

DATAFRAMES III

In [1]:

```
import pandas as pd
```

In [3]:

```
df = pd.read_csv("Iris.csv")
df
```

Out[3]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

Visualización de frecuencias usando CROSSTAB

In [4]:

```
pd.crosstab(df.SepalLengthCm, df.Species)
```

Out[4]:

	Species	Iris-setosa	Iris-versicolor	Iris-virginica
SepalLengthCm	4.3	1	0	0
4.4	3	0	0	0
4.5	1	0	0	0
4.6	4	0	0	0
4.7	2	0	0	0
4.8	5	0	0	0
4.9	4	1	1	1
5.0	8	2	0	0
5.1	8	1	0	0
5.2	3	1	0	0
5.3	1	0	0	0
5.4	5	1	0	0
5.5	2	5	0	0
5.6	0	5	1	1
5.7	2	5	1	1
5.8	1	3	3	3
5.9	0	2	1	1
6.0	0	4	2	2
6.1	0	4	2	2
6.2	0	2	2	2
6.3	0	3	6	6
6.4	0	2	5	5
6.5	0	1	4	4
6.6	0	2	0	0
6.7	0	3	5	5
6.8	0	1	2	2
6.9	0	1	3	3
7.0	0	1	0	0
7.1	0	0	1	1
7.2	0	0	3	3
7.3	0	0	1	1
7.4	0	0	1	1
7.6	0	0	1	1
7.7	0	0	4	4
7.9	0	0	1	1

In [5]:

```
pd.crosstab(df.SepalLengthCm, df.Species).plot(kind="bar")
```

Out[5]:<AxesSubplot:xlabel='SepalLengthCm'>

In [6]:

```
features = ["SepalLengthCm", "SepalWidthCm"]
for feature in features:
    pd.crosstab(df[feature], df["Species"]).plot(kind="bar")
```

Uso del filtrado usando ".LOC"

In [7]:

```
df
```

Out[7]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [8]:

```
# Mostrar datos desde la fila 4 hasta la fila 20 incluida y todas las columnas
df.loc[4:20, :]
```

Out[8]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa

In [9]:

```
df.loc[5]
```

Out[9]:

Id	6
SepalLengthCm	5.4
SepalWidthCm	3.9
PetalLengthCm	1.7
PetalWidthCm	0.4
Species	Iris-setosa
Name: 5, dtype: object	

In [10]:

```
# Mostrar por index concreto
df.loc[df.index[5], :]
```

Out[10]:

Id	6
SepalLengthCm	5.4
SepalWidthCm	3.9
PetalLengthCm	1.7
PetalWidthCm	0.4
Species	Iris-setosa
Name: 5, dtype: object	

In [12]:

```
# Mostrar columnas concretas
df.loc[4:20, "Species"]
```

Out[12]:

4	Iris-setosa
5	Iris-setosa
6	Iris-setosa
7	Iris-setosa
8	Iris-setosa
9	Iris-setosa
10	Iris-setosa
11	Iris-setosa
12	Iris-setosa
13	Iris-setosa
14	Iris-setosa
15	Iris-setosa
16	Iris-setosa
17	Iris-setosa
18	Iris-setosa
19	Iris-setosa
20	Iris-setosa
Name: Species, dtype: object	

In [13]:

```
df.loc[:, "Species"]
```

Out[13]:

0	Iris-setosa
1	Iris-setosa
2	Iris-setosa
3	Iris-setosa
4	Iris-setosa
...	...
145	Iris-virginica
146	Iris-virginica
147	Iris-virginica
148	Iris-virginica
149	Iris-virginica
Name: Species, Length: 150, dtype: object	

In [14]:

```
df.loc[:, ["PetalLengthCm", "Species"]]
```

Out[14]:

	PetalLengthCm	Species
0	1.4	Iris-setosa
1	1.4	Iris-setosa
2	1.3	Iris-setosa
3	1.5	Iris-setosa
4	1.4	Iris-setosa
...
145	5.2	Iris-virginica
146	5.0	Iris-virginica
147	5.2	Iris-virginica
148	5.4	Iris-virginica
149	5.1	Iris-virginica

150 rows × 2 columns

In [17]:

```
# Filtrar una especie: ejemplo mostrar todos los valores de "Iris-versicolor"
df.loc[df["Species"] == "Iris-versicolor"]
```

Out[17]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
50	51	7.0	3.2	4.7	1.4	Iris-versicolor
51	52	6.4	3.2	4.5	1.5	Iris-versicolor
52	53	6.9	3.1	4.9	1.5	Iris-versicolor
53	54	5.5	2.3	4.0	1.3	Iris-versicolor
54	55	6.5	2.8	4.6	1.5	Iris-versicolor
55	56	5.7	2.8	4.5	1.3	Iris-versicolor
56	57	6.3	3.3	4.7	1.6	Iris-versicolor
57	58	4.9	2.4	3.3	1.0	Iris-versicolor
58	59	6.6	2.9	4.6	1.3	Iris-versicolor
59	60	5.2	2.7	3.9	1.4	Iris-versicolor
60	61	5.0	2.0	3.5	1.0	Iris-versicolor
61	62	5.9	3.0	4.2	1.5	Iris-versicolor
62	63	6.0	2.2	4.0	1.0	Iris-versicolor
63	64	6.1	2.9	4.7	1.4	Iris-versicolor
64	65	5.6	2.9	3.6	1.3	Iris-versicolor
65	66	6.7	3.1	4.4	1.4	Iris-versicolor
66	67	5.6	3.0	4.5	1.5	Iris-versicolor
67	68	5.8	2.7	4.1	1.0	Iris-versicolor
68	69	6.2	2.2	4.5	1.5	Iris-versicolor
69	70	5.6	2.5	3.9	1.1	Iris-versicolor
70	71	5.9	3.2	4.8	1.8	Iris-versicolor
71	72	6.1	2.8	4.0	1.3	Iris-versicolor
72	73	6.3	2.5	4.9	1.5	Iris-versicolor
73	74	6.1	2.8	4.7	1.2	Iris-versicolor
74	75	6.4	2.9	4.3	1.3	Iris-versicolor
75	76	6.6	3.0	4.4	1.4	Iris-versicolor
76	77	6.8	2.8	4.8	1.4	Iris-versicolor
77	78	6.7	3.0	5.0	1.7	Iris-versicolor
78	79	6.0	2.9	4.5	1.5	Iris-versicolor
79	80	5.7	2.6	3.5	1.0	Iris-versicolor
80	81	5.5	2.4	3.8	1.1	Iris-versicolor
81	82	5.5	2.4	3.7	1.0	Iris-versicolor
82	83	5.8	2.7	3.9	1.2	Iris-versicolor
83	84	6.0	2.7	5.1	1.6	Iris-versicolor
84	85	5.4	3.0	4.5	1.5	Iris-versicolor
85	86	6.0	3.4	4.5	1.6	Iris-versicolor
86	87	6.7	3.1	4.7	1.5	Iris-versicolor
87	88	6.3	2.3	4.4	1.3	Iris-versicolor
88	89	5.6	3.0	4.1	1.3	Iris-versicolor
89	90	5.5	2.5	4.0	1.3	Iris-versicolor
90	91	5.5	2.6	4.4	1.2	Iris-versicolor
91	92	6.1	3.0	4.6	1.4	Iris-versicolor
92	93	5.8	2.6	4.0	1.2	Iris-versicolor
93	94	5.0	2.3	3.3	1.0	Iris-versicolor
94	95	5.6	2.7	4.2	1.3	Iris-versicolor
95	96	5.7	3.0	4.2	1.2	Iris-versicolor
96	97	5.7	2.9	4.2	1.3	Iris-versicolor
97	98	6.2	2.9	4.3	1.3	Iris-versicolor
98	99	5.1	2.5	3.0	1.1	Iris-versicolor
99	100	5.7	2.8	4.1	1.3	Iris-versicolor

In [18]:

```
df.loc[lambda df: df["Species"] == "Iris-setosa"]
```

Out[18]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
30	31	4.8	3.1	1.6	0.2	Iris-setosa
31	32	5.4	3.4	1.5	0.4	Iris-setosa
32	33	5.2	4.1	1.5	0.1	Iris-setosa
33	34	5.5	4.2	1.4	0.2	Iris-setosa
34	35	4.9	3.1	1.5	0.1	Iris-setosa
35	36	5.0	3.2	1.2	0.2	Iris-setosa
36	37	5.5	3.5	1.3	0.2	Iris-setosa
37	38	4.9	3.1	1.5	0.1	Iris-setosa
38	39	4.4	3.0	1.3	0.2	Iris-setosa
39	40	5.1	3.4	1.5	0.2	Iris-setosa
40	41	5.0	3.5	1.3	0.3	Iris-setosa
41	42	4.5	2.3	1.3	0.3	Iris-setosa
42	43	4.4	3.2	1.3	0.2	Iris-setosa
43	44	5.0	3.5	1.6	0.6	Iris-setosa
44	45	5.1	3.8	1.9	0.4	Iris-setosa
45	46	4.8	3.0	1.4	0.3	Iris-setosa
46	47	5.1	3.8	1.6	0.2	Iris-setosa
47	48	4.6	3.2	1.4	0.2	Iris-setosa
48	49	5.3	3.7	1.5	0.2	Iris-setosa
49	50	5.0	3.3	1.4	0.2	Iris-setosa

Creado por: