Quantitative Data Analysis
For
Development Evaluatation

National Evaluation Capacities (NEC) Conference 2019
No One Left Behind

Hurghada, Egypt

October 20-24, 2019.

Bassirou Chitou, Ph. D.

Presentation Outline

Module I: Basic Concepts

Module II: Summary Statistics Essentials

Module III: Bivariate Analysis and Hypothesis

Module IV: Data Visualization Basics

Module I

Basic Concepts and Dummy Tables

Basic concepts

- What are the **evaluation questions** (EQ)?
- What are the **evaluation hypothesis** (**EH**)?
- What is the **main outcome (s)** ?
- What are the **covariates**?

Dummy Table: Definition and Example

Definition

- “blank mock tables”, or “blank table shells”;
- variable names;
- labels of statistical measures;
- absolutely NO data;
- constructed before data collection

Example

Table I: Participants Sociodemographic Traits

Sociodemographic Characteristics	n	Percent (%)
Gender		
Male		
Total		

How Useful Are Dummy Tables?

True Merits of Dummy Tables

- Template for systematic steps in the analysis
- Ensure correct data were collected
- help to visualize the data in relationship to the evaluation overall goal
- **help you test the evaluation hypotheses**
- help you stay focused on relevant analyses
- powerful communication tool
- Advance planning tool for various analysis
- centralized record of analyses, results, and decisions

Basic Types of Dummy Tables

- Table of participants' baseline socio-demographic characteristics;
- Table of bi-variate analysis of main outcome and key covariate(s);
- Table of subgroup analysis, for example, male vs. female;
- Table of regression analysis or other models building.

Go to

https://d.docs.live.net/43296abb124de8a9/Documents/UNDP_Egypt/workshop/word/nss_sam_dummy_tables_oct_21_2019.docx

How Do We Analyze Data Effectively?

Data Analysis Steps

Step 1: Prepare the data

Step 2: Describe your sample

Step 3: Assess “Difference” and “ Significance”

Step 4: Explore relationships

Step 5: Built meaningful models

Step 6: Organize and Present Findings

Step 7: Validate Findings with Key Stakeholders

Data Analysis Planning Worksheet (DAP)

DAP Worksheet Template

Resources	What you have	What you need	How to get what you need or work within resources limitation
Funding			
Time			
Staff			
Materials and equipment			

DAP Features

- good communication tool;
- help secure the necessary resources;
- ensure your accountability.

Module II

Descriptive Analysis: Summary Statistics

Level Of Measurement

Definition

- scale that defines and identifies a given variable;
- determines the appropriateness and the use of a certain statistical method.

Six Types of Level Of Measurement

- Binary;
- Nominal;
- Ordinal;
- Interval;
- Ratio;
- Likert scale

Binary and Nominal Variable

Binary

- 2 unique values or categories;
- Puts each unit in one and only category
- Sex: male / female
- Did you eat today: yes / no

Nominal

- 2 or more distinct categories or classes;
- puts each unit in one and only one category;
- marital status: single / married /separated/divorced/widowed

Interval and Ratio

Interval

- “difference” or “interval” makes sense;
- “division” or “ratio” does NOT;
- “0” does NOT mean “Absence”;
- Example: Temperature

Ratio

- BOTH “difference” and “ratio” make perfect sense;
- “0” = “ABSENCE;
- Example: Age, height, income, revenue.

Likert Scales

Agreement Scale

- Measure respondent's opinion on a particular topic;
- Extent to which participant "agrees" or "disagrees";
- Extent of which a respondent is "satisfied" or "dissatisfied";

Rating / Ranking

- Ask participant to rate or rank a particular statement;
- "On a scale of 1 to 5 how would you rate the WFP Food Assistance you received in the past 3 months"

Importance of Level of Scale

- The scale of measurement determines
- the correct statistical analysis;
- The inferences or conclusions that may or not be drawn

Measure of Central Tendency

- Single value describes **center** of the data;
- Characterizes **typical behavior** of the data;
- facilitates **comparisons** between data.

Measures of Central Tendency

Central Tendency Measures

Mean:

- Arithmetic average;
- Easy to use;
- Most popular.
- Symmetric distribution
- Affected by extreme values: NOT robust

Mode:

- Most Frequent value;
- Highest Frequency value;
- Skewed distribution
- Robust against extreme values

Median:

- Ordered data;
- Middle value
- Skewed distribution.
- Robust against extreme values

Measure of Variation

- Describes the extent to which the data is **spread out**, **stretched** or **squeezed** around the central tendency value.
- Characterizes **dispersion** of the data;
- Help understand how individual scores or values behaves;
- enhances **comparisons** between data.

Measures of Central Tendency

Measures Of Variation

Standard Deviation (SD):

- Describes how far or close a value is from the mean;
- Square root of the variance;
- Small is good; Large is of concern

Interquartile Range (IQR):

- Difference between Third Quartile (Q3) and First Quartile Q1;
- Contains approximately 50% of the data.

Coefficient of Variation (CV):

- $CV = {}^{SD}/_{Mean}$; expressed as %;
- The higher the CV, the greater the dispersion;
- The lower CV, the more precise the estimate;
- used to compare 2 different surveys;
- compare variability between 2 distributions.

Module III

Bivariate Analysis & Hypothesis Testing

Go to

https://d.docs.live.net/43296abb124de8a9/Documents/UNDP_Egypt/workshop/word/intro_to_quant_anal_undp_nec_2019_sept_26_2019.docx

Module IV

Bivariate Analysis & Hypothesis Testing

Go to

https://d.docs.live.net/43296abb124de8a9/Documents/UNDP_Egypt/workshop/word/intro_to_quant_anal_undp_nec_2019_sept_26_2019.docx

Thank You A Million