

#### Link 8.3 Answers to Cloze Exercise 8.1

Dirty data are data that are of poor QUALITY and may lead to poor business decisions. Dirty data are data that are of may be inaccurate, incomplete, INCONSISTENT, and violate the business rules of data INTEGRITY. Causes of dirty data include ACCIDENTAL or malicious acts, incomplete data fields, and staff working around existing quality controls. A data value is INVALID if the value is not within the value range defined for the attribute. This can be prevented by using VALIDATION checks when data are captured. A data value can be intrinsically valid but INACCURATE due to dependencies between attributes. This can be prevented by using dependency checks. Cleaning data involves identifying and correcting problems, BACKFLUSHING the corrected data and implementing PROCEDURES to ensure the data stay clean.

Data FRAGMENTATION occurs when an entity is physically stored in different locations and may be used to improve data CONFIDENTIALITY. The division of entities into subsets of instances is referred to as HORIZONTAL fragmentation. VERTICAL fragmentation is based on subsets of attributes processed by different transactions. Data in different information systems may use the same attribute name to mean different things or different attribute names to describe the same thing, requiring data RECONCILIATION. This involves standardizing attribute names, attribute structures, units of measure and CODING systems. Two reports may present CONFLICTING information due to how the reports were generated and how the DATA used in the reports were generated. For example, different data QUERIES may have been used to gather the data reported.

Data may be CORRUPTED accidentally or deliberately due to natural, human or manmade actions. For example, data may be lost during a flood, by a computer virus or an electrical fault. Electronic TRANSACTIONS create a digital FOOTPRINT that can be used in computer forensics to investigate cybercrime. TECHNOLOGY can be used to perform or plan a criminal act, or may be the target of the crime. When an incident occurs the situation must be contained to preserve the evidence. Information maps and the information ARCHITECTURE can be used to identify the information that may have been affected by an incident. Data RECOVERY involves identifying the state of the data before the incident occurred. Information MANAGEMENT aims to maintain the quality of information as the organization and its requirements for information change.