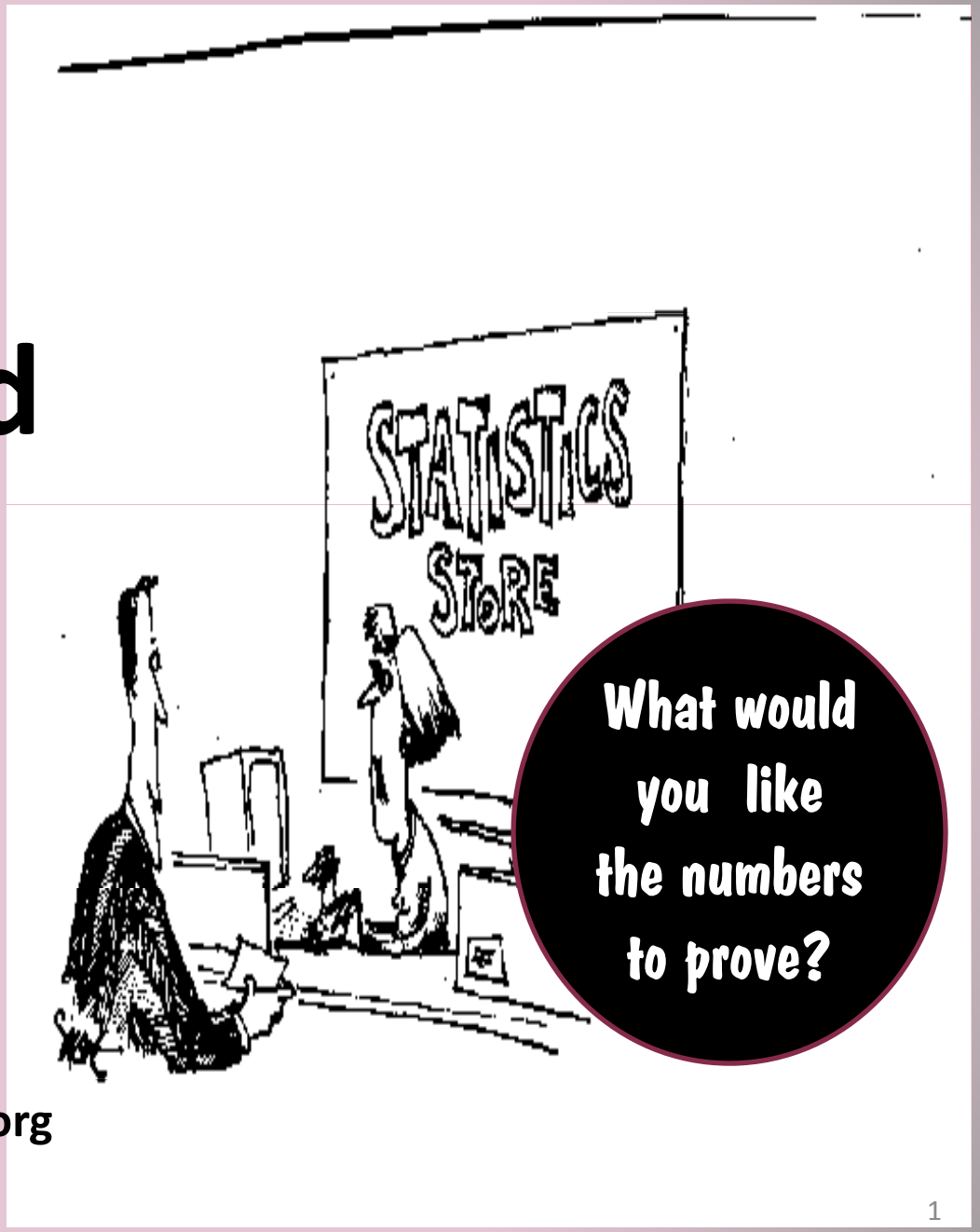
Statistical Literacy and Research Methods

Donna Kotsopoulos
Wilfrid Laurier University
dkotsopo@wlu.ca

www.mathematicsresearcher.org



9: 30 – 10:30	Statistical literacy
10:30 – 10:45	Break
10:45 – 11:30	Statistical literacy
11:30 – 12:30	Lunch
12:00 – 1:00	Weft (data analysis software of transcriptions) http://www.pressure.to/qda/;
1:00 – 1:30	Survey Monkey (online survey tool) http://www.surveymonkey.com/
1:30 – 2:00	Transana (video data analysis software) http://www.transana.org/
2:00 – 3:30	Refworks and EndNote (digital reference manager)

With a partner . . .

Develop a list of “statistical” terms, words, or phrases that you have heard used. These words may be familiar or unfamiliar to you. That is, you may not know the meaning but have heard the term, word, or phrases used.

Five parts of the statistical knowledge base

(Gal, 2002)

1. Knowing why data are needed and how data can be produced;
2. Familiarity with basic terms and ideas related to descriptive statistics;
3. Familiarity with basic terms and ideas related to graphical and tabular displays;
4. Understanding basic notions of probability;
5. Knowing how statistical conclusions or inferences are reached.

1. Knowing why data are needed and how data can be produced

London Central Secondary School is located in downtown London. The school offers a full selection of academic courses. In response to recent media reports about aboriginal students' school achievement, the administration planned to survey some aboriginal students at LCS to assess their achievement outcomes in relation to the results posted in the media.

2. Familiarity with basic terms and ideas related to descriptive statistics (describing, summarizing, and comparing)

Three broad categories of statistics:

Descriptive statistics answer “what” questions;

Explanatory statistics answer “why” questions;

Inferential statistics answer “why” questions AND can be generalized (hypothesis testing).

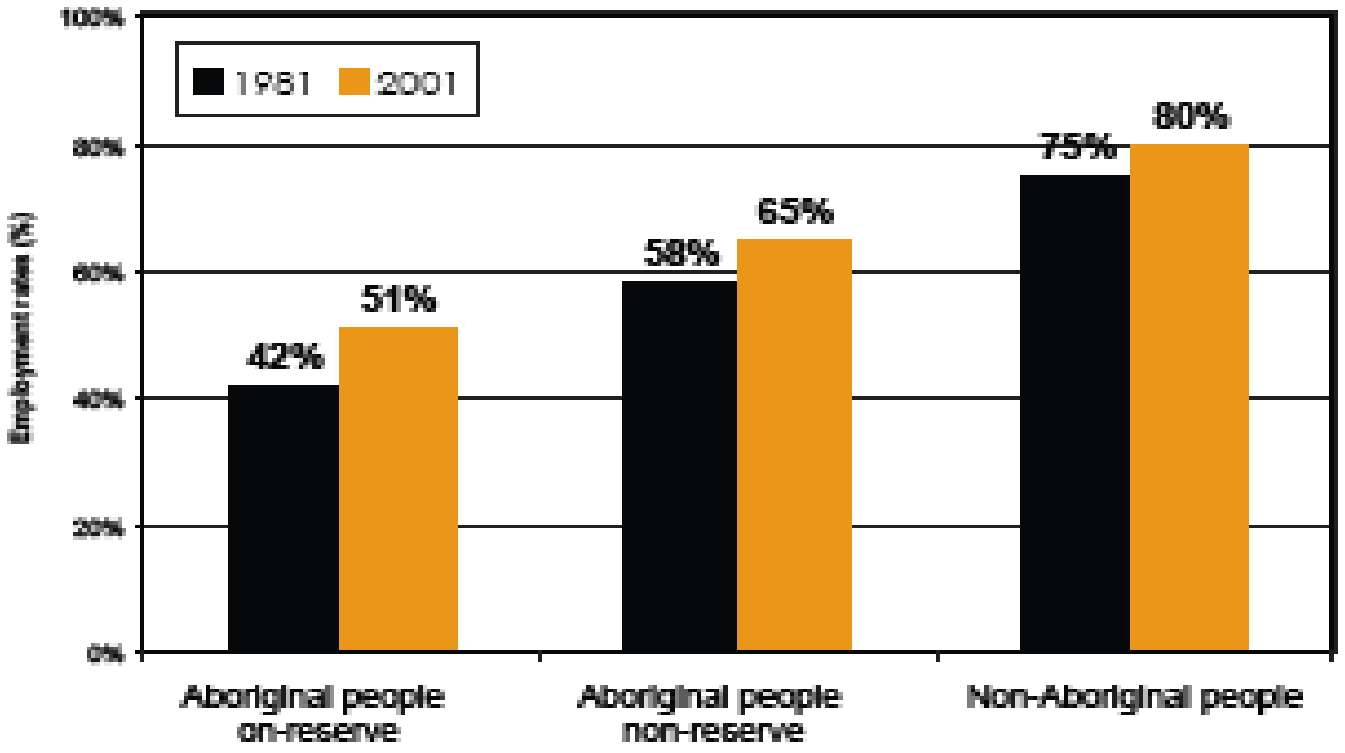
Descriptive Statistics

- **Percents**;
- **Frequency** (how many times);
- **Measures of central tendency** (most common occurrence or “center” of the data):
  - **Mean** (average)
  - **Median** (middle number)
  - **Mode** (number most frequently appearing)
- Types of distribution(normal, bimodal, skewed);
- **Measures of association** (correlations);
- **Measures of dispersion** (numerical representations of how much data points differ from one another).

3. Familiarity with basic terms and ideas related to graphical and tabular displays

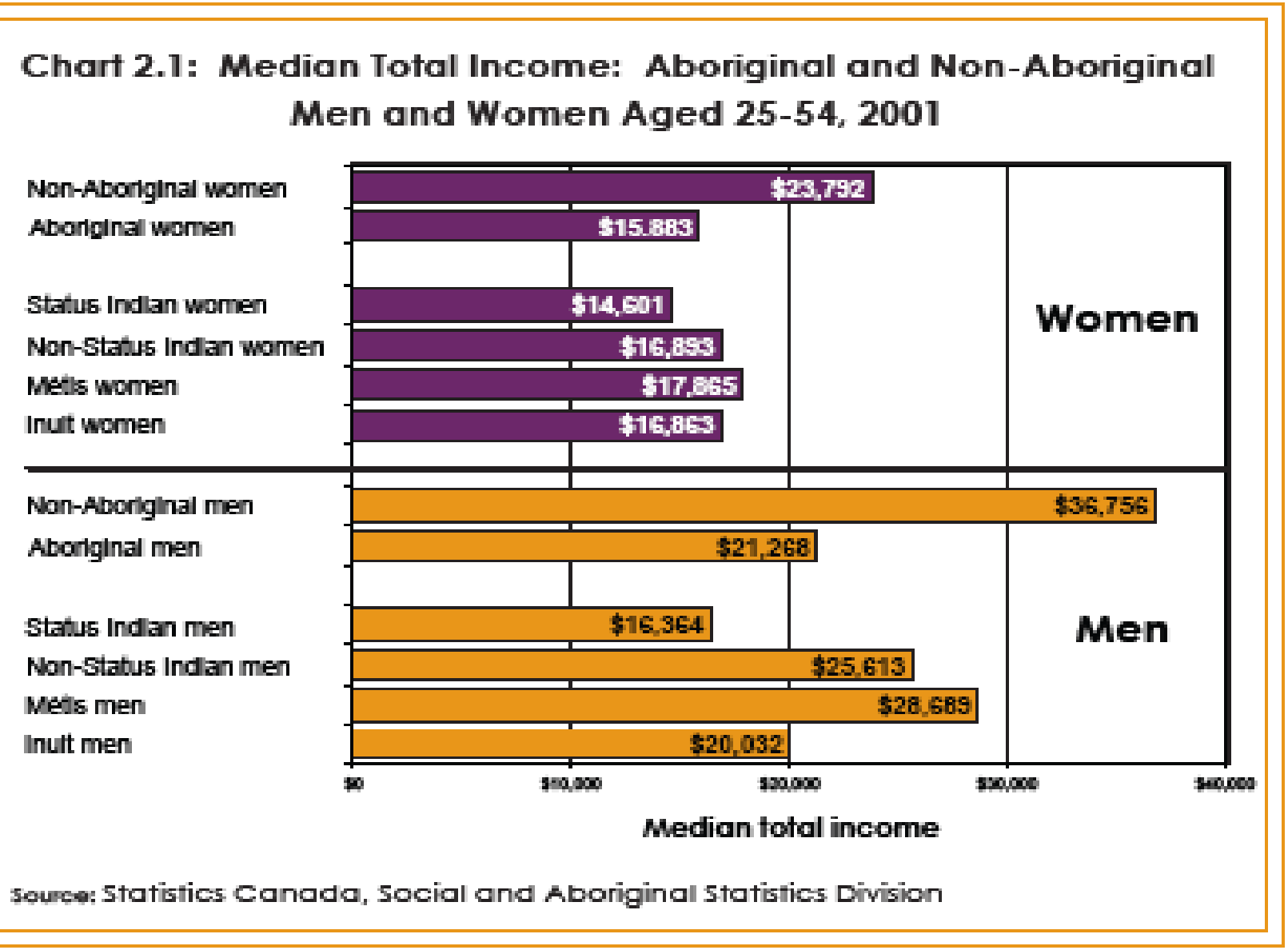
- Data in tables or graphs;
- “Reading between the data” and “reading beyond the data” ;
- Overall patterns and not only at specific points in a graph or a table;
- Different graphs and tables may yield different (and possibly conflicting) views of the phenomena under investigation;
- Graphs can be intentionally created to mislead or highlight/hide a specific trend or difference

Chart 3.1: Employment Rates, Aboriginal and Non-Aboriginal Persons Aged 25-54, 1981 and 2001

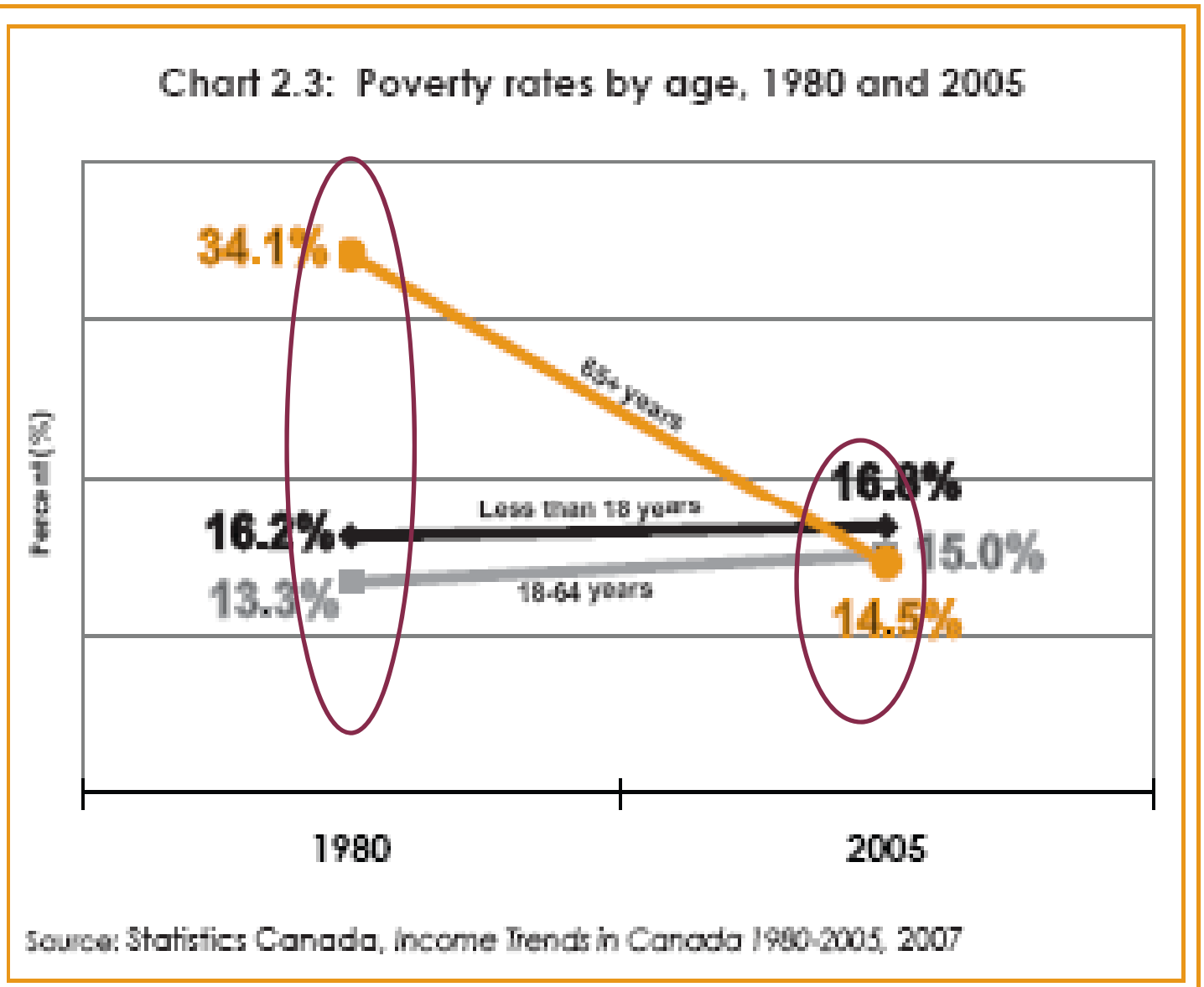


Source: Statistics Canada, Social and Aboriginal Statistics Division

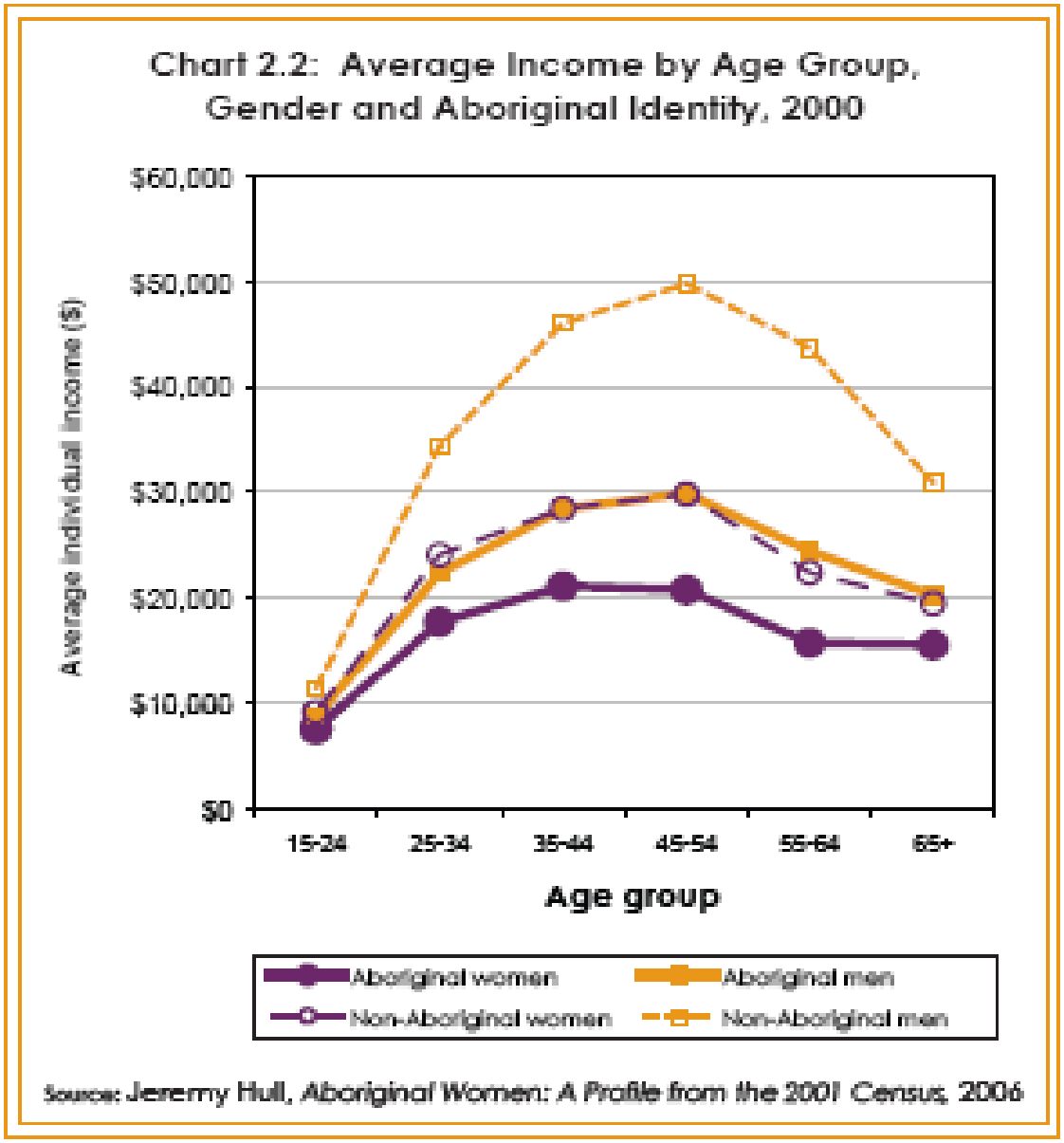
National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.



National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.



National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.



National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.

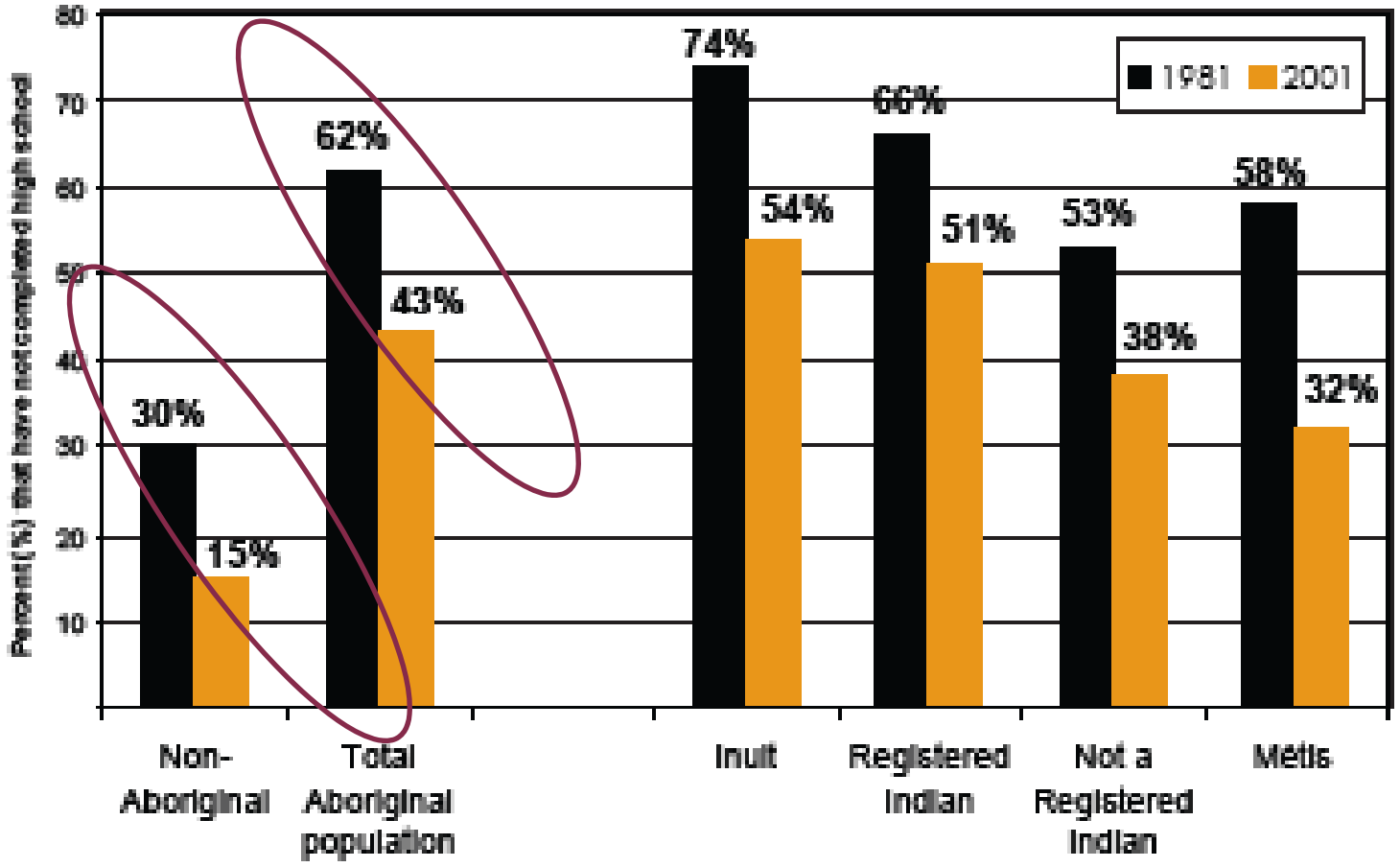
Table 2.1: Comparison of Income or Social Assistance Coverage Rates— On-reserve Population and General Population, 1997 and 2003

Province/ Territory and Canada	1997			2003		
	Total Population at March, 1997	Number of People on IA or SA at March, 1997	Percentage	Total Population at March, 2003	Number of People on IA or SA at March, 2003	Percentage
	General Population (excluding on-reserve population)					
Canada	29,819,070	2,774,900	9.3%	31,543,355	1,745,600	5.5%
British Columbia	3,931,016	321,300	8.1%	4,135,769	180,700	4.3%
Manitoba	1,135,851	79,100	6.9%	1,159,917	59,900	5.1%
Atlantic	2,373,558	247,300	10.4%	2,342,835	165,800	7.0%
Ontario	11,180,472	1,149,600	10.2%	12,193,256	673,900	5.5%
	On-reserve Population					
Canada	369,163	152,746	41.4%	423,631	147,300	34.8%
British Columbia	72,448	22,749	31.4%	80,103	18,009	22.5%
Manitoba	62,554	29,853	47.7%	74,038	32,483	43.9%
Atlantic	15,645	12,634	80.8%	18,420	11,818	64.2%
Ontario	73,109	18,004	24.6%	82,774	18,615	22.5%

Source: INAC, *Income Assistance Reform.*

National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.

Chart 4.1: Young Aboriginal and Non-Aboriginal Adults Aged 20-24 Years Who Have Not Completed High School



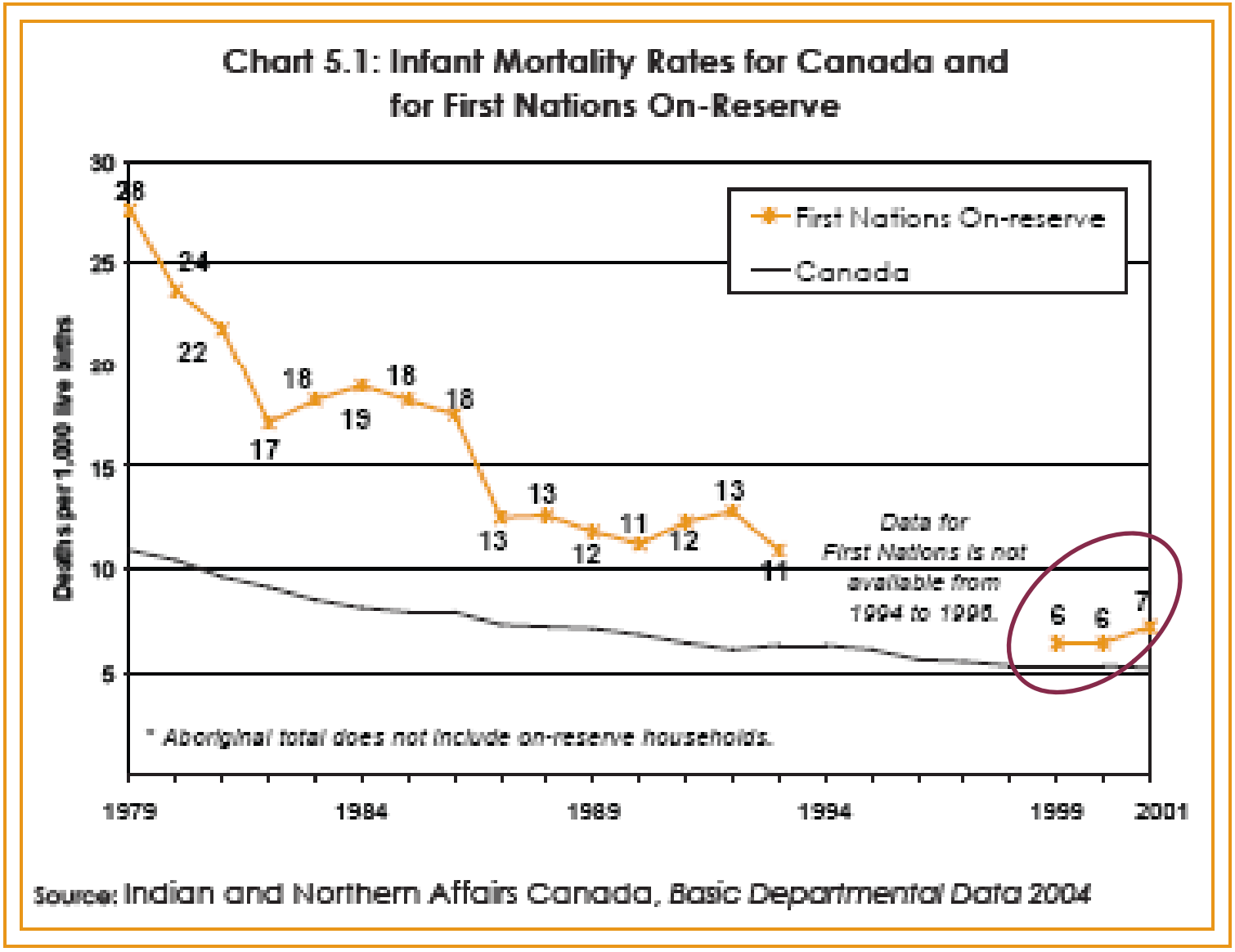
Source: Statistics Canada, Social and Aboriginal Statistics Division

What information does this table NOT give us that would be important?

Table 8.1: Aboriginal Young Persons Admitted to Correctional Services, 2003/04

Jurisdiction	Remand	Sentenced Custody		Probation	Aboriginal youth as % of total youth population
		Open Custody	Secure Custody		
	% of young persons who are Aboriginal				
TOTAL	27%	30%	27%	17%	5%
Newfoundland and Labrador	3%	9%	8%	4%	5%
Prince Edward Island	...	...	...	...	1%
Nova Scotia	6%	8%	0%	6%	3%
New Brunswick	5%	4%	7%	8%	3%
Quebec	...	...	...	...	2%
Ontario - 12 to 15 years old	...	6%	1%	3%	2%
- 16 to 17 years old	11%	10%	16%	8%	2%
Manitoba	71%	82%	77%	56%	19%
Saskatchewan	...	84%	75%	65%	19%
Alberta	39%	44%	37%	29%	8%
British Columbia	36%	35%	32%	29%	7%
Yukon	91%	100%	...	83%	26%
Northwest Territories	88%	100%	83%	...	63%
Nunavut	100%	100%	100%	...	95%

National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.



National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada.

Measures of association: Correlations

- Aware that observed differences or trends may exist;
- May not necessarily be large or stable enough to be important;
- Can be caused by chance processes.

Correlation tables

Table 2

Correlations among survey variables

Variable	1	2	3	4	5	6	7	8	9
1. Impact of inquiry on learning to teach	----								
2. Readiness to teach (RT)	-.29	----							
3. RT-planning	.11	.78**	----						
4. RT-positive learning environment	-.18	.52**	.51**	----					
5. RT- implementing instruction	-.19	.60**	..47**	. 60**	----				
6. RT-adapting instruction	.-.03	.57**	.68**	.54**	.41**	----			
7. RT-evaluating	-.10	.63**	. 55**	.37*	.50**	.57**	----		
8. RT-professional responsibilities	-.15	.67**	. 56**	. 60*	.53**	.55**	. .69**	----	
9. RT - technology	-.08	.43**	. 37*	. 32*	. 17	. 55*	. 37*	.41**	----

* p < .05
** p < .01
*** p < .001

These scores tell us that these results were not random.

Correlations

Between -1 and 1 (from Blaikie, 2003);

- -1.0 Perfect, negative.
- -0.75 – 0.99 Very strong, negative;
- -0.60 – -0.74 Strong, negative;
- -0.30 – - 0.59 Moderate, negative;
- -0.10 – -0.29 Weak, negative;
- -0.01 – -0.09 Negligible, negative;
- 0.00 None;
- 0.01 – 0.09 Negligible;
- 0.10 – 0.29 Weak;
- 0.30 – 0.59 Moderate;
- 0.60 – 0.74 Strong;
- 0.75 – 0.99 Very strong;
- 1.0 Perfect, positive.

Figure 2. Positive or direct relationships between variables

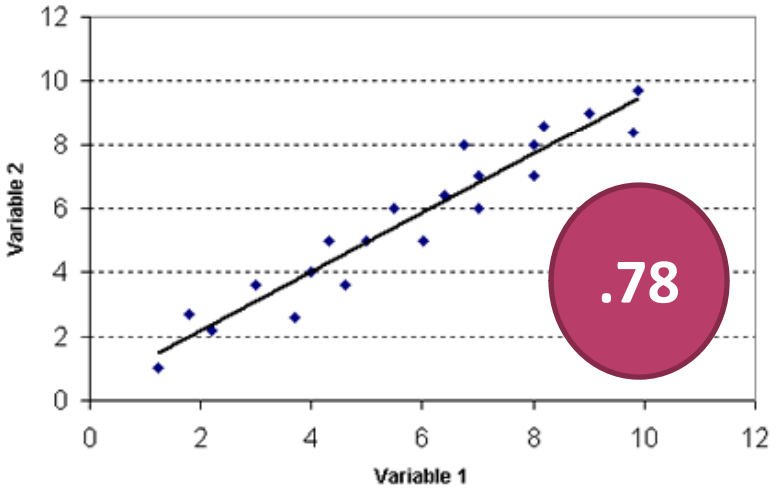


Figure 3. Negative or inverse relationships

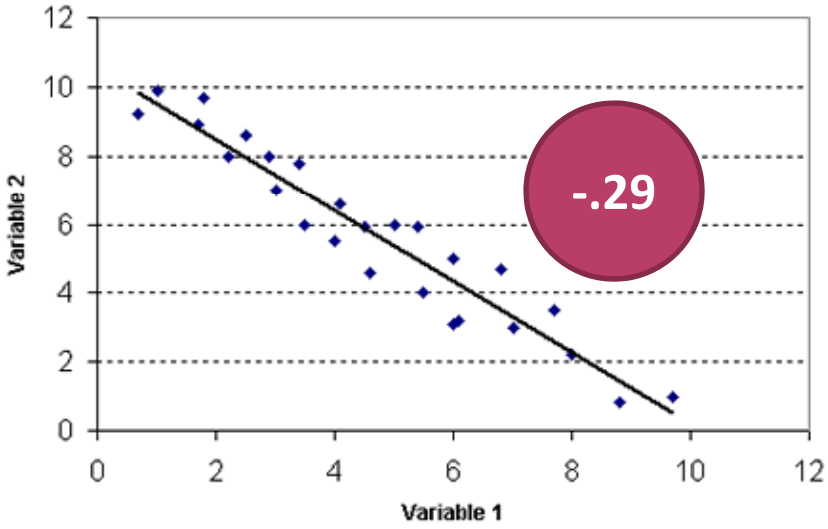
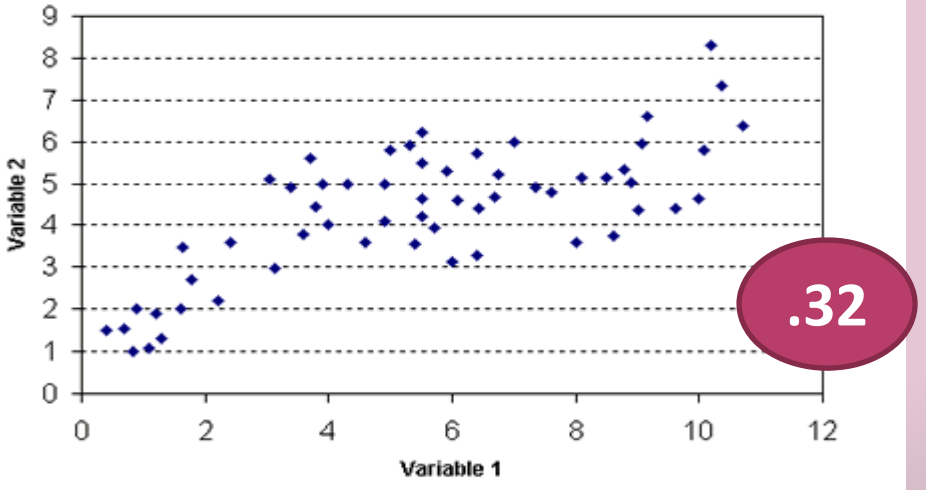
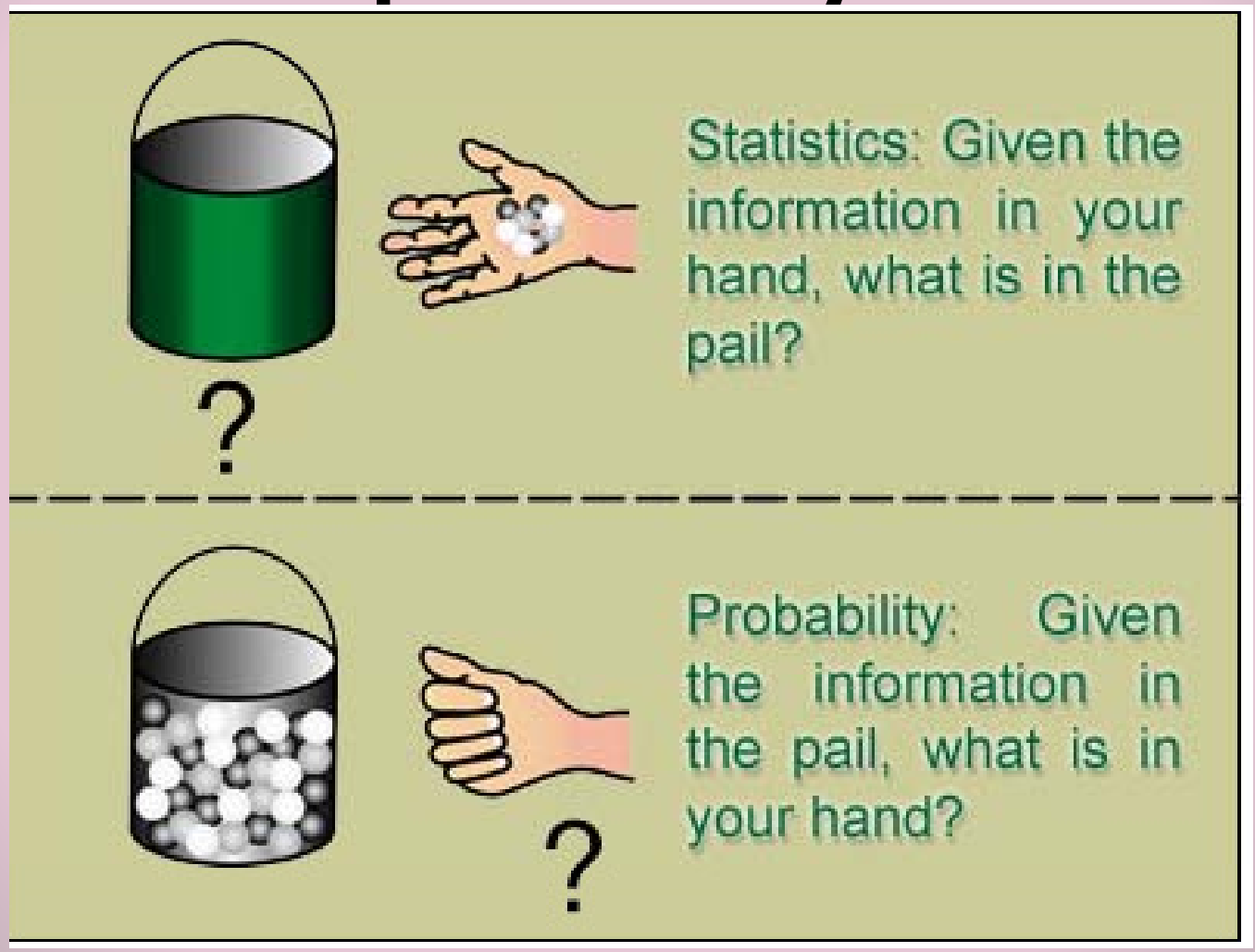
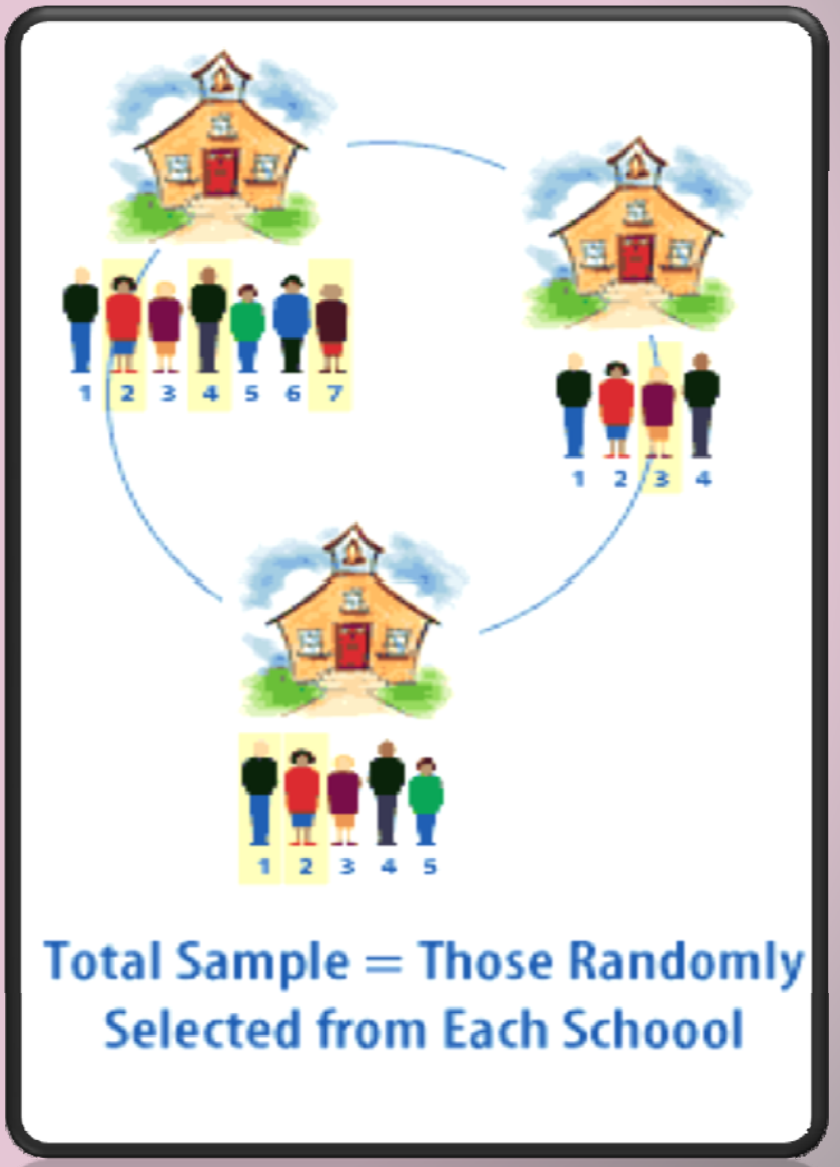


Figure 5. Very low or zero correlation



4. **Understanding basic notions of probability**



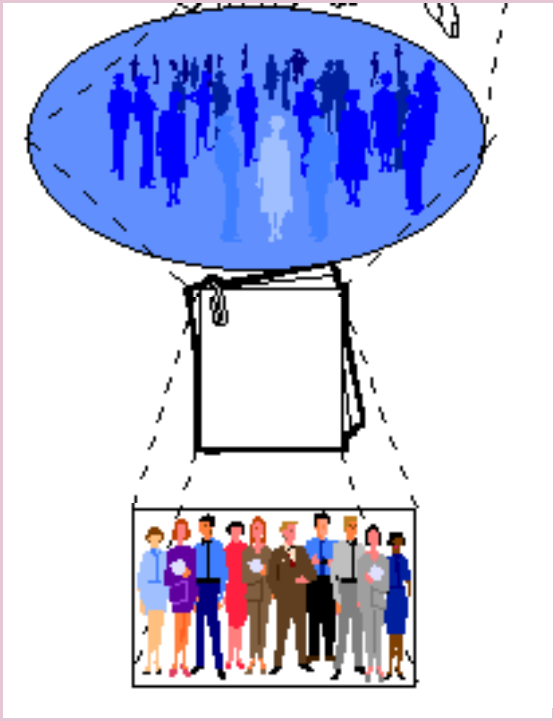

1 2 3 4 5 6 7
1 2 3 4
1 2 3 4 5
Total Sample = Those Randomly Selected from Each Schoool

What do you want to generalize to?

What population can you get access to?

How can you get access to them?

Who is in your study?



The theoretical population.

The study population.

The sampling frame.

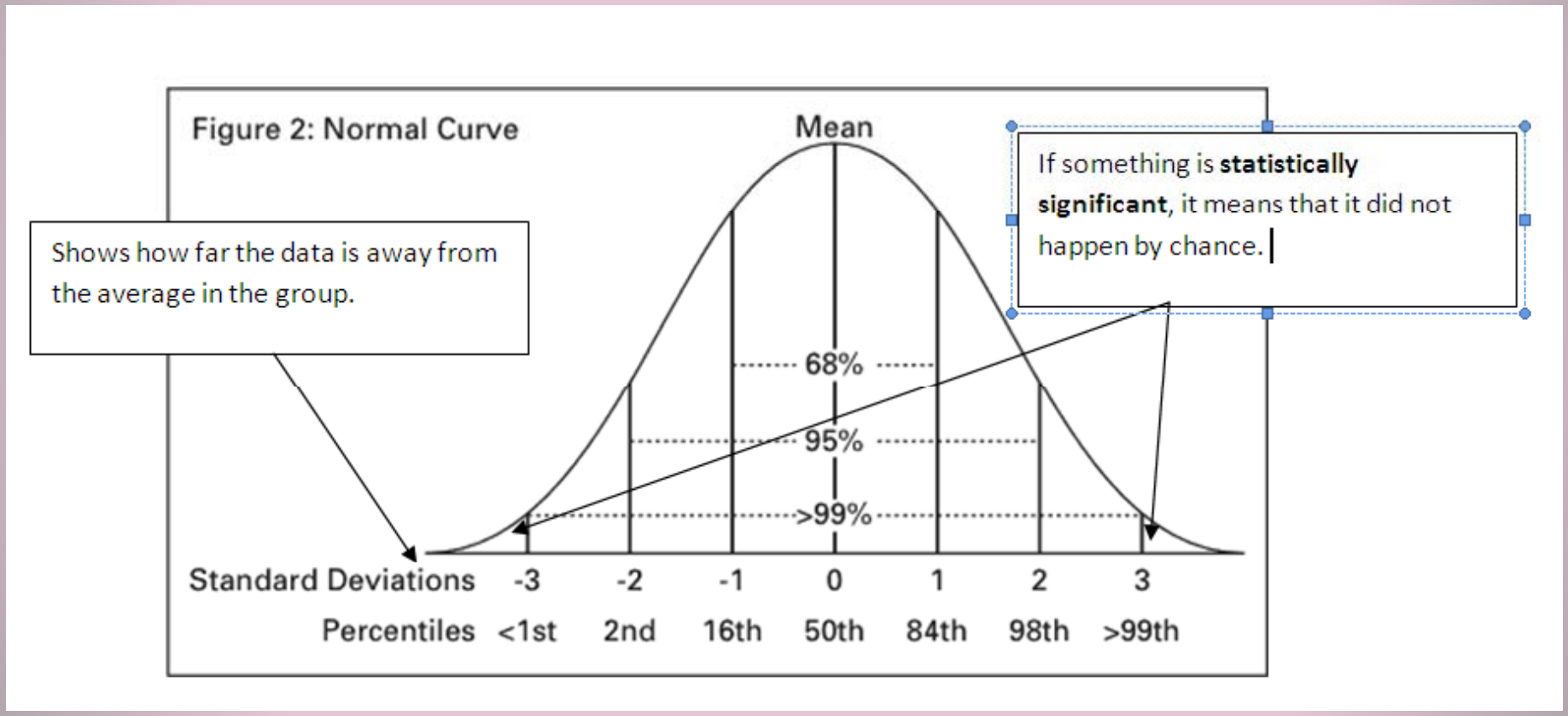
The sample.

5. Knowing how inferences are reached

Inferential statistics answer "why" and "how" questions and can be generalized (hypothesis testing).

- **Hypothesis statement** – always the beginning point;
- Must be a **random sample**;
- **Significance** – "trueness" of a difference between groups;
  - Did not happen by chance;
  - Not a value judgment of good or bad findings;
  - Requires attention to the size of the groups studied, to the quality of the sampling process, and the possibility that a sample is biased.

5. Knowing how statistical conclusions or inferences are reached



What we want to know from inferential analysis is whether what we have found in the sample also exists in the general population.

“WORRY questions ” about statistical messages

Where did the data (on which this statement is based) come from?

What kind of study was it?

Is this kind of study reasonable in this context?

Was a sample used? How was it sampled? How many people did actually participate? Is the sample large enough?

Did the sample include people/units which are representative of the population? Is the sample biased in some way?

Overall, could this sample reasonably lead to valid inferences about the target population?

How reliable or accurate were the instruments or measures (tests, questionnaires, interviews) used to generate the reported data?

What is the shape of the underlying distribution of raw data (on which this summary statistic is based)? **Does** it matter how it is shaped?

Are the reported statistics appropriate for this kind of data, e.g., was an average used to summarize ordinal data; is a mode a reasonable summary?

Could outliers cause a summary statistic to misrepresent the true picture?

Is a given graph drawn appropriately, or does it distort trends in the data?

How was this probabilistic statement derived? Are there enough credible data to justify the estimate of likelihood given?

Overall, are the claims made here sensible and supported by the data? e.g., is correlation confused with causation, or a small difference made to loom large?

Should additional information or procedures be made available to enable me to evaluate the sensibility of these arguments?

Is something missing? e.g., did the writer "conveniently forget" to specify the base of a reported percent-of-change, or the actual sample size?

Are there alternative interpretations for the meaning of the findings or different explanations for what caused them, e.g., an intervening or a moderator variable affected the results?

Are there additional or different implications that are not mentioned?

Weft

http://www.pressure.to/qda/;

- Free;
- Must have a transcribed file, or document in “rtf” format;
- Basic search and comparative mechanisms

Survey Monkey
(online survey tool)
http://www.surveymonkey.com/

- Free;
- Up to 10 questions (some options with multiple parts);
- Up to 100 surveys can be created;
- Secure – one response per connect to the email account from one location.

http://www.transana.org/

- Allows you to transcribe and code from video;
- Requires that you have a video file of the video;
- Benefits;
- Limitations;
- Ethical considerations;
- Free resources (Window's Movie Maker);
- Resources available through UWO
  (Video camera, firewire, video transfer software);

References

Blaike, N. (2003). *Analyzing quantitative data. Thousand Oaks, California: Sage Publications, Inc.*

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*(1), 1-51.

National Council of Welfare (2007). *National Council of Welfare Reports: First Nations, Métis and Inuit children and youth: Time To Act, Volume 127.* Author, Ottawa, Canada. Retrieved from: http://www.ncwcnbes.net/documents/researchpublications/ResearchProjects/FirstNationsMetisInuitChildrenAndYouth/2007Report-TimeToAct/ReportENG.pdf

Rowntree, D. (1981). *Statistics without tears: A primer for Non-mathematicians. London, England: Pelican Books.*

Statistics Canada http://www.statcan.ca/english/edu/power/ch9/scattergraphs/scatter.htm#link1

Web sites:

Statistics Canada http://www.statcan.ca/english/edu/power/ch9/scattergraphs/scatter.htm#link1

Psychology Statistics http://www.uwsp.edu/psych/stat/1/expdes.htm

EndNote

Versus RefWorks