

✓ Iris

✓ Introduction:

This exercise may seem a little bit strange, but keep doing it.

Step 1. Import the necessary libraries

```
import pandas as pd
```

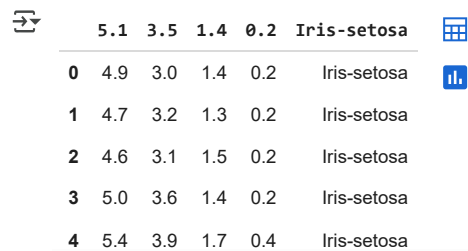
✓ Step 2. Import the dataset from this [address](https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data).

```
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
```

✓ Step 3. Assign it to a variable called iris

```
iris = pd.read_csv(url)
```

```
iris.head()
```



	5.1	3.5	1.4	0.2	Iris-setosa
0	4.9	3.0	1.4	0.2	Iris-setosa
1	4.7	3.2	1.3	0.2	Iris-setosa
2	4.6	3.1	1.5	0.2	Iris-setosa
3	5.0	3.6	1.4	0.2	Iris-setosa
4	5.4	3.9	1.7	0.4	Iris-setosa

Next steps:

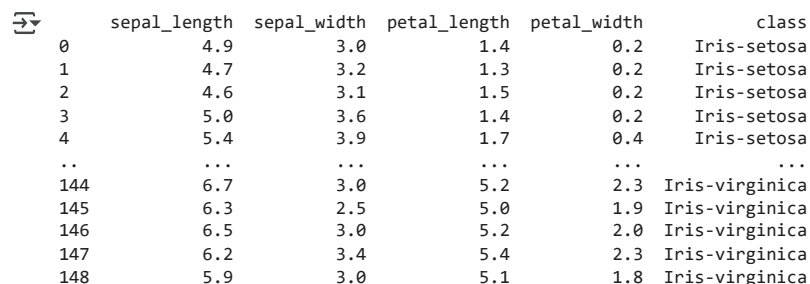
[Generate code with iris](#)
[View recommended plots](#)
[New interactive sheet](#)

✓ Step 4. Create columns for the dataset

```
# 1. sepal_length (in cm)
# 2. sepal_width (in cm)
# 3. petal_length (in cm)
# 4. petal_width (in cm)
# 5. class
```

```
iris.columns = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'class']
```

```
print(iris)
```



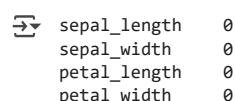
	sepal_length	sepal_width	petal_length	petal_width	class
0	4.9	3.0	1.4	0.2	Iris-setosa
1	4.7	3.2	1.3	0.2	Iris-setosa
2	4.6	3.1	1.5	0.2	Iris-setosa
3	5.0	3.6	1.4	0.2	Iris-setosa
4	5.4	3.9	1.7	0.4	Iris-setosa
..
144	6.7	3.0	5.2	2.3	Iris-virginica
145	6.3	2.5	5.0	1.9	Iris-virginica
146	6.5	3.0	5.2	2.0	Iris-virginica
147	6.2	3.4	5.4	2.3	Iris-virginica
148	5.9	3.0	5.1	1.8	Iris-virginica

[149 rows x 5 columns]

✓ Step 5. Is there any missing value in the dataframe?

```
missing_values = iris.isnull().sum()
```

```
print(missing_values)
```



```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
```

```
class      0
dtype: int64
```

✓ Step 6. Lets set the values of the rows 10 to 29 of the column 'petal_length' to NaN

```
iris.loc[10:29, 'petal_length'] = pd.NA
```

```
iris.head(29)
```

	sepal_length	sepal_width	petal_length	petal_width	class
0	4.9	3.0	1.4	0.2	Iris-setosa
1	4.7	3.2	1.3	0.2	Iris-setosa
2	4.6	3.1	1.5	0.2	Iris-setosa
3	5.0	3.6	1.4	0.2	Iris-setosa
4	5.4	3.9	1.7	0.4	Iris-setosa
5	4.6	3.4	1.4	0.3	Iris-setosa
6	5.0	3.4	1.5	0.2	Iris-setosa
7	4.4	2.9	1.4	0.2	Iris-setosa
8	4.9	3.1	1.5	0.1	Iris-setosa
9	5.4	3.7	1.5	0.2	Iris-setosa
10	4.8	3.4	NaN	0.2	Iris-setosa
11	4.8	3.0	NaN	0.1	Iris-setosa
12	4.3	3.0	NaN	0.1	Iris-setosa
13	5.8	4.0	NaN	0.2	Iris-setosa
14	5.7	4.4	NaN	0.4	Iris-setosa
15	5.4	3.9	NaN	0.4	Iris-setosa
16	5.1	3.5	NaN	0.3	Iris-setosa
17	5.7	3.8	NaN	0.3	Iris-setosa
18	5.1	3.8	NaN	0.3	Iris-setosa
19	5.4	3.4	NaN	0.2	Iris-setosa
20	5.1	3.7	NaN	0.4	Iris-setosa
21	4.6	3.6	NaN	0.2	Iris-setosa
22	5.1	3.3	NaN	0.5	Iris-setosa
23	4.8	3.4	NaN	0.2	Iris-setosa
24	5.0	3.0	NaN	0.2	Iris-setosa
25	5.0	3.4	NaN	0.4	Iris-setosa
26	5.2	3.5	NaN	0.2	Iris-setosa
27	5.2	3.4	NaN	0.2	Iris-setosa
28	4.7	3.2	NaN	0.2	Iris-setosa


Next steps:

[Generate code with iris](#)
[View recommended plots](#)
[New interactive sheet](#)


✓ Step 7. Good, now lets substitute the NaN values to 1.0

```
iris.fillna(1.0, inplace=True)
```

```
iris.head(29)
```



	sepal_length	sepal_width	petal_length	petal_width	class
0	4.9	3.0	1.4	0.2	Iris-setosa
1	4.7	3.2	1.3	0.2	Iris-setosa
2	4.6	3.1	1.5	0.2	Iris-setosa
3	5.0	3.6	1.4	0.2	Iris-setosa
4	5.4	3.9	1.7	0.4	Iris-setosa
5	4.6	3.4	1.4	0.3	Iris-setosa
6	5.0	3.4	1.5	0.2	Iris-setosa
7	4.4	2.9	1.4	0.2	Iris-setosa
8	4.9	3.1	1.5	0.1	Iris-setosa
9	5.4	3.7	1.5	0.2	Iris-setosa
10	4.8	3.4	1.0	0.2	Iris-setosa
11	4.8	3.0	1.0	0.1	Iris-setosa
12	4.3	3.0	1.0	0.1	Iris-setosa
13	5.8	4.0	1.0	0.2	Iris-setosa
14	5.7	4.4	1.0	0.4	Iris-setosa
15	5.4	3.9	1.0	0.4	Iris-setosa
16	5.1	3.5	1.0	0.3	Iris-setosa
17	5.7	3.8	1.0	0.3	Iris-setosa
18	5.1	3.8	1.0	0.3	Iris-setosa
19	5.4	3.4	1.0	0.2	Iris-setosa
20	5.1	3.7	1.0	0.4	Iris-setosa
21	4.6	3.6	1.0	0.2	Iris-setosa
22	5.1	3.3	1.0	0.5	Iris-setosa
23	4.8	3.4	1.0	0.2	Iris-setosa
24	5.0	3.0	1.0	0.2	Iris-setosa
25	5.0	3.4	1.0	0.4	Iris-setosa
26	5.2	3.5	1.0	0.2	Iris-setosa
27	5.2	3.4	1.0	0.2	Iris-setosa
28	4.7	3.2	1.0	0.2	Iris-setosa



Next steps:

[Generate code with iris](#)[View recommended plots](#)[New interactive sheet](#)

Step 8. Now let's delete the column class

```
iris = iris.drop('class', axis=1)
```

```
iris.head()
```



	sepal_length	sepal_width	petal_length	petal_width
0	4.9	3.0	1.4	0.2
1	4.7	3.2	1.3	0.2
2	4.6	3.1	1.5	0.2
3	5.0	3.6	1.4	0.2
4	5.4	3.9	1.7	0.4



Next steps:

[Generate code with iris](#)[View recommended plots](#)[New interactive sheet](#)

Step 9. Set the first 3 rows as NaN

```
iris.loc[:2] = pd.NA
```

```
print(iris)
```

```
↗
   sepal_length  sepal_width  petal_length  petal_width
0             NaN           NaN           NaN           NaN
1             NaN           NaN           NaN           NaN
2             NaN           NaN           NaN           NaN
3             5.0           3.6           1.4           0.2
4             5.4           3.9           1.7           0.4
..           ...           ...           ...           ...
144           6.7           3.0           5.2           2.3
145           6.3           2.5           5.0           1.9
146           6.5           3.0           5.2           2.0
147           6.2           3.4           5.4           2.3
148           5.9           3.0           5.1           1.8
```

[149 rows x 4 columns]

✓ Step 10. Delete the rows that have NaN

```
iris = iris.dropna()
```

```
print(iris)
```

```
↗
   sepal_length  sepal_width  petal_length  petal_width
3             5.0           3.6           1.4           0.2
4             5.4           3.9           1.7           0.4
5             4.6           3.4           1.4           0.3
6             5.0           3.4           1.5           0.2
7             4.4           2.9           1.4           0.2
..           ...           ...           ...           ...
144           6.7           3.0           5.2           2.3
145           6.3           2.5           5.0           1.9
146           6.5           3.0           5.2           2.0
147           6.2           3.4           5.4           2.3
148           5.9           3.0           5.1           1.8
```

[146 rows x 4 columns]

✓ Step 11. Reset the index so it begins with 0 again

```
iris = iris.reset_index(drop=True)
```

```
print(iris)
```

```
↗
   index  sepal_length  sepal_width  petal_length  petal_width
0       3             5.0           3.6           1.4           0.2
1       4             5.4           3.9           1.7           0.4
2       5             4.6           3.4           1.4           0.3
3       6             5.0           3.4           1.5           0.2
4       7             4.4           2.9           1.4           0.2
..     ...           ...           ...           ...
141    144           6.7           3.0           5.2           2.3
142    145           6.3           2.5           5.0           1.9
143    146           6.5           3.0           5.2           2.0
144    147           6.2           3.4           5.4           2.3
145    148           5.9           3.0           5.1           1.8
```

[146 rows x 5 columns]

```
iris = iris.drop(columns=['index'])
```

```
print(iris)
```

```
↗
   sepal_length  sepal_width  petal_length  petal_width
0             5.0           3.6           1.4           0.2
1             5.4           3.9           1.7           0.4
2             4.6           3.4           1.4           0.3
3             5.0           3.4           1.5           0.2
4             4.4           2.9           1.4           0.2
..           ...           ...           ...           ...
141           6.7           3.0           5.2           2.3
142           6.3           2.5           5.0           1.9
143           6.5           3.0           5.2           2.0
144           6.2           3.4           5.4           2.3
145           5.9           3.0           5.1           1.8
```

[146 rows x 4 columns]

✓ BONUS: Create your own question and answer it.

Delete line 139,140

```
iris = iris.drop(index=139)
```

```
iris = iris.drop(index=140)
```

```
print(iris[130:145])
```

```
↕
```

	sepal_length	sepal_width	petal_length	petal_width
130	6.1	2.6	5.6	1.4
131	7.7	3.0	6.1	2.3
132	6.3	3.4	5.6	2.4
133	6.4	3.1	5.5	1.8
134	6.0	3.0	4.8	1.8
135	6.9	3.1	5.4	2.1
136	6.7	3.1	5.6	2.4
137	6.9	3.1	5.1	2.3
138	5.8	2.7	5.1	1.9
141	6.7	3.0	5.2	2.3
142	6.3	2.5	5.0	1.9
143	6.5	3.0	5.2	2.0
144	6.2	3.4	5.4	2.3
145	5.9	3.0	5.1	1.8