

Connecting Content | Driving Performance

Data Quality Assessment Overview

Data supports sophisticated analyses and decision making processes to makes companies more profitable and improve their products and services. The true value of data depends on data quality. The Dayhuff Group provides Business Intelligence (BI) solution leadership and implementation to its clients with an on-site Data Quality Assessment. The purpose of the baseline assessment is to determine the calculated quality of approximately 25-30 data fields that are of key importance to the business. This assessment will play a critical part and:

- ▶ Facilitate the identification of 25-30 data fields of significance
- ▶ Calculate the random sample size for the assessment
- ▶ Specify the stratified random sampling technique
- ▶ Write specifications for the data sample extraction
- ▶ Calculate the quality of each of the identified data fields
- ▶ Calculate multiple data field consistency

Our approach in assessment engagements is to facilitate and guide our client through a rapid, repeatable and comprehensive survey of the business need and technical solution. This cross collaboration prepares our greater team of client business users and technical developers along with Dayhuff Group BI consultants for a successful BI solution.

Approach

Random Sampling

The Dayhuff Group will engage these sampling techniques:

- ▶ Use a stratified random sample
- ▶ Stratified random sampling techniques are generally used when the population is heterogeneous, or dissimilar, where certain homogeneous, or similar, sub-populations can be isolated (strata).
- ▶ Simple random sampling is most appropriate when the entire population from which the sample is taken is homogeneous.
- ▶ A stratified random sample is obtained by taking random samples from each stratum or sub-group of a population
- ▶ When sampling a population with several strata, it is generally required that the proportion of each stratum in the sample should be representative of the population

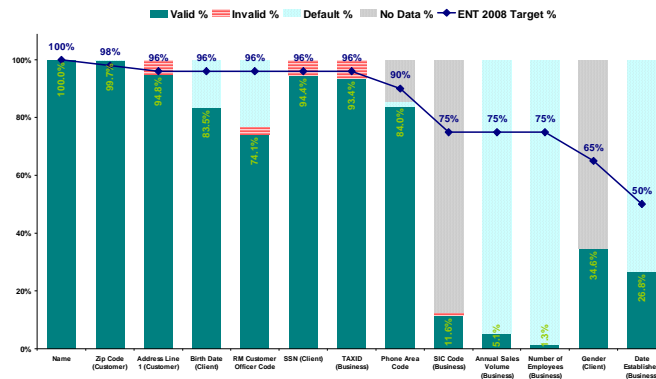
Measurement Types

The Dayhuff Group will generate statistics for the 25-30 data fields. Specifically, five measurements will be calculated:

- ▶ Population – No values, default values, non-default values
- ▶ Validity – Adherence to specific rules or reference codes already defined.
- ▶ Consistency – Comparison of specific fields to other related fields
- ▶ Completeness – A calculation of the percent of records with all 13-22 customer fields populated
- ▶ Duplication – Rate of record redundancy using definitions supplied by the business areas.

Presentation

Each data field will have a data quality chart representing the current state and desired state of quality. The charts will look similar to the one below:



- ▶ Charts show normalized measures, 0%-100%
- ▶ Measures include Validity, Completeness and Consistency
- ▶ Contain Compliance UCL, Experienced-Based LCL, Business-Defined Improvement Target and Actual
- ▶ Charts can display a single data element, composite data elements, compliance groups or Business-defined groups
- ▶ Charts can show repeated samples or new customers or policies tracked over time

Assessment Deliverables

Data Quality Assessment Findings – This document describes quality issues identified by business and IT users. An important part is determining the business impact of poor data quality. A comparison of user perception against mathematical metrics is provided. Data fields are compared against 4 metrics; population, validity, consistency and completeness. The detail includes stratification of these four metrics across as many as ten strata.

Executive Summary – This document, appropriate for executive sponsors, highlights the significant finds of the data quality assessment.

Statistical Sampling Technique – Specific attention is given to examples and specifically how the sample is derived.

Assessment Services

Assessments start at \$60,000 for an approximately 5 week engagement.

Contact Information

The Dayhuff Group
 740 Lakeview Plaza Boulevard, Suite 300
 Worthington, OH 43085
 888-DAYHUFF or 614 854-9999 ext 113
www.dayhuffgroup.com/contact