B8136: Intro to Databases for Business Analytics
Spring 2019 (B Term)

Mattan Griffel
Email: mattan.griffel@columbia.edu
Office Hours: Available upon request

Section 1: Thursdays 4:00PM – 5:30PM
Section 2: Thursdays 6:00PM – 7:30PM

NOTE: Both sections are a hybrid online/offline course consisting of one and a half (1.5) hours of online video and one and a half (1.5) hours of in-person class time per week. Watching the online videos is mandatory and must be completed before the in-person class each week.

Course Description

We don’t think about databases much, right? At least not when they’re working right. But they’re all around us. They’re in every product we use. And when they don’t work (think about the iCloud, LinkedIn, or Ashley Madison data breaches in which hundreds of millions of emails and passwords were exposed) the consequences can be extreme.

Every modern company stores their data in a database (it’s like a really big version of Excel), and if you want to analyze the data, you may be expected to know how to access it yourself. In fact, at many companies are requiring even their business leaders to have an understanding of databases. At the very least, knowing how to set up and interact with databases will improve your ability to GSD (get stuff done), strengthen your understanding of how technology works, and make you less of a pain for developers to work with.

In this class, we’ll explore basic SQL (the most common database language) for business analytics. At the end of the course, should should have a deeper understanding of how databases work, how they fit into the general technology stack, how to connect to databases, and know how to browse and exporting data from databases.

Required Prerequisites

This course assumes no previous knowledge of programming or code.

Required Course Material

- This course does not use a textbook.
- Students must have a laptop that they can bring to class – Mac or PC is fine, as long as your operating system is up to date (at least Windows 7 and Mac OS 10.8).
- Slides and files will be uploaded to Canvas after class.

Online Video

Each week, students will be expected to watch approximately one and a half hours of additional online video content before attending class. Material in the class will build on the content covered online, and students should be prepared to answer questions related to online material.

Video content will be provided via Canvas.
# Course Roadmap/Schedule

<table>
<thead>
<tr>
<th>Session</th>
<th>Topic</th>
<th>Assignment Due</th>
</tr>
</thead>
</table>
| **Class 1** Thursday, Mar 28 | **Querying Bootcamp:**  
  - Overview of the course  
  - What is SQL?  
  - Installing SQLite, text editor, and command line  
  - Command line crash course  
  - SQLite vs Postgres vs MySQL  
  - SELECT  
  - Math  
  - WHERE  
  - AND  | Pre-work for Class 1 Due |
| **Class 2** Thursday, Apr 4 | **Querying Bootcamp pt. 2:**  
  - Renaming columns  
  - Concatenating data  
  - String searches using LIKE  
  - Matching multiple values using IN  
  - Searching by dates & times  
  - Distinct, order by, limit,  
  - Case (if...then statements)  
  - Intro to aggregate functions  
  - Subqueries  
  - Joins  | Pre-work for Class 2 Due  
Assignment 1 Due |
| **Class 3** Thursday, Apr 11 | **Querying Bootcamp pt. 2:**  
  - Truth tables  
  - Aggregate functions  
  - Group by  
  - Extracting from datetime  
  - Having  | Pre-work for Class 3 Due  
Assignment 2 Due |
| **Class 4** Thursday, Apr 18 | **Creating Databases:**  
  - Setting up a database instance on Amazon  
  - Creating a new database  
  - Creating tables  
  - Loading CSV data  
  - Inserting, Updating and Deleting Data  
  - Cleaning Data  
  - Primary Keys and Constraints  
  - Updating Tables  
  - Multi-relational tables  
    - One-to-Many  
    - Many-to-Many  | Pre-work for Class 4 Due  
Assignment 3 Due |
Class 5  
Thursday, Apr 25  
Web Apps/Security + Data Analysis:  
- Connecting a web app to a database  
- SQL Injection  
- Data Analysis & Tools  
- Yammer Case  
Pre-work for Class 5 Due  
Assignment 4 Due

Class 6  
Thursday, May 2  
Data Analysis pt. 2  
- Yammer Case Continued  
Pre-work for Class 6 Due

Due  
Sunday, May 5 Midnight  
Final Project  
SQL Cheat Sheet Due

Grading

Final grades in the class will be calculated as follows:

- Participation (30%): Students will contribute in class by sharing solutions to challenges on Slack (an online instant messaging and chatroom tool).
- Assignments (40%): There will be four homework assignments that should be completed individually.
- Final Project (30%): There is a take-home final project that should be completed with a individually.

Late assignments will be accepted with a 20% penalty any time before the final class.