Assignment 17

1. Using a graph to illustrate slope and intercept, define basic linear regression.

Ans: The slope indicates the steepness of a line and the intercept indicates the location where it intersects an axis.

2. In a graph, explain the terms rise, run, and slope.

Ans: The vertical change between two points is called the rise, and the horizontal change is called the run. The slope equals the rise divided by the run: Slope =riserun Slope = rise run

3. Use a graph to demonstrate slope, linear positive slope, and linear negative slope, as well as the different conditions that contribute to the slope.

Ans: If the graph of a line rises from left to right, the slope is positive. If the graph of the line falls from left to right the slope is negative.

4. Use a graph to demonstrate curve linear negative slope and curve linear positive slope.

Ans: In order to graph a negative slope, it is important to know that a negative number in the ratio will mean to move down along the y-axis or to the left along the x-axis. A positive number in the ratio will mean to move up along the y-axis or to the right along the x-axis.

5. Use a graph to show the maximum and low points of curves.

Ans: To find the maximum/minimum of a curve you must first differentiate the function and then equate it to zero. This gives you one coordinate. To find the other you must resubstitute the one already found into the original function.

6. Use the formulas for a and b to explain ordinary least squares.

Ans: In all cases the formula for OLS estimator remains the same: ^β = (XTX)−1XTy; the only difference is in how we interpret this result.

7. Provide a step-by-step explanation of the OLS algorithm.

Ans: The ordinary least squares (OLS) method is a linear regression technique that is used to estimate the unknown parameters in a model

8. What is the regression's standard error? To represent the same, make a graph.

Ans: The standard error of the estimate is a measure of the average deviation of the errors, the difference between the ^y -values predicted by the multiple regression model and the y -values in the sample.

9. Provide an example of multiple linear regression.

Ans: an analyst may want to know how the movement of the market affects the price of ExxonMobil (XOM).

10. Describe the regression analysis assumptions and the BLUE principle.

Ans: The relationship between X and the mean of Y is linear. Homoscedasticity: The variance of residual is the same for any value of X

11. Describe two major issues with regression analysis.

Ans: Heteroskedasticity: variance of error term is not constant. F-test is unreliable. Standard error underestimated.

12. How can the linear regression model's accuracy be improved?

Ans: Distribution and Residual plots confirm that there is a good overlap between predicted and actual charges. However, there are a handful of predicted values that are way beyond the x-axis and this makes our RMSE is higher. This can be reduced by increasing our data points i.e. collecting more data.

13. Using an example, describe the polynomial regression model in detail.

Ans: Polynomial regression is one of the machine learning algorithms used for making predictions. For example, it is widely applied to predict the spread rate of COVID-19 and other infectious diseases.

14. Provide a detailed explanation of logistic regression.

Ans: Logistic regression is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set.

15. What are the logistic regression assumptions?

Ans: Basic assumptions that must be met for logistic regression include independence of errors, linearity in the logit for continuous variables, absence of multicollinearity, and lack of strongly influential outliers.

16. Go through the details of maximum likelihood estimation.

Ans: This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable.