

```
In [5]: import numpy as np
import statsmodels.api as sm
from statsmodels.formula.api import ols
```

```
In [12]: import pandas as pd
```

```
In [66]: Anova_data = pd.read_clipboard('clipboard')
```

```
In [67]: Anova_data.head(3)
```

Out[67]:

	WORK	TEACHER	Hours
0	1	1	6
1	1	1	8
2	1	1	4

```
In [68]: import statsmodels.api as sm
from statsmodels.formula.api import ols
```

Perform two-way ANOVA

```
In [69]: model = ols('Hours ~ C(TEACHER) + C(WORK) + C(TEACHER):C(WORK)', data=Anova_data).fit()  
sm.stats.anova_lm(model, typ=2)
```

Out[69]:

	sum_sq	df	F	PR(>F)
C(TEACHER)	23.765625	1.0	7.592346	7.750323e-03
C(WORK)	107.640625	1.0	34.387687	2.069678e-07
C(TEACHER):C(WORK)	13.140625	1.0	4.198003	4.485264e-02
Residual	187.812500	60.0	NaN	NaN

```
In [70]: print(model.summary())
```

```

                                OLS Regression Results
=====
Dep. Variable:                  Hours    R-squared:                  0.435
Model:                            OLS    Adj. R-squared:              0.407
Method:                 Least Squares    F-statistic:                  15.39
Date:                Sun, 14 Apr 2024    Prob (F-statistic):          1.54e-07
Time:                18:03:12    Log-Likelihood:              -125.26
No. Observations:                64    AIC:                          258.5
Df Residuals:                    60    BIC:                          267.2
Df Model:                        3
Covariance Type:                nonrobust
=====
                                coef    std err          t      P>|t|      [0.025     0.975]
-----
Intercept                    6.0625     0.442     13.706     0.000     5.178     6.947
C(TEACHER)[T.2]             -0.3125     0.626     -0.500     0.619    -1.564     0.939
C(WORK)[T.2]                 3.5000     0.626     5.595     0.000     2.249     4.751
C(TEACHER)[T.2]:C(WORK)[T.2] -1.8125     0.885     -2.049     0.045    -3.582    -0.043
=====
Omnibus:                    0.460    Durbin-Watson:              1.984
Prob(Omnibus):              0.794    Jarque-Bera (JB):           0.608
Skew:                      0.078    Prob(JB):                   0.738
Kurtosis:                  2.549    Cond. No.                   6.85
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Post-Hoc Test

```
In [71]: from statsmodels.stats.multicomp import pairwise_tukeyhsd
```

```
In [76]: import numpy as np
```

```
In [94]: tukey=pairwise_tukeyhsd(endog=Anova_data['WORK'], groups=Anova_data['TEACHER'],alpha=0.05)
```

```
In [95]: print(tukey)
```

Multiple Comparison of Means - Tukey HSD, FWER=0.05

```
=====
group1 group2 meandiff p-adj lower upper reject
-----
      1      2      0.0   1.0 -0.2539 0.2539 False
-----
```

Data Used is here

```
In [60]: print(Anova_data)
```

	WORK	TEACHER	Hours
0	1	1	6
1	1	1	8
2	1	1	4
3	1	1	6
4	1	1	4
..
59	2	2	9
60	2	2	7
61	2	2	4
62	2	2	8
63	2	2	10

[64 rows x 3 columns]

In [62]: Anova_data.info(all)

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 64 entries, 0 to 63
Data columns (total 3 columns):
#   Column      Non-Null Count  Dtype
---  -
0   WORK        64 non-null    int64
1   TEACHER     64 non-null    int64
2   Hours       64 non-null    int64
dtypes: int64(3)
memory usage: 1.6 KB
```

In []: