Relational Databases in Python

Data Wrangling and Relational Databases

In the context of data wrangling, we recommend that databases and SQL only c
gathering data or storing data. That is:

- **Connecting to a database and importing data** into a pandas DataFrame
structure in your preferred programming language), then assessing and cl
- **Connecting to a database and storing data** you just gathered (which cou
a database), assessed, and cleaned

These tasks are especially necessary when you have large amounts of data, whi
other databases excel over flat files.

The two scenarios above can be further broken down into three main tasks:

- Connecting to a database in Python
- Storing data ***from*** a pandas DataFrame ***in*** a database to which you're conn
- Importing data ***from*** a database to which you're connected ***to*** a pandas Da

This Lesson

For the example in this lesson, we're going to do these in order:

1. Connect to a database. We'll connect to a SQLite database using SQLAlche
for Python.
2. Store the data in the cleaned master dataset in that database. We'll do this
DataFrame method.
3. Then read the brand new data in that database back into a pandas DataFr
pandas' `read_sql` function.

The third one isn't necessary for this lesson, but often in the workplace, instead
files, scrape web pages, hit an API, etc., you're given a database right at the begi

All three of these tasks will be introduced and carried out in the Jupyter Noteboo
quizzes. All of the code is provided for you. Your job is to read and understand e
of code, then run the code.