Introduction

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

In [5]:
```
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("../datasets/winemag-data-130k-v2.csv", index_col=0)

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

```
Setup complete.
```

Exercises

View the first several lines of your data by running the cell below:

In [6]:
```
reviews.head()
```

Out[6]:

	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	

1.

`region_1` and `region_2` are pretty uninformative names for locale columns in the dataset. Create a copy of `reviews` with these columns renamed to `region` and `locale`, respectively.

In [ ]:
```
# Your code here
renamed = ____

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

In [ ]:
```
#_COMMENT_IF(PROD)_
q1.hint()
#_COMMENT_IF(PROD)_
q1.solution()
```

2.

Set the index name in the dataset to `wines`.

In [ ]:
```
reindexed = ____

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

In [ ]:
```
#_COMMENT_IF(PROD)_
q2.hint()
#_COMMENT_IF(PROD)_
q2.solution()
```

3.

The Things on Reddit dataset includes product links from a selection of top-ranked forums ("subreddits") on reddit.com. Run the cell below to load a dataframe of products mentioned on the */r/gaming* subreddit and another dataframe for products mentioned on the *r//movies* subreddit.

In [ ]:
```
gaming_products = pd.read_csv("../input/things-on-reddit/top-things/top-thir
gaming_products['subreddit'] = "r/gaming"
movie_products = pd.read_csv("../input/things-on-reddit/top-things/top-thing
movie_products['subreddit'] = "r/movies"
```

Create a `DataFrame` of products mentioned on *either* subreddit.

In [ ]:
```
combined_products = ____

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

In [ ]:
```
#_COMMENT_IF(PROD)_
q3.hint()
#_COMMENT_IF(PROD)_
q3.solution()
```

4.

The Powerlifting Database dataset on Kaggle includes one CSV table for powerlifting meets and a separate one for powerlifting competitors. Run the cell below to load these datasets into dataframes:

In [ ]:
```
powerlifting_meets = pd.read_csv("../input/powerlifting-database/meets.csv")
powerlifting_competitors = pd.read_csv("../input/powerlifting-database/openp
```

Both tables include references to a `MeetID`, a unique key for each meet (competition) included in the database. Using this, generate a dataset combining the two tables into one.

In [ ]:
```
powerlifting_combined = ____

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

In [ ]:
```
#_COMMENT_IF(PROD)_
q4.hint()
#_COMMENT_IF(PROD)_
q4.solution()
```

Congratulations!

You've finished the Pandas micro-course. Many data scientists feel efficiency with Pandas is the most useful and practical skill they have, because it allows you to progress quickly in any project you have.