

# Data for Impact Evaluation



## EVALUATION

- ☒ **Outstanding**
- ☐ **Very Good**
- ☐ **Satisfactory**
- ☐ **Marginal**
- ☐ **Unsatisfactory**

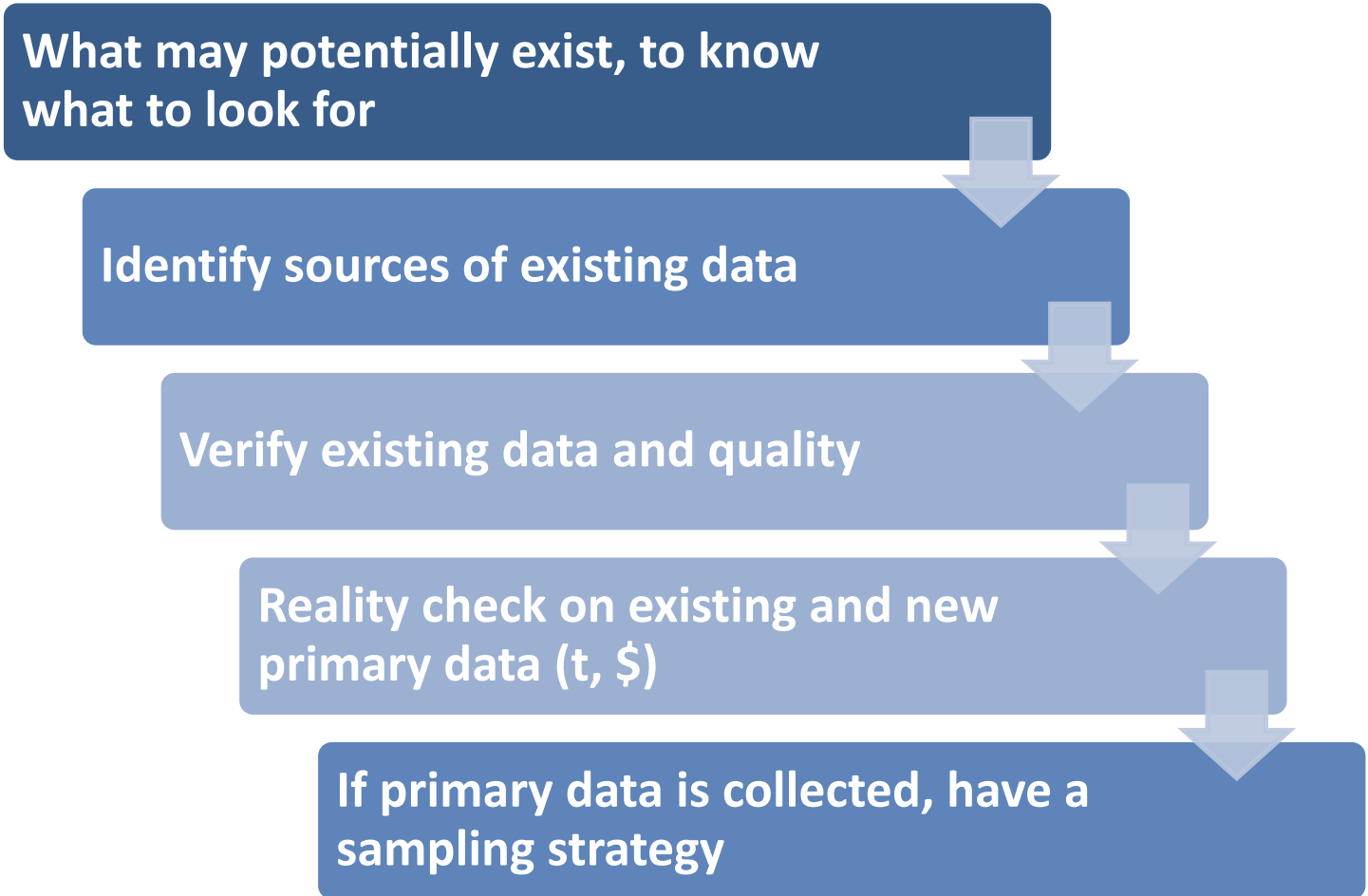
**Susana Puerto**  
**Puerto-Gonzalez@ilo.org**

# OUTLINE

1. What data we need
2. Sources of data
3. Data quality
4. Qualitative vs. quantitative data collection methods
5. Sampling Strategy

# 1. WHAT DATA WE NEED

What may potentially exist, to know what to look for



```
graph TD; A[What may potentially exist, to know what to look for] --> B[Identify sources of existing data]; B --> C[Verify existing data and quality]; C --> D[Reality check on existing and new primary data (t, $)]; D --> E[If primary data is collected, have a sampling strategy];
```

Identify sources of existing data

Verify existing data and quality

Reality check on existing and new primary data (t, \$)

If primary data is collected, have a sampling strategy

# 1. WHAT DATA WE NEED

**a. INDICATORS**



**b. LEVEL OF  
OBSERVATION**



**c. TIMING OF DATA  
COLLECTION**



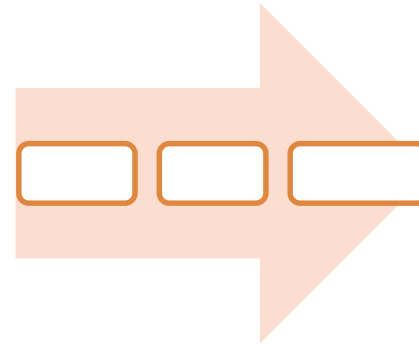
**d. PRIMARY VS.  
SECONDARY DATA**



## a. INDICATORS

- **Theory of change**

- Outcomes of interest
- Outputs
- External factors



- **Challenges in measuring indicators**

- Validity and
- Reliability

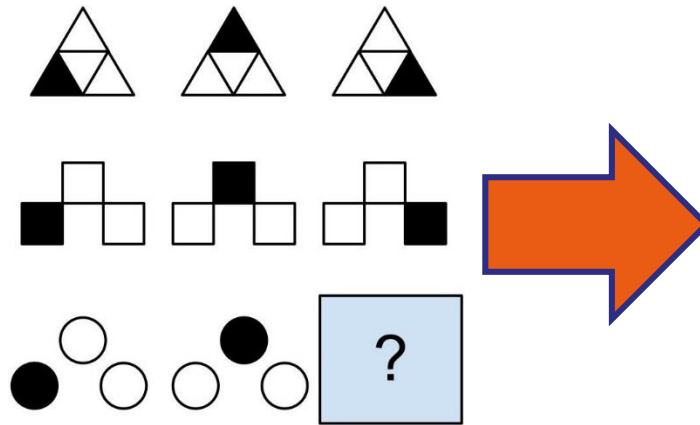


- **What is hard to measure?**

# Validity

## In theory:

How well does the indicator map to the outcome?  
(IQ tests)



## Intelligence

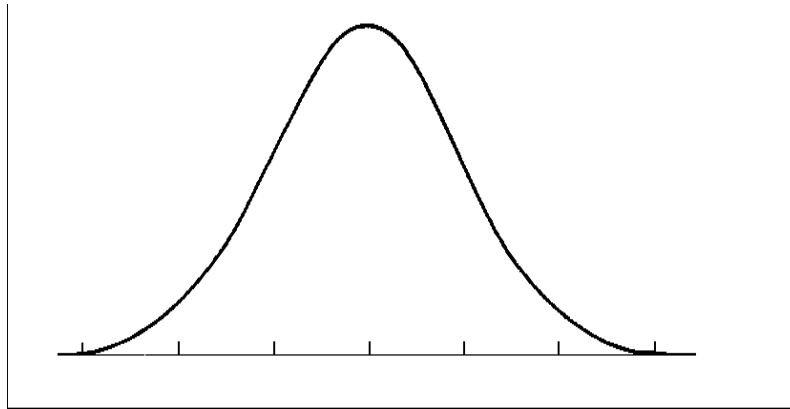
## In practice:

- Are your survey questions unbiased?
- Potential biases:
  - Social desirability bias
  - Demand bias (response bias)
  - Recall bias

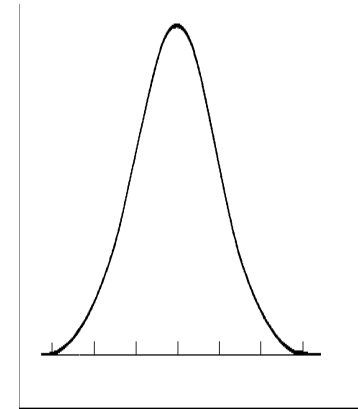
# Reliability

## In theory:

How much do the measurements of our indicator vary?



vs.



## In practice:

- Length, fatigue
- Ambiguous wording (definitions, relationships, recall period)
- Answer choice (open/closed, Likert,...)

# What is hard to measure?

- (1) Things people do not know very well
  - (2) Things people do not want to talk about
  - (3) Abstract concepts
  - (4) Things that are best directly observed
- ⇒ If we can't measure it, we can't evaluate it**

# (1) Things people do not know very well

## What:

Anything to **estimate**, particularly across time.

Prone to recall error and poor estimation

- **Examples:** profit, income, income, plot size, number of hours worked per month,...

## Strategies:

- **Consistency checks:** – How much did you spend in the last week on x? How much did you spend in the last 4 weeks on x?
- **Multiple measurements of same indicator** – How many minutes does it take to walk to the health center? How many kilometers away is the health center?

## (2) Things people don't want to talk about

### **What:**

Anything socially “risky” or something painful

- **Examples:** sexual activity, alcohol and drug use, domestic violence, conduct during wartime, mental health

### **Strategies:**

- Don't start with the tricky questions!
- Consider asking question in third person
- Always ensure comfort and privacy of respondent

# (3) Abstract Concepts

## What:

- **Examples:** empowerment, bargaining power, social cohesion, risk aversion

## Strategies:

- Three key steps when measuring “abstract concepts”
  - Define what you mean by your abstract concept
  - Choose the outcome that you want to serve as the measurement of your concept
  - Design a good question to measure that outcome
- Often choice between a self-reported measure and a behavioural measure – both can add value!

## (4) Things that are best directly observed

### **What:**

Behavioral preferences, anything that is more believable when done than said

### **Strategies:**

- Develop detailed protocols
- Ensure data collection of behavioral measures done under the same circumstances for all individuals

# Choosing fruit vs. chocolate

Read and van Leeuwen (1998)

Choosing Today

Eating Next Week

*Time*

If you were  
deciding **today**,  
would you choose  
fruit or chocolate  
for **next week**?



# Choosing fruit vs. chocolate

Choosing Today

Eating Next Week

*Time*

**Today**, subjects typically choose fruit for **next week**.

74%



# Choosing fruit vs. chocolate

Choosing and Eating  
Simultaneously

*Time*

If you were  
deciding **today**,  
would you choose  
fruit or chocolate  
for **today**?



# Choosing fruit vs. chocolate

Choosing and Eating  
Simultaneously

*Time*

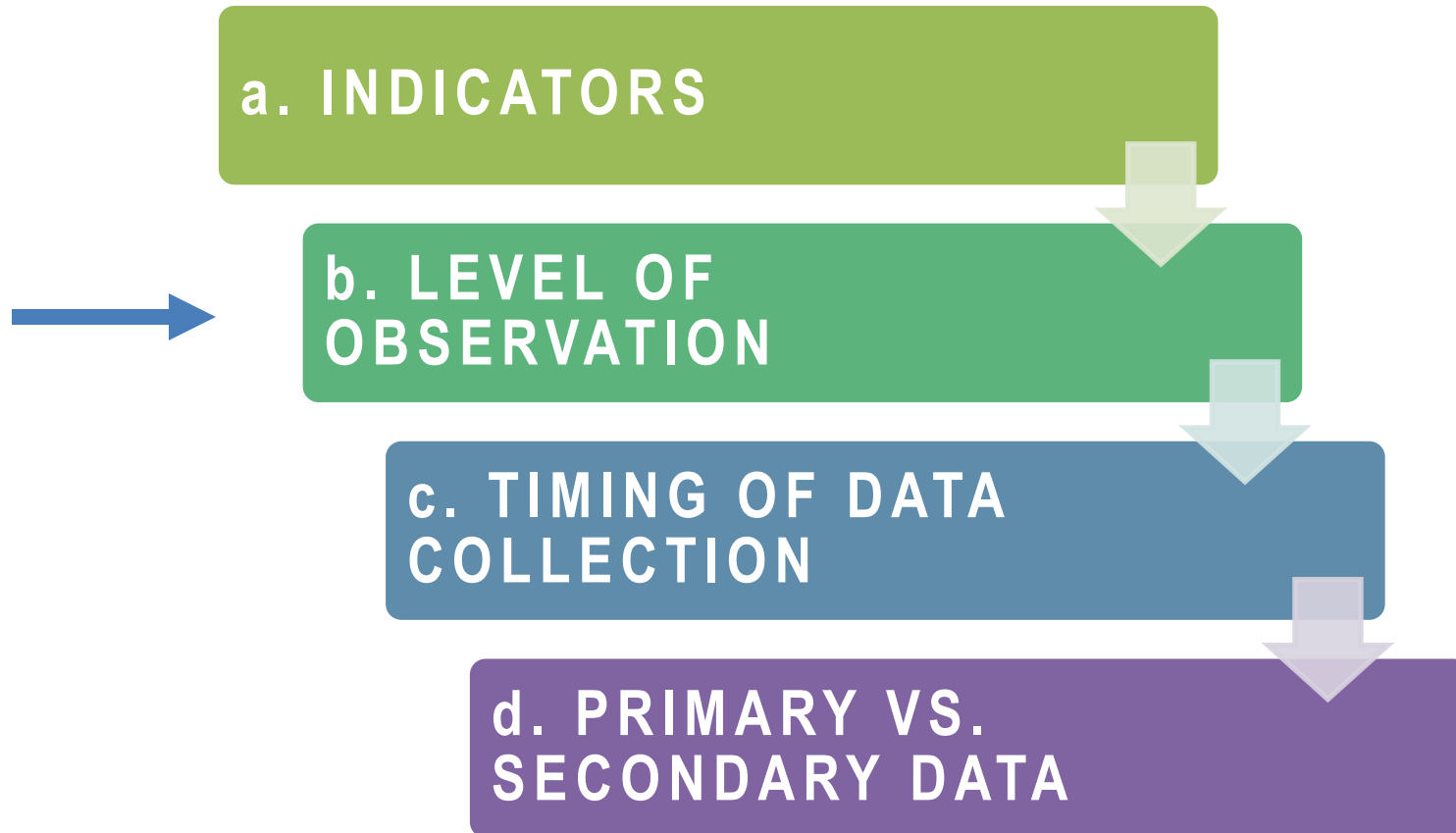
**Today**, subjects  
typically choose  
Chocolate for  
**today**.



70%



# 1. WHAT DATA WE NEED



## b. LEVEL OF OBSERVATION

- Individual (youth)
- Household
- Community/Village
- Facility (school)
- Firm/Enterprise

## c. TIMING OF DATA COLLECTION

- At least two points in time for rigorous IE (progress, trends)
- Timing for a baseline
- Timing for a follow-up

**Implication:** *The evaluation has to go hand in hand with the implementation →*

The method depends on the rules of operation

## c. TIMING OF DATA COLLECTION

|                                 | Targeted<br>(eligibility cutt-off)                                                   |                                                                                                 | Universal<br>(no eligibility cutt-off)                                                                        |                                                                                                                          |
|---------------------------------|--------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
|                                 | Limited Resources                                                                    | Fully Resourced                                                                                 | Limited Resources                                                                                             | Fully Resourced                                                                                                          |
| Phased Implementation Over Time | <ul style="list-style-type: none"> <li>■ Random Assignment</li> <li>■ RDD</li> </ul> | <ul style="list-style-type: none"> <li>■ Random Assignment (phase in)</li> <li>■ RDD</li> </ul> | <ul style="list-style-type: none"> <li>■ Random Assignment</li> <li>■ Matching with Diff. in Diff.</li> </ul> | <ul style="list-style-type: none"> <li>■ Random Assignment (phase in)</li> <li>■ Matching with Diff. in Diff.</li> </ul> |
| Immediate Implementation        | <ul style="list-style-type: none"> <li>■ Random Assignment</li> <li>■ RDD</li> </ul> | <ul style="list-style-type: none"> <li>■ Random Promotion</li> <li>■ RDD</li> </ul>             | <ul style="list-style-type: none"> <li>■ Random Assignment</li> <li>■ Matching with Diff. in Diff.</li> </ul> | <ul style="list-style-type: none"> <li>■ Random Promotion</li> </ul>                                                     |

## d. PRIMARY VS. SECONDARY DATA

- Efficiency of existing data
- \$ available for new data
- Piggyback

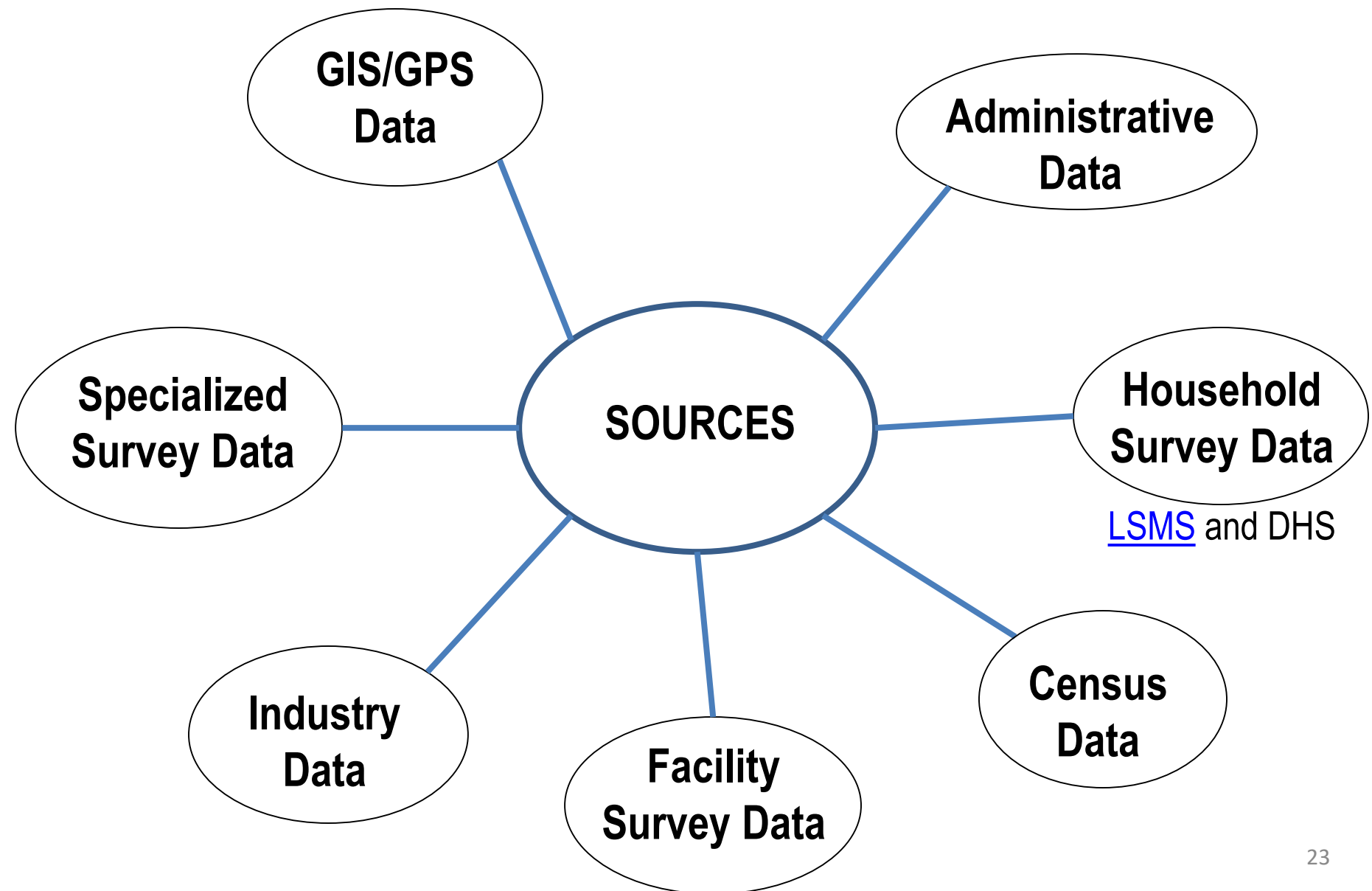
### Potential trade-offs:

1. **Sampling frame:** insufficient overlap (often case with youth).
2. **Timing:** timing of the data vis-à-vis implementation.
3. **Quality:** Check quality of existing data

# OUTLINE

1. What data we need
2. Sources of data
3. Data quality
4. Qualitative vs. quantitative data collection methods
5. Sampling Strategy

## 2. POTENTIAL SOURCES OF DATA



### 3. DATA QUALITY

#### Questionnaire Design

- Team Work
- Question wording and sequencing
- Consistency checks
- Item non-response
- Copying former questionnaires
- Field Testing

#### Quality Control

- Fieldwork and data management
- Electronic field survey forms
- Checking secondary data

#### Data Auditing

- Internal consistency of individual observations
- Consistency with external data sources
- Consistency among subsets of a dataset
- Statistical Checks

## 4. QUANTITATIVE VS. QUALITATIVE DATA COLLECTION METHODS

Quantitative Methods



Qualitative Methods



# Collecting Data



## Qualitative methods

## Quantitative methods

|                                                                                    |                                                           |                                                               |
|------------------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------|
| <b>Aim</b>                                                                         | Complete, detailed description                            | Classify & count; construct models                            |
| <b>Research approach</b>                                                           | <b>Inductive:</b> only know roughly what to look for      | <b>Deductive:</b> knows clearly what to look for              |
| <b>Data</b>                                                                        | Words, pictures, objects:<br>Interpretation of Data       | Numbers & statistics<br>Statistical analysis of data          |
|  | Rich data allows to encompass unexpected outcomes         | Able to test hypothesis and to generalise                     |
|  | Time consuming to analyse and less able to be generalised | May miss contextual details – answers only to questions posed |

# Collecting Data



## **Qualitative data sources**

Focus Group Discussions

Self-Recording Video

Key Informant Interviews

Observation

...

## **Quantitative data sources**

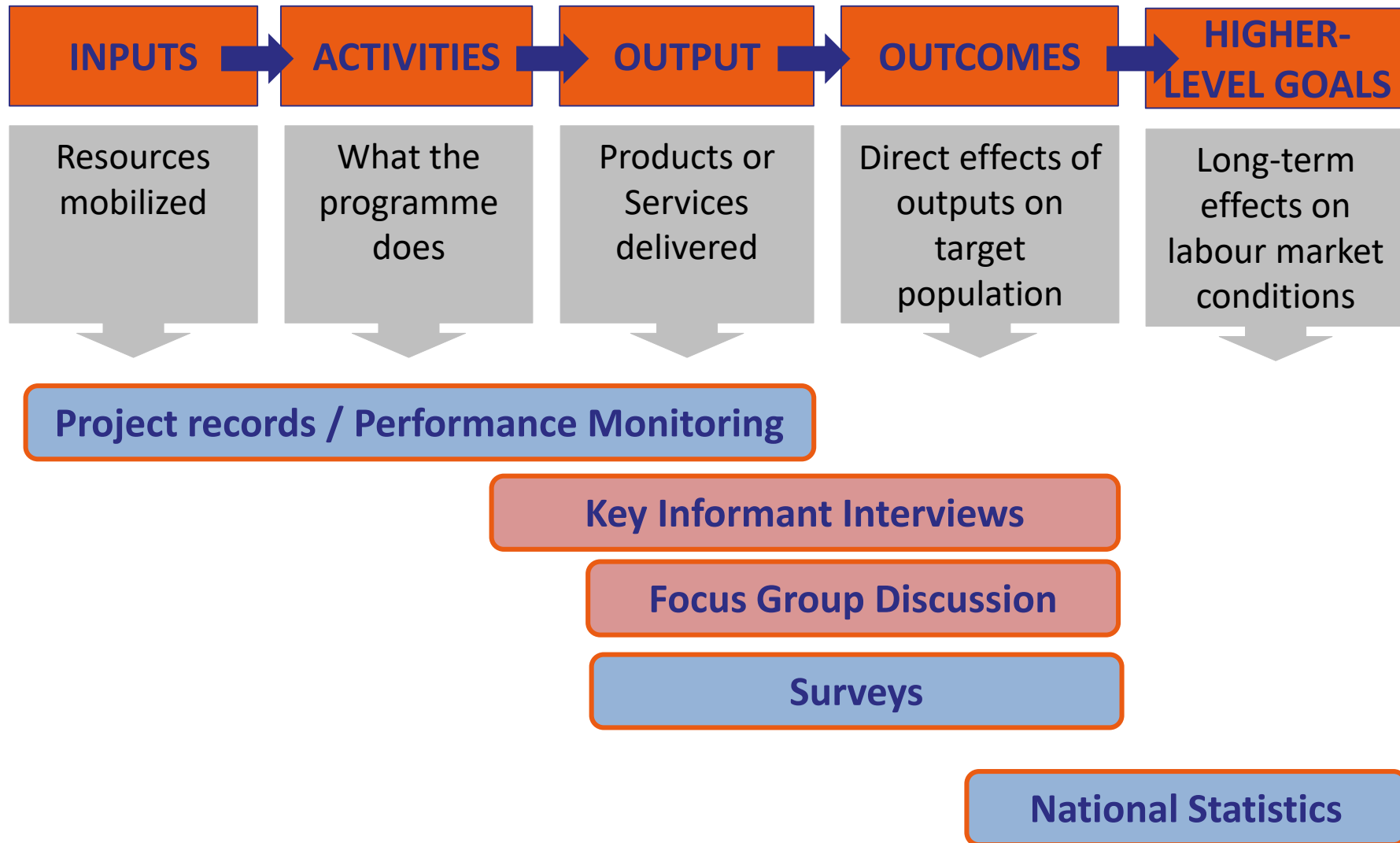
Free Listing

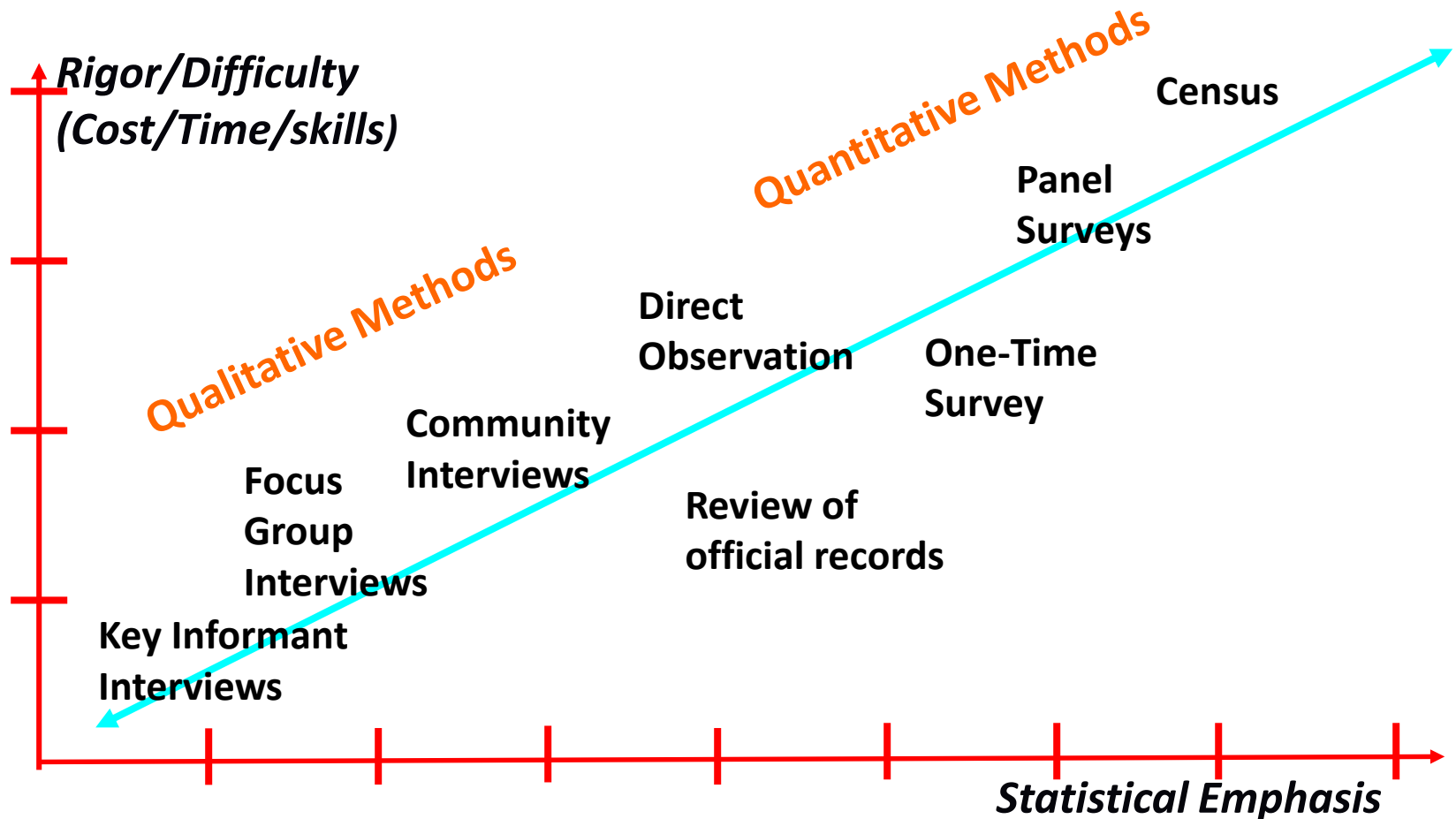
Performance Monitoring Form

Close-ended surveys

Secondary Sources

# Results Chain and Data Collection tools





⇒ Choice depends

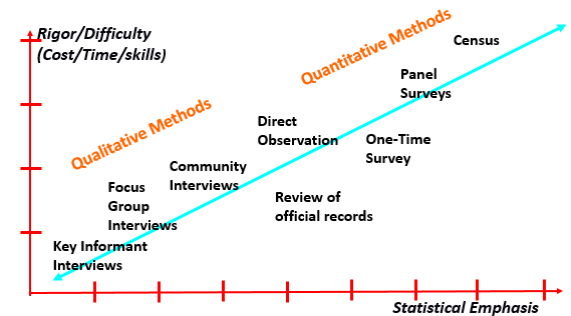
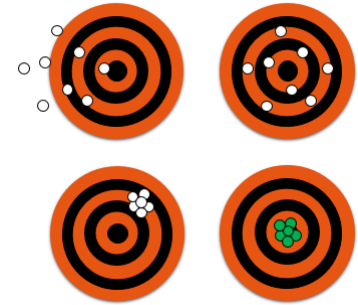
1. Issue to be examined
2. Quality of the information needed
3. Time frame in which information is needed
4. Budget

# Using Mixed Methods

- Combining both qualitative and quantitative methods to enrich the research findings
- Different rationales / types of mixed methods:
  1. **Triangulation:** study the *same* phenomenon to improve reliability of results and validate findings
  2. **Complementarity:** seeking clarification of the quantitative results from qualitative methods
  3. **Development:** using quantitative data to build on or explain initial qualitative findings

# Key Lessons: Collecting Quality Data

- Data collection is about obtaining **valid** and **reliable** estimates for our indicators on all levels of the results chain
- Among the plethora of data collection methods out there pick the one(s) the suit your programme best
- Combine (qualitative and quantitative) methods to double check and enrich findings



## 5. SAMPLING STRATEGY

### Why use samples?

- Because we can't clone.
- Because we don't have enough money or time to study the entire population.
- Because we want to understand the impact of a program among a real population.

# What do we want to know?

- Experiments are expensive. Do we have a chance of finding an impact?
- How big a sample do we need in order to be likely to be able to detect an impact?
- How do different experimental designs affect our chance to find an impact?

# The sample frame

- Start with the population of interest (the group you want to learn about or to which you want to apply general lessons):
  - General population
  - Poor households
  - Young women
- Define the sub-population that you can in principle study (those you have access to)
  - Young women in region x, at time  $t$ ....
- Make a list of those that you can actually contact
  - Names and contact information

*This is your sample frame*

# The sample frame

- Recent ready-made lists provide the ideal sample frame:
  - Census
  - Telephone book
  - Voter rolls

But what if you don't have one of these?

- Conduct an initial enumeration – go into your study areas and make your own lists.

# The sample frame

**So what if we don't have the frame?**

**In principle:**

- with a sampling frame, you can generalize to the sub-population from which you drew the sample frame;
- without the sample frame, you can't generalize beyond the sample population.

# Sampling

## Selecting a sample from the sample frame:

- Probability samples (with known probabilities of selection)
  - *Simple Random* samples (choose randomly from entire list).
  - *Systematic* samples (choose first one randomly, then every  $n^{\text{th}}$  person from the list).
  - *Stratified* samples (divide the population into unique groups  $i$  such that  $\sum_{i=1}^I n_i = N$  and sample within each group).
  - *Cluster* samples (randomly sample groups [census tracts, schools] first, then measure all units within the selected groups).

# Sampling

- **Selecting a sample from the sample frame:**
- Non-probability samples (subjective judgments)
  - Convenience (observations are selected based on their availability - e.g. clients at a health facility on one day)
  - The most similar cases (observations are selected based on the assumption that they represent similar conditions)
  - Typical cases (observations are selected based on available information that indicates if they are useful and not extremes)

# Annex:

## Overview on data collection instruments

# Focus Group Discussion

## What:

- *Collect* in depth info about an issue from a small group (8-15) of people
- *Promote* discussion among group members
- *Provide* understanding of why changes are taking place

## How:

- Moderator knows what to discuss & context of intervention
- Carefully think about **whom** to recruit for **what** kind of questions



# Focus Group Discussion

We invited you all to join our focus group because you were the *easiest to track down* and therefore should provide us with *the best feedback*



# Key Informant Interview

## What:

- Loosely structured interviews with individuals with first-hand knowledge of the intervention
- Select 2-5 individuals, such manager of youth centre, mentors,...

## How:

- List of issues to be discussed but resembles more a conversation than an interview
- Pick a diverse set of key informants to obtain broad range of perspective



# Observation

## What:

- Quick assessment of activity effectiveness through field visits or (informal) observations with programme participants / stakeholders

## How:

- Undertaken during regular work
- Prepare checklist before visit!
- Be familiar with intervention and situation on the ground



# Self-recording video

## What:

- A group of programme participants is asked to answer a set of (pre-selected) questions on a flipchart
- What is most important for participants?

## How:

- Participants are given full control over the camera and should not be interrupted



# Free Listing

**What:**

**Main labour market difficulties faced by youth living in X**

|     | Frequency | Percent |
|-----|-----------|---------|
| 1.  |           |         |
| 2.  |           |         |
| 3.  |           |         |
| 4.  |           |         |
| ... |           |         |

**How:**

- Ask participants about all items they can think of that relate to the topic
- Plot frequency and position of responses
- Highly cost effective!

# Performance Monitoring Forms

## What:

- Collect quantitative information at output level, e.g.
  - **Quantity:** Number of services delivered
  - **Take-Up:** Number of beneficiaries reached
  - **Access:** Characteristics of beneficiaries (socio-economics background, age, gender,...)

## How:

Weekly Report of \_\_\_\_\_ (Date) \_\_\_\_\_ (Facilitator) \_\_\_\_\_ (Location)

1. Participants

|                            |            |             |             |
|----------------------------|------------|-------------|-------------|
| Number of participants     | _____ boys | _____ girls | _____ total |
| Number of NEW participants | _____ boys | _____ girls | _____ total |
| Age range of participants  | from _____ | to _____    |             |

2. Programme of the day

|                            |          |
|----------------------------|----------|
| Improving my Business Plan | Comments |
| -                          |          |
| -                          |          |
| Calculating Living Costs   | Comments |
| -                          |          |
| -                          |          |

# Surveys

## What:

- Collects data from a large number of people, using a *standardised* set of questions
- Gives quantifiable, reliable data on a wide range of issues which can be statistically analysed.
- Surveys are often resource intensive!

## How:

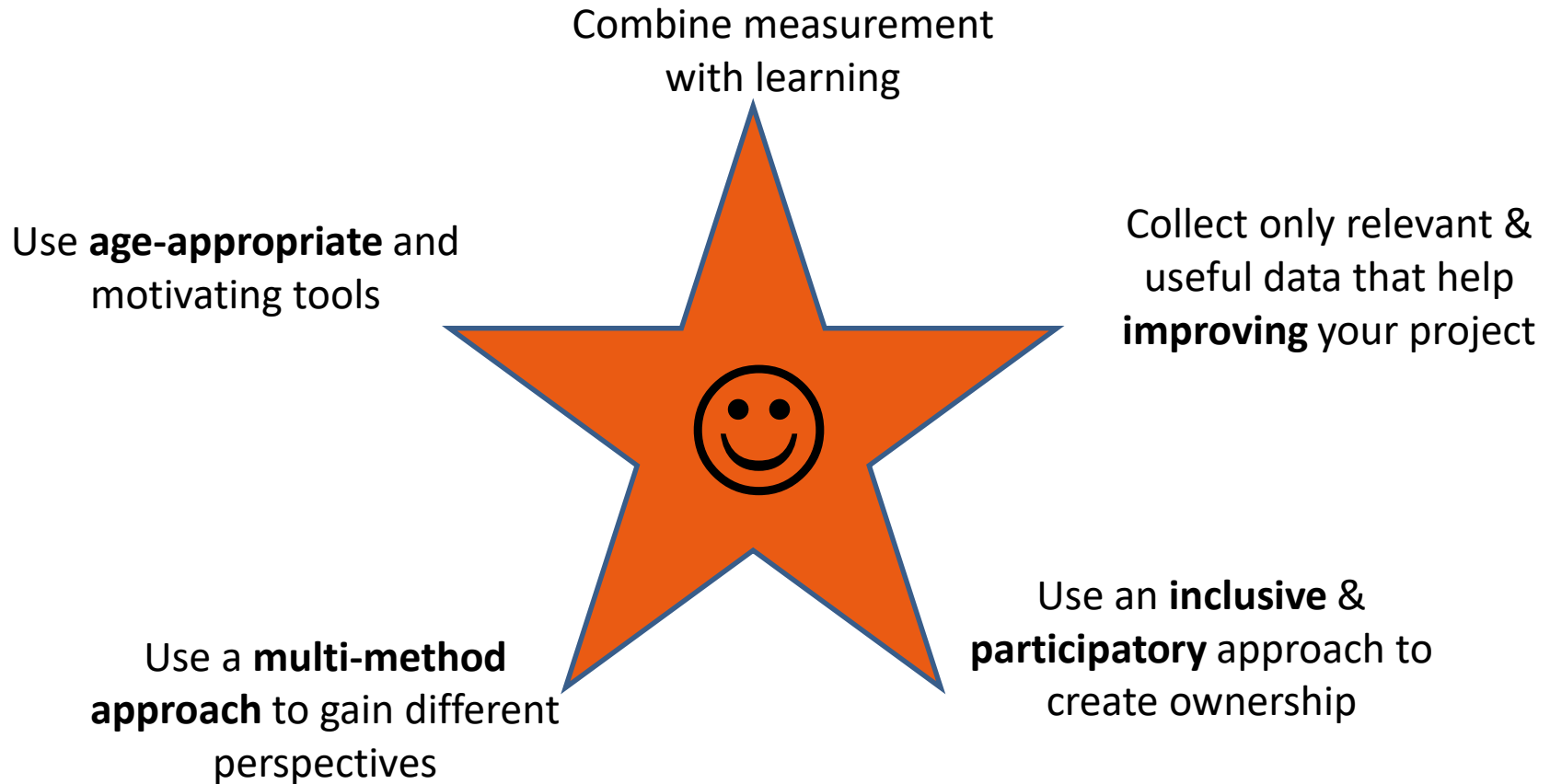
- Keep it simple, clear, easy and short
- Compare your survey (questions) with other research: there is a lot of guidance on (validated) questions out there...!
- More on survey design: Lecture 7

# Secondary Data

## What:

- Data gathered for official purposes for example by national statistical offices:
  - Household survey data: info about living standards, labour market conditions, health, demographics,...
  - Census data: typically only limited amount of information but covers potentially the whole population
  - Facility survey data: data at the level of service provision (schools/ vocational training centres/ ...)
- Can provide important background/control variables
- Can be useful for *planning* an intervention (understanding the context better)
- Attention: check data quality!

# When choosing your data collection tools...



# Further Resources

## Proprietary

---

|       |                                                                     |
|-------|---------------------------------------------------------------------|
| Stata | <a href="http://www.stata.com/">http://www.stata.com/</a>           |
| SAS   | <a href="http://www.sas.com/">http://www.sas.com/</a>               |
| SPSS  | <a href="http://www.spss.com/">http://www.spss.com/</a>             |
| Splus | <a href="http://www.insightful.com/">http://www.insightful.com/</a> |

---

## Free

---

|                                                |                                                                                                                                                         |
|------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| PS: Power<br>and Sample<br>Size<br>Calculation | <a href="http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize">http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize</a> |
| Optimal<br>Design<br>Software                  | <a href="http://sitemaker.umich.edu/group-based/optimal_design_software">http://sitemaker.umich.edu/group-based/optimal_design_software</a>             |

---

## Web-based, interactive

---

<http://statpages.org/>

---

[http://www.dssresearch.com/toolkit/spcalc/power\\_a2.asp](http://www.dssresearch.com/toolkit/spcalc/power_a2.asp)

---