Week 3, Session 1

**Sharda 7.1 – 7.4, Articles: 1) Big Data: Concepts, Technologies and Applications , 2) When it's time to Hadoop, 3) Big data applications in clinical medicine**

1. Provide one or more examples of business problems/opportunities that could be addressed through big data analytics.
2. Research a cloud-base service provider of big data and/or analytics. Provide a summary and a critical evaluation.
3. Research a NoSQL database. Provide a summary and a critical evaluation.
4. What is Hadoop?  How does it work?
5. What is MapReduce? How does it work? Discuss one or more examples of business applications that could benefit from MapReduce’s data processing framework.
6. Compare and contrast end users, business analysts, BI analysts, and data scientists.
7. Pick a vendor of big data technologies, and discuss their product/service offerings.
8. According to the experts, what are the distinct advantages/capabilities of Hadoop and data warehousing? Do you agree? Why or why not?
9. "The more she wants Hadoop to provide database capabilities, the more the data warehouse is the answer."  Discuss.
10. Does secondary analysis of electronic health records replace, complement or supplement randomized clinical trials?  Discuss.
11. Visit [https://mimic.physionet.org/gettingstarted/access/ (Links to an external site.)Links to an external site.](https://mimic.physionet.org/gettingstarted/access/), and review the tables and data details (i.e, the metadata). Answer any one of the following questions:
    * Comment on the breadth and depth of data available for analysis.
    * List two queries that you would like to get answers for from this database.  Why are these queries relevant or interesting to you?
    * List two analytical tasks that could be performed with this database?  Why are these tasks relevant or interesting to you?