Introduction

In this set of exercises we will work with the [Wine Reviews dataset].

Run the following cell to load your data and some utility functions (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))

import pandas as pd

reviews = pd.read_csv("../../pandas/datasets/winemag-data-130k-v2.csv", inde

from learntools.core import binder; binder.bind(globals())
from learntools.pandas.indexing_selecting_and_assigning import *
print("Setup complete.")
```

```
Setup complete.
```

Look at an overview of your data by running the following line.

In [3]:
```
reviews.head()
```

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

Out[3]:

	country	description	designation	points	price	province	region_1	region_2	t
0	Italy	Aromas include tropical fruit, broom, brimston...	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN	
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	NaN	NaN	
2	US	Tart and snappy, the flavors of lime flesh and...	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN	
4	US	Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	

Exercises

1.

Select the `description` column from `reviews` and assign the result to the variable `desc`.

In [4]:
```
# Your code here
desc = reviews.description

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

Correct

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

```
Out[4]: 0        Aromas include tropical fruit, broom, brimston...
        1        This is ripe and fruity, a wine that is smooth...
        2        Tart and snappy, the flavors of lime flesh and...
        3        Pineapple rind, lemon pith and orange blossom ...
        4        Much like the regular bottling from 2012, this...
                                     ...
        65494    Made from young vines from the Vaulorent porti...
        65495    This is a big, fat, almost sweet-tasting Caber...
        65496    Much improved over the unripe 2005, Fritz's 20...
        65497    This wine wears its 15.8% alcohol better than ...
        65498    A unique take on Manzanilla Sherry, which is o...
        Name: description, Length: 65499, dtype: object
```

```
In [5]: type(desc)
```

```
Out[5]: pandas.core.series.Series
```

Follow-up question: what type of object is `desc`? If you're not sure, you can check by calling Python's `type` function: `type(desc)`.

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

        # q1.solution()
```

Solution:

desc = reviews.description
or

desc = reviews["description"]
`desc` is a pandas `Series` object, with an index matching the `reviews` DataFrame. In general, when we select a single column from a DataFrame, we'll get a Series.

2.

Select the first value from the description column of `reviews`, assigning it to variable `first_description`.

```
In [6]: first_description = reviews.description[0]

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

Correct:

first_description = reviews.description.iloc[0]
Note that while this is the preferred way to obtain the entry in the DataFrame, many other options will return a valid result, such as `reviews.description.loc[0]`, more!

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

Out[6]:
```
"Aromas include tropical fruit, broom, brimstone and dried herb. The palate
isn't overly expressive, offering unripened apple, citrus and dried sage al
ongside brisk acidity."
```

In [6]:
```
# q2.hint()

# q2.solution()
```

Solution:

first_description = reviews.description.iloc[0]

Note that while this is the preferred way to obtain the entry in the DataFrame, many other options will return a valid result, such as `reviews.description.loc[0]`, `reviews.description[0]`, and more!

3.

Select the first row of data (the first record) from `reviews`, assigning it to the variable `first_row`.

In [13]:
```
first_row = reviews.iloc[0]

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

Correct

Out[13]:
```
country                                                      Italy
description             Aromas include tropical fruit, broom, brimston...
designation                                           Vulkà Bianco
points                                                          87
price                                                          NaN
province                                          Sicily & Sardinia
region_1                                                      Etna
region_2                                                       NaN
taster_name                                          Kerin O’Keefe
taster_twitter_handle                                  @kerinokeefe
title                                Nicosia 2013 Vulkà Bianco  (Etna)
variety                                               White Blend
winery                                                    Nicosia
Name: 0, dtype: object
```

In [15]:
```
first_row = reviews.loc[0]

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

Correct

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

Out[15]:
```
country                                                    Italy
description             Aromas include tropical fruit, broom, brimston...
designation                                          Vulkà Bianco
points                                                         87
price                                                         NaN
province                                          Sicily & Sardinia
region_1                                                    Etna
region_2                                                     NaN
taster_name                                        Kerin O'Keefe
taster_twitter_handle                                 @kerinokeefe
title                                Nicosia 2013 Vulkà Bianco  (Etna)
variety                                             White Blend
winery                                                  Nicosia
Name: 0, dtype: object
```

In [12]:
```
# q3.hint()

q3.solution()
```

Solution:

first_row = reviews.iloc[0]

4.

Select the first 10 values from the `description` column in `reviews`, assigning the result to variable `first_descriptions`.

Hint: format your output as a pandas Series.

In [8]:
```
first_descriptions = reviews.description.iloc[:10]

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

Correct:

first_descriptions = reviews.description.iloc[:10]
Note that many other options will return a valid result, such as `desc.head(10)` and `reviews.loc[:9, "description"]`.

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

Out[8]:
```
0    Aromas include tropical fruit, broom, brimston...
1    This is ripe and fruity, a wine that is smooth...
2    Tart and snappy, the flavors of lime flesh and...
3    Pineapple rind, lemon pith and orange blossom ...
4    Much like the regular bottling from 2012, this...
5    Blackberry and raspberry aromas show a typical...
6    Here's a bright, informal red that opens with ...
7    This dry and restrained wine offers spice in p...
8    Savory dried thyme notes accent sunnier flavor...
9    This has great depth of flavor with its fresh ...
Name: description, dtype: object
```

In [9]:
```
first_descriptions = reviews.description.loc[:9]

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

Correct:

first_descriptions = reviews.description.iloc[:10]
Note that many other options will return a valid result, such as `desc.head(10)` and `reviews.loc[:9, "description"]`.

Out[9]:
```
0    Aromas include tropical fruit, broom, brimston...
1    This is ripe and fruity, a wine that is smooth...
2    Tart and snappy, the flavors of lime flesh and...
3    Pineapple rind, lemon pith and orange blossom ...
4    Much like the regular bottling from 2012, this...
5    Blackberry and raspberry aromas show a typical...
6    Here's a bright, informal red that opens with ...
7    This dry and restrained wine offers spice in p...
8    Savory dried thyme notes accent sunnier flavor...
9    This has great depth of flavor with its fresh ...
Name: description, dtype: object
```

In [3]:
```
# q4.hint()

q4.solution()
```

Solution:

first_descriptions = reviews.description.iloc[:10]
Note that many other options will return a valid result, such as `desc.head(10)` and `reviews.loc[:9, "description"]`.

5.

Select the records with index labels `1`, `2`, `3`, `5`, and `8`, assigning the result to the variable `sample_reviews`.

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.jsowing DataFrame:

	country	description	designation	points	price	province	region_1	region_2	taster_name	taster_twitter_handle
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	NaN	NaN	Roger Voss	@vossroger
2	US	Tart and snappy, the flavors of lime flesh and...	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Gregutt	@paulgwine
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN	Alexander Peartree	NaN
5	Spain	Blackberry and raspberry aromas show a typical...	Ars In Vitro	87	15.0	Northern Spain	Navarra	NaN	Michael Schachner	@wineschach
8	Germany	Savory dried thyme notes accent sunnier flavor...	Shine	87	12.0	Rheinhessen	NaN	NaN	Anna Lee C. Iijima	NaN

In [17]:
```
sample_reviews = reviews.loc[[1, 2, 3, 5, 8]]

# Check your answer
q5.check()
sample_reviews
```

Correct

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

Out[17]:

	country	description	designation	points	price	province	region_1	region_2
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	NaN	NaN
2	US	Tart and snappy, the flavors of lime flesh and...	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN
5	Spain	Blackberry and raspberry aromas show a typical...	Ars In Vitro	87	15.0	Northern Spain	Navarra	NaN
8	Germany	Savory dried thyme notes accent sunnier flavor...	Shine	87	12.0	Rheinhessen	NaN	NaN

In [18]:
```
sample_reviews = reviews.iloc[[1, 2, 3, 5, 8]]

# Check your answer
q5.check()
sample_reviews
```

Correct

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

Out[18]:

	country	description	designation	points	price	province	region_1	region_2
1	Portugal	This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	NaN	NaN
2	US	Tart and snappy, the flavors of lime flesh and...	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley
3	US	Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN
5	Spain	Blackberry and raspberry aromas show a typical...	Ars In Vitro	87	15.0	Northern Spain	Navarra	NaN
8	Germany	Savory dried thyme notes accent sunnier flavor...	Shine	87	12.0	Rheinhessen	NaN	NaN

In [16]:
```
# q5.hint()

q5.solution()
```

Solution:

```
indices = [1, 2, 3, 5, 8]
sample_reviews = reviews.loc[indices]
```

6.

Create a variable `df` containing the `country`, `province`, `region_1`, and `region_2` columns of the records with the index labels `0`, `1`, `10`, and `100`. In other words, generate the following DataFrame:

	country	province	region_1	region_2
0	Italy	Sicily & Sardinia	Etna	NaN
1	Portugal	Douro	NaN	NaN
10	US	California	Napa Valley	Napa
			Finger Lakes	Finger Lakes

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

In [12]:
```
df = reviews.loc[[0,1,10,100],['country', 'province', 'region_1', 'region_2'

# Check your answer
q6.check()
df
```

Correct

Out[12]:

	country	province	region_1	region_2
0	Italy	Sicily & Sardinia	Etna	NaN
1	Portugal	Douro	NaN	NaN
10	US	California	Napa Valley	Napa
100	US	New York	Finger Lakes	Finger Lakes

In [19]:
```
# q6.hint()

q6.solution()
```

Solution:

```
cols = ['country', 'province', 'region_1', 'region_2']
indices = [0, 1, 10, 100]
df = reviews.loc[indices, cols]
```

7.

Create a variable `df` containing the `country` and `variety` columns of the first 100 records.

Hint: you may use `loc` or `iloc`. When working on the answer this question and the several of the ones that follow, keep the following "gotcha" described in the tutorial:

> `iloc` uses the Python stdlib indexing scheme, where the first element of the range is included and the last one excluded. `loc`, meanwhile, indexes inclusively.

> This is particularly confusing when the DataFrame index is a simple numerical list, e.g. `0,...,1000`. In this case `df.iloc[0:1000]` will return 1000 entries, while `df.loc[0:1000]` return 1001 of them! To get 1000 elements using `loc`, you will need to go one lower and ask for `df.iloc[0:999]`.

In [13]:
```
df = reviews.loc[0:99,['country','variety']]
# Check your answer
```

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

```
q7.check()
df
```

Correct:

```
cols = ['country', 'variety']
df = reviews.loc[:99, cols]
```

or

```
cols_idx = [0, 11]
df = reviews.iloc[:100, cols_idx]
```

Out[13]:

	country	variety
0	Italy	White Blend
1	Portugal	Portuguese Red
2	US	Pinot Gris
3	US	Riesling
4	US	Pinot Noir
...	...	...
95	France	Gamay
96	France	Gamay
97	US	Riesling
98	Italy	Sangiovese
99	US	Bordeaux-style Red Blend

100 rows × 2 columns

In [14]:

```
df = reviews.iloc[0:100,[0,11]]

# Check your answer
q7.check()
df
```

Correct:

```
cols = ['country', 'variety']
df = reviews.loc[:99, cols]
```

or

```
cols_idx = [0, 11]
df = reviews.iloc[:100, cols_idx]
```

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

Out[14]:

	country	variety
0	Italy	White Blend
1	Portugal	Portuguese Red
2	US	Pinot Gris
3	US	Riesling
4	US	Pinot Noir
...	...	...
95	France	Gamay
96	France	Gamay
97	US	Riesling
98	Italy	Sangiovese
99	US	Bordeaux-style Red Blend

100 rows × 2 columns

In [16]:
```
# q7.hint()

q7.solution()
```

Solution:

```
cols = ['country', 'variety']
df = reviews.loc[:99, cols]
```

or

```
cols_idx = [0, 11]
df = reviews.iloc[:100, cols_idx]
```

8.

Create a DataFrame `italian_wines` containing reviews of wines made in `Italy`.
Hint: `reviews.country` equals what?

In [18]:
```
italian_wines = reviews[reviews.country == 'Italy']

# Check your answer
q8.check()
italian_wines
```

Correct

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

Out[18]:

	country	description	designation	points	price	province	region_1	reg
0	Italy	Aromas include tropical fruit, broom, brimston...	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	
6	Italy	Here's a bright, informal red that opens with ...	Belsito	87	16.0	Sicily & Sardinia	Vittoria	
13	Italy	This is dominated by oak and oak-driven aromas...	Rosso	87	NaN	Sicily & Sardinia	Etna	
22	Italy	Delicate aromas recall white flower and citrus...	Ficiligno	87	19.0	Sicily & Sardinia	Sicilia	
24	Italy	Aromas of prune, blackcurrant, toast and oak c...	Aynat	87	35.0	Sicily & Sardinia	Sicilia	
...	...	...	...	...	...	...	...	
65466	Italy	Earthy truffle, porcini mushroom, herb and gam...	NaN	88	70.0	Tuscany	Brunello di Montalcino	
65474	Italy	Made of 70% Syrah, 15% Sangiovese and 15% Merl...	Taneto	88	25.0	Tuscany	Toscana	
65476	Italy	Rose, violet, sour berry and tilled earth arom...	Prugnolo	88	25.0	Tuscany	Rosso di Montepulciano	
65477	Italy	Made of 65% Merlot, 25% Cabernet Sauvignon, 5%...	Ruit Hora	88	30.0	Tuscany	Bolgheri	

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

	country	description	designation	points	price	province	region_1	reg
65478	Italy	Aromas suggesting French oak, coconut and spic...	NaN	88	36.0	Tuscany	Vino Nobile di Montepulciano	

10005 rows × 13 columns

In [19]:
```
# q8.hint()

q8.solution()
```

Solution:

```
italian_wines = reviews[reviews.country == 'Italy']
```

9.

Create a DataFrame `top_oceania_wines` containing all reviews with at least 95 points (out of 100) for wines from Australia or New Zealand.

In [21]:
```
top_oceania_wines = reviews.loc[(reviews.country.isin(['Australia', 'New Zea
& (reviews.points >= 95)  ]

# Check your answer
q9.check()
top_oceania_wines
```

Correct

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

Out[21]:

	country	description	designation	points	price	province	region_1	regio
345	Australia	This wine contains some material over 100 year...	Rare	100	350.0	Victoria	Rutherglen	N
346	Australia	This deep brown wine smells like a damp, mossy...	Rare	98	350.0	Victoria	Rutherglen	N
348	Australia	Deep mahogany. Dried fig and black tea on the ...	Grand	97	100.0	Victoria	Rutherglen	N
349	Australia	RunRig is always complex, and the 2012 doesn't...	RunRig	97	225.0	South Australia	Barossa	N
356	Australia	Dusty, firm, powerful: just a few apt descript...	Georgia's Paddock	95	85.0	Victoria	Heathcote	N
360	Australia	Bacon and tapenade elements merge easily on th...	Descendant	95	125.0	South Australia	Barossa Valley	N
365	Australia	The Taylor family selected Clare Valley for it...	St. Andrews Single Vineyard Release	95	60.0	South Australia	Clare Valley	N
14354	Australia	This wine's concentrated dark fruit shows in t...	Old Vine	95	60.0	South Australia	Barossa Valley	N
16538	Australia	Rich, dense and intense, this is a big, muscul...	The Family Tree	95	65.0	South Australia	Barossa Valley	N
28573	Australia	Astralis has	Astralis	95	350.0	South Australia	Clarendon	N

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

	country	description	designation	points	price	province	region_1	regio
		of Australia's top col...						
34502	Australia	This prodigious wine showcases Barossa's abili...	The Relic	98	135.0	South Australia	Barossa Valley	N
34506	Australia	If Standish's Relic is the feminine side of Sh...	The Standish Single Vineyard	96	135.0	South Australia	Barossa Valley	N
38988	Australia	Penfolds Bin 707 has leapt in quality over the...	Bin 707	95	200.0	South Australia	South Australia	N
39059	Australia	The Taylor family selected Clare Valley for it...	St. Andrews Single Vineyard Release	95	60.0	South Australia	Clare Valley	N
39961	Australia	As unevolved as they are, the dense and multil...	Grange	96	185.0	South Australia	South Australia	N
39962	Australia	Seamless luxury from stem to stern, this 'baby...	RWT	95	70.0	South Australia	Barossa Valley	N
45809	Australia	The 2007 Astralis impresses for its combinatio...	Astralis	95	225.0	South Australia	Clarendon	N
56953	Australia	This inky, embryonic wine deserves to be cella...	Grange	99	850.0	South Australia	South Australia	N
56956	Australia	You may have to scour the country to	Andelmonde	97	95.0	South Australia	Barossa Valley	N

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

	country	description	designation	points	price	province	region_1	regio
56957	Australia	Thorn Clarke has taken its Shiraz to a new lev...	Ron Thorn Single Vineyard	96	89.0	South Australia	Barossa	N
56959	Australia	Is this the Yin to Grange's Yang? The wines ar...	Hill of Grace	96	820.0	South Australia	Eden Valley	N
59977	Australia	This is a top example of the classic Australia...	The Peake	96	150.0	South Australia	McLaren Vale	N
59984	Australia	This is a throwback to those brash, flavor-exu...	One	95	95.0	South Australia	Langhorne Creek	N

In [23]:
```
# q9.hint()

# q9.solution()
```

Solution:

```
top_oceania_wines = reviews.loc[
    (reviews.country.isin(['Australia', 'New Zealand']))
    & (reviews.points >= 95)
]
```

Keep going

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