

# SQL for Business and Marketing

*Preliminary; subject to change*

## Summary

These courses provide an introduction to business-oriented SQL with an emphasis on queries.

Classes mostly consist of brief topic introductions followed by extensive interactive problem solving in the form of writing, executing, and checking SQL queries and other statements running on cloud and local databases. This course emphasizes foundations for business-oriented queries that support a wide range of marketing and other analysis. Students will use primary sources (actual SQL documentation for specific databases) as well as supporting material. As time allows, we'll discuss related topics like data modeling, query performance, geospatial data, data privacy, and data security.

## Format

Class will be interactive and will sometimes take a "reverse" approach, with problems worked in class and assigned reading assumed. Sometimes we'll use [Colab](#) or [Jupyter](#) notebooks to work with example SQL databases. Students will probably each have a personal SQL database running in [AWS](#), and we'll share read-only access to a common database.

We'll use [Learning SQL](#) as a guide, but we'll also use primary sources (like [this one](#)).

We'll use a little Python for some tasks, but this course neither teaches Python (beyond some basics as needed) nor expects students to know Python. Python knowledge is not a prerequisite.

The final grade will be based on homework, two quizzes, and two tests. (There is no final.)

## Topics

(\* as time allows)

1. SQL: background, state of affairs
2. What's in SQL
  - a. Basic functionality
  - b. Standards and non-standards
  - c. Specific databases
  - d. What a SQL database can do for and to you

- e. Competitors
- 3. Class logistics
  - a. Accessing our shared database
  - b. Accessing your personal database
  - c. Submitting assignments (using SQL!)
  - d. Online notebooks and other UIs
  - e. Resources
- 4. Basic types (varieties of numbers, text; tables)
- 5. DQL: The SELECT statement
  - a. The SELECTed expression
  - b. FROM
  - c. WHERE
  - d. GROUP BY
  - e. ORDER BY
  - f. LIMIT
  - g. HAVING
  - h. Subqueries
  - i. Functions
  - j. Joins
  - k. Common table expressions\*
  - l. Boolean logic
  - m. Window queries\*
  - n. Query planner and query plans\*
- 6. DDL: Creating (and modifying and removing) structures
  - a. Tables (including constraints, data modeling)
  - b. Indexes
  - c. Other\*
- 7. DML: Creating and changing data
  - a. INSERT
  - b. Importing data
  - c. UPDATE
  - d. DELETE
- 8. Other topics (time permitting)
  - a. Access control\*
  - b. Transactions\*
  - c. Data modeling\*
  - d. "NoSQL"
  - e. Geospatial data\*
  - f. Unstructured data\*
  - g. JSON and structured data\*
  - h. Parallel queries, partitioning\*
  - i. Data security, GDPR, etc\*
  - j. Using SQL from Python, R, Excel, etc

# Schedule

By week:

1. Everything up to SELECT and then start SELECT
  - a. Survey during class
  - b. Assignment 1 due before next class
2. More SELECT and start on joins
  - a. Assignment 2 due before next class
3. More joins; rest of SELECT
  - a. Test at beginning of class
  - b. Assignment 3 due before next class
4. DDL and DML
  - a. Assignment 4 due before next class
5. Other topics
  - a. Test at beginning of class