



SEARCH



RESOURCES

CONCEPTS



17. JSON Files in Python



18. Mashup: APIs, Downloading Fil...



19. Mashup Solution



20. Flashforward 2



21. Storing Data



22. Relational Database Structure



23. Relational Databases in Python



24. Other File Formats



25. You Can Iterate



26. Gathering Summary

27. Conclusion



Mentor Help

Ask a mentor on our Q&A platform



Peer Chat 3

Chat with peers and alumni

Relational Databases in Python

Data Wrangling and Relational Databases

In the context of data wrangling, we recommend that databases and SQL only be used for 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 cleaning the data
- **Connecting to a database and storing data** you just gathered (which could be from a database), assessed, and cleaned

These tasks are especially necessary when you have large amounts of data, which 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 connected
- Importing data *from* a database to which you're connected *to* a pandas DataFrame

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 [SQLAlchemy](#) for Python.
2. Store the data in the cleaned master dataset in that database. We'll do this using the `DataFrame.to_sql()` method.
3. Then read the brand new data in that database back into a pandas DataFrame using pandas' `read_sql()` function.

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

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