

# PRACTICE SET 3

## 1. Point out the wrong statement.

- a) The big volume indeed represents Big Data
- b) The data growth and social media explosion have changed how we look at the data
- c) Big Data is just about lots of data
- d) All of the mentioned

Answer: c

Explanation: Big Data is actually a concept providing an opportunity to find new insight into your existing data as well guidelines to capture and analysis your future data.

## 2. Which of the following step is performed by data scientist after acquiring the data?

- a) Data Cleansing
- b) Data Integration
- c) Data Replication
- d) All of the mentioned

Answer: a

Explanation: Data cleansing, data cleaning or data scrubbing is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database.

**3. 3V's are not sufficient to describe big data.**

- a) True
- b) False

Answer: a

Explanation: IBM data scientists break big data into four dimensions: volume, variety, velocity and veracity.

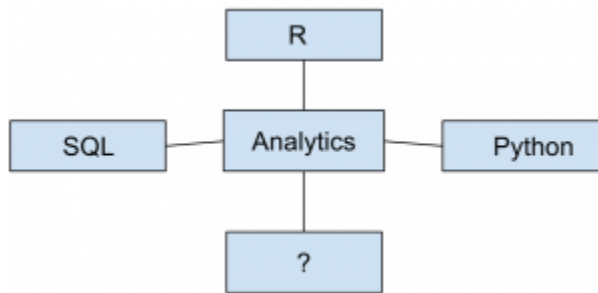
**4. Which of the following focuses on the discovery of (previously) unknown properties on the data?**

- a) Data mining
- b) Big Data
- c) Data wrangling
- d) Machine Learning

Answer: a

Explanation: Data munging or data wrangling is loosely the process of manually converting or mapping data from one "raw" form into another format that allows for more convenient consumption of the data with the help of semi-automated tools.

**5. Which of the following language should be replaced with the question mark in the below figure?**



- a) Java
- b) PHP
- c) COBOL
- d) None of the mentioned

Answer: a


Explanation: Java is used for processing data in Big data Analytics.

**6. Beyond Volume, variety and velocity are the issues of big data veracity.**

- a) True
- b) False

Answer: a

Explanation: Data Veracity is uncertain or imprecise data.



**END OF  
PRACTICE SET  
3**