

Part 1

2. Consider only one feature `sqft_lot`. Run the linear regression algorithm using only this feature and set $\eta=0.005$. Then, evaluate on the instances in `test.csv`. What is the root mean-square error on these instances?

335813.66057

3. Create subsets of the training data containing the first 2000, 5000, 7000, 10000 instances in `train.csv`. Create and submit a plot where the X-axis is the size of the training set (2000, 5000, 7000, 10000, full) and the Y-axis is the root mean-square error on the `test.csv` instances using only the `sqft_lot` feature.

Training data size	error (as shown on Kaggle)
2000	336757.32835
5000	336427.18851
7000	336019.77801
10000	335965.38488
full (17613)	335813.66057

