Question-1:

List down at least three main assumptions of linear regression and explain them in your own words. To explain an assumption, take an example or a specific use case to show why the assumption makes sense.

Answer-1:

Below are the three assumptions of Linear Regression

1. In linear regression, the relationship between the independent and dependent variables should be linear.
   Explanation – The relationship is explained via a straight line having equation

2. There should be no or little multicollinearity in the data. Multicollinearity occurs when the independent variables are too highly correlated with each other.
   Explanation - If there is multicollinearity in the data that means one assumption about data can be explained by one or more variables, which will not lead us to any concrete relation between independent and dependent variable

3. Homoscedasticity - meaning the residuals should be equal across the regression line
   Explanation – If the residuals are unequal across the regression line that means our regression line is not fit for all the data and hence fails to explain the relationship between independent and dependent variable.

Question-2:

Explain at least three regression model evaluation metrics.

Answer 2:

Below are three regression model evaluation metrices

1. Root Mean Square Error must be low. If this is low that means error in our regression model is low. In solution model RMSE is 0.11.
2. Adjusted R-Squared – High adjusted r squared value close to 1 means our model is good. The adjusted R-squared increases only if the new term improves the model more than it would be expected by chance. It decreases when a predictor improves the model by less than expected by chance. Adjusted R-Squared in solution model is 0.695
3. Low p-Value and VIF – Low p-Value means variables in the models are significant and low VIF means there is no or little multicollinearity.