**NAME: Rohitashwa Chakraborty**

Question-1.

Table

Description automatically generated

**Intercept = 10.888; Slope = 0.037**

Interpretation:

The sales was e10.888, even before we started collecting data in Jan 2003.

The slope is statistically significant, and it means log (sales) increases month over month the value.

Question-2.

**R2 : 0.8367; RMSE : 0.33415**

|  |
| --- |
| The two most important statistic are: 1) R-Square: Expresses the amount of variance in the target variable explained by the predictors 2) RMSE: Measure of the Average Squared error of the predictions.   A low RMSE and high R-Squared is desired and our model is at 0.33415, 83.67% respectively.  Question-3 |
|  |
|  |
|  |
|  |
|  |

Graphical user interface, application, table

Description automatically generated with medium confidence

Graphical user interface, application, table, Word

Description automatically generated

**L: No.** The residuals do not follow a constant level across the months

**H: Yes.** The squared values of residuals are constant throughout the range

**I: No.** The residuals are not independent

**N: Yes.** The residual generated are normal

Table

Description automatically generated

Question-4.

M12 leads to multicollinearity . Thus, removing it.

Change in SS (Sum of squares) = From 38.32066 to 39.84667

Wald statistic = (39.84667- 38.32066)/( 0.10634 \* 11) = 1.3057

Decision: **reject** the null hypothesis that the all-new predictors have 0 explanatory power

Table

Description automatically generatedGraphical user interface, application, table

Description automatically generated

Question-5.

Table

Description automatically generated

Change in SS (Sum of squares) = From 38.32066 to 14.87827

Wald statistic = abs(13.29028 - 14.87727)/( 0.06673 \* 12) = 1.982

Decision: **Fail to reject the null hypothesis**: the new features have explanatory power

Question-6.

Table

Description automatically generated

1. Equation:

Log(Sales) = 4.71716 + 0.0068\*month + 0.30824\* Log(lag\_sales\_12) + 0.3029\* Log(lag\_Sales\_1)

The p value of month predictor is 0.253 > 0.05,

Fail to reject the null hypothesis that the estimate value can be 0.

1. Increase in SS = The base model SS is 38.32066 and for the best model is 14.37
2. (38.32066-14.37)/38.32066 = 0.6250

Question-7.

**R-Sq. = 0.8050**

Table

Description automatically generated

Question-8.

**Number of predictors = 2**

* Log lag Sales 1
* Log lag Sales 12

Question-9

1. RESET

Table

Description automatically generated

All 3 powers are passing the 0.05 significance level.

1. White’s Homoscedasticity Test

Table

Description automatically generated with medium confidence

It is passing the 0.05 significance level.

1. Durbin-Watson

Table

Description automatically generated

It is passing the Durbin Watson test at 0.05 significance levels

10.

Forecast

= e(Intercept + coeff\_log\_lag\_sales1\*log(lag\_sales1)+ coeff\_log\_lag\_sales12\*log(lag\_sales12))

= e(3.10113+ 0.3630\*log(531951.44) + 0.040401\*log(508784.22)) = 257,140.922