Exercise 1:-

a) We would add variables that are known to be correlated with diabetes progression, such as glucose levels (variable **`glu`**), insulin levels (**`ins`**), and blood pressure (**`bp`**). These variables are commonly associated with diabetes and may provide additional information for the prediction model.

b) Adding more variables may improve the model's performance by capturing more information about the underlying relationships between predictors and the target variable. We can compute performance metrics such as RMSE and R2 score to compare the model's performance with and without the additional variables.

d) Adding even more variables may or may not improve the model's performance further. It depends on the relevance and strength of the relationships between the additional variables and the target variable. We can assess this by iteratively adding more variables and evaluating the model's performance metrics.

Exercise 2:-

Choose appropriate variables to predict company profit. Justify your choice:-
d) Based on the correlation matrix and domain knowledge, select variables that have a strong correlation with profit.

Exercise 3:-

e) Search optimal value for alpha (in terms of R2 score) by fitting the models with training data and computing the score using testing data: This step is already implemented in the above code. The R2 scores are computed and stored in `ridge_scores` and `lasso_scores`.

g) Identify the value for alpha which gives the best score:
You can examine the plot to determine the alpha value that corresponds to the highest R2 score for each regression method.