

# Homework 1A: Foundation Concepts (I)

COMS W4111: Introduction to Databases

Sections 002

Spring 2025

(v 1; 2025-JAN-28)

## **Notes:**

- HW 1 is due on Sunday, 08-FEB at 11:59 PM.
- There are two parts to HW 1 – part A and part B.
- This document defines HW 1A.
- HW 1B will require material from the 31-JAN lecture and will be published before the 31-JAN lecture.
- Both the programming and non-programming tracks complete HW 1A.

## Submission and Overview

### Grading and Scope

Total points for homework assignments and exams determine final grade. The final point total is between 0 and 100. HW 1 is worth 5 points. HW 1A is worth 2.5 points and HW 1B is worth 2.5 points.

The scope of material for this HW1 is:

- The material in lecture 1.
- The material in lecture 2.
- The slides associated with the recommended textbook for
  - Chapter 1.
  - Chapter 2.
  - Chapter 3, slides 3.1 to 3.36.
  - Chapter 6, slides 6.1 to 6.24.

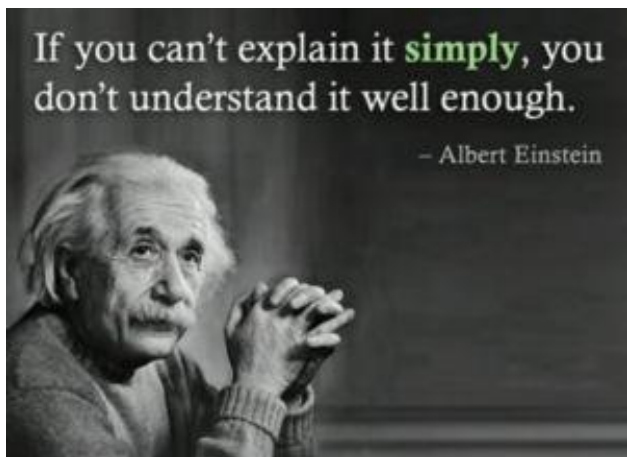
## Submission

**Due date: 2025-Feb-09, 11:59 PM EDT on GradeScope.**

You submit on GradeScope. You upload a PDF of this document with your answers entered. You must [assign pages to the questions](#) in the outline in GradeScope.

There is a [post/mega-thread](#) on Ed Discussions that we will use to resolve questions and issues with respect to homework 1.

## Brevity



**Keep your answers focused, brief and succinct.** The answers to written questions only require 3 or 4 sentences/bullet points. If you ramble and bloviate hoping to get something correct, we will deduct more points.

## Questions

### General Knowledge Questions

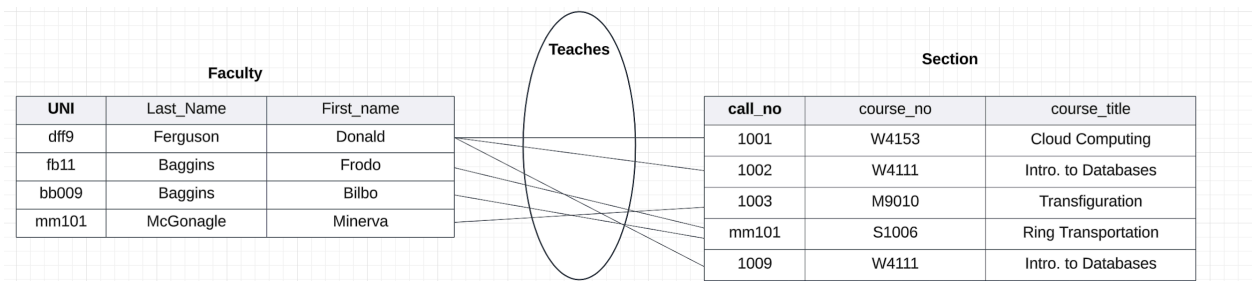
1. Is data in a spreadsheet unstructured, semi-structured or structured. Briefly explain your answer?
2. Professor Ferguson can export SSOL data for his classes to a spreadsheet. Despite the data being spreadsheet-like, Columbia uses an application and database to manage the data. An alternative would be to make shared spreadsheets available to everyone and

let students, faculty and staff collectively edit the spreadsheets. This would be chaos. List 4 functions/capabilities of a database management system that makes using a database and application superior to sharing spreadsheets.

3. What are the three levels of abstraction for data that a DBMS provides? What are two disadvantages of having a user or developer directly use the lowest level and directly access the data without going through the higher abstractions?
4. Is a full stack web application a two-tier or three-tier application? Is a Jupyter notebook a two-tier or three-tier application? Explain your answer. We saw an example of a two-tier application/product in lecture 1. What was the product?
5. What are the four types/categories of database users? Which type of user is most likely to use DDL?

## Entity Relationship Model

6. Consider the entity set with relationships below. Assume that the bold attribute/column is the primary key. Write the relational model schema definitions for representing the entity sets *Faculty* and *Section*.



7. Using the approach for documenting relationship sets in Lecture 1's slides, write down the relationship set *Teaches* for the diagram in question 6.
8. For the diagram in question 6, draw the *conceptual model* Crow's Foot Diagram using [Lucidchart](#).

## Relational Algebra

9. Using the diagram from question 6, what is the result of the relational algebra expression  $\Pi_{\text{course\_no}, \text{course\_title}}(\text{Section})$ ?
10. Using the relation below, what is the result of the expression  $\pi_{b,c}(\sigma_{a>3}(R))$

R.a	R.b	R.c
1	'a'	'd'
3	'c'	'c'
4	'd'	'f'
5	'd'	'b'
6	'e'	'f'

## SQL

11. What is the SQL equivalent to the relational algebra expression  $\pi_{\text{name}, \text{dept\_name}}(\sigma_{\text{dept\_name}='Comp. Sci.'}(\text{instructor}))$ .
12. Translate the following relational model definition into SQL. You may assume that the data type of all columns is TEXT.
- section(courseNo, sectionNo, semester, year, courseName, enrollment)
13. Using the diagram from question 6, translate the relational algebra statement  $\Pi_{\text{course\_no}, \text{course\_title}}(\text{Section})$  into SQL. What is the result of executing the SQL statement?
14. Using DataGrip and the data from the recommended text book that you loaded in homework 0, write a SQL statement that selects the ID, name, dept\_name and tot\_cred of students. Your result should include only students that are in the 'Comp. Sci.' department or who have tot\_cred > 50. Place the text of your query and a screen

capture of the execution below.

15. Using DataGrip and the data from the recommended text book that you loaded in homework 0, write a SQL statement that returns a table containing only dept\_name. The table should not have duplicated. Place the text of your query and a screen capture of the execution below.