

Syllabus
CSC230 Database Technologies for
Analytics
Professor Leon Tabak
Block 3
October 18, 2021 to November 10, 2021

Contents

| | |
|--|----------|
| 1 What we will learn | 2 |
| 2 Our meeting times and places | 2 |
| 3 Attendance | 2 |
| 4 Textbook | 2 |
| 5 Other resources, including online courses | 3 |
| 6 Software | 3 |
| 7 Our Plan | 4 |
| 8 Etiquette for the Classroom | 4 |
| 9 Goals | 6 |
| 10 Policies | 6 |
| 11 Grades | 7 |

1 What we will learn

Together we will learn principles and methods for the design of relational databases. We will learn how to create, update, and query databases using the Structured Query Language (SQL). We will examine alternatives to SQL and the relational model. We will practice collecting, filtering, and formatting data. Our work in the laboratory will introduce us to several popular and powerful software tools. We will develop a familiarity with the origins of computing technologies and the most important questions that define computer science. We will explore predictions for the future of computing.

2 Our meeting times and places

- My office is in West Science 211.
- You may call me in my office at (319) 895-4294.
- You may send me electronic mail at ltabak@cornellcollege.edu.
- I will be in my office and available to meet with you in person Monday through Friday from 3:00 p.m. until 3:30 p.m. (Exceptions: I will attend meetings of the Academic Standing Committee on the first Wednesday and the second, third, and fourth Mondays of the block. These meetings begin at 3:15 p.m.)

| | Where | When |
|-----------|-----------------------|----------------------|
| Classroom | West Science Hall 202 | 9:00 a.m.—11:00 a.m. |
| Classroom | West Science Hall 202 | 1:00 p.m.—3:00 p.m. |

Exceptions: We will not meet in the afternoon on the first and last Wednesdays of the block.

3 Attendance

I will record your attendance each day. Cornell College now requires all instructors to record students' attendance. The law requires colleges to keep these records.

4 Textbook

- *Head First SQL: A Brain-Friendly Guide*, Lynn Beighley, O'Reilly Media 2007, ISBN 978-0-596-52684-9

5 Other resources, including online courses

Assignments will also direct you to some of these resources. I have listed other resources here for those who might want to learn more on their own and for those who might find another author's perspective and approach helpful to learning.

- [“SQL for Web Nerds,”](#) by Philip Greenspun
- [Introduction to Databases](#), from Coursera and taught by Professor Jennifer Widom (Stanford University) in a series of mini-courses
- [HackerRank](#)—Learn SQL by solving puzzles and simultaneously build a record that you can show to employers.
- [SQL for Data Science](#) (free course offered by Coursera)
- [Introduction to NoSQL Databases](#) (free course offered by Coursera)
- [Introduction to Relational Databases](#), from Udacity and taught by Karl Krueger
- [Learn SQL](#), from w3schools
- [Learn SQL](#), from Codecademy
- [MongoDB Learn MongoDB](#), from MongoDB
- [MongoDB White Papers](#)

6 Software

We will make use of a server on which I have already installed software for you.

Some of you may wish to install software on your own computers. I suggest that you wait until we meet and discuss options.

We will use MySQL to create databases.

- [MySQL Community Downloads](#)

We will all learn a little about other software products. Some students might want to take a deeper look at some of these products on their own.

- [MongoDB: The database for modern application](#)
- [PostgreSQL: The World's Most Advanced Open Source Relational Database](#)
- [SQLite](#) Download instructions and documentation.

7 Our Plan

Head First SQL differs from most textbooks.

Its pages contain more illustrations, more white space, and fewer words than do most books. You will be able to read a page in this book more quickly than you can read a page in most other books.

However, to learn from this book you will have to do more than read—you will have to answer the questions and solve the problems. There are questions and problems on almost every page.

You will find the answers in this book—but will learn more if you make an effort to produce your own solutions peeking at the author's. We learn by doing. We learn through our fingers.

The book has twelve chapters and a little more than 500 pages—each chapter has 40 or 50 pages. We will read about one chapter of our book each day.

We will also read articles, respond in writing to what we read, and share with one another what we have learned in oral presentations.

I will add examples and exercises to those that we find in *Head First SQL*.

8 Etiquette for the Classroom

Please show respect to your classmates, to your instructor, and to the seriousness of our enterprise by behaving courteously.

Courtesy is an important aspect of professionalism. Courteous and professional conduct will open for you opportunities to work with greater autonomy, to take on greater responsibility, to exercise greater creativity and leadership, and to solve more challenging problems and thereby serve your clients in more significant ways.

While we are studying online, many software engineers are working online. Let us develop the discipline and habits that will help us succeed in the work that will follow our studies at Cornell College.

How might we adapt rules for etiquette in a classroom and office for our online collaboration?

How can each of us be . . .

- a member of the team who delivers work in full and on time?
- a member of the team who delivers products of the highest quality?
- a member of the team to whom others turn for help?

- a member of the team whose critiques and suggestions others welcome?

Here are guidelines that I shared with my students in past years. I would like to hear your opinion: does this advice still apply in our new circumstance? Is there anything here that we should add or subtract?

- Please give your attention to whomever is speaking. You cannot view unrelated pages on the Web and be part of our class' discussion at the same time.
- You learn from your classmates. Be generous in offering help to classmates in the laboratory. Take interest in your classmates' work. Encourage them. Compliment them for work that is well done. Give them a good audience when they stand at the front of the room to present their work. Show these courtesies to all of your classmates.
- Please do not interrupt the class by late entries or early departures. If you anticipate a need to be absent from all or part of one of our meetings, please notify me in advance of your anticipated absence.
- You may listen to music while working in the laboratory so long as you are still able to hear your name when called and you do not disturb neighbors.
- Please refrain from bringing food or drink into the classroom or laboratory. We can make reasonable exceptions for eating that is not noisy and foods that do not have strong smells.
Acceptable beverages and foods include water, tea, and granola bars. Bringing breakfast to class is not courteous.
Please clean up crumbs and spills. Please dispose of empty containers and leftovers.
- Please dress as you might for an employer in the software engineering industry. This does not mean fancy dress—you do not need to buy new clothes. The dress in most workplaces is casual. Just be neat.
Please keep your shoes on. Wearing hoods, hats, or sunglasses (except when there is a medical reason for shielding the eyes) that hide your face is not courteous.
- Imagine that you are seeking employment. How will you present yourself to your prospective employer?
Imagine that you are now employed in a software engineering firm. How will you speak to your teammates, the head of your team, and your company's clients?
Imagine that your grandmother has purchased the company for which you work. She has joined you in the company's conference room to hear and

see you walk through the code that you have written for the company (her company).

Are there some words that you will keep out of your vocabulary during this hour?

9 Goals

We will give special attention to three of Cornell College's [Educational Priorities and Outcomes](#):

- Reasoning — Students will evaluate evidence; interpret data; and use logical, mathematical, and statistical problem-solving tools.
- Communication — Students will speak and write clearly, listen and read actively, and engage with others in productive dialogue.
- Ethical behavior — Students will recognize personal, academic, and professional standards and act with integrity.

Mathematics and logic are foundations of computer science. We will learn how to organize and retrieve data. We will use logic to compose the queries we use to retrieve data. We will discover a mathematical basis for our organization of data and for our operations on that data.

Success in a rapidly developing technological field requires collaboration. That in turn requires effective communication. It also means meeting obligations and supporting teammates. In our classroom, we will learn with and from one another, practicing the kind of cooperation and shared responsibility that will surely be a key to success your work after graduation.

Organizations of all kinds are finding great value in the data that they hold. Data leaks can harm consumers and citizens who expected organizations to protect their privacy. Our discussions will include considerations of the need to prevent the loss of data and to guard against the unauthorized use of data.

10 Policies

Cornell College is committed to providing equal educational opportunities to all students. If you have a documented learning disability and will need any accommodation in this course, you *must* request the accommodation(s) from the instructor of the course no later than the third day of the term. Additional information about the policies and procedures for accommodation of learning disabilities is available on [Cornell College's Web site](#).

Please also familiarize yourself with the college’s statement on [academic honesty](#) and its [policies for dropping courses](#).

11 Grades

You can earn all the points for “Daily Work” by arriving on time for each of our meetings, arriving prepared, and keeping your focus on the work before us. Of course, there might sometimes be good reasons for an absence (for example, illness)—in those cases, just notify me of your need to be absent. There is no competition for these points—I would like to award every student 20 points for daily work!

Experience presenting work to peers will be a central part of the course. Practice asking your teammates questions during their presentations, critiquing their decisions, and suggesting improvements to their code will also be an important part of your education during this term.

You will learn more if you draw upon the knowledge of classmates and generously share what you know with your classmates. Ask questions. Offer suggestions and help. Work together.

In each project, you will demonstrate your understanding of the concepts that we have introduced during the week by answering questions. You will also share the project that you will have developed during the week.

Effort counts. Each student will be starting from a different place. How far you go from your starting point will determine your score for projects. The quality of your presentation, as much as the quality and sophistication of your software, will determine your score. Earn points by explaining clearly what you have learned. Earn points by explaining what we can learn by studying your work.

“Graded Exercises” will include a mixture of work that you produce at home and questions that you answer in writing in the classroom.

| Activity | Points |
|---|---------------|
| Daily work | 20 |
| Graded Exercise 0 (Friday, 22 October 2021) | 20 |
| Graded Exercise 1 (Friday, 29 October 2021) | 20 |
| Graded Exercise 2 (Friday, 05 November 2021) | 20 |
| + Graded Exercise 3 (Wednesday, 10 November 2021) | 20 |
| | <hr/> 100 |

References

- [1] John Akred and Anjali Samani. Your data is worth more than you think. *MIT Sloan Management Review*, January 18 2018.
[Available online.](#)
- [2] Alan Beaulieu. *Learning SQL (Second Edition)*. O'Reilly Media, 2009. (ISBN 978-0-596-52083-0) See the [website](#).
- [3] Vannevar Bush. As we may think. *The Atlantic Monthly*, July 1945.
[Available online.](#)
- [4] Edgar F. (Ted) Codd.
A biographical sketch and Codd's Turing Award Lecture [on the website of the Association for Computing Machinery](#).
- [5] Peter J. Denning. The great principles of computing. *American Scientist*, 98(5):369–372, 2010.
[Available online.](#)
- [6] Andrei Hagiu and Julian Wright. When data creates competitive advantage (and when it doesn't). *Harvard Business Review*, February 2020.
[Available online.](#)
- [7] Brian Hayes. The Britney Spears problem. *American Scientist*, 96(4):274–279, 2008.
[Available online.](#)
- [8] Martin Heller. 8 databases supporting in-database machine learning. *InfoWorld*, April 28 2021.
[Available online.](#)
- [9] Kirk L. Kroeker. Weighing Watson's impact. *Communications of the ACM*, 54(7):13–15, 2011.
- [10] Paul McFedries. The data gold rush [technically speaking]. *IEEE Spectrum*, 48(12):26, 2011.
- [11] Paul McFedries. We're all data geeks now. *IEEE Spectrum*, 52(8):29, 2015.
- [12] Joe McKendrick. Is today's data market outgrowing the relational database model? *Database Trends and Applications*, August 8 2018.
[Available online.](#)
- [13] David Metz. Why Hadoop is the future of the database. *Wired*, February 25 2013.
[Available online.](#)
- [14] Evgeny Morozov. The planning machine. *New Yorker*, 90(31):119–1, 2014.

- [15] Alex “Sandy” Pentland. The data-driven society. *Scientific American*, 309(4):78–83, 2013.
- [16] Alex “Sandy” Pentland. Saving big data from itself. *Scientific American*, 311(2):65 – 67, 2014.
- [17] Tiernan Ray. Michael Stonebraker explains Oracle’s obsolescence, Facebook’s enormous challenge. *Barron’s*, March 30 2015.
[Available online.](#)
- [18] Daniel A. Reed and Jack Dongarra. Exascale computing and big data. *Communications of the ACM*, 58(7):56–68, 2015.
- [19] Nigel Shadbolt and Tim Berners-Lee. Web science emerges. *Scientific American*, 299(4):76–81, 2008.
- [20] Bradley Shimmin. In-database machine learning: Can’t we all just get along? *Towards Data Science*, March 27 2020.
[Available online.](#)
- [21] Michael Stonebraker.
A biographical sketch and Stonebraker’s Turing Award Lecture a biographical sketch and a link to Stonebraker’s Turing Award Lecture [on the website of the Association for Computing Machinery.](#)
- [22] Sherry Tiao. Why you should use your database for machine learning. *Oracle Big Data Blog*, October 7 2019.
[Available online.](#)
- [23] Joanie Wexler. The future of databases and application development. *CIO*, August 24 2020.
[Available online.](#)