**BUSINESS REPORT**

Terro’s Real Estate Agency

1. **The first step to any project is understanding the data. So for this step, generate the summary statistics for each of the variables. What do you observe?**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *CRIME\_RATE* | |  | *AGE* | |  | *INDUS* | |
|  |  |  |  |  |  |  |  |
| Mean | 4.871976 |  | Mean | 68.5749 |  | Mean | 11.13678 |
| Standard Error | 0.12986 |  | Standard Error | 1.25137 |  | Standard Error | 0.30498 |
| Median | 4.82 |  | Median | 77.5 |  | Median | 9.69 |
| Mode | 3.43 |  | Mode | 100 |  | Mode | 18.1 |
| Standard Deviation | 2.921132 |  | Standard Deviation | 28.14886 |  | Standard Deviation | 6.860353 |
| Sample Variance | 8.533012 |  | Sample Variance | 792.3584 |  | Sample Variance | 47.06444 |
| Kurtosis | -1.18912 |  | Kurtosis | -0.96772 |  | Kurtosis | -1.23354 |
| Skewness | 0.021728 |  | Skewness | -0.59896 |  | Skewness | 0.295022 |
| Range | 9.95 |  | Range | 97.1 |  | Range | 27.28 |
| Minimum | 0.04 |  | Minimum | 2.9 |  | Minimum | 0.46 |
| Maximum | 9.99 |  | Maximum | 100 |  | Maximum | 27.74 |
| Sum | 2465.22 |  | Sum | 34698.9 |  | Sum | 5635.21 |
| Count | 506 |  | Count | 506 |  | Count | 506 |
|  |  |  |  |  |  |  |  |
| *TAX* | |  | *PTRATIO* | |  | *AVG\_ROOM* | |
|  |  |  |  |  |  |  |  |
| Mean | 408.2372 |  | Mean | 18.45553 |  | Mean | 6.284634 |
| Standard Error | 7.492389 |  | Standard Error | 0.096244 |  | Standard Error | 0.031235 |
| Median | 330 |  | Median | 19.05 |  | Median | 6.2085 |
| Mode | 666 |  | Mode | 20.2 |  | Mode | 6.127 |
| Standard Deviation | 168.5371 |  | Standard Deviation | 2.164946 |  | Standard Deviation | 0.702617 |
| Sample Variance | 28404.76 |  | Sample Variance | 4.686989 |  | Sample Variance | 0.493671 |
| Kurtosis | -1.14241 |  | Kurtosis | -0.28509 |  | Kurtosis | 1.8915 |
| Skewness | 0.669956 |  | Skewness | -0.80232 |  | Skewness | 0.403612 |
| Range | 524 |  | Range | 9.4 |  | Range | 5.219 |
| Minimum | 187 |  | Minimum | 12.6 |  | Minimum | 3.561 |
| Maximum | 711 |  | Maximum | 22 |  | Maximum | 8.78 |
| Sum | 206568 |  | Sum | 9338.5 |  | Sum | 3180.025 |
| Count | 506 |  | Count | 506 |  | Count | 506 |

ANS.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *NOX* | |  | *DISTANCE* | |
|  |  |  |  |  |
| Mean | 0.554695 |  | Mean | 9.549407 |
| Standard Error | 0.005151 |  | Standard Error | 0.387085 |
| Median | 0.538 |  | Median | 5 |
| Mode | 0.538 |  | Mode | 24 |
| Standard Deviation | 0.115878 |  | Standard Deviation | 8.707259 |
| Sample Variance | 0.013428 |  | Sample Variance | 75.81637 |
| Kurtosis | -0.06467 |  | Kurtosis | -0.86723 |
| Skewness | 0.729308 |  | Skewness | 1.004815 |
| Range | 0.486 |  | Range | 23 |
| Minimum | 0.385 |  | Minimum | 1 |
| Maximum | 0.871 |  | Maximum | 24 |
| Sum | 280.6757 |  | Sum | 4832 |
| Count | 506 |  | Count | 506 |
|  |  |  |  |  |
| *LSTAT* | |  | *AVG\_PRICE* | |
|  |  |  |  |  |
| Mean | 12.65306 |  | Mean | 22.53281 |
| Standard Error | 0.317459 |  | Standard Error | 0.408861 |
| Median | 11.36 |  | Median | 21.2 |
| Mode | 7.79 |  | Mode | 50 |
| Standard Deviation | 7.141062 |  | Standard Deviation | 9.197104 |
| Sample Variance | 50.99476 |  | Sample Variance | 84.58672 |
| Kurtosis | 0.49324 |  | Kurtosis | 1.495197 |
| Skewness | 0.90646 |  | Skewness | 1.108098 |
| Range | 36.24 |  | Range | 45 |
| Minimum | 1.73 |  | Minimum | 5 |
| Maximum | 37.97 |  | Maximum | 50 |
| Sum | 6402.45 |  | Sum | 11401.6 |
| Count | 506 |  | Count | 506 |

FIRST I GO TO DATA ANALYSIS TAB THEN I GO IN DESCRIPTIVE ANALYTIC TAB THEIR I GAVE INPUT RANGE AND I ALSO GIVE THEIR OUTPUT RANGE AFTER THAT I CHECK ON SUMMERY STATISTICS THEN I GOT THESE RESULTS .

THESE RESULT GIVES

MEAN WHICH IS AVERAGE OF VARIABLES

MODE WHICH IS MOST REPETITIVE VARIABLE

MEDIAN WHICH IS MIDDLE OF ALL VARIABLE AFTER ASSENDING OR DECENDING THEM

MIN SMALLEST NUMBER

MAX MAXIMUM NUMBER

SUM SUM OF ALL VARIBALE

COUNT TOTAL NUMBER OF VARIABLE

RANGE MAX-MIN

STANDARD DEVIATION SPREED IN DATA SEET FROM THE MEAN

VARIANCE SQUARE OF STANDARD DEVIATION

1. **Plot the histogram of the Avg\_Price Variable. What do you infer?**

|  |  |
| --- | --- |
| *Bin* | *Frequency* |
| 0-10 | 24 |
| 11-20 | 191 |
| 21-30 | 207 |
| 31-40 | 53 |
| 41-50 | 31 |
|  |  |

I GO TO DATA ANALYTICS TAB THEN I GO ON HISTOGRAM WHERE I GIVE HIM INPUT RANGE AND ALSO BIN RANGE WHICH IS MANNUALY CREATED BY ME WITH THE DIFFERNCE OF 10.

FROM THIS GRAPH WE CAN OBSERVE THAT THE MOST OF THE VARIABLES LIES BETWEEN 11 TO 30 WHICH IS TOTAL 191+207=398.

1. **Compute the covariance matrix. Share your observations.**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *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 |

FIRST I GO TO DATA ANALYSIS TAB THEN I GO IN COVARIANCE OPTION WHERE I PUT ALL INPUT RANGE AND ALSO OUTPUT RANGE THEN I GOT THIS RESULT. COVARIANCE ,  IS THE MEASUREMENT OF RELATIONSHIPS BETWEEN TWO RANDOM VARIABLES.

1. **Create a correlation matrix of all the variables as shown in the Videos and various case studies. State top 3 positively correlated pairs and top 3 negatively correlated pairs.**

ANS.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *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 |

TOP 3 POSITIVELY CORRELATED PAIR TOP 3 NEGATIVELY CORRELATED PAIR

DISTANCE TAX (0.910228) LSTAT AVG\_PRICE(-0.73766)

INDUS NOX (0.763651) AVG\_ ROOM LSTAT(-0.61381)

AGE NOX(0.73147) PTRATIO AVG\_PRICE(-0.50779)

1. **Build an initial regression model with AVG\_PRICE as the y or the Dependent variable and LSTAT variable as the Independent Variable. Generate the residual plot too.**

ANS.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.737663 |  |  |  |  |  |  |  |
| R Square | 0.544146 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.543242 |  |  |  |  |  |  |  |
| Standard Error | 6.21576 |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 23243.91 | 23243.91 | 601.6179 | 5.08E-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 | 3.7E-236 | 33.44846 | 35.6592247 | 33.44846 | 35.6592247 |
| LSTAT | -0.95005 | 0.038733 | -24.5279 | 5.08E-88 | -1.02615 | -0.87395051 | -1.02615 | -0.8739505 |

ANS. FOR THIS I HAVE TO GO TO DATA ANALYSIS TAB AND THEIR I HAVE TO GO REGRESSION TAB THAN AFTER THAT I HAVE TO GIVE X AND Y RANGE VALUES AND ALSO OUTPUT CELL VALUE AND THEN I CHECKED RESIDUALS PLOT OPTION THEN I GET RESULT.

1. What do you infer from the Regression Summary Output in terms of variance explained, coefficient value, Intercept and the Residual plot?

Ans. P values ​​and coefficients in the regression analysis worked together to tell us which relationship in the model is statistical significant and also nature. Coefficients values described the numerical relationship between the independent variable and the dependent variable or x and y. And the p value indicates and tell is these relationships are statistically significant or not.

1. Is LSTAT variable significant for the analysis based on your model?

Ans. YES, This is significant model because p value is less than 0.05.

1. **Build another instance of the Regression model but this time including LSTAT and AVG\_ROOM together as Independent variables and AVG\_PRICE as the dependent variable.**

ANS.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.7991 |  |  |  |  |  |  |  |
| R Square | 0.638562 |  |  |  |  |  |  |  |
| Adjusted R Square | 0.637124 |  |  |  |  |  |  |  |
| Standard Error | 5.540257 |  |  |  |  |  |  |  |
| Observations | 506 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 2 | 27276.99 | 13638.49 | 444.3309 | 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 | -1.35827 | 3.172828 | -0.4281 | 0.668765 | -7.5919 | 4.875355 | -7.5919 | 4.875355 |
| AVG\_ROOM | 5.094788 | 0.444466 | 11.46273 | 3.47E-27 | 4.22155 | 5.968026 | 4.22155 | 5.968026 |
| LSTAT | -0.64236 | 0.043731 | -14.6887 | 6.67E-41 | -0.72828 | -0.55644 | -0.72828 | -0.55644 |

1. **Write the Regression equation. If a new house in this locality has 7 rooms (on an average) and has a value of 20 for L-STAT, then what will be the value of AVG\_PRICE? How does it compare to the company quoting a value of 30000 USD for this locality? Is the company Overcharging/ Undercharging?**

Ans. LINEAR E UATION Y=m1x1+m2x2+c

Y= dependent variable

X1=independent variable1(avg-room)

X2=independent variable2(LSTAT)

M1= coefficient independent variable1

M2= coefficient independent variable2

Y= (5.094788\*7)+(-0.64236\*20)+( -1.35827)

Y=21.4538 (ans.)

1. **Is the performance of this model better than the previous model you built in Question 5? Compare in terms of adjusted R-square. Explain.**

Ans. Adjusted r square = tells how data fit in the regression model

Current model r square = (0.63712)

Previous model r square=(0.54324)

**7. Now, build a Regression model with all variables. AVG\_PRICE shall be the Dependent Variable. Interpret the output in terms of adjusted R-square, coefficient and Intercept values, Significance of variables with respect to AVG\_price. Explain.**

ANS.FIRST I GO TO DATA ANALYSIS TAB THEN I GO IN REGRESSION OPTION IN DATA ANLYSIS TAB THERE I CHOOSE AVG\_PRICE AS Y OR DEPENDENT VARIABLE AND IN INPUT X RANGE I GAVE ALL OTHER VARIABLE THEN I GOT THIS RESULT.

WE KNOW THAT LESS THAN 0.05 P-VALUES INDICATES THAT WE CAN REJECT THE NULL HYPOTHESIS.

IN THIS MODEL ONLY CRIME RATE SIGNIFICANCE VALUE IS GREATER THAN 0.05 AND ALL OTHER VALUES ARE LESS THAN 0.05 OR WE CAN SAY THAT P-VALUE OF 5% OR LOWER IS CONSIDERED SATISTICALLY SIGNIFICANT.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 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 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 29.24132 | 4.817126 | 6.070283 | 2.54E-09 | 19.77683 | 38.7058 | 19.77683 | 38.7058 |
| CRIME\_RATE | 0.048725 | 0.078419 | 0.621346 | 0.534657 | -0.10535 | 0.202799 | -0.10535 | 0.202799 |
| AGE | 0.032771 | 0.013098 | 2.501997 | 0.01267 | 0.007037 | 0.058505 | 0.007037 | 0.058505 |
| INDUS | 0.130551 | 0.063117 | 2.068392 | 0.039121 | 0.006541 | 0.254562 | 0.006541 | 0.254562 |
| NOX | -10.3212 | 3.894036 | -2.65051 | 0.008294 | -17.972 | -2.67034 | -17.972 | -2.67034 |
| DISTANCE | 0.261094 | 0.067947 | 3.842603 | 0.000138 | 0.127594 | 0.394593 | 0.127594 | 0.394593 |
| TAX | -0.0144 | 0.003905 | -3.68774 | 0.000251 | -0.02207 | -0.00673 | -0.02207 | -0.00673 |
| PTRATIO | -1.07431 | 0.133602 | -8.0411 | 6.59E-15 | -1.3368 | -0.81181 | -1.3368 | -0.81181 |
| AVG\_ROOM | 4.125409 | 0.442759 | 9.317505 | 3.89E-19 | 3.255495 | 4.995324 | 3.255495 | 4.995324 |
| LSTAT | -0.60349 | 0.053081 | -11.3691 | 8.91E-27 | -0.70778 | -0.49919 | -0.70778 | -0.49919 |

**8. Pick out only the significant variables from the previous question. Make another instance of the Regression model using only the significant variables you just picked.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 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 |

1. **Interpret the output of this model.**

ANS. IN LAST QUESTION I MAKE A REGRESSION MODEL OF ALL VARIBLE. IN ALL THESE VARIBLE ONLY CRIME RATE P-VALUE IS GREATER THAN 0.05 SO THAT WAS NOT SIGNIFICANT IN CRIME RATE I ALREADY HIGHLIGHTED THE VALUE IN LAST QUESTION MODEL. SO THIS TIME ITAKE ALL VARIABLES OTHER THAN CRIME RATE. SO I GOT THIS MODEL.

1. **Compare the adjusted R-square value of this model with the model in the previous question, which model performs better according to the value of adjusted R-square?**

ANS. PREVIOUS ADJUSTED R SQUARE :-0.688294

CURRENT ADJUSTED R SQUARE :- 0.688684

IN PREVIOUS QUESTION ADJUSTED R SUARED VALUE IS 0.688299 AND THIS QUESTION ADJUSTED R SQUARE IS 0.688684 WHICH IS NOT SO MUCH DIFFERNCE IN BOTH THE MODEL ADJUSTED R SQURE VALUES SO I THINK BOTH MODELS WILL PERFORM APPROX EQUALLY.

1. **Sort the values of the Coefficients in ascending order. What will happen to the average price if the value of NOX is more in a locality in this town?**

|  |  |
| --- | --- |
|  | *Coefficients* |
| NOX | -10.27271 |
| PTRATIO | -1.071702 |
| LSTAT | -0.605159 |
| TAX | -0.014452 |
| AGE | 0.032935 |
| INDUS | 0.13071 |
| DISTANCE | 0.2615064 |
| AVG\_ROOM | 4.125469 |
| Intercept | 29.428473 |

ANS. NOX HAVE TO MUCH NEGATIVE SLOPE.

NOX IS INDEPENDENT VARIABLE AND AVG\_PRICE IS DEPENDENT VARIABLE IF NOX IS MORE IN A LOCALITY IN THIS TOWN THEN THE AVERAGE PRICE WILL DECREASE BECAUSE  A NEGATIVE COEFFICIENT TELLS THAT WHEN THE INDEPENDENT VARIABLE INCREASES THEN THE DEPENDENT VARIABLE HAVE TO DECREASE.

1. **Write the regression equation from this model.**

ANS. REGRSSSION EQUATION FIND OUT THE RELATIONSHIP EXISTS BETWEEN THE DATASETS.

Y=MX+C

Y=-10.27271x+29.42847 (WHERE M=-10.7271 AND M=29.42847)