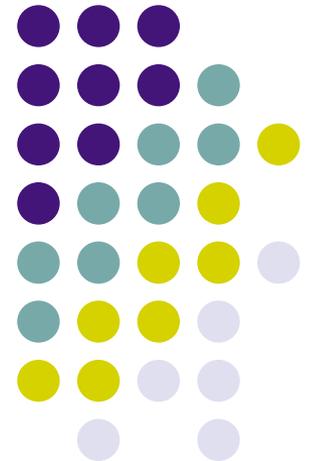
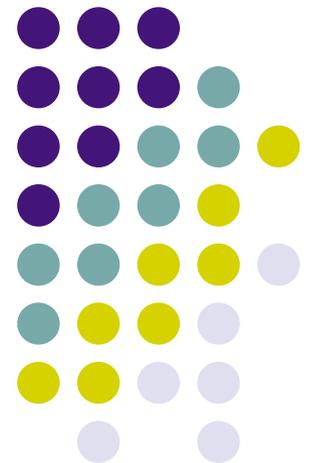


# Métodos de Pesquisa

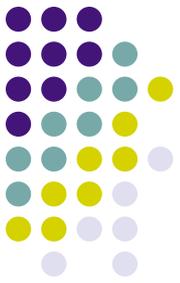
---



# Problemas e Hipóteses

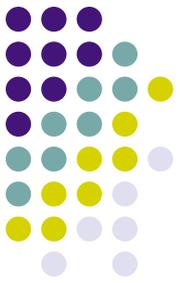


# Problemas



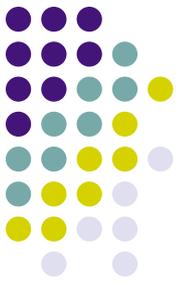
- Boa formulação:
  - Relação entre duas ou mais variáveis
    - Estaria A relacionado com B?
    - Quais são os efeitos dos incentivos no desempenho das pessoas?
  - Forma interrogativa
  - Possibilidade de teste empírico, mensurável

# Hipóteses



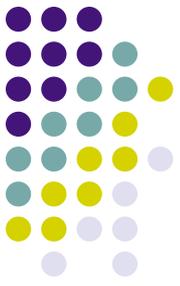
- É um enunciado conjectural da relação entre duas ou mais variáveis..
- Critérios:
  - relações
  - implicações claras para o teste da relação formulada
- Importância:
  - instrumentos de trabalho da teoria
  - podem ser testadas e julgadas como provavelmente verdadeiras ou falsas
  - instrumentos poderosos para avançar o conhecimento

# Variáveis



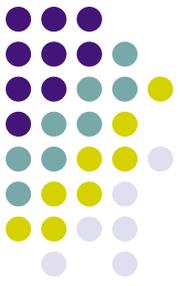
- são os construtos ou propriedades que são estudados
- variável é um símbolo ao qual se atribuem números ou valores
- pode ser dicotômica ou contínua
- constitutiva: é a definição de um constructo com outros constructos
- operacional: é definição que especifica as operações necessárias para medir uma variável

# Dez passos para planejar uma boa pesquisa



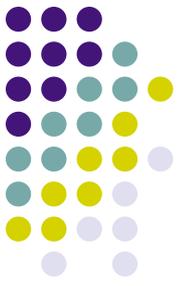
- Dificuldade básica – O que é que desperta o seu interesse ou desperta uma questão em sua mente?
- Base racional (fundamental) e teórica – Isso pode ser ajustado dentro de um modelo conceitual que dá um ponto de vista estruturado?
- Afirmação do propósito ou problema – O que é que você planeja para investigar? Quais são as metas principais do estudo? Defina o problema
- Questões para serem respondidas – Quando a pesquisa estiver terminada, quais são as perguntas que respostas razoáveis podem ser esperadas?

# Dez passos para planejar uma boa pesquisa



- Afirmação de hipóteses ou objetivos – Formule uma pesquisa de hipóteses que você irar testar ou os objetivos específicos que a pesquisa aponta. Seja concreto e claro.
- Design e procedimento.
- Suposição – Quais suposições você fez sobre a natureza dos comportamentos que você está investigando, sobre as condições sob as quais o comportamento ocorre, sobre seus métodos e medidas ou sobre o relacionamento desse estudo com outras pessoas e situações?

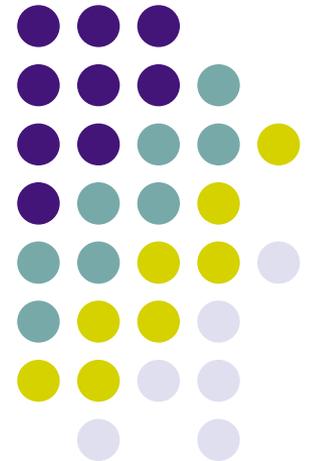
# Dez passos para planejar uma boa pesquisa



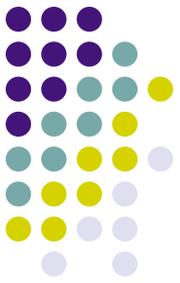
- Limitações – Quais são as limitações que estão em volta do seu estudo e em quais conclusões deve ser reduzido?
- Delimitações – Como você estreitou, arbitrariamente, o escopo do seu trabalho?
- Definição dos termos – Liste e defina os principais termos que você usará, particularmente, quando os termos tem significados diferentes para diferentes pessoas. Deve-se dar ênfase nas definições operacionais e comportamentais.

# Análise de Dados

---

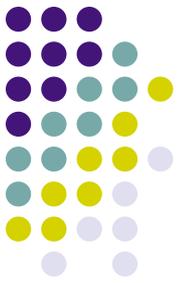


# Lógica da análise de dados



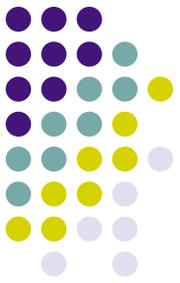
- dada uma relação entre duas variáveis, que se poderá depreender da introdução de uma terceira variável na análise?

# Relação entre variáveis

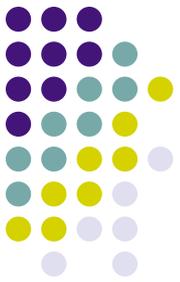


- Simétricas: nenhuma tem ação sobre a outra
- Recíprocas: não é possível dizer qual é a causa e qual é o efeito
- Assimétricas: uma variável é responsável pela outra

# Relação entre variáveis

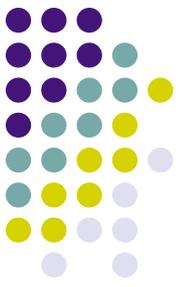


- **Simétricas - nenhuma das variáveis influencia a outra**
  - indicadores alternativos de um mesmo conceito, ex.: suor nas mãos e taquicardia
  - efeitos de uma causa comum, ex.: nível medicina e frequência viagem aérea (espúria)
  - interdependência funcional dos elementos de uma mesma unidade, ex.: normas e hierarquia.



# Relação entre variáveis

- **Simétricas - nenhuma das variáveis influencia a outra**
- os elementos se associam como partes de um “sistema” ou “complexo” comum. Ex.: membro de clube e opera
- fortuitas, ex.: rock e era espacial

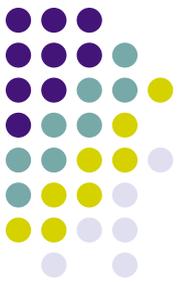


# Relação entre variáveis

- **Recíprocas - ambas podem influenciar-se uma à outra.**
- difícil afirmar qual é variável independente (causa) e qual é variável dependente (efeito).

Ex.:

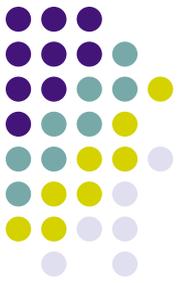
- Temperatura e termostato
- Partido político e comunicação do partido
- Investimento e lucro
- Desemprego e venda varejo



# Relação entre variáveis

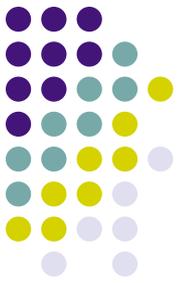
- **Assimétrica** - uma variável pode influenciar a outra
- crítico - falamos de causa quando uma alteração externa produz alteração em alguma unidade da outra

# Relação entre variáveis

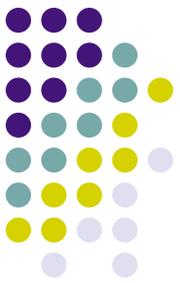


- **Assimétrica** - uma variável pode influenciar a outra
- suscetibilidade a influência: sexo e interesse político.
  - Ordem temporal (namoro e felicidade conjugal) e
  - Fixidez ou alterabilidade da variável

# Tipos de Assimétricas

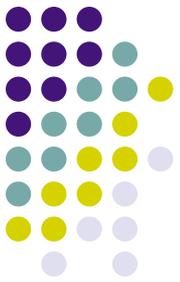


- Associação entre um estímulo e uma resposta. Ex.: ver tv e hábitos de estudo
  - seletividade
- Associação entre uma disposição e uma resposta. Ex.: preconceito e discriminação
- Associação entre uma propriedade e um ato ou disposição. Ex.: estado civil e hábitos de compra, gênero e preferência política

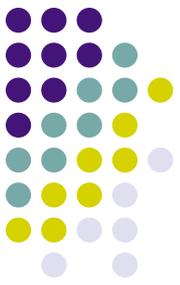


# Tipos de Assimétricas

- condição necessária para um efeito. Ex.: desenvolvimento tecnológico e armas nuclear
- relação imanente entre duas variáveis Ex.: burocracia e papéis
- associação entre fins e meios. Ex.: padronização e custos



# FATOR DE TESTE



# Fator de teste

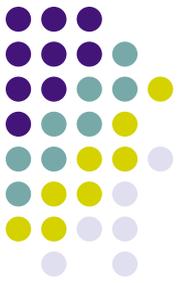
- Introduzido para aumentar a compreensão da relação original entre duas variáveis
- O objetivo da análise é o de determinar que a relação entre  $X$  ( a variável independente) e  $Y$  (a variável dependente) se deve a  $Z$  (o fator de teste)
- Exercer controle ou manter constante o fator de teste, eliminando sua influência sobre a relação

# Tabela 2-1- Idade e audiência de programas religiosos



Ouvem Programas Religiosos	Ouvintes Jovens	Ouvintes idosos
Sim	17%	26%
Não	83%	74%
Porcentagem Total	100	100

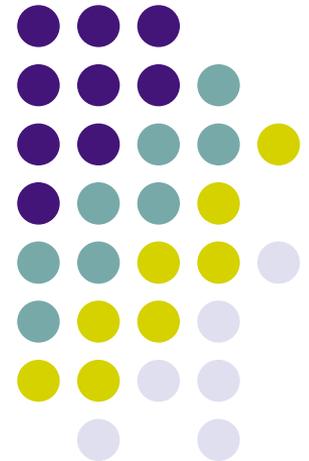
# Tabela 2-2- Idade e audiência de programas religiosos, considerada a educação



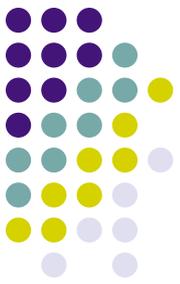
Ouvem Programas Religiosos	Boa educação		Pouca educação	
	Jovens	Idosos	Jovens	Idosos
Sim	9%	11%	29%	32%
Não	91%	89%	71%	68%
Porcentagem Total	100	100	100	100

# Em qual sentido deve se calcular as porcentagens?

Regra de causa e efeito



## Quadro 1 – mortalidade por câncer, segundo a raça, nos EUA (números)



Raça	Causa da morte		
	Câncer	Outras	Total
Branca	136,627	1,055,804	1,195,431
Negra	9,182	169,391	178,573
Total	148,809	1,225,195	1,374,004

## Quadro 2 – Mortalidade por câncer, segundo a raça, nos EUA (porcentagem na vertical)



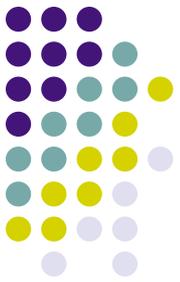
Raça	Causa da morte		
	Câncer %	Outras %	Total %
Branca	93,8	86,2	87,0
Negra	6,2	13,8	13,0
Total	100,0	100,0	100,0

## Quadro 3 – Mortalidade por câncer, segundo a raça, nos EUA (porcentagem na horizontal)

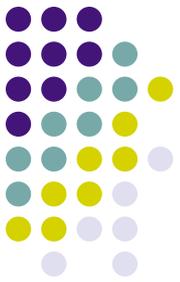


Raça	Causa da morte		
	Câncer %	Outras %	Total %
Branca	11,7	88,3	100,0
Negra	5,1	94,9	100,0
Total	10,7	89,3	100,0

# Comentários

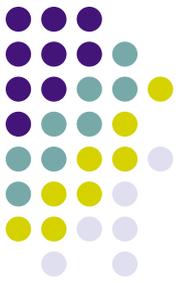


- O conteúdo dos 2 quadros é o mesmo
- Calculando a porcentagem em diferentes sentidos, se acentuam diferentes distribuições e se oferecem diferentes comparações
- O quadro 2 indica a proporção de negros e brancos entre os dois grupos: os que morreram de câncer e os que morreram de outras causas.
- Mas se queremos estudar os efeitos que a raça (causa) tem na proporção de mortes por câncer, o quadro 3 é mais apropriado: ele mostra que a incidência de morte por câncer entre os negros é menos da metade que entre os brancos. Estas proporções não podem ser encontradas no quadro 2.



- A porcentagem pode ser calculada em qualquer sentido

# comentários

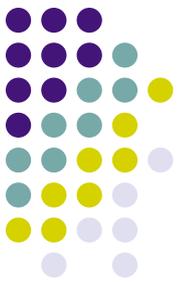


- A regra de causa e efeito nos leva a conclusão que de a porcentagem pode ser calculada em qualquer sentido
- O quadro a seguir mostra uma amostra representativa de audição de várias categorias de renda segundo programa preferido.

# Quadro II-4 - Programas favoritos de rádio segundo nível econômico



Programa	Classe de Renda				Total
	A	B	C	D	
Variedades	85	288	432	202	<b>1,007</b>
Séries	33	108	175	75	<b>391</b>
Noticiários	44	90	87	32	<b>253</b>
Teatro	15	66	79	30	<b>190</b>
Música clássica	25	65	66	6	<b>162</b>
Temas religiosos	4	15	23	11	<b>53</b>
Música de Baile	107	331	443	201	<b>1,082</b>
<b>Total</b>	<b>313</b>	<b>963</b>	<b>1,305</b>	<b>557</b>	<b>3,138</b>



# Comentários

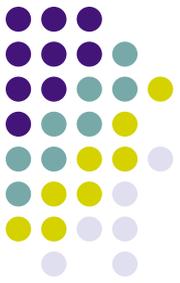
- Se desejamos conhecer o efeito da posição econômica (causa) na preferência por programas de rádio, então a porcentagem deve ser calculada verticalmente, no sentido da posição econômica, i.e, no sentido da causa hipotética.

# Quadro II-5 - Programas favoritos de rádio segundo nível econômico



Programa	Classe de Renda				Total
	A	B	C	D	
Variedades	27	30	33	36	32
Séries	11	11	13	14	12
Noticiários	14	9	7	6	8
Teatro	5	7	6	5	6
Música clássica	8	7	5	1	5
Temas religiosos	1	1	2	2	2
Música de Baile	34	35	34	36	35
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Número de casos	313	963	1,305	557	3,138

# Comentários



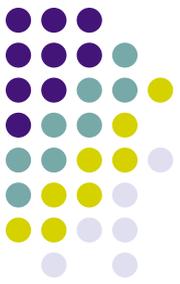
- Partido desse quadro é possível tirar algumas conclusões:
  - Os programas de variedade e de séries aumentam em popularidade a medida que diminui a posição econômica,
  - Mas os noticiários e os programas de música clássica mostram uma tendência contrária, quanto mais alto o nível econômico mais gostam
- Entretanto a porcentagem pode ser calculada no outro sentido. O tipo de programa poderia ser considerado a causa: atrai distintas proporções de ouvintes de outros grupos econômicos.
- Não seria inteligente entender que o programa de rádio fosse a causa que afeta a situação econômica dos ouvintes, unicamente sugerimos que as proporções de audição de cada um dos programas são resultados da atração de distintos programas.

# Quadro II-6 – Programa favorito segundo o nível econômico



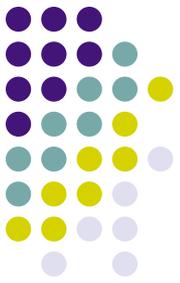
Renda	Baile	Variedades	Séries	Noticiários	Teatro	Música Clássica	Temas Religiosos
A	10	8	8	17	8	15	8
B	30	29	28	36	35	40	28
C	41	43	45	34	41	41	43
D	19	20	19	13	16	4	21
Total	100	100	100	100	100	100	100
Número	1,082	1,007	391	253	190	162	53

# Comentários



- No quadro II-6 mostra-se claramente que a composição econômica dos ouvintes para cada tipo de programa:
  - Os noticiários e a música clássica se encontram entre os programas que tem um atrativo relativamente mais alto para as classes sociais superiores
  - As séries, o teatro, música de baile e os programas religiosos tem uma audiência maior entre os grupos sociais de menor recursos econômicos.

# Outro exemplo, qualquer um pode ser causa



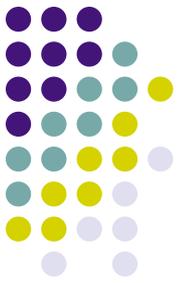
- Amostra aleatória de mulheres foi perguntada quais remédios que preferiam para quatro doenças escolhidas.

# Quadro II-7 Remédios para quatro doenças selecionadas



Doenças	Remédios				
	AA	BB	CC	Outros	Total
Neuralgia	107	47	198	24	376
Resfriado	98	41	401	30	570
Azia	302	60	-----	23	385
Dor de cabeça	-----	242	210	26	478
Nº total de remédios mencionados	507	390	809	103	1809

# Comentários



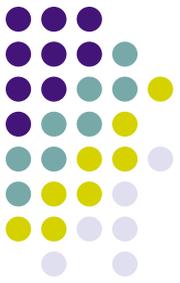
- O quadro seguinte foi analisado a doença como causa provável das distintas proporções dos diversos remédios empregados para cada doença.
- Este quadro responde a pergunta: Qual é a importância relativa de cada remédio para cada doença?
  - Em 53% dos casos de neuralgia utilizou-se o remédio CC e no mercado de resfriados que é o maior, este remédio alcança 71%.
  - Metade dos que tem dor de cabeça utilizam o BB

# Quadro II-8 Remédios preferidos para quatro doenças



Remédios	Neuralgia %	Resfriado %	Azia %	Dor de cabeça %
AA	28	17	78	-----
BB	13	7	16	50
CC	53	71	-----	44
Outros	6	5	6	6
Total	100	100	100	100

# Comentários



- Mas é possível considerar os dados de um outro ponto de vista: o do próprio remédio com suas qualidades e efeitos inerentes; pode se analisar como causas prováveis de sua escolha para determinar as doenças.
- O quadro responde a pergunta: Qual é a importância de cada doença para cada remédio?
  - Os deptos de MKT e de pesquisa estariam interessados nesta forma de quadro.
  - O remédio AA é utilizado principalmente para acidez e nada para dor de cabeça
  - O medicamento BB para dor de cabeça e etc..

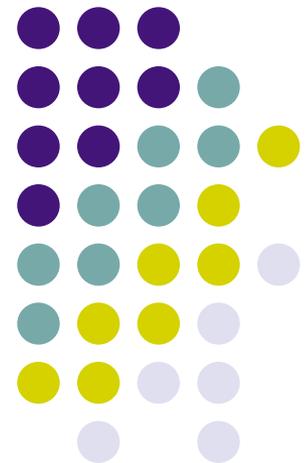
# Quadro II-9 Doenças para as quais se empregam os remédios



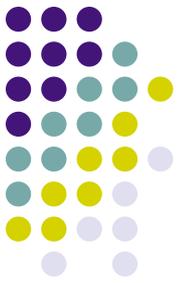
Doenças	Remédios			
	AA %	BB %	CC %	Outros %
Neuralgia	21	12	24	23
Resfriado	19	11	50	29
Azia	60	15	---	22
Dor de cabeça	---	62	26	26
<b>Total</b>	100	100	100	100

---

Nem sempre a regra de causa e  
efeito é aplicável

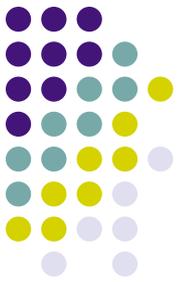


# Exemplo que a regra de causa e efeito não é aplicável



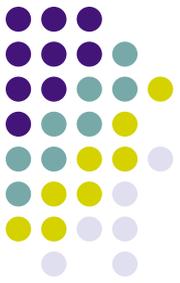
- Com a finalidade de poder calcular a força relativa de dois partidos políticos antes de uma eleição, realizou um pesquisa de intenção de votos nos dois estados
- Em cada estado se realizaram 8000 entrevistas e o resultado foi o seguinte

# Quadro II-14 Voto por determinado partido em dois estados



Estado	Partido XX	Partido YY	Total
A	2500	5500	8000
B	3500	4500	8000
Total	6000	10000	16000

# Quadro II-15 Importância dos partidos para o estado

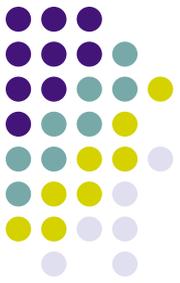


Calculou-se a porcentagem na direção do estado. Supõe-se que a força relativa do partido é o **estado**, i.é., a estrutura social e econômica e atitude política dos seus habitantes.

O quadro indica que o partido YY é mais forte nos dois estados e relativamente mais forte no estado A.

Partido	Estado A %	Estado B %
XX	31	44
YY	69	56
Total	100	100
Nº de votos	8000	8000

# Comentários

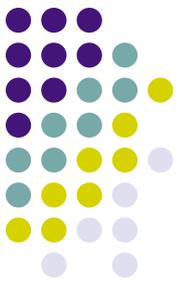


- O quadro seguinte parece indicar que 42% dos votos a favor do partido X vem do estado A e que 55% dos votos do mesmo estado são favoráveis ao partido Y

# Quadro II-16 Cálculo das porcentagens no outro sentido (vertical)



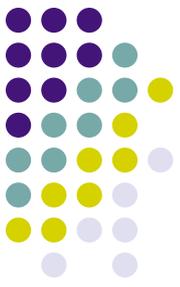
Estado	Partido XX %	Partido YY %
A	42	55
B	58	45
Total	100	100
Nº de casos	6000	10000



# Comentários

- Mas uma consideração indica que o quadro não pode demonstrar que pretende, ou seja a contribuição relativa dos estados A e B para a força total dos partidos.
- É verdade que 42% dos votos da amostra são favoráveis ao partido XX no estado A mas este 42% depende, na realidade de um fator arbitrário: o tamanho relativo da amostra nos estados
- Para esclarecer este ponto, vamos supor que foram feitas 8000 entrevistas no estado A e que a população votante do estado B é duas vezes maior do que a de A e que fizeram 16000 entrevistas no lugar de 8000
- Suponhamos também que com a amostra aumentada não há diferença na distribuição do partido.
- O quadro seguinte ficaria da seguinte forma.

# Quadro II-17 Importância do estado para o Partido



Nesse quadro podemos ver que o partido XX obtém 26% dos seus votos no estado A, comparado com o 42% dos quadro anterior.

Com isso se demonstra que no tocante as porcentagens destes quadros, existe o controle do número relativo de entrevistas por estado.

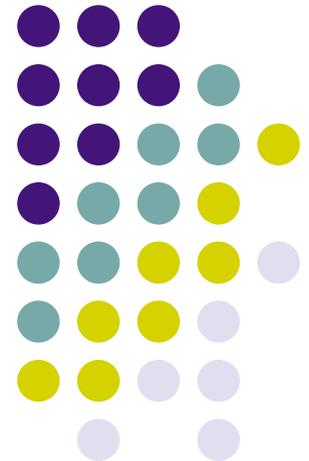
No caso do quadro II-15, independente do número de entrevistas, as proporções não mudariam materialmente.

Os números do quadro II-16 são enganosos porque foram calculados no sentido que a amostra não é representativa da amostra

Estado	Partido XX		Partido YY		Total	
	N	%	N	%	N	%
A	(2500)	26	(5500)	38	(8000)	33
B	(7000)	74	(9000)	62	(16000)	67
Total	(9500)	100	(14500)	100	(24000)	100

# Tabela cruzada refina os conceitos

---

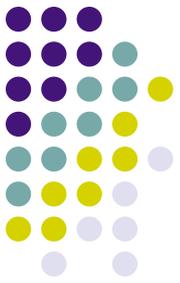


# TABELA VIII – 1. Amostra de votação no condado X antes da eleição presidencial



<b>Votarão no:</b>	<b>Porcentagem</b>
<b>Candidato Republicano</b>	52%
<b>Candidato Democrático</b>	48%
<b>Total</b>	100%
<b>No de Casos</b>	5.160

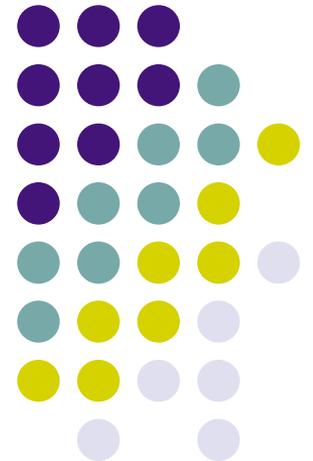
## TABELA VIII – 2. Votação antes da eleição no condado X, segundo o status econômico



	<b>Status econômico</b>	
	<b>Alto %</b>	<b>Baixo %</b>
<b>Republicano</b>	<b>60</b>	45
<b>Democrático</b>	40	<b>55</b>
<b>Total</b>	100	100
<b>No de casos</b>	2.604	2.556

---

# Outro exemplo



## TABELA VIII – 3. Taxa de acidentes entre motoristas



	<b>%</b>
<b>Nunca teve um acidente enquanto dirigia</b>	<b>62</b>
<b>Teve pelo menos um acidente enquanto dirigia</b>	<b>38</b>
<b>Total</b>	<b>100</b>
<b>No de casos</b>	<b>14.030</b>

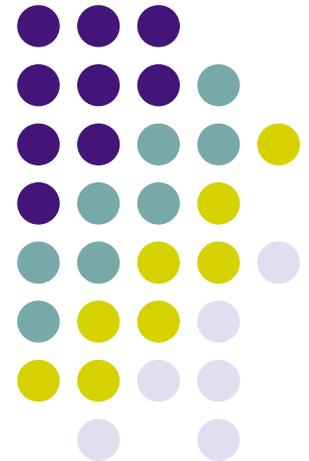
## TABELA VIII – 4. Taxa de acidentes de automóveis por sexo



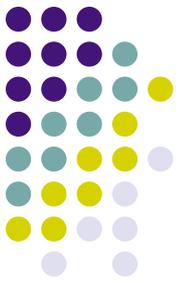
	<b>Homens</b> <b>%</b>	<b>Mulheres</b> <b>%</b>
<b>Nunca teve um acidente enquanto dirigia</b>	56	68
<b>Teve pelo menos um acidente enquanto dirigia</b>	44	32
<b>Total</b>	100	100
<b>Nº de casos</b>	7.080	6.950

# Tabela cruzada refina os conceitos

Tipo de tabulação cruzada

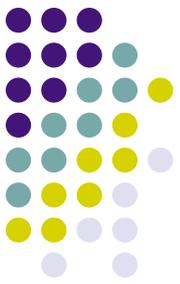


# Tipo de tabulação cruzada



- A análise por um fator único é a forma mais simples de tabulação cruzada
- O procedimento correto é introduzir cada fator adicional não como uma alternativa para, mas simultaneamente com os outros fatores, de modo que todas as inter-relações possíveis dos outros fatores possam ser estudadas

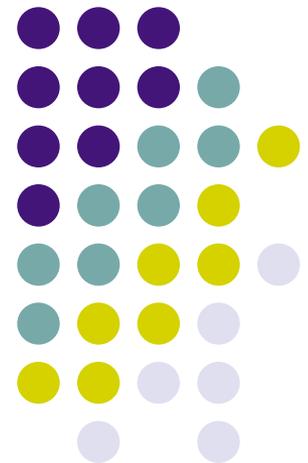
# Tipos de tabulação cruzada

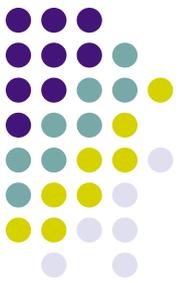


- A introdução simultânea pode produzir os seguintes efeitos:
  - Pode refinar os resultados da simples tabulação cruzada
  - Pode não lograr refinar os resultados da simples tabulação cruzada, mas pode revelar o efeito independente de um terceiro fator.
  - Pode explicar os resultados da simples tabulação cruzada:
    - Confirmando a interpretação original
    - Identificando essa interpretação original como espúria.

---

# Um fator adicional refina a correlação

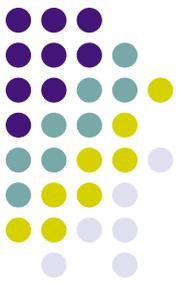




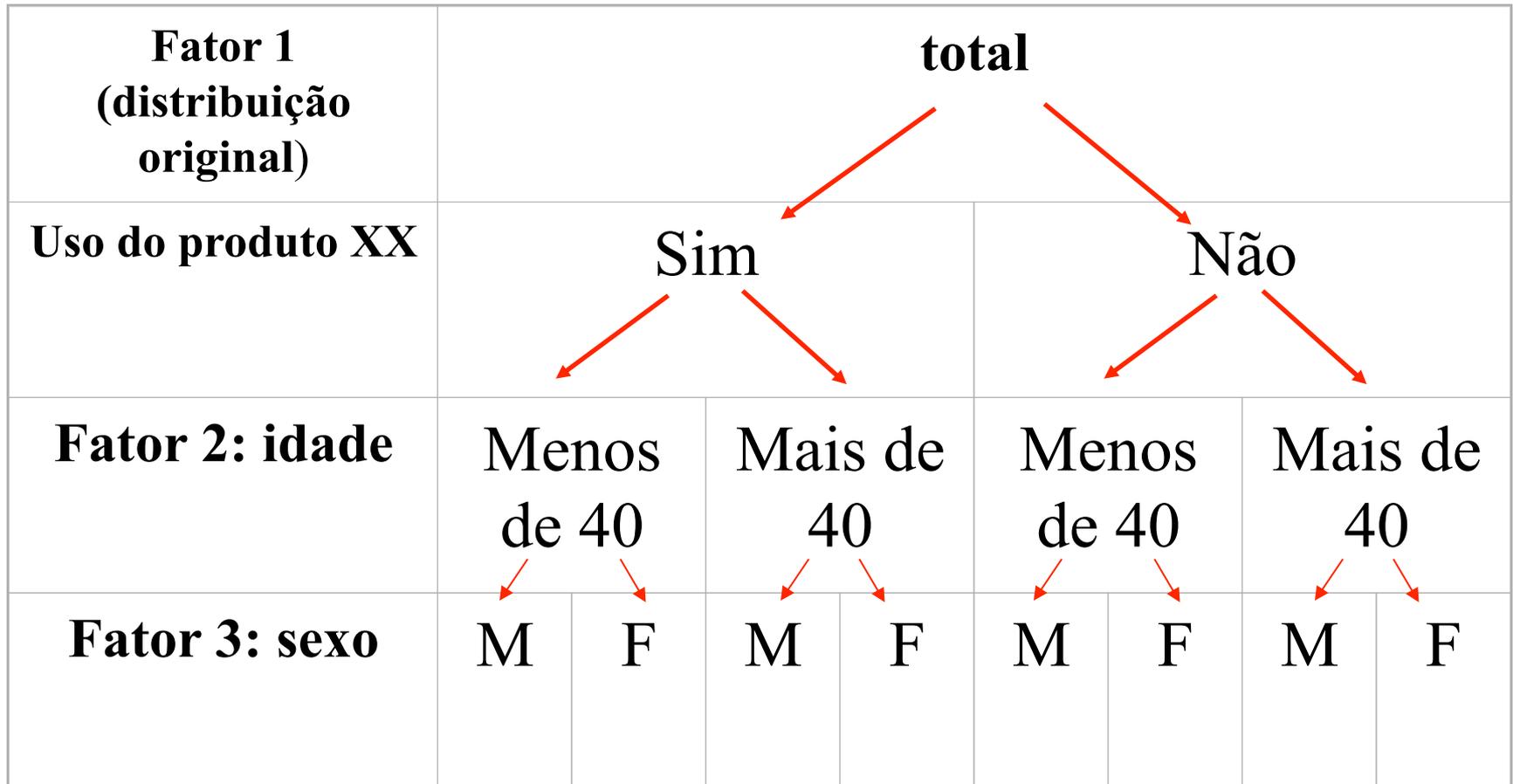
## TABELA VIII – 5. Uso de alimento XX no desjejum, segundo a idade:

	<b>Menos de 40 anos</b> %	<b>Mais de 40 anos</b> %
<b>Usam XX</b>	28	20
<b>Não usam XX</b>	72	80
<b>total</b>	100	100
<b>Nº de casos</b>	1.224	952

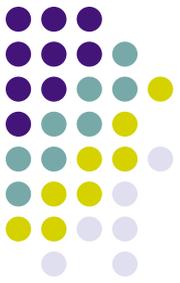
# Esquema de separação



O investigador imaginou o sexo como fator adicional que poderia ter influência sobre o uso do alimento em foco



# TABELA VIII – 6. Uso do alimento XX por Idade e Sexo



A tabela III-5 mostrou uma relação entre idade e uso de XX

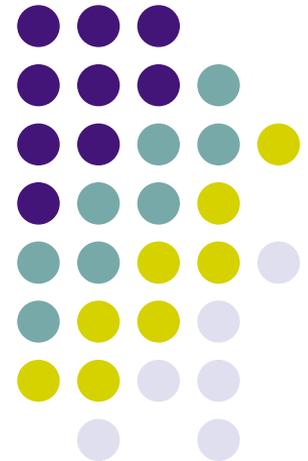
A tabela III-6 refina este conhecimento, mostrando como isto difere para os dois sexos.

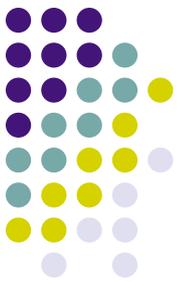
Ela indica que a correlação é maior para os homens do que para mulheres

	<b>Homens</b>		<b>Mulheres</b>	
	<b>Menos de 40</b>	<b>Mais de 40</b>	<b>Menos de 40</b>	<b>Mais de 40</b>
<b>Comem XX</b>	<b>36%</b>	<b>23%</b>	<b>20%</b>	<b>17%</b>
<b>Nº de casos</b>	<b>619</b>	<b>480</b>	<b>605</b>	<b>472</b>

---

# Correlação próxima de zero



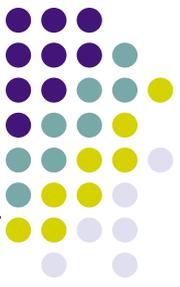


## TABELA VIII -7. Audição de música segundo a idade

Os casos em que a correlação é zero ou próximo de zero são de especial interesse. Somente pela introdução de um terceiro fator pode tornar-se visível a inter-relação dos fatores envolvidos

Não aparece correlação entre idade e música clássica, contudo quando se introduziu na análise nível educacional como fator adicional, obteve-se a tabela na página seguinte.

	<b>Abaixo de 40 anos</b>	<b>40 anos ou mais</b>
<b>Nº de casos</b>	603	676
<b>Ouvem musica clássica</b>	64%	64%



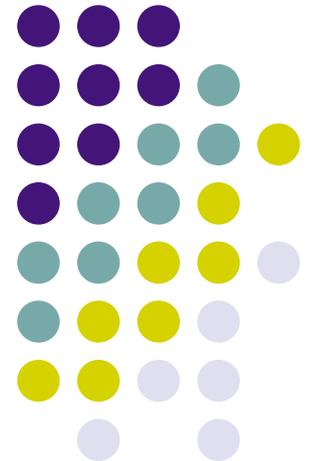
## TABELA VIII- 8. Audição de música clássica segundo a idade e o nível educacional.

A introdução do nível educacional mostra que existe de fato uma correlação entre idade e audição. As pessoas instruídas ouvem mais música clássica quando são mais velhas e acontece justamente o contrário com pessoas de baixo nível educacional: ouvem mais quando são jovens

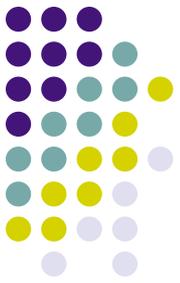
	Nível educacional alto		Nível educacional baixo	
	-40	+40	-40	+40
Ouvem música clássica	73%	78%	61%	56%
Nº de casos	224	251	379	425

---

# Um fator adicional revela condições limitantes



# Definição



- O refinamento trazido pelo terceiro fator, algumas vezes consiste em revelar
  - que determinada correlação tende a desaparecer sob condições especiais e,
  - correspondentemente, a aumentar na ausência dessas condições

## TABELA VIII – 11. Taxa anual de suicídio por milhão, segundo a religião

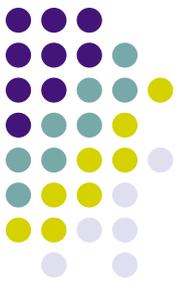


O número significa um em milhão. Mostra que os católicos tendem a apresentar taxas de suicídio mais baixa que os protestantes

Todavia as taxas foram computadas para as duas religiões separadamente para os habitantes de zona rural e urbana e temos o quadro da página seguinte

Católicos	Protestantes
198	396

# TABELA VIII – 12. Taxa anual de suicídio por milhão, segundo a religião e o tamanho da comunidade



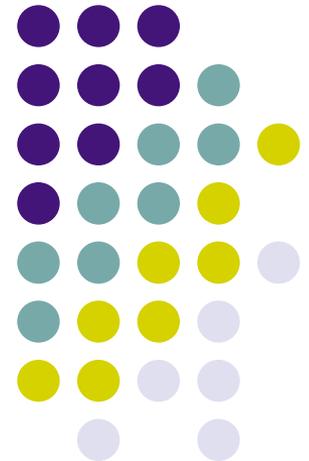
A introdução do tipo de comunidade refinou nosso conhecimento: a diferença entre as taxas do suicídio é bem maior nas áreas rurais que em comunidades urbanas

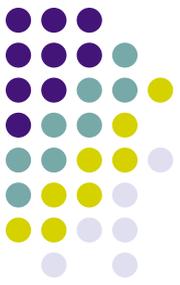
A taxa de suicídio, entre católicos, é grandemente afetada pelo tipo de comunidade: aumenta bastante nas áreas urbanas. A taxa entre os protestantes é apenas ligeiramente influenciada pelo tamanho da comunidade.

Rural		Urbana	
Católicos	Protestantes	Católicos	Protestantes
88	414	309	378

---

Fator adicional tem um efeito independente





# Conceito

- Algumas vezes a introdução de um terceiro fator não tem efeito na correlação original.
- No entanto, o terceiro fator pode ter uma influencia independente sobre considerado como o efeito na correlação original
- Voltamos aos quadros já discutidos

# TABELA VIII – 1. Amostra de votação no condado X antes da eleição presidencial



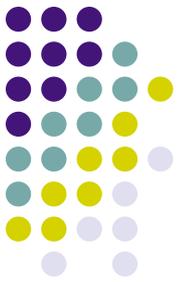
<b>Votarão no:</b>	<b>Porcentagem</b>
<b>Candidato Republicano</b>	52%
<b>Candidato Democrático</b>	48%
<b>Total</b>	100%
<b>No de Casos</b>	5.160

## TABELA VIII – 2. Votação antes da eleição no condado X, segundo o status econômico



	<b>Status econômico</b>	
	<b>Alto %</b>	<b>Baixo %</b>
<b>Republicano</b>	<b>60</b>	45
<b>Democrático</b>	40	<b>55</b>
<b>Total</b>	100	100
<b>No de casos</b>	2.604	2.556

# Comentários



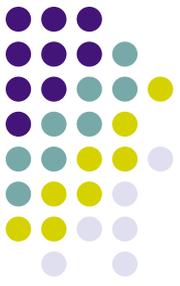
- Se introduzirmos a religião na tabela anterior (votação segundo o status econômico) como um terceiro fator, obtemos a seguinte tabela



# TABELA VIII – 13. Votação antes das eleições no Candidato X, segundo o status econômico e a religião:

Voto provável	Status econômico			
	Alto		Baixo	
	Católicos	Protestantes	Católicos	Protestantes
<b>Republicano</b>	27%	69%	19%	52%
<b>Democrata</b>	73%	31%	81%	48%
<b>Casos</b>	538	2.066	587	1.969

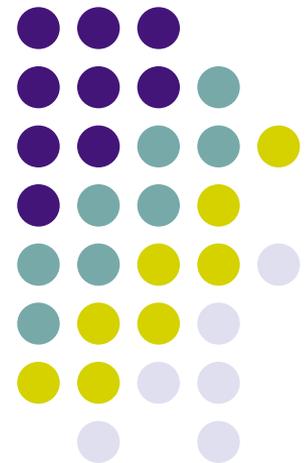
# Comentários

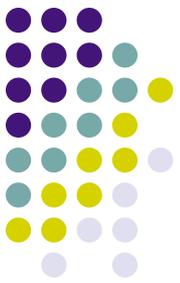


- Há menos probabilidade de os católicos votarem no candidato republicano, do que os protestantes
- Mas a correlação original entre status econômico e a votação é confirmada para ambos os grupos religiosos

---

# Sumário

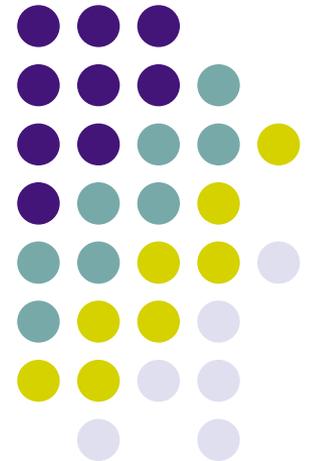




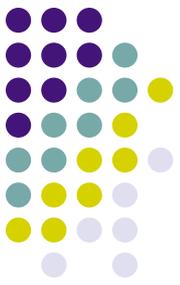
- O refinamento produzido pela introdução de um terceiro fator resultado do fato de que a correlação original é analisada sob diferentes condições
- O refinamento consiste no fato que a reação original mudou sob essas diferentes condições

# Tabulação cruzada

Explicações verdadeiras e  
espúrias



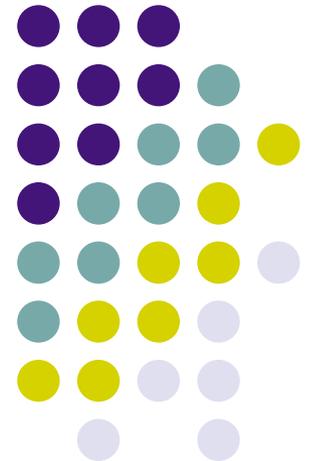
# Comentários



- Os exemplos estudados até agora nenhum caso o fator adicional se relaciona com o fator causal na correlação original
- A falta de correlação do terceiro fator com o fator causal na correlação original é o critério essencial para a análise do refinamento
- Nos casos em que o terceiro fator explica a correlação original, surge a relação inversa: em todos os casos que envolvem uma explicação o fator adicional relaciona-se também com o fator considerado como causal na correlação original
- A função explanatória do terceiro fator introduzido consiste no fato real de que ele se relaciona com ambos os fatores da correlação original e não apenas com um deles

# Tabulação cruzada

O fator adicional explica a correlação



## TABELA VIII – 3. Taxa de acidentes entre motoristas



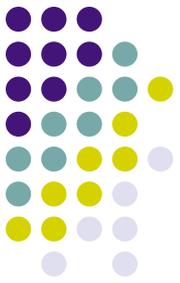
	<b>%</b>
<b>Nunca teve um acidente enquanto dirigia</b>	<b>62</b>
<b>Teve pelo menos um acidente enquanto dirigia</b>	<b>38</b>
<b>Total</b>	<b>100</b>
<b>No de casos</b>	<b>14.030</b>

## TABELA VIII – 4. Taxa de acidentes de automóveis por sexo



	<b>Homens</b> <b>%</b>	<b>Mulheres</b> <b>%</b>
<b>Nunca teve um acidente enquanto dirigia</b>	56	68
<b>Teve pelo menos um acidente enquanto dirigia</b>	44	32
<b>Total</b>	100	100
<b>Nº de casos</b>	7.080	6.950

# Comentários



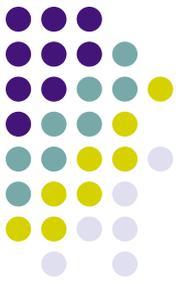
- Isto parece contradizer a opinião amplamente difundida de que um homem ao volante constitui menos perigo que uma mulher.
- Procurando uma explicação para esta estranha correlação, pode-se introduzir um terceiro fator: quantidade de milhas dirigidas

# TABELA IX – 1. Acidente de automóvel de motoristas de sexo masculino e feminino, segundo o número de milhas percorridas por ano.



	<b>Homens</b>		<b>Mulheres</b>	
	Guiaram mais de 1000 milhas %	Guiaram 1000 milhas ou menos %	Guiaram mais de 1000 milhas %	Guiaram 1000 milhas ou menos %
<b>Pelo menos 1 acidente enquanto dirigiam</b>	52	25	52	25
<b>Nunca tiveram um acidente</b>	48	75	48	75
<b>Total</b>	100	100	100	100
<b>N.º de casos</b>	5010	2070	1915	5035

# Comentários



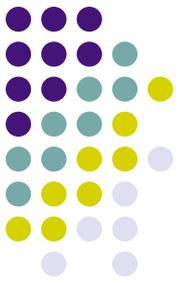
- A correlação entre sexo e taxa de acidente desapareceu. Se compararmos as taxas de acidente dos motoristas que guiaram a mesma quilometragem, não observamos diferença.
- Mas outra correlação se impõe: motoristas com maior média de quilometragem anual tem maior possibilidade de acidente
- Como foi possível? Se observamos a tabela seguinte observamos que o número de motoristas com maior quilometragem é maior no grupo masculino do que no feminino.



## TABELA IX – 1b – sexo do motorista segundo a quilometragem anual

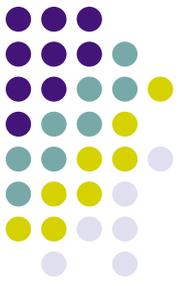
	Homens		Mulheres	
	Nº	%	Nº	%
<b>Guiaram mais de 1.000 milhas</b>	5010	71	1915	28
<b>Guiaram menos de 1.000 milhas</b>	2070	29	5035	72
<b>Total</b>	7080	100	6950	100

# Comentários



- Sexo e quilometragem estão correlacionados e isto reflete tres correlações interdependentes:
  - O sexo está correlacionado com taxa de acidentes
  - As taxas de acidente estão correlacionadas com quilometragem percorrida
  - O sexo está correlacionado com a quilometragem média

# Por que os homens tem mais acidentes?

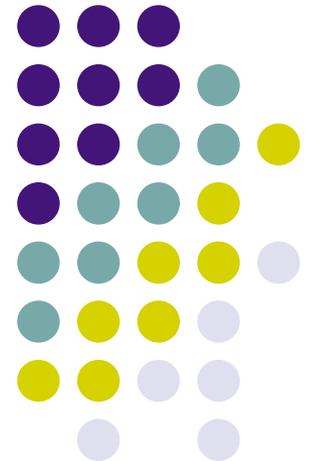


- A quilometragem maiores levam a mais acidentes e
- Comparativamente há mais homens que mulheres que guiam maiores quilometragens.
- Por causa da coincidência dessas duas relações, o quadro original apresentou correlação entre sexo e taxa de acidente
  - Toda a diferença original entre taxa de acidente e sexo foi eliminada com a quilometragem.
  - Não existem diferenças quando se comparam homens e mulheres que dirigiram a mesma quilometragem

# Tabulação cruzada

---

Explicação parcial





## TABELA IX –2. Proporção de ausência em dias de trabalho

As mulheres casadas registram taxa de faltas no trabalho mais alta que as solteiras.

É possível que as mulheres casadas tenham mais serviço doméstico a fazer?

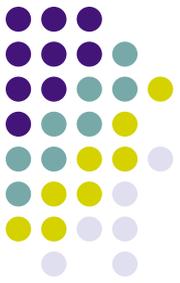
	<b>Mulheres casadas %</b>	<b>Mulheres solteiras %</b>
<b>Ausentes</b>	<b>6,4</b>	<b>2,3</b>
<b>Presentes</b>	93,6	97,7
<b>Total</b>	100	100
<b>No de casos</b>	6496	10560

## TABELA IX – 3. Proporção de ausência em dias de trabalho, segundo o estado civil e a quantidade de trabalho doméstico.



	<b>Mulheres casadas</b> %	<b>Mulheres Solteiras</b> %
<b>Têm muito trabalho doméstico</b>	<b>7</b> (5680)	<b>5,7</b> (1.104)
<b>Têm pouco ou nenhum trabalho doméstico</b>	2,2 (816)	1,9 (9.126)
<b>Total de operárias</b>	6,4 (6496)	2,3 (10.560)

# Comentários



- A tabela anterior indica que o aumento do absenteísmo provém do fato de se ter mais trabalho doméstico não do fato de ser casada
- Qual o motivo?
- A tabela seguinte mostra que as mulheres casadas tem maior trabalho doméstico

# TABELA IX – 3 a . Estado civil e trabalho Doméstico

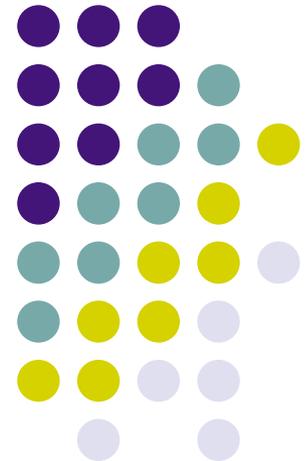


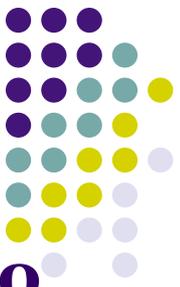
	Mulheres casadas %		Mulheres Solteiras %	
	Nº	%	Nº	%
Têm muito trabalho domestico	(5680)	88	(1.104)	10
Têm pouco ou nenhum trabalho domestico	(816)	12	(9.126)	90
Total de operárias	(6496)	100	(10.560)	100

# Tabulação cruzada

---

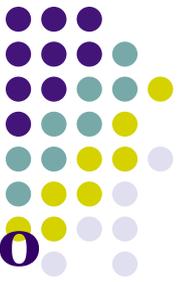
Correlações espúrias





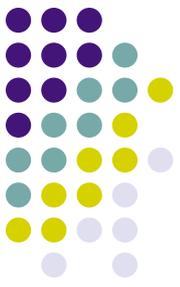
## **TABELA IX – 4. Porcentagem dos que consomem produtos de confeitaria, segundo o estado civil**

	<b>Solteiras</b>	<b>Casadas</b>
<b>Consumem regularmente</b>	75%	63%
<b>Nº de casos</b>	999	2010



## **TABELA IX – 5. Porcentagem do que consomem produtos de confeitaria, segundo a idade:**

	<b>Até 25 anos</b>	<b>25 anos ou mais</b>
<b>Consumem produtos de confeitaria regularmente</b>	80%	58%
<b>Nº de casos</b>	1302	1707



# Comentários

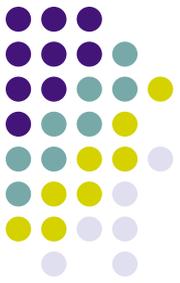
- Estudando as tabelas o leitor verá:
- As solteiras comem mais doces que as casadas
- As mais jovens comem mais doces do que as mais velhas
- Questão: as pessoas comem menos porque são casadas ou por são mais velhas?
- A tabela seguinte mostra a resposta

## TABELA IX – 6. Porcentagem de pessoas que comem produtos de confeitaria por idade e estado civil.



	Casadas		Solteiras	
	Até 25 anos	25 anos ou mais	Até 25 anos	25 anos ou mais
Consumem produtos de confeitaria regularmente	81	58	79	60
Nº de casos	503	1507	799	

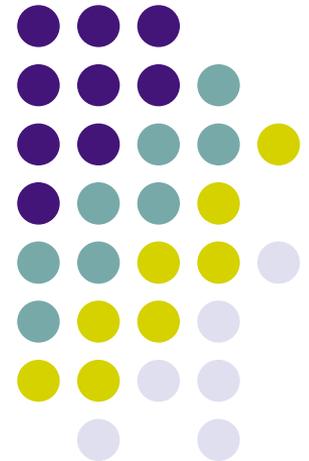
# Comentários



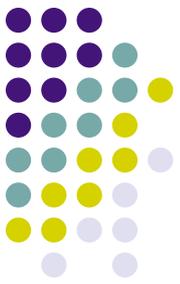
- A relação ser casada e comer menos doces pode ser completamente explicada pelo fato de que as pessoas casadas são, proporcionalmente mais velhas que as solteiras e as pessoas mais velhas comem menos produtos de confeitaria.
- Se compararmos pessoas casadas e solteiras de idade igual a suposta correlação desaparece

# Tabulação cruzada

A correlação se inverte



# TABELA IX – 9. Proporção de maus créditos segundo o preço dos artigos comprados.



Parece indicar que as vendas de artigos de alto preço constituem um menor risco de crédito. Mas considerando o valor da entrada?

	Preço inferior a \$60 %	Preço superior a \$60 %
Maus créditos	12.5	9.9
Bons créditos	87.5	90.1
Total	100	100
Nº casos	4303	4088

# TABELA IX – 10. Proporção de maus créditos, segundo o preço do artigo e a quantia da entrada.

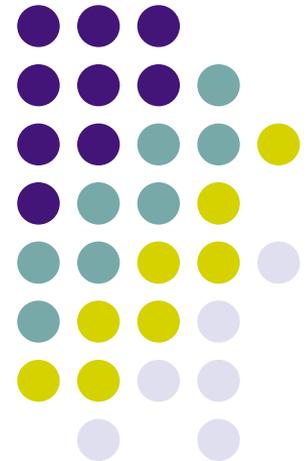


O preço elevado indica pequeno risco na medida em que é acompanhado por uma entrada elevada, como usualmente acontece, sendo que quando a prestação permanece constante, os preços altos indicam maior risco

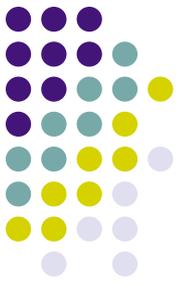
Entrada	Preço inferior a \$60		Preço superior a \$60	
	Menos de \$15	Maior de \$15	Menos de \$15	Maior de \$15
Maus créditos	13.6	6.2	15.1	8.5
Bons créditos	86.4	93.8	89.9	91.5
Total	100	100	100	100
Nº casos	3.655	648	890	3.198

# Tabulação cruzada

sumário



# A introdução de fatores adicionais na análise



- Pode refinar os resultados da simples tabulação cruzada
- Pode não refinar os resultados da simples tabulação cruzada mas pode revelar um efeito independente que esse terceiro fator exerce
- Confirmando o resultado original, ou
- Mostrando ser espúrio o resultado original
- Limitações
  - A redução do tamanho da amostra faz com que as diferenças de porcentagem percam sua significação estatística
  - Os resultados suplementares podem não constituir fundamento para a questão prática a que o projeto de pesquisa deve responder, Nesse caso é desnecessário maior refinamento