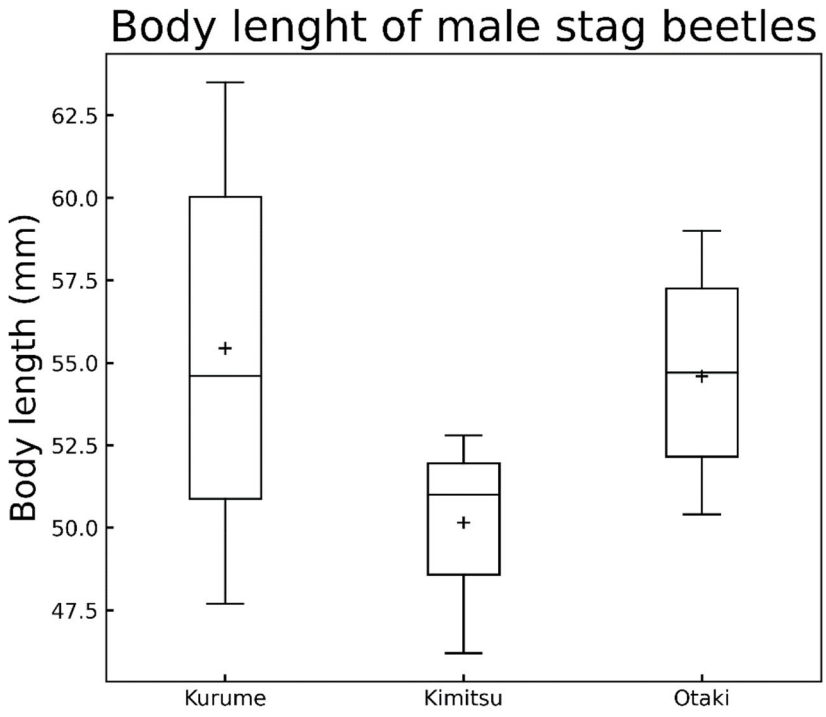
Python seminar Homework for Chap. 5.4

1. Please plot a box-and-whisker diagram with testing significant differences between the data. It is acceptable to use your own research data or any data you are interested in.
(1) Make a box-and-whisker diagram of the data of several groups (preferably three or more groups).
(2) After creating the box-and-whisker diagram, determine whether there is a significant difference in the means between the groups using ANOVA or other methods. If there is a significant difference between the groups, please do multiple comparisons if you are able to do.
When selecting a test method, it is necessary to check the distribution for normality.
For the case, you do not have any good data, an example data is listed in below.

Body length of stag beetles captured at 3 stations (Kurume, Kimitsu, Otaki).

```
kurume = np.array([52.3,60.4,47.7,48.7,58.9,54.2,63.5,55.0,63.3,50.4])
kimitsu = np.array([49.7,51.3,46.2,52.8,52.1,51.8,46.7,52.0,48.2,50.7])
otaki = np.array([57.9,50.4,53.2,51.8,55.3,59.0,57.7,50.6,55.9,54.1])
BLdata = pd.DataFrame({'kurume':kurume, 'kimitsu':kimitsu, 'otaki':otaki})
print(BLdata)
```

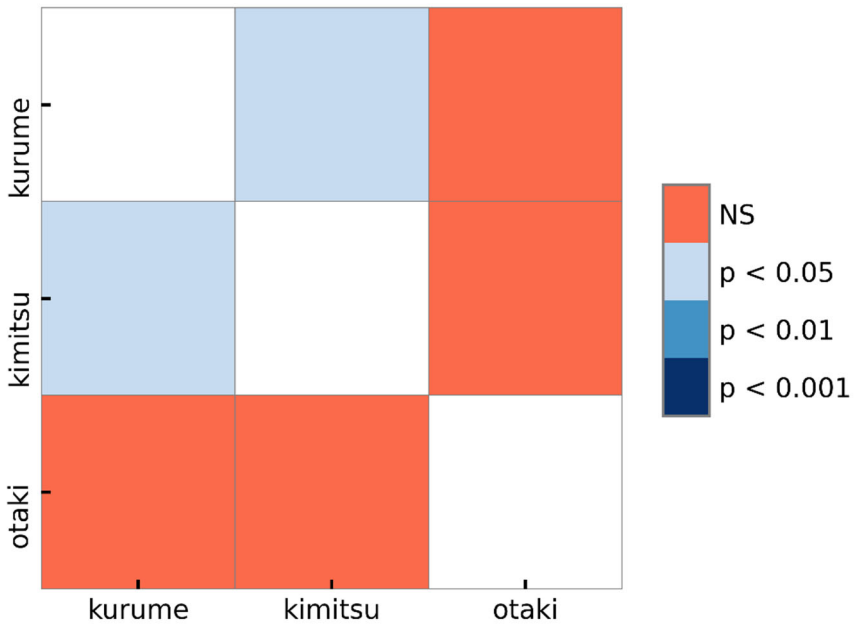


normal distribution test for kurume, p = 0.394291
The null hypothesis cannot be rejected: normal distribution (D'Agostino and Pearson's test)
normal distribution test for kimitsu, p = 0.423392
The null hypothesis cannot be rejected: normal distribution (D'Agostino and Pearson's test)
normal distribution test for otaki, p = 0.456277
The null hypothesis cannot be rejected: normal distribution (D'Agostino and Pearson's test)

Normal distribution test cannot reject the normal distribution. Therefore, parametric ANOVA was performed.
ANOVA, p = 0.0148994

ANOVA confirmed significant difference. Therefore, Tukey-Kramer test was performed.
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
kimitsu	kurume	5.29	0.0183	0.8071	9.7729	True
kimitsu	otaki	4.44	0.0526	-0.0429	8.9229	False
kurume	otaki	-0.85	0.8819	-5.3329	3.6329	False



We must pay attention whether the data distribution is a normal distribution or not? The test cannot reject the null hypothesis. However, data number is very limited and the actual distribution is questionable.
In addition,
I am tending to use non-parametric test for these data.