## CS/INFO 1305: Computation and Culture in the Digital Age

## Assignment: Introduction to statistical computing AKA understand your Piazza usage and class social network.

## Date: 27th July 2011

**Objective:** The primary objective of this assignment is to understand how to process data using computing technologies. Your role would be to process, analyze and understand the data as it happens (on a larger, much more complex scale) in real life research!

**Dataset:** piazza.xslx is an Excel file and is provided to you.

**Tools:** Anything under the sun! Every question prompt is doable using Excel and/or calculators. If you can do them in Matlab then it's awesome! If you cannot, then it's awesome too!

**Questions:**

1. What is the minimum number of days that a user has spent online?
2. What is the minimum number of posts that a user has viewed when he/she is online?
3. What is the minimum number of posts that a user has commented on when he/she is online?
4. What is the maximum number of days that a user has spent online?
5. What is the maximum number of posts that a user has viewed when he/she is online?
6. What is the maximum number of posts that a user has commented on when he/she is online?
7. What is the average number of days that a user has spent online?
8. What is the average number of posts that a user has viewed when he/she is online?
9. What is the average number of posts that a user has commented on when he/she is online?
10. Which user is the best poster?
11. Which user is the best "lurker"? (i.e. the ratio of views to posts is maximum)
12. Which user has the highest betweenness metric? (use the following formula and directions)

   Take the outputs from questions 7, 8 and 9. Lets call these outputs A, B and C. Now, use them in the formula:

   Betweenness = $(C \wedge 2) + 2A / (3B)$

   **Hint:** For those who would rather use Matlab. You can use loops and functions which you have learnt earlier to complete this very easily.