

In-Class Activity L8-1: Instructor and School Learning Scores

□ A large manufacturing company operates three regional training schools for mechanics, one in each of its operating districts. The schools have two instructors each, who teach classes of about 15 mechanics in three-week sessions. The company was concerned about the effect of school (factor A) and instructor (factor B) on the learning achieved. To investigate these effects, classes in each district were formed in the usual way and then randomly assigned to one of the two instructors in the school. This was done for two sessions, and at the end of each session a suitable summary measure of learning for the class was obtained.

Determine the ANOVA table from the data and draw conclusion.

Score	School	Instructor	Session
25	1	1	1
29	1	1	2
14	1	2	1
11	1	2	2

```
In [17]: import pandas as pd
data_trainingschool = pd.DataFrame({"Score": [25, 29, 14, 11, 11, 6, 22, 18, 17, 20, 5, 2],
                                     "School": [1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3],
                                     "Instructor": [1, 1, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2],
                                     "Session": [1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2]})
```

```
In [18]: data_trainingschool
```

Out[18]:

	Score	School	Instructor	Session
0	25	1	1	1
1	29	1	1	2
2	14	1	2	1
3	11	1	2	2
4	11	2	1	1
5	6	2	1	2
6	22	2	2	1
7	18	2	2	2
8	17	3	1	1
9	20	3	1	2
10	5	3	2	1
11	2	3	2	2

```
In [19]: from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm
formula = 'Score ~ C(School) + C(Instructor) + C(Instructor):C(School) '
lm = ols(formula, data=data_trainingschool).fit()
aov_table = anova_lm(lm)
aov_table
```

Out[19]:

	df	sum_sq	mean_sq	F	PR(>F)
C(School)	2.0	156.5	78.25	11.178571	0.009473
C(Instructor)	1.0	108.0	108.00	15.428571	0.007731
C(Instructor):C(School)	2.0	459.5	229.75	32.821429	0.000587
Residual	6.0	42.0	7.00	NaN	NaN

```
In [13]: SSb_a= 108.0 + 459.5
df_ssb_a = 3
MS_b_a = (SSb_a / df_ssb_a) / 7
```

```
In [16]: from scipy import stats
p_value = stats.f.sf(MS_b_a, 3, 6)
p_value
```

Out[16]: 0.0006970134862605147

จากตารางวิเคราะห์ความแปรปรวนพบว่า Instructor(School)ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

```
In [ ]:
```