import seaborn as sns

import pandas as pd

import numpy as np

2 Tue

3 Mon

4 Sun

5 Wed

6 Mon

Suma de reg x dia / group by dia

================================

df['day'].value\_counts()

Sat 85

Sun 75

Mon 13

Tue 64

Promedios x dia

===============

df['day'].value.count() / df['day'].value.count().sum() \* 100

Sat 14.5

Sun 37.5

Mon 13.8

Tue 6.9

Promedios registros x sexo

===========================

df.groupby('sex').mean()

Group by sex, promedio edad

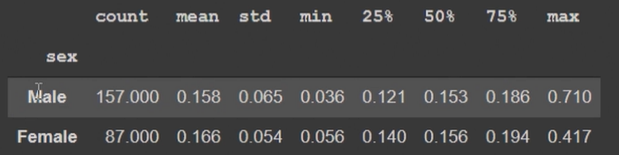
===========================

df.groupby('sex')[['edad']].mean()

Describir con percentiles la agrupacion

========================================

df.groupby('sex')[['edad']].describe()

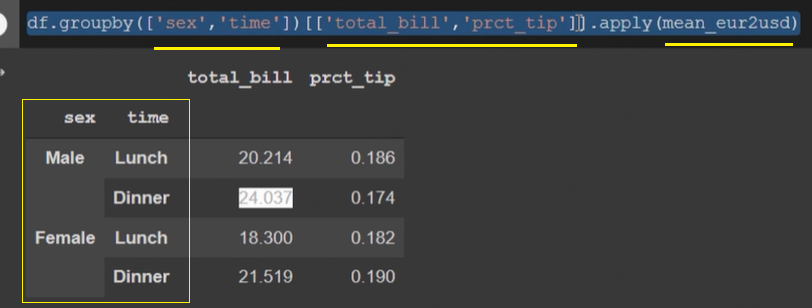


Nuevo campo

===========

df['price\_prom'] = df['pr']/df['tot']

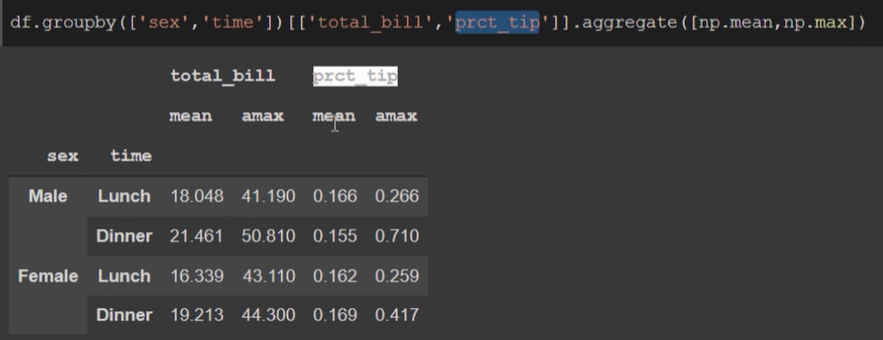




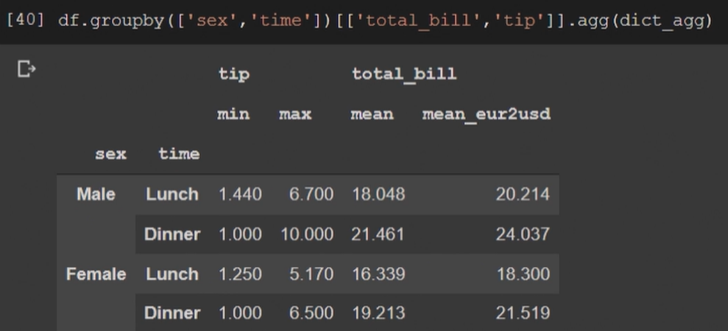
Tambien se puede con



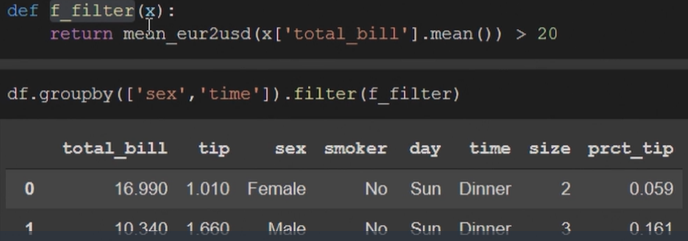
Mean y Max





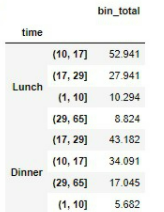


Con filter

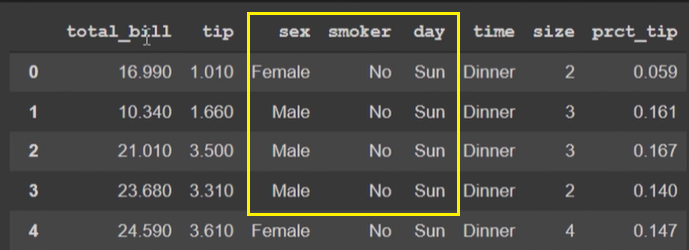


Las agrupaciones que hace el profesor usando la columna auxiliar ‘ones’ para tener proporciones también se puede hacer con value\_counts(normalize=True).

df.groupby(['time'])['bin\_total'].value\_counts(normalize=True).to\_frame()\*100



Data



Crear variable para si fuma 1, 0

df['fuma'] = np.where(df['smoker']!= 'Smoke', 1, 0)

df.groupby( [‘sex’] )