Reading in Database Files

The steps to read in a database file using the sqlite library are:

- create a path variable that references the path to your database
- create a connection variable that references the connection to your database
- create a query variable that contains the SQL query that reads in the data table from your database
- create an observations variable to assign the read_sql functions from pandas package
- create a tables variable to read in the data from the table sqlite_master

JSON files are a standard way to store data across platforms. Their structure is similar to Python dictionaries.

NoSQL databases are not relational and vary more in structure. Most NoSQL databases store data in JSON format.

In [ ]:

Reading SQL data

SQL data imports

import sqlite3 as sq3 import pandas as pd

Initialize path to SQLite database

path = "...db "

Create connection sql database

con = sq3.Connection(path)

Write query

query = ' ' ' SELECT * FROM table_name; ' ' '

Execute query

data = pd.read_sql(query, con)

In [ ]:

Reading NoSQL Data ¶

SQL Data Imports

from pymongo import MongoClient

Create a Mongo connection

con = MongoClient()

Choose database (con.list_database_names() will display available databases)

db = con.database_name

Create a cursor object using a query

cursor = db.collection_name.find(query)

Expand cursor and construct DataFrame

df = pd.DataFrame(list(cursor))

In [ ]: