**Data Science  
In-Class Exercise: Assessing Trustworthiness of Data Sources**

**Objective:** Assess the reliability of a data source.  
  
**Learning Outcomes:**

* Identify potentially unreliable data and information within a data source.
* Differentiate between questionable data and questionable information.
* Differentiate between reliable and unreliable sources.

**Step 1: Explore - individual**

1. Visit these two sites:  
   <http://maps.google.com> and <http://www.yelp.com>
2. Browse each site to get a feel for what kind of data is available from each one.
3. Try out a few searches on each site. For example, search for “Temple University” on both sites and look at what information is displayed. Make sure you read the reviews!  
     
   After you’ve searched for Temple, compare some other places across both sites.
4. As you do your searches, make some notes regarding these questions:
   1. What are ways in which the data and information from each site might not be reliable?
   2. What data and information do you trust from each site? What data and information don’t you trust?
   3. How do you differentiate? Are there “signs” of unreliable data or information?

**Step 2: Prepare – group – breakout session**

1. In groups, compare and discuss your notes.
2. Discuss these items:
   1. What are the ways to spot unreliable data? Unreliable information?
   2. How can unreliable data from these sites lead to unreliable information?
   3. When you see data (or information) you doubt, how would you figure out if it was accurate?

**Step 3: Class Discussion**

Each group will briefly report out with their answers to the questions in Step 2.