Leer DATOS CON PANDAS

Import libraries

In [1]:
```
#importar packages
import pandas as pdsmn
```

Leer datos Cargar datos

In [2]:
```
# Cargar el archivo CSV desde la carpeta cartwheel
df = pd.read_csv(r'C:\VisualProyectsPC\TC1002S\llddl\23\TC1002S\NotebooksStudents\A01642529\Iris
```

In [3]:
```
# Mostrar las primeras filas del DataFrame para verificar que los datos se hayan cargado correct
df.head()
```

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

In [4]:
```
#Get the number of rows and columns
df.shape
```

Out[4]:
```
(150, 6)
```

In [5]:
```
#get number of rows
num_rows = df.shape[0]

#Print the number of rows
print('The number of rows is :' + str(num_rows))
```

```
The number of rows is :150
```

In [6]:
```
#Get number of cos
Ncols = df.shape[1]

#Print the number of columns
print('The number of columns is :' + str(Ncols))
```

```
The number of columns is :6
```

EXPLORAR DATAFRAME

In [7]:
```
df.columns
```

Out[7]:
```
Index(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm',
       'Species'],
      dtype='object')
```

In [8]:
```
#Get the data types of each column
df.dtypes
```

Out[8]:
```
Id                int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species           object
dtype: object
```

In [9]:
```
#Obtener un nuevo dataframe con algunas columnas
df2 = df[['SepalLengthCm', 'SepalWidthCm', 'Species']]

df2.head()
```

Out[9]:

	SepalLengthCm	SepalWidthCm	Species
0	5.1	3.5	Iris-setosa
1	4.9	3.0	Iris-setosa
2	4.7	3.2	Iris-setosa
3	4.6	3.1	Iris-setosa
4	5.0	3.6	Iris-setosa

In [10]:
```
df2.head(10)
```

Out[10]:

	SepalLengthCm	SepalWidthCm	Species
0	5.1	3.5	Iris-setosa
1	4.9	3.0	Iris-setosa
2	4.7	3.2	Iris-setosa
3	4.6	3.1	Iris-setosa
4	5.0	3.6	Iris-setosa
5	5.4	3.9	Iris-setosa
6	4.6	3.4	Iris-setosa
7	5.0	3.4	Iris-setosa
8	4.4	2.9	Iris-setosa
9	4.9	3.1	Iris-setosa

In [11]:
```
df2.tail(3)
```

Out[11]:

	SepalLengthCm	SepalWidthCm	Species
147	6.5	3.0	Iris-virginica
148	6.2	3.4	Iris-virginica
149	5.9	3.0	Iris-virginica

Estadisticas de las variables

Calcular y visualizar estadisticas descriptivas

In [12]:
```
#Calcular y visualizar estadisticas descriptivas
df2.describe()
```

Out[12]:

	SepalLengthCm	SepalWidthCm
count	150.000000	150.000000
mean	5.843333	3.054000
std	0.828066	0.433594
min	4.300000	2.000000
25%	5.100000	2.800000
50%	5.800000	3.000000
75%	6.400000	3.300000
max	7.900000	4.400000

In [13]:
```
df.describe()
```

Out[13]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

In [14]:
```
#IDENTIFICAR ROWS/cols woth mission information

df.isnull()
```

Out[14]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
145	False	False	False	False	False	False
146	False	False	False	False	False	False
147	False	False	False	False	False	False
148	False	False	False	False	False	False
149	False	False	False	False	False	False

150 rows × 6 columns

In [15]:
```
#IMPRIMIR ESTADISTICA DESCRIPTIVA
df.describe()
```

Out[15]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000