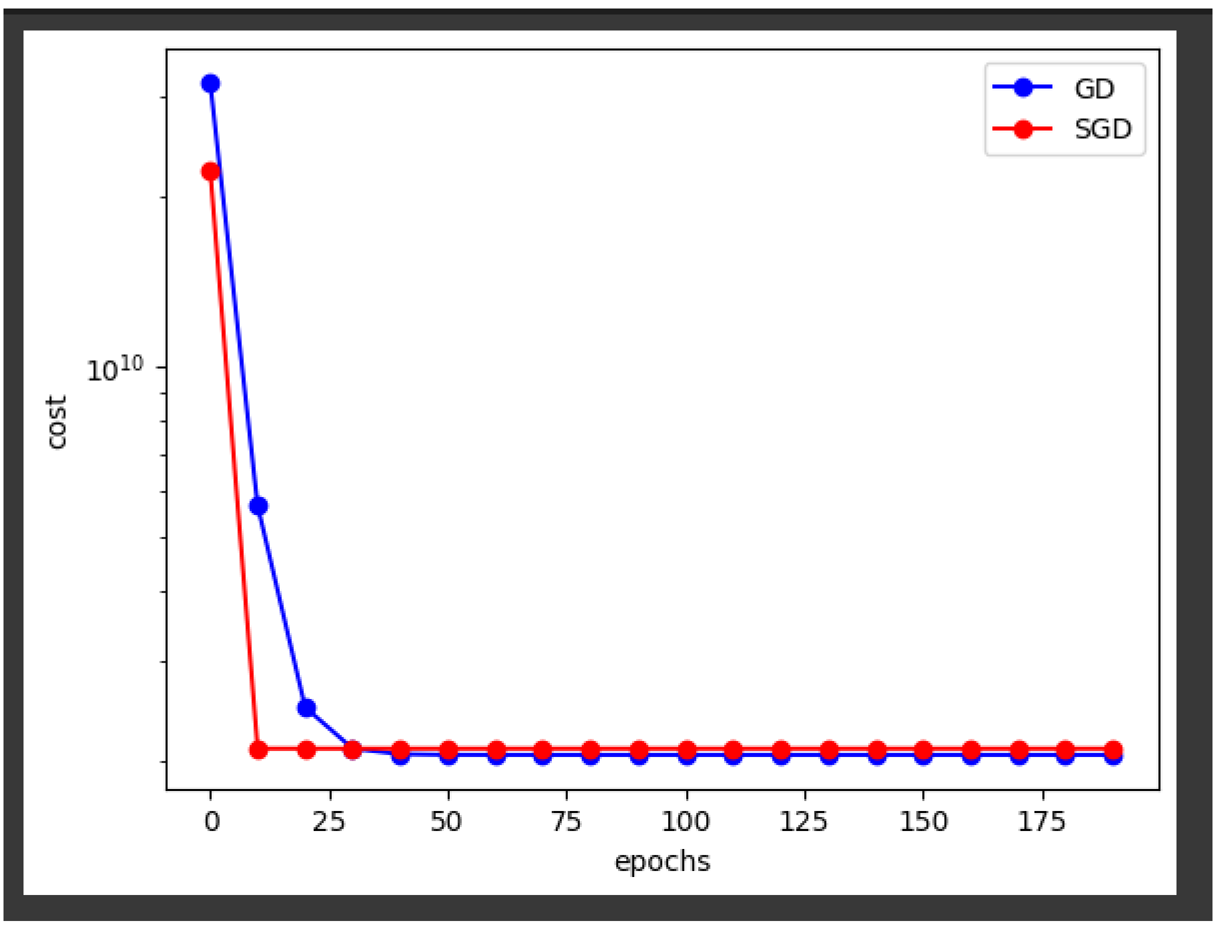
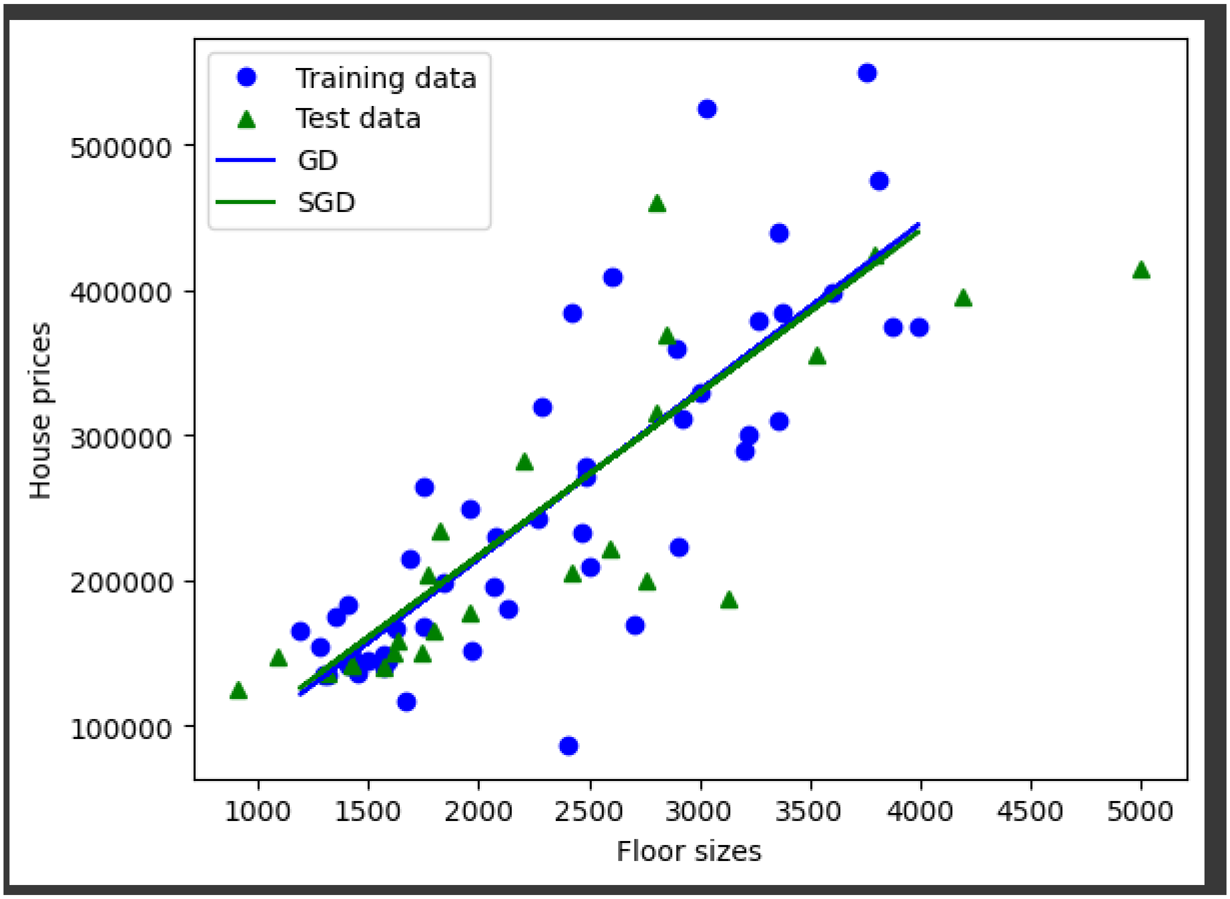
Report: Comparison of Gradient Descent (GD) and Stochastic Gradient Descent (SGD) for Linear Regression

Simple Output:

```
Params GD:  [254449.99982048  93308.92004027]
Params SGD:  [255138.8717082   90684.56621139]
Training RMSE (GD): 64083.51.
Training cost (GD): 2053348364.32.
Training RMSE (SGD): 64140.93.
Training cost (SGD): 2057029253.11.
Test RMSE (GD): 65773.19.
Test cost (GD): 2163056350.22.
Test RMSE (SGD): 64683.16.
Test cost (SGD): 2091955816.68.
```

Plots:

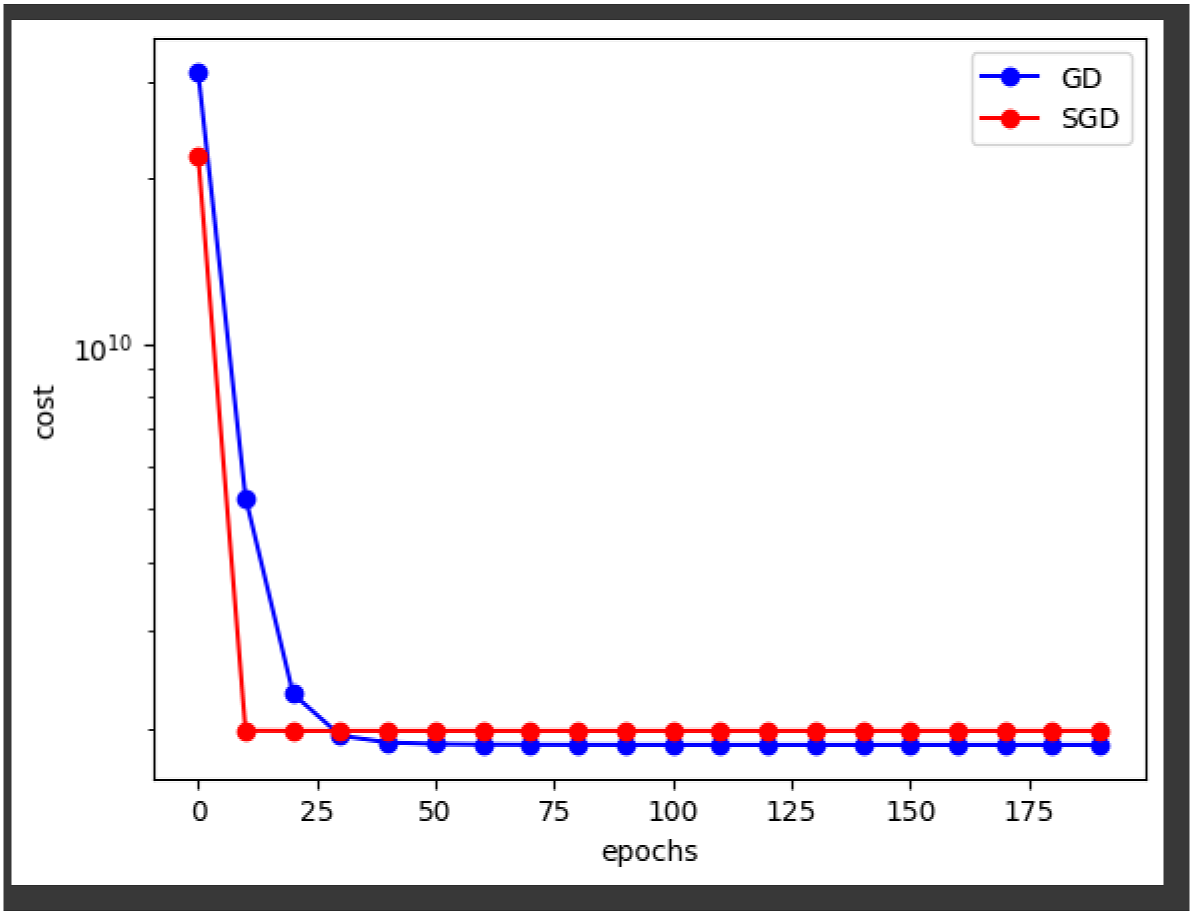




Multiple Output:

```
Params GD:  [254449.99982048  78079.18106675  24442.5758378    2075.95636731]
Params SGD:  [255277.40615601  76431.09645851  22488.74314279   1627.00666084]
Training RMSE: 61070.62.
Training cost: 1864810304.94.
Test RMSE: 58473.59.
Test cost: 1709580288.69.
Training RMSE SGD: 61164.73.
Training cost SGD: 1870562156.35.
Test RMSE SGD: 57361.06.
Test cost SGD: 1645145407.30.
```

Plots:



Discussion:

Simple Output

Both GD and SGD provide similar results in terms of training and test performance metrics. However, SGD tends to converge faster due to its stochastic nature.

Multiple Output

SGD slightly outperforms GD in terms of both training and test performance. This could be attributed to SGD's ability to handle large datasets more efficiently.

Overall, both GD and SGD perform reasonably well for linear regression tasks. The choice between the two depends on factors such as dataset size, computational resources, and convergence speed requirements.