ASSUMPTIONS OF MULTIPLE LINEAR REGRESSION

Provide one sample dataset that you will use for this activity. Ensure that your sample data is distinct from those of your classmates.

Here are the key points for developing a detailed handout:

1. Introduce your sample data, and provide a brief discussion about its purpose, structure, and a description of each variable.
2. Build a multiple linear regression model using all the independent variables, regardless of their significance based on p-values, and name this model "fullmodel." Then, create a second model named "reducedmodel" that includes only the significant independent variables. Discuss each variable that is significant in predicting the dependent variable, explaining its impact. For example, discuss the sign of the beta coefficient and reference relevant articles that support your discussion.
3. Use R software to determine if the assumptions of multiple linear regression are satisfied for the "reducedmodel." Discuss each test used and include the relevant code (you may present a screenshot of the software to support your discussion). It is also better to present some graphical methods to determine whether the assumptions are satisfied.
4. If there is any violation of the assumptions, indicate this in your paper and suggest possible remedial measures, but you do not need to perform the remedial actions.
5. At the end of the detailed handout, before the references, include the following table as a summary.

Sample:

Assumptions	**Method of Detection**			**Satisfied (✓) or Violated (×)**	**Possible Remedial Measures**
	Graphical	*Statistical*	*Both*		
1.			✓	✓	
2.	✓			×	Perform the Box-Cox transformation.