Georgia State University
J. Mack Robinson
COLLEGE of BUSINESS

Data Analysis Assignment

"Every day I wake up and ask, 'how can I flow data better, manage data better, analyse data better?'" Rollin Ford, WalMart

This may be completed individually or in groups up to 5 students. Notify the instructor at least one week before the assignment is due of the names of the members in your group or whether you are completing the assignment by yourself. This will help to ensure that the dropboxes are set up before the submission is due.

Data Mining

As the amount of data stored in corporate databases and warehouses continues to grow, there is an increasing emphasis on analyzing and interpreting the data and using the results to make improvements and innovations in business processes. The objective of this assignment is to perform a simple data mining exercise in order to appreciate the challenges and usefulness of finding patterns in data and acting upon these patterns.

Data mining, in general, strives to identify valid, novel, potentially useful, and ultimately understandable patterns in data. The challenge is to identify the patterns and decide whether it is possible to act upon them.

Answer the following questions.

1. [40 Points] Identify 5 patterns in the provided *Coffee Example.xlsx* data file. Compare, for example, the size of the coffee and the length of time the customer is in the Coffee Shop to find a pattern (relationship), if any, between these two variables. This is a 2-variable comparison. At least one of the patterns should include three or more variables. List the patterns you find and describe how you identified them.

2. [40 Points] For each pattern identified, propose a change to the Coffee Shop to make it more profitable and explain your recommendation. If it is not possible to act on a

pattern, then state, "no change" and justify why you consider the pattern to be irrelevant or not useful. These could be presented in a table.

Note that the recommendations must make sense based on the patterns identified.

3. [20 Points] Describe the implications of this data mining exercise. What else would you need to know if this were a real-world application? What, if anything would you do differently to perform the data mining task?