**Project #1 – Data Analysis Tool**

**Purpose**

The purpose of this project is to get a glimpse of the available tools for data analysis and visualization, as well as the areas that they can be applied. Also, it gives you the opportunity to learn a new tool of your choice from the list below. You are expected after learning the tool to present it in the classroom and explain its capabilities and how it works.

**Tasks and Deliverables**

Task 1

Spend some time investigating the tools below and decide which tool looks interesting to learn. Then find one or more students and form a group to learn this tool (Groups should include 2-3 students only).

**Deliverable 1** – A team representative should provide in a Word document the team members’ names and the name of the tool to be learned.

**Deliverable 1 Deadline** - January 22 using Blackboard (BB)

Task 2

This task includes the presentation of the tool in the classroom to your instructor and classmates. This should be in the form of a tutorial. Expect to answer questions. Prepare this presentation for 15 minutes and 5 minutes for Q/A (*This time may change depending on the number of projects we have for presentation*).

**Deliverable 2** – In class presentation/tutorial. In addition, the team representative should provide this presentation to the instructor. Also, the team representative should provide (using BB) 5 questions to the instructor. These may be used as part of your Final Exam section B. The instructor reserves the right to change these questions.

**Deliverable 2 Deadline** – The class presentation depends on the tool you are working on. These deadlines will be announced after January 22 when all teams have submitted their tool interest.

**Data Analysis Tools**

1. RapidMiner – (Predictive Analytics platform) - <https://rapidminer.com/>
2. Gephi  (Network modeling and calculations) - [www.gephi.org](http://www.gephi.org)
3. NoteXL (Visualization and Analysis of networks and relationships) - <http://nodexl.codeplex.com/>
4. [Tableau](http://tableausoftware.com/)(Data visualization) - <http://www.tableau.com/>
5. R (for statistical computing and basic visualization) - <https://www.r-project.org/>
6. [Python](http://python.org/) (data science libraries) -  <http://www.numpy.org/> and <http://scipy.org/>