

```
In [ ]: !pip install statsmodels
!pip install scikit-learn
!pip install scikit-posthocs
!pip install openpyxl
!pip install scikit-posthocs
!pip install bioinfokit
```

Librerias

```
In [14]: import pandas as pd
import statsmodels.api as sm
from statsmodels.formula.api import ols
from statsmodels.stats.multicomp import pairwise_tukeyhsd
import scikit_posthocs as sp
from bioinfokit.analys import stat
```



1. TOROS6

```
In [12]: # Leer el archivo Excel
df = pd.read_excel("toros6.xlsx")
df['Toros'] = df['Toros'].astype(str)

# ANOVA
modelo = ols('Peso ~ C(Toros)', data=df).fit()
print(sm.stats.anova_lm(modelo, typ=2))

# Duncan
#print(sp.posthoc_duncan(df, val_col='Peso', group_col='Toros'))

# Tukey
tukey = pairwise_tukeyhsd(df['Peso'], df['Toros'])
print(tukey)
```

```

              sum_sq    df      F      PR(>F)
C(Toros)    22.883333    5.0  1.842953  0.119954
Residual   134.100000   54.0      NaN      NaN
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj   lower  upper  reject
-----
    1     2     -0.4 0.9927 -2.4822  1.6822  False
    1     3      0.8 0.8645 -1.2822  2.8822  False
    1     4     -0.2 0.9997 -2.2822  1.8822  False
    1     5      1.3 0.4468 -0.7822  3.3822  False
    1     6      0.8 0.8645 -1.2822  2.8822  False
    2     3      1.2  0.536 -0.8822  3.2822  False
    2     4      0.2 0.9997 -1.8822  2.2822  False
    2     5      1.7 0.1703 -0.3822  3.7822  False
    2     6      1.2  0.536 -0.8822  3.2822  False
    3     4     -1.0 0.7155 -3.0822  1.0822  False
    3     5      0.5   0.98 -1.5822  2.5822  False
    3     6      0.0   1.0 -2.0822  2.0822  False
    4     5      1.5 0.2886 -0.5822  3.5822  False
    4     6      1.0 0.7155 -1.0822  3.0822  False
    5     6     -0.5   0.98 -2.5822  1.5822  False
-----

```



2. ALFALFA

```

In [18]: df = pd.read_excel("alfalfa.xlsx")
df['Parcela'] = df['Parcela'].astype(str)

# ANOVA
modelo = ols('Dosis ~ C(Parcela)', data=df).fit()
print(sm.stats.anova_lm(modelo, typ=2))

# Duncan
#print(sp.posthoc_duncan(df, val_col='Dosis', group_col='Parcela'))

# Tukey
tukey = pairwise_tukeyhsd(df['Dosis'], df['Parcela'])
print(tukey)

```

```

              sum_sq    df      F      PR(>F)
C(Parcela)  2000.16    4.0 265.978723 4.947594e-17
Residual      37.60   20.0      NaN      NaN
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj   lower  upper  reject
-----
    1     2      4.2 0.0008  1.6051  6.7949   True
    1     3     15.0   0.0 12.4051 17.5949   True
    1     4     14.6   0.0 12.0051 17.1949   True
    1     5     25.4   0.0 22.8051 27.9949   True
    2     3     10.8   0.0  8.2051 13.3949   True
    2     4     10.4   0.0  7.8051 12.9949   True
    2     5     21.2   0.0 18.6051 23.7949   True
    3     4     -0.4  0.99 -2.9949  2.1949  False
    3     5     10.4   0.0  7.8051 12.9949   True
    4     5     10.8   0.0  8.2051 13.3949   True
-----

```

✓ 2. TOROS5

```
In [19]: df = pd.read_excel("toros5.xlsx")
df['Toros'] = df['Toros'].astype(str)

# ANOVA
modelo = ols('Peso ~ C(Toros)', data=df).fit()
print(sm.stats.anova_lm(modelo, typ=2))

# Duncan
#print(sp.posthoc_duncan(df, val_col='Peso', group_col='Toros'))

# Tukey
tukey = pairwise_tukeyhsd(df['Peso'], df['Toros'])
print(tukey)
```

| | sum_sq | df | F | PR(>F) |
|----------|--------|------|----------|----------|
| C(Toros) | 777.6 | 4.0 | 2.778113 | 0.038046 |
| Residual | 3148.9 | 45.0 | NaN | NaN |

Multiple Comparison of Means - Tukey HSD, FWER=0.05

```
=====
```

| group1 | group2 | meandiff | p-adj | lower | upper | reject |
|--------|--------|----------|--------|----------|---------|--------|
| 1 | 2 | -7.4 | 0.2931 | -18.0299 | 3.2299 | False |
| 1 | 3 | -7.5 | 0.2803 | -18.1299 | 3.1299 | False |
| 1 | 4 | -12.2 | 0.0172 | -22.8299 | -1.5701 | True |
| 1 | 5 | -5.4 | 0.6035 | -16.0299 | 5.2299 | False |
| 2 | 3 | -0.1 | 1.0 | -10.7299 | 10.5299 | False |
| 2 | 4 | -4.8 | 0.7027 | -15.4299 | 5.8299 | False |
| 2 | 5 | 2.0 | 0.9832 | -8.6299 | 12.6299 | False |
| 3 | 4 | -4.7 | 0.7187 | -15.3299 | 5.9299 | False |
| 3 | 5 | 2.1 | 0.9799 | -8.5299 | 12.7299 | False |
| 4 | 5 | 6.8 | 0.3764 | -3.8299 | 17.4299 | False |

```
-----
```

✓ 4. CEBADA

```
In [20]: df = pd.read_excel("cebada.xlsx")
df['Cebada'] = df['Cebada'].astype(str)

# ANOVA
modelo = ols('Cosecha ~ C(Cebada)', data=df).fit()
print(sm.stats.anova_lm(modelo, typ=2))

# Duncan
#print(sp.posthoc_duncan(df, val_col='Cosecha', group_col='Cebada'))

# Tukey
tukey = pairwise_tukeyhsd(df['Cosecha'], df['Cebada'])
print(tukey)
```

```
      sum_sq    df          F    PR(>F)
C(Cebada) 24775.0   3.0  12.234568  0.000206
Residual  10800.0  16.0          NaN      NaN
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj    lower    upper  reject
-----
      1      2       4.0 0.9947  -43.0114  51.0114  False
      1      3     -84.0 0.0005 -131.0114 -36.9886   True
      1      4     -22.0 0.553  -69.0114  25.0114  False
      2      3     -88.0 0.0003 -135.0114 -40.9886   True
      2      4     -26.0 0.4155  -73.0114  21.0114  False
      3      4      62.0 0.0081   14.9886 109.0114   True
-----
```