

This linear regression model predicts listing prices of apartments in Manhattan using data from the Douglas Elliman website.

Since square feet is the most highly correlated feature variable with the target variable price, I will start by creating a single variable linear regression using sq ft to predict listing price.

The linear model is as follows:

$$\text{Listing price} = -992435.75 + (\text{square feet}) * 2658.59$$

This model achieves an  $R^2$  value of 0.705 on the test data.

Here is a graph of sq ft versus price including the regression line.

