**Linear Regression Homework 5**

**Name**: Eric Yuan

**UNI**: qy2205

1. **Covariance matrix of**

Since is a constant matrix. So we have

1. **Show that and are equivalent in the sense that the where is the and is the**

Full model:

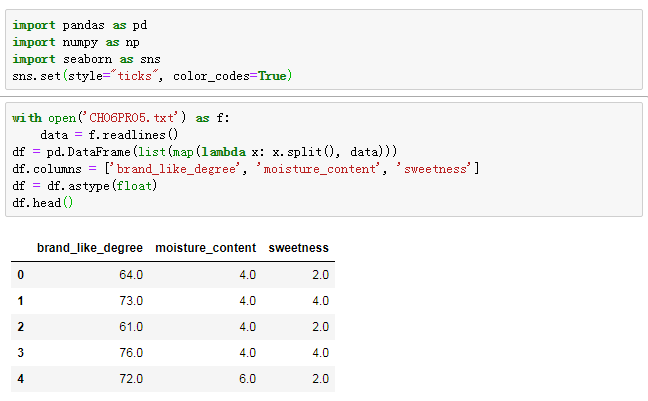
Where:

Restricted model:

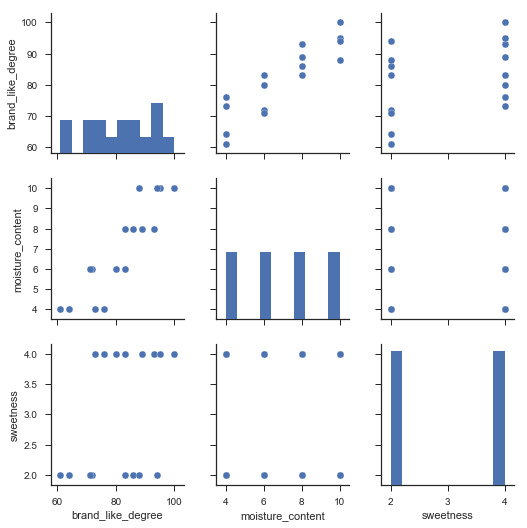
Since

Thus

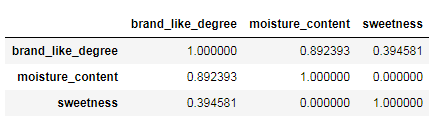
1. **Chapter 6 Problem a, b, c**
2. Basic data cleaning with Python



Scatter plot matrix

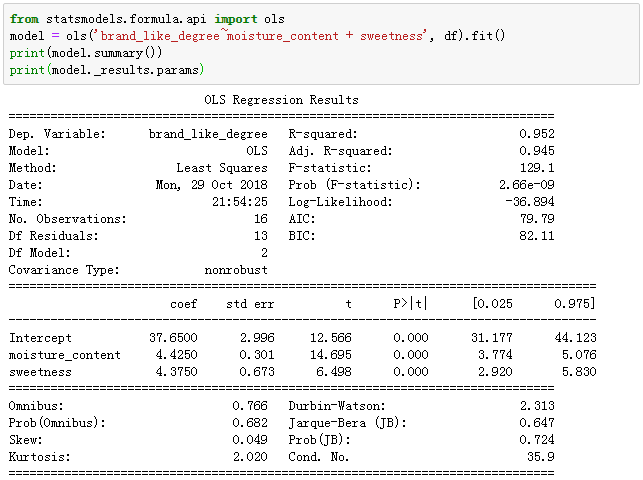


Correlation Matrix



Based on the results, we could see that the correlation between moisture content and degree of brand liking is very high. There is no correlation between moisture content and sweetness.

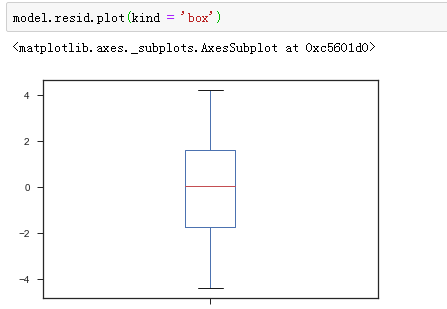
1. Based on 6.1



Based on the result, the regression function is:

is the coefficients of regression model, which means when improve 1, degree of brand liking improve 4.425, when sweetness improve 1, degree of brand liking improve 4.375.

1. Residual



We could see the residual almost follow the normal distribution and also don’t have large variance, which means the assumptions of model are correct.

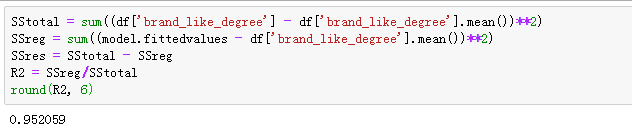
1. **Chapter 6 Problem 7**

(1)

C:\Users\ADMINI~1\AppData\Local\Temp\WeChat Files\56891bc5510b9e528a9aa5ac6fa4621.png

R square is 0.952, which means the factor sweetness and moisture content have strong relationship with degree of brand liking.

(2)

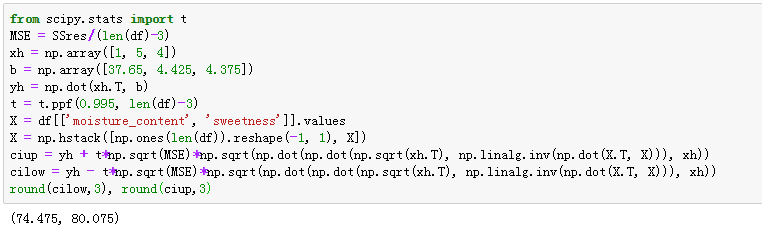


Based on the result, the multiple and single determination are the same.

1. **Chapter 6 Problem 8**

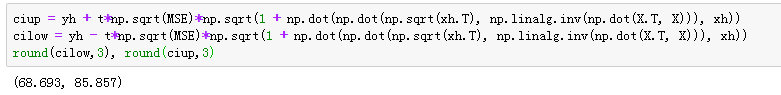
(a)

The confidence interval for is



The confidence interval is:

(b)



Confidence interval is [68.693, 85.857]

1. **Chapter 6 Problem 25**

Set , Use the model

to get the fitted line. Thus, we have