This notebook is an exercise in the Pandas (https://www.kaggle.com/learn/pandas) course. You can reference the tutorial at this link (https://www.kaggle.com/residentmario/creating-reading-and-writing).

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 [ ]:
```
import pandas as pd
pd.set_option('max_rows', 5)
from learntools.core import binder; binder.bind(globals())
from learntools.pandas.creating_reading_and_writing import *
print("Setup complete.")
```

Exercises

1.

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

	Apples	Bananas
0	30	21

In [19]:
```
# Your code goes here. Create a dataframe matching the above diagram and
assign it to the variable fruits.
fruits = pd.DataFrame({'Apples': [30], 'Bananas': [21]})

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

Correct

Out[19]:

	Apples	Bananas
0	30	21

2.

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

	Apples	Bananas
2017 Sales	35	21
2018 Sales	41	34

In [18]:
```
# Your code goes here. Create a dataframe matching the above diagram and
assign it to the variable fruit_sales.
fruit_sales = pd.DataFrame({'Apples': [35, 41],
                            'Bananas': [21, 34]},
                           index=['2017 Sales', '2018 Sales'])

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

Correct

Out[18]:

	Apples	Bananas
2017 Sales	35	21
2018 Sales	41	34

3.

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

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

In [20]:

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

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

Correct

Out[20]:

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

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 `../input/wine-reviews/winemag-data_first150k.csv`. The first few lines look like:

```
,country,description,designation,points,price,province,region_1,region_2,varie
ty,winery
0,US,"This tremendous 100% varietal wine[...]",Martha's Vineyard,96,235.0,Cali
fornia,Napa Valley,Napa,Cabernet Sauvignon,Heitz
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 [21]:
```
reviews = pd.read_csv("../input/wine-reviews/winemag-data_first150k.csv", index_col=0)

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

Correct

Out[21]:

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

150930 rows × 10 columns

5.

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

In [22]:
```
animals = pd.DataFrame({'Cows': [12, 20], 'Goats': [22, 19]}, index=['Year 1', 'Year 2'])
animals
```

Out[22]:

	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 [23]:
```
# Your code goes here
animals.to_csv ('cows_and_goats.csv', header=True)
# Check your answer
q5.check()
```

Correct

Keep going

Move on to learn about **indexing, selecting and assigning (https://www.kaggle.com/residentmario/indexing-selecting-assigning)**.

In [ ]:
```

```

Have questions or comments? Visit the Learn Discussion forum (https://www.kaggle.com/learn-forum/161299) to chat with other Learners.