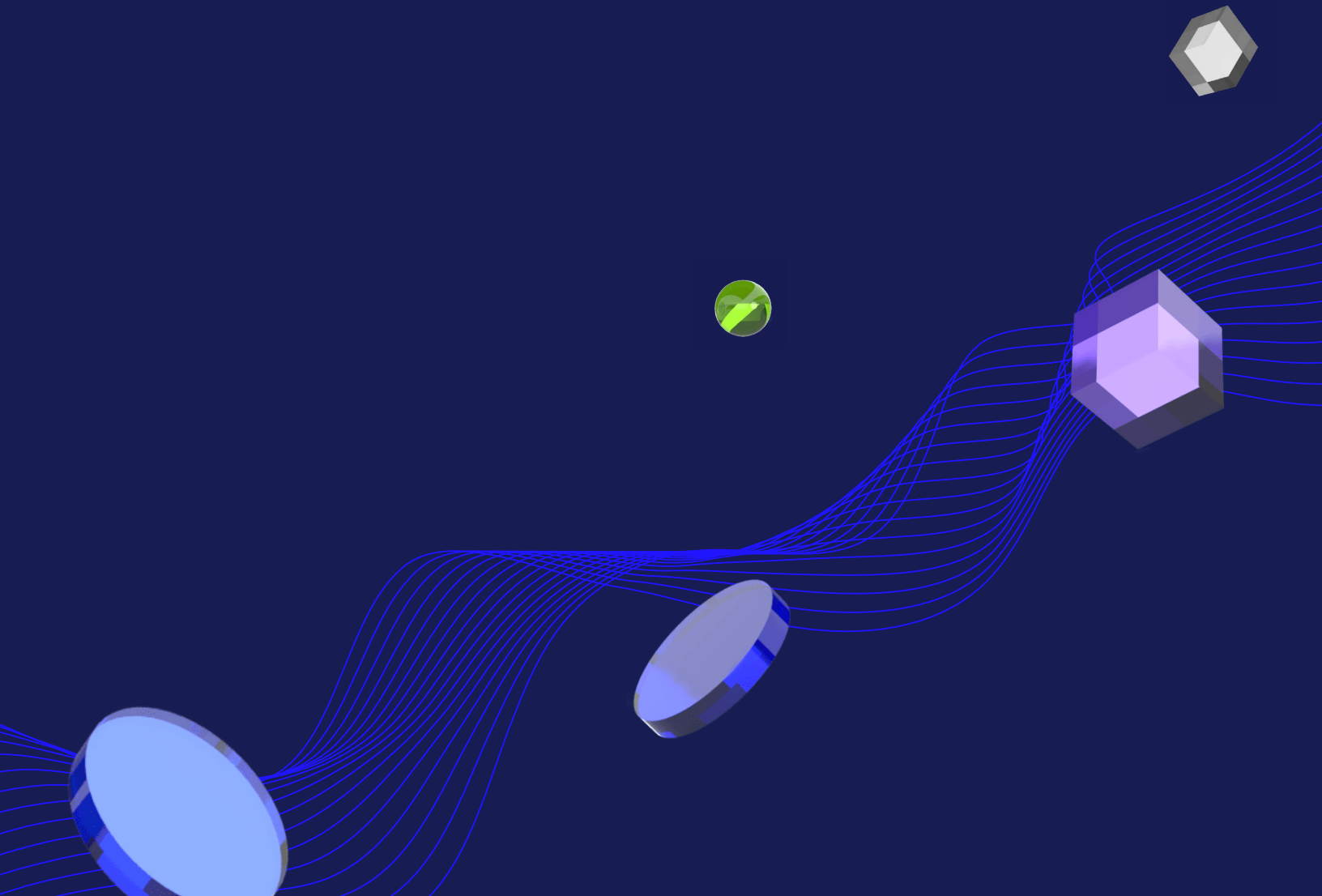U UDACITY

SCHOOL OF DATA SCIENCE

SQL

Nanodegree Program Syllabus

Overview

Perform analysis on data stored in relational and non-relational database systems to power strategic decision-making. Learn to determine, create, and execute SQL and NoSQL queries that manipulate and dissect large scale datasets. Begin by leveraging the power of SQL commands, functions, and data cleaning methodologies to join, aggregate, and clean tables, as well as complete performance tune analysis to provide strategic business recommendations. Finally, apply relational database management techniques to normalize data schemas in order to build the supporting data structures for a social news aggregator.

Program information

⧗ Estimated Time

2 months at 10hrs/week*

▮ Skill Level

Beginner

⚏ Prerequisites

A well-prepared learner should have basic computer skills and understanding of and ability to write SQL and NoSQL languages, plus:

- Familiarity with Operating systems (Windows and MacOS)
- Familiarity and comfort using word processing programs (Microsoft Word, Google Docs, PDF)
- Comfort using presentation software (PowerPoint, Keynote, Google Slides)
- Comfort using Spreadsheet programs (Microsoft. Excel, Google Spreadsheets)

🧑‍💻 Required Hardware/Software

Learners need access to the internet and a 64-bit computer. They will use SQL, NoSQL, Postgres, SQL DDL, and SQL DML.

*The length of this program is an estimation of total hours the average student may take to complete all required coursework, including lecture and project time. If you spend about 5-10 hours per week working through the program, you should finish within the time provided. Actual hours may vary.

Introduction to SQL

When it comes to extracting insights from stored data, SQL is one of the most versatile tools available. Learn how to execute core SQL commands to define, select, manipulate, control access, aggregate and join data and data tables. Understand when and how to use subqueries, several window functions, as well as partitions to complete complex tasks. Clean data, optimize SQL queries, and write select advanced JOINs to enhance analysis performance. Explain which cases one would want to use particular SQL commands, and apply the results from queries to address business problems.

Course Project

Deforestation Exploration

SQL is most commonly used to manipulate and analyze data to inform decision making. In this project, learners will act as a data analyst for an organization on a mission to reduce deforestation around the world and to raise awareness about this important environmental topic. First, learners will clean any erroneous values in a table, join that table to another lookup table to bring in a new categorical and quantitative variable, and return a new view of all categories greater than a reference value. Then, learners will create and execute SQL queries to perform calculations using variables from those disparate data sets to answer questions for stakeholders. Their analysis will help them better understand which countries and regions around the world seem to have forests that have been shrinking in size, and also which countries and regions have the most significant forest area. Lastly, learners will compile their answers and summarize their analysis into a report that can be shared to a leadership team.

Lesson 1

Basic SQL

- Write common SQL commands including SELECT, FROM, and WHERE.
- Use logical operators like LIKE, AND, and OR.

Lesson 2

SQL JOINs

- Write JOINs in SQL to combine data from multiple sources to answer more complex business questions.
- Understand different types of JOINs and when to use each type.

Lesson 3

SQL Aggregations

- Write common aggregations in SQL including COUNT, SUM, MIN, and MAX.
- Write CASE and DATE functions, as well as work with NULLs.

Lesson 4

SQL Subqueries & Temporary Tables

- Write subqueries to run multiple queries together.
- Learn the types of subquery placement and formatting.
- Use temp tables to access a table with more than one query.

Lesson 5

SQL Data Cleaning

- Learn and apply the basics of data cleaning strategies in SQL to normalize or create a column from existing data.
- Perform the appropriate data cleaning methodology based on goals for further analysis.

Lesson 6

SQL Window Functions

- Apply core window functions to tackle analysis tasks that require further targeting or segmentation.
- Use other window functions including RANK, NTILE, LAG, LEAD new functions along with partitions to complete complex tasks.

Lesson 7

SQL Advanced JOINs & Performance Tuning

- Learn how and when to use advanced joins (e.g., self joins) to write queries that run quickly across giant datasets.
- Learn the high-level tradeoffs with queries, including performance and what one can do to optimize them.

Management of Relational & Non-Relational Databases

Databases need to be structured properly to enable efficient and effective querying and analysis of data. Build normalized, consistent, and performant relational data models. Use SQL Database Definition Language (DDL) to create the data schemas designed in Postgres and apply SQL Database Manipulation Language (DML) to migrate data from a denormalized schema to a normalized one. Understand the tradeoffs between relational databases and their non-relational counterparts, and justify which one is best for different scenarios. With a radical shift of paradigms, learn about MongoDB and Redis to get an understanding of the differences in behaviors and requirements for non-relational databases.

Course Project

Uddidit, A Social News Aggregator

Many of today's most popular web applications have supporting database structures that allow them to customize and aggregate information within seconds. In this project, learners will build the supporting data structures for Uddidit, a social media news aggregator site. First, learners will investigate the provided data model for potential errors such as lack of normalization, consistency rules, and proper indexing. Then, learners will create a new, normalized database using DDL based on the denormalized one that is provided. Lastly, learners will write DML queries to migrate the data from the denormalized schema to their normalized schema.

Lesson 1

Normalizing Data

- Organize data in a format suitable for relational databases.
- Get a grasp on database normal forms.

Lesson 2

Data Definition Language (DDL)

- Write common SQL commands with CREATE TABLE and ALTER TABLE.
- Use different data types to model real-world situations.

Lesson 3

Data Manipulation Language (DML)

- Write common SQL commands with INSERT, UPDATE, and DELETE.
- Use SQL functions to manipulate numbers, strings, and dates.

Lesson 4

Consistency with Constraints

- Implement business rules at the database level using SQL commands with CONSTRAINT, UNIQUE, PRIMARY KEY, and CHECK.
- Formalize the relations between tables using SQL FOREIGN KEY and its variations.

Lesson 5

Performance with Indexes

- Fix some slow SQL queries by introducing database indexes with the SQL command CREATE INDEX.
- Introspect SQL queries through the query planner with EXPLAIN and EXPLAIN ANALYZE.
- Assess whether a use case is a good candidate for indexing.

Lesson 6

Intro to Non-Relational Databases

- Articulate why non-relational databases were created, and what are their tradeoffs compared to relational databases.
- Add, modify, and query data in a MongoDB database.
- Use the right MongoDB design patterns for various real-life situations.
- Add, modify, and query data in a Redis database.
- Use Redis as a standalone database to build the data part of a small application.

Meet your instructors.
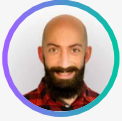
Ziad Saab

Software Developer

Ziad is a seasoned software developer who loves mentoring and teaching. Currently working as an independent contractor, he previously co-founded and taught full stack web development at DecodeMTL, Montreal's first web development bootcamp. In addition to developing software, Ziad also enjoys workng with his clients on their product development.

Malavica Sridhar

Product Manager at Waymo

Mal Sridhar is currently a product manager at Waymo where she works on the scalability team to think through the data and ML needs as the organization scales. Prior to her current role, she was a senior product manager at CircleUp, an investment platform that leverages data and ML to invest and lend to early-stage consumer packaged goods companies. Mal started her career as a management consultant at McKinsey & Company and is passionate about leveraging data to transform industries.

Udacity's learning experience
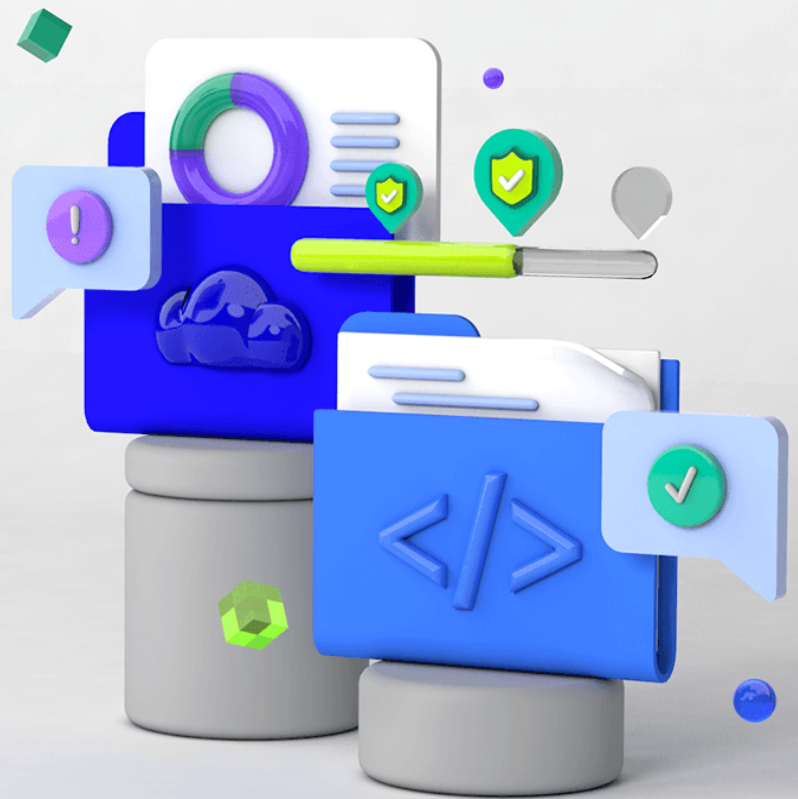
Hands-on Projects

Open-ended, experiential projects are designed to reflect actual workplace challenges. They aren't just multiple choice questions or step-by-step guides, but instead require critical thinking.

Knowledge

Find answers to your questions with Knowledge, our proprietary wiki. Search questions asked by other students, connect with technical mentors, and discover how to solve the challenges that you encounter.

Workspaces

See your code in action. Check the output and quality of your code by running it on interactive workspaces that are integrated into the platform.

Quizzes

Auto-graded quizzes strengthen comprehension. Learners can return to lessons at any time during the course to refresh concepts.
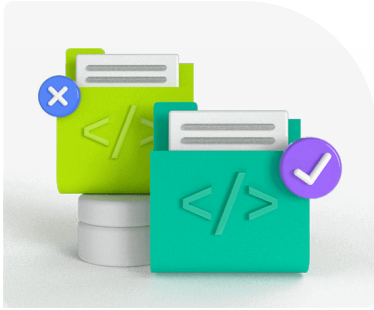
Custom Study Plans

Create a personalized study plan that fits your individual needs. Utilize this plan to keep track of movement toward your overall goal.

Progress Tracker

Take advantage of milestone reminders to stay on schedule and complete your program.

Our proven approach for building job-ready digital skills.

Experienced Project Reviewers

Verify skills mastery.

- Personalized project feedback and critique includes line-by-line code review from skilled practitioners with an average turnaround time of 1.1 hours.

- Project review cycle creates a feedback loop with multiple opportunities for improvement—until the concept is mastered.

- Project reviewers leverage industry best practices and provide pro tips.
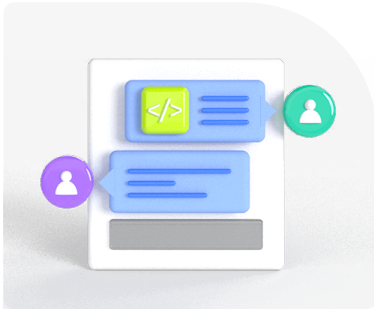
Technical Mentor Support

24/7 support unblocks learning.

- Learning accelerates as skilled mentors identify areas of achievement and potential for growth.

- Unlimited access to mentors means help arrives when it's needed most.

- 2 hr or less average question response time assures that skills development stays on track.
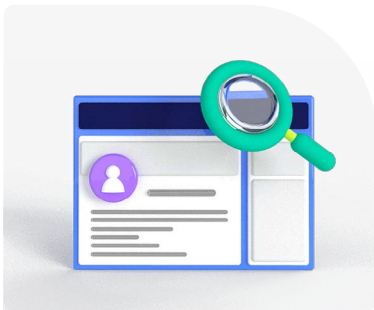
Personal Career Services

Empower job-readiness.

- Access to a Github portfolio review that can give you an edge by highlighting your strengths, and demonstrating your value to employers.*

- Get help optimizing your LinkedIn and establishing your personal brand so your profile ranks higher in searches by recruiters and hiring managers.
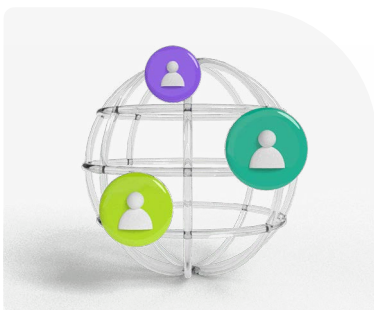
Mentor Network

Highly vetted for effectiveness.

- Mentors must complete a 5-step hiring process to join Udacity's selective network.

- After passing an objective and situational assessment, mentors must demonstrate communication and behavioral fit for a mentorship role.

- Mentors work across more than 30 different industries and often complete a Nanodegree program themselves.

*Applies to select Nanodegree programs only.

UDACITY

Learn more at

www.udacity.com/online-learning-for-individuals →