Introduction to creating relational databases and SQL

ETHAN YOO, DATA SCIENCE GRADUATE SPECIALIST

FEBRUARY 14, 2023

ETHAN.YOO@RUTGERS.EDU

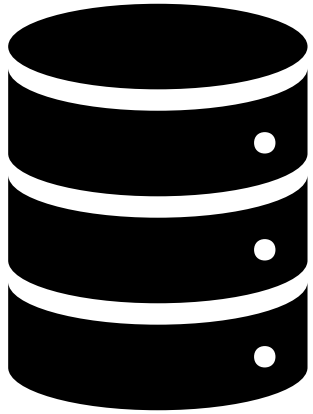
Outline

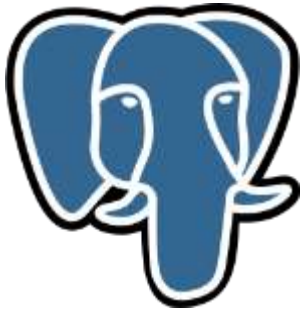
- Terminology
- Introduction to relational databases
- Differences between spreadsheets and relational databases
- SQLite
- Demonstration/walkthrough
- Additional resources

Terminology

- “**SQL** (Structured Query Language) is a descriptive computer language designed for updating, retrieving, and calculating data in table-based databases” (MDN Web Docs)
- Table
- Database engine
- Create, Read, Update, and Delete (CRUD)

Introduction to relational databases





Spreadsheets vs. relational databases

Data types	• CSV is a text-based file format • Binary data (e.g., images) can be stored in databases
Data integrity	• Spreadsheets will not enforce data standards (e.g., non-empty cells or redundancy) • Databases will store only the specified data type in a given column
Data volume	• There is a limit to the number of rows a spreadsheet can store • There is no limit to the number of tables, rows, or columns in a database
Relationships	• Finding related information requires formulas and relies on appropriate data entry • Related tables can be merged to perform queries

SQLite

- Embedded database and database engine
- Public domain
- Digital format recommended by the Library of Congress for datasets
- Major known users from among millions of applications
  - **Apple**: Many native applications across devices
  - **Google**: The Android operating system and Chrome web browser
  - **Microsoft** Windows 10 and other Microsoft products
  - **Mozilla** Firefox (web browser) and Thunderbird (email client)
  - All **Python** distributions since Python 2.5

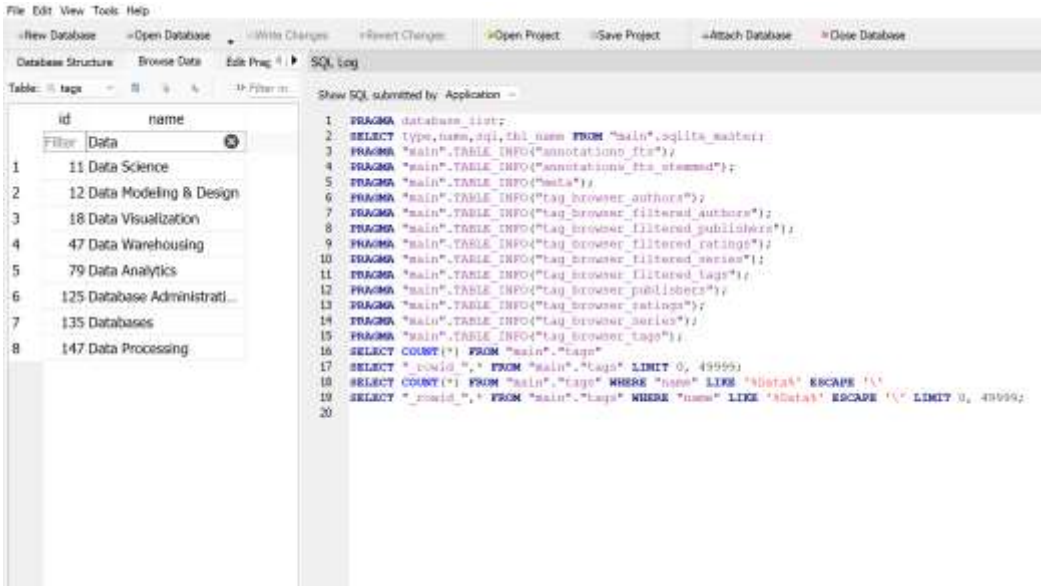
How SQLite stores data

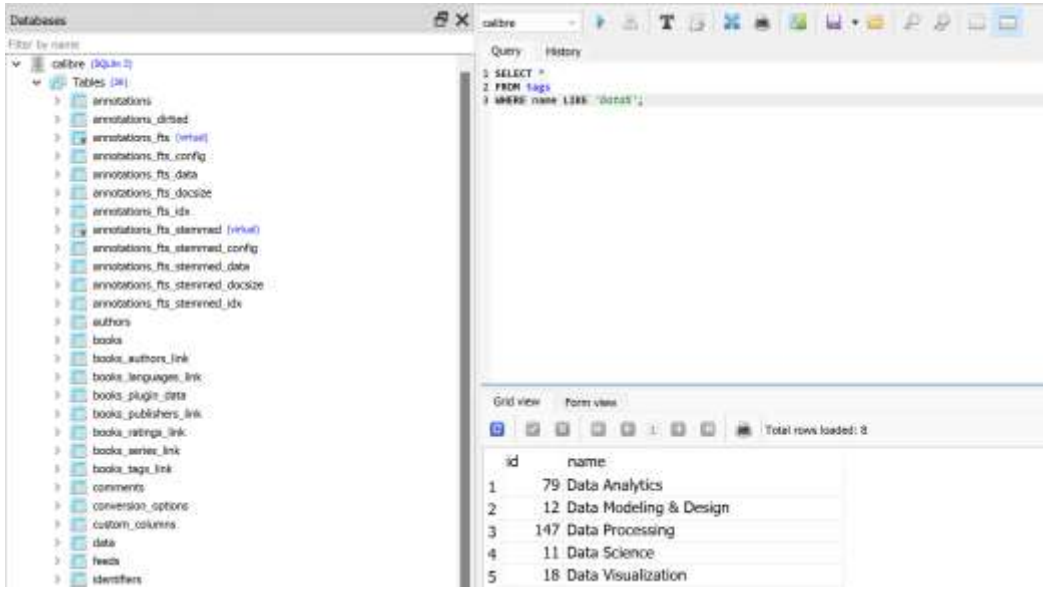
- **Data types (flexible)**
  - `NULL` (missing or unknown values)
  - `INTEGER` (positive or negative whole numbers)
  - `REAL` (real numbers with decimal values)
  - `TEXT` (character data)
  - `BLOB` (binary large object)
- **File type (extensions)**
  - `.db`
  - `.db3`
  - `.sqlite`
  - `.sqlite3`

Database management tools

DB Browser for SQLite

SQLiteStudio





Additional Resources

- **Tutorials**
  - **Data Carpentry (uses DB Browser for SQLite)**
  - **SQLBolt (browser-based)**
  - **SQL Zoo (browser-based)**
- **SQLite documentation**
- **Library of Congress: SQLite, Version 3**
  - **Technical details**
  - **Useful references**