MIS 584 Assignment: Segmentation and Profiling

**You will complete this assignment with your partner (i.e., groups of two students).**

An Excel file named “Wholesale customers data.xlsx” is uploaded on Canvas. It contains information about the clients of a wholesale distributor and the annual spending of each client in monetary units on different types of products (e.g., milk, grocery, and detergents paper). More information about the dataset are provided here: <http://archive.ics.uci.edu/ml/datasets/Wholesale+customers>

In this assignment, you will perform segmentation and profiling on the dataset.

1. Segmentation: You will first use the data related to spending on the six products (columns c – h) to segment (cluster) the clients. You can simply use Tableau (and Excel if needed) to perform the segmentation process; however, if you are so inclined, you can use other tools and languages such as R or Python to implement a clustering algorithm and answer the questions. If you use Tableau, you can use the automatically determined number of clusters by Tableau (or you can manually change the number of clusters if you have a reason for that).

Then, answer the following questions:

* 1. How many segments are created?
  2. How many clients are in each segment?

1. Profiling: You will next characterize each segment using different attributes. To do so, you will answer the following questions and provide statistical details and screenshots of the corresponding graphs that you use to answer each question:
   1. How are the segments different in terms of the types and volumes of purchased products?
   2. How are the segments different in terms of regions? In other words, is there any significant association between specific regions and specific segments of clients?
   3. How are the segments different in terms of channels? In other words, is there any significant association between specific channels and specific segments of clients?

**What to submit**: Each 2-person group must turn in answers to the questions in the form of a word document submitted to Canvas. Again, for each question, provide screenshots (and statistical details, if applicable) that show that you have done the task properly. If you use any other tools or languages to answer the questions, make sure to provide sufficient explanations on what you have done. For example, if you use R or Python, provide the code as well. Make sure to include your name in your word file. This will be due at 5:00pm November 6th.

**Videos to watch:**

* Visual Analytics
  + Clustering