Introduction

The first step in most data analytics projects is reading the data file. In this exercise, you'll create Series and DataFrame objects, both by hand and by reading data files.

Run the code cell below to load libraries you will need (including code to check your answers).

```
In [2]:  import sys
         from pathlib import Path
         learntools_dir = Path().absolute().parents[1]
         sys.path.append(str(learntools_dir))
         from learntools.core import binder; binder.bind(globals())
         from learntools.pandas.creating_reading_and_writing import *

         import pandas as pd

         print("Setup complete.")
```

Setup complete.

Exercises

1.

In the cell below, create a DataFrame `fruits` that looks like this:

	Apples	Bananas
0	30	21

```
In [3]:  # Your code goes here. Create a dataframe matching the above diagram and ass
         fruits = pd.DataFrame({'Apples': [30], 'Bananas': [21]})

         # Check your answer
         q1.check()
         fruits
```

Correct

Out[3]:

	Apples	Bananas
0	30	21

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

```
In [4]:  # q1.hint()
         # q1.solution()
```

Solution:

fruits **=** pd**.**DataFrame([[30, 21]], columns**=**['Apples', 'Bananas'])

2.

Create a dataframe `fruit_sales` that matches the diagram below:

	Apples	Bananas
2017 Sales	35	21
2018 Sales	41	34

```
In [5]:  # Your code goes here. Create a dataframe matching the above diagram and ass
         fruit_sales = pd.DataFrame({'Apples': [35,41], 'Bananas': [21,34]}, index=['

         # Check your answer
         q2.check()
         fruit_sales
```

Correct

```
Out[5]:              Apples   Bananas

         2017 Sales      35        21

         2018 Sales      41        34
```

```
In [7]:  # q2.hint()

         # q2.solution()
```

Solution:

fruit_sales **=** pd**.**DataFrame([[35, 21], [41, 34]], columns**=**['Apples', 'Bananas'],
          index**=**['2017 Sales', '2018 Sales'])

3.

Create a variable `ingredients` with a Series that looks like:

```
        Flour       4 cups
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

```
        Eggs        2 large
        Spam            1 can
        Name: Dinner, dtype: object
```

In [22]:
```
ingredients = pd.Series(['4 cups', '1 cup', '2 large', '1 can'],
                        index=['Flour', 'Milk', 'Eggs', 'Spam'], name='Dinne

# Check your answer
# q3.check()
ingredients
```

Out[22]:
```
Flour       4 cups
Milk         1 cup
Eggs       2 large
Spam         1 can
Name: Dinner, dtype: object
```

In [10]:
```
# q3.hint()

q3.solution()
```

Solution:

quantities **=** ['4 cups', '1 cup', '2 large', '1 can']

items **=** ['Flour', 'Milk', 'Eggs', 'Spam']

recipe **=** pd**.**Series(quantities, index**=**items, name**=**'Dinner')

4.

Read the following csv dataset of wine reviews into a DataFrame called `reviews`:

	country	description	designation	points	price	province	region_1	region_2	variety	winery
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	Napa	Cabernet Sauvignon	Heitz
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	NaN	Tinta de Toro	Bodega Carmen Rodríguez
...	...	...	...	...	...	...	...	...	...	...
150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	NaN	Champagne Blend	Gosset
150929	Italy	More Pinot Grigios should taste like this. A r...	NaN	90	15.0	Northeastern Italy	Alto Adige	NaN	Pinot Grigio	Alois Lageder

The filepath to the csv file is `../datasets/winemag-data_first150k.csv`. The first few lines look like:

```
,country,description,designation,points,price,province,region_1,reg
0,US,"This tremendous 100% varietal wine[...]",Martha's
Vineyard,96,235.0,California,Napa Valley,Napa,Cabernet
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

```
1,Spain,"Ripe aromas of fig, blackberry and[...]",Carodorum
```

```
        Selección Especial Reserva,96,110.0,Northern
        Spain,Toro,,Tinta de Toro,Bodega Carmen Rodríguez
```

In [17]:
```python
reviews = pd.read_csv("../datasets/winemag-data_first150k.csv", index_col=0)

# Check your answer
q4.check()
reviews
```

Correct

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

Out[17]:

	country	description	designation	points	price	province	region_1
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro
2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley
3	US	This spent 20 months in 30% new French oak, an...	Reserve	96	65.0	Oregon	Willamette Valley
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol
...	...	...	...	...	...	...	...
150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne
150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino
150928	France	A perfect salmon shade, with peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

	country	description	designation	points	price	province	region_1	
150929	Italy	More Pinot Grigios should taste like this. A r...	NaN	90	15.0	Northeastern Italy	Alto Adige	

150930 rows × 10 columns

In [14]:
```python
# q4.hint()
q4.solution()
```

Solution:

reviews **=** pd**.**read_csv('../../pandas/datasets/winemag-data_first150k.csv', index_col**=**0)

5.

Run the cell below to create and display a DataFrame called `animals`:

In [23]:
```python
animals = pd.DataFrame({'Cows': [12, 20], 'Goats': [22, 19]}, index=['Year 1
animals
```

Out[23]:

	Cows	Goats
Year 1	12	22
Year 2	20	19

In the cell below, write code to save this DataFrame to disk as a csv file with the name `cows_and_goats.csv`.

In [24]:
```python
# Your code goes here
animals.to_csv("cows_and_goats.csv")
# Check your answer
q5.check()
```

Correct

In [19]:
```python
# q5.hint()
q5.solution()
```

Solution:

animals**.**to_csv("cows_and_goats.csv")

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

Keep going

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js