Ans 1

Yes

Ans 2:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| -0.07 | -0.19 | -0.08 | 0.03 | -0.19 | -0.02 |

Ans 3:

Calculate beta on training set

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0.01 | 0.19 | -0.07 | 0.64 | -0.06 | 0.00 |

Use it in test set in excel file New\_Linear-Regression-Forecasting-Excel-Document in cell O7 to T7

0.6750 Answer is in cell “X8”

Ans 4

0.7378 Answer is in cell “Z4”

Ans 5: **the Standard deviation of model error, in Dollars, for the Test Set?**

3885.14 Answer is in cell “Z7”

Ans 6: **he 90% confidence interval, in dollars, for the Test Set?**

6390.49 Answer is in cell “Z10”

Ans 7:

27.7% Answer is in cell “Z14”

Ans 8:

With the help of beta values of training set make 2 columns for training set and calculate Actual minus Estimate (Residual), Squared Errors (residuals) as done for test sets . make charts as made for test sets

0.895 Answer is in cell “Z21”

Ans 9:

Use the value in B219 cell where we have calculated the linest (first column,third row) “0.655302185” in cell “AD4” to get result in cell “AB6”

0.5871 is in cell “AB6”

Ans 10: **the standard deviation of model error in dollars on the Training Set?**

Use the value in ans 9 i.e. 0.6750 to calculate as per test formula in cell in cell “Z7”

$3,379.35Answer is in cell “Z24”

Ans 11: **the 90% confidence interval, in dollars, on the Training Set?**

$5,558.54Answer is in cell “Z27”

Ans 12:

37.5% Answer is in cell “AB14”

Ans 13:

We use excel New Data for Final Project.xls sheet “Training Set” . Enter customer data in cells C207 to H207 , drag L205 to Q205 to get the standardized data

Standardize each input variable separately, using the appropriate mean and standard deviation from the Data for Final Project Workbook [mean, cells C207:H207 and standard deviation, cells C209:H209].

Then multiply each individual z-score by its “beta” coefficient from the original Excel “Linest” Calculation on the (standardized) Training Set.

Then sum the results. That sum is the point forecast of profitability, expressed as a standardized output (z-score). The correct z-score estimate for y = 1.525059.

Multiply the z-score by the standard deviation of profits ($5755.91), then add the mean profit ($1,905.51). You should get $10,683.61.

Ans 14: **With 50% confidence, what is the range of profitability?**

The mean of the interval is $10,683.61 (from Question 13).

This is a left-sided 25% confidence interval.

The normsinv(p = .25) = -0.67448975.

The interval from p= .25 to p = .75 is +- plus or minus

(the standard deviation of error as a fraction of the standard deviation of profitability)\*(normsinv(.25))\*(standard deviation of profits)

= (.675)\*(0.67448975)\* ($5755.91)

= +- $2,620.55.

The 50% confidence interval range is from ($10,683.61 + 2,620.55) to ($10,683.61 - $2,620.55).

13034.16 to 8063.06

Ans 15: **With 99.5% confidence, what is the range of profitability?**

The mean of the interval is $10,683.61 (from Question 13).

The left-sided confidence interval is .5% or .005. The normsinv(p = .005)

= -2.575829304.

The interval from p= .005 to p = .995 is:

+- plus or minus

(the standard deviation of error as a fraction of the standard deviation of profits)\*(normsinv(p = .005))\*(standard deviation of profits)

= (.675)\*(2.575829304)\* ($5755.91)

= +- $10,007.71.

The range is from ($10,683.61 + $10,007.71)

to ($10,683.61 - $10,007.71).

20691.32 to 675.90

Ans 16:

The actual change was from $3,379.36 to $3,885.14, an increase of less than 15%.

Increased by less than 20% which suggest minimal overfitting.