W4111 – Introduction to Databases
Sections 002, V002; spring 2022

Homework 1 – Written Assignment

Instructions

- The homework submission date/time is 06-Feb-2022 at 11:59 PM.
- Submission format is a PDF version of this document with your answers. Place your answers in the document after the questions.
- The name of your PDF must be <UNI>_S22_W4111_HW1_Written.pdf. For example, mine would be dff9_S22_W4111_HW1_Written.pdf
- You must use the Gradescope functions to mark the location of your questions/answers in the submitted PDF. Failure to mark pages will cause point deductions.
- You can use online sources but you must cite your sources. You may not cut and paste text..
- Questions typically require less than five sentences for an answer. You will lose points if your answer runs on and wanders.

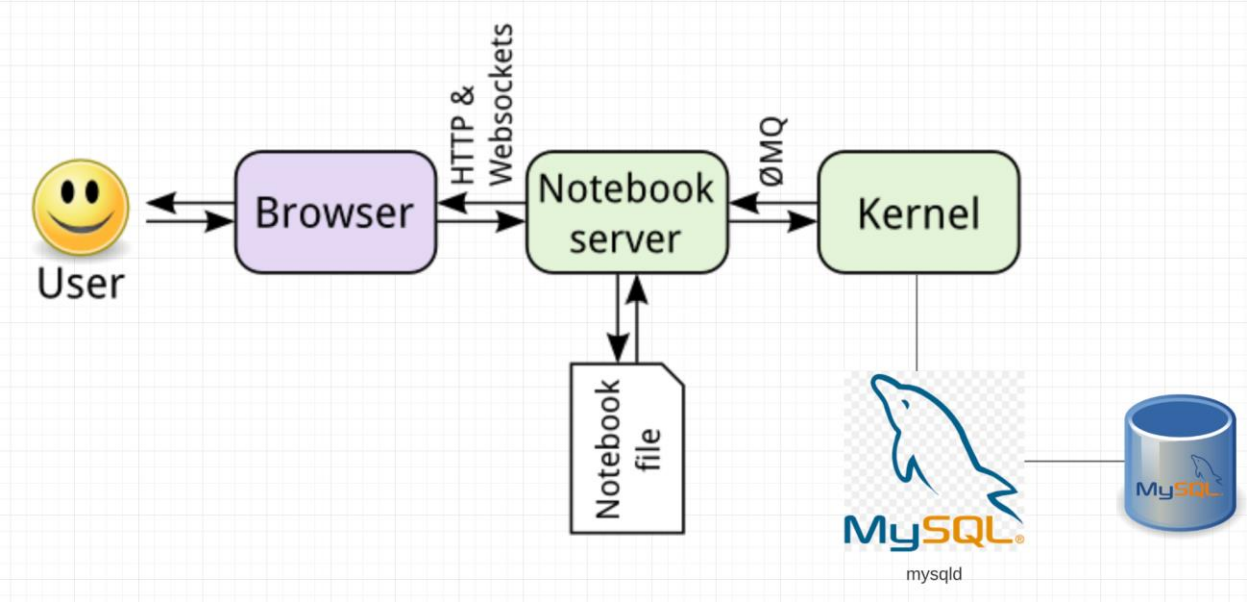
  "Verbosity wastes a portion of the reader's or listener's life."

Questions

Question 1: Briefly explain the terms *structured data, semi-structured data* and *unstructured data.* Give an example of each type.

Question 2: Briefly explain the concept of *metadata*. For a presentation (PowerPoint, Google Slides), what would be some examples of metadata?

Question 3: The following diagram is an overview of Jupyter Notebook's runtime model when the notebook is using MySQL. Is this a 2-tier application or 3-tier application? Briefly explain why.



Question 4: Briefly define and explain procedural and declarative languages. Is SQL procedural or declarative?

Question 5: List 4 advantages/differences of database management systems (DBMS) compared to programs and files for data processing. List two disadvantages of DBMS?

Question 6: In a relational DBMS, columns/attributes should be *atomic.* Briefly explain what this means. If a table has a column *name* of the form “last name, first name”, is this atomic?

Question 7: Attributes/columns have *types,* e.g. int, varchar(128), timestamp. An attribute/column values must be from a *domain*? What is the difference between a type and a domain (hint: domain constraints)?

Question 8: There are four common types of people that interact with a database management system. List and briefly explain each of the four types.

Question 9: Briefly explain the concepts of database *instance* and *schema*?

Question 10: Explain the concept of *physical data independence* and the importance of the concept.