compare the adjusted R-square value of this model with the model in the previous question I.e 68.83%, both the model performs better according to the value of adjusted R-square.
2. . Sort the values of the Coefficients in ascending order. What will happen to the average price if value of NOX is more in a locality in this town

Ho- the null hypothesis states that the value of NOX is low in the locality in this town

H1- the alternate hypothesis states that the value of NOX is more in the locality in this town.

Thus we can reject the null hypothesis the p value is less than 0.05

So we can say that the average price will be high when the value of nox is more in a locality in this town.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.427321 |  |  |  |  |  |  |  |
| R Square | 0.182603 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.180981 |  |  |  |  |  |  |  |
| Standard Error | 8.323348 |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 7800.126 | 7800.126 | 112.5915 | 7.07E-24 |  |  |  |
| Residual | 504 | 34916.17 | 69.27811 |  |  |  |  |  |
| Total | 505 | 42716.3 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 41.34587 | 1.811192 | 22.828 | 9.87E-80 | 37.78746 | 44.90429 | 37.78746 | 44.90429 |
| NOX | -33.9161 | 3.196337 | -10.6109 | 7.07E-24 | -40.1958 | -27.6363 | -40.1958 | -27.6363 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

1. Write the regression equation from this model.

Y = 0.032935x1 + 0.130710006682182x2 -10.2727050815094x3 +0.261506423001819x4 -0.0144523450364819x5 -1.07170247269449x6 +4.12546895908474x7 -0.605159282035406x8 + 29.42847