**Dakota State University**

**INFS 774 Big Data Analytics**

Summer 2020

Guidelines for week 1

Relevant learning objectives:

1. Develop a broad understanding of the terms “Big Data” and “Big Data Analytics”
2. Explore representative example scenarios and case studies of industry specific applications highlighting Big Data issues – volume, variety, velocity, veracity and value
3. Gain an appreciation of the challenges in analyzing Big Data and an overview of current Big Data technologies
4. Review Python

Activities - Required:

1. In week 1’s video, Jun discussed the syllabus. Please watch the video.
2. Read Chapters 1 and 2 from the book “Harness the Power of Big Data” by IBM scientists D. Corrigan, D. deRoos, J. Giles, K. Parasuraman, P. Zikopoulos, and T. Deutsch – (IBM has posted the book online.<ftp://public.dhe.ibm.com/software/pdf/at/SWP10/Harness_the_Power_of_Big_Data.pdf>)
3. Chapter 1. “What is Big Data”, where IBM proposed the famous 5Vs definition of Big Data
4. Chapter 2. “Applying Big Data to Business Problems: A Sampling of Use Cases” that provides a flavor for different applications of Big Data analytics.
5. Watch the following videos that supplement the above readings
6. BSI: Teradata Case of the Tainted Lasagna <https://www.youtube.com/watch?v=-fB3rYtm-tw>
7. BSI: Case of the Retail Turnaround <https://www.youtube.com/watch?v=e36v7hzfvjs>
8. If you are not very familiar with python and iPython notebook, you need to finish the python tutorial videos posted under “Content” -> “Week1”. The notebook used in the tutorial is “Python+Tutorial.ipynb”.

Activities – Optional:

1. Please finish the numpy tutorial and the pandas tutorial posted under “Content” -> “Week 1”.