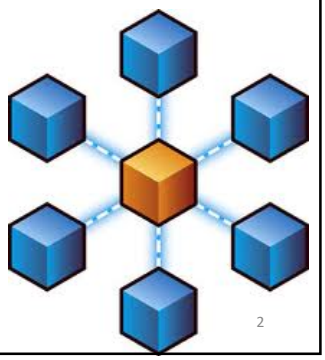
Assessment and Analysis of Test Data

Presented by SARF

Presenter:
Ron Berkers

1

Module C
Data Management

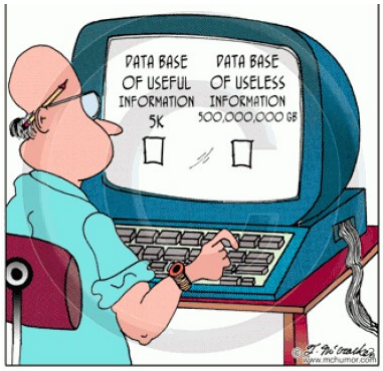


2

Data

Data:

- **Data means a set of quantities, figures and statistics, relating to any event or product.**
- **A set of data can be either numerical or non-numerical.**
- **Numeric data are numbers (like age, cost, mass, etc.).**
- **Non-numeric data are not real numbers (like name, address, etc.).**



Data Collection

Data Collection:

- **The process of gathering and measuring information on variables of interest.**
- **Data collection is done in an established, systematic fashion that enables one to answer routine data management questions, test hypotheses, and evaluate outcomes.**

Data Collection

Data Integrity:

- **Accurate data collection is essential to maintaining the integrity of data quality.**
- **Both the selection of appropriate data collection methods and clearly delineated instructions for their correct use reduce the likelihood of errors occurring.**

5

Data Collection

Data Quality Assessment:

- **Data quality assessment is a multi-step procedure for determining according to prescribed procedures whether or not a data set is suitable for its intended purpose.**
- **The assessment is a statistical evaluation of data to determine if it is of the type, quantity, and quality needed.**
- **Data quality assurance and quality control are two approaches that can preserve data integrity and ensure the scientific validity of the investigation or control results.**
- **Quality assurance activities take place before data collection begins.**
- **Quality control activities take place during and after data collection.**

6

Data Collection

Data Quality Assurance:

- The main focus of data quality assurance is prevention.
- Data quality assurance tries to forestall problems with data collection.
- An important component of quality assurance is developing a rigorous and detailed training plan.

7

Data Collection

Data Quality Control:

- Quality control activities such as the monitoring, detection and actions) occur during and after data collection.
- A clearly defined communication structure is a necessary pre-condition for establishing monitoring systems.
- While site visits may not be appropriate in all applications, failure to perform regular audit records, whether quantitative or qualitative, will make it difficult for investigators to verify that data collection is proceeding according to standard procedures.

Examples of data collection problems that require prompt action include:

- Errors in individual data items.
- Systematic errors.
- Violation of protocol.
- Problems with individual staff or site performance.
- Fraud or scientific misconduct.

8

Data Collection

Data reviewing prior to analysis:

Steps:

- Thorough review of the results (measurements).
- Editing (only in the case of blunders or transcription errors).
- Analysis.
- Reporting

Common mistakes:

- Simplification: arithmetic mean is inside specifications while 50% of test results can be below specified minimum.
- Alteration of data: attempts can be made to change data to compensate for poor quality control and/or experimental design.
- Projection of data: test range or region is not adequately represented by the results.

9

Data Collection

Interpretation:

Data interpretation is the ability to understand and to extract maximum information from a quantity of data that is being determined by measurement and presented in various forms e.g. a single value, groups of data, tables, histograms, charts etc.

Requirements:

- The data collector must understand the data as thoroughly as possible in terms of:
  - ✓the purpose for which the data source was selected or designed to address.
  - ✓what, if any, surprises the data may have produced, and
  - ✓what, if anything, such surprises may imply
- The data must be reported to the client in a manner that is as clear, complete, and intuitively compelling as possible.

10

Preparation of data

- Data should be arranged and presented in a systematic or organized format, so that some useful deductions for the population under consideration can be drawn from the data set.
- This is important as it is usual practice to base decisions on the quality of a population or a lot on the outcomes of tests on relatively small numbers of samples of the lot or population.

11

Data Quality

There is no such thing as a Perfect Measurement

Measurement Category:

- **Direct measurement:** measurements that can actually be taken with gauges, instruments or by counting.

  For example: height of a person (measured with a ruler)

- **Indirect measurement:** a technique that uses proportions to find a measurement when direct measurement is not possible.

  For example: the radiation received by a nuclear gauge is measured and will be an indication of the soil density.

12

Data Quality

Instrumentation Reliability:

Instrumentation Reliability is defined as the extent to which an instrument consistently measures what it is supposed to, i.e. The same results will be obtained again if the measurement is performed again.

Types of Reliability:

- **Internal consistency (Consistency of the items) is typically a measure based on the correlations between different items on the same test.**
- **Test-retest Reliability (Consistency over time). Test-retest is the variation in measurements taken by a single person or instrument on the same item and under the same conditions when tested one after the other.**
- **Interrater Reliability (Consistency between raters). Interrater Reliability describes the degree of agreement among raters.**

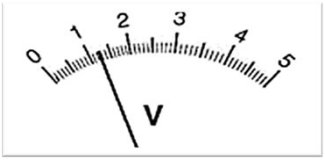
13

Data Quality

Analogue vs. Digital

Analogue:

When reading of an analogue display the precision of readings is limited by our ability to read them.



Digital:

Digital meters normally have sufficient digits to show values more precisely than it is possible to read an analogue display

14

Data Quality

Weighing Scales

Three types of weighing scales:

- The spring scale: measures weight by the distance a spring deflects under its load.

- The strain gauge scale: an electronic version of a spring scale.

- The balance scale:

15

Data Quality

Weighing Scales

Sources of error:

- Error in mass of reference weight.
- Air gusts, even small ones, which push the scale up or down.
- Settling airborne dust contributing to the weight.
- Calibration out of specification over time, due to drift in the circuit's accuracy, or temperature change.
- Condensation of atmospheric water on cold items.
- Evaporation of water from wet items.
- Vibration and seismic disturbances; for example, the rumbling from a passing truck.

16

Data Quality

Test Method:

A test method is a definitive procedure for the identification, measurement, and evaluation of one or more qualities, characteristics, or properties of a material, product, system or service that produces a test result.

A test method must specify :

- the control over such factors as the test equipment.
- the test environment.
- the qualifications of the operator.
- the preparation of test specimens.
- the operating procedure for using the equipment in the test environment.
- the number of test specimens required.
- how measurements on the test specimens are to be combined to provide a test result.

17

Data Quality

A quantitative test method may have three distinct stages:

- The direct measurement or observation of dimensions or properties.

  An observation or observed value (direct reading or interpolation) should be interpreted as the most elemental single reading or corrected reading obtained in the process of making a measurement.

- The arithmetical combination of the observed values to obtain a single determination.

  A test determination may be described as the process of calculating a property of a single test specimen from one or more observations.

- The arithmetical combination of a number of determinations to obtain the test result of the test method.

  A test result is the value obtained by carrying out the complete protocol of the test method once, being either a single test determination or a specified combination of a number of test determinations.

18

Data Quality

Test method types

Direct Test Method:

Test method where the results are directly related to the property of the materials, i.e. Sieve analysis

Indirect Test Method:

Test method where the test results are converted to the property of the material, i.e. when using a Schmidt Hammer, the rebound of the plunger is measured and converted to strength according to a certain empirical relationship

19

Data Quality

Test method types

Basic:

The test results are determined strictly in accordance with a definition of a value i.e. mass per volume (kg/m³) or the compression strength of concrete by testing a concrete cube which reports force per unit area (kN/mm²).

Rational:

Test results of this nature are expressed as a ratio where the basis of reference is a predefined standard, i.e. the degree of compaction of a layer; it is a ratio of the layer density expressed as a percentage of the MDD which is the basis of reference (95% of 100%in labs).

Conditional:

Test results which depend on certain conditions, i.e. the test result density depends on factors such as gradation, moisture content, etc. and is therefore conditional.

20

Data Quality

Test method types

Relational:

The property the test actually measured is related to another property which will be the test result of interest, i.e. the value of a bitumen spray rate is related to the mean of all the measurements of the least chip dimension (in millimetres) of a sample and the test result of interest (spray rate) is in terms of litres per square metre - not in millimetres anymore.

Interpolative:

The test result is obtained from a function and is not directly obtained from the test carried out, i.e. the MDD value is a good example of an interpolated test value. A moisture content versus dry density curve is constructed and from this relationship the Optimum Moisture Content is the moisture content where the relationship is maximum. The point at maximum density is called the Maximum Dry Density.

21

Data Quality

Test method types

Multiple:

Test values can be of more than one type, for example the CBR value at a specified compaction is rational, conditional, and interpolative

Data Quality

Test result reliability

Reliability:

- Reliability refers to the consistency of a measure.
- A test is considered reliable if we get the same result repeatedly.
- The precision and accuracy of measurements are often collectively referred to as reliability.
- Reliability of data is a complex subject and two of the major elements are inherent, variation and human related errors.
- Variability is the result of changes in the conditions under which observations are made.

23

Data Quality

Test result reliability

Variability regarding road building materials and construction test results is defined formally as follows:

$$E_t = E_m + E_i + E_s + E_g + E_e + E_o + E_r$$

Where:

E_t = total variation
E_m = variation due to inherent material variability.
E_i = Induced variation
E_s = variation due to sampling.
E_g = variation due to test method.
E_e = variation due to testing environment.
E_o = variation due to test operator.
E_r = random error.

24

Data Quality

Test result reliability

Total variation:

- In most cases with unsystematic statistical comparison or reproducibility studies the variation obtained is actually the total variation.
- The actual degree or magnitude of the effect of the individual variation as mentioned above is normally unknown to the analyst.
- The degree of variation of a certain property does not have to be significant while the variation of another property can be significantly high.
- Thus the degree of interaction of the individual variation factors varies from situation to situation and must be determined every time a statistical analysis is carried out.

- For example, in a case of a comparison study carried out among testing organisations it is senseless to make a statement that one or more of the laboratories differ significantly when the magnitude of the material variability is not known

25

Data Quality

Test result reliability

Variability due to the material:

- The material variation is subdivided into components such as source and sample variation.
- Material variability is usually the major source of variation if not taken properly care of.

Induced variability:

Induced variability is caused by certain construction actions and preparation for testing.

For example:

- In the case of coarsely graded material the gradation of the material in a layer after compaction does not match up with the material in a stockpile.
- The gradations of the material prepared for certain tests such as the MDD- and CBR-test are not the same as that of the layer it represents.

26

Data Quality

Test result reliability

Sampling:

- Materials for road construction work are not perfectly uniform in terms of the properties to be measured.
- Hence, sampling variability is, in certain cases, the main source of variability among test results.

Test method:

- The execution of a test method involves an interpretation of the test method by a specific test operator.
- Experimental factors - known and unknown - that can change the outcome of a test method are potential sources of variability in test results (Variations in reduction of material, variation in test apparatus etc.).

27

Data Quality

Test result reliability

Test environment:

- Certain properties of certain materials are sensitive to environmental factors such as temperature, humidity, atmospheric pressure.
- The test method must describe or define a tolerable environmental condition.

Test operator:

- Every effort must be made in preparing a standard test method to eliminate the possibility of serious differences in interpretation.
- The test operator must be able to interpret test data correctly. To interpret data correctly requires knowledge and skill.

Random error:

- Random error or unexplained variation is the input of "mother-nature" or the lack of identifying all the factors contributing to the final variation.
- In most cases the effect of random error is usually trivial.

28

Data Quality

Proficiency Schemes

Proficiency Testing Programs (PTP) are statistical quality assurance programs that enable laboratories to assess their performance in conducting test methods within their own laboratories when their data are compared against other laboratories that participate in the same program.

- In the design of any proficiency scheme some of the contributing factors to variation must be kept constant while the factor of interest is assigned as a variable.
- Normally, the factors E_m(material variation), E_i (induced variation) and E_s (sampling variation) are kept constant while E_g (test method variation), E_e (test environment variation), and E_o (operator variation) are variables.
- Before a proficiency program can be executed a repeatability study must be carried out first to establish whether E_m and E_s in particular are constant and what are the respective ranges of the variables.

29

Data Quality

Proficiency Schemes

- In such a repeatability study the magnitude of E_r (random error) must also be determined.
- The reference value(s) must also be agreed beforehand.
- All the samples must be prepared according to a set of rules which addresses all the factors that contribute to variability.
- Any procedure within a test method open to personal interpretation must be addressed in a list of instructions.
- All the participating laboratories must follow the required test according to a step-styled program of actions.
- Valid certificates of all equipment, where a calibration is required, must be available.

30

Data Quality

Specification:

- To attain an acceptable level of quality of a product or construction, it is required to comply with the provisions prescribed in a specification.
- A specification is a set of conditional statements and rules prescribing the quality required.
- For every property quality specified there must be a test method available to verify it.

False precision:

- False precision arises when numerical data are presented in a manner that implies better precision than is actually the case.
- False precision commonly arises when high-precision and low-precision data are combined, and in conversion of units.

31

Data Quality

Precision and Accuracy

Precision:

- The precision of a measurement system, also called reproducibility or repeatability, is the degree to which repeated measurements under unchanged conditions show the same results.
- Precision is how close the measured values are to each other.
- Precision of an experiment/object/value is a measure of the reliability of the experiment.

Accuracy:

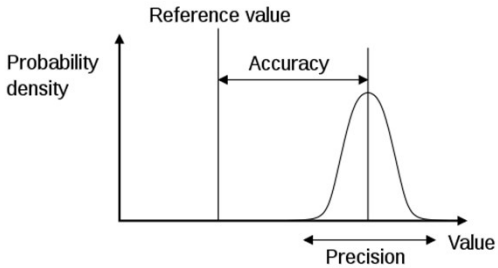
- Accuracy refers to the agreement between a measurement and the true or correct value.
- Accuracy is how close a measured value is to the actual (true) value.
- Accuracy refers to the closeness of a measured value to a standard or known value.

32

Data Quality

Precision and Accuracy

Low Accuracy **High Precision**	**High Accuracy** Low Precision	**High Accuracy** **High Precision**



33

Data Quality

Precision and Accuracy

- **A measurement system can be accurate but not precise, precise but not accurate, neither, or both.**
- **If an experiment contains a systematic error, then increasing the sample size generally increases precision but does not improve accuracy.**
- **Eliminating the systematic error improves accuracy but does not change precision.**

A measurement system is called valid if it is both accurate and precise.

34

Data Quality

Precision and Accuracy

Example:

The true mass of a gem is less than one millionth of a gram away from 5 grams, i.e. the mass of the gem is 4,999999 grams.

If the measurement is reported as 5 grams ± 0,000 003 grams, then the measurement has both high accuracy and high precision.

If it is reported as 5 grams ± 2 grams; then the measurement has high accuracy but low precision.

If it is reported as 8 grams ± 0,000 003 grams; then the measurement has high precision but low accuracy.

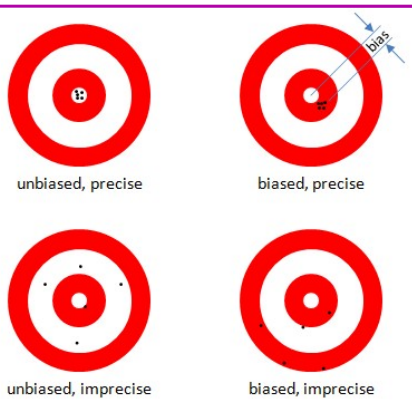
If it is reported as 8 grams ± 2 grams; then the measurement has both low precision and low accuracy.

Data Quality

Bias

Bias:

- **Bias is the consistent deviation of analytical results from the "true" value caused by systematic errors in a procedure.**
- **Bias is a systematic (built-in) error which makes all measurements wrong by a certain amount.**



36

Data Quality

Bias

Examples:

- **The scales read "1 kg" when there is nothing on them**
- **You always measure your height wearing shoes with thick soles.**
- **A stopwatch that takes half a second to stop when clicked.**

37

End of Chapter 3

38