



# Introduction to Data Analysis

Greg Keeble  
UNESCO Institute for Statistics



# Overview

- Objective & Audience
- Forms of presentation
- Tables dimensions
- Components of a table
- Layout of a Table
- Presenting Data
- Purpose of graphs & charts
- Components of graphs & charts
- Types of graphs/charts
- Data presentation



# Objective and Audience

- The type of statistics depends on the main objective of the presentation and the target audience
- Consider how best to present the data and indicators:
  - What am I trying to communicate?
  - Who are my audience?
  - What kind of presentation will be most effective?
  - What will help my audience to better understand the data?



# Forms of presentation !

Communicate to your audience the meaning of the data using the summary statistics in an informative and interesting manner that is easy to understand:

- **Tables** are useful for presenting data and statistics in numeric form
- **Charts and graphs** may be used to highlight key patterns and trends in a graphical form
- **Descriptive text** can describe and summarize findings in verbal form



# Dimensions of a summary table

- Summary tables are used to present counts of students, teachers and schools by a categorical variable
- e.g. number of teachers by qualification  
number of students by grade.
- Summary tables can be:
  - simple one-dimensional table  
with one categorical variable
  - multi-dimensional table  
with two or more categorical variables.



# One-dimensional Summary Table

The table shows the number and percentage of student enrollments by grade in high school.

## Student Enrollments by High School Grade

Grades	Count	Percent
Grade 9	813	30.88%
Grade 10	704	26.74%
Grade 11	575	21.84%
Grade 12	541	20.55%
<b>Total</b>	<b>2,633</b>	<b>100.00%</b>



# Multi-dimensional Summary Table

## Student Enrollments by Elementary Grade by Gender

Grades	Gender		Female No.	Female %	Total No.	Total %
	Male No.	Male %				
Grade 1	647	14.0%	555	13.0%	1,202	13.5%
Grade 2	684	14.8%	557	13.0%	1,241	13.9%
Grade 3	637	13.7%	609	14.2%	1,246	14.0%
Grade 4	586	12.6%	565	13.2%	1,151	12.9%
Grade 5	608	13.1%	524	12.2%	1,132	12.7%
Grade 6	528	11.4%	550	12.8%	1,078	12.1%
Grade 7	478	10.3%	481	11.2%	959	10.8%
Grade 8	469	10.1%	442	10.3%	911	10.2%
<b>Total</b>	<b>4,637</b>	<b>100.0%</b>	<b>4,283</b>	<b>100.0%</b>	<b>8,920</b>	<b>100.0%</b>





# Multi-dimensional tables

Multi-dimensional tables:

- can show detailed data patterns and complex relationships
- can become very complicated if too many data values are presented
- need to consider the ability of the reader to understand and interpret multi-dimensional tables
- enable in-depth analysis of the patterns of school participation by presenting data about the distribution of students by grade, age and gender





# Multi-dimensional Summary Table

**Student Enrollments by Elementary Grade by Age and Gender**

Age	G 1		G 2		G 3		G 4		G 5		G 6		G 7		G 8		Total	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F		
<=5	168	147	9	6	2	2												334
6	367	335	147	108	13	11	1	1										983
7	75	47	374	315	129	118	7	6		1		2	1					1,075
8	22	15	99	92	295	290	92	111	19	11		5					1	1,052
9	10	3	31	23	111	115	277	279	93	97	11	20	1					1,071
10	5	3	12	10	48	39	117	91	232	243	88	111	15	16	1	3		1,034
11		2	7	1	20	23	50	46	135	93	194	221	78	92	7	12		981
12		1	3	1	7	7	22	21	67	42	110	104	192	190	70	90		927
13		1			5	3	15	8	37	26	80	60	113	109	178	191		826
14		1	2		1		3	1	15	7	27	13	53	54	110	90		377
15+					1		2		10	4	18	14	25	20	103	55		252
<b>Total</b>	<b>647</b>	<b>555</b>	<b>684</b>	<b>556</b>	<b>632</b>	<b>608</b>	<b>586</b>	<b>564</b>	<b>608</b>	<b>524</b>	<b>528</b>	<b>550</b>	<b>478</b>	<b>481</b>	<b>469</b>	<b>442</b>		<b>8,912</b>



# Components of a table

A properly presented table should include:

- Title
- Headings
- Rows and columns
- Units of measurement
- Degree of accuracy
- Footnotes
- Source of data

Rates of transition to, and participation in, secondary education, 1999 and 2006, worldwide and by region

	Transition rates from primary to secondary education (median)			Secondary education			
	School year ending in 2005			Gross enrolment ratios		Net enrolment ratios	
				School year ending in 1999	School year ending in 2006	School year ending in 1999	School year ending in 2006
	Total (%)	Male (%)	Female (%)	(%)	(%)	(%)	(%)
World	93	92	94	60	66	52	58
Developing countries	88	93	83	52	60	45	53
Developed countries	99	...	...	100	101	88	91
Countries in transition	100	100	99	90	89	83	82
Sub-Saharan Africa	62	66	57	24	32	18	25
Arab States	92	90	93	60	68	52	59
Central Asia	99	99	99	83	91	78	83
East Asia and the Pacific	...	...	...	65	75	61	69
East Asia	91	...	...	64	75	61	69
Pacific	...	...	...	111	107	70	66
South and West Africa	87	90	83	45	51	39	45
Latin America and the Caribbean	93	...	...	53	57	44	40
Caribbean	94	...	...	53	57	44	40
Latin America	92	92	92	81	91	59	71
N. America/W. Europe	99	99	99	100	101	88	91
Central/Eastern Europe	98	98	99	87	88	80	81

Source: UNESCO (2009) EFA Monitoring Report 2009. Paris: UNESCO.

Common heading

Unit of measurement

Title

**Example 4. The use of title, headings and units of measurement**  
Pre-primary enrolment and gross enrolment ratios by region, 1999 and 2006

Heading	Total enrolment			Gross enrolment ratios		
	School year ending in		Change between 1999 and 2006	School year ending in		Change between 1999 and 2006
	1999	2006		1999	2006	
	(millions)	(millions)	(%)	(%)	(%)	(%)
World	112	139	24	33	41	26
Developing countries	80	106	32	27	36	32
Developed countries	25	26	3	73	79	9
Countries in transition	7	7	2	46	62	36

Note: Change is computed using non-rounded figures.

Source: Annex, Statistical Table 3B.

Footnote



# Layout

A clear, well-structured layout makes it easier for the reader to interpret and understand the information in the table:

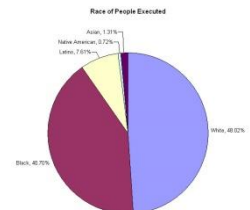
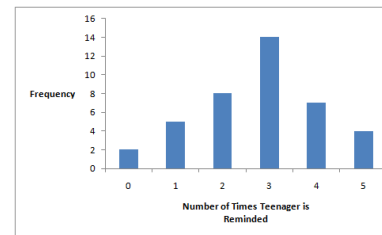
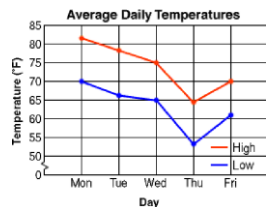
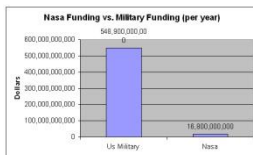
- Font style
- Ordering of rows and columns
- Numbers
- Consistent appearance
- Number the table
- Unnecessary distraction

Note: Complex tables that should logically appear together should be placed in the appendices.



# Graphs and charts:

- represent and summarize statistical information in a visual manner to show patterns and trends in the data
- are useful for highlighting and presenting important information
- an ideal method for presenting statistical information to non-technical audiences

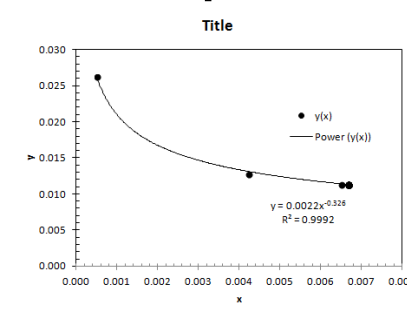
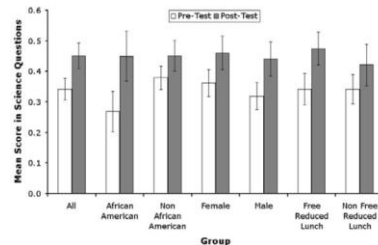




# Purpose of charts and graphs

Charts and graphs are used:

- to visually represent information that cannot be easily read and interpreted from a table
- to show trends and changes in statistical data
- to make comparisons between two different set of data or when making to make predictions and forecasts.

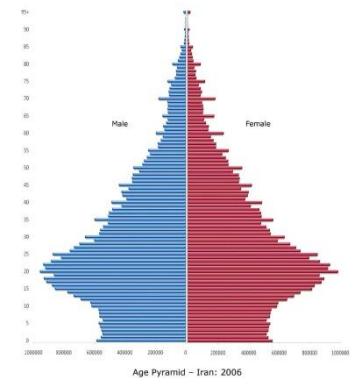
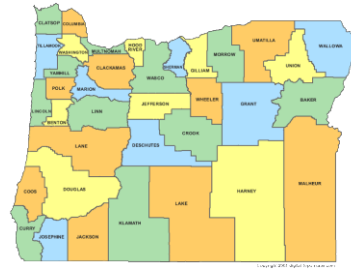




# Advantages of charts and graphs

- easier to understand than a table of numbers
- highlight patterns and trends in the data
- makes comparisons and analysis easy
- representation of data using different types of graphs and charts
- allow for special designs

e.g. age-pyramids  
and thematic maps





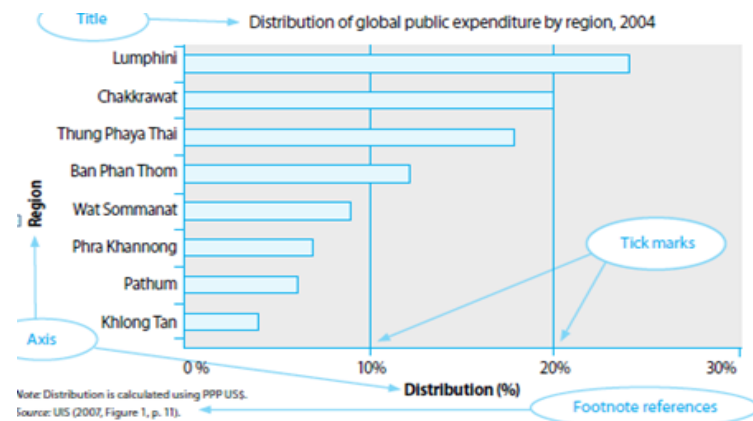


# Components of charts and graphs

To understand and interpret the data represented in a graph or chart :-

Graphs and charts should have:

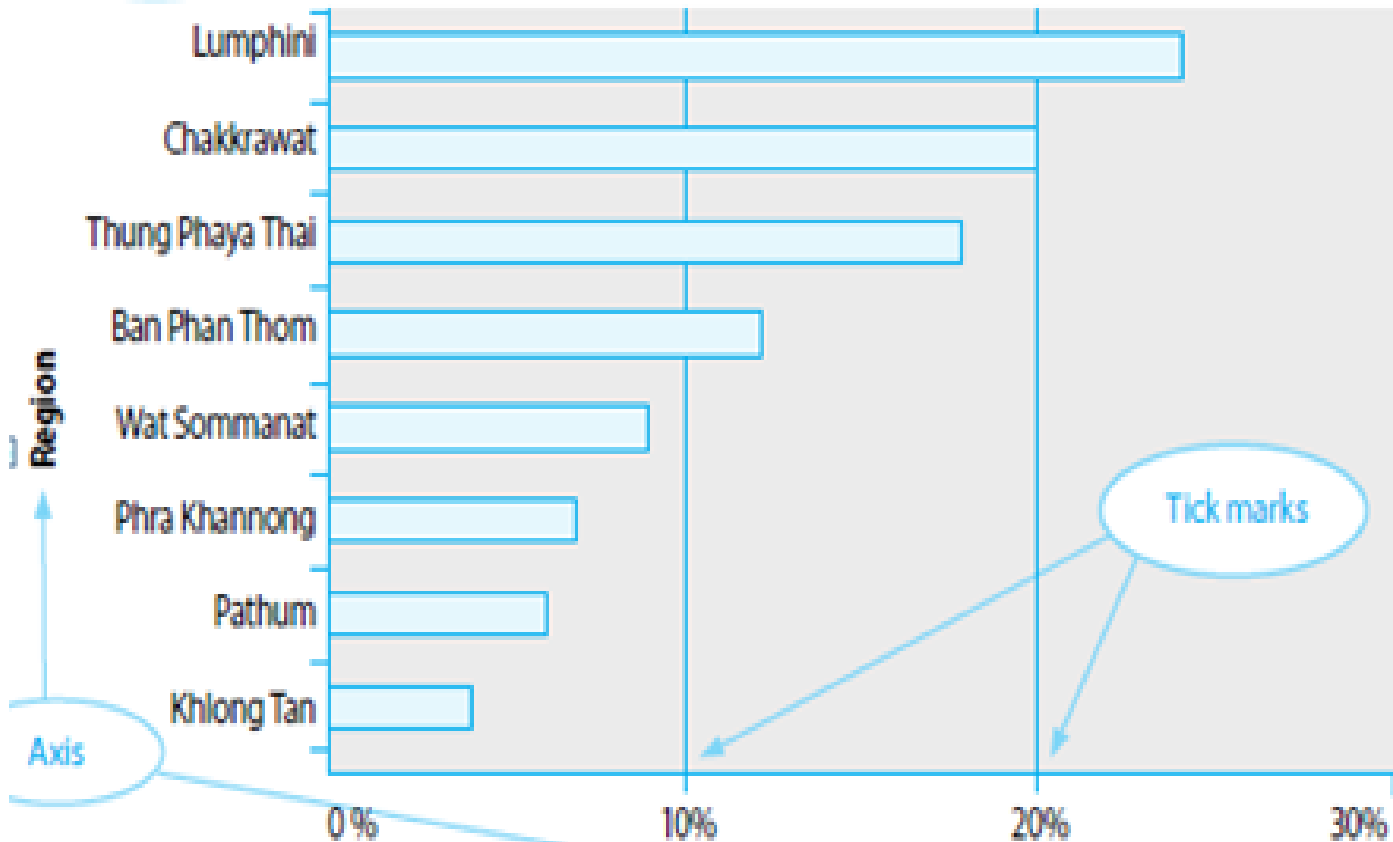
- a title
- axis labels
- labels for subgroups
- footnotes
- references to source data





Title

Distribution of global public expenditure by region, 2004



Region

Axis

Tick marks

Distribution (%)

Footnote references

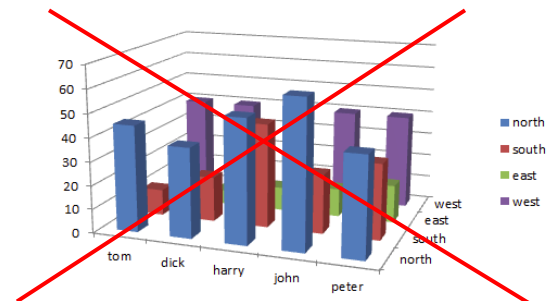
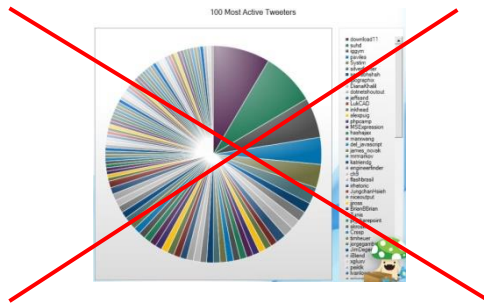
Note: Distribution is calculated using PPP US\$.

Source: UIS (2007, Figure 1, p. 11).



# Characteristics of charts and graphs

- present a key message
- have a clear objective
- use an appropriate type of presentation
- have a simple and clear design



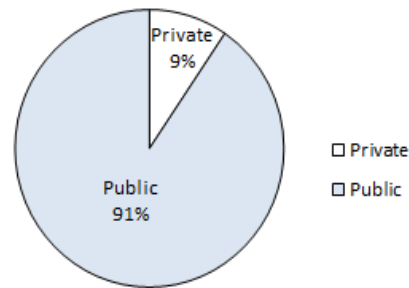


# Types of graphs and charts

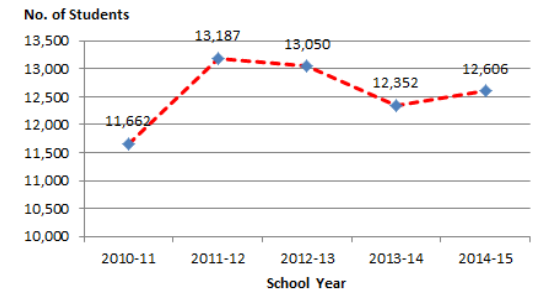
There are many different types of graphs and charts including:

- Pie chart
- Line graph
- Bar chart
- Area graph
- Scatter plot
- Maps

**SY 2014-15 Enrollments by School Type**

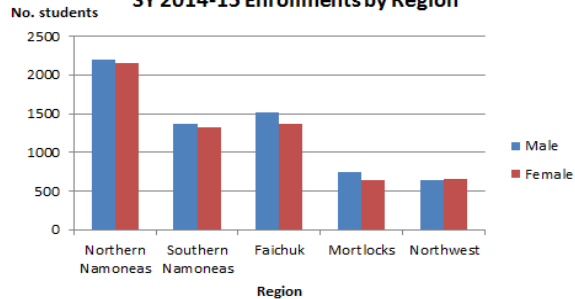


**All Students by School Year 2010-2014**

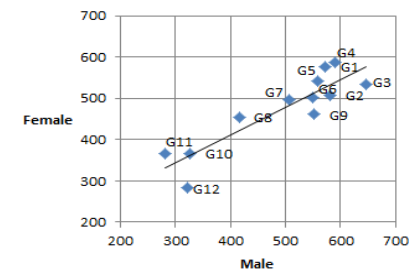


Footnote: Data for the School Year 2010-11 is incomplete  
Source: Chuuk State EMIS System

**SY 2014-15 Enrollments by Region**



**Grade Enrollment by Gender**





# Process of creating a graph or chart

Steps to create a graph or chart:

1. Organize and present data in a table
2. Calculate percentages, ratios and indicators
3. Create graph or chart to illustrate the data

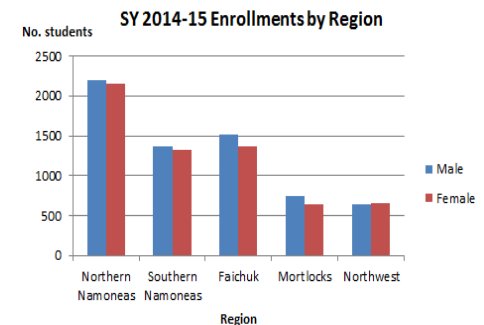
## STEP 1

Count	Gender		Grand Total
	Male	Female	
Northern Namoneas	2197	2159	4356
Southern Namoneas	1367	1327	2694
Faichuk	1509	1369	2878
Mortlocks	751	640	1391
Northwest	636	651	1287
<b>Grand Total</b>	<b>6460</b>	<b>6146</b>	<b>12606</b>

## STEP 2

Percent	Gender		Grand Total
	Male	Female	
Northern Namoneas	34.0%	35.1%	34.6%
Southern Namoneas	21.2%	21.6%	21.4%
Faichuk	23.4%	22.3%	22.8%
Mortlocks	11.6%	10.4%	11.0%
Northwest	9.8%	10.6%	10.2%
<b>Grand Total</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

## STEP 3





# Practical Exercise

- Create Pivot Table from your EMIS dataset
- Calculate a percentage table
- Create a chart/graph from percentage table