Data Science Batch Assignment 1

**Problem Statement 1**

1. H0 – There is no significant difference for the marital status in response to the bank marketing campaign. HA – There is a significant difference for the marital status in response to the bank marketing campaign.

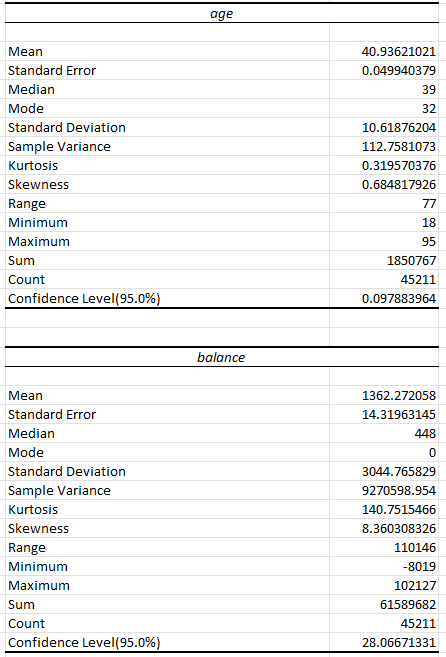
A screenshot of a graph

Description automatically generated

Random Sample

**A screenshot of a calculator

Description automatically generated**

****

**A screenshot of a graph

Description automatically generated**

**Problem Statement 2**

1. A graph of a line

   Description automatically generated
2. A screenshot of a table

   Description automatically generated
3. The 90% confidence interval for the slope of the regression line y on x is (18.362, 21.920).

Lower Bound: Slope – (1.645 \* Standard Error of the Slope)

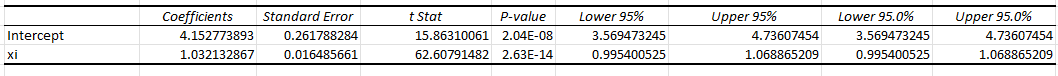
Upper Bound: Slope + (1.645 \* Standard Error of the Slope)

1. The P value is 2.622867 x 10^-14 which is almost 0 and the alpha value is 0.05. Since the p value is much smaller than the alpha value, we reject the null hypothesis in favour of the alternative hypothesis.

A screenshot of a spreadsheet

Description automatically generated

1. So from this table here:



We can see that the intercept is 4.1527 and the slope is 1.0321. So we use these to calculate the predicted value. The predicted value is:

Predicted Value = 4.1527 + (1.0321 x 12) = 16.5272 – Equation is Confidence Interval = Predicted Value +- (Critical Value x Standard Error)

For calculating the predicted value, we have:

Confidence Interval = Predicted Value +- (Critical Value x Standard Error)

Confidence Interval = 16.5271 +- (1.96 x 0.016486)

Lower Bound: 16.52718 – (1.96 x 0.016486) == 16.4942

Upper Bound: 16.52718 + (1.96 x 0.016486) == 16.5602

The confidence interval of the x value at 12 is (16.4942, 16.5602).

1. We can use a test statistic to determine whether the mean response for the y value at x = 12 is significantly more than 16.5 at the 1% level of significance. From the following information:

Sample Mean: 16.5272

Hypothesised Mean: 16.5

Standard Error: 0.0165

Significance Level (A) : 0.01

Degrees of Freedom (df): 10

t = 16.5272 – 16.5 / 0.0165 == 1.6364

c (t-distribution) = 2.7639

The test statistic (1.6364) is less than the critical value (2.7639), we fail to reject the null hypothesis, meaning there is not enough evidence to conclude the mean response for the y value at x = 12 is significantly more than 16.5 at the 1% level of significance. We not have enough statistical evidence to support the claim that the mean response is significantly greater than 16.5.