ISL Assignment -1

Requirements: For the first three questions, please submit a PDF file (e.g., a scanned copy of your handwritten answers). For the last two questions, please submit an ipynb (Jupyter) file that contains your Python code.

1. What is linear regression and why is it used for data analysis and predictive analysis? (Please explain briefly in your own words)

2. a) Calculate the slope (β_1) and simple linear regression equation for below sample training dataset. (intercept (β_o) = 3)

Y (response)	13	15	17	19	21
X (predictor)	5	6	7	8	9

   b) Calculate RSE(Residual Standard Error) and R^2(Coefficient of Determination) on below testing dataset

Y (response)	24	27	30	33	36
X (predictor)	10	11	12	13	14

3. a) Calculate multiple linear Equation and show value of estimated coefficients.

	X1	**X2**	**X3**	**Y**
1	25	15	20	350
2	30	20	15	420
3	20	18	25	310
4	35	22	22	460
5	28	17	18	380

   b) Calculate the adjusted R-squared (R^2) value for this regression model.
   c) Predict Y for the following values X1 = 32, X2 = 19 , X3 = 21

4. Perform simple linear regression on Credit dataset in ISLR library by considering independent variables as “age” and dependent variable as “Income.”
   a) Import the dataset.
   b) Test functions like head(), tail(), nrow(), ncol(), dim() etc.,
   c) Create simple linear regression model using sklearn or statsmodels packages
   d) Find RSE (Residual Standard Error)
   e) Create a plot using the matplotlib package to display the model

5. Perform Multiple linear regression on the “Auto” dataset in ISLR package and answer above questions by considering independent variables as “Horsepower”, “Cylinders”,” Displacement” and dependent variable as “MPG”