01 Which of the following statements is true regarding the differences between Generation 3 of Data Curation approaches w.r.t. previous generations:
o Generation 3 does not address any of the Dara Curation tasks faced in previous generations. Instead, it only focuses addressing scalability and automation issues.
X One of the goals of Generation 3 approaches is to help in the automation of some of the data curation tasks, specially those that can be easily tackled with the use of machine learning.
o Generation 3 does not take into account data streaming. In fact, Generation 2 has fully addressed and solved this problem.
o None of the above

02 Which of the following is an example scenario of a low data quality problem if we take the Believability dimension into account:
o The data needed to run our big data project is hard to access and bits of data are always missing
o The records are represented in formats that are not supported by our big data algorithms
o The sources providing data to our big data project cannot be relied on and content are of low importance
X We are not fully sure if data collected by our system truly represent the real-world facts we are interested in.

03 Which of the statements below is false about Data Curation:
X Data Curation is a one-off task that happens early in the big data process.
o Data Curation typically deals not only with removing/fixing errors from data but also with other tasks such as data enrichment.
o In the context of Big Data, Data Curation needs to take scalability issues into account
o All of the above

04 Data Curation is a technical/engineering task and therefore needs to be planned, designed and executed without taking into account the business requirements and context:
o True
X False

05 Which of the following statements is true about the data quality dimension Value-added?
o Value-added is a data quality dimension that helps in assessing whether we can add more value to data collected from various sources
o Value-added refers to whether data has been enriched at the source (e.g., using annotations)
o Value-added refer to the extent to which more data has been added to the original data collected from the source
X Value-added refers to how much value we can get in return from the usage of the data