Analytics Jumpstart

Working with sqlite

Nashville Software School

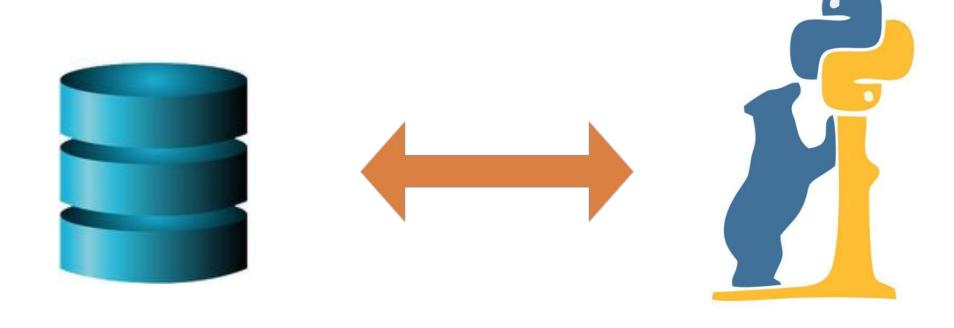


For today

- Using a SQLite database in python
 - create a connection
 - create a cursor
 - execute SQL statement
 - fetchall()



SQLite is an embedded, file-based relational database management system (RDBMS).





```
# load the sqlite3 library
import sqlite3 as sql
# load the database
db = "./data/weather.db"
# create a connection, declare a cursor, and execute a select statement
con = sql.connect(db)
mycursor = con.cursor()
mycursor.execute("SELECT name FROM sqlite_master WHERE type='table' ORDER BY
name;")
# retrieve the data stored in the cursor
tables = mycursor.fetchall()
```



The interactions with the database are done using a query, where you select data based on different criteria.

SELECT name
FROM sqlite_master
WHERE type='table'
ORDER BY name;

-> Which value you want returned

—> Where to look for it

WHERE type='table' —> A criteria to filter the result

-> How you want the result formatted

Many other keywords to help with creating the perfect query:

https://www.codecademy.com/articles/sql-commands

You can write a function like this one to execute a query

```
def get_query(select, db=db):
     "Executes a query and returns results and column/field names."
     with sql.connect(db) as conn:
                                                     declare a cursor
         c = conn.cursor()
                                                     execute a query
         c.execute(select)
         col_names = [str(name[0]).lower() for name in c.description]
     return c.fetchall(), col_names
                                                          return the results of the query
                                                          along with the column names
                                                          The results can be used to
                                                          construct a df; fetchall() gets
                                                          the data and col names gets
                                                          the columns for the df
```

grab column names



But...pandas makes loading the results of a query to a DataFrame easier

df = pd.read_sql_query("SELECT * from my_table;", conn)

And the specified connection

Using the specified query



Questions?

