

You have 70 minutes to complete this exam. Time begins after we have all read through the exam questions together and time ends promptly 70 minutes later. You may not read the exam questions before we read them together, and you may not write anything while we are reading those questions together.

Please record your answers in a Word file, in a text file, or on a piece of paper (which could be the last page of this exam or any other paper). You can answer some questions one way (e.g., as Word) and others the other way (e.g., by writing on paper). At the top of the first page of your answers, write your name and the date. If you answer any questions using a file, submit that file on ELMS and also email it to both oard@umd.edu and jyothikv@umd.edu. If you answer any questions on paper, turn in that paper. Make sure your name is on everything that you turn in! And if you use both paper and a file, make a note on each about what can be found in the other so that we don't miss any of your answers.

You may use any information and software that existed before the start of this exam. This means (among other things) that you may search the Web. You may NOT communicate with any other person other than one of the instructors for any purpose during the exam period, either in person or in any other way, and you may not post anything to any location for any purpose during the exam period. Note that this means you may not have skype or any instant messaging application active on any device that you use during the exam, and that that even if you leave the exam room early you may not talk with anyone about anything, you may not send or receive any email, etc. until the exam period ends at 3:15 PM.

Hand write and sign (or, if you type your answers, hand type – no cut and paste – followed by your name) the honor pledge on this exam. (For reference, the honor pledge as stated at <http://www.umd.edu/catalog/index.cfm/show/content.section/c/27/ss/1583/s/1604>, is: “I pledge on my honor that I have not given or received any unauthorized assistance on this exam.”)

As strategies for completing the exam, keep the following in mind:

- If you find a question to be ambiguous, you may come to the front of the room to ask about it, but please do so in a way that other students can't hear. If you don't get an answer that resolves your question, then please explain your confusion and any reasonable assumptions that you have made in order to answer the question and include those with your answer so that they can be considered during grading.
- You are more likely to get partial credit for an incorrect answer if you show your work.
- **Be careful not to spend too much time on any one question.** Plan ahead, and don't devote more time to a question than it is worth.

*** WRITE YOUR NAME ***

Everyone must answer question 1.

1. [60 points] Design and build a database for a shipping company such as Federal Express in which packages are shipped from one customer (the sender) to another customer (the recipient). Any customer may ship any number of packages to any other customer. For each customer, record contact information such as their name, their telephone number, and their address. For each package, record which customer is the sender and the recipient. For full credit, create a two-table database using Access or Base. Enter data for 5 shipped items so that two packages are shipped by one specific customer (e.g., both are shipped by Doug), and two other packages are shipped to some other specific customer (e.g., both are shipped to Jyothi). Then create one query that lists the name and address of the sender and the recipient for each package. Be sure to save your query. Submit the .odb (for Base) or .mdb (for Access) file by ELMS and email. For partial credit, you could instead show your design on paper. If you choose to do that, you should specify as much of what you would have done using Access or Base as possible, including the table design, the table contents, and the query design. If you submit an .odb file or an .mdb file you do not need to also submit anything on paper (but you may).

Answer any two of the following three questions. If you answer all three, only the first two will be graded.

2. [20 points] State one advantage of using a database to store just the content of each Web page when using a Web Content Management System (CMS) that could not be achieved as easily by having the CMS store the HTML for each complete Web page.
3. [20 points] One problem with linked data is that some of the available data may be incorrect, either because of inadvertent human error or because of malicious behavior. Explain the consequences of this fact for how linked data can productively be used.
4. [20 points] One challenge in data mining is that a pattern might be explained by an almost unlimited number of possible “models.” For example, if we train a model with two facts: Sally (a woman) is 5’6” and Jim (a man) is 5’5”, then the system might learn that women are taller than men, or that people with longer names are taller than people with shorter names, or that people whose names begin with S are 5’6” tall. Of course, we would train real models with more data than that, but there are always a very large number of possible models. Recall the “1, 1, 2, ...” example from class as another example of how many different models might explain the data that we used to learn a model. Explain how a real data mining system can choose which model to use from among the possible models.

*** WRITE AND SIGN THE HONOR PLEDGE ***