IS 510 Databases for Business Analytics

Fall 2017

Course Description

This course introduces databases and data management in three parts. The first part is an introduction to database terminology, with hands-on instructions in Structured Query Language (SQL) SELECT queries needed to extract data from a relational database. The second part is about SQL Data Definition and Manipulation queries, accompanied by the basics of data modeling and normalization needed to ensure data integrity. Specific instruction will be given to the use of SQL data sources in Python/Pandas. The course concludes in the third part with a comprehensive data engineering project that gives each student the opportunity to integrate and apply their new knowledge and skills. (Prerequisite: IS 505 or equivalent; Credit hours: 3)

Learning Objectives

- To introduce and reinforce fundamentals of data modeling and relational database systems
- To provide students with hands-on knowledge of Structured Query Language and the basics of relational database administration
- To provide students with an understanding of design and implementation concepts such as entity-relationship modeling and normalization of database tables
- To enable each student to apply course concepts through a database-based business analytics project
- To integrate with the Python/Pandas materials covered in IS505

Instructor

Christopher L. Huntley
Dolan School of Business, Room 1122
Email: chuntley@fairfield.edu

Office Hours

- Tuesdays 11am-noon & 5:00-6:15pm
- Fridays 2:00-3:30pm
- Other times are by appointment only. However, email works well, especially when the instructor is otherwise unavailable.

Resources

- **Textbook:** Kroenke and Auer, Database Processing: Fundamentals, Design, Implementation, 14-th Edition, Pearson, 2016.
- **Online Tutorials:** DataCamp's online SQL and Python Database courses.
- **Software:** MySQL Server (Relational Database), MySQL Workbench (DB Admin), Lucidchart (diagramming tool), GitHub Desktop (source control), Atom (code editor), and Anaconda/Jupyter (Python environment)
- **Reference Docs:** The official MySQL Manual covers the MySQL server and coding. When in doubt, RTFM.
- **Hardware:** Each student is expected to bring to class a recent-vintage laptop to class. The student must have admin rights to install software, manage ports, and set file permissions. The laptop should have at least 5GB of free space. In addition, they are expected to bring headphones / earbuds for doing online tutorials in class.
- **Websites:**
  - Class documents (including this syllabus) are Dr. Huntley's is510-docs GitHub repository.
  - Programing assignments are posted and managed in GitHub Classroom.

Student Expectations

This is a graduate class intended for serious professionals:

- Expect to spend at least 12 hours per week attending class and completing your assignments.
- Do your own work. There is no professional benefit to pretending that somebody else's work is your own. We will follow the university's academic honesty policy to the letter.
- Be a good teammate and class citizen. Free-riding and other unprofessional behavior will result in immediate consequences, which may include failing the course.
- The course has IS 505 Python for Analytics as a pre-requisite. However, this pre-requisite may be waived for MBA and MSBA students who matriculated before Fall 2017. If you are one of these students, please notify the instructor to arrange for proper accommodation.

If you cannot abide by these policies then please take another course.

Assignments and Course Grade

Quizzes (50% of Course Grade)

There will be five quizzes, with the lowest grade dropped from your quiz average. They will be given on the days shown in the attached schedule of classes. Quizzes are closed book and closed notes. They are designed to last about 25 minutes, though students will be permitted to work on them up to 40 minutes. The remainder of the class time will be devoted to new material. Please notify the instructor if you require extra time or support services.

Final Project (40% of Course Grade)

The final project is meant to a demonstration that the course objectives have been met. Teams will work in teams of 2-3 to complete a project of intermediate complexity and scope. Project details and requirements will be distributed in the fifth week of class.

Professionalism (10% of Course Grade)

Given the large class enrollment and the technical nature of the subject, it is critical that everyone take a professional approach to learning the material. Students are expected to come to class prepared and then participate in all activities to the best of their abilities. Cheating, free-riding, and other unprofessional behavior will not be tolerated.

Schedule

- Dates and assignments are subject to change, usually with one week's notice.
- **Classwork** is to be completed before leaving class.
- **Homework** is to be completed before the next class.
- Saturday class dates are shown **in bold**.
- **K/A #** refers to Kroenke and Auer Chapter #

Date	Topics	Classwork	Homework
Oct 31	Course Introduction Software Setup Database Concepts (K/A 1)	Software Installations Account Registrations System Check	K/A 1 DataCamp

Date	Topics	Classwork	Homework
Nov 4	Quiz 1: Concepts SQL Select Queries (K/A 2): Joins, Functions, Aggregates, Subqueries	Deals DB (part 1)	K/A 2 (skip pps 37-44, 48-58) DataCamp
Nov 7	SQL Select (continued) Relational Model and Normalization (K/A 3)	Movies Tonight (part 1)	K/A 3
Nov 14	Quiz 2: SQL Select Queries Entity-Relationship Modeling (K/A 5)	Movies Tonight (part 2)	K/A 5,6
Nov 18	Quiz 3: Relational Data Modeling ERD → Table Designs (K/A 6) SQL Data Definition (K/A 7)	Movies Tonight (part 3)	K/A 7
Nov 21	SQL Data Manipulation (K/A 7) Python, Pandas, and SQL Jupyter %sql magics	Movies Tonight (part 4),Deals DB (Part 2)	DataCamp
Nov 28	Quiz 4: SQL DDL/DML SQLAlchemy	Movies Tonight (part 5) Project Kickoff	DataCamp
Dec 5	Data Scrubbing and ETL	Project Workshop	
Dec 9	Quiz 5: Python Integration	Project Workshop	
Dec 12	Project Presentations: 10 minutes per team plus Q&A		