**Nafiur Rahman Niloy**

**22-46459-1**

**Report**

**Results of running the code**

At first I tried to exicute the code in vs code but my vs code was not working properly so that I have exicuted the codes in google colab also I have provided the source code in the py files that you have provided.

**Simple:  
output:** A screenshot of a computer

Description automatically generated

**Plot:**

**A screenshot of a computer screen

Description automatically generated**

**Multiple:**

**Output:** **A screenshot of a computer

Description automatically generated**

**Plot:**

**A graph with red and blue lines

Description automatically generated**

**discussion/explanation:**

Simple linear regression involves predicting a target variable based on a single predictor variable, assuming a linear relationship represented by a straight line with parameters like slope and intercept. In contrast, multiple linear regression extends this concept to include two or more predictor variables, accommodating a multidimensional relationship between predictors and the target variable represented by a hyperplane. While simple linear regression is appropriate for scenarios with only one predictor variable, multiple linear regression allows for a more comprehensive analysis when there are multiple predictors influencing the target variable, offering insights into complex relationships between variables.

**Simple:**

For the GD model, parameter estimates are recorded as [254449.99982048, 93308.92004027], culminating in a training root mean square error (RMSE) of 64083.51 and a training cost of 2053348364.32. During the testing phase, the GD model manifests a test RMSE of 65773.19 and a test cost of 2163056350.22.

Conversely, the SGD model yields parameter estimates of [271373.5639219, 109566.48145122]. In the training domain, the SGD model registers a training RMSE of 68245.23 and a training cost of 2328706022.67. Subsequent testing reveals a test RMSE of 81574.42 and a test cost of 3327192645.02 for the SGD model.

Comparative analysis indicates that while the GD model tends to achieve lower training and testing RMSE values compared to its SGD counterpart, the latter incurs significantly higher computational costs in both training and testing endeavors. This observation underscores the nuanced interplay between computational efficiency and predictive accuracy inherent in the selection of optimization algorithms within linear regression modeling contexts.

**Multiple:**

For the GD model, the estimated parameters are presented as [254449.99982048, 78079.18106675, 24442.5758378, 2075.95636731]. The training phase yields a root mean square error (RMSE) of 61070.62 and a training cost of 1864810304.94. Subsequent testing reveals a test RMSE of 58473.59 and a test cost of 1709580288.69 for the GD model.

In contrast, the SGD model produces parameter estimates of [260645.91010201, 93312.2435386, 23747.2305936, 16104.42747514]. During training, the SGD model demonstrates a training RMSE of 64474.34 and a training cost of 2078470327.11. Evaluation in the testing phase unveils a test RMSE of 74448.45 and a test cost of 2771285624.96 for the SGD model.

Upon analysis, it is evident that the GD model generally attains lower training and testing RMSE values compared to the SGD model. However, the computational costs associated with training and testing are notably lower for the GD model as opposed to the SGD model. This comparison underscores the intricate balance between predictive accuracy and computational efficiency inherent in the selection of optimization algorithms within multiple linear regression modeling frameworks.

**Plots as image:**

A graph of a test

Description automatically generated with medium confidence A graph with red and blue lines

Description automatically generatedA graph with numbers and lines

Description automatically generated