**ASSIGNMENT 2**

**DS2306**

**TOLULOPE OLAJOBI**

QUESTION 21.

d) Both a) and b) are true statements. 𝛽₀, 𝛽₁, ..., 𝛽ᵣ are the regression coefficients, and linear regression aims to determine the best predicted weights using the method of ordinary least squares.

QUESTION 22.

d) The value 𝑅² = 1, which corresponds to SSR = 0, indicates a perfect fit in linear regression. This means that all of the variation in the dependent variable is explained by the regression model.

QUESTION 23.

b) B0 represents the value where the estimated regression line crosses the 𝑦-axis.

QUESTION 24.

a) The bottom-left plot represents an underfitted model. It shows a low degree of fit to the data, suggesting that the model is too simple to capture the underlying relationship.

QUESTION 25.

b) The correct order of the steps for implementing linear regression is: e) Create a regression model and fit it with existing data, d) Import the packages and classes that you need, b) Provide data to work with, and eventually do appropriate transformations, a) Check the results of model fitting to know whether the model is satisfactory, c) Apply the model for predictions.

QUESTION 26.

b) fit intercept, c) normalize, d) copy\_X, e) n\_jobs are the optional parameters to LinearRegression in scikit-learn.

QUESTION 27.

c) Polynomial regression requires transforming the array of inputs to include nonlinear terms such as 𝑥².

QUESTION 28.

c) You should choose statsmodels over scikit-learn when you need more detailed results. Statsmodels provides more comprehensive statistical analysis and detailed output compared to scikit-learn.

QUESTION 29.

b) Numpy is a fundamental package for scientific computing with Python that offers comprehensive mathematical functions, linear algebra routines, and more.

QUESTION 30.

b) Seaborn is a Python data visualization library based on Matplotlib that provides a high-level interface for drawing attractive and informative statistical graphics, closely integrating with pandas data structures.