Levels of Data Quality

1. **Inaccurate data** (also called invalid data): Involves data that are incorrect or not reflecting what they're meant to reflect.

   Examples:_____
   _____
   _____
   _____

2. **Unreliable data**: Data that are not recorded in the same way every time. Perhaps one staff member is not recording data correctly, even though all of the other staff members are doing so. However, because of that one staff member, the entire dataset might not be reliable.

   Examples:_____
   _____
   _____
   _____

3. **Incomplete data**: May result from not recording every student, service, or outcome as needed (or required).

   Examples:_____
   _____
   _____
   _____

4. **Imprecise data**: Might involve providing data that are not detailed enough.

   Examples:_____
   _____
   _____
   _____

5. **Untimely data**: Data are not available in time for program needs or reporting requirements. Timeliness of data is an important aspect of data for decision making, since the data have to be available for the decisions to be made in a timely manner.

    Examples:_____
    _____
    _____
    _____

6. **Data integrity**: Data must have integrity and must be protected from manipulation or deliberate bias.

    Examples:_____
    _____
    _____
    _____

Types of Data Errors

1. **Transposition errors:** An example of a transposition error is when the correct number is 1998, but you write 1989 by mistake.

2. **Missing data:** This type of error may include data that are truly missing, a sum that has not been appropriately disaggregated, or a "false" zero (i.e., when a zero is recorded to signify a true zero when in fact the data are missing).

3. **Copy errors:** These include writing that is not easy to read and extra zeros that should not be there (e.g., "10" written as "100").

Steps in Data Cleaning

Data cleaning is a process that starts during data collection and verification. In general, data cleaning is a two-step process: identifying potential problems, and correcting or verifying any potential problems. To identify problems, you need to look for:

1. Transposition errors, missing data, and copy errors

2. Numbers that are out of the range of what is normal

3. An incorrect sum

4. Logical impossibilities