

AULA 21

Tipos de desenhos de pesquisas

Ernesto F. L. Amaral

15 de outubro de 2010
Metodologia (DCP 033)

Fonte:

Babbie, Earl. “Métodos de Pesquisas de *Survey*”. 1999. Belo Horizonte: Editora UFMG. pp.93-111.

Wooldridge, Jeffrey M. “Introdução à econometria: uma abordagem moderna”. São Paulo: Cengage Learning, 2008. pp.1-17.

TIPOS DE DESENHOS DE PESQUISA

- Há uma tendência a considerar a análise de dados de *survey* como mais desafiadora e interessante do que o desenho do *survey* e a coleta de dados.
- Durante a análise, começa-se a ganhar entendimento do assunto estudado e pode-se compartilhar as descobertas com colegas.
- No entanto, é preciso que haja a correta elaboração do desenho de pesquisa e sua execução, para que não haja problemas na etapa da análise.
- Portanto, o desenho é igualmente desafiador e importante.
- Há vários tipos de pesquisa de *survey*, com objetivos, custos, tempo e escopo diferenciados.

FINALIDADES DA PESQUISA DE *SURVEY*

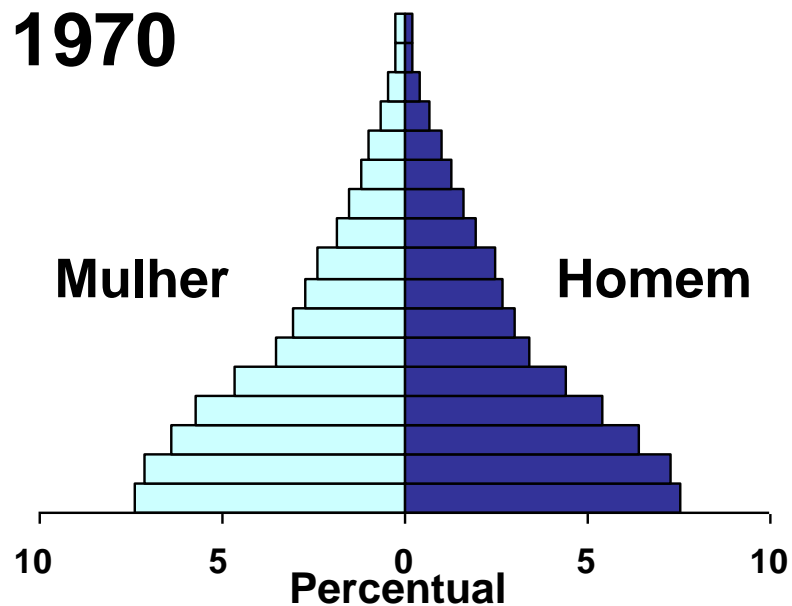
- Uma pesquisa de *survey* pode ser realizada para atender diferentes objetivos:
 - Intenção de voto.
 - Venda de produtos.
 - Projetar infra-estrutura.
 - Modificar programa social.
- De uma forma geral, três objetivos gerais permeiam os interesses de pesquisas de *survey*:
 - Descrição.
 - Explicação.
 - Exploração.

DESCRIÇÃO

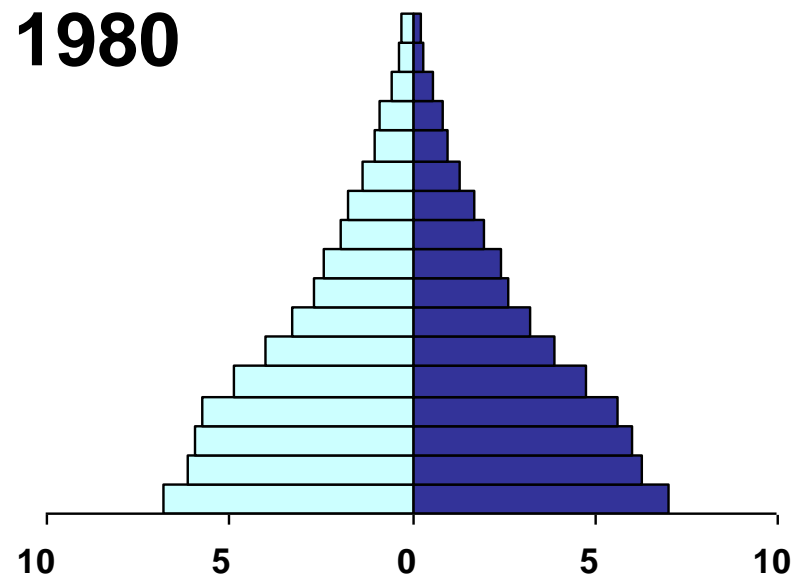
- *Surveys* são frequentemente realizados para descobrir a distribuição de certos traços e atributos da população, permitindo análise descritiva.
- Há preocupação em **como** é a distribuição de determinada variável (descrição), e não no **porquê** (explicação).
- Exemplos de dados descritivos:
 - Distribuições por idade e sexo.
 - Taxas de criminalidade no país.
 - Índice de Desenvolvimento Humano (IDH) dos municípios brasileiros.
- Além de descrever a amostra total (e inferir para a população total), pesquisadores descrevem subgrupos.

ESTRUTURA ETÁRIA NO BRASIL, 1970-2000

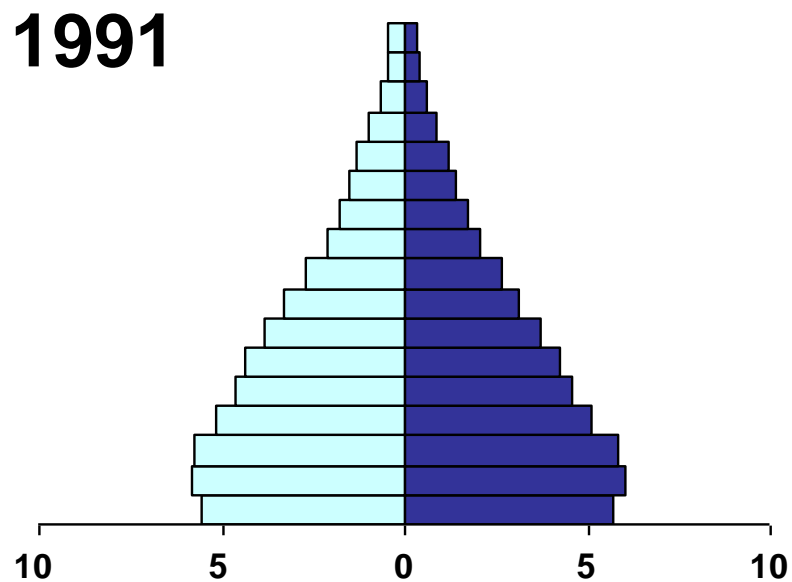
1970



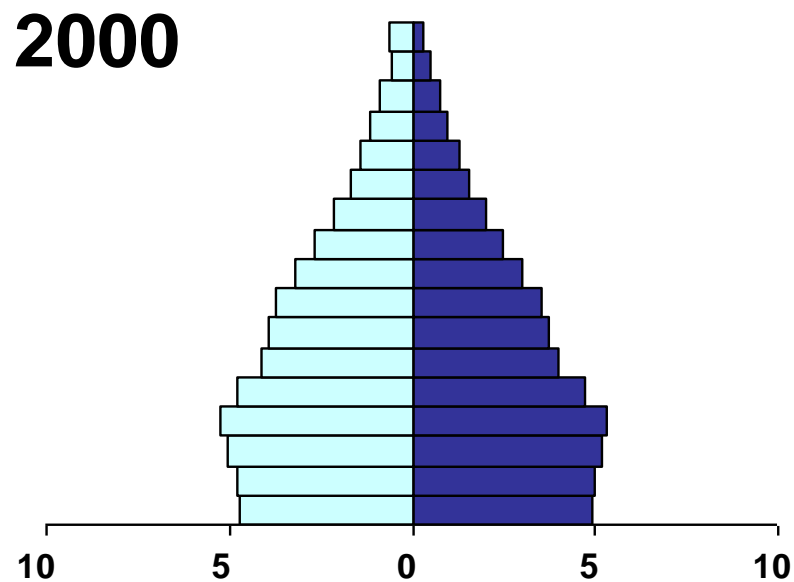
1980



1991

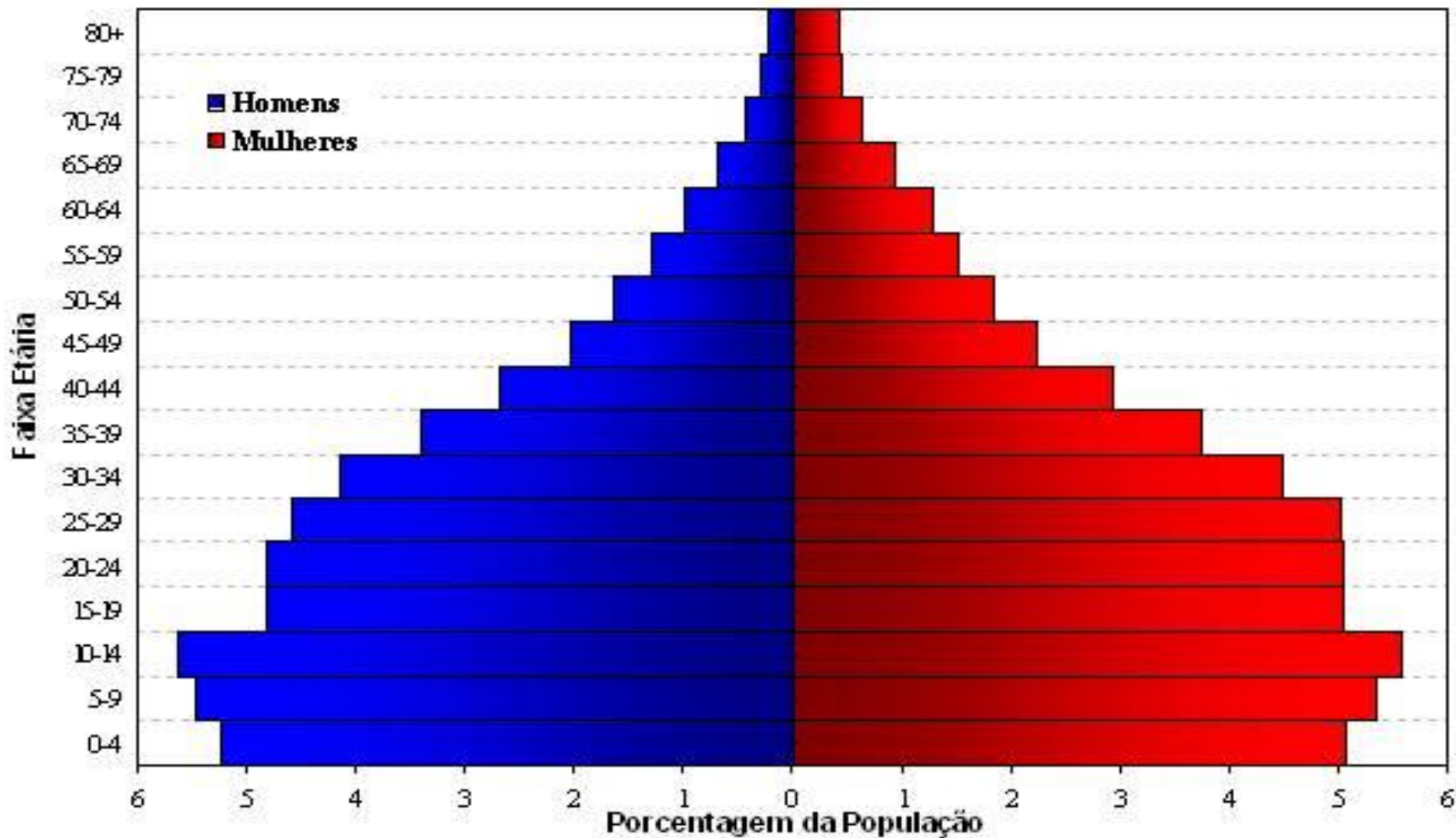


2000



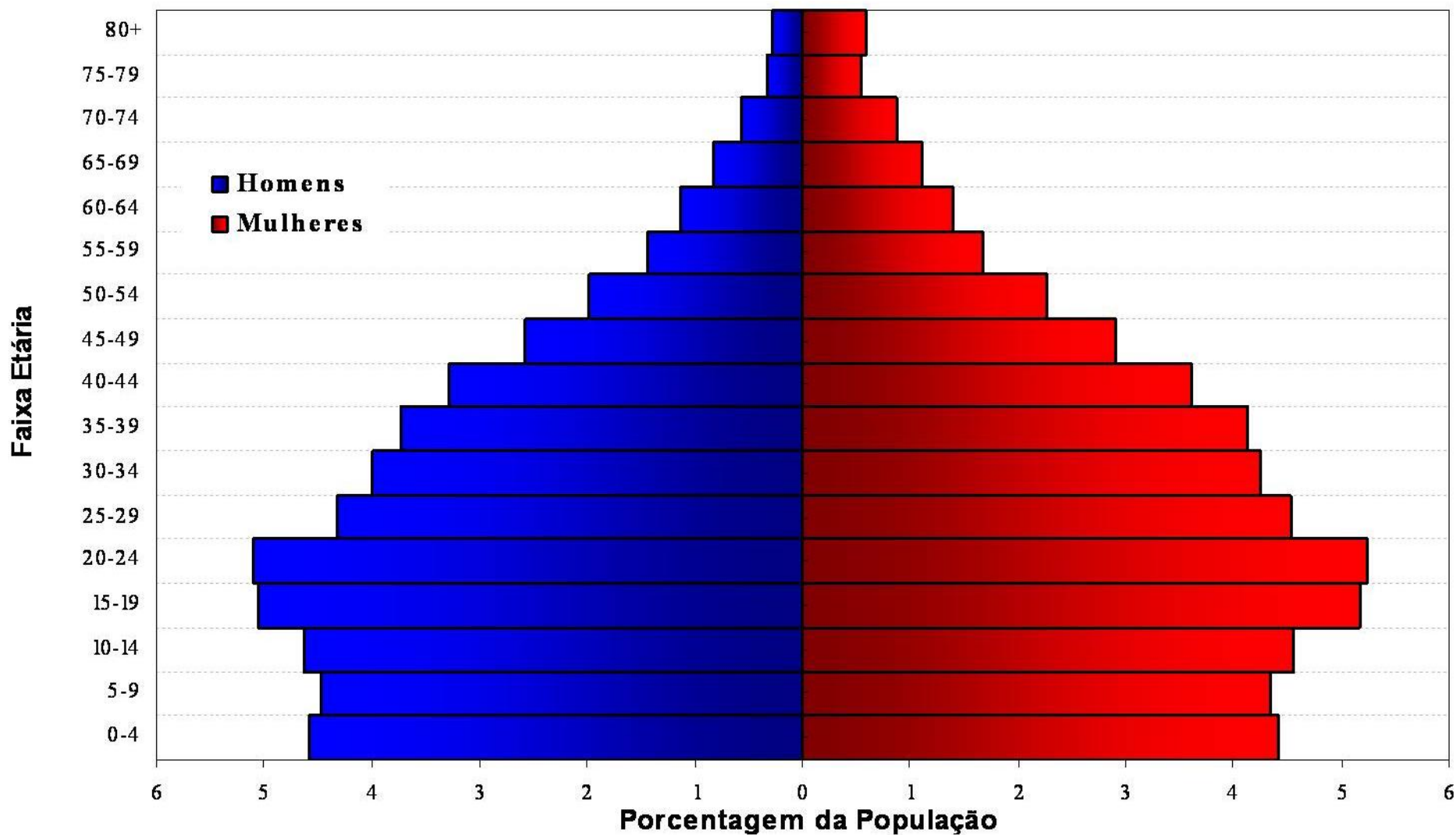
Fonte: Microdados do Censo Demográfico do Brasil (IBGE).

ESTRUTURA ETÁRIA POR SEXO DA RMBH 1991



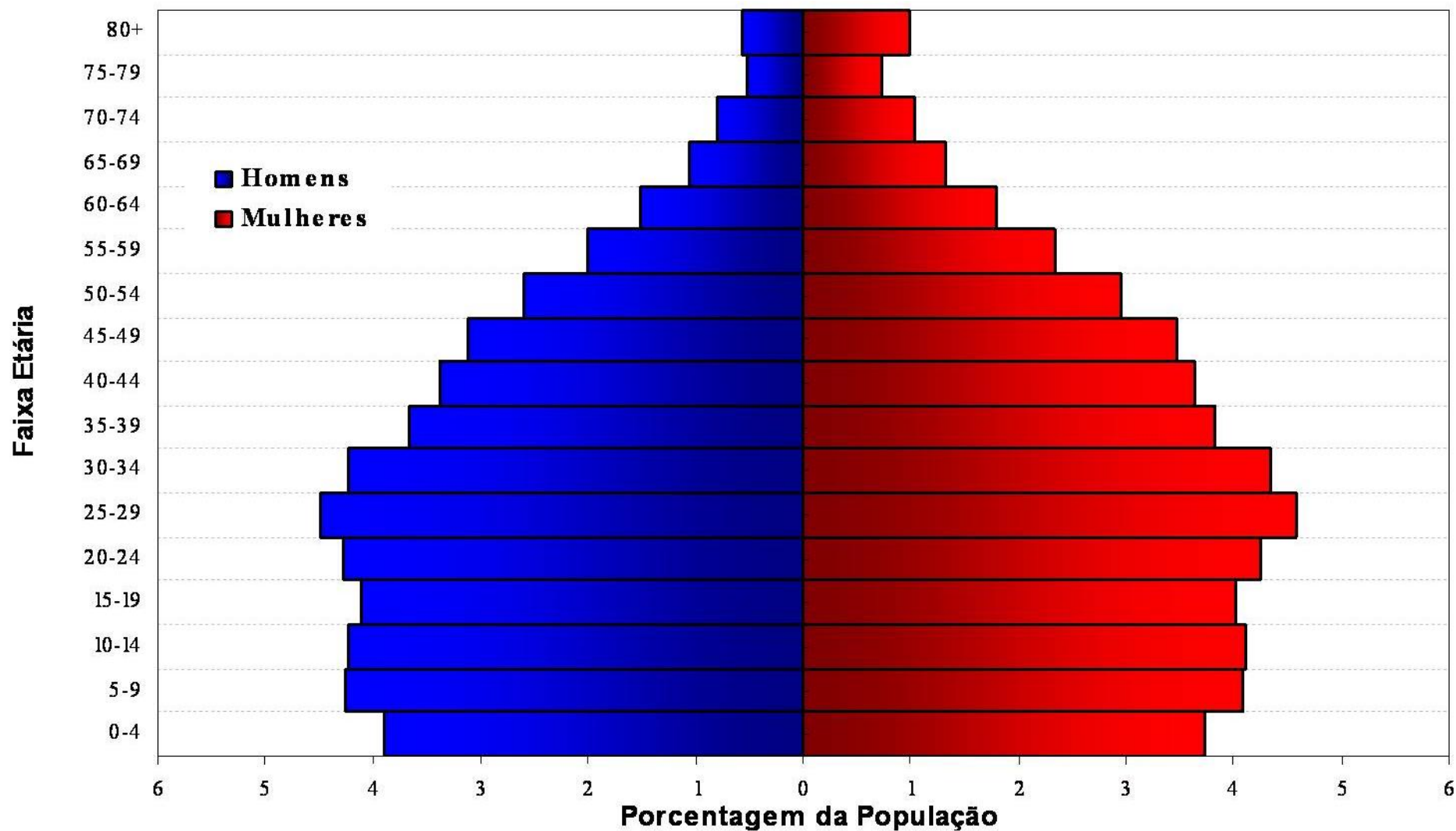
Fonte dos dados básicos: IBGE, PNAD 1996-2006 e Censo 1991-2000.

ESTRUTURA ETÁRIA POR SEXO DA RMBH 2000



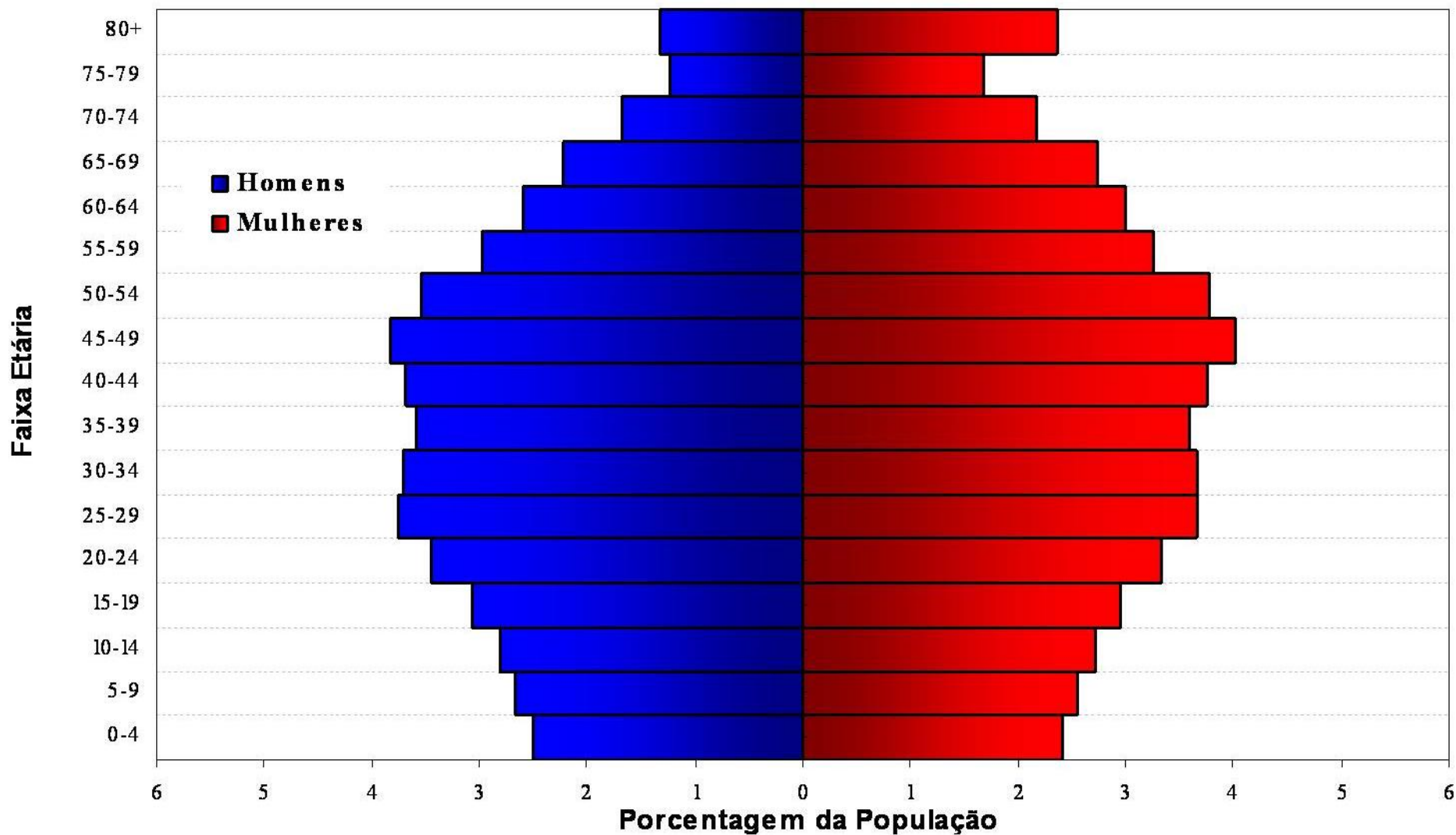
Fonte dos dados básicos: IBGE, PNAD 1996-2006 e Censo 1991-2000.

ESTRUTURA ETÁRIA POR SEXO DA RMBH 2010



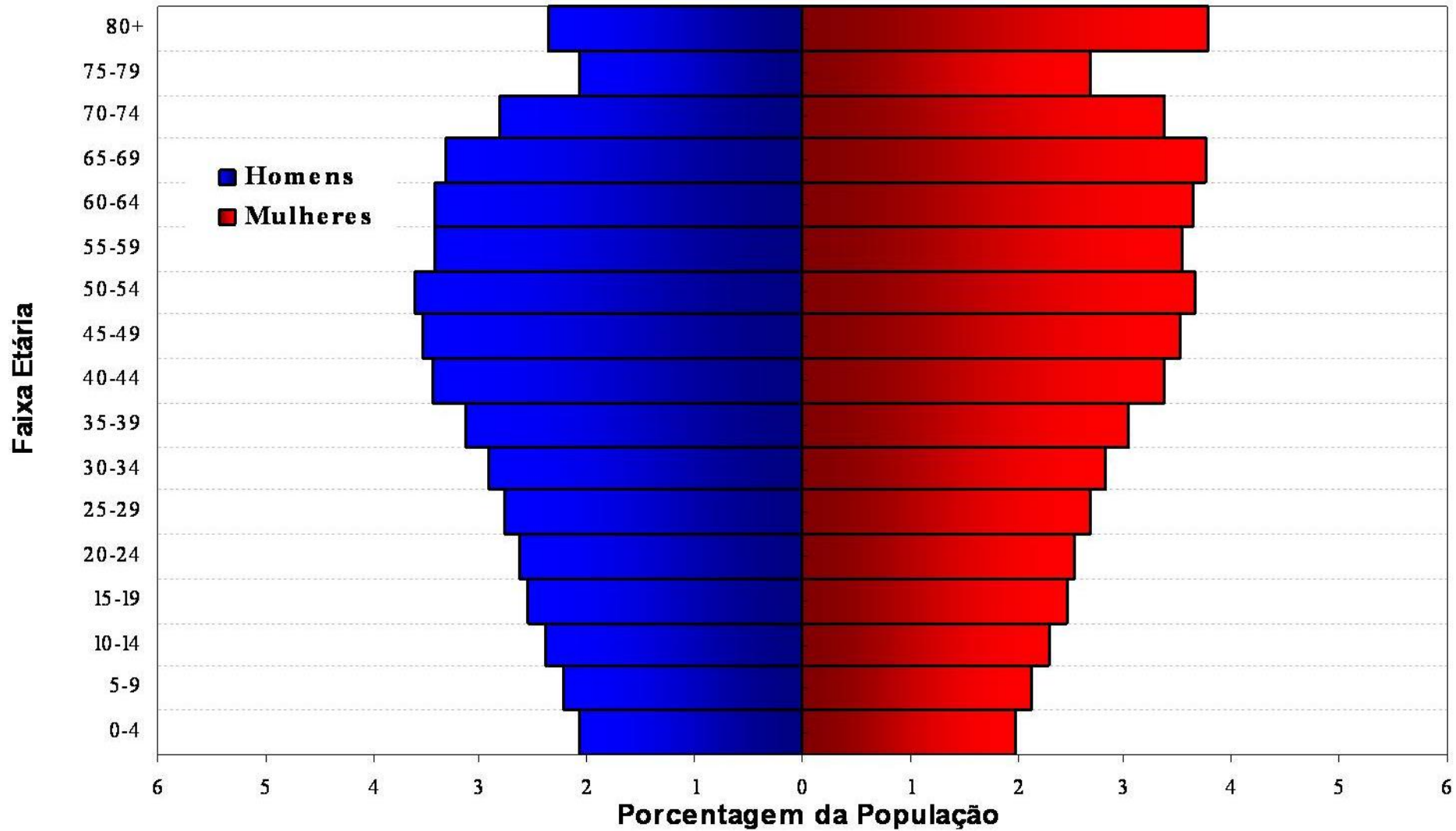
Fonte dos dados básicos: IBGE, PNAD 1996-2006 e Censo 1991-2000.

ESTRUTURA ETÁRIA POR SEXO DA RMBH 2030



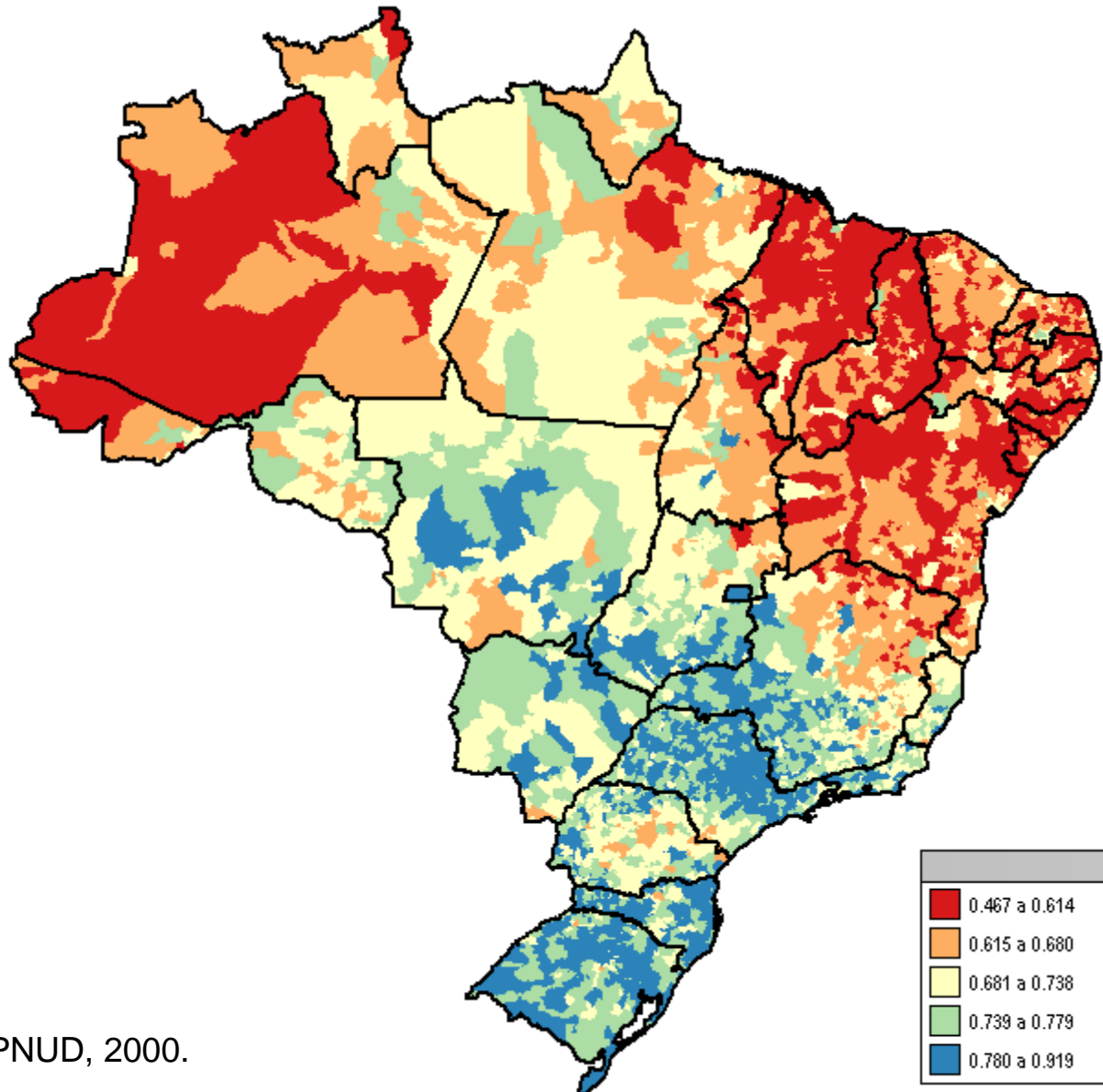
Fonte dos dados básicos: IBGE, PNAD 1996-2006 e Censo 1991-2000.

ESTRUTURA ETÁRIA POR SEXO DA RMBH 2050



Fonte dos dados básicos: IBGE, PNAD 1996-2006 e Censo 1991-2000.

ÍNDICE DE DESENVOLVIMENTO HUMANO MUNICIPAL BRASIL, 2000



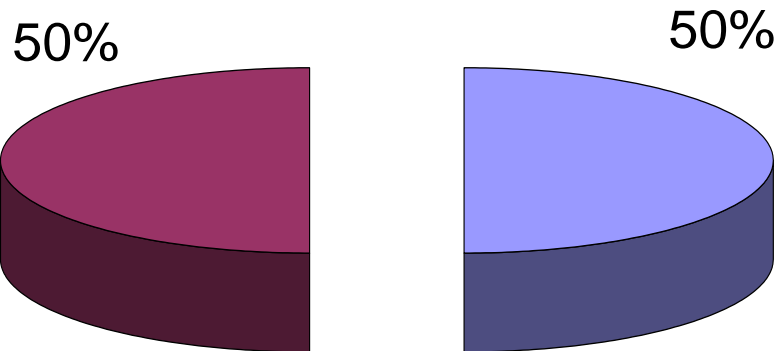
Legenda		
0.467 a 0.614	(1105)	
0.615 a 0.680	(1112)	
0.681 a 0.736	(1098)	
0.739 a 0.779	(1119)	
0.780 a 0.919	(1073)	

Fonte: Atlas do PNUD, 2000.

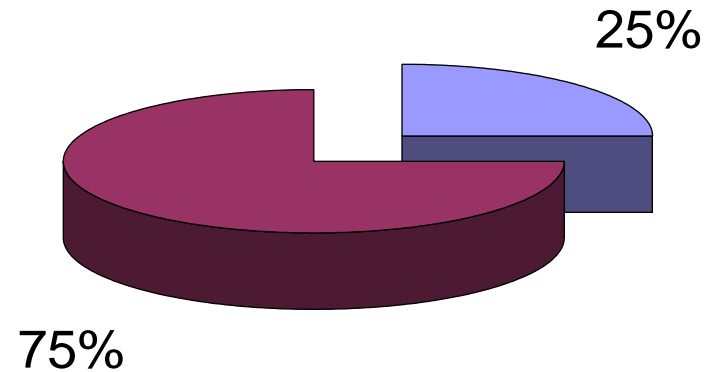
CONCENTRAÇÃO DOS HOMICÍDIOS EM 2000 NAS CIDADES BRASILEIRAS MAIS POVOADAS

– Menos de 1% dos municípios brasileiros concentrou 50% dos homicídios e 25% da população nacional em 2000.

Concentração dos homicídios



Concentração da população



 27 cidades mais violentas

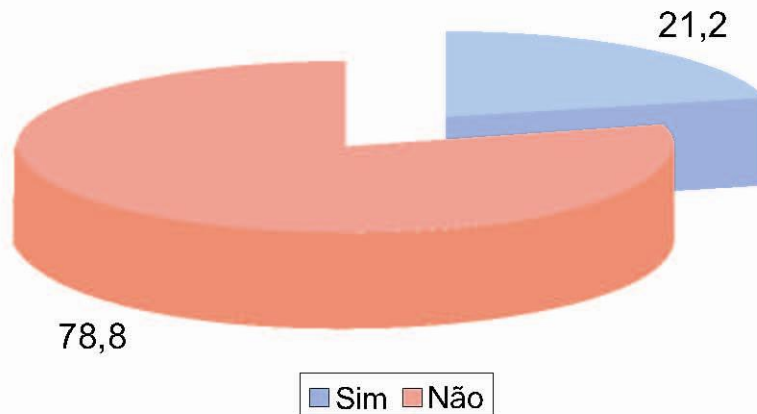
 5480 outras cidades

Fonte dos dados: DATASUS.

Fonte do slide: http://www.unodc.org/pdf/brazil/pp_1_diagn_introd_pt.pps.

PRESENÇA DE ARMAS EM BELO HORIZONTE, 2002

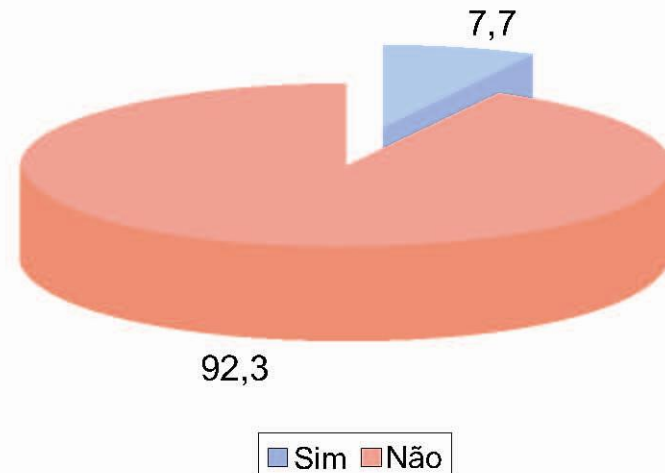
Constatação de Pessoas Armadas na Vizinhança



Um entre cada cinco pessoas já percebeu a presença de pessoas andando armadas na sua vizinhança

Cerca de 10% da população possui arma de fogo

Possui Arma de Fogo



Fonte: CRISP/UFMG - Belo Horizonte (2002)

Fonte dos dados: CRISP/UFMG.

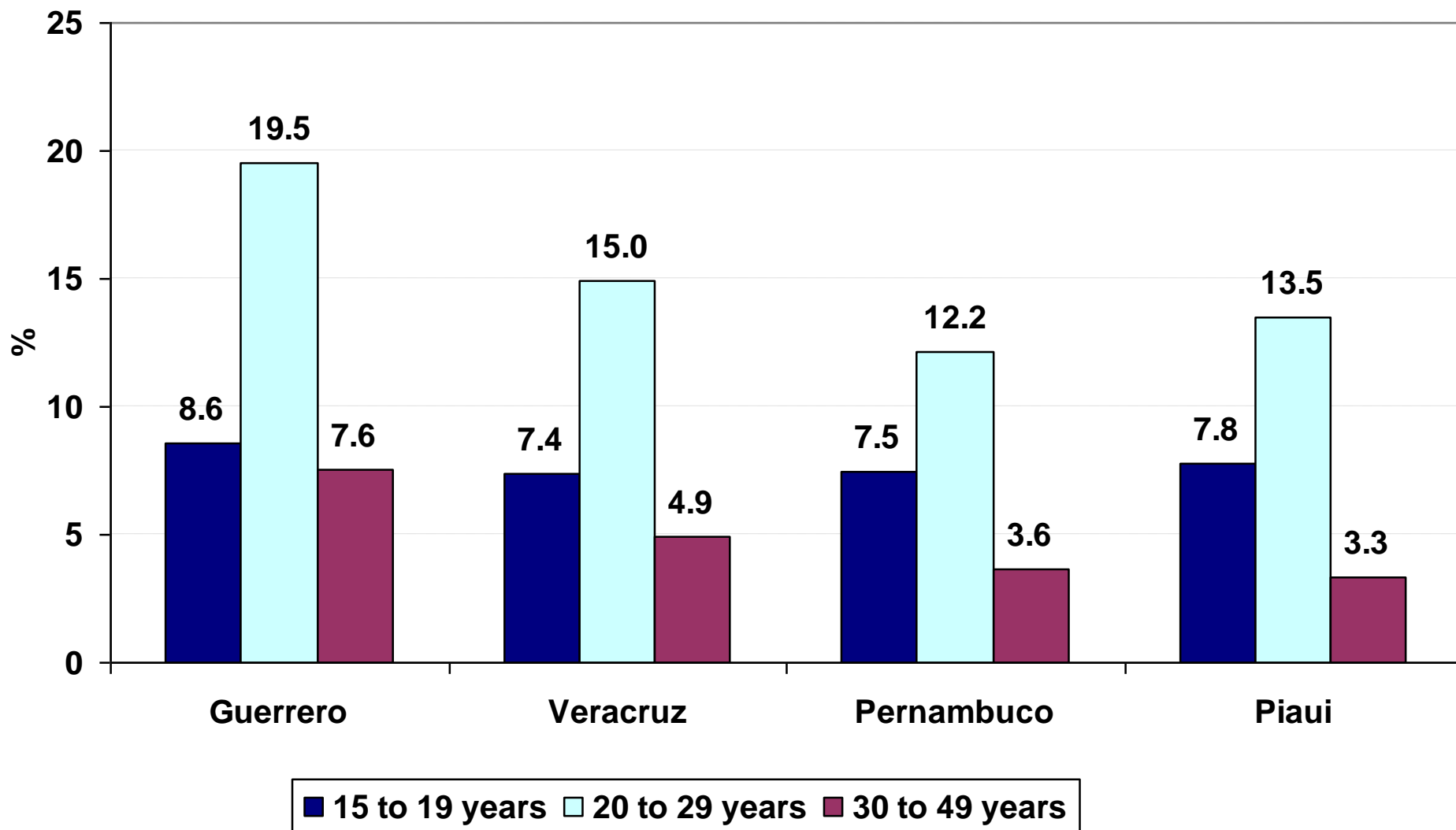
Fonte do slide:

<http://www.mj.gov.br/Senasp/data/Pages/MJCF2BAE97ITEMID472C81CEA9EF40649B47CE414AD28132PTBRIE.htm>.

EXPLICAÇÃO

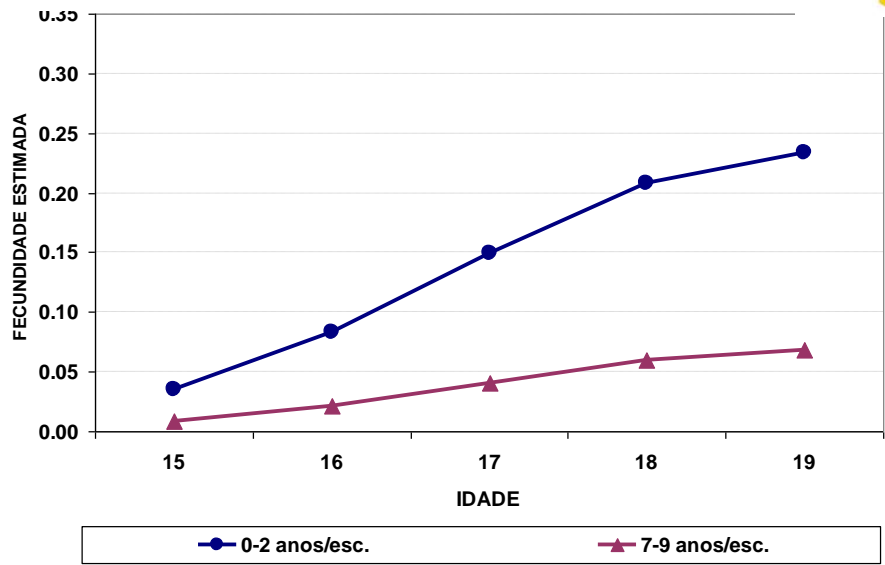
- Muitas pesquisas de *survey* têm o objetivo adicional de fazer análises explicativas sobre a população, procurando relações de causa e efeito.
- Explicar quase sempre requer análise multivariada, que é a análise do impacto simultâneo de duas ou mais variáveis (independentes) sobre outra variável (dependente).
- Há uma preocupação em explicar o porquê de determinada distribuição da variável, geralmente utilizando modelos estatísticos.
- Por exemplo, podemos querer explicar taxas de fecundidade por idade das mulheres, local de residência e características do domicílio...

MULHERES COM FILHO NASCIDO VIDO NO ÚLTIMO ANO, MÉXICO E BRASIL - 2000

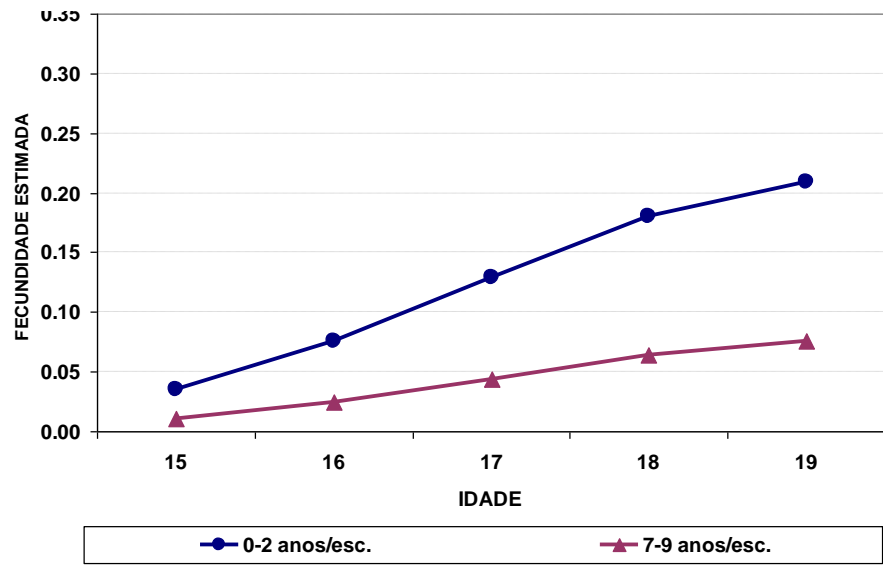




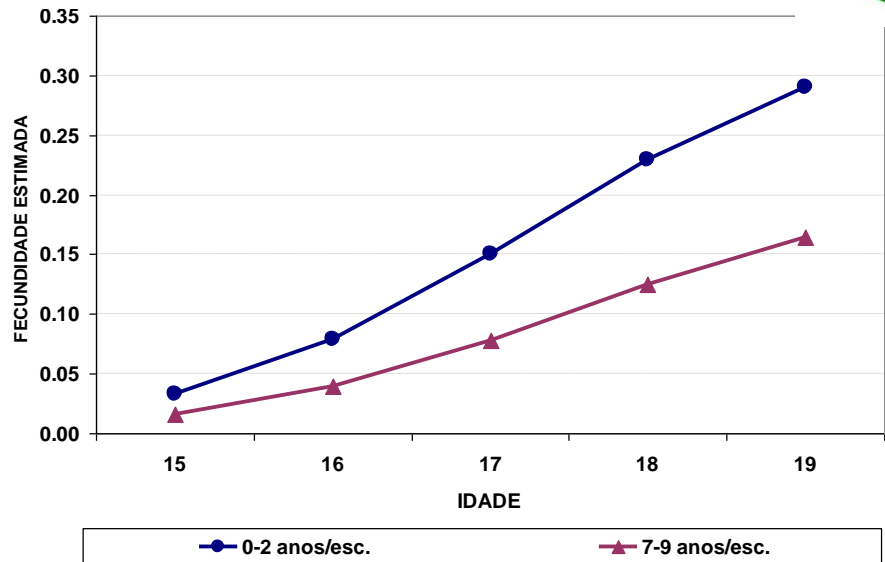
PIAUI - BRASIL



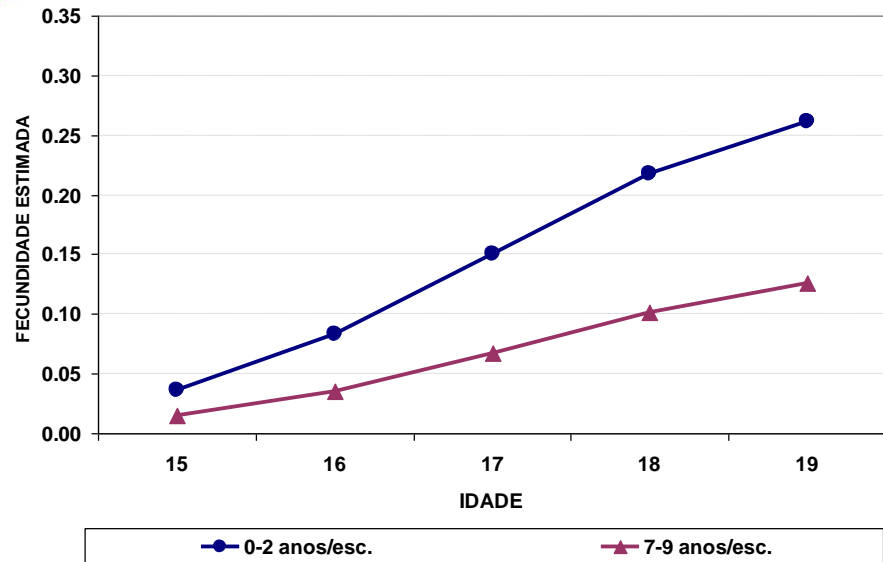
PERNAMBUCO - BRASIL



GUERRERO - MÉXICO



VERACRUZ - MÉXICO

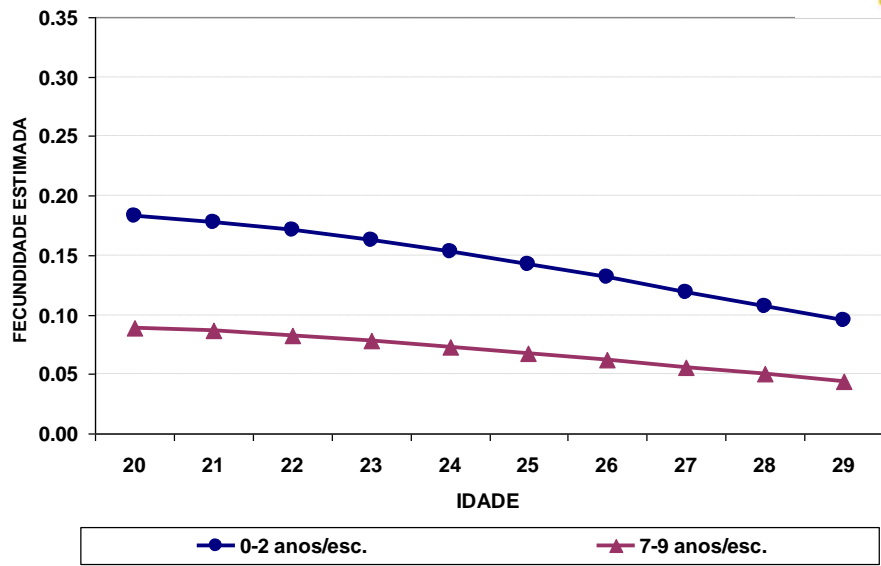


MULHERES DE 20-29 ANOS

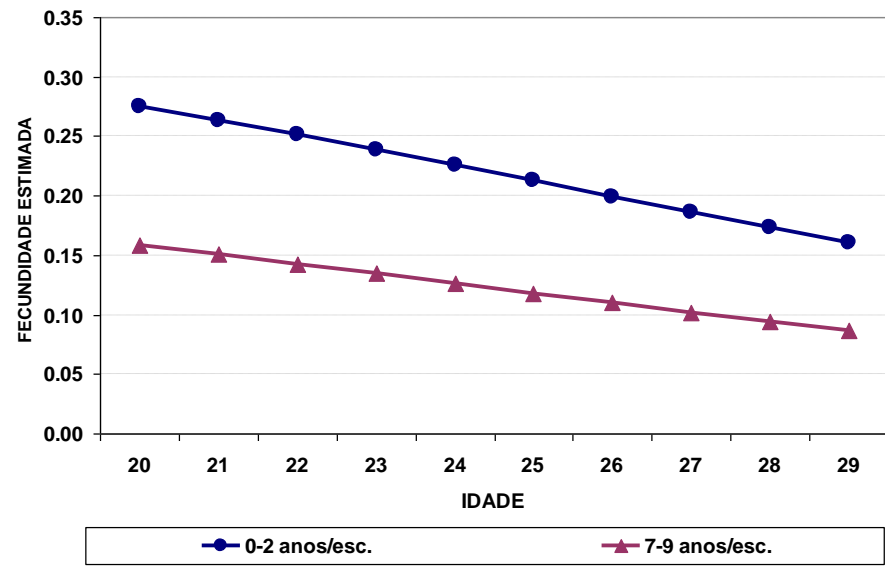


MULHERES COM 3 FILHOS OU MAIS⁷

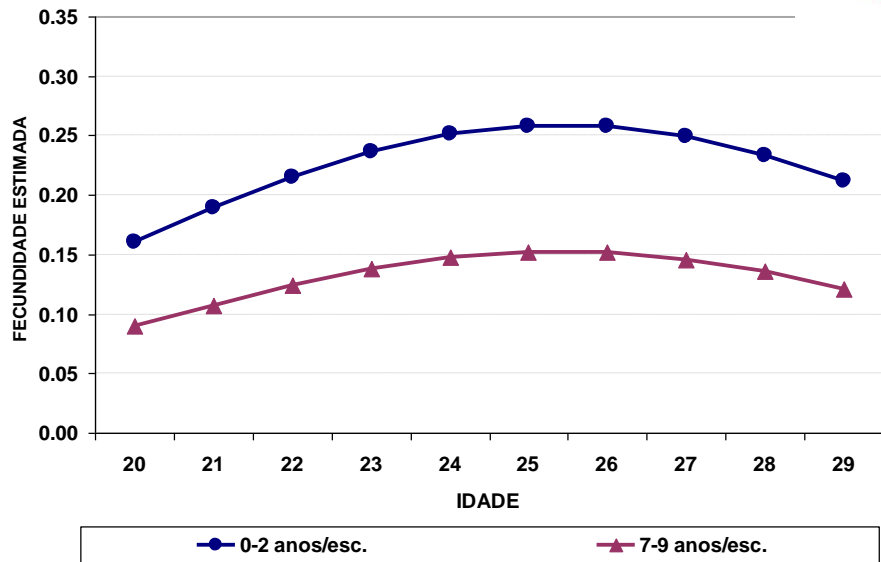
PIAUI - BRASIL



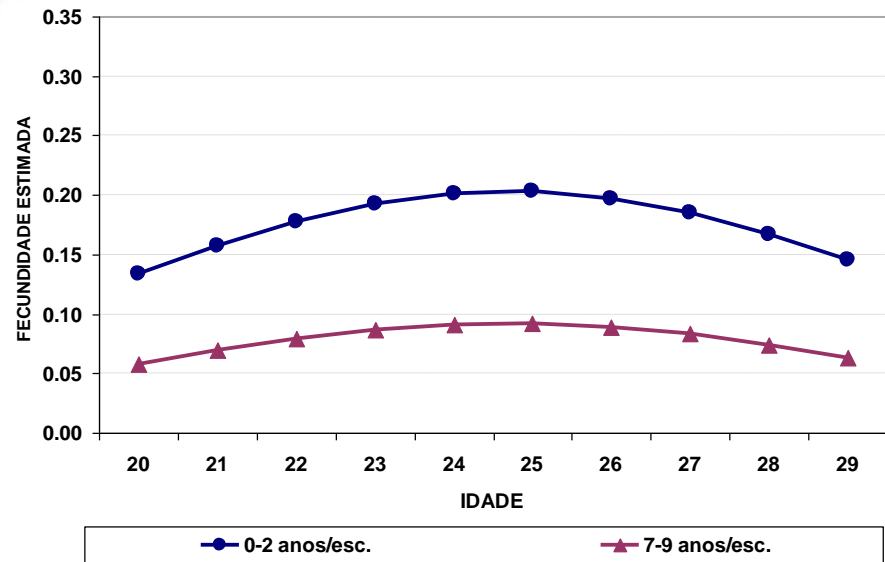
PERNAMBUCO - BRASIL



GUERRERO - MÉXICO



VERACRUZ - MÉXICO

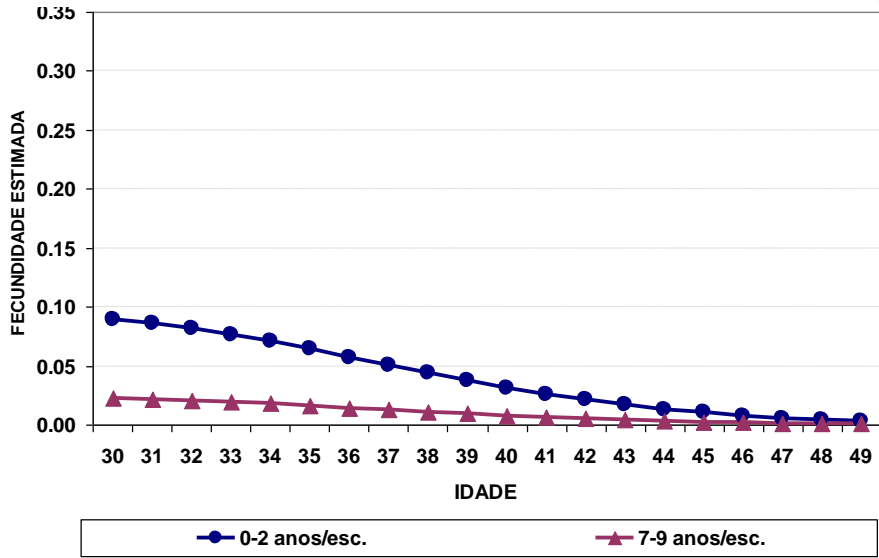


MULHERES DE 30-49 ANOS

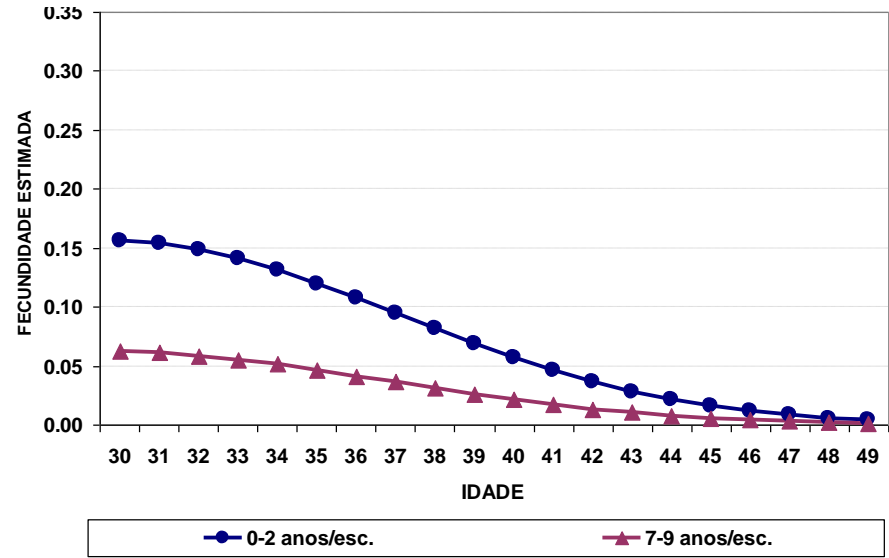


MULHERES COM 3 FILHOS OU MAIS⁸

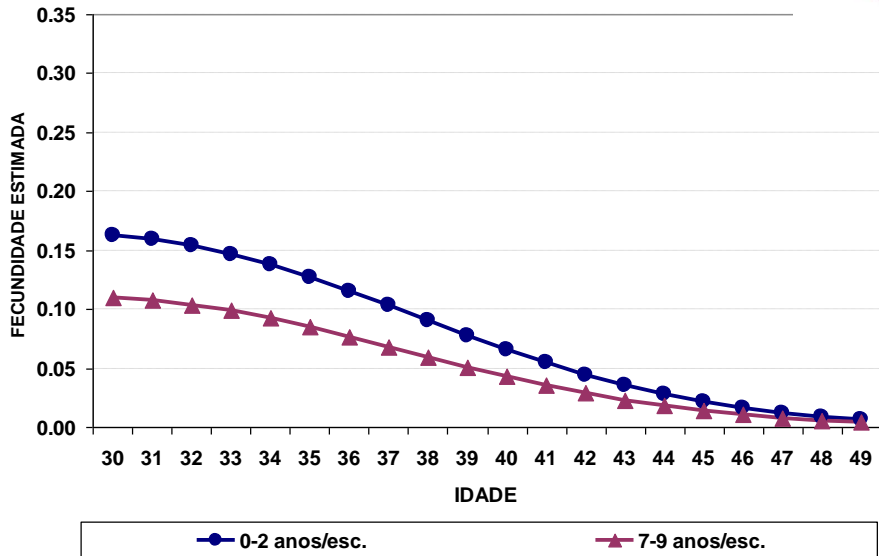
PIAUI - BRASIL



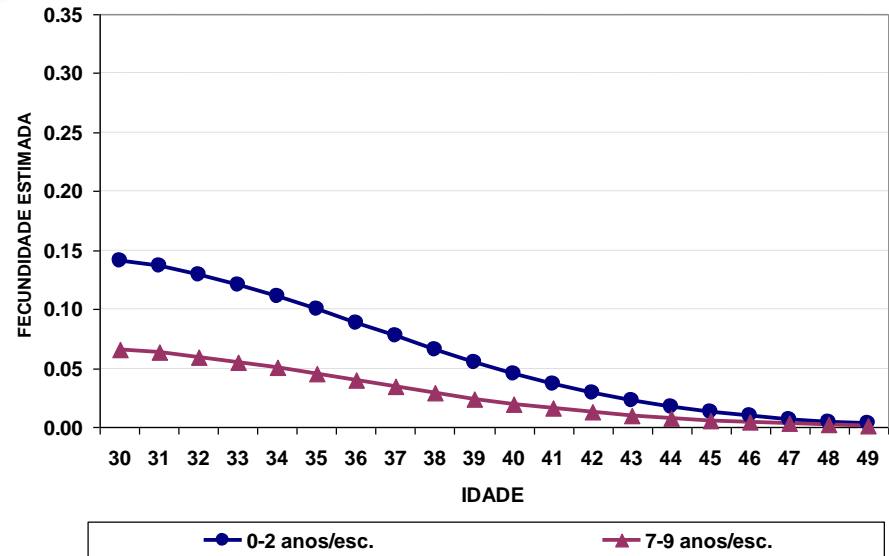
PERNAMBUCO - BRASIL



GUERRERO - MÉXICO



VERACRUZ - MÉXICO



EXPLORAÇÃO

- Pesquisadores podem fazer pesquisa sobre um tema que está começando a ser investigado, e não se sabe exatamente suas causas e consequências.
- Primeiramente, pode ser realizada entrevista em profundidade (qualitativa) para captar o máximo de informações possíveis do público em estudo:
 - Não é necessário ter amostra representativa.
 - Não é preciso coletar dados com questionário padronizado.
- Posteriormente, as informações são analisadas para melhor entender os indivíduos, sem preocupação explicativa, mas visando a elaboração de questionário padronizado.

UNIDADES DE ANÁLISE

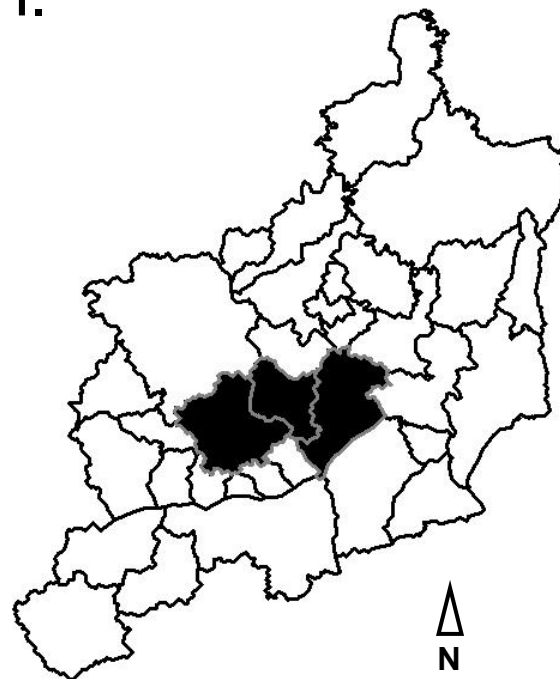
- Em uma pesquisa de *survey*, dados são coletados para descrever cada unidade que está sendo analisada.
- As unidades de análise são tipicamente pessoas, mas podem ser domicílios, bairros, municípios, Estados, países, empresas, universidades...
- Os dados são agregados e manipulados, permitindo descrever a amostra estudada, além de realizar análises explicativas.
- Um mesmo *survey* pode envolver mais de uma unidade de análise, como é o caso da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE), que coleta dados de pessoas e domicílios.

POPULAÇÃO DA RMBH EM 2009

Município da RMBH	População em 2009
310500 Baldim	8,580
310620 Belo Horizonte	2,452,612
310670 Betim	441,749
310900 Brumadinho	34,388
311000 Caeté	41,042
311250 Capim Branco	9,275
311787 Confins	6,073
311860 Contagem	625,390
312410 Esmeraldas	59,416
312600 Florestal	6,203
312980 Ibirité	157,441
313010 Igarapé	33,768
313220 Itaguara	12,955
313370 Itatiaiuçu	9,370
313460 Jaboticatubas	16,514
313665 Juatuba	20,979
313760 Lagoa Santa	48,211
314015 Mário Campos	12,024
314070 Mateus Leme	26,862
314110 Matozinhos	35,229
314480 Nova Lima	76,611
313660 Nova União	5,654
314930 Pedro Leopoldo	59,064
315390 Raposos	15,521
315460 Ribeirão das Neves	349,306
315480 Rio Acima	8,695
315530 Rio Manso	5,266
315670 Sabará	126,195
315780 Santa Luzia	231,610
316292 São Joaquim de Bicas	23,985
316295 São José da Lapa	19,233
316553 Sarzedo	25,579
316830 Taquaraçu de Minas	3,944
317120 Vespasiano	101,844
Total	5,110,588

– Dos 34 municípios da RMBH, 26 possuem menos de 100 mil habitantes, os quais correspondem a 12% da população da RMBH.

– Belo Horizonte, Contagem e Betim agregam 69% do total da população da RMBH.



FALÁCIA ECOLÓGICA

- Unidades de análise num *survey* podem ser descritas com base em seus componentes.
- Se o objeto da pesquisa for descrever os municípios e agregar as várias descrições para descrever todos municípios, a unidade básica de análise é o município.
- Em outro exemplo, pesquisador pode selecionar unidade de análise incorreta, incorrendo em falácia ecológica:
 - Ao analisar taxas de criminalidade por raça, o correto seria ter dados por indivíduos.
 - Se utilizarmos dados municipais, poderíamos concluir que negros possuem maiores taxas de criminalidade.
 - No entanto, os crimes podem estar sendo realizados por brancos que residem em municípios com maioria negra.

IDENTIFICAÇÃO DAS UNIDADES DE ANÁLISE

- É importante que as unidades de análise sejam identificadas no momento da montagem dos bancos de dados.

- O mais seguro é que em pesquisas que possuam mais de uma unidade de análise, sejam criados arquivos separados de dados para cada unidade.

DESENHOS BÁSICOS DE *SURVEY*

- Após especificar os objetivos e unidades de análise da pesquisa, é preciso escolher entre diversos desenhos diferentes:
 - *Surveys* interseccionais (*cross-sectional*).
 - *Surveys* longitudinais (tendências, coortes ou painel).
 - *Surveys* interseccionais servindo como longitudinais.
- Wooldridge (2008) classifica os dados econômicos em:
 - Dados de corte transversal = *surveys* interseccionais.
 - Cortes transversais agrupados = estudos de tendências.
 - Dados de séries de tempo = estudos de coortes.
 - Dados de painel ou longitudinais = estudos de painel.

DADOS DE CORTE TRANSVERSAL (Wooldridge)

SURVEYS INTERSECCIONAIS (Babbie)

- Um conjunto de dados de corte transversal consiste em uma amostra de uma unidade de análise, tomada em um determinado ponto no tempo.
- Esses dados são muito utilizados em economia e em outras ciências sociais.
- Dados em um determinado ponto do tempo são importantes para testar hipóteses e avaliar políticas.
- Dados podem ter problemas de seleção amostral, no caso de determinados indivíduos não revelarem informações acuradas.
- Amostragem deve ser realizada de forma acurada para evitar que coleta se concentre em unidades com características semelhantes.

EXEMPLO DE DADOS DE CORTE TRANSVERSAL

– Conjunto de dados de corte transversal para o ano de 1976 de 526 trabalhadores (Wooldridge 2008):

Número da observação	Salário por hora	Anos de escolaridade	Anos de experiência no mercado de trabalho	Feminino	Estado civil (casado)
1	3,10	11	2	1	0
2	3,24	12	22	1	1
3	3,00	11	2	0	0
4	6,00	8	44	0	1
5	5,30	12	7	0	1
...
525	11,56	16	5	0	1
526	3,50	14	5	1	0

CORTES TRANSVERSAIS AGRUPADOS (Wooldridge)

ESTUDOS DE TENDÊNCIAS (Babbie)

- Uma população pode ser amostrada e estudada em ocasiões diferentes.
- Um mesmo conjunto de variáveis é coletado em diferentes períodos do tempo, em **distintas** amostras aleatórias de uma mesma população (Censo Demográfico, Pesquisa Nacional por Amostra de Domicílios – PNAD).
- Agrupar cortes transversais de diferentes anos é eficaz para analisar os efeitos de uma política pública.
- O ideal é coletar dados de anos anteriores e posteriores a uma importante mudança de política governamental.
- Além de aumentar o tamanho da amostra, a análise de corte transversal agrupada é importante para estimar como uma relação fundamental mudou ao longo do tempo.
- Geralmente são utilizados dados secundários, coletados por outros pesquisadores ou instituições.

EXEMPLO DE CORTES TRANSVERSAIS AGRUPADOS

– Conjunto de dados sobre os preços da moradia em 1993 e 1995 nos Estados Unidos (Wooldridge 2008):

Número da observação	Ano	Preço comercializado	Impro	Arquad	Quantidade de dormitórios	Quantidade de banheiros
1	1993	85.500	42	1.600	3	2,0
2	1993	67.300	36	1.440	3	2,5
3	1993	134.000	38	2.000	4	2,5
...
250	1993	243.600	41	2.600	4	3,0
251	1995	65.000	16	1.250	2	1,0
252	1995	182.400	20	2.200	4	2,0
253	1995	97.500	15	1.540	3	2,0
...
520	1995	57.200	16	1.100	2	1,5

DADOS DE SÉRIES DE TEMPO (Wooldridge)

ESTUDOS DE COORTES (Babbie)

- Um conjunto de dados de séries de tempo consiste em observações sobre variáveis ao longo do tempo.
- Como eventos passados podem influenciar eventos futuros, o tempo é uma dimensão importante em um conjunto de dados de séries de tempo.
- A análise desses dados pode ser dificultada, porque observações econômicas não são independentes ao longo do tempo (variáveis possuem padrões sazonais).
- Há uma série de frequências possíveis: diárias, semanais, mensais, trimestrais, anuais, decenais...
- Estes dados são também chamados de estudos de coorte, em que mesma população é analisada, mas amostras estudadas podem ser diferentes:
 - Pessoas com 10 anos em 2000, 20 anos em 2010, 30 anos em 2020, 40 anos em 2030...

EXEMPLO DE DADOS DE SÉRIES DE TEMPO

– Conjunto de dados de séries de tempo sobre efeitos do salário mínimo em Porto Rico (apud Wooldridge 2008):

Número da observação	Ano	Salário mínimo médio no ano	Taxa de trabalhadores cobertos pela lei de salário mínimo	Taxa de desemprego	Produto Nacional Bruto (PNB)
1	1950	0,20	20,1	15,4	878,7
2	1951	0,21	20,7	16,0	925,0
3	1952	0,23	22,6	14,8	1.015,9
...
37	1986	3,35	58,1	18,9	4.281,6
38	1987	3,35	58,2	16,8	4.496,7

DADOS DE PAINEL OU LONGITUDINAIS (Wooldridge) ESTUDOS DE PAINEL (Babbie)

- Um conjunto de dados de painel consiste em uma série de tempo para **cada** membro do corte transversal.
- Os dados de painel são distintos dos dados de corte transversal agrupados (tendências) e de séries de tempo (coortes), porque as **mesmas** unidades são acompanhadas ao longo de um determinado período.
- Dados de painel podem ser coletados para indivíduos, domicílios, instituições ou unidades geográficas.
- Esses dados são os mais sofisticados para fins explicativos, mas são mais difíceis e caros de se obter.
- Pode haver problema de grande número de não respostas nas últimas ondas de entrevistas.
- A análise dos dados pode se tornar complicada quando se tentar avaliar as mudanças dos indivíduos no tempo.

EXEMPLO DE DADOS DE PAINEL OU LONGITUDINAIS

- Conjunto de dados de painel sobre crime e estatísticas relacionadas em 1986 e 1990 em 150 cidades nos Estados Unidos (Wooldridge 2008):

Número da observação	Cidade	Ano	Homicídios	População	Desemprego	Polícia
1	1	1986	5	350.000	8,7	440
2	1	1990	8	359.200	7,2	471
3	2	1986	2	64.300	5,4	75
4	2	1990	1	65.100	5,5	75
...
297	149	1986	10	260.700	9,6	286
298	149	1990	6	245.000	9,8	334
299	150	1986	25	543.000	4,3	520
300	150	1990	32	546.200	5,2	493

CORTE TRANSVERSAL USADO COMO LONGITUDINAL

- Alguns mecanismos podem ser utilizados num *survey* interseccional (corte transversal) para aproximar o estudo de processo ou mudança (longitudinal).
- Podem ser realizadas perguntas referentes ao passado (renda no ano anterior, local de residência anterior):
 - Há problemas de erro de memória.
 - Os dados devem ser interpretados como amostra da população atual, e não de população passada.
- Por exemplo, é possível utilizar um único banco de dados de corte transversal para comparar pessoas de diferentes idades (jovens e idosos) e coortes (calouros e veteranos).

VARIAÇÕES DOS DESENHOS BÁSICOS

- Os desenhos básicos de pesquisa apresentados anteriormente podem ser modificados para se enquadrarem aos objetivos de um estudo:
 - **Amostras paralelas:** amostras separadas de populações diferentes, utilizando mesmo questionário (exemplo é a pesquisa sobre preconceito na UFMG).
 - **Estudos contextuais:** uso de dados sobre o ambiente ou meio da pessoa para descrever o contexto do indivíduo.
 - **Estudos sociométricos:** intenção é de observar as inter-relações entre membros da população estudada (redes de amizades, por exemplo).

ESCOLHENDO O DESENHO APROPRIADO

- **Dados de corte transversal** são mais apropriados se objetivo é descrição de tempo único.
- **Mudanças ao longo do tempo** são mais difíceis de realizar, porque dados de painel exigem tempo e recursos:
 - É possível utilizar dados de corte transversal e comparar pessoas que passaram por uma experiência no passado, com aqueles que não passaram.
- **Estudos de painel** são mais viáveis economicamente quando o fenômeno estudado tem duração curta (por exemplo, opinião de voto durante uma campanha eleitoral).
- **Estudos de tendências** podem ser realizados quando dados antigos são complementados com dados coletados pelo pesquisador.