ASSIGNMENT 4

Q1

Errors in each technique:

```
For sklearn LinearRegression :

RMSE:  0.6207498618060853
MAE:  0.5126346094502968
--------------------------
For linear regression using normal equations :

RMSE:  0.6207498618060854
MAE:  0.5126346094502968
--------------------------
For linear regression using SVD :

RMSE:  0.6207498618060853
MAE:  0.5126346094502968
--------------------------
```

Time for each technique :

```
Time taken by Sklearn's implementation: 0.0027539730072021484 seconds
Time taken by normal equations: 0.0002944469451904297 seconds
Time taken by SVD: 0.004835128841796875 seconds
```

Q2

```
Batch Gradient Descent with manual gradient computation for unregularized objective :
 Batch size= 1 , RMSE:  0.7905046278690018
 Batch size= 1 , MAE:  0.6314635711163878
---------------------------
Rigde regression with manual gradient computation :
 Batch size= 1 , RMSE:  0.8254707844072325
 Batch size= 1 , MAE:  0.6304058871895001
---------------------------
Batch gradient descent with jax gradient computation for unregularized objective :
 Batch size= 1 , RMSE:  0.7995486600057372
 Batch size= 1 , MAE:  0.6487486828744025
---------------------------
Ridge regression with jax gradient computation :
 Batch size= 1 , RMSE:  0.7917161721091319
 Batch size= 1 , MAE:  0.6298493002999888
---------------------------
LASSO :
 Batch size= 1 , RMSE:  0.832308215077365
 Batch size= 1 , MAE:  0.6957399363046195
---------------------------
SGD with momentum for manual gradient computation on unregularized objective :
 Batch size= 1 , RMSE:  0.8055490409646227
 Batch size= 1 , MAE:  0.6268142945465978
---------------------------
Ridge regression on SGD with momentum  using manual computation :
 Batch size= 1 , RMSE:  0.9785720532637601
 Batch size= 1 , MAE:  0.8140122430647128
---------------------------
SGD with momentum for jax gradient on unregularized:
 Batch size= 1 , RMSE:  0.8892927066451223
 Batch size= 1 , MAE:  0.7061438121083694
---------------------------
```

```
---------------------------
SGD with momentum for jax gradient on unregularized:
 Batch size= 1 , RMSE:  0.8892927066451223
 Batch size= 1 , MAE:  0.7061438121083694
---------------------------
Ridge regression using jax using SGD :
 Batch size= 1 , RMSE:  0.7946100588309402
 Batch size= 1 , MAE:  0.6417867355312361
---------------------------
LASSO using SGD:
 Batch size= 1 , RMSE:  0.8280695290241143
 Batch size= 1 , MAE:  0.6708981730103792
Final data:
                  Regressor Method of calculating gradient   Penalty type  \
0    Batch gradient descent                            manual  unregularized
1    Batch gradient descent                            manual  unregularized
2    Batch gradient descent                            manual  unregularized
3    Batch gradient descent                            manual  unregularized
4    Batch gradient descent                            manual  unregularized
..                      ...                               ...            ...
395       SGD with momentum                               jax             l1
396       SGD with momentum                               jax             l1
397       SGD with momentum                               jax             l1
398       SGD with momentum                               jax             l1
399       SGD with momentum                               jax             l1

     Learning rate  Lambda Momentum      RMSE       MAE  Training time
0            0.001   0.001       NA  0.869777  0.681868       0.169184
1            0.001   0.010       NA  1.204946  1.018111       0.170525
2            0.001   0.100       NA  1.497003  1.245115       0.158658
3            0.001   1.000       NA  1.508113  1.172299       0.177218
4            0.010   0.001       NA  0.805932  0.648113       0.168406
```

```
396          1.000   1.000      -3      NaN       NaN      1.438138
397          1.000   1.000      -2      NaN       NaN      1.431901
398          1.000   1.000      -1      NaN       NaN      1.490301
399          1.000   1.000       0      NaN       NaN      1.710637

[400 rows x 9 columns]
-------------------------------------

Optimal hyperparameters for batch gradient descent based on RMSE:
                   Regressor Method of calculating gradient Penalty type  \
58  Batch gradient descent                             jax           l2

     Learning rate  Lambda Momentum    RMSE      MAE  Training time
58             0.1     0.1       NA  0.792424  0.618377      2.312205
-------------------------------------

Optimal Hyperparameters for batch gradient descent based on MAE
                   Regressor Method of calculating gradient Penalty type  \
58  Batch gradient descent                             jax           l2

     Learning rate  Lambda Momentum    RMSE      MAE  Training time
58             0.1     0.1       NA  0.792424  0.618377      2.312205
-------------------------------------

Optimal hyperparameters for SGD with momentum based on RMSE
              Regressor Method of calculating gradient Penalty type  \
296  SGD with momentum                             jax           l2

      Learning rate  Lambda Momentum    RMSE      MAE  Training time
296            0.01     0.1       -3  0.790582  0.62824      1.553014
-------------------------------------

Optimal hyperparameters for SGD with momentum based on MAE
              Regressor Method of calculating gradient   Penalty type  \
250  SGD with momentum                             jax  unregularized

      Learning rate  Lambda Momentum    RMSE      MAE  Training time
250            0.1     0.1       -1  0.800185  0.617581      0.979765
```
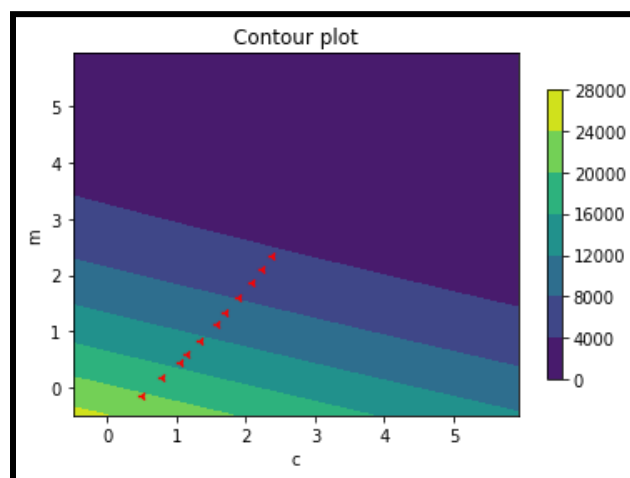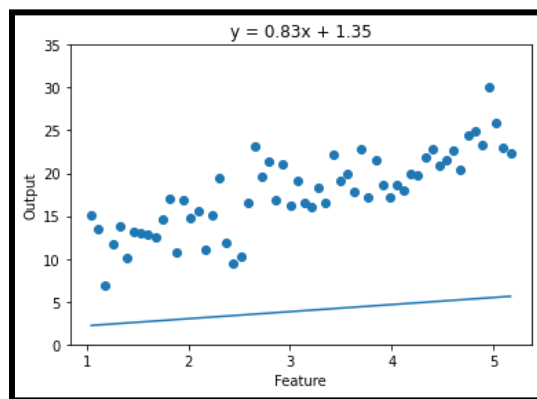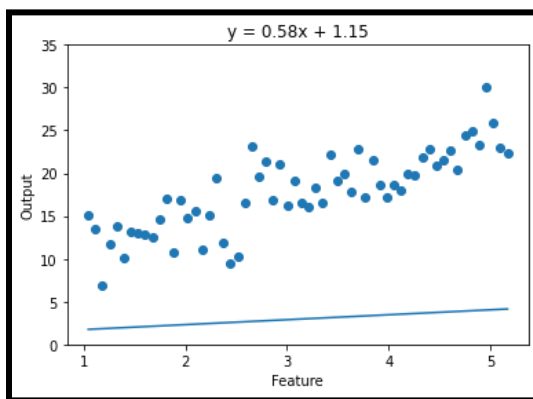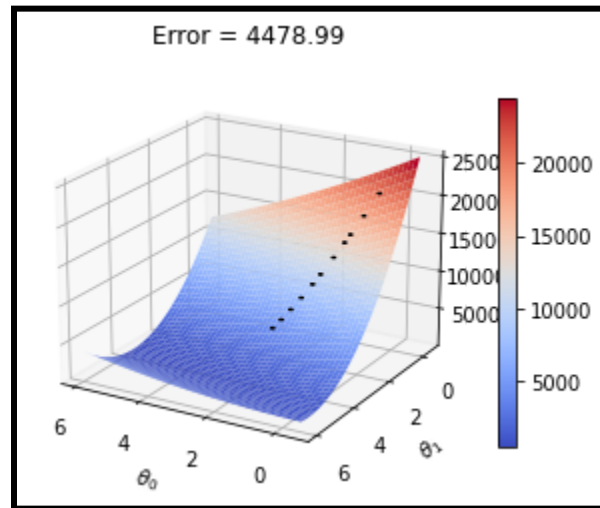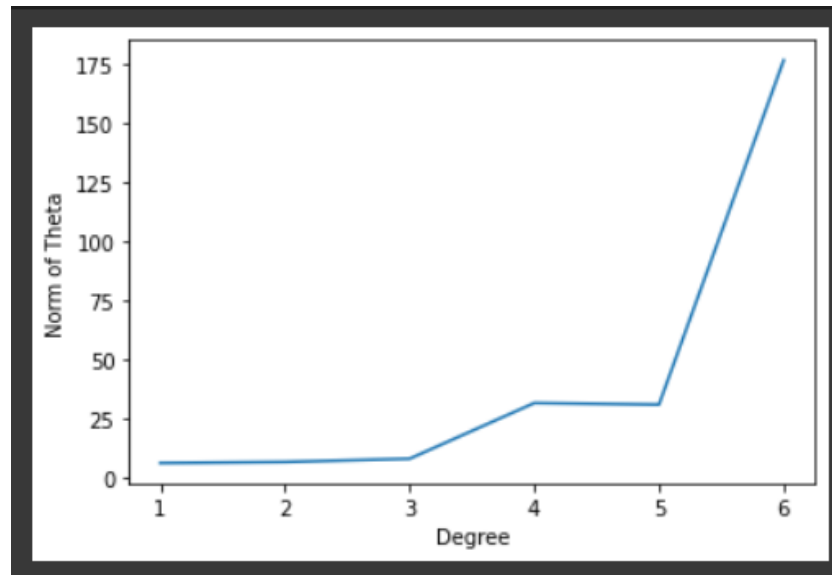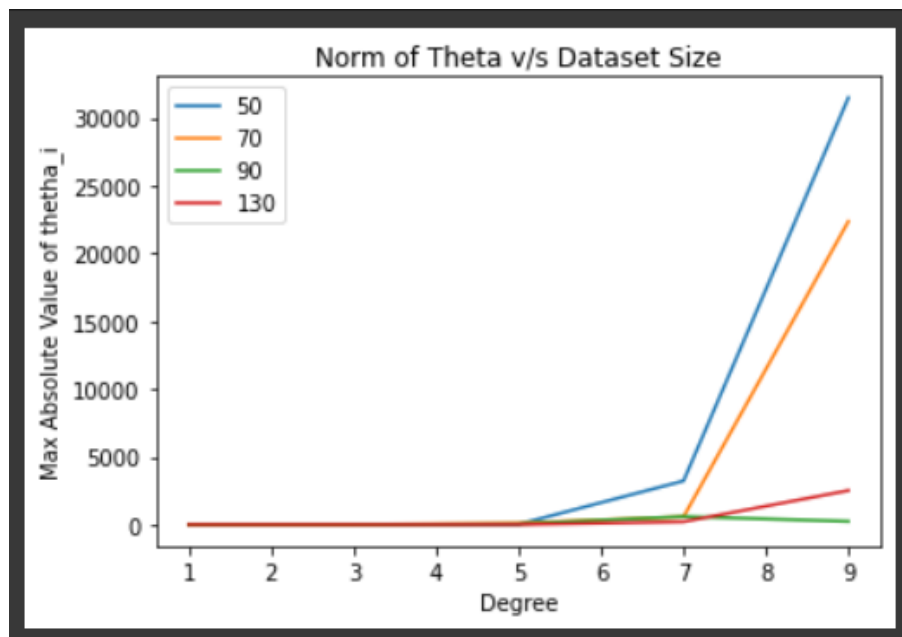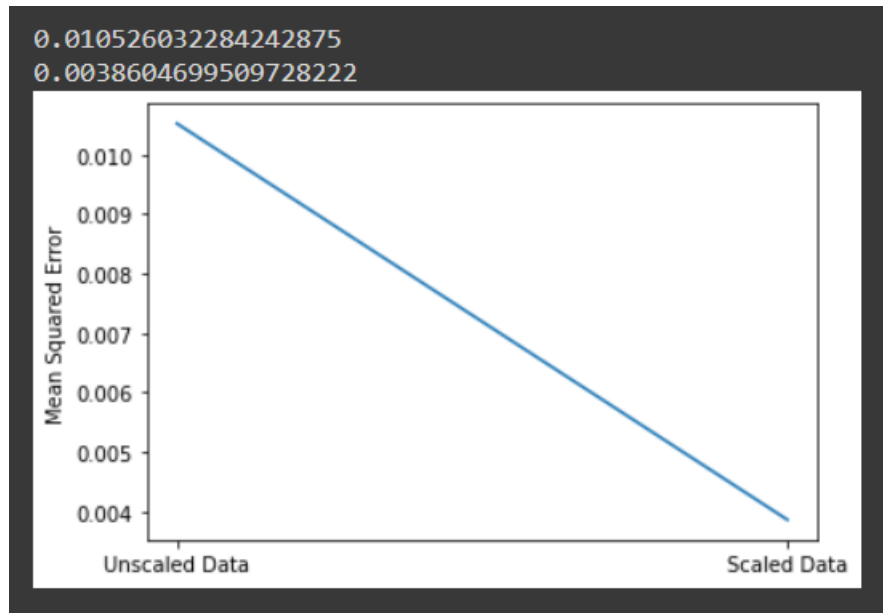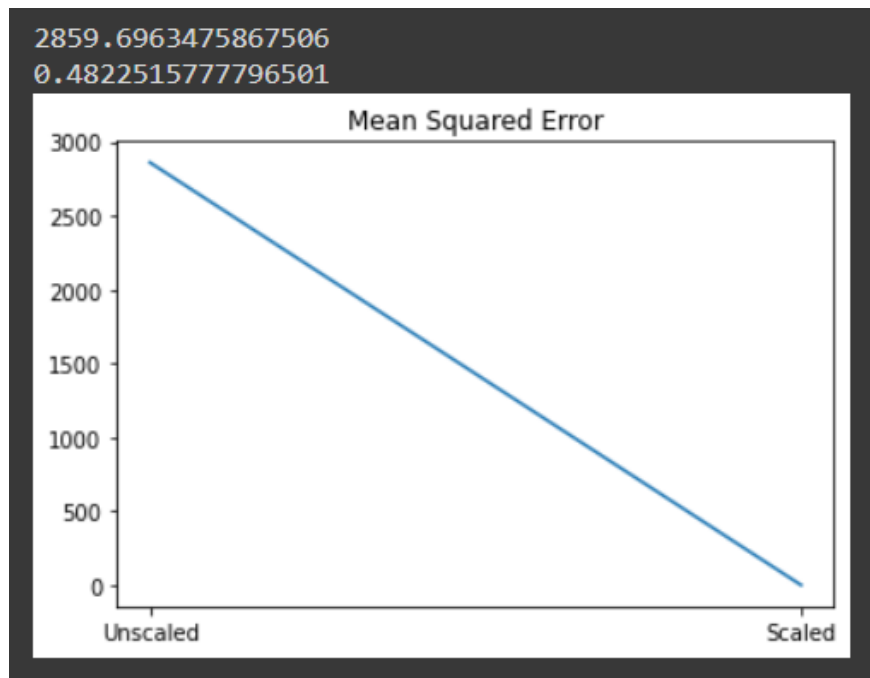
Q3   Gif uploaded on Git

Error = 4478.99

y = 0.58x + 1.15

y = 0.83x + 1.35

Contour plot

Q4

Norm vs Degree

Q5

Q6

For random dataset:


```
0.010526032284242875
0.0038604699509728222
```

For diabetes dataset:


```
2859.6963475867506
0.4822515777796501
```

Q7

Train RMSE: 2.619
Test RMSE: 2.377