

Data Frames. Ejercicio

A partir del archivo csv de viviendas en venta ('realestate.csv'), deseamos localizar los que tienen más de 1000 pies cuadrados, tres dormitorios y dos o más baños, situados en una ciudad distinta de "Sacramento" y tener un precio entre 100.000\$ y 150.000\$. Nos interesa averiguar la siguiente información:

1. El listado completo
2. De ellos, obtén un listado por ciudades y número de baños
3. El número total de pisos en nuestra tabla
4. La media de los precios de los pisos de nuestra tabla

In [1]:

```
# Ejercicio sobre Dataframes

import pandas

realestate = pandas.read_csv('realestate.csv')

cond = (realestate["sq__ft"] > 1000) \
      & (realestate["beds"] == 3) \
      & (realestate["baths"] >= 1) \
      & (realestate["city"] != "SACRAMENTO") \
      & (100000 <= realestate["price"]) \
      & (realestate["price"] <= 150000)

# 1. Listado completo

nuestra_seleccion = realestate[cond]

nuestra_seleccion
```

Out[1]:

	street	city	zip	state	beds	baths	sq__ft	type	sale_date	price	latitud
19	113 LEEWILL AVE	RIO LINDA	95673	CA	3	2	1356	Residential	Wed May 21 00:00:00 EDT 2008	121630	38.68999
20	6118 STONEHAND AVE	CITRUS HEIGHTS	95621	CA	3	2	1118	Residential	Wed May 21 00:00:00 EDT 2008	122000	38.70785
25	3828 BLACKFOOT WAY	ANTELOPE	95843	CA	3	2	1088	Residential	Wed May 21 00:00:00 EDT 2008	126640	38.70974
24	6768	NORTH	95660	CA	3	2	1152	Residential	Wed May 21	124555	38.60116

In [2]:



```
# Ejercicio sobre Dataframes

import pandas

# 2. Listado por ciudades y número de baños

print(nuestra_seleccion.groupby(['city', 'baths']).size())
```

city	baths	
ANTELOPE	2	2
CARMICHAEL	2	1
CITRUS HEIGHTS	2	1
ELVERTA	2	3
GALT	1	1
	2	2
NORTH HIGHLANDS	1	4
	2	5
RANCHO CORDOVA	1	1
	2	1
RIO LINDA	2	1

dtype: int64

In [3]:



```
# 3. El número total de pisos en nuestra tabla

print(nuestra_seleccion.count()) # OJO: cuenta los totales de todas las columnas...

print()

print(nuestra_seleccion["city"].count()) # OJO: cuenta los totales de una columna (La prime
```

street	22
city	22
zip	22
state	22
beds	22
baths	22
sq__ft	22
type	22
sale_date	22
price	22
latitude	22
longitude	22

dtype: int64

22

In [4]:



```
# 4. La media de los precios de los pisos de nuestra tabla

print(nuestra_seleccion["price"])

print()

print(nuestra_seleccion["price"].mean())
```

```
19      121630
20      122000
25      126640
31      134555
33      138750
36      147308
174     106716
176     111000
178     120108
183     126000
189     140000
357     112500
365     119250
372     122000
380     130000
383     139500
384     140000
625     116000
631     124413
880     105000
884     110000
902     142000
Name: price, dtype: int64

125244.09090909091
```