

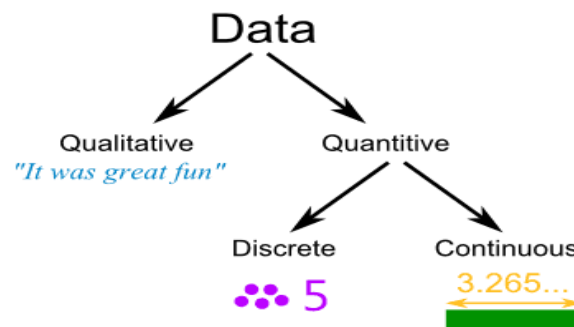
LO1. Monitor data conversion

1.1. Defining concepts of data conversion and Data Terminologies

- **Data** is raw facts or unorganized things (such as alphabets, numbers, or symbols) that refers to, or represent, conditions, ideas, or objects.

It can be qualitative or quantitative.

- **Qualitative data** is descriptive information (it *describes* something)
- **Quantitative data** is numerical information (numbers).



- **Discrete data** can only take certain values (like whole numbers)
- **Continuous data** can take any value (within a range)

Put simply: **Discrete data** can be counted, **Continuous data** can be measured

Example:



Qualitative:

- It is brown and black
- It has long hair
- It has lots of energy

Quantitative:

- Discrete:
 - It has 4 legs
 - It has 10 fingers
- Continuous:
 - It weighs 25.5 kg
 - It is 565 mm tall

- **Data conversion** is the conversion of one file or database from one format (from one physical environment) to another.

Often, when data is moved from one system to another, some form of data conversion is required to convert the data to a format the receiving system can interpret.

Types of conversion:

- Database conversion (SQL, MySQL, MS Access, XLS, XML etc)
- File format conversion (PDF to Word)
- Image conversion (GIF to JPG, TIFF, PNG etc)
- Character or string conversion(numeric to alphabet or viceversa)

1.2. Reading and Analyzing Existing Data Conversion Documents

The data conversion process can often be a complex and difficult task during an implementation. When performing data conversions, you must include analysis of your source data and continues through to system testing and user acceptance.

Throughout the conversion process, we perform quality control checks to ensure correctness of the conversion.

1.3. Understanding Data and Its Characteristics

1.3.1. Data Conversion Systems and Tools

Data Conversion Tool allows you to convert data both from and to (both sides are supported) a wide variety of formats, including:

- SQLServer Tables
- Oracle Tables
- ODBC Tables
- OleDb Tables
- Microsoft Access Tables
- XML Files

Once a conversion type is defined, it can be saved and reused either in a future conversion or as a step within a batch conversion.

1.3.2. Data Modeling Methodologies

Data modeling is the formalization and documentation of existing processes and events that occur during application software design and development.

Data modeling techniques and tools capture and translate complex system designs into easily understood representations of the data flows and processes, creating a blueprint for construction or re-engineering.

A data model can be thought of as a diagram or flowchart that illustrates the relationships between data.

There are several different approaches of data modeling, including:

- **Conceptual Data Modeling** - identifies the highest-level relationships between different entities.
- **Logical Data Modeling** - illustrates the specific entities, attributes and relationships involved in a business function.
- **Physical Data Modeling** - represents an application and database-specific implementation of a logical data model.

1.3.3. Data Conditioning and cleaning

Data conditioning (Pre-processing) is the use of data management and optimization techniques which result in the intelligent routing, optimization and protection of data for storage or data movement in a computer system.

Data cleaning is the act of detecting and removing or correcting dirty data (i.e.: data that is incorrect, out-of-date, redundant, incomplete, or formatted incorrectly).

Data Cleaning helps to increase the overall efficiency of your data management systems and leads to an increase in the productivity of the organization.

1.3.4. Data Transformation and integration

Data transformation is one of the collective processes known as extract, transform or load which is one of the most important processes in data warehouse implementation from different data sources.

Data Integration is the process of combining heterogenous data sources in to a single queriable schema so as to get a unified view of these data.

1.3.5. Sorting, updating, exporting and convert data

Sorting data

Sorting data is the process of arranging items into meaningful order so that you can analyze it more effectively.

Example:

- sort text data into alphabetical order
- sort numeric data into numerical order

Updating Data

The modification of data that is already in the database is referred to as updating. The update operation allows you to change an existing database record in a logical or physical file. You can update individual rows, all the rows in a table. Each column can be updated separately without affecting other columns.

```
UPDATE table_name
SET column1=value, column2=value2, ...
WHERE some_column=some_value
```

To perform an update, you need three pieces of information:

1. The name of the table and column to update,
2. The new value of the column,
3. Which row(s) to update?

Exporting data

You can export data from one application to another application using the Export Wizard.

Exporting lets you share data from one application by providing a copy of data.

1.3.6. Ensuring Data Quality

DBMS provides a set of features that enable you to ensure the quality of data that is moved from source systems to your data destination. Data profiling is a feature that enables you to analyze the content and structure of your data to determine inconsistencies, anomalies, and redundancies in the data.

1.4. Collection, Organization and Analysis of Data and Information

The data is typically organized to model relevant aspects of reality (for example, the availability of rooms in hotels), in a way that supports processes requiring the information (for example, finding a hotel with vacancies).

Organization of data

Organization of data is any one of the data management conventions for physical and spatial arrangement of the physical records of a data set.

Analysis of data

Analysis of data is the process of evaluating data using analytical and logical reasoning to examine each component of the data provided. Data from various sources is gathered, reviewed, and then analyzed to form some sort of finding or conclusion.

Information

Information is processed data that can affect behaviour, a decision or outcome.

It is valuable that is:

- accurate and timely
- specific and organized for a purpose
- presented within a context that gives it meaning and relevance
- Can lead to an increase in understanding and decrease in uncertainty.

1.5. Validating Data Conversion Systems

1.5.1. Data Accuracy

Data accuracy is generally expressed as a confidence interval (CI). This means more of the information collected can be trusted as valid and free of confounding variables.

1.5.2. Data Integrity

Data integrity (also known as data validity) refers to the overall completeness, accuracy and consistency of data. This can be indicated by the absence of alteration (unchanged) between two instances or between two updates of a data record.

Data integrity can be maintained through the use of various error checking methods and validation procedures.

1.5.3. Back-up before Conversion

Before and after the process of converting your data, it is strongly recommended that you perform a full data backup.

1.6. Identifying and Confirming Data Conversion Tools

- Software
- Hardware
- Environmental Pre-Requisites (Dust, Heat, Extreme Cold, Temperature Stability, Air Circulation and Moisture)

Dust and Dirty environment increases the overheating problem and mostly affects

- The Motherboard.
- The Processor and Power Supply fan.
- The CD-drive's Lens and Floppy drive's head.
- The Add-in card connection.
- The cable connection.
- The mouse and keyboard.

LO2. Support data conversion

- Results must be verified based on the relevant checklist.
- Verified data must be presented and approved by appropriate persons.
- Back-up copies of conversion files must be maintained and documented according to requirements
- Developing clear and coherent technical documentation