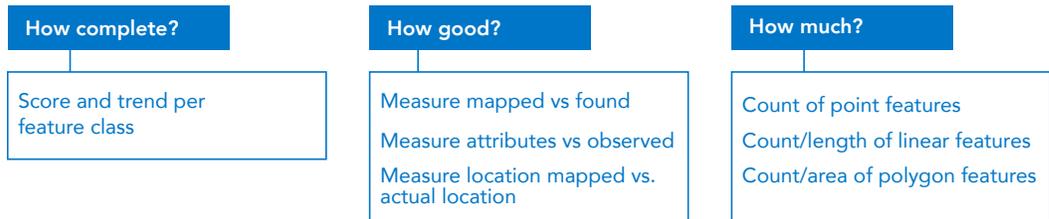


How do you measure and track GIS data quality and completeness?

It's a challenging question for any organization: "How confident are you regarding the quality and completeness of your GIS data?" Answers vary widely and include process, source and usage, but it's an essential question. Efficiency and productivity depend on good and thorough data documentation.

Measuring and tracking GIS data quality and completeness is a complex process, but Open Spatial brings you the rundown in two straightforward steps: First, record and measure your data. Second, look at the overall health of your GIS data. How good is it? How much do you have? And ultimately, is the data quality improving? These aspects combine to indicate the overall condition of GIS data.

Single Measure of Data Health



Record and Measure

The first step in tracking data quality and completeness is to record it! Once data is recorded, you will be able to measure it. Begin by simply recording your data on a consistent basis, and don't worry about having an "exact number" to record. The benefit in recording data regularly is that it allows you to track and show relative change over time based on the same measuring technique.

Thus, measuring your data is important. Open Spatial's geospatial suite allows setup of procedures to run summary statistics on the values you think are important for each feature. For example, what percentage of your water pipes have information on diameter, material, owner and installation date? Store this summary information as a quality measure and date it in a table. When running the summary procedure again, say next month or next quarter, you will have a second set of measures and can compare if water pipe data has improved or deteriorated.

You can also use an average value as a single indicator for the feature at the time of measure. For example, the average completeness of water pipe data (average value of the three items measured) may have changed from 55 percent to 62 percent. Applying this approach for all features and all datasets will give you an overview of your data completeness.

Overall Health

How good is your data? GIS data quality covers both spatial and attribute accuracy. A simple way to improve data quality is to include a column to record the data source and a column for data quality on each feature. Recording the data source is important for reference purposes and gives some indication of spatial accuracy. Ultimately, spatial accuracy is best assessed by comparing the location as recorded in the GIS with an accurate field measurement such as a survey or high-resolution Global Positioning System. When your GIS data management gets to this level, doing a series of sample tests on different datasets and recording the actual value versus the recorded value creates a new level of confidence and measures the spatial accuracy of your data.

Attribute accuracy is involved, but it's manageable to record one measure of quality for each feature. The American Society of Civil Engineers developed a National Consensus Standard, ASCE CI 38-02, which is also an American National Standard Institute standard used by Subsurface Utility Engineering (SUE) consultants. Commonly referred to as the SUE standard, this provides a simple four-point measure of quality for subsurface utility data, which is often used in a wider context. Alternatively, you can use your own descriptive measure and assign each record a data quality value. For example:

SUE data quality measures

Quality Level D – Based on record info or recollections

Quality Level C – Based on field verified surface features and or aerials

Quality Level B – Electronically designated and field surveyed

Quality Level A – Visually verified, surveyed and documented

Simple point descriptive measure of data quality

01-UNVERIFIED

02-LOCAL KNOWLEDGE

02-SKETCH

04-DESIGN DRAWING

05-TENANT/CONSULTANT DRAWING

06-RECORD DRAWING

07-AS-BUILT DRAWING

08-SURVEY

09-SURVEY GRADE GPS

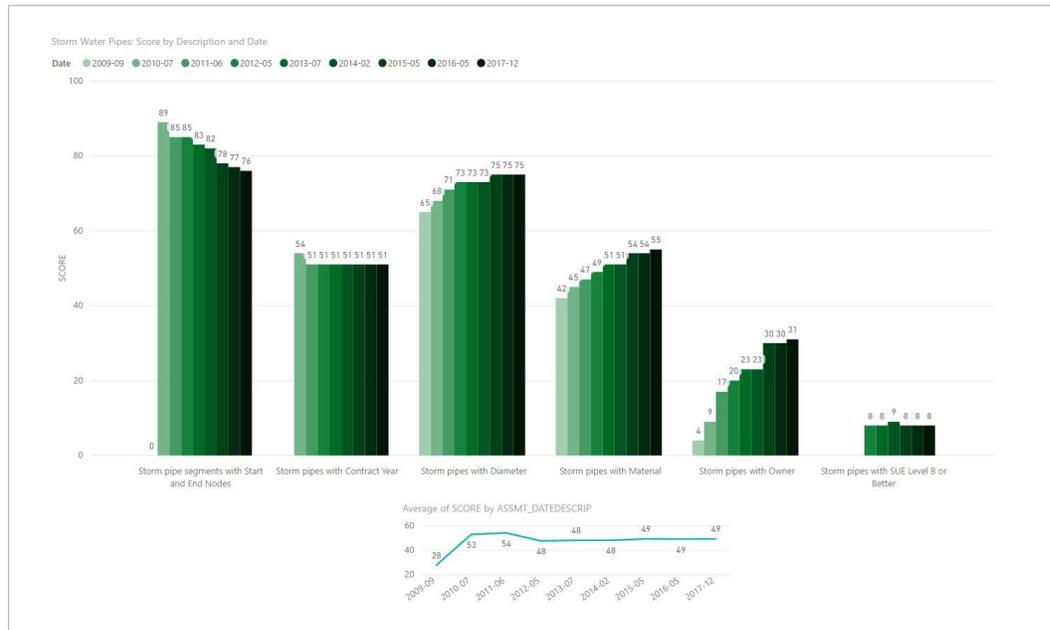
Quality Improvement

In a similar way, how much data you have can be measured. How much is recorded as a simple count and/or total length and area summary totals per feature. You can make the overall measures more useful by including a weighting factor for each measure. The amount of data you have is used as an indicator of overall progress and change. It also helps to interpret trends in data quality as sometimes a new batch of data may not be as complete as existing data, and may bring down measures of quality.

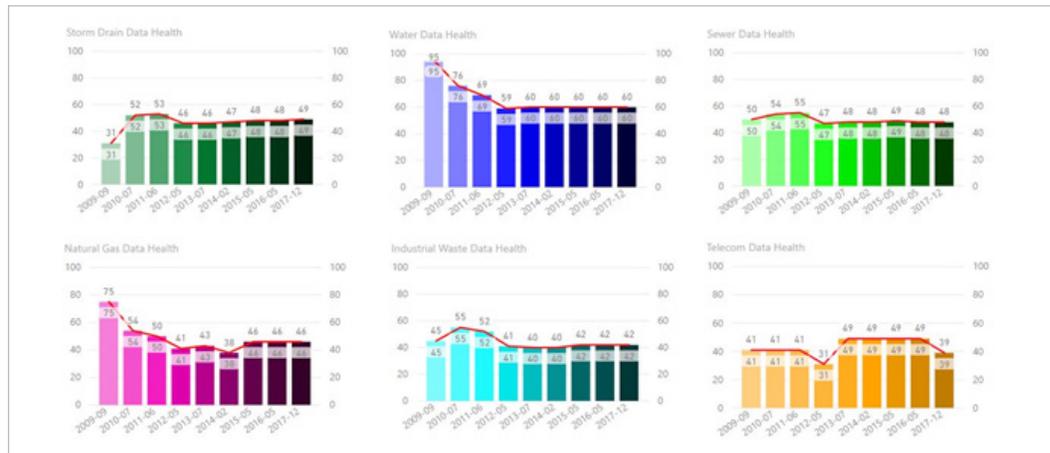
For instance, the graphic below shows relative measures of overall GIS data quality and data health for an organization over 10 years. As new data was added, it was not as complete and caused the overall data quality health to decline slightly.

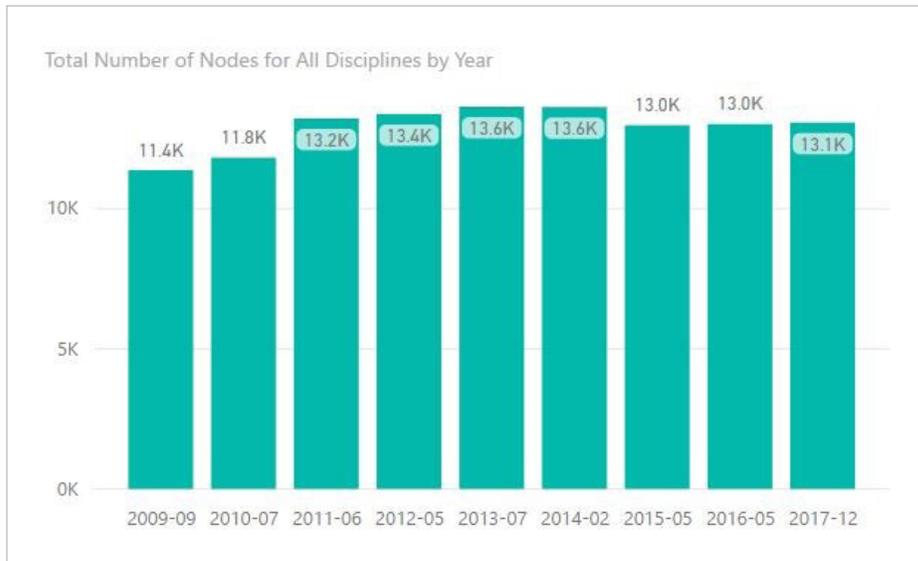
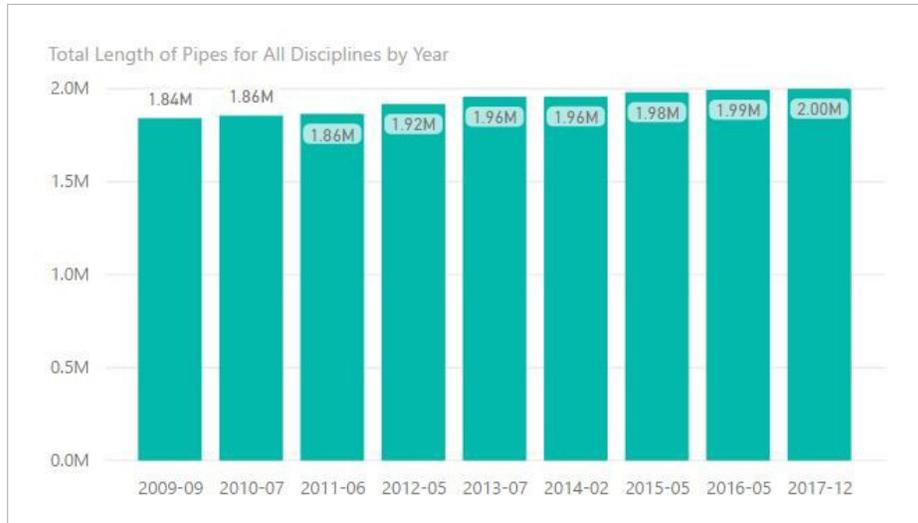
Performing the summary statistical analysis on a regular basis and storing the values for each assessment allows you to track data quality over time and have a simple overview of changes in quality and a relative measure of overall data health.

Single feature class data completeness measures over 10 years



Overall score and trend per utility





Trend of data health as a single measure for all utilities.

