The purpose of this database assignment is to give you the basic ideas of database as well as some hands-on experience with Database Management System (DBMS). As you have learned in class, the main goal of a DBMS is to enable people storing, accessing, and manipulating large amount of data. To give you a rough idea of “large amount of data”, you are provided with a data set consisting of more than 7 thousands of records. We will use Microsoft Access 2003 as the DBMS for this assignment (you could also use later version of Access).

The data set is adopted from Damdarn[1] which records individual company’s corporate finance and valuation within the US market in 2006. Some of the columns are discarded for the sake of simplicity (you could find original data set at [1]).

The data set is in Microsoft Excel Spreadsheet format (xls), and you need to convert it into Microsoft Access database format. To do this, please refer to the URL below for instructions of how to import Excel data into Access:


Once you have a working database (a table, to be precise), you could then analyze the data and start to answer interesting question. You will need to use queries to extract information from the table.

(1) For starter, please create queries in Access to provide answers to the following questions:

1. Make a list with all the columns but grouped records with respect to Industry Name.

2. Sort the table by ascending order of Trading Volume.

3. How many companies are listed in the data set? How many industry categories (the Industry Name column) are there? How many companies under each category?
4. Taking the Computer Software/Services industry as example, make a list of all the information of companies in this industry by a query.

5. What is the average (unweighted) stock price of companies within the Computer Software/Services industry? What is the total amount of Trading Volume?

6. Pick out the companies in the Computer Software/Services industry which have the Growth in Revenue-last year values, and count the number of companies have positive/negative growth percentage respectively, using queries.

For this part, please include the screenshots of your queries into the submitted report, the long lists generated are not necessary.

(2) After completing the above small problems, you shall have some working knowledge of queries already. Then, please think of another 3 similar problems based on the data and provide the solution queries for them.

Please clearly state your problem first, and then explain how you solve it.

For this part, your report should include the following information:

- The description of your problems.
- The screenshots of your solution queries.

(3) The final part of this project is to do a preliminary analysis and draw some simple conclusions about a topic of your interest. Based on the data, you could restrict your attention to a particular industry, and give a general profile of the corporate finance information of companies within that industry.

To do this, you will first need to think of some small questions similar as Part (1) and (2). Those questions will lead to you a set of information pieces. Then, you could integrate those pieces of information together and provide some conclusions.

As a more concrete guideline, you could find an industry segment and pick up some performance metric of your interest. Then you could make some queries to compare companies within that industry segment based on that performance metric. In your report, you need to discuss what is interesting about that metric as well as what the differences of that metric between companies may suggest. (Also, you could pick up a set of metrics to evaluate as well as compare multiple industry segments, which may lead to more interesting conclusions.)
For this part, your report should contain the information below:

- Identifying the specific metric and relevant data within the dataset.
- The reason to choose that (those) metric(s). Why is it interesting?
- How do you use that metric to evaluate companies?
- What are your evaluation results? What may your results suggest?
- The summary of your conclusion.

References