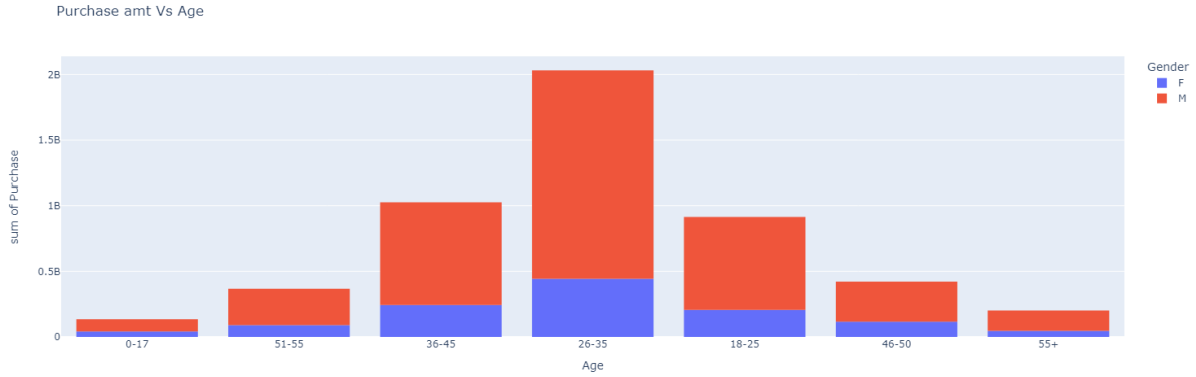
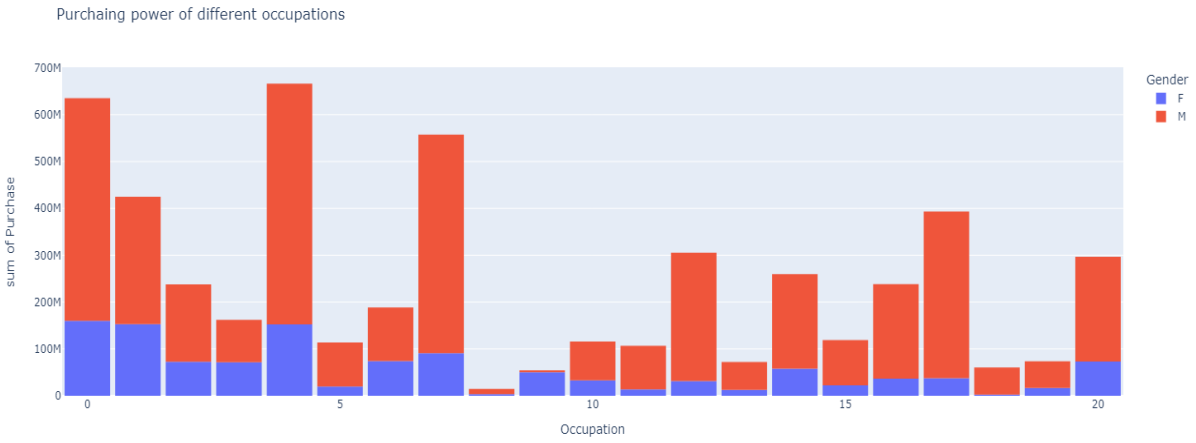
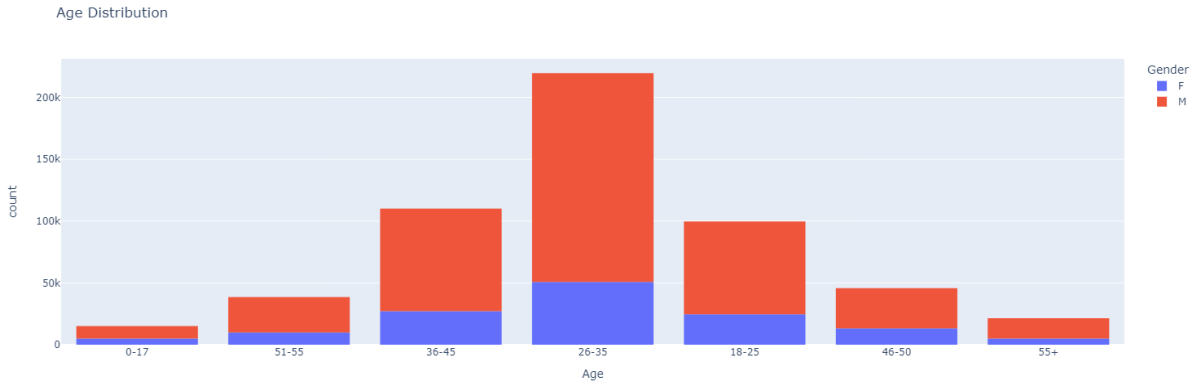
Assignment 2
Divyanshu Vaibhav
21BT10014

Experiment 1(EDA and correlation heatmaps)

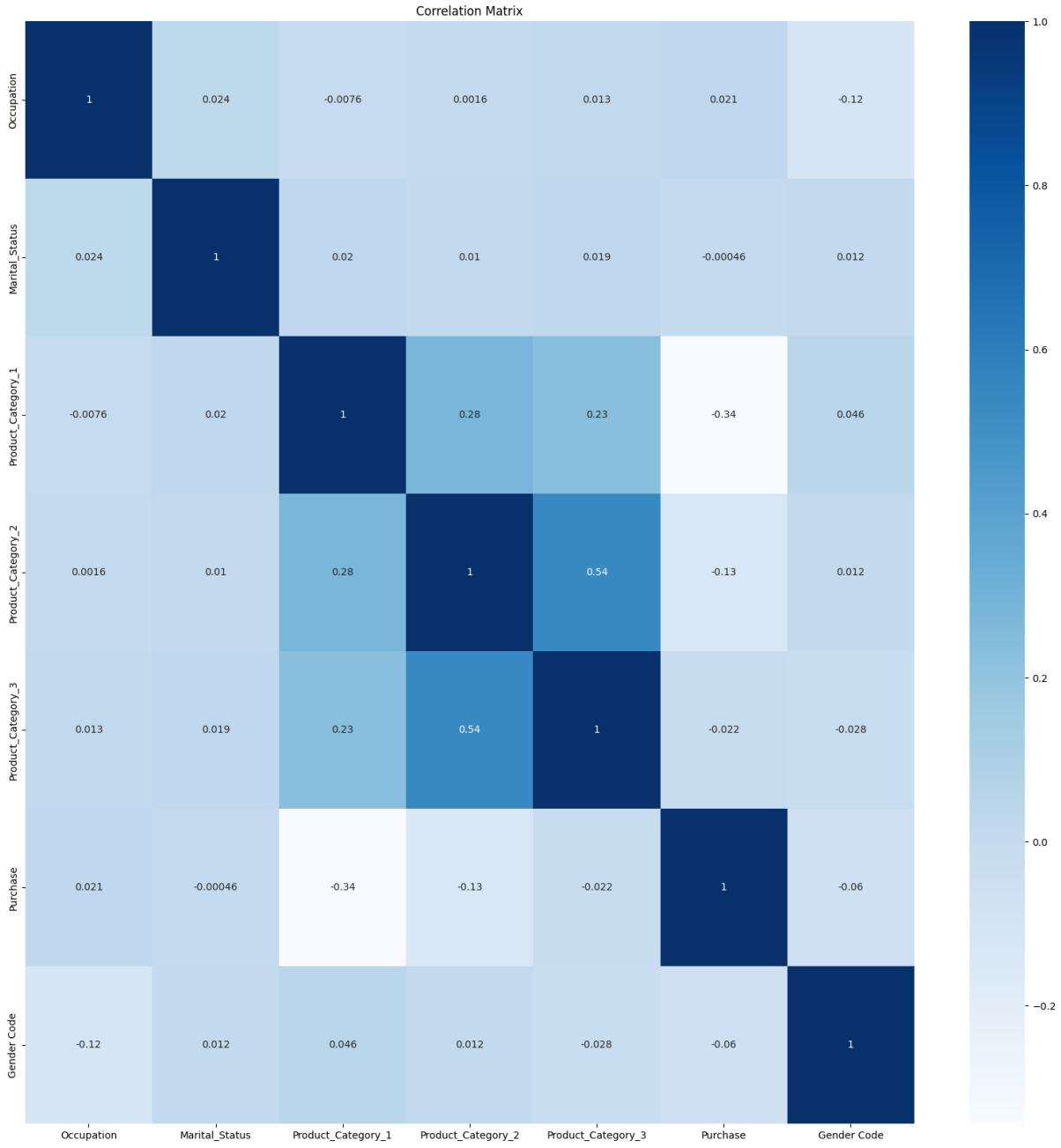






Assignment 2
Divyanshu Vaibhav
21BT10014

Experiment 1(EDA and correlation heatmaps)



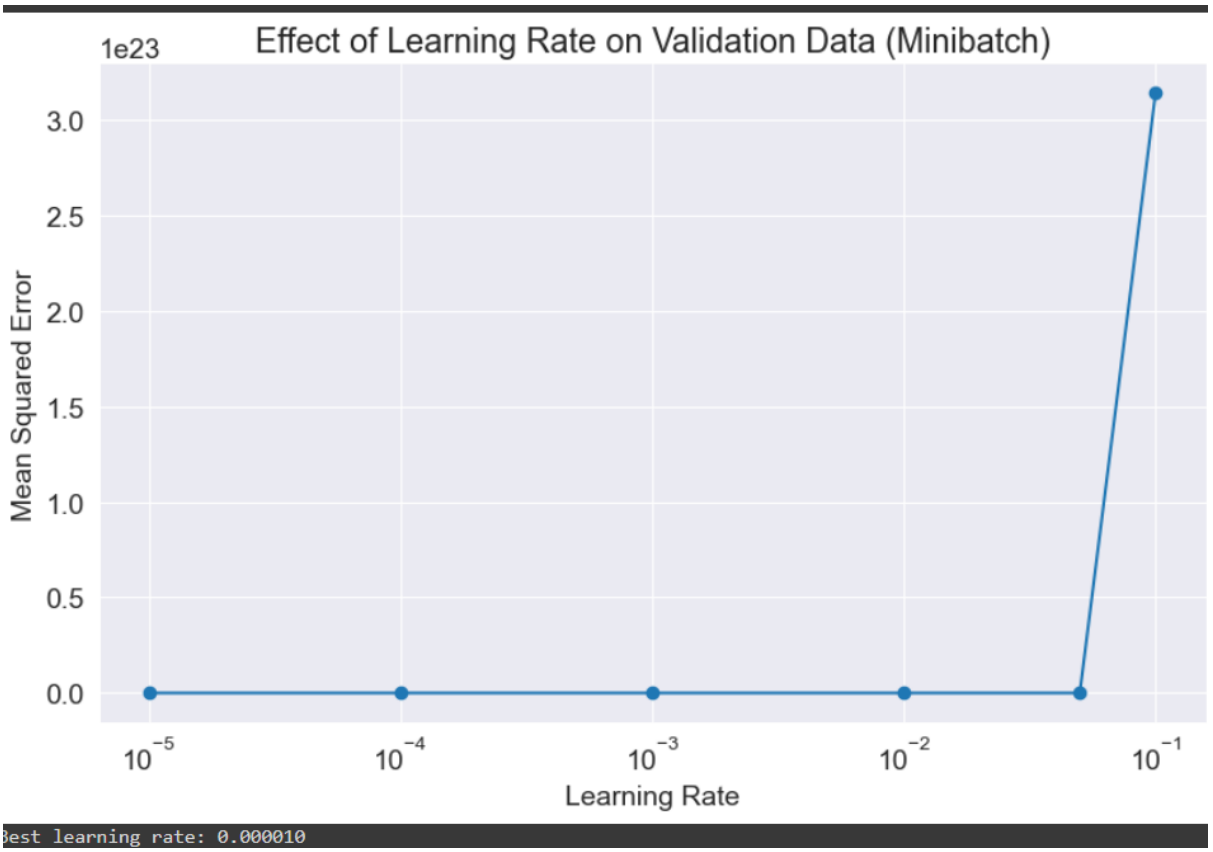
Assignment 2
Divyanshu Vaibhav
21BT10014

Experiment 1(EDA and correlation heatmaps)

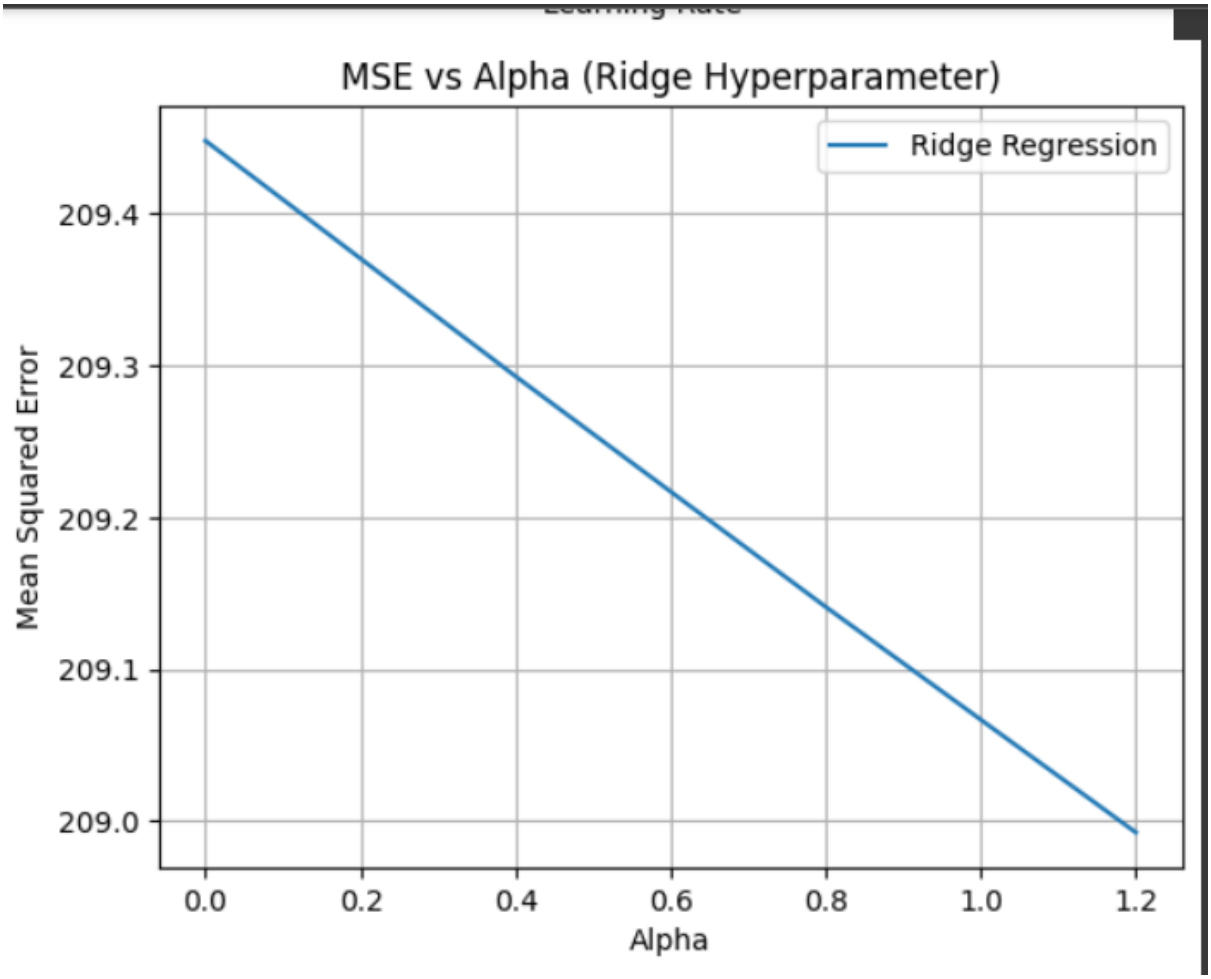
Experiment 2(Performance values with and without scaling)

Scaled	YES	NO
MSE	`MSE value without scaled data: 199.7851273260234`	`MSE value with scaled data: 197.18436915017926`

Experiment 3 and 4(Linear model grad and Linear model ridge) performance metrics



The best learning rate is 0.00001.



The best value of MSE is 209.2 at Alpha=0.6

Experiment 5: Comparison for MSE values for all three models

LIN_CLOSED	LIN_GRAD	LIN_RIDGE
199.78	230	209.05

- As the Alpha is increased the regularisation also decreases so MSE increases. It should be as low as possible.
- We need to make a balance between learning factor
- Lower Alpha: When alpha is set to a very small value or even zero, ridge regression behaves like ordinary linear regression. The model will try to fit the data as closely as possible, which can lead to overfitting if the data has noise or outliers. In this case, the training MSE might be low, but the model might perform poorly on new, unseen data (test data).

- Higher Alpha: As alpha increases, the model's emphasis on keeping coefficients small becomes stronger. This can help in reducing overfitting and generalising better to new data. However, as alpha becomes very large, the model may underfit the training data and perform poorly on both the training and test sets.