The purpose of this database assignment is to give you the basic idea of what a database is, as well as some hands-on experience with Database Management Systems (DBMS). As you have learned in class, the main goal of a DBMS is to enable people to store, access, and manipulate large amounts of data. To give you a rough idea of a “large amount of data”, you are provided with a dataset consisting of more than 6,800 records. We will use Microsoft Access 2003 as the DBMS for this assignment (you can use a later version of Access if you prefer).

The data set is adopted from Damdarn[1] which records data relating to all publicly traded firms that had a market capitalization of more than $10 million dollars. 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 questions. You will need to use queries to extract information from the table.

1) Please create queries in Access to answer the following questions (provide the query only for your answer):

1. Make a list with all the columns but group records (sort) with respect to Industry Name.
2. Sort the table in descending order of Total Debt.
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 Information Services industry as example, display all information in the table for the companies in this industry.
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. What Information Services companies had a positive Growth in Revenue last year? What companies had a negative Growth in Revenue last year? What companies had more than 50% Growth in Revenue last year?

2) After completing the above small problems, you should have some working knowledge of queries already. Now, come up with 3 queries on your own. Please be sure to include the following information:

- Screenshots of your query results in Access
- Clear description of your problem
- The query you used to solve your problem
3) Finally, you will analyze and find a conclusion on a topic of your interest. You could choose a specific industry and then create a profile of that industry based on specific metrics, or you could show that certain related industries have comparable metrics (i.e. the growth rate of *Industry A* is similar to that of *Industry B*).

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

As a more concrete guideline/example, 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 (i.e. choose 20 companies and show how they respond to your performance metrics). 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)

You will be graded on the following:
- Identify the metric/performance criteria and relevant data in the dataset.
- Your reasoning for choosing the metric/performance criteria. What makes it interesting?
- How do you use that metric to evaluate companies?
- What are your evaluation results? What may your results suggest?
- The summary of conclusion.
- Include all your queries
- Include a screenshot of any relevant results to your analysis

References