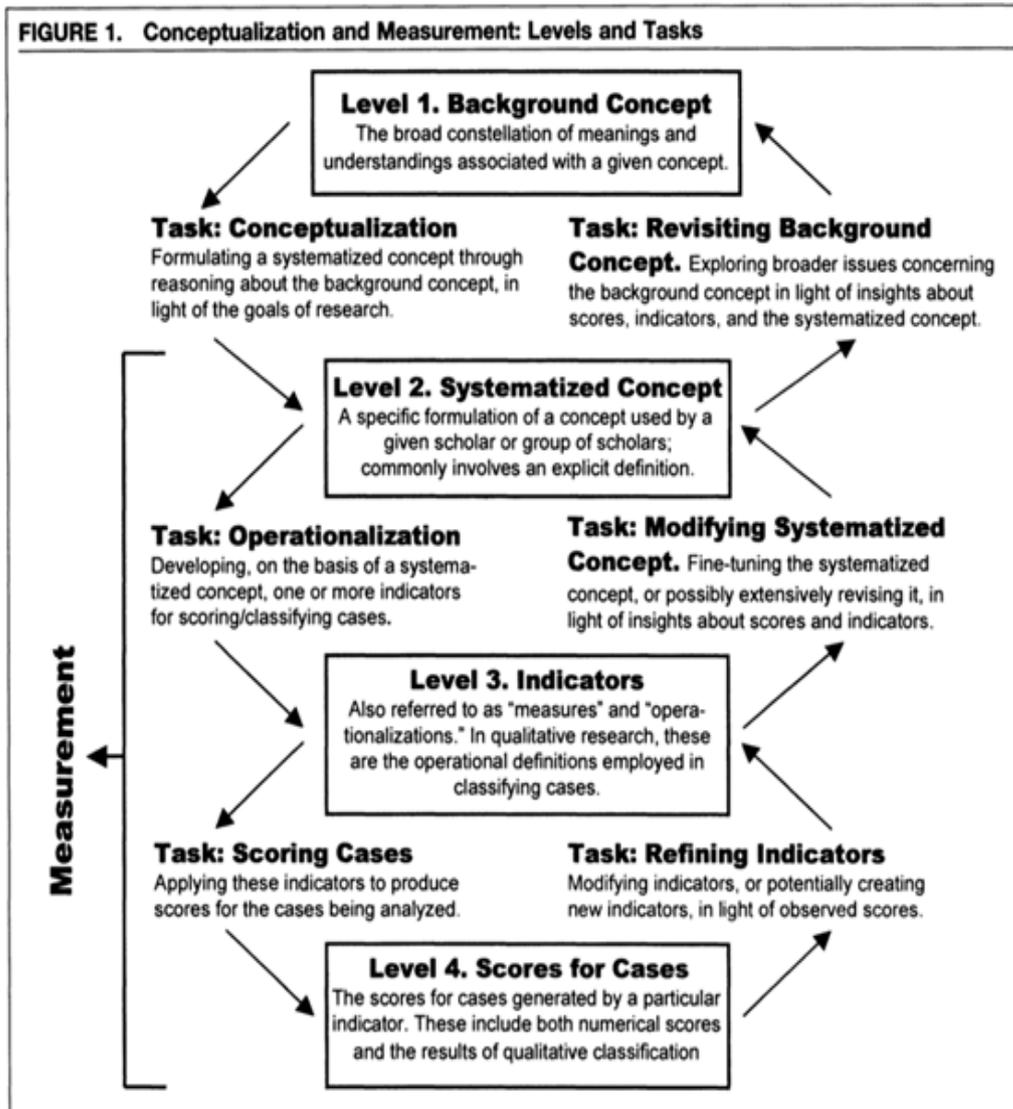


# A complexidade da mensuração em CS



## **Indicadores e variáveis**

Operacionalização: tarefa através da qual o pesquisador especifica que observações empíricas podem ser tomadas como indicadores adequados de seus conceitos (conceitos => indicadores => variáveis).

Indicador: observação empírica considerada um “reflexo” de um conceito.

- relação entre conceitos e indicadores é *probabilística*;

Variáveis: conjuntos lógicos de atributos; resultado da partição dos objetos segundo um ou vários critérios;

Tipos de variáveis: **quantitativa** (nível intervalar); **qualitativa** ou **categórica** (nominal ou ordinal).

---

---

## Matriz de dados

de ordenar los datos de manera que sea particularmente visible la forma tripartita:

Tabla 1.1.3. La matriz de datos

$$M = \begin{matrix} & S_1 & S_2 & S_3 & \dots & S_j & \dots & S_n \\ \begin{matrix} O_1 \\ O_2 \\ O_3 \\ \vdots \\ O_i \\ \vdots \\ O_m \end{matrix} & \left\{ \begin{array}{cccccccc} R_{11} & R_{12} & R_{13} & \dots & R_{1j} & \dots & R_{1n} \\ R_{21} & R_{22} & R_{23} & \dots & R_{2j} & \dots & R_{2n} \\ R_{31} & R_{32} & R_{33} & \dots & R_{3j} & \dots & R_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ R_{i1} & R_{i2} & R_{i3} & \dots & R_{ij} & \dots & R_{in} \\ \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ R_{m1} & R_{m2} & R_{m3} & \dots & R_{mj} & \dots & R_{mn} \end{array} \right. \end{matrix}$$

En este esquema,  $R_{ij}$  es la respuesta que da la unidad  $nt_i$  al estímulo  $nt_j$ ; o, en otras palabras, el valor que tiene la unidad  $nt_i$  en la variable  $nt_j$ .

### Forma tripartite:

- i) Unidade de análise;
- ii) Variáveis;
- iii) Valores.

### Qualidade da matriz:

- i) Princípio da comparabilidade;
- ii) Princípio da classificação;
- iii) Princípio da integridade.

## Escalas e índices

O que escalas e índices têm a ver com mensuração?

- relação *probabilística* entre indicador e conceito/variável => quanto mais indicadores válidos, mais precisa será a mensuração.

Índices e escalas são “medidas sintéticas” que resumem informações sobre os casos.

Diferenças entre índices e escalas:

- índice: somatório simples dos escores dos indicadores;
  - escalas: atribuição de escores a padrões de resposta.
- 
-

## *Construindo índices e escalas*

Etapas:

1. Escolha dos indicadores da medida sintética;

- validade aparente;

2. Exame das correlações entre os itens (análise bivariada e multivariada);

3. Atribuições de escores;

- amplitude;

- peso.

---

---

## Construindo índices e escalas (II)

**Revidar violência policial**

Não	Não	Não	Não	Sim
Não	Não	Não	Sim	Sim
Não	Não	Sim	Sim	Sim
Não	Sim	Sim	Sim	Sim
<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

**Interromper o tráfego**

**Acampar em praça pública**

**Carregar cartazes**

# Análise dos dados quantitativos

## 1. Análise univariada:

- proporções (atributos de variáveis nominais ou ordinais);
- medidas de tendência central (atributos de variáveis intervalares ou de razão).

**Tabela 1 – Frequência absoluta e relativa de homens e mulheres na amostra:**

	N	%	% válida
<b>Homem</b>	482	49,5	49,5
<b>Mulher</b>	492	50,5	50,5
<b>Total</b>	974	100,0	100,0

**Tabela 2 – Frequência absoluta e relativa dos indivíduos por faixa etária**

	N	%	% válida	% acumulada
<b>18-24</b>	185	19,0	19,0	19,0
<b>25-34</b>	265	27,2	27,2	46,2
<b>35-44</b>	218	22,4	22,4	68,5
<b>45-54</b>	144	14,8	14,8	83,3
<b>55-64</b>	84	8,6	8,6	92,0
<b>65+</b>	78	8,0	8,0	100,0
<b>Total</b>	974	100,0	100,0	

## *Análise dos dados (II)*

2. Análise bivariada: cruzamento entre duas variáveis nominais ou ordinais (tabela de contingência);
  - O que significa dizer que duas variáveis estão relacionadas (correlacionadas) entre si?  
=> alterações no valor de uma variável são acompanhadas por alterações no valor da outra variável;

Três tipos de relação:

i) simétrica;

ii) recíproca;

iii) assimétrica:

- ordem temporal;

- sentido dominante de influência.

- modalidade de relação assimétrica: relação entre propriedades dos agentes e comportamentos (nível de instrução e participação política).

---

---

# Tabelas de contingência (exemplos)

**Tabela 3 - Cruzamento das variáveis sexo por escolaridade**

Escolaridade	Sexo		
	Homem	Mulher	Total
Nunca Frequentou	1 0,3%	1 0,4%	2 0,3%
Fundamental incompleto	146 41,0%	86 35,8%	232 38,9%
Fundamental completo	25 7,0%	16 6,7%	41 6,9%
Médio incompleto	42 11,8%	22 9,2%	64 10,7%
Médio completo	78 21,9%	49 20,4%	127 21,3%
Superior incompleto	20 5,6%	18 7,5%	38 6,4%
Superior completo	44 12,4%	48 20,0%	92 15,4%
<b>Total</b>	<b>356 100,0%</b>	<b>240 100,0%</b>	<b>596 100,0%</b>

**Tabela 4 – Frequência absoluta e relativa dos que gostam ou não de rock por escolaridade**

	escolaridade do respondente							Total
	Nunca frequentou	Fundamental incompleto	Fundamental completo	Médio incompleto	Médio completo	Superior incompleto	Superior completo	
<b>Gosta de rock</b>	0 0,0%	54 23,2%	9 21,4%	26 41,3%	58 45,7%	19 50,0%	57 61,3%	223 37,3%
<b>Não gosta de rock</b>	2 100,0%	179 76,8%	33 78,6%	37 58,7%	69 54,3%	19 50,0%	36 38,7%	375 62,7%
<b>Total</b>	2 100,0%	233 100,0%	42 100,0%	63 100,0%	127 100,0%	38 100,0%	93 100,0%	598 100,0%

**Tabela 5 – Cruzamento entre o índice de práticas culturais e a escolaridade dos respondentes**

	Escolaridade do respondente							Total
	Nunca frequentou	Fundamental incompleto	Fundamental completo	Médio incompleto	Médio completo	Superior incompleto	Superior completo	
<b>Baixo</b>	2 100,0%	191 82,0%	25 59,5%	25 39,7%	54 42,5%	7 17,9%	7 7,5%	311 51,9%
<b>Médio</b>	0 0,0%	38 16,3%	15 35,7%	27 42,9%	52 40,9%	15 38,5%	18 19,4%	165 27,5%
<b>Alto</b>	0 0,0%	4 1,7%	2 4,8%	11 17,5%	21 16,5%	17 43,6%	68 73,1%	123 20,5%
<b>Total</b>	2 100,0%	233 100,0%	42 100,0%	63 100,0%	127 100,0%	39 100,0%	93 100,0%	599 100,0%

## *Análise dos dados (III)*

### 3. Análise multivariada:

- descrição estratificada da amostra (comparações de subgrupos);
- técnica de inferência causal.

Uma correlação empírica é condição suficiente e necessária de uma relação causal?

A causalidade pode ser observada diretamente nos dados?

- Não é condição suficiente: uma terceira variável (extrínseca) pode ser a causa de ambas;
  - Interpretação de uma correlação: supõe a introdução de outras variáveis (não-espuriidade).
  - Causalidade: derivada da teoria.
- 
-

## Uma concepção de causalidade

- Dependência “robusta”: a questão é estabelecer *se, e em que medida*, uma associação observada entre duas variáveis, X e Y, sendo X *temporalmente anterior* a Y, pode ser tomada como uma relação em que X é causalmente significativa para Y.
- Diferenciar associação de causação: associação não é eliminada caso novas variáveis (Z) sejam inseridas na relação original ( $X \Rightarrow Y$ ).

Três condições de inferência causal:

- i) X é anterior a Y;
  - ii) o sentido dominante da influência é *de X para Y*;
  - iii) a relação entre X e Y não se deve a uma variável comum a ambas e extrínseca à relação.
- 
-

## Exemplo retirado de “O Suicídio”

QUADRO XXI  
FRANÇA (1889-91)  
*Suicídios cometidos por 1.000 habitantes de cada grupo de idade e de estado civil*  
– Média anual

Idades	Solteiros	Casados	Viúvos	Coeficiente de preservação dos		
				Casados		Viúvos
				Com relação aos solteiros	Com relação aos viúvos	Com relação aos solteiros
<i>Homens</i>						
15-20 .....	113	500		0,22		
20-25 .....	237	97	142	2,40	1,45	1,66
25-30 .....	394	122	412	3,20	3,37	0,95
30-40 .....	627	226	560	2,77	2,47	1,12
40-50 .....	975	340	721	2,86	2,12	1,35
50-60 .....	1.434	520	979	2,75	1,88	1,46
60-70 .....	1.768	635	1.166	2,78	1,83	1,51
70-80 .....	1.983	704	1.288	2,81	1,82	1,54
Acima .....	1.571	770	1.154	2,04	1,49	1,36
<i>Mulheres</i>						
15-20 .....	79,4	33	333	2,39	10	0,23
20-25 .....	106	53	66	2,00	1,05	1,60
25-30 .....	151	68	178	2,22	2,61	0,84
30-40 .....	126	82	205	1,53	2,50	0,61
40-50 .....	171	106	168	1,61	1,58	1,01
50-60 .....	204	151	199	1,35	1,31	1,02
60-70 .....	189	158	257	1,19	1,62	0,77
70-80 .....	206	209	248	0,98	1,18	0,83
Acima .....	176	110	240	1,60	2,18	0,79

## O método da elaboração (I)

Elaboração: tornar interpretável a relação ente duas variáveis;

Três tipos:

1. Explicação: eliminar a associação inserindo uma variável que se assuma seja comum a ambas e extrínseca à relação original;

Relações entre idade e o interesse por três tipos de transmissões radiofônicas		
	Jovens	Idosos
	%	%
Programas religiosos	17	26
Programas políticos	34	45
Programas de música clássica	30	29
N (absoluto)	1000	1300

Interesse pelos programas religiosos em função da idade e nível de instrução		
Nível de instrução superior	Jovens	9%
	Idosos	11%
Nível de instrução inferior	Jovens	32%
	Idosos	29%

## Método da elaboração (II)

2. Interpretação: reconstruir a cadeia causal que conecta X e Y (variáveis intervenientes e antecedentes);
3. Especificação: condições de ocorrência da relação original;

Relações entre idade e o interesse por três tipos de transmissões radiofônicas		
	Jovens	Idosos
	%	%
Programas religiosos	17	26
Programas políticos	34	45
Programas de música clássica	30	29
N (absoluto)	1000	1300

Interesse por programas de música clássica em função da idade e do nível de instrução		
Nível de instrução superior	Jovens	32%
	Idosos	52%
Nível de instrução inferior	Jovens	28%
	Idosos	19%

## Influência conjunta

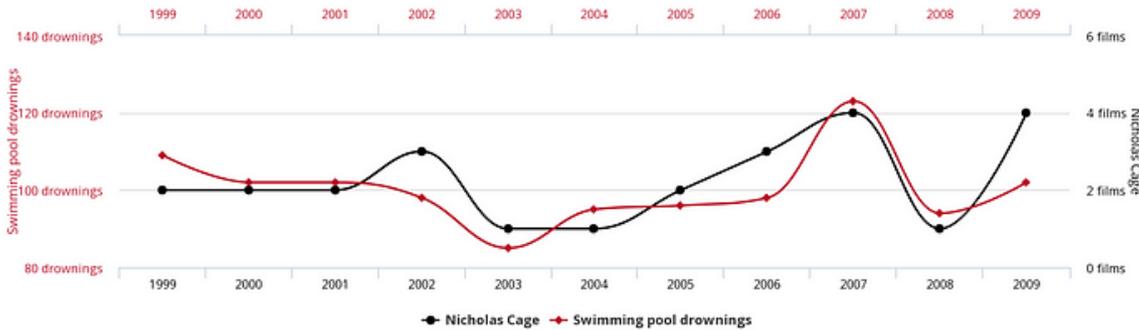
- duas ou mais variáveis têm efeitos independentes sobre outra variável;

Relações entre idade e o interesse por três tipos de transmissões radiofônicas		
	Jovens	Idosos
	%	%
Programas religiosos	17	26
Programas políticos	34	45
Programas de música clássica	30	29
N (absoluto)	1000	1300

Interesse pelos programas políticos em função da idade de nível de instrução		
Nível de instrução superior	Jovens	40%
	Idosos	55%
Nível de instrução inferior	Jovens	25%
	Idosos	40%

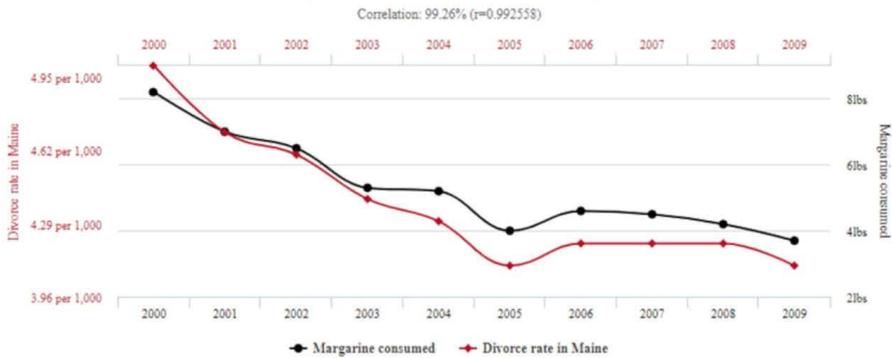
# Exemplos de correlação espúria

Number of people who drowned by falling into a pool  
correlates with  
Films Nicolas Cage appeared in



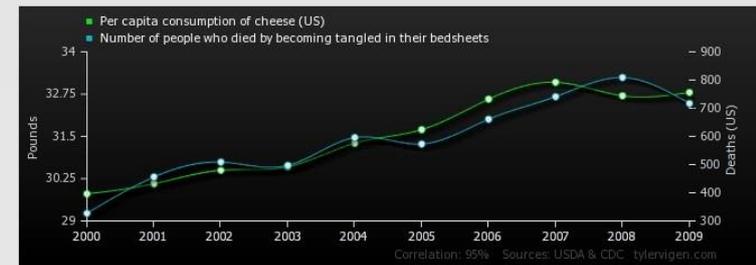
tylervigen.com

Divorce rate in Maine  
correlates with  
Per capita consumption of margarine



tylervigen.com

Per capita consumption of cheese (US)  
correlates with  
Number of people who died by becoming tangled in their bedsheets



Upload this chart to Imgur

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Per capita consumption of cheese (US) Pounds (USDA)	29.8	30.1	30.5	30.6	31.3	31.7	32.6	33.1	32.7	32.8
Number of people who died by becoming tangled in their bedsheets Deaths (US) (CDC)	327	456	509	497	596	573	661	741	809	717

**Correlation: 0.947091**

Permalink - Mark as interesting - Not interesting