METODOLOGI PENELITIAN
data collection and analysis

Ir. Nurly Gofar, MSCE, PhD
Program Studi Magister Teknik Sipil
Program Pascasarjana
Universitas Bina Darma
Sem Ganjil 2019/2020

Data collection process

- Identify the type of information needed to define the problem
- Select data collection methods best suited for this type of information
- Define the specific target data you intend to collect with each appropriate technique
- Collect the data
- Analyze the data and present in understandable form
- Establish a method to confirm the analysis, such as an experiment or more focused data collection process

Data collection

Source of data

- Experimental
- Survey
- Existing data

Data Collection Methods

- Survey questionnaires
- Interview
- Observation
- Experiment
- Existing data

Survey Questionaires

- Open-ended question (what h/p you use?)
- Closed-ended question
  - Scale question (strongly agree, agree, disagree, etc)
  - Listing / choice question (I use sony, nokia, etc)
  - Yes/No question
  - Ranking question
  - Category question

Interviews

Types of interview:

- Structured
- Unstructured
- Semi-structured

Questions to be asked in the initial planning of SURVEY and INTERVIEW instrumentation.

- What variables shall be investigated ?
- What control must be exterted on the experiment ?
- What range of the primary variables will be necessary to describe the phenomena under study ?
- How many data points should be taken in the range ?
- What instrument accuracy required for each measurement ?
- If a dynamic measurement is involved, what frequency response must the instrument have ?
- Are the instrument available in market
- What safety precautions are necessary in case of hazardous operation ?
- Financial ?
- What provision have been made for recording the data ?

Observation

- No measurement e.g.:
  - What are the activities that students do in a classroom?
  - What are the arrival and departure time of students?

Experiments

- Types of experimental / Survey
  - Realistic setting (survey, field experiment, case study
  - Artificial setting (Laboratory, use standard procedure)
  - Parametric study / Numerical

Observations vs. Experiments

- Both types of studies may seek to the understanding of relationships between variables
- Experimental studies are
  - better at determining **cause-and-effect**
  - less likely to be confounded by hidden variables

Observation

- Does not attempt to manipulate the independent variable (s) because the variable (s) is (are) inherently can not be manipulated for some reason or another.
- Researcher interference is minimum if studying events as they normally occur : correlational study i.e : to identify the important factors "associated with" the problem
- Types of Qualitative Design :
  - Exploratory Study
  - Descriptive/Predictive

Exploratory Study

- Develop concept, establish priorities, develop operational definitions
- Secondary data analysis
- Experience surveys
- Focus groups
- Two-stage design

Descriptive/Predictive

- Surveys – uses some sort of questionnaire to describe and/or predict some phenomenon
- Type – paper – and – pencil questions, personal interviews

Experiment

- An experiment is an activity or *process*, a combination of activities which produces events, possible outcomes.
  - e.g. tossing a coin or tossing a dice
- The research manipulates or in some way controls the independent variable (s) and measures the effect on the dependent variable (s) of interest
- An experiment is designed and it occurs in the future.
- If data have been collected in the past → ex-post-facto
  - e.g. developing empirical correlation model from which forecast can be made

Experimental

Types of Variables

- independent((IV)
- dependent (DV)
- Moderating variables – has a strong effect on relationship between independent and dependent variables
- Intervening variables – a function of the independent variables → helps to explain the influence of IV on DV

Types of Correlation

Causal : Essential element of causation e.g. A "produces" B or A "forces" B to occur

Causal relationship

Symmetrical – two variables fluctuate together

Reciprocal – two variables are mutually influence

Asymmetrical – changes in one variable is responsible for changes in another variable

Experimental Control

- How to vary an independent variable to assess its interference on the dependent variable
- How to collect the data (Data Collection)
- How to control extraneous variable that may influence the dependent variable or the developed relationship (Data Analysis)

Experimental data

- No simplifying assumption necessary if tests are run on an actual / prototype system. The true behavior is revealed.
- Accurate measurement necessary. The characteristic of all devices should be thoroughly understood
- Actual system or a scaled model is required. If scaled model, similarity of all significant features must be preserved.
- Consideration of time required for design of test.

Engineering Design

To design or modify a product

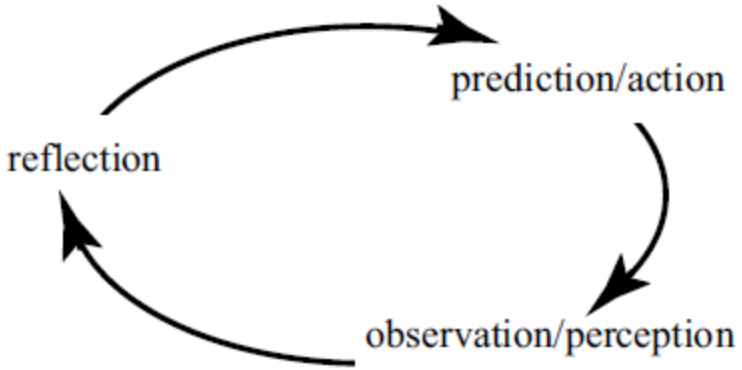
- Equipment
- Computer program
- Subroutine
- Procedure
- Guidelines
- Modeling

Modeling

Although there are many other types of modeling techniques available such as : theoretical model, empirical model, constitutive model, geological model and classification model, **physical model and numerical model** have been adopted more frequently in engineering research.

Physical Model

Physical modeling is performed in order to validate theoretical or empirical hypotheses. Poorly designed physical modeling is mere data gathering. However, good model should fulfill the reflective practice cycle which is not stuck in the *prediction / observation part of the loop* but also closed by the need for precedent and subsequent *reflection*.



Physical Model

Two types of Physical Model

- Physical models: full-scale

Full scale testing is usually performed to evaluate geotechnical processes which it is believed may be so dependent on the detail of actual structure that it is imperative to use real material as prepared by nature. Disadvantages of full scale modeling lead to the use of small scale model.

- Physical models: small-scale

Problems relates to scale. Some quantities cannot be scaled down (e.g. gravity) → centrifuge model

Data Collection

DETERMINE SAMPLE SIZE

consider confidence level for the results and practical aspects – based on timeframe and cost

Example:

Objective: To compare the quality of ready mix concrete from two suppliers by considering water cement ratio, maximum aggregate size and amount of plastisizer

Number of experiment :

assume you want to take only high and low values of each independent variables, then

$2 \times 3 \times 2 \times 2 = 24$.

Data Collection

CONDUCT THE EXPERIMENT/SURVEY

collect data accurately, keep and maintain testing conditions for validity and consistency of results.

SETTING

- Realistic setting

  Survey, Field experiments, Case study

- Artificial setting (Laboratory, Field Test)

  use standard procedure (BS., ASTM. etc)

- Parametric Study / Numerical

  should be verified

ESTABLISH EXPERIMENTAL CONTROL

Data Analysis

- use appropriate techniques to analyze the results --- aimed at answering the research questions/hypothesis
- provide justifications of the analytical procedures
- use statistics or other tools
  - REGRESSION : Trend equation, R^2 value, Standard of Deviation, Variance, Covariance, Coefficient of Correlation
  - QUANTITATIVE HYPOTHETICAL TEST:
    One tail test AND Two tail test
  - ANALYSIS OF VARIANCE
  - MODEL TEST

Data Analysis

- ✓Comparisons
- ✓Differences
- ✓Extreme value
- ✓Percentage etc.
- ✓Effect

RESULTS OF DATA ANALYSIS

- Trend/Tendency
- Hypothetical Test

Ensuring Validity

Requirements of a valid results

- It is important to be sure of the validity of the work – the confidence which someone may have in the findings – how findings may be used in future research or in practical application
- Validity is concerned with limiting research errors so the results are accurate and usable when delivered.
- There are two types of validity
  - Internal validity – the degree to which the results of the study can be relied upon as being correct.
  - External Validity – the degree to which the study's results can be generalized across populations, setting and other, similar conditions.

Internal Validity is affected by the following factors:

- History – event outside affect the results
- Maturation – changes occurs in respondents over time - aging
- Testing – effects the process
- Instrumentation – changes in measurement
- Selection procedure
- Mortality – respondents may be lost from the study.

External validity relate closely to generalization → be careful because

- False consensus effect → based on one sample
- Small, biased, non-representative sample (statistically incorrect) → DANGER (depression, Stereotyping, Foolishness)

- Ensuring validity is closely related to minimizing errors Errors could occur during Planning, Data Collection and Data Analysis
- Minimizing Planning Error
  - Write Good Proposals
  - Develop Good Research Design
  - Use STANDARDIZED experimental technique (BS, ASTM, Eurocode, SNI, etc)
  - Careful with identification and measurement of data
  - Define assumptions clearly
  - Prepare Hypothesis (relationship between IV and DV) carefully

Minimizing Data collection error

- Variable
  - Variable must be identified (definition, measurement)
  - Assumptions must be clearly defined
  - Hypothesis
    (relationship between IV and DV)

  Example : The method of programming construction project employed by a contractor influenced the project performance and hence participant satisfaction of the project.

  HYPOTHESIS

  programming – performance

  performance – satisfaction

- by good measurement
  - accuracy
  - probability of error (randomized change, selected range of variable (e.g. high low), control IV rigidly
  - Control Replication!!!!!

Instrument

Evaluation of instruments :

Feasibility (Usually established through a pilot study)

- the appropriateness of a particular instrument to a particular need
- can collect the needed information
- respondents understand it
- the interviewer understand it
- can be completed in time

Validity

- the degree to which a particular instrument is useful in measuring
- if not accurate , it is worthless

Reliability

- the degree to which a particular test or instrument provides consistent or trustworthy measures.

Instrument

Elements in a data logging system :

- Measured Medium
- Primary sensing element (strain gages, transducer, etc.)
- Variable conversion element (tape recorder, digital, etc)
- Variable manipulation element (computer, etc.
- Presentation of data element
- Observer

Sources of error in Experimental data

Mistake

1) reading taking, reading recording, zero error, calculation mistake can be avoided by be careful and repeating reading.

Human error

1) - parallax error,
2) Limit of equipment / instrumentation
3) the quality of measuring instrument, range of environmental condition, mechanism of measurement e.g friction.

Observation error

1) e.g connecting pressure gauge to a tyre will reduce a bit the pressure in the tyre.
2) Driver will reduce speed when noticing someone doing traffic survey .

Sources of error in Experimental data

External effect

- Wind can affect the reading of a sensitive weighing machine
- Impurites can affect the chemical reaction
- Temperature can affect the result of some mechanical and chemical equipment

Statistical fluctuation

- Error due to unrepresentative sample.
- When taking a sample for measurement to conclusion for the whole population, the selected sample must represent the population.

Dealing with error

While executing the test, do these things to ensure that the test produce good results:

- Record anything about the test in a note book
- State measurement, observations and comments directly into the note books.
- Do not write in a loose paper
- Record the date and time of the experiment
- Write down all mathematical calculation on the note book, do not just doing calculation in your mind. For example: to find the difference in weight of two objects, write down the weight of each the object, do not just write the calculated difference.
- When taking a set of readings, fill in them in a table form.

Minimizing Data Collection Error

- By Increasing Accuracy
  - Increasing the size of experiment by replication or incorporating more treatments
  - Refining the experimental method or techniques to minimize error
  - Arranging experimental units into homogeneous groups.

Example :

Objective: To compare the quality of ready mix concrete from two suppliers by considering water cement ratio, maximum aggregate size and amount of plastisizer

- Number of experiment :
- assume you want to take only high and low values of each independent variables, then

Sampling and Data Collection

- Identify Variables
  - Independent variables
  - Dependent variables

- Determine sample size
  - Consider confidence level for the results and practical aspects – based on timeframe and cost

Sampling

- Census - obtaining sample from the whole population
- Survey – obtaining sample from portion of population (reliability is different)
- Advantages of sampling: Reduce cost, Greater speed, Greater scope, Greater accuracy
- use Inferential Techniques to make conclusions about the population based on the data in the sample

Type of Sampling

- Random Sampling
  - Simple sampling – every member has an equal chance
  - Stratified sampling – sample selected in proportion
  - Cluster sampling – selection of group sample
- Non-random sampling
  - Systematic sampling – every N th. sample in a population.
  - Convenience – samples who are readily accessible
  - Purposive – samples selected based on criteria

Sample Quality

Study should be designed to achieve a high quality sample

- High quality samples are representative of the population
- Poor quality samples produce misleading results (biased)
- To avoid biased sampling, use impersonal chance mechanisms as the basis for selection i.e. Simple Random Sample (SRS)
  - Each individual in population has the same chance of being selected.
  - Every possible sample has an equal chance to be chosen.

Example of Poor Quality Sampling Designs

Voluntary response sampling

Allows individuals to choose to be in the study
e.g., Call-in polls

Example : Survey to study the efficiency of office administration in a government agency.

- 4.5% responded
- the researcher used those responses to write report.

Weakness of the survey method:

- respondents "were fed up with the admin and eager to criticize …"
- "the anger became the theme of the report…"
- "angry staff are more likely" to respond.

- This technique favors certain outcomes and are therefore biased

Example of Poor Quality Sampling Designs

Convenience Sampling

- Individuals that are easiest to reach are selected
- e.g., Interviewing at the mall

Example : Sampling mice from a large cage to study how a drug affects physical activity

- Lab assistant reaches into the cage to select the mice one at a time until 10 are chosen
- Which mice will likely be chosen?
  - The less active mice will be selected
  - The sample yields biased results

- This technique favors certain outcomes and are therefore biased

Limitation of sampling

- Ability to measure the precision of the sample result
- No more representative than the list
- **Principle steps in a sample survey**
- Objective of the survey
- Population to be sampled
- Data to be collected
- Degree of precision desired ? The larger the sample, the more the precision
- Method of measurement

Minimizing Reporting Error

- conclusion should have maximum validity (concerning depth and breadth)
- experimental technique should be kept simple
- use assistance of unbiased and knowledgeable reviewers

Final Remarks

- There is no single, correct design for a research problem
- Design research to answer the research problem
- All research design represents a compromise
- A research design is not a framework to be blindly and without deviation