

## Characteristics of Quality Data

People depend on the systems they use to contain high quality data. If they find the data to be wrong, outdated or incomplete, they begin to distrust the system and will likely stop using it. If the data in the system is personally important to the individual, such as the data in payroll or medical systems, then there is a strong need to have it corrected as quickly as possible.

What are the characteristics of good quality data? There are a variety of characteristics, but we will focus on the six listed here. Let's look at the data that may be in a payroll system and how each of the characteristics of quality data are important, and consider an example of each.

- Accuracy – Is the information correct? For example, is the annual salary correct?
- Completeness – Is all the information there? For example, if overtime was worked, is it included?
- Timeliness – Does the payroll data pertain to the current pay period? For example, is old, outdated data used which could change the amount paid to the employee?
- Uniqueness – Does the payroll record apply to a specific individual? For example, does a specific payroll record pertain to a specific employee?
- Validity – Is the data in the payroll record based on the accepted ranges? For example, is the hourly rate within the accepted range and is the number of hours worked reasonable?
- Consistency – Is the data in the payroll record correctly aligned with other data in the system, such as employee status (e.g., hourly or salaried worker)? For example, if the employee is a salaried employee does the payroll record reflect that vs. hourly employee data?

It is important that the data in information systems is of high quality. As systems are developed, the testing should include ensuring that the quality of data is maintained throughout the system, from its source to the final output. Therefore, data needs to have these characteristics when it is entered into the system. The data entry process should include validation that it meets these quality attributes, and then it needs to be protected as it resides in and flows through the system. If any of these characteristics are missing, the system must be analyzed to discover where the problem lies. The correction may be as simple as fixing an individual record, or, if it is not clear where the problem lies, the system may be considered unreliable overall and may need to be taken offline until corrections are applied.