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()
```

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:
        c = conn.cursor()
        c.execute(select)
        col_names = [str(name[0]).lower() for name in c.description]
    return c.fetchall(), col_names
```

declare a cursor

execute a query

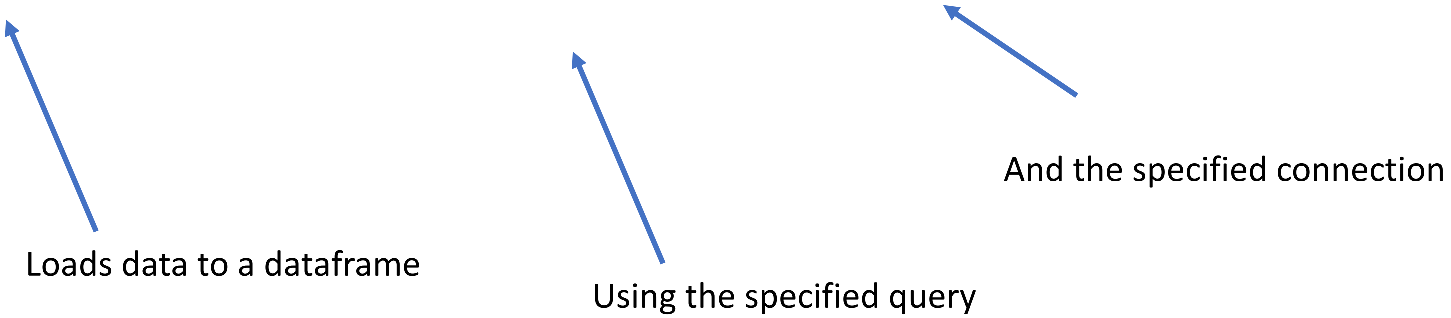
grab column names

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

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

Loads data to a dataframe

Using the specified query

Questions?