



Programming for Data Analytics

- Databases
- MySQL (and Mongo DB)
- Sam Weiss
- sweiss@cct.ie

cct

College Dublin
Computing • IT • Business

Databases: Introduction

- A database is an organised collection of data
- They provide a few advantages over csv files:
 - Access Control
 - Concurrency
 - Auditing
- Two main types:
 - SQL, Relational
 - MySQL
 - NoSQL, Non-relational
 - Mongo DB

Databases

- Structured Query Language (SQL, “sequel”)
 - Is made up of tables
 - Analogous to a Pandas DataFrames.
 - Data type for each column is defined on creation
- Mongo DB
 - Is a collection of JSON objects
 - JSON = JavaScript Object Notation
 - Is like nested Python dictionaries
 - APIs tend to give data in this format
 - We’ll do APIs next week

CRUD

- The basic operations on a database are CRUD:
- Create
 - Creates new things
 - Can create databases, tables and entries (rows)
 - Does not alter existing things, just adds new ones
- Read
 - Can get data from the database
 - Does not change anything
- Update
 - Updates existing things
- Destroy
 - Deletes existing things

SQL Basic Commands

- Create

- CREATE DATABASE [name]
- CREATE TABLE [name] (col1 TYPE, col2 TYPE, col3 TYPE)
 - Like Python, SQL has data types – more on this later
- INSERT INTO [table] (col1, col2, col3...) VALUES (val1, val2, val3...)
 - db.commit() to save

- Read

- SELECT [col1, col2,...] FROM [table]
- SELECT * FROM [table]
 - Selects everything
- SELECT * FROM [table] WHERE [condition]
 - % is wildcard: %Python% will match with anything containing the substring “Python”
- SELECT * FROM [table] ORDER BY [column] {DESC}
 - {DESC} is optional, will give descending order instead

SQL Basic Commands

- Update

- UPDATE [table] SET [col] = [val] WHERE [condition]
 - db.commit() to save
 - If there is no condition, will update every row!
- ALTER TABLE [table] ADD COLUMN [col] [TYPE]

- Destroy

- DELETE FROM [table] WHERE [condition]
 - db.commit() to save
 - If there is no condition, will delete every row!
- DROP TABLE [table]

SQL Other Commands

- IF EXISTS
 - Will only run if the thing exists
 - Useful when creating databases/tables
 - CREATE DATABASE IF NOT EXISTS test
- NOT NULL
 - Forces all values in this column to have a value
- PRIMARY KEY
 - Forces all rows to have a unique value in this column
- AUTO_INCREMENT
 - Will automatically increment this column, so you don't have to set it
- LIMIT x
 - Only returns the first x results
- OFFSET x
 - Skips the first x results

SQL Data Types

- String: VARCHAR(n)
- Integer: INT
- String: TEXT
- Date: DATE
- More at:
 - <https://dev.mysql.com/doc/refman/8.0/en/data-types.html>
 - https://www.w3schools.com/mysql/mysql_datatypes.asp



Our Dedication.
Your Journey.

cct

College Dublin
Computing • IT • Business