**<https://github.com/marsja/jupyter/blob/master/Python_ANOVA/One_Way_Python_ANOVA.ipynb>**

**SE CALCULA POR 3 FORMAS**

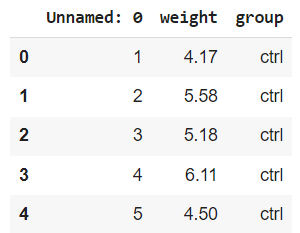
**LUEGO SE COMPARA EN PARES <- IMPORTANTE**

import pandas as pd

datafile = 'https://vincentarelbundock.github.io/Rdatasets/csv/datasets/PlantGrowth.csv'

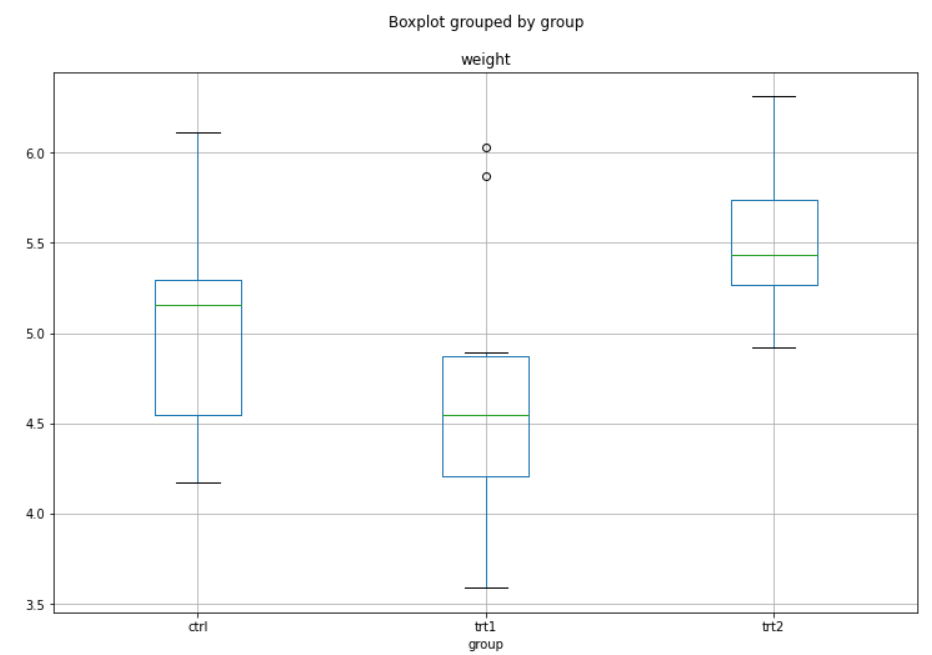
df = pd.read\_csv(datafile)

df.head()



#Explore, create a boxplot

df.boxplot('weight', by='group', figsize=(12, 8))



**Group: ctrl, trt1, trt2**

**Weight: 3.5 - 6**

ctrl = df['weight'][df.group == 'ctrl']

grps = pd.unique(df.group.values)

d\_data = {grp:df['weight'][df.group == grp] for grp in grps}

### Anova usig Scipy

**from** scipy **import** stats

# Son 3 grupos de datos

F, p **=** stats**.**f\_oneway(d\_data['ctrl'],

d\_data['trt1'],

d\_data['trt2'])

print(F, p)

4.84608786 0.01590995

**Como 0.01 < 0.05 hay significancia que las medias sin similares**

**OTRA OPCION DE CALCULO**

### Degrees of Fredoom

k = len(pd.unique(df.group))   # 3 number of conditions (3 groups)?

N = len(df.values)   # 10 conditions times participants

n = df.groupby('group').size()[0] # 30 Participants in each condition

DFbetween **=** k **-** 1

DFwithin **=** N **-** k

DFtotal **=** N **–** 1

### Sum of Squares Between, WIthin, and Total

# Between

SSbetween = (sum(df.groupby('group').sum()['weight']\*\*2)/n) \

    - (df['weight'].sum()\*\*2)/N

# Within

sum\_y\_squared = sum([value\*\*2 for value in df['weight'].values])

SSwithin = sum\_y\_squared - sum(df.groupby('group').sum()['weight']\*\*2)/n

# Total

SStotal = sum\_y\_squared - (df['weight'].sum()\*\*2)/N

### Means Square Errors and F-value

MSbetween = SSbetween/DFbetween

MSwithin = SSwithin/DFwithin

F = MSbetween/MSwithin

p = stats.f.sf(F, DFbetween, DFwithin)

print(F, p)

4.8460878 0.015909958

**Da nuevamente los mismos valores**

**Como 0.01 < 0.05 hay significancia que las medias sin similares**

### Calculate Effect sizes: Eta Squared and Omega Squared:

eta\_sqrd = SSbetween/SStotal

om\_sqrd = (SSbetween - (DFbetween \* MSwithin))/(SStotal + MSwithin)

results = {'sum\_sq':[SSbetween, SSwithin],

           'df':[DFbetween, DFwithin],

           'F':[F, ''],

            'PR(>F)':[p, ''],

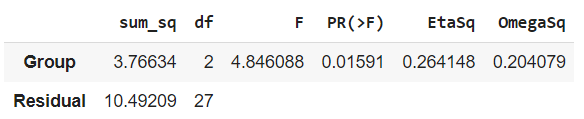
          'EtaSq':[eta\_sqrd, ''],

          'OmegaSq':[om\_sqrd, '']}

aov\_table1 = pd.DataFrame(results,

                          index=['Group', 'Residual'])

aov\_table1

****

**Otra forma de calcularlo**

import statsmodels.api as sm

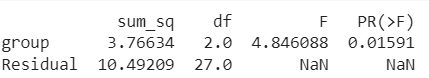
from statsmodels.formula.api import ols

mod = ols('weight ~ group',

                data=df).fit()

aov\_table = sm.stats.anova\_lm(mod, typ=2)

print(aov\_table)

****

**Tercera forma de calcular**

esq\_sm = aov\_table['sum\_sq'][0]/(aov\_table['sum\_sq'][0] + aov\_table['sum\_sq'][1])

aov\_table['EtaSq'] = [esq\_sm, 'NaN']

print(aov\_table)

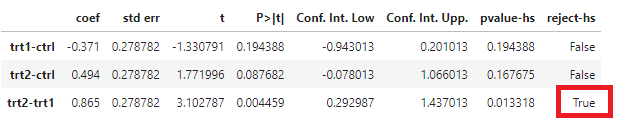
**A picture containing text

Description automatically generated**

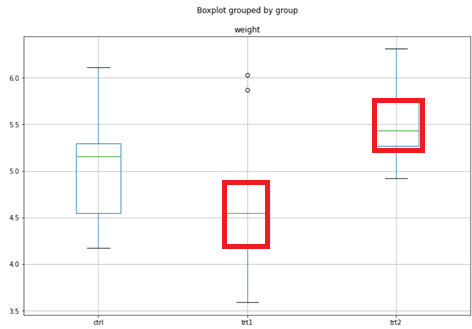
## Pairwise Comparisons

pair\_t = mod.t\_test\_pairwise('group')

pair\_t.result\_frame

****

**En los 3 grupos no se rechaza, pero si se compara trt2 – trt1 si se ve diferencia**

****