Quiz 1 – Why Big Data and Where Did it Come From?

1. Which of the following is an example of big data utilized in action today?

- The Internet
- Wi-Fi Networks
- **Social Media**
- Individual, Unconnected Hospital Databases

2. What reasoning was given for the following: why is the "data storage to price ratio" relevant to big data?

- Companies can't afford to own, maintain, and spend the energy to support large data storage unless the cost is sufficiently low.
- Larger storage means easier accessibility to big data for every user because it allows users to download in bulk.
- It isn't, it was just an arbitrary example on big data usage.
- **Access of larger storage becomes easier for everyone, which means client-facing services require very large data storage.**

3. What is the best description of personalized marketing enabled by big data?

- Marketing to each customer on an individual level and suiting to their needs.
- **Being able to use the data from each customer for marketing needs.**
- Being able to obtain and use customer information for specific groups and utilize them for marketing needs.

4. Of the following, which are some examples of personalized marketing related to big data?

- A survey that asks your age and markets to you a specific brand.
- **Facebook revealing posts that cater towards similar interests.**
- News outlets gathering information from the internet in order to report them to the public.

5. What is the workflow for working with big data?

- Extrapolation -> Understanding -> Reproducing
- **Big Data -> Better Models -> Higher Precision**
- Theory -> Models -> Precise Advice

6. Which is the most compelling reason why mobile advertising is related to big data?

- Mobile advertising in and of itself is always associated with big data.
- Since almost everyone owns a cell/mobile phone, the mobile advertising market is large and thus requires big data to contain all the information.
- **Mobile advertising benefits from data integration with location which requires big data.**
- Mobile advertising allows massive cellular/mobile texting to a wide audience, thus providing large amounts of data.

7. What are the three types of diverse data sources?

- Information Networks, Map Data, and People
- Machine Data, Map Data, and Social Media
- **Machine Data, Organizational Data, and People**
- Sensor Data, Organizational Data, and Social Media

8. What is an example of machine data?

- Sorted data from Amazon regarding customer info.
- **Weather station sensor output.**
- Social Media

9. What is an example of organizational data?

- Satellite Data
- Social Media
- **Disease data from Center for Disease Control.**

10. Of the three data sources, which is the hardest to implement and streamline into a model?

- **People**
- Machine Data
- Organizational Data

11. Which of the following summarizes the process of using data streams?

- **Integration -> Personalization -> Precision**
- Big Data -> Better Models -> Higher Precision
- Theory -> Models -> Precise Advice
- Extrapolation -> Understanding -> Reproducing

12. Where does the real value of big data often come from?

- Having data-enabled decisions and actions from the insights of new data.
- **Combining streams of data and analyzing them for new insights.**
- Size of the data.
- Using the three major data sources: Machines, People, and Organizations.

13. What does it mean for a device to be "smart"?

- **Connect with other devices and have knowledge of the environment.**
- Having a specific processing speed in order to keep up with the demands of data processing.
- Must have a way to interact with the user.

14. What does the term "in situ" mean in the context of big data?

- Accelerometers.
- The sensors used in airplanes to measure altitude.
- **Bringing the computation to the location of the data.**
- In the situation

15. Which of the following are reasons mentioned for why data generated by people are hard to process?

- They cannot be modeled and stored.
- **The velocity of the data is very high.**
- **Very unstructured data.**
- **Skilled people to analyze the data are hard to come by.**

16. What is the purpose of retrieval and storage; pre-processing; and analysis in order to convert multiple data sources into valuable data?

- To enable ETL methods.
- Since the multi-layered process is built into the Neo4j database connection.
- Designed to work like the ETL process.
- **To allow scalable analytical solutions to big data.**

17. Which of the following are benefits for organization generated data?

- **Improved Safety**
- **Better Profit Margins**
- **Higher Sales**
- High Velocity
- **Customer Satisfaction**

18. What are data silos and why are they bad?

- **Data produced from an organization that is spread out. Bad because it creates unsynchronized and invisible data.**
- A giant centralized database to house all the data production within an organization. Bad because it hinders opportunity for data generation.
- Highly unstructured data. Bad because it does not provide meaningful results for organizations.
- A giant centralized database to house all the data produces within an organization. Bad because it is hard to maintain as highly structured data.

19. Which of the following are benefits of data integration?

- **Reduce data complexity.**
- **Adds value to big data.**
- Monitoring of data.
- **Increase data collaboration.**
- **Increase data availability.**
- **Unify your data system.**