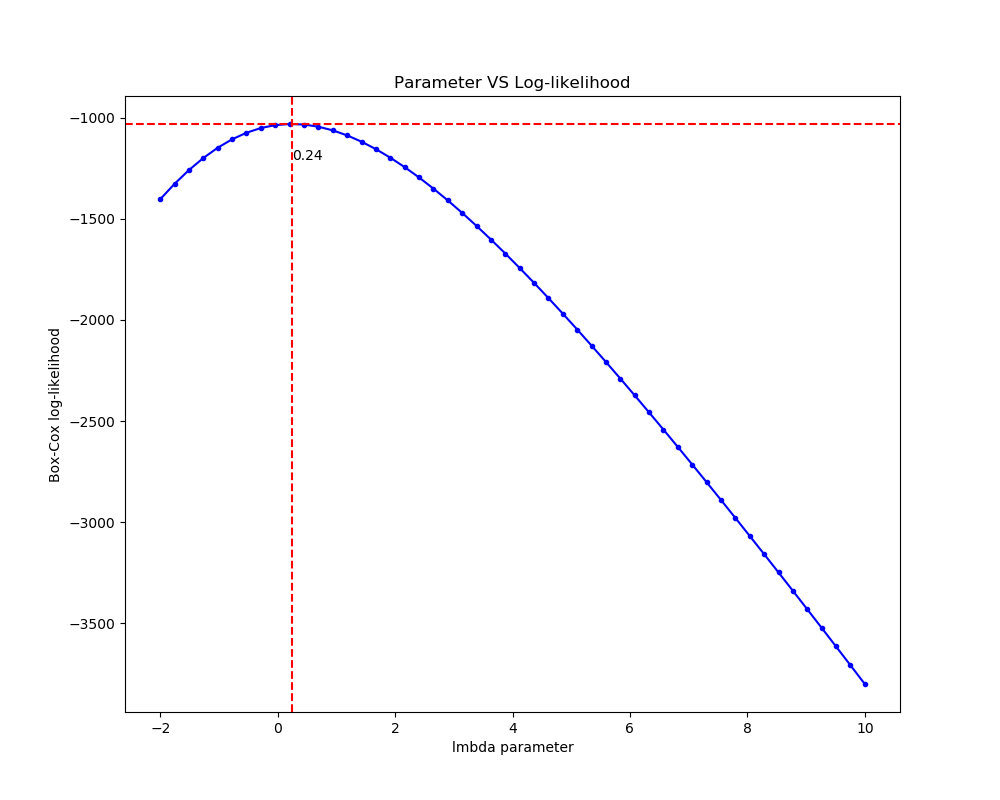
# Homework 6

1. **List all the points (row numbers) you removed (indexed on the original dataset) as outlier points:**

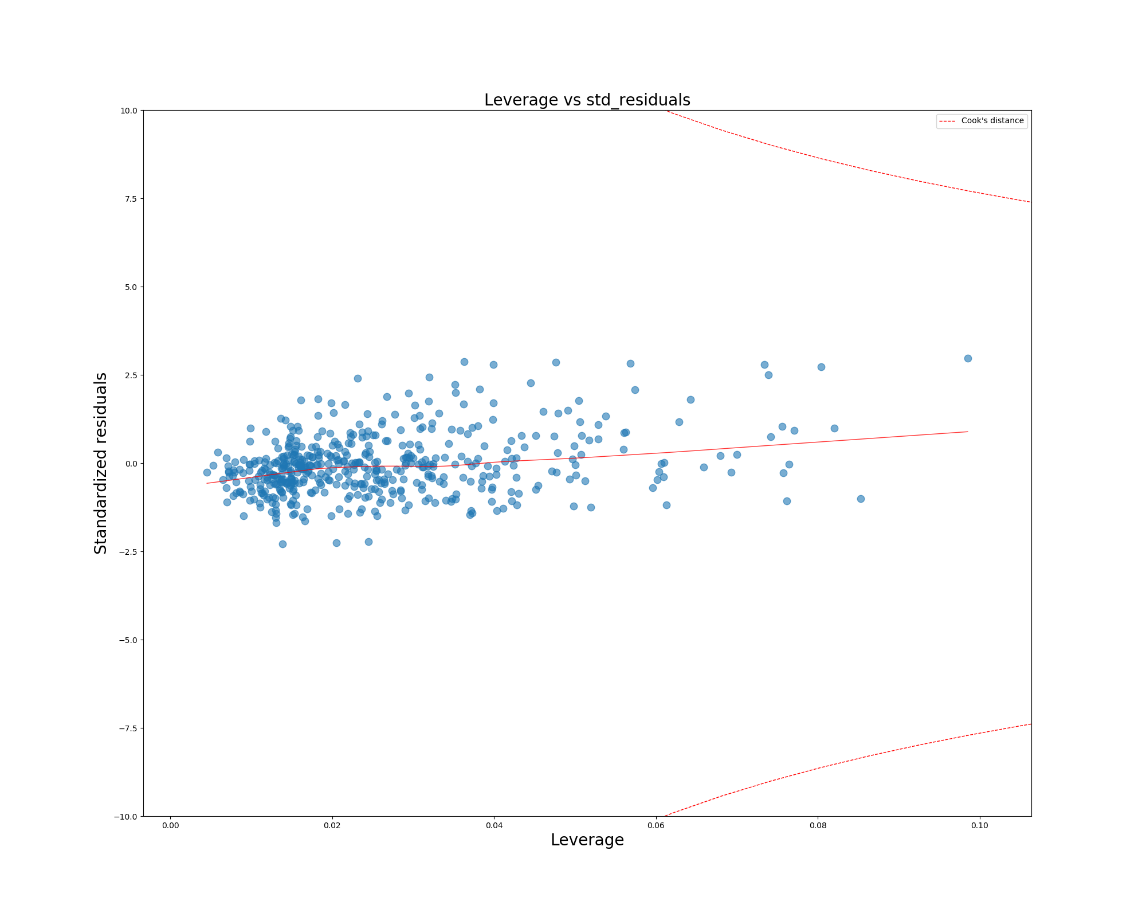
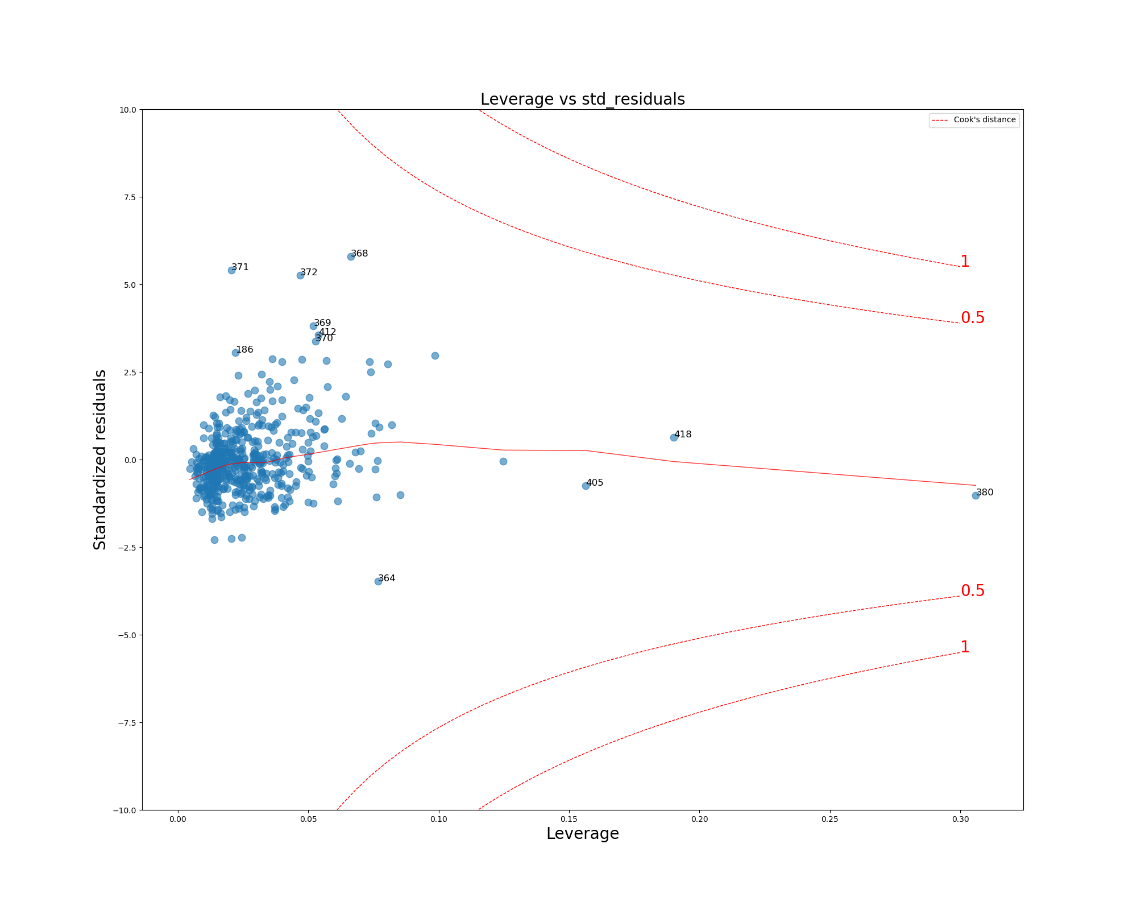


1. **Box-Cox Transformation – Plot the Box-Cox transformation curve (Log-likelihood vs Parameter value). What is the best value of the parameter you got?**

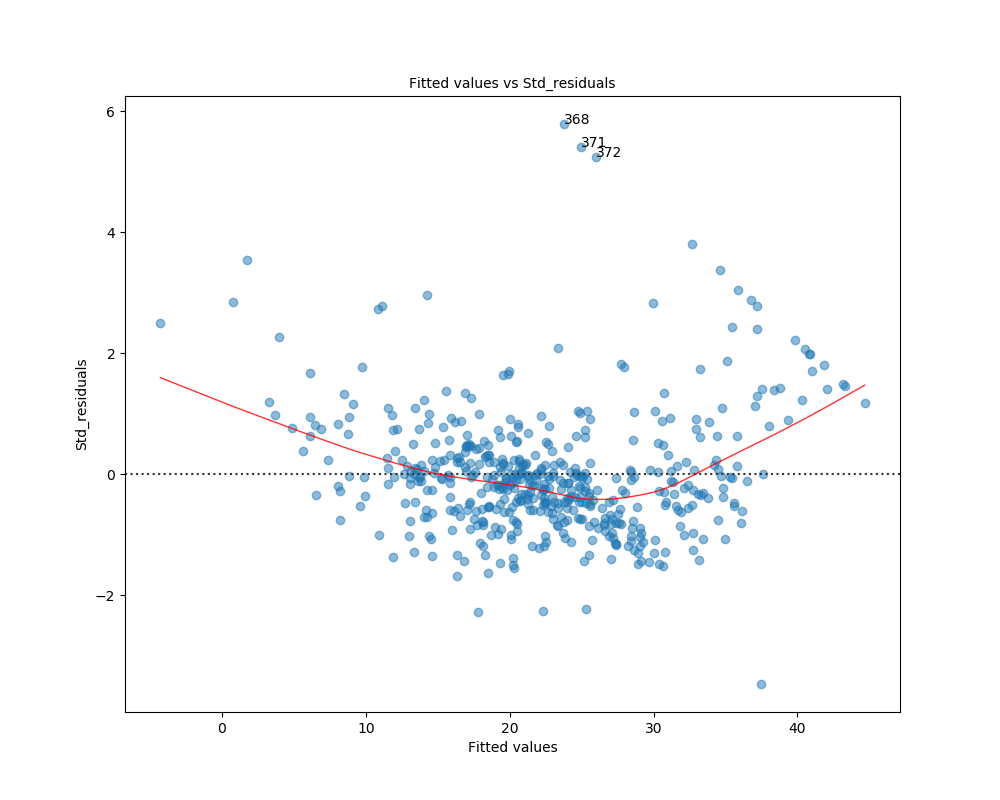


****

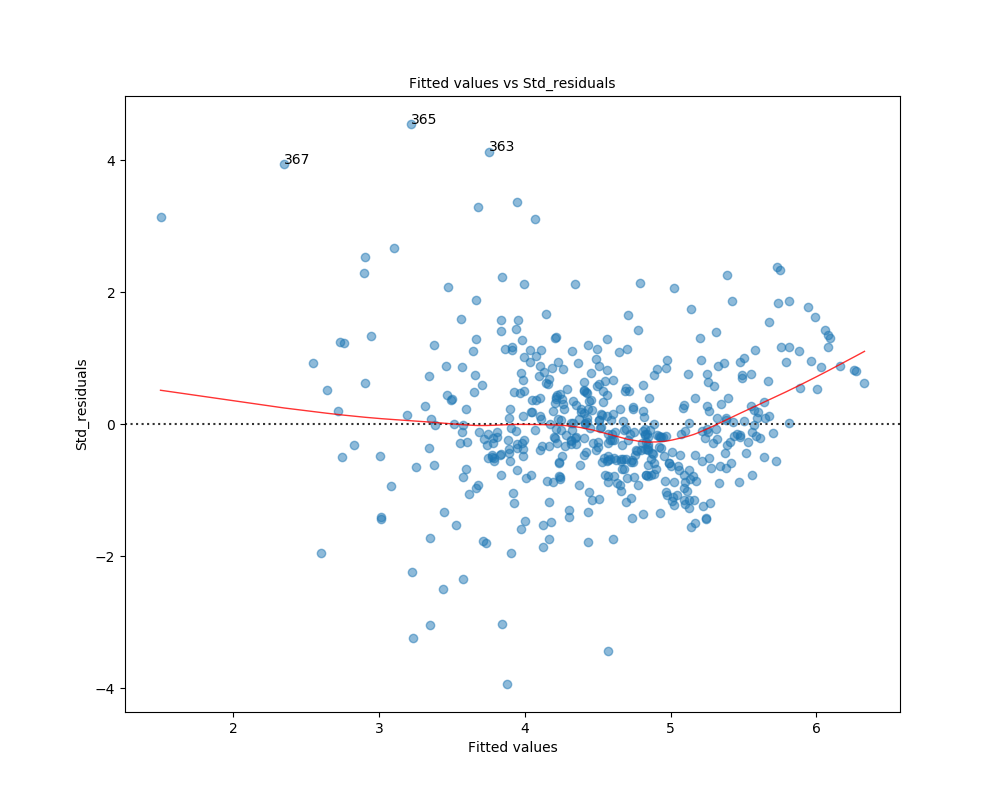
1. **Diagnostic plots used for identification of outliers. Please only include the Standard residuals vs Leverage vs Cook’s distance plots (do not put other 3 plots you obtain for R). The final diagnostic plot obtained after removing all outliers should also be included:**

****

1. **Plot of Standardized residuals vs Fitted values for the linear regression model obtained without any transforms (like removing outliers or transforming dependent variables).**

****

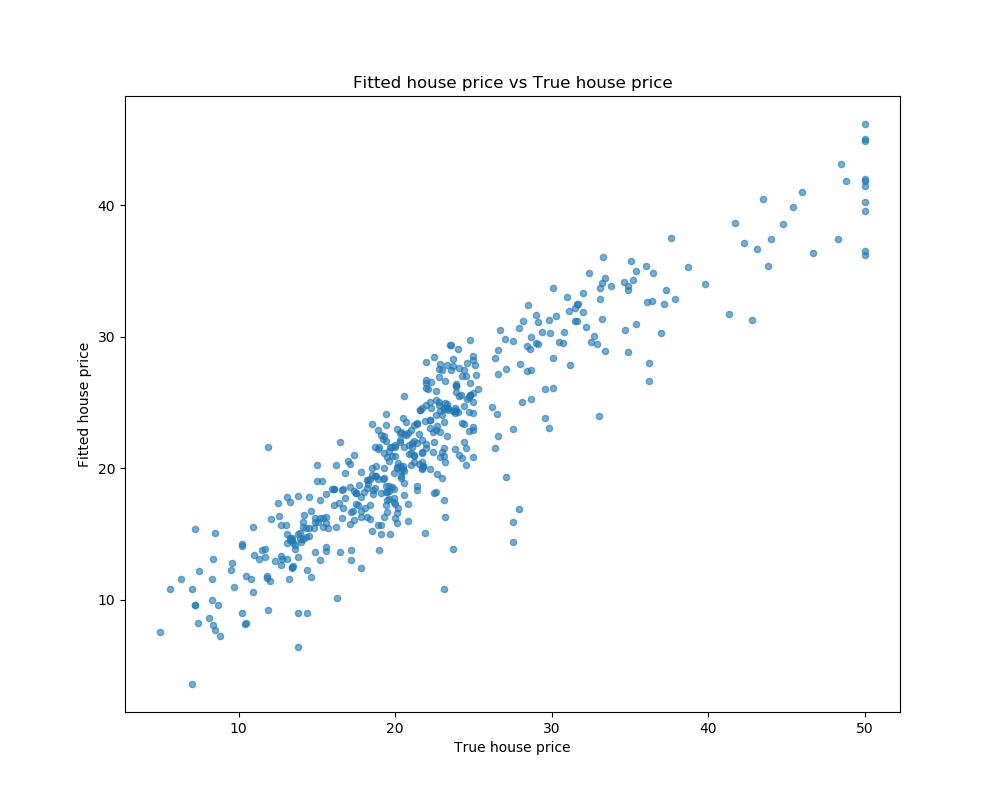
1. **Plot of Standardized residuals vs Fitted values for the final linear regression model obtained after removing all outliers and transforming the dependent variable.**

****

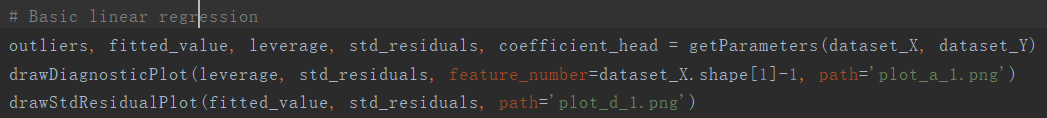
1. **Compare the two plots. What do you observe?**

After removing all outliers and transforming the dependent variable, the data concentrate more in smaller scale and the range of fitted values becomes smaller.

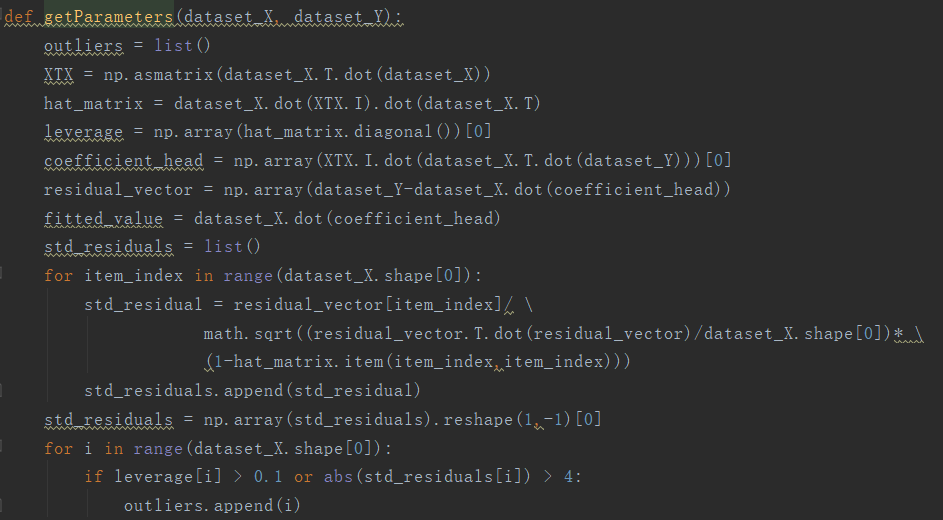
1. **Final plot of Fitted house price vs True house price. What do you observe.**

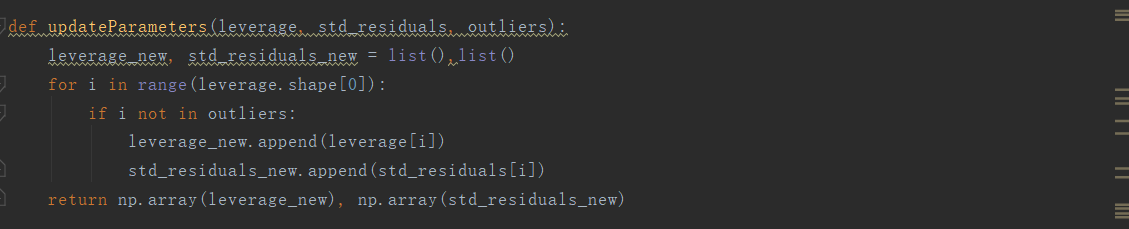
****

1. **Code screenshot**
2. **Linear regression**

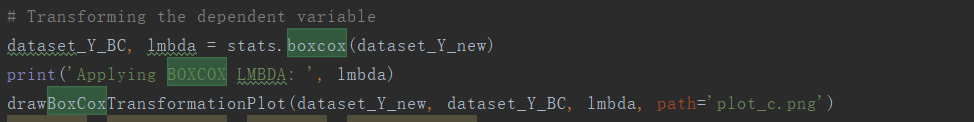


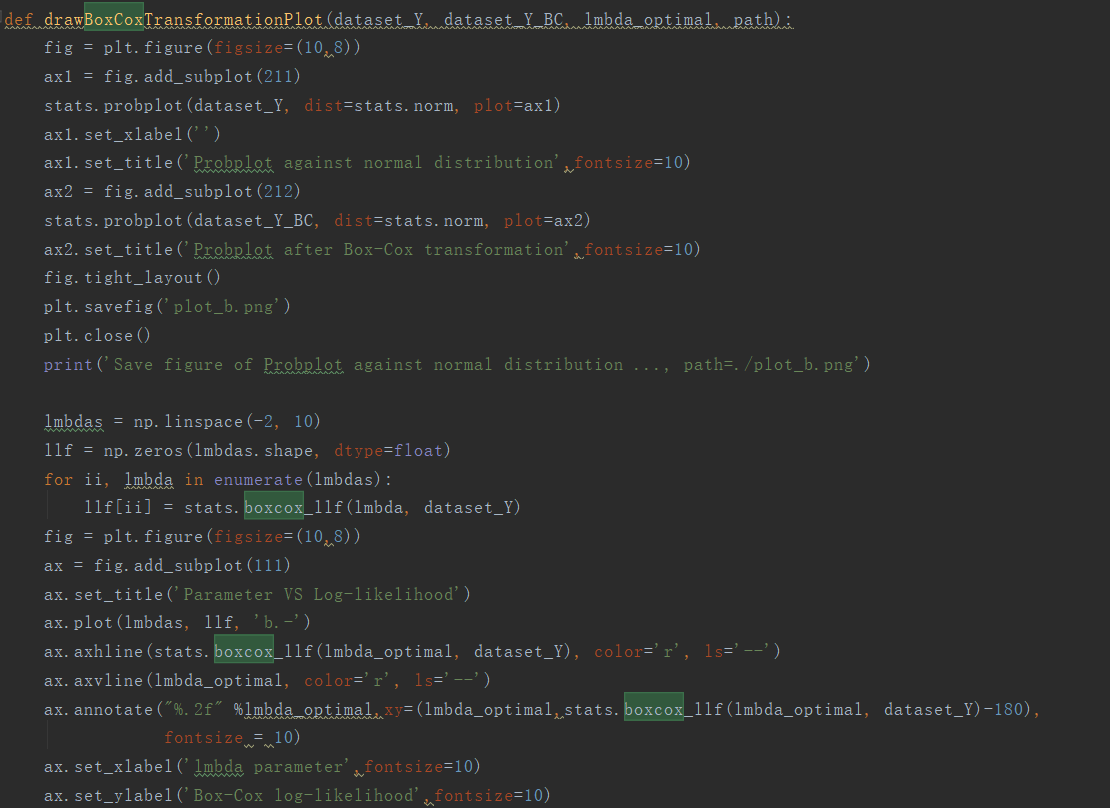




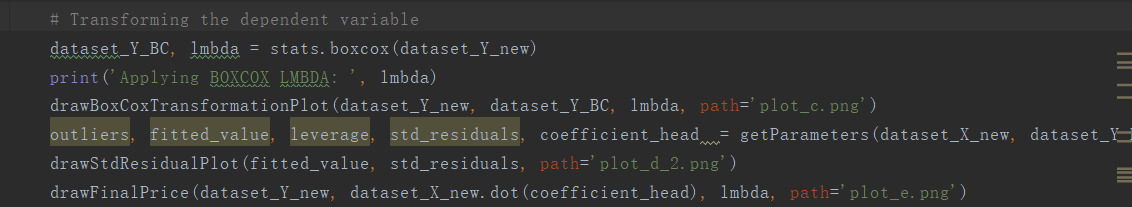


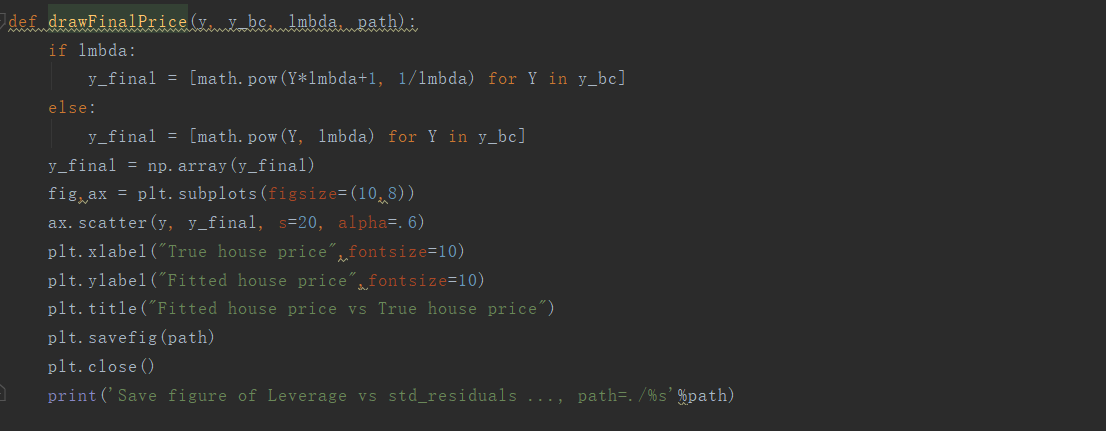
1. **Box-Cox transformation**





1. **Used the parameter value to transform the dependent variable**





1. **Entire code**

