3.1 Explore Scikit-learn Dataset

3.1.1 Get n_features and n_samples

```
Number of features in the Boston dataset is: 13
Number of samples in the Boston dataset is: 506
```

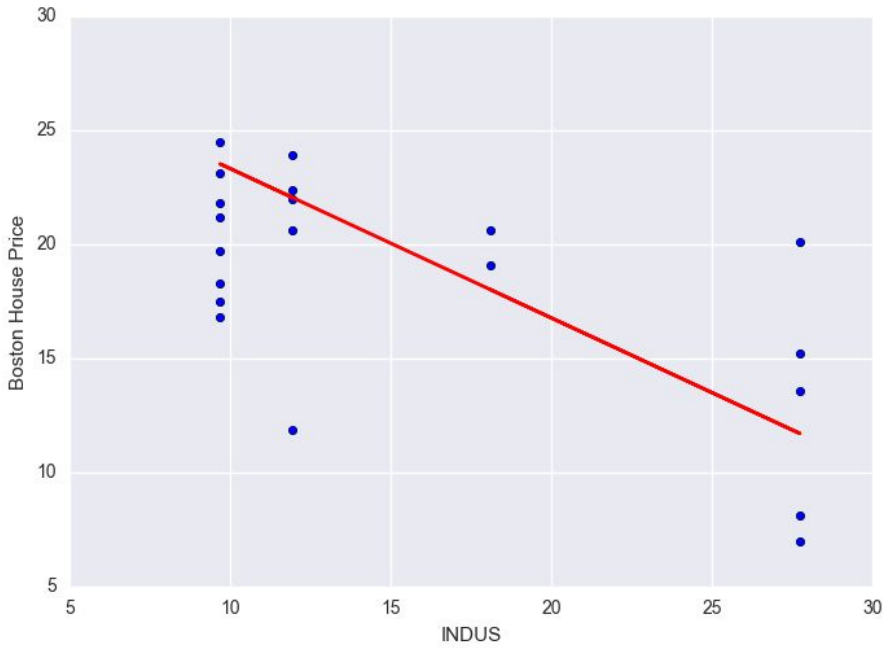
3.1.2 Find best fitted feature

```
Best fitted feature name is: INDUS
Best fitted model score is: 0.205969
```

3.1.3 Calculate the loss function

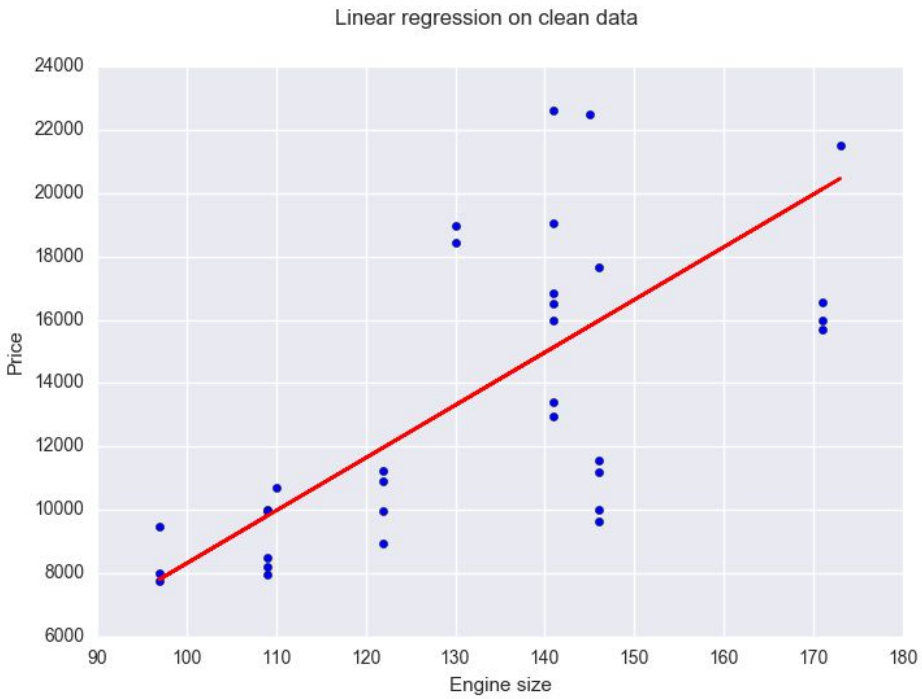
```
Value of the loss function for the best fitted model is: 18.564536
```

3.1.4 Plot the predictions and test data



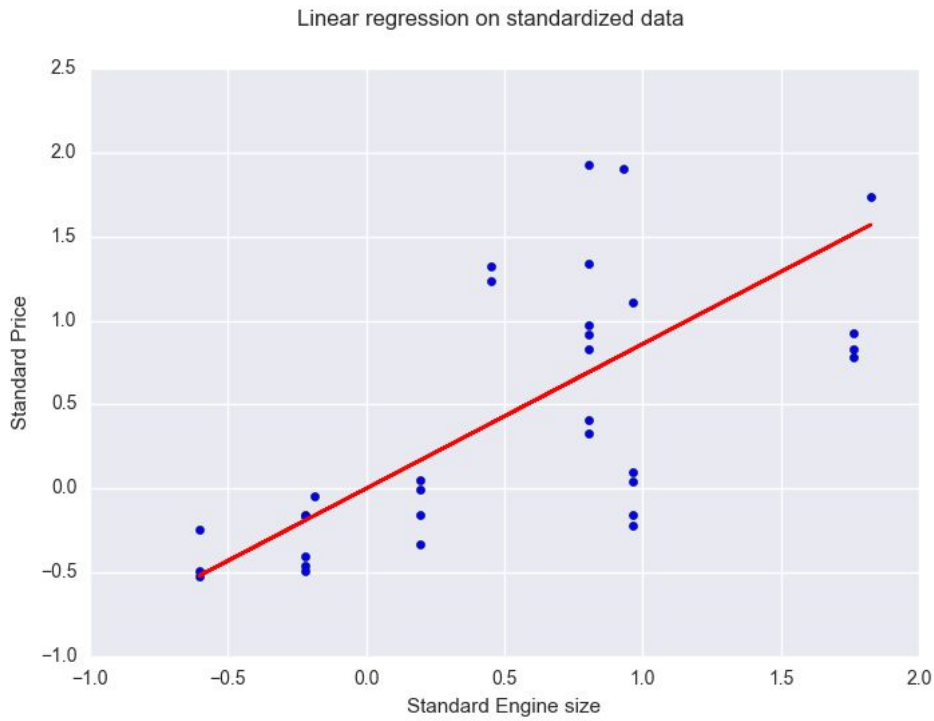
3.2 Explore Raw Dataset

3.2.3 Linear regression on the cleaned data




```
Price prediction for engine size equals to 175 is: 20793.53
```

3.2.4 Linear regression on the standardized data



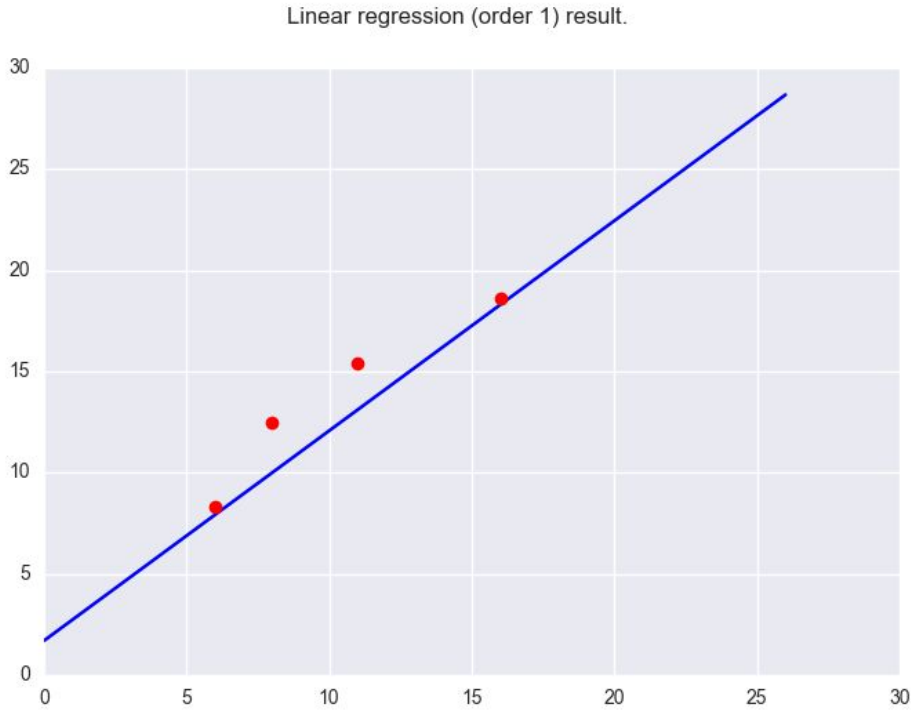
3.2.5 Linear regression with multiple features

```
Parameter theta calculated by normal equation: 0.000, 0.862, 0.074
Parameter theta calculated by SGD: 0.003, 0.723, -0.009
```

3.3 Understand Regularization

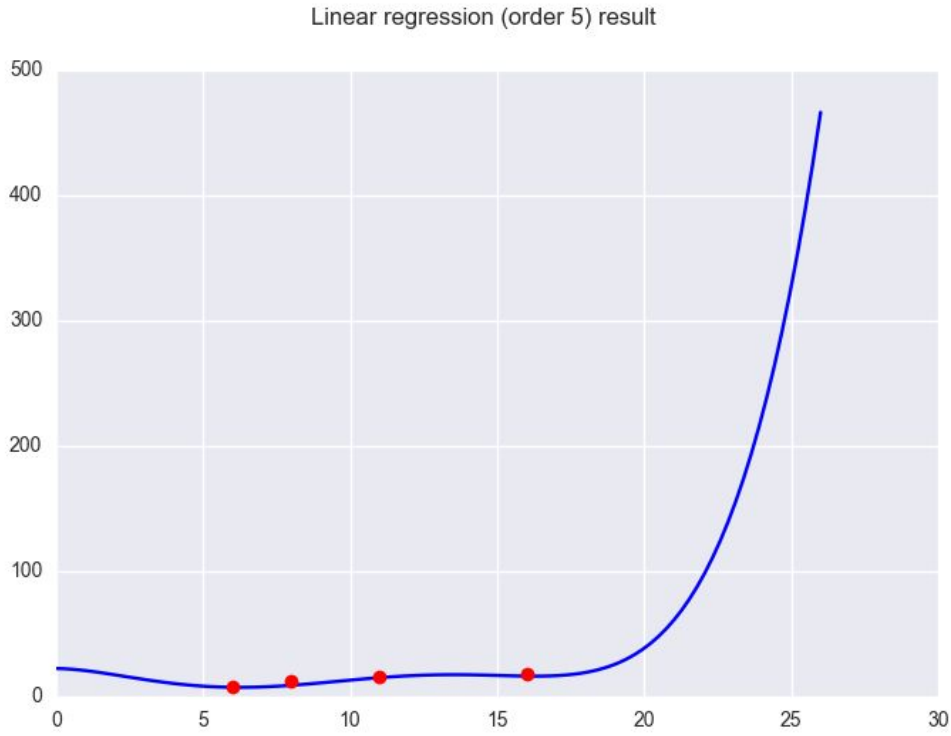
3.3.1 LR regression on polynomial data

```
y1 = 1.70 + 1.04x
Linear regression (order 1) model score is: 0.796
```



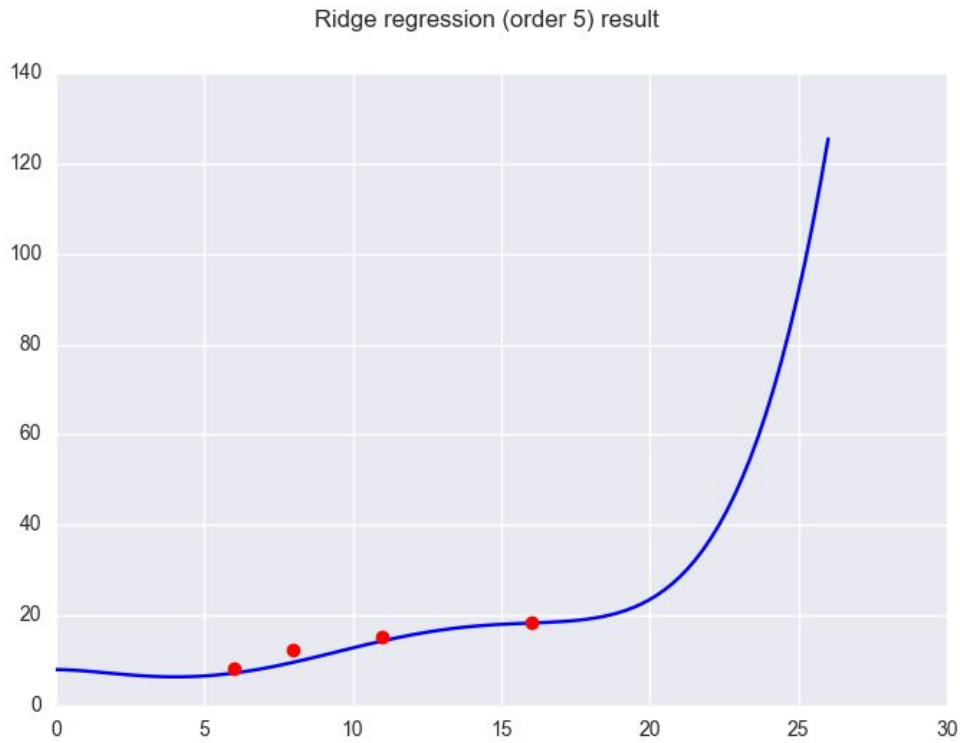
3.3.2 Polynomial regression on training data

```
y2 = 22.51 - 0.345x  - 1.69x*x + 0.341x*x*x - 0.0228x*x*x*x  +
0.000509x*x*x*x*x
Linear regression (order 5) score is: 0.706
```



3.3.3 Ridge Regression

```
Ridge regression (order 5) score is: 0.821
y3 = 8.048 - 0.0708x  - 0.345x*x  +  0.0879x*x*x  - 0.00615x*x*x*x +
0.000136x*x*x*x*x
```



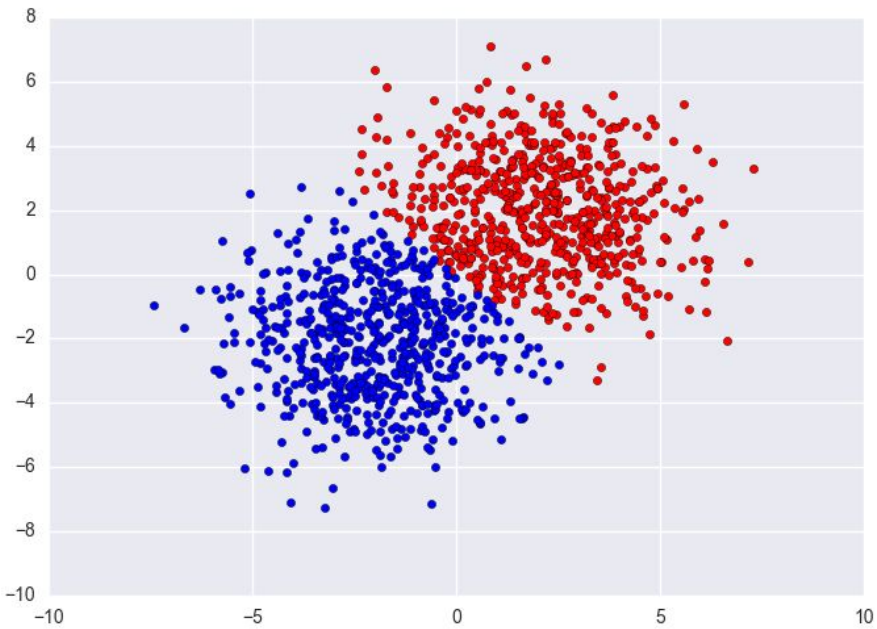
3.3.4 Comparisons

```
The model with the highest score is: Ridge model
Ridge model can prevent over-fitting: yes
Ridge model is nearly equivalent to LR model (order 5) if alpha=0: yes
A larger alpha results in a larger coefficient for x*x*x*x*x: no
```

4 Linear Discriminant/Classification

4.1 Binary Classification

```
The predictions only have 0 and 1: yes
```



4.2 Classification Statistics

```
Number of wrong predictions is: 73
```