

# PRESIDENCY UNIVERSITY, BENGALURU



Presidency University, Bengaluru

**Course Code : CSE2015**

**Course Name : Data Analysis and  
Visualization**

**Faculty Name: Dr. Ragaventhiran**

**Designation : Associate Professor**

**Department : Computer Science &  
Engg.**



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Module 1: Introduction to Data Visualization



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Introduction

- Data Collection Strategies
- Characteristics of Good Measures
- Quantitative and Qualitative Data
- Tools for Collecting Data



# Data Collection Strategies

- No one best way: decision depends on:
  - What you need to know: *numbers or stories*
  - Where the data reside: *environment, files, people*
  - Resources and time available
  - Complexity of the data to be collected
  - Frequency of data collection
  - Intended forms of data analysis



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Rules for Collecting Data

- Use multiple data collection methods
- Use available data, but need to know
  - how the measures were defined
  - how the data were collected and cleaned
  - the extent of missing data
  - how accuracy of the data was ensured



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Rules for Collecting Data

- If must collect original data:
  - be sensitive to burden on others
  - pre-test, pre-test, pre-test
  - establish procedures and follow them (protocol)
  - maintain accurate records of definitions and coding
  - verify accuracy of coding, data input



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Structured Approach

- All data collected in the same way
- Especially important for multi-site and cluster evaluations so you can compare
- Important when you need to make comparisons with alternate interventions



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Use Structured Approach When:

- need to address extent questions
- have a large sample or population
- know what needs to be measured
- need to show results numerically
- need to make comparisons across different sites or interventions



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Semi-structured Approach

- Systematic and follow general procedures but data are not collected in exactly the same way every time
- More open and fluid
- Does not follow a rigid script
  - may ask for more detail
  - people can tell what they want in their own way



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Use Semi-structured Approach when:

- conducting exploratory work
- seeking understanding, themes, and/or issues
- need narratives or stories
- want in-depth, rich, “backstage” information
- seek to understand results of data that are unexpected



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Characteristics of Good Measures

- Is the measure relevant?
- Is the measure credible?
- Is the measure valid?
- Is the measure reliable?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Relevance

Does the measure  
capture what  
matters?

Do not measure  
what is easy  
instead of what is  
needed



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Credibility

Is the measure believable?

Will it be viewed as a reasonable and appropriate way to capture the information sought?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Internal Validity

How well does the measure capture what it is supposed to?

Are waiting lists a valid measure of demand?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Reliability

A measure's precision and stability- extent to which the same result would be obtained with repeated trials

How reliable are:

- birth weights of newborn infants?
- speeds measured by a stopwatch?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Quantitative Approach

- Data in numerical form
- Data that can be precisely measured
  - age, cost, length, height, area, volume, weight, speed, time, and temperature
- Harder to develop
- Easier to analyze



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Qualitative Approach

- Data that deal with description
- Data that can be observed or self-reported, but not always precisely measured
- Less structured, easier to develop
- Can provide “rich data” — detailed and widely applicable
- Is challenging to analyze
- Is labor intensive to collect
- Usually generates longer reports



# Which Data?

*If you:*

- want to conduct statistical analysis
- want to be precise
- know what you want to measure
- want to cover a large group

*Then Use:*

Quantitative

- want narrative or in-depth information
- are not sure what you are able to measure
- do not need to quantify the results

Qualitative

# Obtrusive vs. Unobtrusive Methods

## Obtrusive

data collection methods that directly obtain information from those being evaluated

e.g. interviews, surveys, focus groups

## Unobtrusive

data collection methods that do not collect information directly from evaluatees

e.g., document analysis, GoogleEarth, observation at a distance, trash of the stars



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# How to Decide on Data Collection Approach

- Choice depends on the situation
- Each technique is more appropriate in some situations than others
- Caution: All techniques are subject to bias



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Triangulation to Increase Accuracy of Data

- Triangulation of methods
  - collection of same information using different methods
- Triangulation of sources
  - collection of same information from a variety of sources
- Triangulation of evaluators
  - collection of same information from more than one evaluator



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Data Collection Tools

- Participatory Methods
- Records and Secondary Data
- Observation
- Surveys and Interviews
- Focus Groups
- Diaries, Journals, Self-reported Checklists
- Expert Judgment
- Delphi Technique
- Other Tools



# Tool 1: Participatory Methods

- Involve groups or communities heavily in data collection
- Examples:
  - community meetings
  - mapping
  - transect walks



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Community Meetings

- One of the most common participatory methods
- Must be well organized
  - agree on purpose
  - establish ground rules
    - who will speak
    - time allotted for speakers
    - format for questions and answers



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Mapping

- Drawing or using existing maps
- Useful tool to involve stakeholders
  - increases understanding of the community
  - generates discussions, verifies secondary sources of information, perceived changes
- Types of mapping:
  - natural resources, social, health, individual or civic assets, wealth, land use, demographics



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Transect Walks

- Evaluator walks around community observing people, surroundings, and resources
- Need good observation skills
- Walk a transect line through a map of a community — line should go through all zones of the community



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Tool 2: Records and Secondary Data

- Examples of sources:
  - files/records
  - computer data bases
  - industry or government reports
  - other reports or prior evaluations
  - census data and household survey data
  - electronic mailing lists and discussion groups
  - documents (budgets, organizational charts, policies and procedures, maps, monitoring reports)
  - newspapers and television reports



# Using Existing Data Sets

Key issues: validity, reliability, accuracy, response rates, data dictionaries, and missing data rates



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Advantage/Challenge: Available Data

Advantages	Often less expensive and faster than collecting the original data again
Challenges	There may be coding errors or other problems. Data may not be exactly what is needed. You may have difficulty getting access. You have to verify validity and reliability of data

# Tool 3: Observation

- See what is happening
  - traffic patterns
  - land use patterns
  - layout of city and rural areas
  - quality of housing
  - condition of roads
  - conditions of buildings
  - who goes to a health clinic



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Observation is Helpful when:

- need direct information
- trying to understand ongoing behavior
- there is physical evidence, products, or outputs than can be observed
- need to provide alternative when other data collection is infeasible or inappropriate



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Degree of Structure of Observations

- Structured: determine, before the observation, precisely what will be observed before the observation
- Unstructured: select the method depending upon the situation with no pre-conceived ideas or a plan on what to observe
- Semi-structured: a general idea of what to observe but no specific plan



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Google Earth

- Maps and satellite images for complex or pinpointed regional searches
- Has an Advanced version and an Earth Outreach version
- Web site for Google Earth
  - <http://earth.google.com/>



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Ways to Record Information from Observations

- Observation guide
  - printed form with space to record
- Recording sheet or checklist
  - Yes/no options; tallies, rating scales
- Field notes
  - least structured, recorded in narrative, descriptive style



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Guidelines for Planning Observations

- Have more than one observer, if feasible
- Train observers so they observe the same things
- Pilot test the observation data collection instrument
- For less structured approach, have a few key questions in mind



# Advantages and Challenges: Observation

**Advantages** Collects data on actual vs. self-reported behavior or perceptions. It is real-time vs. retrospective

---

**Challenges** Observer bias, potentially unreliable; interpretation and coding challenges; sampling can be a problem; can be labor intensive; low response rates

# Tool 4: Surveys and Interviews

- Excellent for asking people about:
  - perceptions, opinions, ideas
- Less accurate for measuring behavior
- Sample should be representative of the whole
- Big problem with response rates



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Structures for Surveys

- Structured:
  - Precisely worded with a range of pre-determined responses that the respondent can select
  - Everyone asked exactly the same questions in exactly the same way, given exactly the same choices
- Semi-structured
  - Asks same general set of questions but answers to the questions are predominantly open-ended



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Structured vs. Semi-structured Surveys

Structured	harder to develop easier to complete easier to analyze more efficient when working with large numbers
Semi-structured	easier to develop: open ended questions more difficult to complete: burdensome for people to complete as a self-administrated questionnaire harder to analyze but provide a richer source of data, interpretation of open-ended responses subject to bias

# Modes of Survey Administration

- Telephone surveys
- Self-administered questionnaires distributed by mail, e-mail, or websites
- Administered questionnaires, common in the development context
- In development context, often issues of language and translation



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Mail / Phone / Internet Surveys

- Literacy issues
- Consider accessibility
  - reliability of postal service
  - turn-around time
- Consider bias
  - What population segment has telephone access? Internet access?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Advantages and Challenges of Surveys

Advantages	Best when you want to know what people think, believe, or perceive, only they can tell you that
------------	---

---

Challenges	People may not accurately recall their behavior or may be reluctant to reveal their behavior if it is illegal or stigmatized. What people <i>think they do</i> or <i>say they do</i> is not always the same as what they <i>actually do</i> .
------------	---

# Interviews

- Often semi-structured
- Used to explore complex issues in depth
- Forgiving of mistakes: unclear questions can be clarified during the interview and changed for subsequent interviews
- Can provide evaluators with an intuitive sense of the situation



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Challenges of Interviews

- Can be expensive, labor intensive, and time consuming
- Selective hearing on the part of the interviewer may miss information that does not conform to pre-existing beliefs
- Cultural sensitivity: e.g., gender issues



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Tool 5: Focus Groups

- Type of qualitative research where small homogenous groups of people are brought together to informally discuss specific topics under the guidance of a moderator
- Purpose: to identify issues and themes, not just interesting information, and not “counts”



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Focus Groups Are Inappropriate when:

- language barriers are insurmountable
- evaluator has little control over the situation
- trust cannot be established
- free expression cannot be ensured
- confidentiality cannot be assured



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Focus Group Process

<i>Phase</i>	<i>Action</i>
1 Opening	Ice-breaker; explain purpose; ground rules; introductions
2 Warm-up	Relate experience; stimulate group interaction; start with least threatening and simplest questions
3 Main body	Move to more threatening or sensitive and complex questions; elicit deep responses; connect emergent data to complex, broad participation
4 Closure	End with closure-type questions; summarize and refine; present theories, etc; invite final comments or insights; thank participants



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Advantages and Challenges of Focus Groups

**Advantages** Can be conducted relatively quickly and easily; may take less staff time than in-depth, in-person interviews; allow flexibility to make changes in process and questions; can explore different perspectives; can be fun

**Challenges** Analysis is time consuming; participants not be representative of population, possibly biasing the data; group may be influenced by moderator or dominant group members



# Tool 6: Diaries and Self-Reported Checklists

- Use when you want to capture information about events in people's daily lives
- Participants capture experiences in real-time not later in a questionnaire
- Used to supplement other data collection



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Guidelines for Diaries or Journals

## *Step Process*

### 1 Recruit people face-to-face

- encourage participation, appeal to altruism, assure confidentiality, provide incentive

### 2 Provide a booklet to each participant

- cover page with clear instructions, definitions, example
- short memory-joggers, explain terms, comments on last page , calendar

### 3 Consider the time-period for collecting data

- if too long, may become burdensome or tedious
- if too short may miss the behavior or event



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Self-reported Checklists

- Cross between a questionnaire and a diary
- The evaluator specifies a list of behaviors or events and asks the respondents to complete the checklist
- Done over a period of time to capture the event or behavior
- More quantitative approach than diary



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Advantages and Challenges of Diaries and Self-reported Checklists

<b>Advantages</b>	<ul style="list-style-type: none"><li>Can capture in-depth, detailed data that might be otherwise forgotten</li><li>Can collect data on how people use their time</li><li>Can collect sensitive information</li><li>Supplements interviews provide richer data</li></ul>
-------------------	--

<b>Challenges</b>	<ul style="list-style-type: none"><li>Requires some literacy</li><li>May change behavior</li><li>Require commitment and self-discipline</li><li>Data may be incomplete or inaccurate</li><li>Poor handwriting, difficult to understand phrases</li></ul>
-------------------	--

# Tool 7: Expert Judgment

Use of experts, one-on-one or as a panel

E.g., Government task forces, Advisory Groups

Can be structured or unstructured

Issues in selecting experts



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Selecting Experts

- Establish criteria for selecting experts not only on recognition as expert but also based on:
  - areas of expertise
  - diverse perspectives
  - diverse political views
  - diverse technical expertise



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013



# Advantages and Challenges of Expert Judgment

Advantages    Fast, relatively inexpensive

---

Challenges    Weak for impact evaluation  
May be based mostly on perceptions  
Value of data depends on how credible the experts are perceived to be

# Tool 8: Delphi Technique

- Enables experts to engage remotely in a dialogue and reach consensus, often about priorities
- Experts asked specific questions; often rank choices
- Responses go to a central source, are summarized and fed back to the experts without attribution
- Experts can agree or argue with others' comments
- Process may be iterative





# Advantages and Challenges of Delphi Technique

**Advantages**

- Allows participants to remain anonymous
- Is inexpensive
- Is free of social pressure, personality influence, and individual dominance
- Is conducive to independent thinking
- Allows sharing of information

---

**Challenges**

- May not be representative
- Has tendency to eliminate extreme positions
- Requires skill in written communication
- Requires time and participant commitment

# Other Measurement Tools

- scales (weight)
- tape measure
- stop watches
- chemical tests :  
i.e. quality of water
- health testing tools:  
i.e. blood pressure
- aptitude and  
achievement tests
- citizen report cards

# Data Collection Summary

Choose more than one data collection technique

No “best” tool

Do not let the tool drive your work but rather choose the right tool to address the evaluation question



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013





# Questions?



**PRESIDENCY  
UNIVERSITY**

Private University Estd. in Karnataka State by Act No. 41 of 2013

