

## Quiz 2 – V for the V's of Big Data

1. Amazon has been collecting review data for a particular product. They have realized that almost 90% of the reviews were mostly a 5/5 rating. However, of the 90%, they realized that 50% of them were customers who did not have proof of purchase or customers who did not post serious reviews about the product. Of the following, which is true about the review data collected in this situation?

- High Valence
- Low Valence
- Low Volume
- High Veracity
- High Volume
- **Low Veracity**

2. As mentioned in the slides, what are the challenges to data with a high valence?

- **Complex Data Exploration Algorithms**
- Difficult to Integrate
- Reliability of Data

3. Which of the following are the 6 V's in big data?

- **Variety**
- **Volume**
- **Valence**
- **Value**
- **Veracity**
- **Velocity**
- Vision

#### 4. What is the veracity of big data?

- The connectedness of data.
- The speed at which data is produced.
- **The abnormality or uncertainties of data.**
- The size of the data.

#### 5. What are the challenges of data with high variety?

- Hard to perform emergent behavior analysis.
- The quality of data is low.
- Hard in utilizing group event detection.
- **Hard to integrate.**

#### 6. Which of the following is the best way to describe why it is crucial to process data in real-time?

- Batch processing is an older method that is not as accurate as real-time processing.
- More accurate.
- **Prevents missed opportunities.**
- More expensive to batch process.

#### 7. What are the challenges with big data that has high volume?

- Storage and Accessibility
- Speed Increase in Processing
- **Cost, Scalability, and Performance**
- Effectiveness and Cost