# INFO 1600 Assignment 2: Data Set and Questions for Dashboard Build

**Due Date: end of Week 5 (Sunday, 11:59pm)**

For this assignment you will apply the decision-making strategies discussed in *The Accidental Analyst* textbook (First 3 Cs). The purpose is first to define a specific problem that you would like to resolve and formulate the questions you need to do so (Choose Your Question). The next step involves identifying the dataset you need to answer those questions, downloading your data (Collect Your Data), and quickly checking the quality in your data (Check Your Data).

This assignment has two sections that will have you engage in this important process and demonstrate that you have done so. For the first section, you will follow the three-step process outlined in *The Accidental Analyst* to choose your questions. The first two steps (Relax, Gather Information) provide good advice for how to proceed. The third step (Select Your Questions) involves a process that you will demonstrate for this assignment. Work through the three stages involved in selecting your questions (Make a List, Refine Your List, Choose Your Best Questions) as defined in the Week 2 PowerPoint and “The First C” chapter in your text.

In the second section, you will download a dataset and begin checking it to prepare for use in your dashboard. You will identify your dataset, download it, inventory it, and integrate it if necessary following the procedures outlined in “The Second C” chapter and Week 3 PowerPoint. You may not need to integrate data from a single download, but it is important to understand when this might be necessary and how to do it. Follow the checking guidelines (for example, sorting and filtering) discussed in the Week 4 PowerPoint and “The Third C” chapter, you will check your dataset for possible errors and missing data.

You will structure your report for this assignment in three sections. The first section will show how you worked through each of the three stages of selecting your questions. The second section will describe where your data is coming from (URL if downloaded from the web, basic description if coming from other source) and what you have done to integrate it (if applicable). The third section will describe the tools you used to quickly check your data for problems and any results these checks returned.

**Submit your report as a Word document to the Assignment 2 dropbox on our eConestoga course site (Course Tools > Assignments > Assignment 2) by the Week 5 deadline**. Your report should be 3 – 4 pages long. Please include your name in the file name.