**MIS2502: Data Analytics
Optional Extra Credit Assignment**

**The deadline is 11:59 PM on Mar 30th, 2022. No late submissions will be accepted!**

**Overview and Purpose:**

Research a current topic on **Data Analytics** that we have not covered extensively in class. You will create a write-up, and the purpose of this assignment is to give you an outlet to display your ability to understand and describe an aspect of data analytics to your current and future employers.

**Possible topics:**

* Big data
* Distributed data technologies (i.e., Hadoop)
* Natural Language Processing
* Text mining and sentiment analysis
* Virtual data cubes and/or in-memory analytics
* The Internet of Things (IoT)
* Cloud-based analytics/Cloud computing
* NoSQL databases
* Deep learning
* Artificial Intelligence
* Social network analysis/analytics

You don’t have to choose one of these, but if you’re unsure of your topic choice ask me first.

**For all students:** Everyone who successfully completes the assignment will receive **1 extra point in your final grade**. This is the equivalent of about one percent of your total grade, and may be the missing percent that you need to bump you up from e.g. a B-minus to a B.

**For MIS majors**, you will get *50 professional achievement points* when you submit your work through the community website

* + - If you don’t know how to submit your work, please review the following page ('class project' section)
		- <http://community.mis.temple.edu/professionalachievement/earn/>

**Deliverable:**

* A Word or PDF file in Canvas->Assignments>Extra Credit and ProPoints.

**Content:**

Your write-up should be **between 300 and 400 words**, not including references (as a reference, this document is about 500 words long).

You should cover the following points in your writeup:

* **A brief overview** of the topic (i.e., what it is and why it is important).
* **How the topic relates to the material we have covered in MIS2502**. How does it build on the concepts covered in the course? What are the related topics in MIS2502?
* **An example** that describes how this tool or technique has been applied in practice.
* **Citation Guidelines**: If you use materials (text, figures, data, etc.) in the write-up that was created by others, you must identify the source and clearly differentiate your work from the materials that you are referencing. Failure to do so will be considered plagiarizing. There are many different acceptable formats that you can use to cite the work of others. The format is not as important as the intent. You must clearly show the reader what is your work and what is a reference to someone else’s work.

**Evaluation:**

To receive credit, your deliverable must satisfy these criteria:

1. The three areas described in the “Content” section must be covered.
2. The text must be in your own words. Paraphrase your sources; do not quote directly from them: **Plagiarism will be checked through a system. Be aware that the system will compare your submission with online sources, academic papers, books, and previous submissions on Canvas.**
3. You should cite all the sources you use in a professionally formatted bibliography at the end of your write-up.
4. The write-up should look professional and be free of grammatical errors and typos.

**There is no partial credit!** If you don’t meet all of these requirements, you will receive no credit for this assignment.

MIS 2502 – Extra Credit
Tarisha Sarker

Social Network Analysis (SNA) collects data of social relationships between individuals and presents them diagrammatically (Robert A. Hanneman, 2005). SNA aids with understanding not only individuals, but also groups, teams and organizations by mapping out the relationships between them to create a network. A network is generally defined as points that are connected with links, and in this case, the points are individuals/groups, and the links are social connections such as friendships, family, financial ties, etc. SNA serves a multitude of real life purposes such as analyzing follower networks on social media (Twitter, Instagram, etc.), connecting individuals with similar tastes together (Twitter, Pinterest, etc.), interaction networks (phone calls, messages, emails, etc.) (Shazia Tabassum, 2018) and much more.

This relates to the topics we have covered in class because it is a form of data visualization that can be used to extract meaningful information about dynamics between people, especially when dealing with large amounts of data (Big Data). From this convoluted network of multiple people and connections, a schema can be created to have more succinct data organization. These databases can be categorized into Structured Query Language (SQL) relational database structure to extract information from the schema. Another example would be that data queries can be written in SQL, data scientists can create an SNA diagram from those queries.

An example of this is from *Social Network Analysis with Content and Graphs* by William M. Campbell, Charlie K. Dagli, and Clifford J. Weinstein, where they gathered large amounts of data from the Institute for the Study of Violent Groups (ISVG). These data queries were designed in SQL in order to extract information about people from documents. From that, networks of documents and individuals were created to see their association with general society. Similarly, they were able to create SNA diagrams that related to already established violent groups, such as Al Qaeda, to understand community dynamics, create behavioral analyses and potentially predict those behaviors in communities.

**Bibliography:**

Hanneman, Robert A. *Introduction to Social Networks*. University of California, 2005.

Tabassum, Shazia, et al. “Social Network Analysis: An Overview.” *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 5, 2018, doi:10.1002/widm.1256.

Campbell, William M, et al. “Social Network Analysis with Content and Graphs.” *Lincoln Laboratory Journal*, vol. 20, no. 1, 2013.