**LP#43 반복수가 같은 일원배치**

In [81]:

**import** scipy.stats **as** stats

**import** pandas **as** pd

**import** urllib

**from** statsmodels.formula.api **import** ols

**from** statsmodels.stats.anova **import** anova\_lm

**import** matplotlib.pyplot **as** plt

**import** numpy **as** np

**%**matplotlib inline

**import** warnings

warnings.filterwarnings('ignore')

In [82]:

data **=** pd.read\_csv('One-way ANOVA.csv')

In [83]:

df **=** pd.DataFrame(data)

In [84]:

df

Out[84]:

|  | **인장강도** | **온도** |
| --- | --- | --- |
| **0** | 8.44000 | 1 |
| **1** | 8.36000 | 1 |
| **2** | 8.28000 | 1 |
| **3** | 8.59000 | 2 |
| **4** | 8.91000 | 2 |
| **5** | 8.60000 | 2 |
| **6** | 9.34000 | 3 |
| **7** | 9.41000 | 3 |
| **8** | 9.69000 | 3 |
| **9** | 8.92000 | 4 |
| **10** | 8.92000 | 4 |
| **11** | 8.74000 | 4 |

In [85]:

model **=** ols('인장강도 ~ C(온도)', df).fit()

​

print(anova\_lm(model))

df sum\_sq mean\_sq F PR(>F)

C(온도) 3.00000 1.97880 0.65960 31.18676 0.00009

Residual 8.00000 0.16920 0.02115 NaN NaN

**ANOVA TABLE**

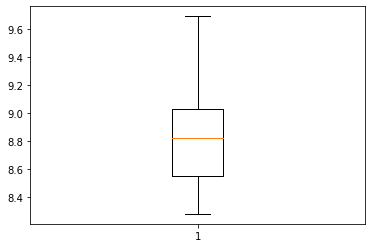
|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(반응온도) | 3.0 | 1.9788 | 0.65960 | 31.186761 | 0.000092 |
| Residual(오차) | 8.0 | 0.1692 | 0.02115 |  |  |

In [9

plot\_data **=** [df['인장강도']]

ax **=** plt.boxplot(plot\_data)

plt.show()

****

**LP#47 반복없는 이원배치**

In [86]:

data **=** pd.read\_csv('Two-way ANOVA.csv')

In [87]:

df **=** pd.DataFrame(data)

In [88]:

df

Out[88]:

|  | **수율** | **반응온도** | **압력** |
| --- | --- | --- | --- |
| **0** | 97.60000 | 1 | 1 |
| **1** | 97.30000 | 1 | 2 |
| **2** | 96.70000 | 1 | 3 |
| **3** | 98.60000 | 2 | 1 |
| **4** | 98.20000 | 2 | 2 |
| **5** | 96.90000 | 2 | 3 |
| **6** | 99.00000 | 3 | 1 |
| **7** | 98.00000 | 3 | 2 |
| **8** | 97.90000 | 3 | 3 |
| **9** | 98.00000 | 4 | 1 |
| **10** | 97.70000 | 4 | 2 |
| **11** | 96.50000 | 4 | 3 |

In [91]:

pd.options.display.float\_format **=** '{:.5f}'.format

formula **=** '수율 ~ C(반응온도)+C(압력)'

lm **=** ols(formula, df).fit()

print(anova\_lm(lm))

df sum\_sq mean\_sq F PR(>F)

C(반응온도) 3.00000 2.22000 0.74000 7.92857 0.01647

C(압력) 2.00000 3.44000 1.72000 18.42857 0.00274

Residual 6.00000 0.56000 0.09333 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(반응온도) | 3.00000 | 2.22000 | 0.74000 | 7.92857 | 0.01647 |
| B(압력) | 2.00000 | 3.44000 | 1.72000 | 18.42857 | 0.00274 |
| Residual(오차) | 6.00000 | 0.56000 | 0.09333 |  |  |

**LP#3-2-2-2-1난괴법**

In [120]:

data **=** pd.read\_csv('RBD.csv')

In [121]:

df **=** pd.DataFrame(data)

In [122]:

df

Out[122]:

|  | **결정화도변화율** | **온도** | **습도** |
| --- | --- | --- | --- |
| **0** | 13.10000 | 1 | 1 |
| **1** | 12.90000 | 1 | 2 |
| **2** | 13.40000 | 1 | 3 |
| **3** | 12.40000 | 2 | 1 |
| **4** | 12.70000 | 2 | 2 |
| **5** | 12.50000 | 2 | 3 |
| **6** | 12.30000 | 3 | 1 |
| **7** | 12.00000 | 3 | 2 |
| **8** | 12.20000 | 3 | 3 |

In [123]:

pd.options.display.float\_format **=** '{:.5f}'.format

formula **=** '결정화도변화율 ~ C(온도)+C(습도)'

lm **=** ols(formula, df).fit()

print(anova\_lm(lm))

df sum\_sq mean\_sq F PR(>F)

C(온도) 2.00000 1.42889 0.71444 16.07500 0.01224

C(습도) 2.00000 0.04222 0.02111 0.47500 0.65299

Residual 4.00000 0.17778 0.04444 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(온도) | 2.00000 | 1.42889 | 0.71444 | 16.07500 | 0.01224 |
| B(습도) | 2.00000 | 0.04222 | 0.02111 | 0.47500 | 0.65299 |
| Residual(오차) | 4.00000 | 0.17778 | 0.04444 |  |  |

**Pooling**

In [127]:

model **=** ols('결정화도변화율 ~ C(온도)', df).fit()

​

print(anova\_lm(model))

df sum\_sq mean\_sq F PR(>F)

C(온도) 2.00000 1.42889 0.71444 19.48485 0.00238

Residual 6.00000 0.22000 0.03667 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(반응온도) | 2.00000 | 1.42889 | 0.71444 | 19.48485 | 0.00238 |
| Residual(오차) | 6.00000 | 0.22000 | 0.03667 |  |  |

**LP#48 반복있는 이원배치**

In [92]:

data **=** pd.read\_csv('1two-way ANOVA.csv')

In [93]:

df **=** pd.DataFrame(data)

In [94]:

df

Out[94]:

|  | **압축강도** | **석고종류** | **첨가량** |
| --- | --- | --- | --- |
| **0** | 305 | 1 | 1 |
| **1** | 302 | 1 | 1 |
| **2** | 335 | 1 | 2 |
| **3** | 337 | 1 | 2 |
| **4** | 366 | 1 | 3 |
| **5** | 364 | 1 | 3 |
| **6** | 372 | 1 | 4 |
| **7** | 374 | 1 | 4 |
| **8** | 376 | 1 | 5 |
| **9** | 373 | 1 | 5 |
| **10** | 348 | 1 | 6 |
| **11** | 350 | 1 | 6 |
| **12** | 322 | 2 | 1 |
| **13** | 325 | 2 | 1 |
| **14** | 350 | 2 | 2 |
| **15** | 348 | 2 | 2 |
| **16** | 326 | 2 | 3 |
| **17** | 324 | 2 | 3 |
| **18** | 330 | 2 | 4 |
| **19** | 330 | 2 | 4 |
| **20** | 327 | 2 | 5 |
| **21** | 330 | 2 | 5 |
| **22** | 310 | 2 | 6 |
| **23** | 308 | 2 | 6 |
| **24** | 320 | 3 | 1 |
| **25** | 322 | 3 | 1 |
| **26** | 342 | 3 | 2 |
| **27** | 344 | 3 | 2 |
| **28** | 338 | 3 | 3 |
| **29** | 336 | 3 | 3 |
| **30** | 348 | 3 | 4 |
| **31** | 348 | 3 | 4 |
| **32** | 350 | 3 | 5 |
| **33** | 350 | 3 | 5 |
| **34** | 330 | 3 | 6 |
| **35** | 328 | 3 | 6 |

In [97]:

pd.options.display.float\_format **=** '{:.5f}'.format

formula **=** '압축강도 ~ C(석고종류) + C(첨가량) + C(석고종류):C(첨가량)'

lm **=** ols(formula, df).fit()

print(anova\_lm(lm))

df sum\_sq mean\_sq F PR(>F)

C(석고종류) 2.00000 3088.22222 1544.11111 694.85000 0.00000

C(첨가량) 5.00000 5548.88889 1109.77778 499.40000 0.00000

C(석고종류):C(첨가량) 10.00000 4825.77778 482.57778 217.16000 0.00000

Residual 18.00000 40.00000 2.22222 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(석고종류) | 2.00000 | 3088.22222 | 1544.11111 | 694.85000 | 0.00000 |
| B(첨가량) | 5.00000 | 5548.88889 | 1109.77778 | 499.40000 | 0.00000 |
| A(석고종류)XB(첨가량) | 10.00000 | 4825.77778 | 482.57778 | 217.16000 | 0.00000 |
| Residual(오차) | 18.00000 | 40.00000 | 2.22222 |  |  |

**LP#3-4 반복없는 삼원배치법**

In [109]:

data **=** pd.read\_csv('Three-way ANOVA.csv')

In [110]:

df **=** pd.DataFrame(data)

In [111]:

df

Out[111]:

|  | **합성률** | **반응압력** | **반응시간** | **반응온도** |
| --- | --- | --- | --- | --- |
| **0** | 74 | 1 | 1 | 1 |
| **1** | 86 | 1 | 1 | 2 |
| **2** | 76 | 1 | 1 | 3 |
| **3** | 72 | 1 | 2 | 1 |
| **4** | 91 | 1 | 2 | 2 |
| **5** | 87 | 1 | 2 | 3 |
| **6** | 48 | 1 | 3 | 1 |
| **7** | 65 | 1 | 3 | 2 |
| **8** | 56 | 1 | 3 | 3 |
| **9** | 61 | 2 | 1 | 1 |
| **10** | 78 | 2 | 1 | 2 |
| **11** | 71 | 2 | 1 | 3 |
| **12** | 62 | 2 | 2 | 1 |
| **13** | 81 | 2 | 2 | 2 |
| **14** | 77 | 2 | 2 | 3 |
| **15** | 55 | 2 | 3 | 1 |
| **16** | 72 | 2 | 3 | 2 |
| **17** | 63 | 2 | 3 | 3 |
| **18** | 50 | 3 | 1 | 1 |
| **19** | 70 | 3 | 1 | 2 |
| **20** | 60 | 3 | 1 | 3 |
| **21** | 49 | 3 | 2 | 1 |
| **22** | 68 | 3 | 2 | 2 |
| **23** | 64 | 3 | 2 | 3 |
| **24** | 52 | 3 | 3 | 1 |
| **25** | 69 | 3 | 3 | 2 |
| **26** | 60 | 3 | 3 | 3 |

In [114]:

pd.options.display.float\_format **=** '{:.5f}'.format

formula **=** '합성률 ~ C(반응압력) + C(반응시간) + C(반응온도) + C(반응압력):C(반응시간)+ C(반응압력):C(반응온도) + C(반응시간):C(반응온도)'

lm **=** ols(formula, df).fit()

print(anova\_lm(lm))

df sum\_sq mean\_sq F PR(>F)

C(반응압력) 2.00000 743.62963 371.81481 164.57377 0.00000

C(반응시간) 2.00000 753.40741 376.70370 166.73770 0.00000

C(반응온도) 2.00000 1380.96296 690.48148 305.62295 0.00000

C(반응압력):C(반응시간) 4.00000 651.92593 162.98148 72.13934 0.00000

C(반응압력):C(반응온도) 4.00000 9.03704 2.25926 1.00000 0.46091

C(반응시간):C(반응온도) 4.00000 56.59259 14.14815 6.26230 0.01384

Residual 8.00000 18.07407 2.25926 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(반응압력) | 2.00000 | 743.62963 | 371.81481 | 164.57377 | 0.00000 |
| B(반응시간) | 2.00000 | 753.40741 | 376.70370 | 166.73770 | 0.00000 |
| C(반응온도) | 2.00000 | 1380.96296 | 690.48148 | 305.62295 | 0.00000 |
| A(반응압력) X B(반응시간) | 4.00000 | 651.92593 | 162.98148 | 72.13934 | 0.00000 |
| A(반응압력) X C(반응온도) | 4.00000 | 9.03704 | 2.25926 | 1.00000 | 0.46091 |
| B(반응시간) X C(반응온도) | 4.00000 | 56.59259 | 14.14815 | 6.26230 | 0.01384 |
| Residual(오차) | 8.00000 | 18.07407 | 2.25926 |  |  |

**Pooling**

In [116]:

pd.options.display.float\_format **=** '{:.5f}'.format

formula **=** '합성률 ~ C(반응압력) + C(반응시간) + C(반응온도) + C(반응압력):C(반응시간) + C(반응시간):C(반응온도)'

lm **=** ols(formula, df).fit()

print(anova\_lm(lm))

df sum\_sq mean\_sq F PR(>F)

C(반응압력) 2.00000 743.62963 371.81481 164.57377 0.00000

C(반응시간) 2.00000 753.40741 376.70370 166.73770 0.00000

C(반응온도) 2.00000 1380.96296 690.48148 305.62295 0.00000

C(반응압력):C(반응시간) 4.00000 651.92593 162.98148 72.13934 0.00000

C(반응시간):C(반응온도) 4.00000 56.59259 14.14815 6.26230 0.00584

Residual 12.00000 27.11111 2.25926 NaN NaN

**ANOVA TABLE**

|  | **df** | **SS** | **MS** | **F비** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| A(반응압력) | 2.00000 | 743.62963 | 371.81481 | 164.57377 | 0.00000 |
| B(반응시간) | 2.00000 | 753.40741 | 376.70370 | 166.73770 | 0.00000 |
| C(반응온도) | 2.00000 | 1380.96296 | 690.48148 | 305.62295 | 0.00000 |
| A(반응압력) X B(반응시간) | 4.00000 | 651.92593 | 162.98148 | 72.13934 | 0.00000 |
| B(반응시간) X C(반응온도) | 4.00000 | 56.59259 | 14.14815 | 6.26230 | 0.00584 |
| Residual(오차) | 12.00000 | 27.11111 | 2.25926 |  |  |