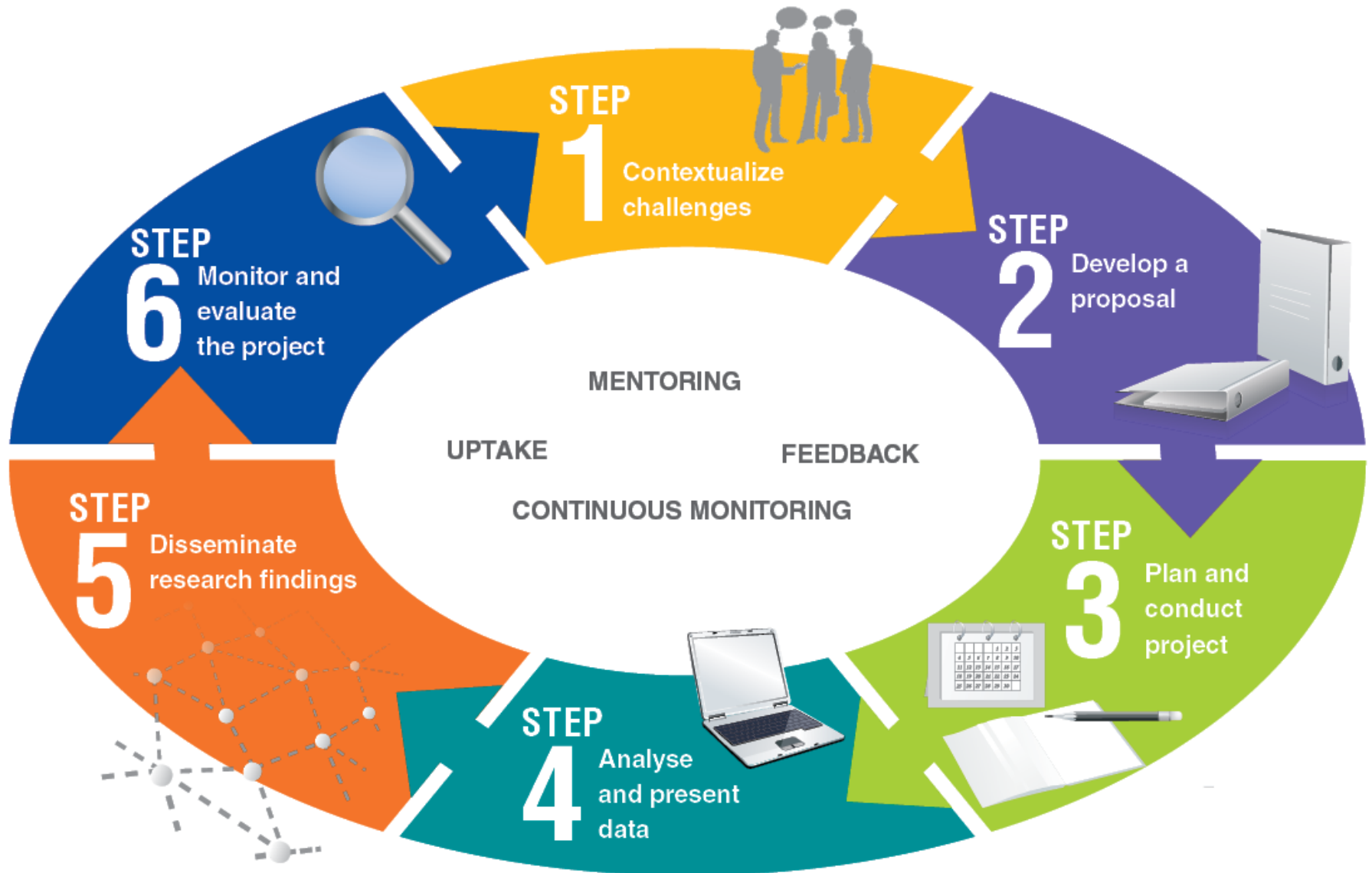# Module 4:
Data analysis and presentation

Six steps in the IR process

Quantitative vs Qualitative

What are the differences between quantitative and qualitative research?

Research questions

Methodological differences

Data analysis

Comparing qualitative and quantitative approaches

	Qualitative	Quantitative
Social theory	Action	Structure
Methods	Observation, interview	Experiment, survey
Question	What is x? How? Why? (classification)	How many xs? (enumeration)
Reasoning	Inductive	Deductive
Sampling	Theoretical	Statistical
Strength	Validity	Reliability

Pope and Mays (1995). Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ:* 311; No. 6996

Module 4a:
Quantitative data analysis and presentation

Presentation outline

Expected outcomes

Key concepts

Data analysis plan

Quantitative data analysis

Data management

Learning Objectives & Expected outcomes

Able to:

Describe data analysis planning processes

Understand appropriate statistical measures

Understand data management approaches

Appreciate the importance of tailored / audience sensitive data presentation

Key concept 1: Data analysis plan

Designing analysis for use IR aims to:

Understand the implementation processes

Communicate the implementation process to stakeholders

"Emphasis on simplicity and interpretability"

Key concept 1: Data analysis plan

Designing analysis for use

Different stakeholders need different information:

Lay people?

Community leaders?

Local government/health service leaders?

Civil society and media personnel?

National policy-makers?

"Emphasis on simplicity and interpretability"

Key concept 1: Data analysis plan

Designing analysis by purpose focuses on the objective of the analysis:

Effectiveness

Efficiency

Equity

Sustainability

Key concept 1: Data analysis plan

Data presentation formats

Data reporting should be presented in both textual and visual formats, such as:

- Tables
- Diagrams
- Graphs
- Infographics
- Maps

Provider education expressed as frequency table

Level of education of private providers	Frequency
Illiterate	106
Basic literacy	74
Primary school certificate	57
Secondary school certificate	11
Higher level qualification	2
Total	250

Joint frequency distributions for two or more variables

Highest level	Men	Women	All
Illiterate	42	64	106
Basic literacy	45	29	74
Primary school certificate	32	25	57
Secondary school certificate	8	3	11
Higher level qualification	1	1	2
Total	128	122	250

Provider education presented as proportion, percentage and cumulative %

Level of education	Proportion	Percentage	Cumulative percentage
Illiterate	0.424	42.4	42.4
Basic literacy	0.296	29.6	72.0
Primary school certificate	0.228	22.8	94.8
Secondary school certificate	0.044	4.4	99.2
Higher level qualification	0.008	0.8	100.0
Total	1.000	100.0	

Row percentages

Highest level	Men	Women	All
Illiterate	39.6	60.4	100.0
Basic literacy	60.8	39.2	100.0
Primary school certificate	56.1	43.9	100.0
Secondary school certificate	72.7	27.3	100.0
Higher level qualification	50.0	50.0	100.0
Total	51.2	48.8	100.0

Column percentages

Highest level	Men	Women	All
Illiterate	32.8	52.5	42.4
Basic literacy	35.2	23.8	29.6
Primary school certificate	25.0	20.5	22.8
Secondary school certificate	6.3	2.5	4.4
Higher level qualification	0.8	0.8	0.8
Total	100.0	100.0	100.0

Provider education expressed as a bar chart

Level of education
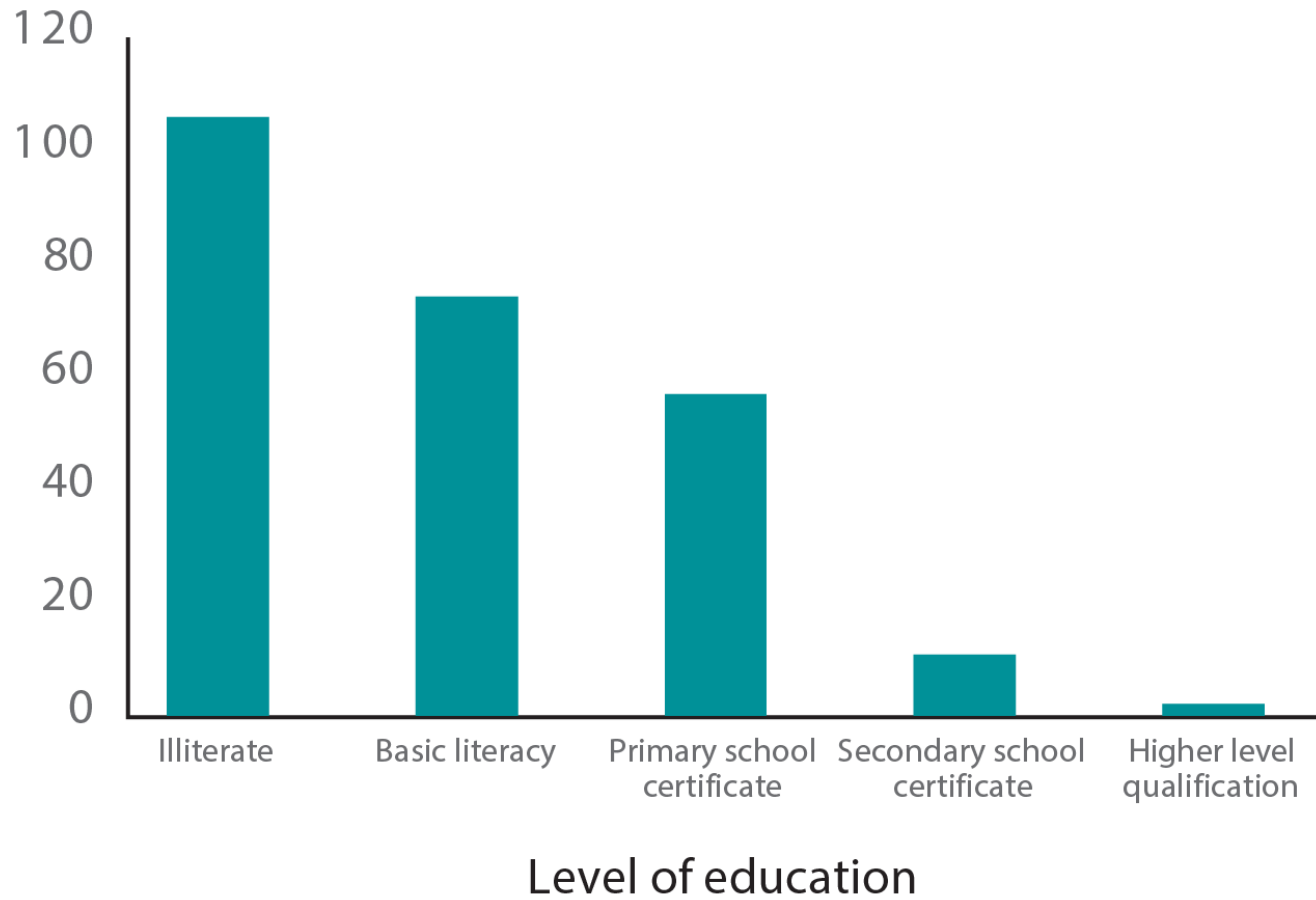
Bar chart for two variables

Level of education

Legend: Men, Women

Categories: Illiterate, Basic literacy, Primary school certificate, Secondary school certificate, Higher level qualification
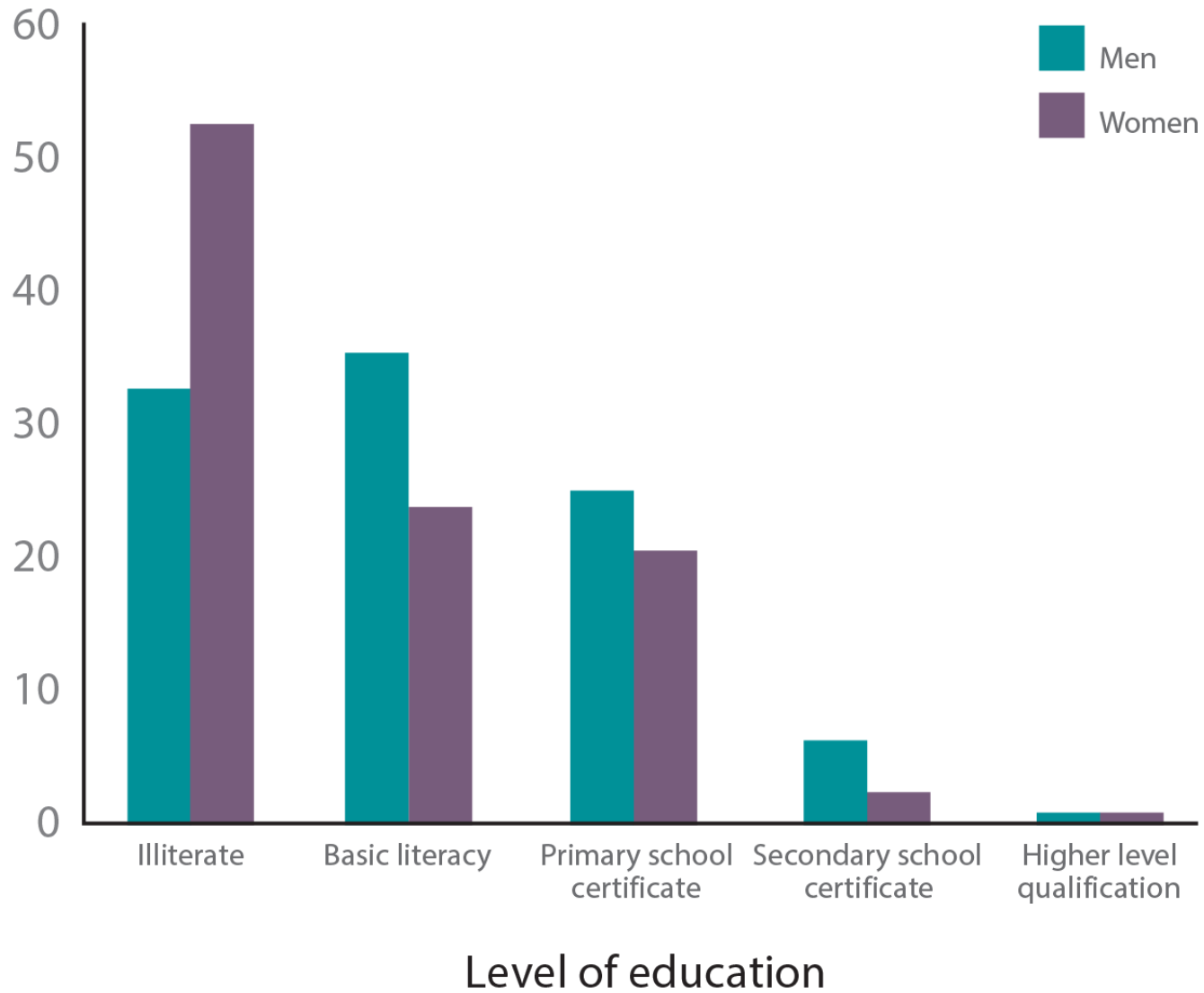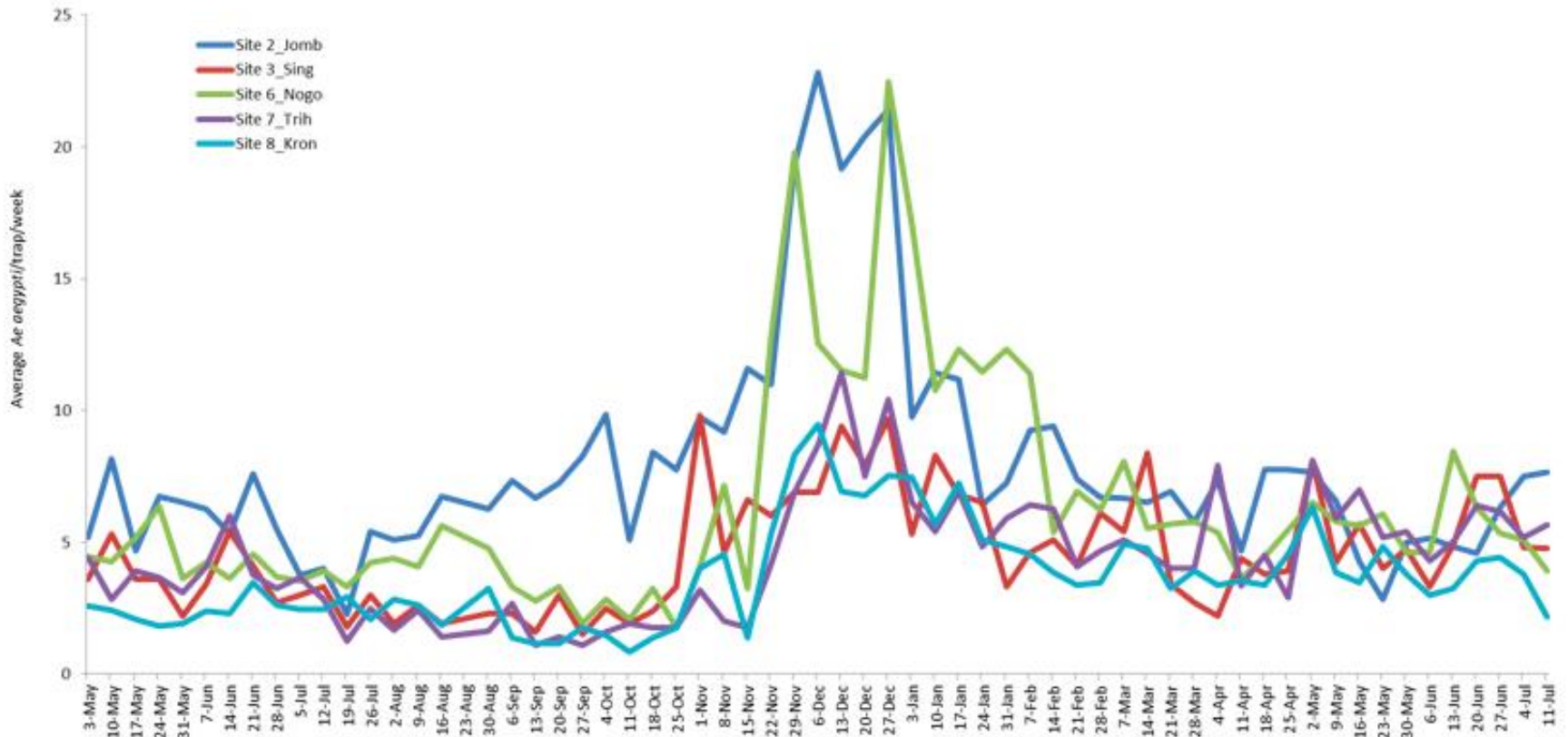
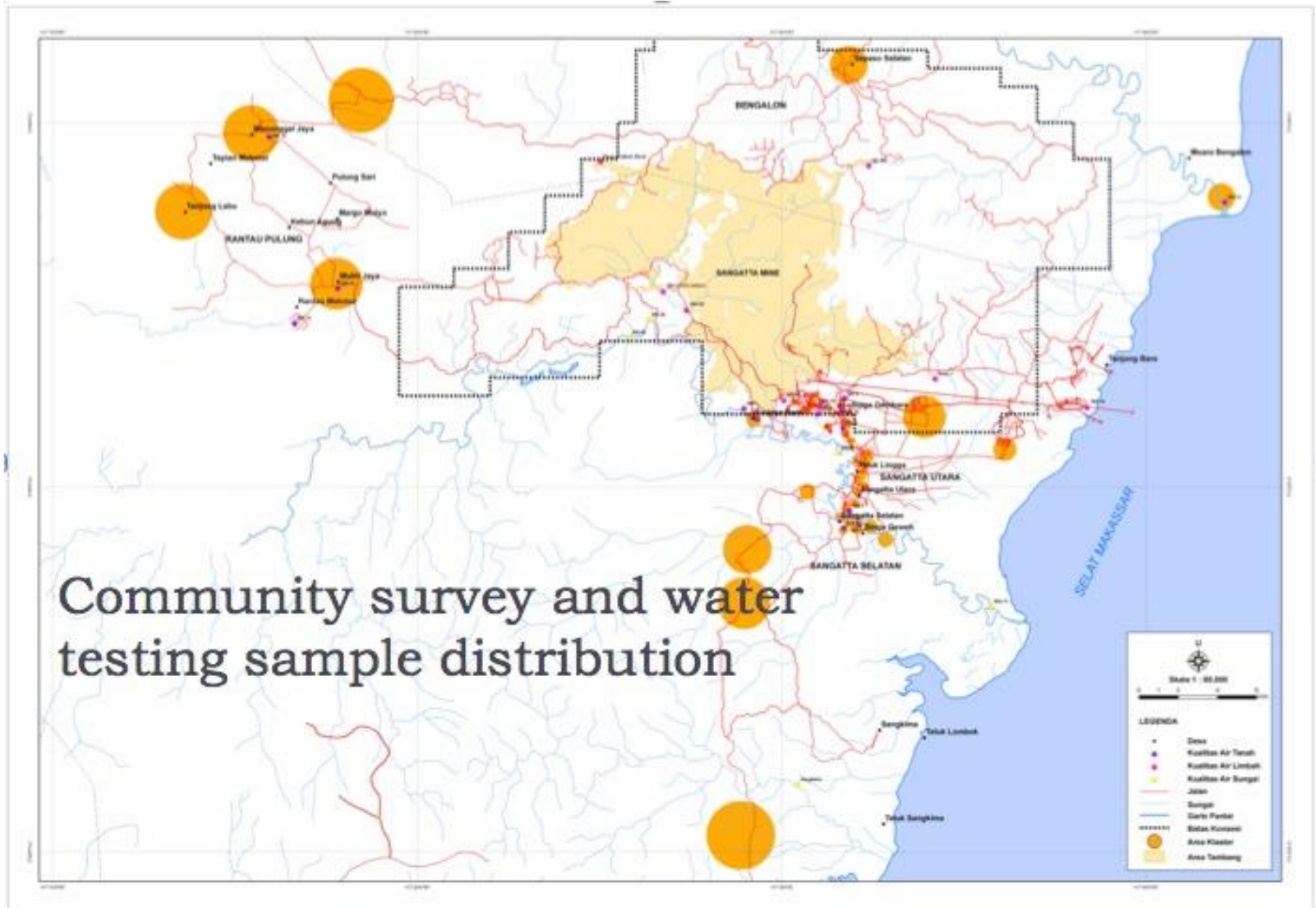# Provider education presented as a pie chart
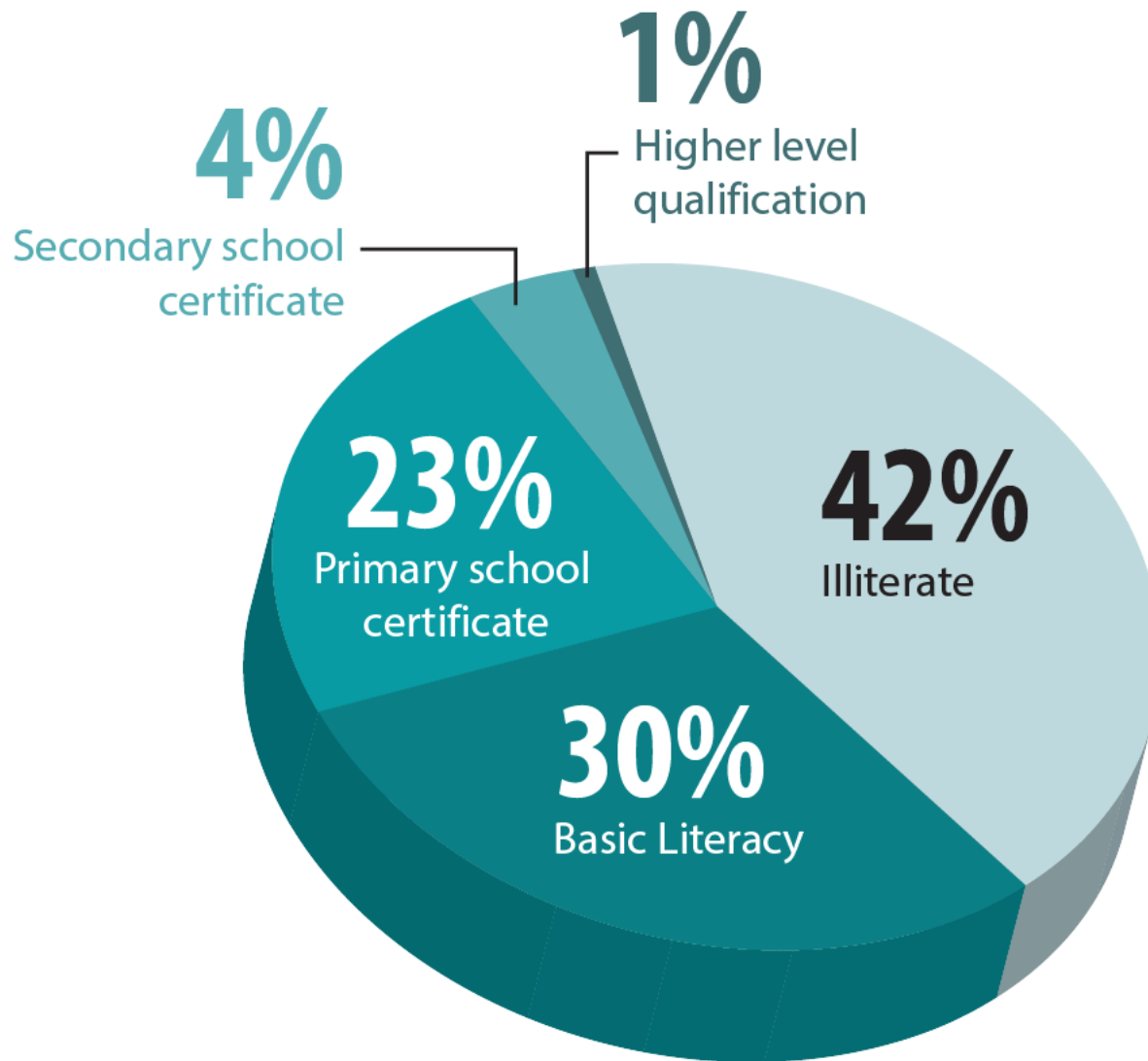
Line graph for trend analysis

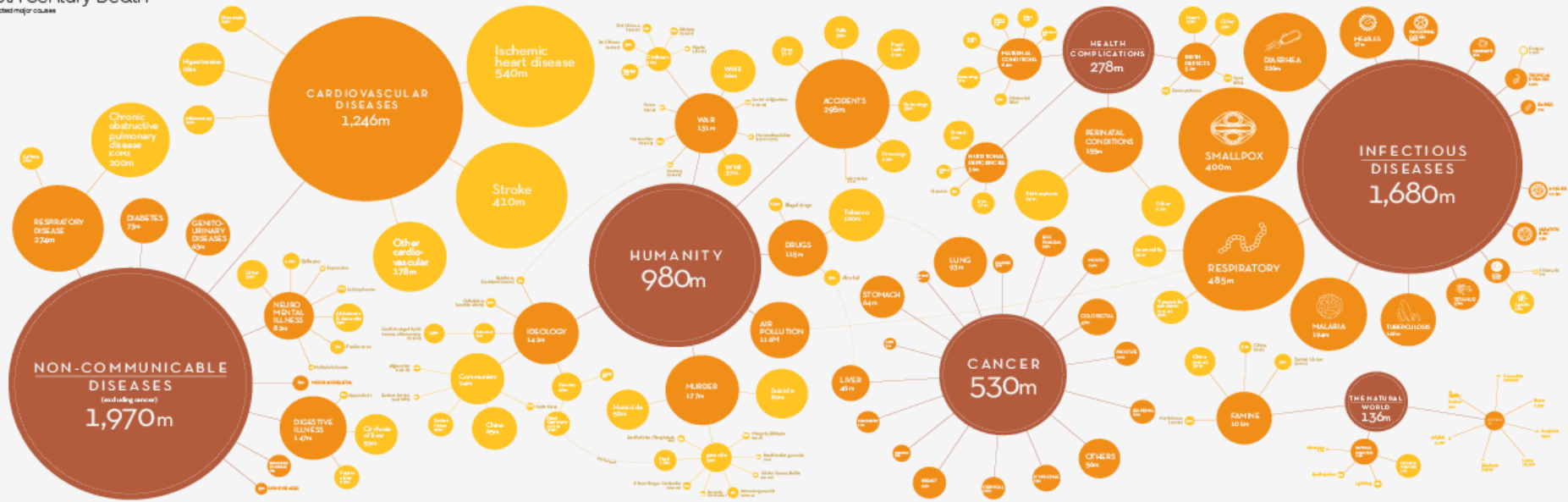Average *Ae. Aegypti* population per week in 5 field study sites

In what other format(s) can this data set be presented?

Community survey and water testing sample distribution

20th century death

Potential tax revenue from drugs

Drug Deal?
Potential tax revenue from legalising drugs in the UK

Most users

Cannabis "Skunk" £461m tax revenue
Cannabis Resin £381m
Cocaine £182m
Ecstacy £8m
Amphetamine £112m
LSD £40m
Heroin £212m
Crack £474m
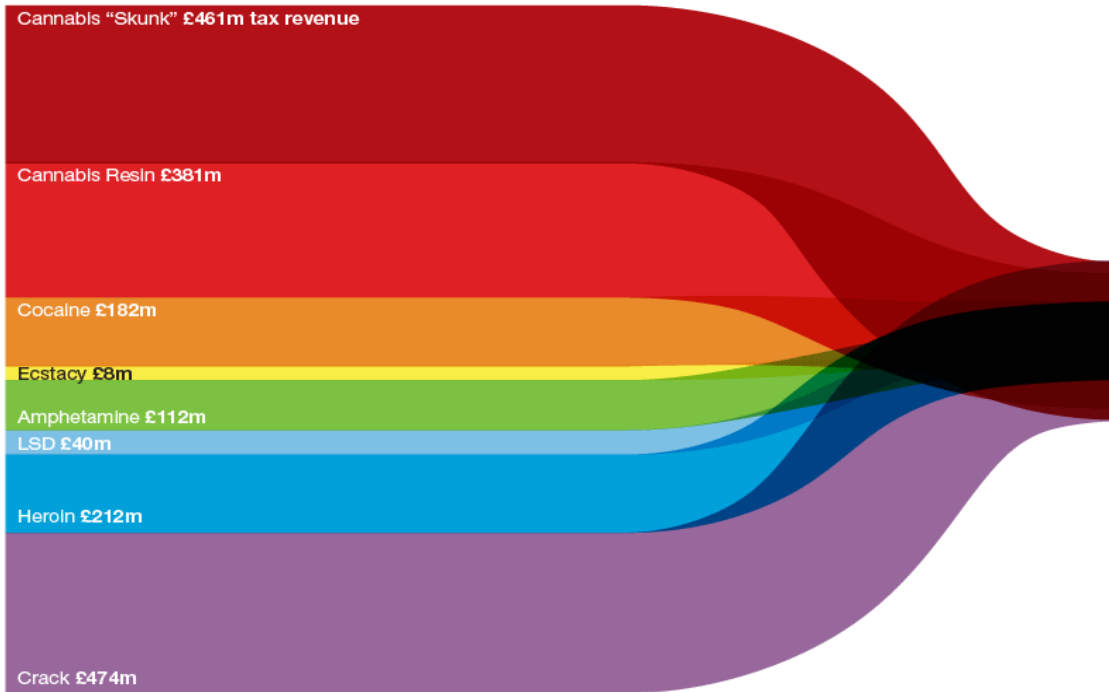
Least users

£5bn Shortfall in university funding 2012

£9bn Government income from alcohol

£5bn Total revenue from legalization of drugs

+ £1.8bn tax
+ £0.3bn VAT
+ £1.7bn savings from societal costs
+ £1.0bn income tax from working users
+ £0.5bn inflation adjustment
− £0.3bn additional societal costs

• law enforcement
• drug-related crime

• new addicts
• health consequences
• "purity police"
• licensing

£10bn government income from tobacco

£7bn Bailout for Irish republic

David McCandless & Joe Swainson
v 1.0 // Dec 10 // InformationIsBeautiful.net

source: Independent Drugs Monitoring Unit (idmu.co.uk)
data: bit.ly/drug_deal

Source: Information is beautiful site
http://www.informationisbeautiful.net/visualizations/
drug-deal-potential-tax-revenue-from-legalized-narcotics/
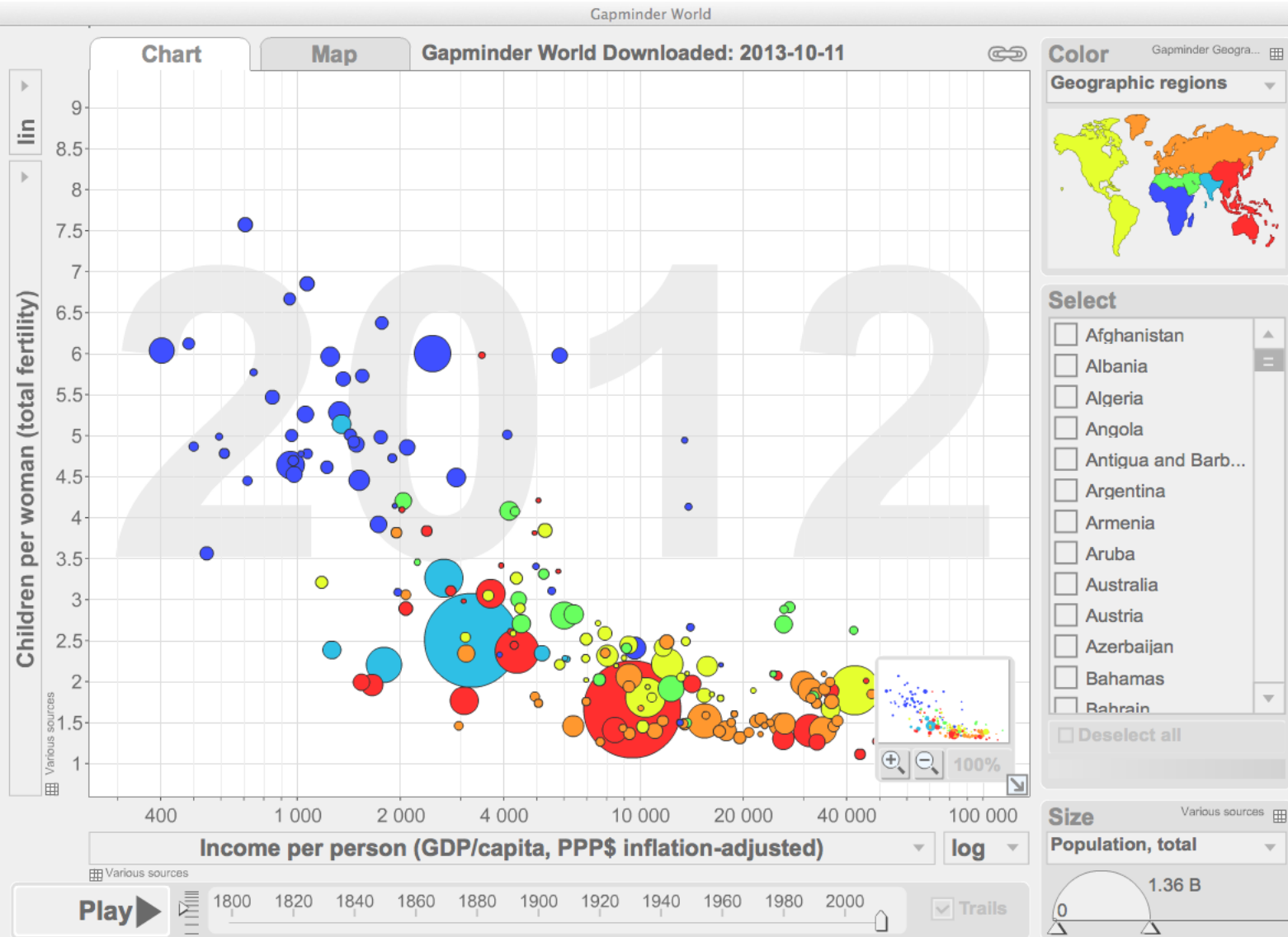
20 oz cola = 16 tsp sugar

Interactive graph

Reflection activity

In your project, discuss the results of the study that you need to disseminate and format of data presentation you will use for different stakeholders.

Key concept 2: Quantitative data analysis

Depending on research question:

Descriptive vs analytic study?

Analytic study, what to find?

Association

Causality

Statistical difference

Key concept 2: Quantitative data analysis

Variables in quantitative analysis are usually classified by their level of measurement:

Rational

Interval

Ordinal

Nominal

Key concept 2: Quantitative data analysis

Descriptive statistics

Distributions and summary measures

Defining intervals for frequency distributions

Frequency distribution and summary statistics

Measures of variation

Key concept 2: Quantitative data analysis

Distributions and summary measures
Advantages of frequency distributions:

useful for all types of variables

easy to explain and interpret

presented graphically and in different formats

Key concept 2: Quantitative data analysis

Defining intervals for frequency distributions

Constructing a frequency distribution requires a choice of intervals:

- Ordinal
- Interval
- Rational

Two conflicting objectives when determining intervals:

- Limiting the loss of information
- Providing a simple, interpretable and useful summary

Key concept 2: Quantitative data analysis

Summary statistics and frequency distributions

A powerful and robust form of analysis. Summary statistics usually focus on: overall location of a distribution or extent of variation within a population.
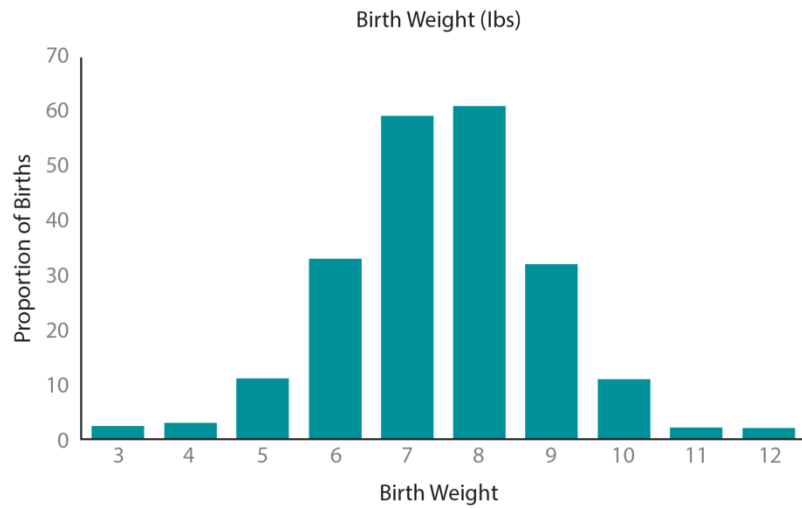
Key concept 2: Quantitative data analysis

Use of mean or median

Mean – the average value

Median – the value in the middle

Use of mean or median

Normal distribution

Birth Weight (lbs)

Skewed distribution

Distribution of land holdings

Key concept 2: Quantitative data analysis

Measures of variation

How much variability?

Low variability

High variability

Key concept 2: Quantitative data analysis

Measures of variation

Choices of measures

Variances

Standard deviations

Alternative measures

Quartiles: divide data into four quarters (Q1 to Q4) – 25% in each

Percentiles: divide the data into two parts

Key concept 2: Quantitative data analysis

Analytical statistics

Group comparison

Association

Causality

Key concept 2: Quantitative data analysis

Measurement scale	Assumption of distribution	Type of group	Analysis
Nominal or Ordinal	-	Independent	Chi square test
	-	Paired	Sign test
Interval or Ratio	Normally distributed	Independent	Independent test
		Paired	Paired test
	Not normally distributed	Independent	Median test
		Paired	Wilcoxon

Key concept 2: Quantitative data analysis

Finding association

Pearson correlation

Ratio/interval scale

Normal distribution of data

Rank correlation

Ratio/interval scale

Non normal distribution of data

Chi Square

Categorical data

Key concept 2: Quantitative data analysis

Causality (regression)

Linear regression

- Continuous variable of both independent and dependent variable
- Normal distribution of data

Logistic regression

- Dichotomous dependent variable
- Continuous and categorical independent variable

Key concept 2: Quantitative data analysis

Causality (regression)

Cox proportional hazard model

Time-dependent outcome (survival model)

Continuous and categorical independent variable

Key concept 2: Quantitative data analysis

Measures of risk

'Risk' and 'odds' used interchangeably, but not the same

'Reduction in risk' is not equivalent to 'reduction in odds'

Key concept 2: Quantitative data analysis

Measures of risk:
The 'denominator problem'

Risk calculation requires calculation of the population at risk

Provide the estimates of both the numerator and denominator alongside any proportion, percentage or risk estimate

Key concept 2: Quantitative data analysis

Sub-group analysis

The outcomes of an intervention may differ among sub-groups.

'Data mining' is useful to formulate new hypotheses but requires great caution in IR.

Reflection activity

In your project, discuss the data analysis that you will do and identify whether the data you are collecting is suitable for the type of analysis you plan.

Key concept 3: Data management

Principle of data management

Data management and study phase

Key concept 3: Data management

Data quality and integrity

Data should be:

  High quality

  Reliable

"No study is better than the quality of its data"

Key concept 3: Data management

Prior to data collection process

ID number
Flow of data collection and handling process
Protocol for quality control
    Checking interviewee response
    Re-interview process
Electronic database development
SOP for data entry process

Key concept 3: Data management

Data collection process

Data collection supervision

Questionnaires/data collection forms storage management

Checking data entry process

Key concept 3: Data management

Post-data collection process

Checking database consistency

Data cleaning

Data coding

Reflection activity

In your project proposal, discuss how to improve the quality of their data management system.

Conclusion

Start from the end

Plan your data analysis according to stakeholders' need for information

Use appropriate statistical tools according to the information needed

Manage your data to ensure the validity of the data collected