

Série Ensino, Aprendizagem e Tecnologias

Pesquisa e Análise de Dados:

problematizando o rural e
a agricultura numa perspectiva
científica

Daniela Dias Kühn



UNIVERSIDADE
FEDERAL DO RIO
GRANDE DO SUL

Reitor

Rui Vicente Oppermann

Vice-Reitora e Pró-Reitora
de Coordenação Acadêmica

Jane Fraga Tutikian

EDITORA DA UFRGS

Diretor

Alex Niche Teixeira

Conselho Editorial

Álvaro Roberto Crespo Merlo

Augusto Jaeger Jr.

Carlos Pérez Bergmann

José Vicente Tavares dos Santos

Marcelo Antonio Conterato

Marcia Ivana Lima e Silva

Maria Stephanou

Regina Zilberman

Tânia Denise Miskinis Salgado

Temístocles Cezar

Alex Niche Teixeira, presidente

Série Ensino, Aprendizagem e Tecnologias

Pesquisa e Análise de Dados:

**problematizando o rural e
a agricultura numa perspectiva
científica**

Daniela Dias Kühn

© dos autores
1ª edição: 2017

Direitos reservados desta edição:
Universidade Federal do Rio Grande do Sul

Coordenação da Série:
Laura Wunsch, Gabriela Trindade Perry, Tanara Forte Furtado e Marcello Ferreira

Capa: Ely Petry
Projeto gráfico: Editora da UFRGS
Editoração eletrônica: Tiago Dillenburg
Revisão: Equipe de Revisão da SEAD

Curso de Graduação Bacharelado em Desenvolvimento Rural (PLAGEDER)
Coordenação Pedagógica: Rumi Regina Kubo
Coordenação de Tutoria: Laura Wunsch
Coordenação Núcleo EAD: Tânia Rodrigues da Cruz
Secretário: Jorge Luis Aguiar Silveira

A grafia desta obra foi atualizada conforme o Acordo Ortográfico da Língua Portuguesa, de 1990, que entrou em vigor no Brasil em 1º de janeiro de 2009.



BACHARELADO EM
DESENVOLVIMENTO RURAL
PLAGEDER



CAPES



UNIVERSIDADE
ABERTA DO BRASIL



Associação Brasileira
das Editoras Universitárias



K457p Kühn, Daniela Dias

Pesquisa e análise de dados: problematizando o rural e a agricultura numa perspectiva científica [recurso eletrônico] / Daniela Dias Kühn ; coordenado pela SEAD/UFRGS. – dados eletrônicos. – Porto Alegre: Editora da UFRGS, 2017.

96 p. : il. ; pdf

(Série Ensino, Aprendizagem e Tecnologias)

Inclui figuras, quadros e tabelas.

Inclui referências.

1. Metodologia científica. 2. Metodologia da pesquisa. 3. Métodos de pesquisa. 4. Pesquisa científica – Análise de dados. 5. Trabalho científico – Organização. 6. Estudos rurais. I. Universidade Federal do Rio Grande do Sul. Secretaria de Educação a Distância. II. Título. III. Série.

CDU 001.89

CIP-Brasil. Dados Internacionais de Catalogação na Publicação.
(Jaqueline Trombin – Bibliotecária responsável CRB10/979)

ISBN 978-85-386-0377-1

SUMÁRIO

Apresentação 7

Unidade I

Problematizando a pesquisa11

- 1.1 Breves apontamentos sobre o histórico da formação do conhecimento
- 1.2 Os tipos de pesquisa
- 1.3 As etapas de uma pesquisa
- 1.4 O “problema” de pesquisa
- 1.5 Os tipos de dados

Unidade II

Buscando os dados: disponibilidade, seleção e limitação dos dados 23

- 2.1 O que são dados e que tipos de dados podem ser reconhecidos
- 2.2 Pesquisa exploratória: primeiro contato sistematizado com o tema de pesquisa
- 2.3 Busca e coleta de dados
- 2.4 Algumas bases de dados disponibilizadas no Brasil e no mundo
- 2.5 Alguns elementos importantes sobre dados e informações

Unidade III

Organizando, configurando e formatando a base de dados47

- 3.1 Configurando a base de dados
- 3.2 *Softwares* de trabalho com base de dados
- 3.3 Limitações e considerações sobre a configuração de base de dados

Unidade IV

Analisando os dados57

- 4.1 Principais medidas de estatística descritiva
- 4.2 Análise da matriz de correlação e a composição de indicadores e índices
- 4.3 Alguns elementos gerais

Apresentando os dados organizados	67
5.1 Elementos gráficos mais utilizados	
5.1.1 Quadros	
5.1.2 Tabelas	
5.1.3 Gráficos	
5.1.4 Organogramas e fluxogramas	
5.1.5 Mapas	
5.2 Outros elementos importantes na apresentação de trabalhos de pesquisa	
5.2.1 A escrita	
5.2.2 A apresentação	
Considerações finais.....	91
Referências.....	93

A realização de uma pesquisa envolve uma série de etapas e de processos que permitem que os estudantes e pesquisadores, a partir do reconhecimento de alguma situação interessante, sejam capazes de configurar um objeto e de analisá-lo, considerando também o contexto em que o interesse emerge. A identificação, seleção, organização e análise dos dados que ajudam a interpretação do problema de pesquisa constituem etapas importantes de um trabalho científico.

Nesse sentido, a origem deste trabalho se dá com a necessidade de sistematização de alguns temas e tópicos que têm como objetivo auxiliar o aluno na etapa de seleção e análise dos dados disponíveis. O volume que ora se apresenta pode ser considerado complementar ao material desenvolvido para a disciplina de Métodos de Pesquisa. Entretanto, atende ao foco específico e restrito em relação aos dados que deverão ser organizados pelos discentes ao realizar trabalhos acadêmicos e o próprio trabalho de conclusão de curso. Pode ser útil comentar com os leitores que, ao aceitar com imensa satisfação o convite para atuar no Bacharelado em Desenvolvimento Rural (PLAGEDER), deparei-me com alguma dificuldade de encontrar um material de referência que atendessem àquilo que estava demandado para essa disciplina de Tópicos Especiais.

A princípio, o material foi organizado em um caderno didático disponibilizado aos alunos da disciplina no ano de 2016. A oportunidade de publicação na série *Ensino, Aprendizagem e Tecnologias* exigiu revisão e atualização do material. Para tanto, foi fundamental a avaliação realizada ao final da disciplina pelos alunos, a quem agradeço imensamente a participação, as ideias e as sugestões enviadas (isentando-os, como de costume, em relação a qualquer inadequação persistente neste volume).

O objetivo desta publicação, em consonância com aquele estabelecido para a disciplina, é o de instrumentalizar o discente para a busca de construção e análise do problema de pesquisa. Além desta introdução e de um breve comentário à guisa de consideração final, o material foi organizado a partir da identificação de cinco temas centrais que correspondem, aqui, às cinco uni-

dades descritas. Em cada unidade, o leitor identificará a apresentação e construção de conceitos e categorias que atendem ao objetivo da unidade, bem como curiosidades, esclarecimentos e questões para a reflexão. Ao final, há um conjunto de referências que auxiliam no aprofundamento dos assuntos tratados ao longo do livro.

A primeira unidade é dedicada a uma retomada de conceitos importantes de temas do ambiente dos Métodos de Pesquisa. Muitos dos conceitos constantes aqui estão sendo retomados (considerando que, ao longo do curso, os alunos já se envolveram no debate geral sobre o tema). Entretanto, considera-se importante que, neste material, os leitores possam situar-se entre os tipos de pesquisa e dados existentes, bem como entre alguns conceitos que permitem, nas unidades seguintes, a identificação daqueles dados que são mais adequados ao tipo de pesquisa que deverá ser (ou que está sendo) realizada.

Elementos que indicam como a busca pelos dados pode ser feita, bem como a disponibilidade e a limitação de algumas bases de dados, são apresentados na segunda unidade de discussão.

A terceira unidade apresenta alguns elementos referentes à configuração e à formatação dos dados disponíveis. É importante reconhecer que, além da busca de dados, a organização dos mesmos facilita e viabiliza a análise do pesquisador. Essa é uma etapa fundamental para a transformação de uma base de dados em informação e conhecimento.

Apresentadas algumas possibilidades de organização da base de dados formada, a quarta unidade apresenta instrumentos que permitem a análise dos dados. São identificados aqui alguns *softwares* capazes de estabelecer relações interessantes entre variáveis. Nessa unidade, são retomados conceitos estatísticos que auxiliam na análise das relações evidenciadas a partir da base de dados organizada.

Por fim, a quinta unidade oferece sugestões para uma boa apresentação dos dados organizados e analisados, tanto de forma escrita como de forma oral.

Certamente, um volume como este não é capaz de esgotar as possibilidades de tratamento de seu tema. Apesar da preocupação na manutenção da escrita no âmbito conceitual, o material pode não ser capaz de manter-se atualizado por muito tempo (como tudo aquilo que utiliza a tecnologia como base). Entretanto, espera-se que seja possível chamar a atenção para a importância e a necessidade de organização que a busca e o tratamento dos dados, que com-

porão todo e qualquer trabalho científico, exige. Nesse sentido, espero que o material possa, além de representar um elemento de estudo para a disciplina, servir como apoio para a realização de trabalhos e da monografia de cada um dos leitores.

Toda e qualquer sugestão e crítica é bem-vinda por meio do endereço eletrônico daniela.kuhn@ufrgs.br.

Um grande abraço e boa leitura a todas e todos!

PROBLEMATIZANDO A PESQUISA

A análise de dados e informações para a organização de um trabalho científico não é um processo aleatório. Antes de qualquer organização de uma base de dados, é preciso reconhecer claramente o problema de pesquisa, ou seja, a pergunta que a análise dos dados identificados e organizados deve ser capaz de responder. Nesse sentido, a unidade que inaugura este volume busca retomar alguns aspectos importantes para a construção daquilo que é reconhecido como a realização de uma pesquisa científica. Assim, o objetivo desta unidade é, em primeiro lugar, o de retomar conceitos envolvidos no estudo de Método de Pesquisa, evidenciando a importância da formulação adequada de um problema de pesquisa para a boa condução da etapa de reunião dos dados; em segundo lugar, procura-se, ao longo do texto, apresentar os tipos de pesquisa e de dados que podem ser utilizados para a realização dos trabalhos.

Para que os objetivos propostos sejam atingidos, a unidade está dividida em 05 (cinco) subseções: a primeira retoma brevemente o histórico da construção científica para a formação do conhecimento na nossa sociedade; a segunda apresenta a classificação dos tipos de pesquisa que podem ser realizadas; a terceira examina as etapas de uma pesquisa a partir dos elementos considerados fundamentais ao método científico; a quarta oferece um breve debate sobre a importância da definição do problema de pesquisa; e a quinta detalha a classificação dos tipos de dados que estão disponíveis aos cientistas para a realização da sua pesquisa.

1.1 BREVES APONTAMENTOS SOBRE O HISTÓRICO DA FORMAÇÃO DO CONHECIMENTO

Há muito tempo o ser humano preocupa-se em reconhecer e explicar eventos e acontecimentos das mais diversas origens. A natureza do questionamento levantado pelas pessoas leva a um conjunto de informações que envolvem a noção de conhecimento. Nesse sentido, pode-se dizer que a evolução da sociedade deriva diretamente desses questionamentos e da reunião de um volume crescente de publicações que apresentem as conclusões aos pesquisadores e interessados. Na sistematização e no reconhecimento dessas informações está o elemento que caracteriza a consolidação da Ciência como o espaço formal para a compreensão das mais diversas relações.

É preciso destacar, entre as características da Ciência, que ela é uma atividade humana (ou seja, desenvolve-se a partir da percepção e da racionalização do ser humano em relação aos problemas/curiosidades que o rodeiam). A Ciência tem um caráter provisório do conhecimento gerado, ainda que cumulativo. Isso significa que o que pode estar certo na Ciência atualmente poderia não ser considerado correto em períodos anteriores e pode ser superado em um futuro mais ou menos distante. Dessa forma se configura e se reconhece uma dimensão que permite a reflexão sobre um campo que pode ser reconhecido como a História da Ciência.

A ideia de Ciência está associada ao conjunto de conhecimentos obtidos de forma sistemática, de uma forma organizada e reconhecida entre os pares de determinada área. Esse conjunto está associado a um tipo de prática e ao ambiente em que o “cientista” (quem faz Ciência) está inserido. A palavra Ciência deriva do latim “*scientia*” que significa “conhecimento”, “sabedoria”.

A partir da insuficiência do caráter místico de algumas explicações referentes ao cotidiano da vida humana, a Ciência organiza-se no sentido de gerar um conhecimento, uma verdade. Esse movimento histórico inicia-se e consolida-se a partir da Modernidade. A passagem da sociedade feudal para a sociedade moderna (da Idade das Trevas ao Capitalismo) envolve a consolidação da Ciência como a organização do conhecimento legítimo da sociedade.

O conhecimento decorre do estabelecimento entre dois elementos: aquele que se conhece, se percebe e aquele que está sendo conhecido. Assim, sendo o conhecimento um processo complexo, é preciso que exista alguma pergunta, alguma questão a ser resolvida para que se configure o processo. O conheci-

mento se constrói, se constitui e evolui a partir da relação entre sujeito e objeto. Dessa forma, é preciso sempre reconhecer a influência dessa relação no “fazer” científico (em especial nas Ciências Sociais).

A percepção da realidade pode ocorrer de formas diferentes. O conhecimento sensível envolve uma relação com o objeto; essa relação se dá a partir dos sentidos externos ou internos. Os sentidos externos são aqueles que reconhecem a relação de contato entre observador e objeto (visão, audição, tato, olfato e paladar). Por exemplo, a partir desse conjunto de conhecimento sensível externo, decidimos se devemos ou não colocar um casaco para sair de casa pela manhã. O conhecimento sensível interno envolve a construção de memórias, da imaginação e de fantasias. Um exemplo desse sentido é o reconhecimento da criança que atende pelo nome ao ser chamada. O conhecimento sensível está presente nos seres vivos e garante sua interação com o meio em que vivem.

O conhecimento intelectual, característico dos seres humanos, permite a formulação de modelos, a formatação de ideias, envolvendo uma série de julgamentos intertemporais e a formulação de leis, por exemplo. A nossa sociedade caracteriza-se (distinguindo-se, portanto, dos demais seres vivos) pela possibilidade de organização e acumulação de conhecimento intelectual.

A percepção do ser humano sobre o conjunto de saberes que ele possui pode, então, ser nomeada de conhecimento. A sistematização e o registro desses saberes deram-se a partir do desenvolvimento da escrita. Ainda que se possa reconhecer a importância da tradição oral para a difusão do conhecimento em ambientes sociais, a sociedade compõe e sistematiza seu conhecimento, atualmente, a partir do registro escrito.

Em relação aos níveis de conhecimento, eles podem ser: empírico, filosófico, teológico/religioso, e científico. O conhecimento empírico é o conhecimento do cotidiano (aquele que nos permite atravessar a rua sem sermos atropelados). Algumas vezes é chamado de conhecimento do “povo”, uma vez que deriva da experiência do dia a dia de cada pessoa, independentemente da sua exposição ao sistema de educação formal. Não se reconhece nesse tipo de conhecimento nenhum método ou conjunto de relações sistematizadas, ou seja, pode ou não ser reproduzível (em condições parecidas ou adversas). Ainda que essas características levem a uma desorganização na sua sistematização, é preciso atentar que esse tipo de conhecimento é o primeiro que todos nós utilizamos, de modo que ele serve, então, de base para todo o saber construído posteriormente.

O conhecimento filosófico envolve a discussão e o debate de ideias e objetos de estudo não materializados. É um tipo bastante antigo de conhecimento. Ele busca relações conceituais e não é reduzível a realidades materiais. Por meio desse tipo de conhecimento são organizados conteúdos reflexivos e lógicos que envolvem a análise da transformação da realidade. É um tipo de conhecimento crítico. Questões como “Qual o sentido da vida?”, “Há liberdade?”, “Há verdade?” envolvem o conhecimento filosófico.

Outro tipo de conhecimento é o conhecimento teológico (também reconhecido como religioso). Esse conjunto de saberes humano envolve a percepção em relação à existência de Deus (*theos*). A partir do conhecimento teológico, são estudados elementos que estão relacionados a Deus e à fé. Tem como pressuposto/base textos sagrados que apresentam a argumentação lógica para a compreensão dos eventos (por exemplo, a ocorrência de milagres e pragas/castigos).

O tipo de conhecimento que será apresentado e discutido neste curso é o conhecimento científico. Esse conhecimento é a base de toda a produção científica da chamada academia (o ambiente das universidades e das instituições de pesquisa). Ele deve ser sistematizado, ou seja, organizado de forma a que cada um dos seus integrantes seja capaz de reconhecer suas etapas e, portanto, replicar a discussão. Esse é o tipo de conhecimento que envolve a percepção de racionalidade, da busca por uma verdade. Ele é auxiliado por um método: o método científico, que, a partir da Idade Moderna, torna-se a forma de “fazer” Ciência. Diversos autores, entre os quais Bachelard, Popper, Kuhn e Edgar Morin, estudaram as bases do processo da geração do conhecimento científico.

O filósofo Gaston Bachelard (1884-1962) apresenta sua análise sobre a constituição da Ciência a partir do contexto da Teoria da Relatividade de Albert Einstein. Esse novo ambiente de análise indica a relativização do objeto, fugindo dos moldes empiricistas da percepção moderna. A partir dessa perspectiva, cada objeto é constituído pelo seu observador e já não pode mais existir como absoluto, apenas em termos relativos. A ordem de construção da Ciência é do racional para o real, e não o contrário.

Popper (1902-1994) conceitua a teoria científica como aquela capaz de descrever uma série de fenômenos a partir da análise de postulados simples, sendo capaz de realizar também previsões. Segundo Popper, o conhecimento científico está constantemente sendo “falseado”, ou seja, deve estar sempre

sendo provado e testado pelo grupo de especialistas. Enquanto o conhecimento não for tornado falso, ele é científico.

Kuhn (1922-1996) apresenta a evolução da Ciência por meio de rupturas, reconhecidas como paradigmas. Segundo essa percepção, o paradigma envolve a forma de ver o mundo. Historicamente, quando as perguntas científicas passam a não mais ser respondidas pela explicação “normal”, consagrada no ambiente acadêmico, é preciso o desenvolvimento de uma nova explicação (KUHNS, 1998, p. 44). Surge, então, o novo paradigma científico. Durante a vigência de um paradigma, a evolução da Ciência seria incremental, complementando e explicando, a cada nova descoberta, uma etapa do próprio paradigma.

Todo o debate em torno da historicidade do fazer científico e do conceito de Ciência envolve Epistemologia. A Epistemologia, que é o estudo da própria produção de Ciência, leva-nos a uma percepção relativa, exigindo a ponderação em torno da relação observador e objeto. Trata-se da problematização da verdade. Nesse sentido, conforme nos alerta Fonseca (2002, p. 12):

O objetivo básico da ciência não é o de descobrir verdades ou de se constituir como uma compreensão plena da realidade. Deseja fornecer um conhecimento provisório, que facilite a interação com o mundo, possibilitando previsões confiáveis sobre acontecimento futuros e indicar mecanismos de controle que possibilitem uma intervenção sobre eles.¹ (FONSECA, 2002, p. 12)

A evolução da Ciência levou a uma percepção das fortes limitações da divisão em áreas do conhecimento. Ainda que essa divisão tenha sido essencial ao progresso científico, ela apresentou seu limite no século XX. No debate científico, surgem e se consolidam as noções de multi e transdisciplinaridade; ao mesmo tempo, começa-se a debater uma teoria da complexidade. Assim, fica para trás a figura do quebra-cabeça, do relógio para marcar o tempo, a fim de emergir o debate em torno do holograma e da relatividade. Essa é uma discussão de ponta da teoria científica: o desenvolvimento de *softwares* permite uma análise muito mais complexa do que aquela desenvolvida há algum tempo atrás. A realidade tecnológica revela a necessidade de uma maior complexidade

¹ Decorre desta percepção um debate sobre a adequação e aderência da forma de fazer ciência a uma distribuição de poder político e econômico, o que não é tema deste volume. Um dos autores pós-modernos que nos levam a esse debate é Zygmunt Bauman. (BAUMAN, Zygmunt. *Modernidade e Ambivalência*. Rio de Janeiro: Zahar, 1999).

para a percepção e análise do ambiente social com o objetivo de construção de hipóteses, verificação das evidências e construção de certezas.

O conceito de multidisciplinaridade envolve a reunião de diversos olhares sobre o mesmo objeto. Portanto, envolve a reunião de vários campos do conhecimento na constituição de um único objeto. É um trabalho conjunto de diversas disciplinas.

A transdisciplinaridade, por outro lado, é uma abordagem científica que busca a reunião e a unidade do conhecimento para a construção do próprio objeto de pesquisa. Busca ultrapassar o limite das disciplinas, gerando um novo tipo de conhecimento e não simplesmente reunindo conhecimentos distintos sobre um objeto. Trata-se de um exercício de reconstrução da realidade de forma mais complexa. A ideia é a de que a soma pode ser maior ou menor do que as partes que a compõem. Há um debate extenso em relação a como esse conhecimento pode ser formado; entretanto, percebe-se um aumento das tentativas de reconstrução de objetos de estudo a partir do olhar transdisciplinar.

1.2 OS TIPOS DE PESQUISA

A busca pelo conhecimento se dá a partir de uma série de curiosidades e perguntas que, sistematizadas, configuram o procedimento chamado de pesquisa. A pesquisa pode ser então definida como “um procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos” (GIL, 2002, p. 17).

Em linhas gerais, conforme Gil (2002, p. 41) e Gerhardt (2009, p. 31), podemos classificar a pesquisa considerando origem e formas de atuação e utilização de recursos e informações. O Quadro 1 identifica algumas características dos principais tipos.

Quadro 1 – Tipos de pesquisa

Tipo	Pesquisa
Exploratória	Proporciona familiaridade com o problema. Aprimora as ideias sobre o tema. Flexível. Atividades: Levantamento Bibliográfico, Entrevistas exploratórias, Reconhecimento de exemplos.
Descritiva	Descrição dos fenômenos. Padronização na coleta de dados (parâmetros claros de comparação). Percepção de frequência de fenômenos conjuntos.
Explicativa	Identificar fatores que influenciam a ocorrência de determinados fenômenos. Procura explicar o porquê dos elementos estudados.
Qualitativa	Aprofundamento da compreensão de determinado comportamento, sem elementos de representatividade numérica. Não há um modelo único.
Quantitativa	Procura a possibilidade de generalização, constituindo amostras representativas. Dados estruturados a partir de um instrumento de pesquisa formal e comparáveis.
Bibliográfica	Baseia-se no levantamento e na análise de material já elaborado por diversos autores em diferentes meios de registro científico. Realizada a partir de fontes secundárias.
Documental	Evidencia características de material escrito ainda não analisado. Fontes dispersas e diversificadas.
De campo	Realiza a coleta de dados primários aprofundados referentes à população em estudo. Utilizada para estudo de grupo social específico.
De levantamento	Identifica e caracteriza informações com grau de comparabilidade para populações em geral e amostras com comportamento que se pretende conhecer.
<i>Ex-post-facto</i>	Investiga a relação de causa e efeito entre os fenômenos estudados, após a ocorrência do fenômeno.
Experimental	Evidencia a relação entre os fenômenos partindo da organização e do controle possível das variáveis identificadas.
De coorte	Acompanha elementos com características em comum que interessam à pesquisa.
Estudo de caso	Apresentação e análise profunda e exaustiva de um (ou poucos) objetos de estudo. Não há base para generalizações.
<i>Survey</i>	Procura obter dados sobre elementos e comportamentos específicos, considerando um grupo de interesse delimitado metodologicamente (população-alvo).
Participante	Ocorre o reconhecimento do envolvimento entre pesquisador e objeto de estudo.
Pesquisa-Ação	Ocorre a partir da participação planejada do pesquisador que transforma a relação do investigado com o objeto.

Básica	Busca desenvolver novas teorias explicativas que colaboram para o avanço da Ciência de uma maneira geral.
Aplicada	Tem o objetivo de desenvolver conhecimentos referentes a problemas específicos, com uma solução prática e aplicável no cotidiano.
Etnográfica	Estudo detalhado de um grupo ou povo, a partir de características específicas.
Etnometodológica	Evidencia como objeto de estudo o comportamento dos sujeitos na vida cotidiana. Estudos sobre a conduta humana.

Fonte: Adaptado pela autora a partir de Gil (2002) e Gerhardt (2009).

Ainda que se possa tentar estabelecer essas diferenças entre os tipos de pesquisas realizáveis, convém chamar a atenção para o fato de que vários projetos de pesquisa os reúnem como etapas da formulação de um documento final. Não é raro que a pesquisa exploratória, por exemplo, antecipe uma pesquisa explicativa sobre o mesmo tema.

A organização de uma pesquisa tem diversas etapas. Quanto mais for possível pensar e sistematizar cada uma das etapas identificadas, melhor será o resultado e a possibilidade de análise.

1.3 AS ETAPAS DE UMA PESQUISA

A realização de uma pesquisa exige uma organização quanto ao caminho a ser seguido para que se alcance o resultado esperado. De origem grega, a palavra “método” significa, literalmente, o “caminho para chegar a um fim”.² Podemos ter caminho, regras e condutas a cada momento da nossa vida. O ato de atravessar uma rua exige de nós um método complexo para perceber o momento em que a ação deve ocorrer.

Assim como no cotidiano, a Ciência trabalha a partir de um conjunto de regras, normas e procedimentos que se consolidaram, historicamente, como o caminho reconhecido para a confirmação de um resultado científico.

Não há um modelo específico a ser seguido: para cada problema de pesquisa é possível identificar uma estrutura capaz de atender aos objetivos pro-

² Aqui se pode estabelecer a diferença entre método e metodologia. O método é o caminho escolhido pelo pesquisador. Metodologia é o estudo da validação do caminho, envolvendo um ambiente mais amplo de análise que reúne, por exemplo, a possibilidade de avaliação da adequação científica entre o arcabouço teórico e o método desenvolvido. Segundo Minayo (1994, p. 16), podemos considerar a metodologia como o “caminho do pensamento”. A metodologia inclui o método, mas não se resume a ele.

postos. Entretanto, Gil (2002, p. 93) cita algumas das etapas mais comuns: formulação do problema, construção das hipóteses, desenvolvimento do instrumento de coleta de dados, coleta de dados, organização dos dados, redação da análise dos resultados finais.

A partir da organização do estudo por esse “caminho”, os resultados são reconhecidos e validados entre os pesquisadores. Ao final deste volume, são indicadas obras que analisam mais detalhadamente as características do método científico para a realização de pesquisa. Por ora, interessa-nos reconhecer que é preciso perceber uma realidade, identificar o caminho para o estabelecimento das relações observadas e encontrar respostas parciais de acordo com os dados selecionados. Portanto, devemos ter claro o objetivo da pesquisa quando reunimos o conjunto de informações de que trataremos ao longo de um volume. Neste sentido, já nos alertava o filósofo Sêneca: “Para o barco que não sabe a que porto quer chegar, nenhum vento poderá ser favorável”.³

1.4 O “PROBLEMA” DE PESQUISA

Segundo Gil (2002, p. 23), a pesquisa só se inicia com uma dúvida, uma pergunta. Esse é o chamado “problema” de pesquisa. Essa pergunta está relacionada a um objeto de discussão que leva em consideração causas e/ou consequências em relação ao fenômeno que se busca compreender.

Pensar a formulação de um problema de pesquisa envolve a possibilidade de estabelecimento da relação entre variáveis. Sendo assim, a pergunta precisa ser composta por elementos que deverão estar (ou não) relacionados. Por exemplo, é possível identificar um processo de melhoria no Índice de Desenvolvimento de determinado município a partir do acesso dele a determinada política nos anos 1990? Ou, pode-se considerar a agricultura familiar como a estrutura produtiva que mais cresce em determinada região? Nesses e em muitos outros casos, é interessante perceber que há possibilidade de estabelecer relações e medições entre variáveis para a verificação da resposta.

O mais comum é que o problema seja formulado em forma de pergunta. Mas esse é um exercício que exige dedicação dos pesquisadores, geralmente é

³ “O nauta, incerto da vida, confia aos ventos as velas, quando uma aura enfuna o frouxo tecido.” (MARCHIORI, Luciano Antonio B. Spinelli. *Hércules furioso de Sêneca: estudo introdutório, tradução e notas*. 2008. 165 f. Dissertação (Mestrado em Letras) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008. p. 60.

necessária a realização de um conjunto de leituras sobre o tema e a conversa com pesquisadores mais experientes, como o orientador, por exemplo. Em linhas gerais, pode-se comentar que a formulação de um problema de pesquisa pode ter como objetivo a avaliação de determinada situação, identificação de causas e consequências diversas e até mesmo a antecipação de eventuais acontecimentos.

1.5 OS TIPOS DE DADOS

Conhecido o problema de pesquisa e estabelecido o caminho por onde a pesquisa deverá se organizar, o trabalho do pesquisador passa a ser o de buscar e organizar dados que o ajudem na resolução do problema e na realização dos objetivos.

As variáveis correspondem a dados e informações que permitem visualizações, análises e medições que respondem ao problema proposto. São elementos que descrevem diferentes medidas, valores e atributos que determinada categoria pode assumir. Assim como o problema, a identificação, organização, tabulação e análise dos dados garantem a qualidade do trabalho a ser realizado.

Possivelmente o problema de pesquisa envolve mais de uma variável, o que se exige na análise relacional. Nesse sentido, pode-se atentar para a variável dependente ou independente. A variável independente é aquela que não sofre influência das demais; a dependente é a que tem seu comportamento alterado a partir de mudanças ocorridas nas outras variáveis. Dessa forma, os modelos de análise comumente verificam como uma ou um conjunto de variáveis independentes é capaz de influenciar o comportamento da variável dependente.

A variável serve para identificar determinado comportamento, característica, medida ou atributo naqueles elementos que compõem a amostra ou a população em estudo. De acordo com o tipo de característica que deve ser comparada e/ou analisada, é possível identificar tipos de variáveis. O Quadro 2 resume as principais das características das variáveis mais comumente encontradas nos estudos na área das Ciências Sociais.

Tipo	Características		Exemplo
Quantitativa	Mensurável numericamente, é representada por números. Possibilidade de ordenamento	Contínua	Assume valores dentro de uma escala contínua de valores, com sentido para as frações obtidas.
		Discreta	Fazem sentido apenas em valores inteiros.
Qualitativa	Apresenta atributos definidos a partir de categorias	Ordinal	Possibilidade de ordenação de categorias pré-estabelecidas
		Nominal	Não pode ser hierarquizada

Fonte: Elaborado pela autora.

A partir da organização das variáveis, é possível estabelecer indicadores e compor índices. Os indicadores representam informações contidas nas variações das variáveis selecionadas. O índice pode representar a variação conjunta de diversas variáveis. A organização e a interação entre os tipos de variáveis e a composição de índices e indicadores são apresentadas e discutidas posteriormente.

Reconhecendo aqueles dados que são importantes para a resposta a ser dada ao problema de pesquisa configurado, pode-se iniciar, então, a busca prática pelas informações. Apresentar alguns elementos capazes de auxiliar os pesquisadores nesse objetivo é o tema da próxima unidade de estudo.

Questões para a reflexão:

- 1) Qual a importância do observador para a construção do objeto e a realização da Ciência?
- 2) A realidade pode ser considerada relativa? Ela depende de quem vê e como ela é processada por quem a percebe. Imagine alguém privado de algum dos sentidos que nos compõem como ser humano (alguém surdo ou cego, por exemplo): será que as percepções da realidade são as mesmas? O que é parecido? O que é diferente? Procure conversar com alguém que perceba o mundo de forma diferente da sua; como isso afeta o cotidiano desta pessoa e sua forma de interpretar a realidade? Se você não conhece alguém que perceba o mundo de forma diferente da sua, procure na Internet histórias que relatem essa realidade. Pode ser surpreendente!

- 3) Qual a importância do acúmulo de conhecimento realizado pela Humanidade? O que ocorreria se, hipoteticamente, a Ciência não evoluísse? Quais as consequências para a nossa vida?
- 4) Considere o desenvolvimento rural como um objeto de estudo. É possível identificar como esse objeto sofre transformações ao longo do tempo? Que ênfases foram dadas ao estudo do desenvolvimento rural? O que leva o desenvolvimento rural a passar de uma análise agrícola para a percepção atual, bem mais complexa, de um rural como modo de vida? Pode-se perceber essa transformação do objeto acompanhando a transformação da própria forma de fazer Ciência ao longo do tempo?
- 5) Formule três problemas de pesquisa relacionados ao processo de desenvolvimento rural da sua região.
- 6) Que dados você reconhece como importantes para a análise do processo de desenvolvimento rural?
- 7) Qual a diferença entre variável e indicador?

Para assistir:

O Ponto de Mutação (Mindwalk) – 1:51:22 (1990)

Apresenta o debate entre as formas de fazer/perceber a Ciência ao longo da história da Humanidade.

Erin Brockovich – Uma mulher de talento – 2:11:11 (2000)

Inspirado em uma história real. Dona de casa começa a investigar a qualidade da água na sua região.

Para ouvir:

Oito Anos

Compositores: Dunga/Paula Toller

Intérprete: Adriana Calcanhoto

Menino tem uma série de dúvidas curiosas sobre o cotidiano.

BUSCANDO OS DADOS: DISPONIBILIDADE, SELEÇÃO E LIMITAÇÃO DOS DADOS

Estabelecer o problema de pesquisa é apenas o começo da busca pelo entendimento de uma realidade contextualizada. O reconhecimento do problema identifica ao pesquisador em que áreas do conhecimento e temas específicos será necessário realizar o esforço de busca por informações. O objetivo desta unidade é o de instrumentalizar o pesquisador, tanto em termos conceituais como em termos práticos, para a busca pelos dados capazes de auxiliar na resposta do seu problema de pesquisa. Nesse sentido, trata-se de uma etapa inicial e primordial do trabalho de pesquisa, merecendo muita atenção dos pesquisadores.

Para isso, esta unidade está dividida em cinco seções, além de apresentar algumas questões para a reflexão ao final. A primeira seção apresenta uma retomada da noção de dados e quais são os tipos de dados com os quais podemos trabalhar. A segunda seção apresenta a importância e uma sugestão de roteiro para a realização de uma etapa exploratória na busca dos dados disponíveis no ambiente virtual. Alguns elementos referentes à busca e à coleta de dados são apresentados na terceira seção.

Em seguida, há um esforço na apresentação das principais bases de dados disponíveis atualmente, tanto nacional como internacionalmente. A quinta e derradeira seção apresenta alguns elementos que precisam ser ponderados, em relação aos dados, no momento da tabulação e da realização da análise das relações percebidas. Brevemente, são apresentadas para os pesquisadores algumas limitações dos diferentes tipos e das diferentes bases existentes.

2.1 O QUE SÃO DADOS E QUE TIPOS DE DADOS PODEM SER RECONHECIDOS

Pensando sobre os primórdios das Ciências, podemos perceber que, se a curiosidade é quem faz com que o cientista se interesse pela solução de determinadas questões, é apenas com a obtenção e organização dos dados que se torna possível sistematizar respostas e soluções para os mais diversos problemas. A forma de fazer Ciência é a mesma: perguntar e procurar informação de modo a adquirir elementos para articular uma resposta.

Segundo Sayão e Sales (2015, p. 80):

o dado de pesquisa é uma “informação criada ou coletada no curso da pesquisa científica, e que são frequentemente formatadas de maneira a torná-las adequadas à comunicação, interpretação e processamento por um computador”.

Os dados são os elementos que permitem ao pesquisador o estabelecimento de relações; são eles que permitem a validação ou não da pesquisa realizada. Uma boa organização e formatação dos dados é fundamental para a validação das análises desenvolvidas. Quanto mais pioneiro for o trabalho, maior a importância da organização e da sistematização dos dados com clareza e objetividade.

Existe um debate que envolve o próprio conceito de dados e a percepção de que eles são gerados para a consolidação de objetivos distintos, por diferentes tipos de pesquisa em diferentes áreas do conhecimento (SAYÃO; SALES, 2015, p. 7). Os dados podem ser identificados de diversas maneiras que envolvem desde a sua origem até sua temporalidade. O Quadro 3 apresenta os principais tipos de dados.

Classificação	Tipos	Características
Temporalidade	Imediatos	Dados com valor de análise imediato, no momento seguinte ao da coleta.
	Futuros	Dados que adquirem valor de análise com o passar do tempo.
	Reprodutíveis	Passíveis de reprodução, por exemplo, através de experimentos.
	Irreprodutíveis	Capturados em momento específico.
Origem	Observacionais	Obtidos pela observação direta do pesquisador, referindo-se a lugares e temas específicos (registros históricos).
	Computacionais	Resultados de modelos computacionais ou simulações. Podem ser replicados ao longo do tempo.
	Experimentais	Resultam de situações controladas em bancadas de laboratórios. Em tese, podem ser reproduzidos.
Fase da Pesquisa	Dados brutos, crus ou preliminares	Obtidos diretamente dos instrumentos científicos.
	Derivados	Resultados de combinação de dados.
	Canônicos ou Referenciais	Coleções de dados consolidados e arquivados em bases sob a gestão e responsabilidade de centros de estudos consolidados.

Fonte: Elaborado pela autora com base em Sayão e Sales (2015).

Alguma atenção deve ser dada às limitações dos tipos de dados: os dados computacionais, por exemplo, ainda que em tese sejam reprodutíveis, precisam ser armazenados de forma adequada e comumente dependem da aplicação de determinados *softwares* que podem deixar de ser utilizados no ambiente acadêmico, por exemplo. Em relação aos dados experimentais, ainda que seja possível a reprodução do experimento, existem casos em que isso se torna inviável por conta das dificuldades e dos custos da realização do experimento (por exemplo, dados recolhidos de sondas enviadas aos limites do Sistema Solar). A coleta de dados precisa ser planejada. Quando o pesquisador procede à coleta de dados, ele deve reconhecer a importância do elemento para o desenvolvimento da sua pesquisa. Um elemento que pode ajudar na escolha dos dados, entre outras coisas, é a realização de uma pesquisa exploratória sobre o tema a ser estudado.

Ainda nesta primeira seção, convém identificar a relação entre dado, informação e conhecimento. Nesse sentido, o desenvolvimento de uma área de pesquisa reconhecida como sistema de gerenciamento de informações nos ajuda a realizar a distinção conceitual. Podemos identificar o dado como o elemento mais bruto correspondente ao registro de eventos ocorridos; a informação, por sua vez, representa o dado configurado para o entendimento de determinada situação, advindo do processo de dar sentido a um conjunto de dados. O conhecimento resulta, então, da interpretação das informações que é realizada a partir da percepção do pesquisador (ROSINI; PALMISANO, 2003).

Podemos considerar que um dos objetivos gerais em relação à realização de atividades e trabalhos de pesquisa é o de transformar dado em informação, contribuindo para a geração de conhecimento na nossa sociedade.

2.2 PESQUISA EXPLORATÓRIA: PRIMEIRO CONTATO SISTEMATIZADO COM O TEMA DE PESQUISA

A partir da identificação do nosso problema de pesquisa (ou seja, da pergunta que se quer responder com o trabalho de busca a ser realizado) é possível iniciar o mapeamento de dados e informações que podem ser utilizados.

Toda pergunta envolve a identificação de um tema de análise. O primeiro passo, então, é o de realizar uma busca geral sobre o tema. Há alguns anos, o método de busca por essa bibliografia geral demandava a previsão de um tempo importante de pesquisa em bibliotecas institucionais, em que o estudante/pesquisador precisava verificar em um grupo de fichas (chamadas fichas catalográficas) o que estava disponível sobre o tema em estudo. Os que vivenciaram esse tempo conseguem identificar que atualmente o desenvolvimento tecnológico e o acesso facilitado a computadores tornou o trabalho de pesquisa sobre um tema muito mais dinâmico e mais rápido.

Mas, se por um lado o trabalho manual podia ser exaustivo e ter como resultado poucas referências em algum tema, o modelo atual nos apresenta em menos de um segundo milhares — quiçá milhões — de direcionamentos na rede de computadores. Existem diversos sites que facilitam a busca por qualquer tipo de informação. Um dos mais utilizados atualmente ainda é o “Google”.

Apenas para apresentarmos um exemplo: a busca pela expressão ‘desenvolvimento rural’ retorna mais de três milhões de resultados e utilizando a expressão ‘Plageder’ temos aproximadamente sessenta e quatro mil (!) direcionamentos. O planejamento dessa pesquisa é útil no sentido de organizar o tempo dedicado à atividade exploratória, uma vez que, com um contingente tão amplo de resultados, é preciso começar a delimitar mais claramente o tema de pesquisa considerando a viabilidade, em termos de tempo e recursos, do trabalho a ser realizado.

A partir do tema geral, evocado pelo problema de pesquisa formulado, é possível ir refinando a pesquisa nas bases de dados. O grande cuidado que deve ser tomado neste momento é o da identificação de fontes confiáveis de informação. Sendo a rede de computadores um ambiente dos mais democráticos da sociedade atual, é possível identificar elementos contraditórios nos mais diversos temas. Assim, há uma quantidade de informações que precisam ser confirmadas e selecionadas, para que não se corra o risco de, num trabalho científico, termos uma informação errada.

Dessa forma, um mapeamento inicial permitirá aos alunos e pesquisadores a identificação de assuntos e conteúdos mais recorrentes, bem como a identificação das principais correntes teóricas que debatem o tema pesquisado e quais são as contradições que envolvem esse tema. Logo, ainda que se possa considerar que o trabalho de pesquisa foi facilitado com o advento e a difusão da rede de computadores, ele demanda hoje um critério muito claro de análise no sentido de que possa ser realizada uma separação entre o conteúdo que de fato auxilia a solucionar o problema de pesquisa e aquele que pode aumentar o caminho a ser percorrido, causando inclusive confusão em relação à resposta.

Alguns sites não trazem elementos capazes de aprofundar os temas tratados já em nível de graduação. Entre eles, podem ser citados: Wikipedia, Info-Escola, ColaWeb e outros. O que se sublinha neste momento é que a busca por informações e referências para a realização de um trabalho acadêmico precisa estar adequada ao nível de dedicação e complexidade que se espera.

Identificado o tema, pode-se pesquisar em ambientes que organizam a busca em material científico, como livros e artigos em revistas especializadas nos mais diversos temas. Por exemplo, o chamado Google Acadêmico (*Google Scholar*) (<https://scholar.google.com.br>), que, a partir do tema, realiza busca em periódicos científicos tanto em Língua Portuguesa e em publicações nacionais

quanto em outras línguas e publicações internacionais. O Portal Periódicos da Comissão de Aperfeiçoamento de Pessoal em Nível Superior (CAPES) (<http://www-periodicos-capes-gov-br.ez45.periodicos.capes.gov.br>) também reúne quantidade importante de publicações de qualidade no Brasil e no exterior sobre os mais diversos temas de pesquisa. A instituição oferta ainda treinamentos, por área do conhecimento, que podem ajudar no manejo e na busca de artigos no ambiente. Algumas editoras científicas também oferecem material de pesquisa que pode ser útil na organização de referências para a realização de pesquisas.

Universidades e instituições de pesquisa organizam e produzem pesquisa nos mais diversos campos, sistematizando a publicação dos resultados em revistas e relatórios específicos. Em relação ao Desenvolvimento Rural, pode-se citar, a título de exemplo, as revistas: REDES, Revista de Economia e Sociologia Rural, Revista Ideas, Revista NERA entre outras. Além das revistas e publicações, cada Universidade possui um ambiente de Repositório Digital, em que a produção científica da instituição de ensino fica disponível para pesquisa. O ambiente de repositório digital da UFRGS, por exemplo, pode ser acessado através de <http://www.lume.ufrgs.br>.

Muito do material consultado será de fundamental importância para a realização da pesquisa. Não raro, o material organizado nesta pesquisa exploratória compõe o próprio referencial teórico do trabalho. Nesse sentido, ainda que não seja exatamente esse o objetivo desta publicação, convém alertar de antemão que qualquer referência textual ao material desenvolvido por outro pesquisador deve, invariavelmente, ser referenciado. Por referência, entende-se aqui a indicação, ao longo do texto, do autor da ideia original. Desse modo, sempre que o trabalho de pesquisa utilizar-se de conceitos e ideias de outros autores, eles devem ser citados.¹ A utilização de trechos de outros pesquisadores constitui plágio. De acordo com o dicionário da Língua Portuguesa (FERREIRA, 1986), o plágio é a “apresentação feita por alguém, como de sua própria autoria, uma obra intelectual etc. produzido por outrem”. Nesse sentido, é preciso bastante cuidado na organização das referências bibliográficas e nas citações que eventualmente venham a ser feitas durante as análises realizadas no trabalho. Toda referência a ideias desenvolvidas por outros autores, pesquisadores e alunos

¹ A Associação Brasileira de Normas Técnicas (ABNT) (www.abnt.org.br) normatiza as formas de referência para os mais distintos materiais de pesquisa.

precisa estar devidamente especificada no sentido de evitar o enquadramento em crime previsto pelos Códigos Civil e Penal.²

2.3 BUSCA E COLETA DE DADOS

Conforme apresentado por Gerhardt (2009, p. 37) no material desenvolvido sobre métodos de pesquisa, os dados podem ser primários (coletados diretamente pelos pesquisadores) ou secundários (quando os pesquisadores utilizam-se de uma organização de dados previamente realizada e disponibilizada aos demais interessados no tema). No caso da coleta primária de dados, há uma série de métodos envolvendo diversas concepções teóricas e procedimentos que podem ser consultados especificamente (ibid., p. 38). A formatação de um banco de dados, além de ser importante para a pesquisa sob a responsabilidade de quem realizou a busca e coleta de dados, pode interessar também a outros pesquisadores. Nesse sentido, é interessante pensar na necessidade e na viabilidade de disponibilização dos dados organizados. Se a pesquisa é financiada por alguma instituição científica, por exemplo, é preciso atentar-se a disponibilização dos recursos não está condicionada exatamente à formação de uma base de dados que deve ficar disponível a outros interessados.³

Neste volume, interessa-nos a utilização das bases de dados secundárias, ou seja, aquelas que pesquisadores e instituições organizaram previamente. Esses dados exigem atenção para a análise e, em muitos casos, uma nova organização a partir dos objetivos que estão previstos para o trabalho de pesquisa. A apresentação da base de dados — bem como das eventuais transformações que a análise da pesquisa específica exige — deve estar detalhadamente exposta em seção (no caso de artigos científicos) ou capítulo (no caso de trabalhos acadêmicos) dedicada à análise metodológica.

A utilização das bases de dados secundárias permite uma economia de tempo e recursos, evitando a realização desnecessária de coletas de dados, muitas vezes caras e exaustivas. Atente-se, entretanto, que essa utilização decorre da pertinência dos dados para a análise do problema de pesquisa e para o cumprimento dos objetivos, bem como da anuência do pesquisador principal ou orientador.

2 A Universidade Federal Fluminense desenvolveu um documento bastante elucidativo sobre o tema para esclarecimento de seus pesquisadores e estudantes. A cartilha pode ser acessada em <http://www.noticias.uff.br/arquivos/cartilha-sobre-plagio-academico.pdf>.

3 Mais detalhes sobre a gestão de arquivos podem ser acessados pelos interessados em Sayão e Sales (2015).

2.4 ALGUMAS BASES DE DADOS DISPONIBILIZADAS NO BRASIL E NO MUNDO

Com a emergência e difusão de novas tecnologias, houve uma maior divulgação dos ambientes de disponibilidade de dados e informações em diversas áreas. Apresentados os problemas de pesquisa, reconhecida a pergunta que cada pesquisador quer responder, é possível buscar elementos que ajudem na resposta.

Serão apresentadas neste material algumas bases de dados que podem ser utilizadas para a solução dos problemas de pesquisa. São elencadas aqui variáveis de produção municipais e que envolvem a possibilidade de avaliação e identificação de processos de desenvolvimento rural. Por questões inerentes ao processo tecnológico, alerta-se que, com o passar dos anos, pode haver alterações no formato das bases em si; entretanto, a lógica de busca e da reflexão sobre aquilo que pode ou não ser utilizado em determinados casos pode ser um pouco mais perene do que o próprio instrumento.

O Instituto Brasileiro de Geografia e Estatística (IBGE) é uma das principais instituições que organiza bases de dados no país. Sua dedicação em reunir e disponibilizar dados é muito anterior ao advento dos recursos computacionais. Segundo Penha (1993, p. 17-19), o IBGE foi criado em 1938 a partir da necessidade da reunião de dados e informações que permitissem aos governantes do país organizar a tomada de decisão, considerando o vasto e diverso território ocupado. Desde então, o IBGE se transformou em referência na organização e busca de dados sobre o Brasil. Por meio da rede de computadores, é possível obter uma série de dados que ajudam na análise das situações rurais (www.ibge.gov.br).

O site da instituição traz diversos indicadores e dados sobre economia, população e geociências. Considerando a página principal da instituição é possível destacar os seguintes temas, sobre os quais é possível encontrar dados e informações:

- Indicadores (informações sobre Trabalho e Rendimento — Pesquisa Nacional por Amostra Continuada, PNAD e Pesquisa Mensal de Emprego, ambas com dados para as regiões metropolitanas do país; Agropecuária — dados de estoque e produção; Pesquisa Mensal de Comércio; Pesquisa

Mensal de Serviços; indicadores referentes às Contas Nacionais; Índices de Preço);

- População (Indicadores Sociais Diversos — saúde, educação etc.; dados sobre Economia Urbana; Censos Demográficos; Pesquisa de Orçamentos Familiares; Pesquisa Nacional de Saúde Escolar; Pesquisa Nacional de Saúde; Projetos e Estimativas da População; Tábuas de Mortalidade);
- Economia (dados sobre Indústria, Construção, Inovação e Empreendedorismo no Brasil, inclusive com a Pesquisa Industrial de Inovação; Censos Agropecuários; dados sobre as empresas do país; dados sobre a finança nacional; Matriz de Insumo-Produto brasileira; além de diversos dados sobre o uso de tecnologias no Brasil);
- Geociências (informações territoriais; sistemas de georreferenciamento; dados sobre Geografia Urbana, redes e fluxo territoriais; publicações sobre os Recursos Naturais, incluindo documentos de Diagnóstico Ambiental e Zoneamento Ecológico Econômico).

O ambiente de pesquisa do IBGE é bastante completo e possui diversas outras informações, relatórios e documentos. Quanto aos indicadores agropecuários, por exemplo, são apresentadas as seguintes informações: estoques, produção agrícola e produção pecuária. Além das informações produtivas, essa aba apresenta uma série de preços e indicadores de trabalho e emprego.

De uma maneira geral, ao clicar em algum dos temas, a página direciona o pesquisador a um ambiente em que o conceito geral das informações disponíveis é apresentado. No lado esquerdo da tela, é possível obter o relatório principal em formato PDF para a leitura, bem como as tabelas que contêm as principais informações sobre o tema. São apresentados a metodologia de coleta de dados e diversos outros dados que podem ser verificados na busca por elementos que ajudem no tratamento da questão teórica e da formação da base de dados. Há possibilidade de acesso a tabelas em formato de planilhas eletrônicas que apresentam os dados quantitativos aos pesquisadores.

Cada planilha pode ser vista em ambiente *online* ou pode ser descarregada no computador em que é feito o acesso, garantindo a possibilidade de reorganização dos dados e das tabelas. É um ambiente rico em dados e informações que merece ser analisado pelos estudantes e pesquisadores.

Acessando as informações é possível estabelecer comparações de um mesmo produto ou região ao longo do tempo, ou entre diferentes regiões, unidades da federação ou municípios. Essas informações podem ser interessantes para a contextualização da realidade a ser estudada e problematizada. O papel do pesquisador após a organização dos dados é identificar elementos que expliquem as variações ou não modificações das situações configuradas (objeto de debate na Unidade IV deste volume).

A formatação dos relatórios é diversificada. As informações sobre a população, por exemplo, são acompanhadas de infogramas, gravuras e mapas que ajudam na compreensão da situação estudada. A grande maioria das informações apresenta a divisão entre a população rural e urbana, o que pode interessar a alguns.

Caso seja utilizada a informação e a comparação entre população urbana e rural, é sempre interessante identificar o conceito de rural e urbano para o IBGE. A própria metodologia indicada no site permite essa diferenciação.

Nesse ambiente, diversas informações e documentos comumente utilizados em pesquisa nas Ciências Sociais estão presentes. Um desses documentos é o Censo Demográfico. Estão disponíveis dados dos dois Censos mais recentes, os de 2000 e 2010. Considerando o problema de pesquisa, cada pesquisador/aluno deve identificar quais são as informações capazes de auxiliar na solução do seu problema de pesquisa. Também é possível, a partir das informações do Censo, obter os dados para regiões, estados e municípios. No caso dos Censos, é possível baixar a base de dados de domicílio e pessoas, tendo-se, assim, acesso aos chamados microdados, que permitem a tabulação dos itens do questionário do Censo a partir da base de dados original (certamente resguardadas todas as normas de sigilo inerentes ao processo). Por tratar-se de uma base bastante extensa, com grande quantidade de informações, os dados devem ser descarregados por estado — se o objetivo for, por exemplo, organizar uma base nacional do Censo.

Conforme já comentado, as tabelas estão organizadas para mesorregiões, microrregiões e municípios. Após visualizar as informações, o pesquisador pode reorganizar as tabelas, separando ou reunindo as informações como mais convém a sua análise.⁴ Realizado o acesso às diversas tabelas, é possível sele-

⁴ Esses elementos de manipulação das variáveis identificadas como importantes para a resolução do problema de pesquisa são o tema do próximo módulo.

cionar e salvá-las para análise em formato de planilhas eletrônicas. O próprio instrumental da planilha permite uma série de ações úteis à análise posterior, incluindo a realização de gráficos, por exemplo. Trataremos desses elementos nas próximas unidades.

Percebam que, ainda que seja possível reunir uma grande quantidade de dados e informações utilizando essas bases disponíveis, deve-se sempre saber o que se quer descobrir. Nesse sentido, é preciso reiterar a necessidade de clareza em relação ao problema de pesquisa e aos objetivos do trabalho. O Censo, por exemplo, representa uma fotografia da situação. Seus dados se referem a um ponto no tempo. É preciso saber se esse é um intervalo temporal suficiente. Caso o problema de pesquisa exija avaliações temporais, caso seja preciso verificar como a variável muda ao longo do tempo, é necessário buscar as informações em outras bases de dados.

Além do Censo Populacional, o IBGE disponibiliza acesso às informações do Censo Agropecuário (a partir do tema Economia). O Censo Agropecuário, assim como o Censo Populacional, é uma fotografia da situação da atividade econômica desse setor no país, mas, diferentemente do documento anterior, seus dados correspondem aos anos de 1996 e 2006 (o mais recente). No documento, assim como na base de dados, é possível obter informações específicas sobre produção, propriedades, estrutura agrária e muitas outras que estão diretamente relacionadas ao ambiente rural. Para aqueles interessados no processo de desenvolvimento rural, esse é um documento indispensável para o debate. Assim como para os demais documentos do IBGE, há um relatório em PDF e são disponibilizadas tabelas em formato de planilha eletrônica que podem ser descarregadas no equipamento de acordo com o interesse do pesquisador. Recomenda-se atenção ao pesquisador/estudante para a verificação daqueles elementos, dentre os vários disponíveis, que, de fato, ajudam na análise do seu problema de pesquisa.

Nas tabelas, algumas informações são indicadas em linhas (por exemplo, os municípios e as Unidades da Federação) e outras em colunas (por exemplo, quantidade produzida). Sendo assim, há uma ordem de apresentação dessas informações que garante a leitura e compreensão dos dados que estão sendo apresentados. Pela verificação da disposição das linhas e das colunas e como as células se encontram, é possível encontrar a informação procurada. Se considerarmos a Figura 1, por exemplo, que apresenta uma das tabelas disponibilizadas

pelos dados do Censo Agropecuário 2016, podemos afirmar que, em 2006, foi registrada a presença de 378.853 estabelecimentos de agricultura familiar no Rio Grande do Sul.

Figura 1 — Exemplo de tabela sobre agricultura familiar no Brasil — Censo Agropecuário/IBGE

Grandes Regiões e Unidades da Federação	Agricultura familiar - Lei nº 9.326		Não familiar	
	Estabelecimento	Área (ha)	Estabelecimento	Área (ha)
Brasil	4 266 267	99 182 634	689 369	293 577 343
Norte	412 666	16 611 277	63 112	38 524 487
6 Rondônia	75 955	2 282 577	8 102	5 942 231
7 Acre	25 114	1 928 287	4 281	2 503 786
8 Amazonas	61 630	14 759 558	4 954	2 393 195
9 Roraima	3 644	423 660	1 447	1 079 614
10 Pará	105 565	6 077 304	26 044	8 847 946
11 Amapá	2 005	108 205	602	742 504
12 Tocantins	42 009	2 008 207	13 750	10 009 062
Nordeste	2 116 131	28 379 852	206 529	41 738 259
13 Maranhão	282 042	4 594 639	24 937	6 558 523
14 Piauí	220 728	3 779 442	24 443	5 147 331
15 Ceará	345 009	2 442 410	39 508	4 455 640
16 Rio Grande do Norte	111 719	1 046 070	8 943	2 361 059
17 Paraíba	161 663	1 566 056	10 137	2 280 169
18 Pernambuco	275 720	2 568 324	29 070	2 867 752
20 Alagoas	101 769	492 495	8 932	1 420 829
21 Sergipe	30 229	788 090	92 278	2 079 586
22 Bahia	465 707	3 948 056	30 778	10 605 604
Sudeste	609 755	12 771 239	222 342	42 846 474
23 Minas Gerais	437 328	8 828 622	98 381	24 347 087
25 Espírito Santo	67 416	968 813	16 347	8 672 241
26 Rio de Janeiro	44 121	489 737	14 722	1 800 095
27 São Paulo	100 890	2 993 267	76 722	14 644 462
Sul	849 693	13 654 931	156 530	28 726 482
29 Paraná	302 620	4 262 039	68 228	9 109 623
30 Santa Catarina	369 852	2 842 241	25 756	2 490 205
31 Rio Grande do Sul	177 221	6 550 651	62 546	11 800 654
Centro-Oeste	217 622	9 258 956	109 476	36 000 820
32 Mato Grosso do Sul	4 007	1 164 237	24 997	23 910 753

Fonte: Censo Agropecuário 2006 (IBGE).

O cuidado que o IBGE e outros pesquisadores e instituições de pesquisa têm na organização e apresentação dos dados é fundamental para a compreensão dos futuros leitores em relação à apresentação dos resultados da pesquisa.

O IBGE disponibiliza também outras formas de acesso aos dados e informações. O Sistema IBGE de Recuperação Automática (SIDRA) permite que o pesquisador reúna as informações de modo a atender necessidades específicas de cruzamento de dados. Nesse ambiente, é possível obter os dados referentes às séries temporais mais longas que venham a interessar para a pesquisa. O sistema exige uma dedicação um pouco maior em relação à navegação; entretanto, é importante para diversos recortes setoriais e temporais. Constantemente o IBGE atualiza e melhora a *interface* do SIDRA no sentido de torná-lo mais intuitivo aos usuários. São encontrados nesse ambiente dados temporais além dos anuais (mensais, trimestrais e quadrimestrais), bem como a organização

dos dados de acordo com desenhos geográficos específicos, como é o caso, por exemplo, dos Territórios da Cidadania.

Outros ambientes disponibilizados pelo IBGE para a pesquisa de dados são o IBGEEst@do e o IBGECid@de, que disponibilizam, respectivamente, um conjunto de informações sobre os estados e municípios. Ao acessar o estado ou município brasileiro para o qual se pretende obter as informações, ao pesquisador é apresentado um grande conjunto de dados que variam desde os populacionais até a estrutura de prestação de serviços que pode ser encontrada na escala territorial acessada (estado ou município). Após digitar o nome do município no IBGECid@des, o pesquisador acessará dados sobre as condições socioeconômicas e históricas daquela localidade, conhecendo inclusive a previsão do tempo para o dia da pesquisa e os próximos. Trata-se de um ambiente extremamente importante para a obtenção de informações e cujo acesso idealmente se poderia recomendar a cada morador de alguma localidade a fim de conhecer melhor onde mora.

O IBGE é uma das principais fontes de dados e informações no que diz respeito à realidade brasileira. Dificilmente um trabalho sobre desenvolvimento rural seria capaz de prescindir da busca de informações nessa instituição. Espera-se, dessa forma, que esta breve menção a algumas alternativas apresentadas no IBGE para a busca de dados seja capaz de incentivar os alunos a uma visita ao ambiente de rede desta instituição.

Existem outras fontes de dados e informações além do IBGE. O debate que compreende a distinção existente entre os processos de crescimento e desenvolvimento econômico levou à análise de variáveis que permitissem uma compreensão mais ampla dos processos reconhecidos como de desenvolvimento.⁵ A literatura sobre esse processo envolve a percepção do desenvolvimento humano, que procura evidenciar aspectos além dos econômicos, como a renda. Algumas instituições se preocuparam, então, em disponibilizar dados e informações capazes de apresentar esse novo interesse de pesquisa. No Brasil, é disponibilizada pelo Programa das Nações Unidas para o Desenvolvimento

5 O processo de crescimento compreende o acréscimo que envolve essencialmente elementos quantitativos de análise, cujo expoente na literatura econômica é o Produto Interno Bruto (PIB). O processo de desenvolvimento é mais abrangente: envolve parâmetros que permitam a percepção de uma melhoria na qualidade de vida das pessoas. Nesse sentido, essencialmente a partir da identificação de uma crise ambiental, além da percepção em relação ao aumento da desigualdade, a década de 1980 marca definitivamente a distinção entre os dois processos. Ainda que se possa considerar o crescimento econômico como um dos elementos importantes do processo de desenvolvimento (principalmente em sociedades desiguais, como no caso brasileiro), esse já não é mais um processo que, por si, leve automaticamente ao desenvolvimento.

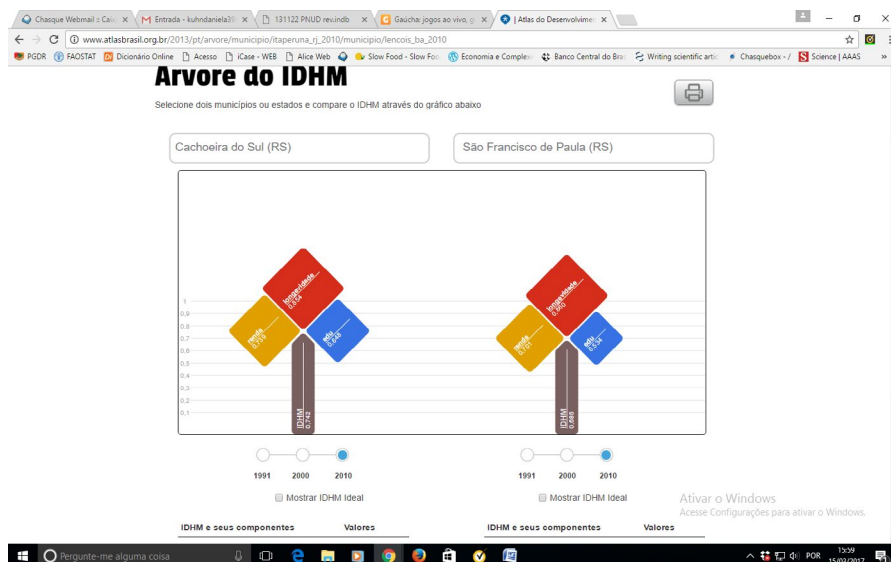
(PNUD) uma base de dados que apresenta o chamado Índice de Desenvolvimento Humano (IDH).

O IDH procura consolidar em um elemento quantitativo três dimensões de análise que poderiam ser utilizadas como parâmetros para a qualidade de vida das pessoas: Renda, Educação e Saúde. Esse índice se configura a partir da reunião de diversas variáveis para cada dimensão. A dimensão saúde é percebida a partir da expectativa de vida ao nascer; a dimensão educação está representada pela média geométrica entre duas variáveis, escolaridade da população adulta e fluxo escolar da população jovem; e a variável de renda utilizada é a renda *per capita*.⁶ O índice é calculado a partir da média geométrica de seus três componentes. Ele varia de 0 a 1: quanto mais próximo de 1, melhor a condição de desenvolvimento do local para onde o índice foi calculado. Valores menores que 0,5 são considerados como de IDH muito baixo.

Os dados e as informações sobre o IDH podem ser acessados em <http://www.atlasbrasil.org.br/2013/>. Além do valor do índice calculado, o PNUD fornece aos pesquisadores outras informações socioeconômicas, como mortalidade infantil, analfabetismo por faixa etária, o Índice de Gini (que indica a desigualdade de renda), entre outros. Assim como no IBGE, o ambiente do Atlas do Desenvolvimento Humano oferece uma série de relatórios que podem servir como fonte de informação e permitem a configuração de tabelas de acordo com o interesse da pesquisa. As consultas podem ser feitas por município, regiões metropolitanas, estados e abrangendo todo o país. De acordo com a consulta escolhida, podem ser organizados mapas e histogramas que permitem uma melhor visualização dos dados e das informações. É possível comparar o IDH de dois municípios com o recurso gráfico da Árvore do IDH, conforme nos mostra a Figura 2.

6 A renda *per capita* de um lugar é calculada a partir do somatório da produção do local, dividido pela população do local (que pode ser uma residência, um bairro, um município, um estado ou um país).

Figura 2 — Exemplo de consulta ao Atlas do Desenvolvimento Humo Árvore do IDH



Fonte: Atlas do Desenvolvimento Humano (PNUD, 2013).
Consulta realizada pela autora.

Além de conhecer um *ranking* municipal em relação às variáveis disponibilizadas, o ambiente do Atlas do Desenvolvimento Humano permite a geração de um relatório por município. Esse relatório pode ser impresso pelo pesquisador ou salvo em formato PDF para posterior impressão/visualização. O ambiente desenvolvido pelo PNUD é interessante e permite reconhecer novos arranjos e dimensões que podem ser considerados importantes para a análise dos processos de desenvolvimento rural.

Existem ambientes que nos permitem consultar dados regionais e estaduais. No Rio Grande do Sul, o site da Fundação de Economia e Estatística (FEE) é um dos mais utilizados por pesquisadores das mais diversas áreas.⁷ O material disponibilizado pela FEE apresenta diversos caminhos para a obtenção de dados. Em relação a séries históricas de indicadores, podem ser encontrados elementos relacionados a estatísticas internacionais, o Índice de Desenvolvimento Socioeconômico (Idese), indicadores da atividade do agronegócio gaúcho, um

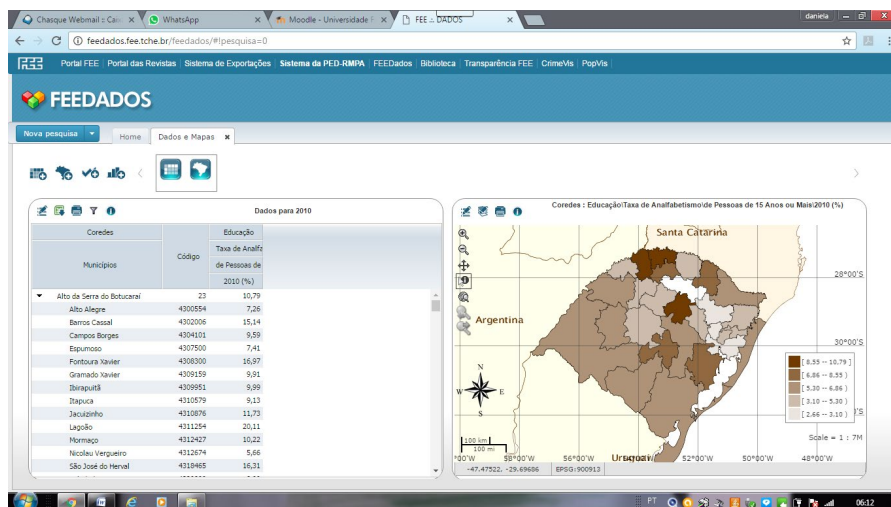
⁷ No contexto de escrita deste material, convém evidenciar que a FEE está em vias de extinção, em virtude de uma alegada necessidade da realização de redução de gastos do governo estadual.

conjunto de indicadores ambientais, índices de exportações e rentabilidade de exportações, índices de venda da indústria e do comércio, entre outros. A FEE é responsável também por calcular a Matriz Insumo-Produto do Estado.

No ambiente do Perfil Socioeconômico, são encontradas diversas informações municipais. Os municípios estão organizados em ordem alfabética para a consulta. Ali é possível reunir informações sobre a situação de cada município. É possível também reunir os dados por Conselhos Regionais de Desenvolvimento (COREDES). Atualmente, o Rio Grande do Sul tem 28 COREDES.

Além do perfil socioeconômico, o ambiente da FEE permite a busca por outras informações que podem ser organizadas e configuradas de acordo com o objetivo da pesquisa que está sendo realizada. A FEEDADOS, com dados para o Rio Grande do Sul, é uma boa base; tem formato acessível e informações atualizadas. É possível organizar informações sobre as seguintes variáveis: agricultura, caracterização do território, comércio, comunicações, contabilidade social, demografia, domicílios particulares permanentes, educação, emprego, energia elétrica, extração vegetal, finanças públicas, instituições financeiras e justiça. As informações podem ser agregadas por estado, mesorregião, microrregião, COREDE ou município. Escolhidas as variáveis e o nível de agregação, a base permite a geração de uma tabela de dados cruzados que atende ao interesse do pesquisador. Além da tabela, o pesquisador tem como resultado da sua consulta um mapa que apresenta graficamente a frequência e distribuição da variável. A maioria das variáveis possui mais de uma observação (ou seja, está disponível na base para mais de um ano). Algumas possuem séries relativamente longas, como, por exemplo, a de população, cujo dado mais antigo é de 1970. A Figura 3 demonstra o resultado da pesquisa sobre analfabetismo de pessoas acima dos 15 anos, em 2010, para os COREDES do Rio Grande do Sul. Após a geração do resultado, o pesquisador tem a opção de gerar uma tabela em planilha eletrônica que pode ser descarregada no seu computador.

Figura 3 — Exemplo de resultado de pesquisa no FEEDADOS



Fonte: FEEDADOS.

Disponível em: <feedados.fee.tche.br/feedados/#!pesquisa=0>.

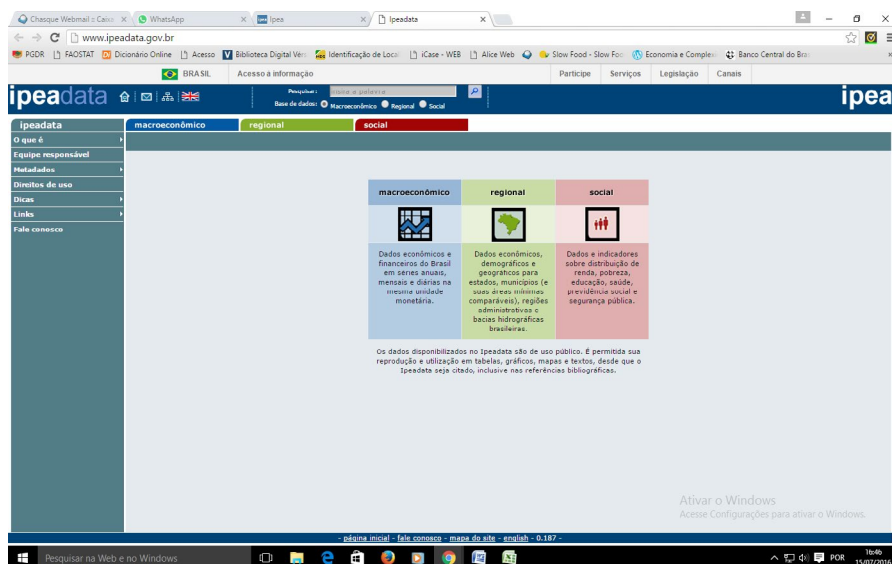
Em relação a essa ou qualquer outra das bases de dados, sugere-se que cada estudante dedique um tempo para verificar a lista de variáveis, pois certamente alguma pode ser utilizada para a realização de parte da pesquisa. Em termos geográficos, o diferencial do FEEDADOS é o fato de ser possível enxergar os dados a partir de diferentes escalas geográficas.

Além dos dados em si, a FEE disponibiliza um importante volume de relatórios e material analítico sobre situações conjunturais e análises setoriais específicas. É uma importante fonte de referências bibliográficas também, além de fonte de dados.

Outra opção para a obtenção de dados especializados é a consulta à base do Instituto de Pesquisa Econômica Aplicada (IPEA). O IPEA é um órgão administrado pelo governo federal. O objetivo da instituição é fornecer suporte técnico ao governo federal para a realização de políticas públicas. Além de manter um serviço que disponibiliza uma série de dados estatísticos, o instituto oferece uma série de textos de análise de diversos setores e dimensões da sociedade brasileira. Possui diversas ferramentas para a exploração dos dados disponíveis.

O IPEADATA apresenta as variáveis disponíveis na base de dados em três dimensões com diversos temas cada uma: dados macroeconômicos (Balanço de Pagamentos, Câmbio, Comércio Exterior, Consumo e Vendas, Contas Nacionais, Correção Monetária, Economia Internacional, Emprego, Estoque de Capital, Finanças Públicas, Informações Financeiras, Indicadores Sociais, Moeda e Crédito, Percepção e Expectativa, População, Preços, Produção, Projeções, Salário e Renda e Transporte), dados regionais (Agropecuária, Comércio Exterior, Consumo, Vendas, Contas Regionais, Eleições, Emprego, Estoque de Capital, Finanças Públicas e Informações Geográficas) e dados sociais (Assistência Social, Demografia, Desenvolvimento Humano, Educação, Emprego, Habitação, Mercado de Trabalho, Renda e Saúde). Os dados são bastante diversos, com séries e fontes diversificadas; por isso, é importante que o pesquisador se detenha com atenção na relação entre o problema de pesquisa e a variável escolhida. A Figura 4 apresenta o ambiente inicial do IPEADATA.

Figura 4 — Ambiente inicial do IPEADATA



Fonte: IPEADATA.

Disponível em: <ipeadata.gov.br>.

Outro instrumento disponível através do IPEA aos pesquisadores é o IPE-AGEO. Essa ferramenta exige um pouco mais de manejo no ambiente virtual,

mas é interessante para a organização de figuras e para a realização de algumas análises estatísticas. O IPEAGEO é um *software* gratuito com funcionalidades de georreferenciamento. A ferramenta já tem as malhas georreferenciadas e cruza esses dados com as informações socioeconômicas.

Análogo ao Atlas do Desenvolvimento Humano do PNUD, o IPEA também disponibiliza o Atlas da Vulnerabilidade Social, com dados específicos sobre vulnerabilidade e exclusão social nos municípios e estados brasileiros. A página identifica o índice de vulnerabilidade social (que varia de 0 a 1, em que 0 é a menor vulnerabilidade e o 1 indica vulnerabilidade máxima). A consulta pode ser realizada por município, por região metropolitana, estado e abrangendo todo o país. As variáveis disponíveis envolvem elementos das dimensões Capital Humano, Renda e Trabalho e Infraestrutura Urbana.

Muitas pesquisas voltadas à análise do processo de desenvolvimento exigem a utilização de dados referentes às políticas públicas e, em especial, em relação a programas de financiamento. No que tange ao processo de desenvolvimento rural, encontramos análises específicas que utilizam, por exemplo, os dados do Programa Nacional de Fortalecimento da Agricultura Familiar (PRO-NAF). Esses dados estão disponíveis, assim como diversos outros dados sobre a realização de políticas públicas (fiscal, monetária, comercial e cambial), no ambiente do Banco Central do Brasil, o BACEN. É possível encontrar, nesse ambiente, dados sobre a Balança de Pagamentos, Importação e Exportação, preços dos produtos, taxas de câmbio, gastos do governo com os diversos tipos de financiamentos — inclusive em séries temporais relativamente longas. Lembre-se de que dados que absorvem impacto decorrente da realização de política monetária, como dados de inflação e preços agrícolas (bem como informações referentes a comparações com o ambiente internacional, a taxa de câmbio), exigem um tratamento temporal diferenciado dado efeitos inflacionários que ocorrem ao longo do tempo. Nesse sentido, é importante considerar que com o passar o tempo a relação de preços e de aquisição, tanto da moeda estrangeira como da moeda nacional, se altera. Sendo assim, é preciso realizar, quando da utilização desses dados, um método de deflação, a fim de validar a comparação de um preço com outro em espaços temporais diferentes. Esse exercício metodológico permite que os dados sejam preparados para descontar efeitos inflacionários e apresentar variações reais e não nominais.

Além do BACEN, outros órgãos de governo têm, em suas páginas na Internet, ambiente de apresentação e disponibilização de dados. Dependendo do tema, é importante fazer uma busca no site do órgão governamental responsável pela política em estudo. São exemplos de páginas de governos que reúnem informações específicas: o site do Ministério da Agricultura, Pecuária e Abastecimento (MAPA); Ministério do Meio Ambiente; a Companhia Nacional de Abastecimento (CONAB); DATASUS, entre outros.

O MAPA, por exemplo, tem um ambiente de estatística que reúne informações importantes para a produção agropecuária brasileira. Dependendo do tema de pesquisa, pode ser interessante buscar informações nesse ambiente também.

Para os alunos com interesse nas políticas territoriais no Brasil, o governo disponibiliza o Sistema de Informações Territoriais (SIT) (<http://sit.mda.gov.br/mapa.php>). A base de dados apresenta informações para os municípios que compõem os Territórios da Cidadania espalhados pelo Brasil. A partir do mapa do Brasil, é possível acessar também uma série de informações municipais gerais e relacionadas a políticas públicas. Ao escolher o território que se quer estudar, é possível obter outras informações sobre a realidade municipal.

Em relação a dados para os países latino-americanos, a Comissão Econômica para a América Latina e o Caribe (CEPAL) disponibiliza uma base de dados que apresenta uma série de variáveis interessantes, bem como as séries temporais, o CEPALSTAT (acessado por meio de http://estadisticas.cepal.org/cepalstat/WEB_CEPALSTAT/Portada.asp). Eventualmente, na necessidade de buscar dados comparativos para determinadas situações, esse pode ser um ambiente bastante útil. Além das variáveis tradicionais (demográficas, socioeconômicas e ambientais), o ambiente apresenta uma série de temas transversais (Objetivos de Desenvolvimento do Milênio, Gênero, Coesão Social, Populações Tradicionais, Juventude, Tecnologias de Informação e Comunicação). Além da configuração de consulta de acordo com o objetivo da pesquisa, é possível acessar os perfis nacionais dos países latino-americanos.

Considerando a necessidade de uma visão mais abrangente de determinada situação ou contexto, dados que permitem a comparação entre países podem ser acessados através da FAO (sigla em inglês para *Food and Agriculture Organization of the United Nations*, Organização das Nações Unidas para a Alimentação e a Agricultura). A instituição disponibiliza dados estatísticos sobre

diversos temas: população, produção, segurança alimentar, questões ambientais e climáticas, entre outros. Eventualmente, pode ser interessante comparar a realidade a ser estudada com a realidade de outros países. Entretanto, a página apresenta as informações em Inglês. O pouco conhecimento da língua pode ser superado com a utilização do Google tradutor, que pode auxiliar na identificação das variáveis mais significativas.

Foram identificadas aqui algumas fontes de dados. Existem ainda outras fontes governamentais e particulares, fundações e institutos que podem auxiliar nas pesquisas, dependendo de cada tema. Um movimento importante que auxilia os pesquisadores é o de organização dos chamados Observatórios Sociais (www.osbrasil.org.br). Esses observatórios estão espalhados pelo país e permitem o acesso a informações em diversos âmbitos em prol do exercício da cidadania.

Temas e assuntos específicos podem, inclusive, ser solicitados separadamente, quando as informações não são disponibilizadas ao público em geral. Em alguns casos, esse acesso é pago, mas, em outros, uma mensagem eletrônica de solicitação permite o acesso a outras tabelas e informações. Não há como esgotar todas as fontes de dados e informações que existem atualmente na rede de computadores; entretanto, é sempre importante reconhecer que a busca por elementos em qualquer uma dessas bases exige que o pesquisador tenha clareza em relação aos seus objetivos de trabalho. Certamente, cada um de nós, enquanto pesquisador, deve ter condições de buscar e explorar as diversas fontes de dados secundárias disponíveis, sempre evidenciando a fonte e a credibilidade da instituição que nos fornece a informação.

2.5 ALGUNS ELEMENTOS IMPORTANTES SOBRE DADOS E INFORMAÇÕES

Há uma quantidade bastante grande de dados e informações disponíveis na Internet e nas instituições. Essa breve unidade de conteúdo apresentou algumas das opções, aquelas mais comuns e mais gerais. Espera-se que a leitura deste material, bem como o acesso aos sites sugeridos, tenha servido para familiarizar alunos e pesquisadores com alguns dos elementos que podem auxiliar na resposta do problema de pesquisa de cada um. Certamente o início do trabalho pode envolver o acesso a uma quantidade grande de informações,

mas a clareza em relação ao direcionamento da pesquisa permite a seleção das variáveis mais significativas e importantes em cada caso. O passo a passo para o acesso às informações foi apresentado para que aqueles leitores não familiarizados com esses ambientes pudessem ter segurança para iniciar seu percurso. Nesse sentido, após a localização daqueles elementos significativos, o trabalho é o de seleção, tratamento e organização, assuntos dos próximos módulos do nosso livro.

Ainda que se possa celebrar o avanço tecnológico e a facilidade de acesso às informações do mundo atual, brevemente convém ponderar sobre a qualidade dos dados que estão disponíveis. O leitor que acompanha este volume pode concordar comigo que, dada a imensidão de temas e regiões que foram contemplados nas bases exemplificadas, reconhecem-se eventuais dificuldades de tabulação; na busca pelos dados, pode haver problemas na qualidade e na frequência dos mesmos.

Em instituições públicas, a demanda por recursos não raramente interrompe séries temporais que podem ser prejudicadas nas análises sazonais, por exemplo. A própria série temporal pode exigir ajustes e modernizações conceituais, com o avanço tecnológico e com a mudança de comportamento das pessoas, o que pode levar a uma dificuldade de comparação entre dados de períodos diferentes. Nesse sentido, recomenda-se, na medida do possível, a atenção às notas técnicas da base de dados secundária que se utiliza.

Existem bases de dados, por exemplo, em que nem todas as observações necessárias são apresentadas. Em bancos de dados que comparam países, por exemplo, ocorre de algumas nações não disponibilizarem os dados, seja por questões políticas ou porque não têm condição de realizar o vasto trabalho de campo necessário. Sendo assim, a base fica incompleta.

Elementos como emancipações também podem trazer a necessidade de ajustes a uma base de dados. Por exemplo, uma série de dados municipais gaúchos que procure trabalhar com variáveis coletadas antes de 1992 até os dias atuais possuiria quantidades de municípios diferentes, uma vez que, no início dos anos 1990, houve uma série de emancipações no Rio Grande do Sul. O ajuste pode ser feito, por exemplo, utilizando-se a composição dos municípios-mãe.

Além de elementos quantitativos, pode-se alertar também quanto à qualidade dos dados obtidos. No Brasil, por exemplo, há uma noção de que os dados

de renda representam uma aproximação, visto que se percebe na realização da pesquisa de campo que pessoas com maior poder aquisitivo podem ter uma tendência a indicar ao entrevistador uma renda menor do que a real, enquanto aqueles mais pobres podem indicar uma renda maior (por várias razões, inclusive constrangimento).

De qualquer forma, a existência do dado e a sua disponibilidade permite ao pesquisador realizar uma série de análises que podem ser consideradas, por hipótese, bastante próximas da realidade.

Questões para a reflexão

- 1) Se tivéssemos o seguinte problema de pesquisa: “É possível evidenciar que o acesso ao crédito através de programas governamentais contribuiu para o processo de desenvolvimento rural no Rio Grande do Sul nos últimos trinta anos?”, que tipo de dados e variáveis você acha que nos ajudariam na análise?
- 2) Que dados poderiam ajudar a compreender nosso problema de pesquisa?
- 3) Das bases de dados apresentadas, qual é aquela que você considera ser capaz de lhe fornecer elementos para a solução do problema? Entre na base e busque os dados e informações que lhe interessam.
- 4) Encontre no site IBGE Cid@des o município em que você mora. Observe as informações disponibilizadas. O que você já sabia? O que aprendeu ao consultar o IBGE? Agora verifique os dados sobre o seu município na base FEEDADOS. Quais informações se repetem em relação àquelas encontradas no IBGE? Quais são as informações novas? Busque no Atlas do Desenvolvimento Humano os dados para o Rio Grande do Sul e veja quais são parecidos e quais são diferentes do Estado. Se quisermos considerar a situação do seu município em relação ao Rio Grande do Sul, ao Brasil e à América Latina, que bases de dados poderiam ajudá-lo?

Escolha uma variável e tente realizar esta análise: qual a situação do seu município em relação a essa variável quando comparada à situação gaúcha e à situação brasileira? É necessário imaginar algum tipo de política pública para melhorar essa situação? Como você argumentaria, agora que tem possibilidade de comparar diferentes situações em relação à mesma variável?

Para assistir:

Série Megadados — Série desenvolvida pelo canal de televisão National Geographic, em 2015, com 12 episódios que apresentam informações sobre os mais diversos temas. Os episódios podem ser assistidos na rede de computadores.

ORGANIZANDO, CONFIGURANDO E FORMATANDO A BASE DE DADOS

A partir da identificação do problema de pesquisa e do reconhecimento dos possíveis elementos que podem nos ajudar a compreender o contexto que tentamos analisar, é preciso organizar e interpretar as variáveis.

A observação com atenção àquilo que os dados podem apresentar representa um importante elemento para a análise dos problemas de pesquisa. Os dados estão relacionados com aquilo que queremos perceber seja em uma população, em alguma situação específica ou ao longo de determinado período. As variáveis, a configuração da base e as medidas que servirão como princípio da análise dos trabalhos científicos devem todas ser apresentadas na seção dedicada à apresentação metodológica, tanto em artigos como em monografias, dissertações e teses. Normalmente, essa etapa da escrita é realizada após a apresentação da base teórica do trabalho. Cada uma das variáveis a ser apresentada deve, então, estar em consonância com a abordagem teórica escolhida para nortear o trabalho e precisa ser justificada na apresentação metodológica. Além disso, é sempre importante lembrar-se de referenciar as fontes e o ano dos dados que estão sendo utilizados. Segundo Han e Kamber (2006, p. 99-100), pode-se imaginar que a organização da base de dados inicia-se por três etapas: limpeza dos dados (redução das inconsistências, verificação da existência de todos os dados da base, etc.); integração (organização dos dados que serão compilados a partir de fontes diferentes) e transformação dos dados (cuidado com as categorias de análise, verificação da necessidade de agrupamento, criação de novas variáveis, etc.). Dependendo do tipo de estudo, pode-se necessitar das três etapas (e mais algumas) ou não. Nesta unidade, aborda-se de uma maneira bem ampla esse tratamento; mais informações e o atendimento da configuração

de bases específicas podem ser obtidos no material referenciado ao final do volume.

Esta unidade do material está dividida em três momentos. A primeira apresenta alguns elementos para a organização das bases de dados. Em seguida, são apresentados brevemente alguns dos *softwares* mais comumente utilizados para a configuração e análise de dados atualmente nas Ciências Sociais. Por fim, algumas considerações e sugestões gerais são apresentadas ao leitor.

3.1 CONFIGURANDO A BASE DE DADOS

Existem diversas formas de organizar os dados e variáveis em análise. Cada pesquisador pode ter preferências sobre como realizar a configuração das suas variáveis. Alguns utilizam fichas por temas; outros poderiam organizar um fichário por variável, entre outras sugestões. Atualmente, a possibilidade de utilização dos computadores facilitou a análise das bases de dados. A partir de alguns *softwares*, é possível obter informações interessantes sobre os dados (por exemplo, a possibilidade de obtenção de médias, máximos e mínimos de um grupo de observações de alguma variável).

A ideia geral para a organização da base de dados é imaginar qual o melhor “desenho” que permite a visualização das situações que se pretende analisar. Basicamente, a disposição para a organização dos dados na base envolve a decisão sobre as informações que serão colocadas nas linhas e nas colunas. Essa composição gera uma matriz de informação em que cada uma das células (encontro entre linhas e colunas) nos informa algo sobre a relação daquilo definido na linha com aquela variável definida na coluna, conforme a Figura 1 da unidade anterior já nos indicava.

Podemos considerar que as informações das linhas (ou seja, organizadas verticalmente na base de dados) nos informam o *locus* em que observamos o dado da coluna. Ao dispor as informações dessa forma, linhas e colunas, podemos obter informações tanto do local em si (por exemplo, os municípios gaúchos), como da distribuição da variável em determinado ambiente (por exemplo, participação da atividade agrícola no PIB municipal).

A sistematização dos dados em planilhas eletrônicas configura-se em forma de matriz, como lógica de organização. Talvez a planilha mais utilizada e conhecida seja o *Excel*, disponível no pacote *Microsoft Office*®.

As planilhas apresentam o formato de linhas e colunas para a inserção dos dados e permitem a utilização de uma série de instrumentos de formatação que facilitam o trabalho com os dados.

Nas planilhas eletrônicas, é possível organizar grandes bases de dados.¹ Os dados devem ser digitados ou importados de outras bases. Se trabalharmos com os municípios, por exemplo, é possível organizar os municípios nas linhas e identificar as diversas variáveis nas colunas. Analisando a linha, é possível perceber a situação do município; olhando a coluna, é possível verificar informações referentes à variável em análise. No conjunto, é possível verificar quais são os melhores e os piores indicadores em cada município e no grupo de municípios de uma maneira geral. Dessa forma, a partir de uma tabela relativamente simples, pode-se obter uma grande quantidade de informações. A Figura 5 apresenta um exemplo de organização de base de dados que servirá para chamar a atenção para alguns elementos importantes quando pensamos na organização das informações para o nosso trabalho de análise.

Observando a Figura 5, podemos verificar que se trata da organização de informações municipais. A primeira linha da tabela identifica as variáveis. Nas linhas estão os municípios e nas colunas os valores observados naquele município em cada uma das variáveis. Por exemplo, em Cachoeira do Sul a população é de mais de 83 mil pessoas, estando mais de 85% dos domicílios localizados na zona urbana do município.

¹ No *Excel* 2010, por exemplo, a base de dados pode ter até 1.048.576 linhas e até 16.384 colunas, o que, na prática, tem se mostrado mais do que suficiente para os mais diversos trabalhos. A organização de bases maiores, ditas “mais pesadas”, exige um aparato tecnológico específico.

Figura 5 — Exemplo de base de dados Excel

Município	População (quant. Pessoas) - Censo Demográfico 2010	Situação do Domicílio - Urbana (% do total de domicílios)	Situação do Domicílio - Rural (% do total de domicílios)	Pessoas com 15 anos ou mais que não sabem ler e escrever	Pessoas com 15 anos ou mais que não sabem ler e escrever (%)	Existência de Secretária Municipal do Meio Ambiente (exclusiva)
1						
2	83.827	85,53	14,47	5.019	5,99%	1
3	2.592	42,25	57,75	181	6,98%	0
4	9.752	66,76	33,24	643	6,59%	0
5	5.182	87,98	12,02	100	1,93%	0
6	39.685	70,84	29,16	2.851	7,18%	0
7	20.537	63,32	36,68	1.160	5,65%	1
8	43.111	56,22	43,78	1.758	4,08%	0
9	34.556	88,29	11,71	1.825	5,28%	0
10	19.250	88,73	11,27	579	3,01%	0
11	23.726	79,92	20,08	867	3,65%	0
12	23.965	79,51	20,49	1.042	4,35%	1
13						
14	Fontes: Ministério do Desenvolvimento Social - Plano Brasil sem Miséria (2013) e sites das prefeituras					
15						
16						
17						
18						
19						
20						
21						
22						
23						

Fonte: Elaborada pela autora.

A formatação da base de dados permite que sejam reunidas informações com unidades de descrição diferentes. No caso da Figura 5, podemos perceber que algumas das variáveis estão indicadas como quantidade de pessoas e que outras estão representadas como percentual do total de população da categoria. Temos aqui duas maneiras de apresentar os dados. Quando nos interessam as quantidades totais de elementos que apresenta determinada característica (por exemplo, a quantidade de crianças em idade escolar para comparar com a quantidade de vagas disponíveis em escolas), o dado é apresentado em unidade que decorre diretamente da contagem referente à característica ou ao elemento em análise. Esses dados são aqueles que encontramos em fontes secundárias de informação, ou seja, as bases disponíveis comentadas na unidade anterior.

Outra maneira de apresentar um dado é a forma percentual da característica em relação ao total. Quando nos interessa perceber características em termos de participação na população total ou variações de uma para a outra, é comum que se apresente na base de dados a relação percentual dessa característica em relação ao total de uma população mais ampla. O cálculo desse valor ocorre com a divisão do número em que a característica é observada em relação ao total da população, multiplicado por 100 (o que nos leva à relação

percentual/porcentual). Se observarmos as informações referentes à população de São Lourenço do Sul, saberemos de quanto era a população do município no Censo Populacional realizado pelo IBGE em 2010 (43.111 pessoas) e quantas dessas pessoas, maiores de 25 anos, não sabem ler e escrever (1.758 pessoas). Com essas informações somos capazes de indicar que 4,08% da população acima dos 25 anos de idade no município de São Lourenço do Sul não sabe ler nem escrever. Observe que essa informação percentual é gerada, a partir da organização das duas informações anteriores, na própria base de dados. Se prestarmos atenção ao conteúdo da célula 8F (correspondente à linha 8, que traz as informações de São Lourenço do Sul, e à coluna F, que indica o percentual da população maior de 25 anos que não sabe ler nem escrever), perceberemos que ela reúne informações de duas colunas anteriores (a E e a B, dividindo os valores um pelo outro, respectivamente). O formato da apresentação do resultado, em termos percentuais, é configurado com a seleção da coluna e a utilização do ícone % na barra de comandos que formata números. Essa instrução e formato podem ser replicados para toda a coluna.

Outro elemento que podemos perceber na Figura 5 é a existência de uma variável que representa, em números, situações de existência ou não (no caso do exemplo, a existência de Secretarias Municipais Exclusivas de Meio Ambiente na estrutura de gestão do município). Essa é uma variável que tenta captar alguma percepção em relação à situação do município na condução de situações que envolvam conflitos ambientais. Essas variáveis, que nos permitem identificar determinados comportamentos, sem representá-las totalmente (ou representando determinadas situações com limitações bastante claras) e evidenciam como os elementos se relacionam, são chamadas de variáveis *proxi*. Esse tipo de variável é utilizado quando aquilo que se pretende estudar ainda é considerado de difícil mensuração, mas é possível guardar alguma relação com o objeto de estudo e aquilo que a variável identifica (metodologicamente é preciso que a escolha e a análise estejam bem embasadas do ponto de vista teórico). No exemplo, prefeituras com a Secretaria de Meio Ambiente independente representam um ganho, em tese, para o debate das questões ambientais nos municípios, considerando que muitas vezes o tema é tratado dentro da Secretaria de Agricultura (que representa, em algumas circunstâncias, posições não coincidentes em relação a alguns temas).

Além de ser uma *proxi*, a variável da coluna G é binária, ou seja, é representada por 0 ou 1. Nesse caso, o número 1 representa a presença da estrutura que trata exclusivamente dos temas ambientais e o 0 representa que esta não é uma estrutura de decisão presente no organograma municipal. Muitas variáveis podem ser representadas desta forma. Outro exemplo é a utilização de variáveis binárias para a identificação da presença de utensílios domésticos em uma residência (cada utensílio presente no domicílio — fogão, refrigerador, etc. — ganha valor 1; aquele que não está presente, 0). Por fim, de uma lista de itens é possível estabelecer um *ranking* em relação à estrutura domiciliar dos participantes da pesquisa.

É possível formar uma base exclusivamente de variáveis binárias, conforme apresentado na Figura 6, um exemplo de base binária que representa um questionário que integrava uma pesquisa referente a aspectos ambientais aplicado a prefeituras gaúchas.

Figura 6 – Exemplo de base de dados binária

	MUNICÍPIO											MIO																													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	
1	Municípios	1	2	3	4	5	6	7	8	9	10	11	MIO	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	FIM	TOTAL			
2	Ajuntaca	0	1	0	0	0	1	1	0	0	0	0	3	1	0	0	1	1	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	16	19
3	Azenha	1	1	1	0	1	1	1	1	0	0	0	7	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	0	1	8	15	
4	Antônio Prado	0	1	0	1	1	0	1	0	0	0	0	4	1	0	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	17	21	
5	Arroio dos Ratos	0	1	0	0	1	0	1	1	0	0	0	4	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	0	0	1	1	1	1	1	1	0	8	12	
6	Bento Gonçalves	1	1	1	1	1	0	1	1	1	1	1	10	1	0	0	0	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	10	20	
7	Bom Princípio	1	1	0	0	1	1	1	1	0	0	0	6	1	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	9	15	
8	Bossoroca	0	1	0	0	0	1	1	1	0	0	0	4	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	25	
9	Brazão	0	1	0	1	1	0	1	1	0	1	0	6	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	26	
10	Braga	0	1	0	0	1	0	1	1	1	0	0	6	1	0	0	1	0	0	0	0	1	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	0	11	17	
11	Cachoeira do Sul	1	1	1	1	0	0	1	1	1	0	0	8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	19	27	
12	Campanas do Sul	1	1	0	0	1	0	1	0	0	0	0	4	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17	21	
13	Candelária	0	1	0	0	1	0	1	1	1	0	0	5	1	1	0	1	0	0	0	1	0	1	0	1	0	1	0	0	1	0	0	1	0	1	0	1	8	13		
14	Canoas	1	1	0	1	1	1	1	1	0	1	0	8	1	0	1	1	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	14	22		
15	Capão Bonito do Sul	1	0	0	0	1	0	1	1	0	0	0	4	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	17		
16	Capão do Cipó	1	1	0	1	0	0	1	1	1	0	0	6	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	26		
17	Coronel Bascio	1	1	0	0	1	1	1	1	1	0	0	8	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	26		
18	Chapadão	0	0	0	0	0	0	1	1	0	0	0	2	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	29		
19	Cruzeiro do Sul	0	1	0	1	1	0	1	1	1	0	0	6	0	0	0	0	1	0	1	1	1	0	1	1	0	1	0	0	0	1	1	1	1	1	1	0	9	15		
20	Novembro	0	1	1	1	0	1	1	0	0	0	0	6	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	22		
21	Fragatas Velas	0	1	0	0	0	1	1	1	0	0	0	5	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	19	24		
22	Fortaleza dos Valos	0	1	1	1	0	0	1	0	0	0	0	5	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	22		
23	Frederico Westphalen	0	1	1	0	1	1	1	1	1	0	0	8	1	0	0	1	1	0	0	1	0	1	0	1	0	1	0	0	0	1	1	1	1	1	1	1	1	19	19	
24	General Câmara	1	1	0	1	0	1	1	1	1	0	0	7	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	17	24		
25	Gaspar	1	1	0	0	1	1	1	1	1	0	0	7	1	0	1	1	1	0	0	1	0	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	18	18	
26	Humaitá	1	1	1	1	0	0	1	1	0	0	0	6	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	24	
27	Imigrante	0	1	0	0	1	0	1	1	0	0	0	5	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	22	27		
28	Itauna	0	1	1	1	0	1	1	0	0	0	0	6	1	0	0	0	1	0	0	1	1	0	1	0	1	0	0	0	1	1	1	1	1	1	1	0	9	15		
29	Itaíba Nova	0	1	1	1	1	1	1	1	0	0	0	6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26		

Fonte: Elaborado pela autora.

Ao organizar os dados dessa forma, pode-se perceber que somando as linhas e as colunas é possível obter informações que permitem uma análise detalhada da situação. Na linha, reconhece-se a situação do município; no caso, quanto maior o somatório, melhor a situação municipal em relação a variáveis

ambientais. Na coluna, é possível reconhecer a situação da variável nos municípios. Qual é a mais presente no Estado? Qual a menos presente?

3.2 SOFTWARES DE TRABALHO COM BASE DE DADOS

Para o trabalho com base de dados talvez o *software* mais comumente utilizado seja o *Excel*, já mencionado neste material. Sua disponibilidade e a forma descritiva e amigável fazem com que muitos pesquisadores trabalhem a partir de bases organizadas nesse formato. Entretanto, existem outros como o *R*, o *Matlab*, o *SAS*, o *Stata* e o *SPSS*, que serão muito brevemente identificados nesta seção.

O *R* é um ambiente voltado para a chamada computação estatística. Suas ferramentas permitem a realização de inferências e simulações a partir da determinação de um processo de modelagem. É um *software* livre, que possibilita importar bases de dados configuradas em outras planilhas eletrônicas.

Outro *software* que permite uma série de interações entre os dados de uma planilha eletrônica é o *MathLab*. Ele tem funções de análise numérica, cálculos com matrizes, correlações, modelagem econométrica. O ambiente é relativamente fácil de ser manipulado e a fluência do analista melhora com a dedicação ao programa. Diferente do *R*, é um pacote de análise pago, o que nem sempre viabiliza sua utilização.

O *SAS* (sigla, em Inglês, para software de análise estatística), por sua vez, visa à aplicação integrada dos dados. A partir dele, é possível realizar acesso, gerenciamento, análise estatística e apresentação de dados para a geração de relatórios. É um pacote pago que atende a necessidade de quem o adquire, valendo-se de módulos de procedimentos (que incluem, entre outros, apenas a título de exemplo: módulo gráfico; módulo de análise de séries temporais, módulo de operação com matrizes, etc.).

Outro pacote estatístico de análise de dados bastante utilizado é o *Stata*, que, ao longo do tempo, vem atualizando suas versões. É um pacote comercial.

O *SPSS* é outro programa apropriado para a realização de análises estatísticas de um conjunto grande de dados, permitindo a geração de relatórios a partir de diversas organizações das variáveis.

Esses *softwares* possuem funções estatísticas que possibilitam a análise de grandes conjuntos de dados, configurando figuras de dispersão, cada um com

suas especialidades e especificidades de manuseio. É importante que o pesquisador saiba da existência das diversas possibilidades; entretanto, ao trabalhar com um dos pacotes, certamente iremos nos familiarizar e seguir trabalhando com ele. A próxima unidade é dedicada à apresentação das possibilidades de análise mais comuns trabalhadas a partir de uma base de dados, indicando como podem ser interpretados alguns desses resultados realizados nos pacotes estatísticos.

3.3 LIMITAÇÕES E CONSIDERAÇÕES SOBRE A CONFIGURAÇÃO DE BASE DE DADOS

Entre as principais recomendações que vale a pena relembrarmos aqui, temos a necessidade de clareza da relação daquilo que o nosso problema de pesquisa exige com o que a variável que compõe nosso estudo possa nos indicar. Dada a grande quantidade de dados disponíveis, ou a possibilidade de obtenção desses dados, é preciso pensar a variável analisada em termos da explicação que buscamos.

Além disso, precisamos estar atentos a alguns elementos gerais. A formação das bases tem muitos recursos de organização nas planilhas eletrônicas; por exemplo, a possibilidade de ordenar as observações das linhas pelo sentido que uma variável específica apresenta na coluna (do menor para o maior, do maior para o menor, em ordem alfabética, etc.). É preciso atentar para que as observações não se “misturem”, o que, além de não ajudar, torna a base de dados errada. Nesse sentido, é sempre preciso considerar que as ordenações tenham como base uma coluna, mas que abarquem toda a tabela, sob pena de tirarmos informações “do lugar” e desfazermos relações observáveis.

Questões para refletir:

- 1) Considere as seguintes variáveis relacionadas a municípios e pense se elas poderiam ser representadas por variáveis binárias em alguma base de dados: densidade demográfica; existência de Fundo de Proteção Ambiental; percentual da população que reside em área rural; participação do PIB agrícola; localização em algum Território da Cidadania.

- 2) Pense em algo que você gostaria de estudar. Seria possível montar uma base de dados? Que configuração ela teria?
- 3) Considere as seguintes informações obtidas no IBGE (Censo 2010) sobre alguns dos municípios que compõem o Território da Cidadania Médio Alto Uruguai no Estado gaúcho:
- Município: Boa Vista das Missões;
Área (km²): 185,7;
População Urbana: 886;
População Rural: 1.228;
Estabelecimentos da Agricultura Familiar: 256;
Pessoas em Extrema Pobreza: 179.
 - Município: Cerro Grande;
Área (km²): 74,7;
População Urbana: 830;
População Rural: 1.587;
Estabelecimentos da Agricultura Familiar: 427;
Pessoas em Extrema Pobreza: 287.
 - Município: Cristal do Sul;
Área (km²): 74,7;
População Urbana: 830;
População Rural: 1.587;
Pescadores: 0;
Estabelecimentos da Agricultura Familiar: 427;
Pessoas em Extrema Pobreza: 287.
 - Município: Engenho Velho;
Área (km²): 73,7;
População Urbana: 599;
População Rural: 928; Pescadores: 0;
Estabelecimentos da Agricultura Familiar: 218;
Pessoas em Extrema Pobreza: 271.

- Município: Frederico Westphalen;
Área (km²): 264,5;
População Urbana: 23.330;
População Rural: 5.510;
Estabelecimentos da Agricultura Familiar: 1.280;
Pessoas em Extrema Pobreza: 720.
- Município: Sagrada Família;
Área (km²): 77,4;
População Urbana: 785;
População Rural: 1.810;
Estabelecimentos da Agricultura Familiar: 612;
Pessoas em Extrema Pobreza: 305.
- Município: Três Palmeiras;
Área (km²): 175,8;
População Urbana: 2.090;
População Rural: 2.291;
Estabelecimentos da Agricultura Familiar: 692;
Pessoas em Extrema Pobreza: 420.
- Município: Lajeado do Bugre;
Área (km²): 73,6;
População Urbana: 706;
População Rural: 1.781;
Estabelecimentos da Agricultura Familiar: 610;
Pessoas em Extrema Pobreza: 287.
- Município: Novo Xingu;
Área (km²): 81;
População Urbana: 554;
População Rural: 1.203;
Estabelecimentos da Agricultura Familiar: 430;
Pessoas em Extrema Pobreza: 32.

Como seria possível organizá-los em uma base de dados? Utilizando uma planilha eletrônica (por exemplo, o *Excel*), organize os dados. Tente colocá-los em ordem alfabética utilizando o comando “dados”, “classificar” da planilha eletrônica.

ANALISANDO OS DADOS

Organizada a base de dados, é hora de analisar os dados organizados. Existem diversas formas de compreender aquilo que o conjunto de dados da base configurada pode nos mostrar. O objetivo desta unidade é iniciar os estudantes nesta aventura de descobrir o que os números nos dizem quando devidamente organizados. Para isso, os temas foram organizados em duas seções. A primeira apresenta como é realizada a estatística descritiva, bem como o significado dos seus principais elementos. A segunda seção evidencia a possibilidade de estabelecer relações entre as variáveis, essencialmente apresentando os elementos básicos da análise de correlação e as possibilidades de análise a partir da configuração de indicadores e índices.

4.1 PRINCIPAIS MEDIDAS DE ESTATÍSTICA DESCRITIVA

A Estatística é a área da Ciência que indica a probabilidade da ocorrência de determinados fenômenos. O manejo da base de dados pode ser feito a partir da identificação de funções estatísticas que permitem a interpretação de um conjunto de observações. As próprias planilhas eletrônicas possuem, no menu, fórmulas prontas ou permitem a inserção de fórmulas e comandos nas células que ajudam na formatação dos resultados. A Estatística Descritiva é aquela em que, a partir de uma base de dados, é possível reconhecer uma série de situações (média, máximos, mínimos, etc.).

A partir de um banco de dados, é possível identificar somatórios tanto nas linhas como nas colunas. Muitas vezes, utilizar o comando soma nos dá informação sobre o que está analisado ou serve para gerar um protocolo de

conferência em relação aos valores em análise. As planilhas eletrônicas possuem uma função específica de soma que é capaz de agregar um conjunto de dados, representado pela letra grega Σ . Além de somar o conjunto de dados, a partir da função já estabelecida, na célula da planilha esse comando também pode ser especificado como = somar (célula A; célula B). Lembre-se de que é sempre importante ter em mente a interpretação do resultado.

As planilhas eletrônicas permitem uma organização dos dados que facilita bastante o trabalho do analista. Além de permitir a visualização e organização dos dados, é possível também gerar gráficos que auxiliam o estudo dos resultados (tema para o nosso próximo, e último, módulo).

Ao abrir a planilha eletrônica, pode-se perceber que, a partir da barra de tarefas inicial, é possível acessar uma barra de tarefas de fórmula. Nessa barra, existe uma série de funções predeterminadas, entre as mais utilizadas para a análise de dados.

Para cada função, o analista precisa indicar na planilha o intervalo de células que está sendo avaliado. Nesse sentido, somatórios, médias, máximos e mínimos são facilmente identificados a partir da função. A média aritmética (a mais comumente utilizada) apresenta, num conjunto de valores, o resultado do somatório dividido pela quantidade de observações, indicando a média do valor para o conjunto de dados. Muitos estudos utilizam a média como o instrumento de comparação para o conjunto. A partir do cálculo da média, é possível identificar se determinada observação está acima ou abaixo dos parâmetros. Além da média aritmética, podem ser calculadas as médias geométricas e ponderadas.

A média é uma medida de tendência central, assim como a moda e a mediana. A moda, que também pode ser identificada para a base de dados como uma função da planilha, indica o valor mais frequente em uma base de dados.

A mediana indica exatamente o valor central quando as observações são colocadas em ordem crescente ou decrescente. De maneira geral, quanto mais próximos forem os valores de média e mediana, melhor a qualidade da média como instrumento de análise para o conjunto de dados. Além das medidas de tendência central, as planilhas permitem a identificação dos valores máximos e mínimos de determinado intervalo de dados, possibilitando dessa forma o reconhecimento da amplitude dos dados em análise.

Considerando uma das planilhas eletrônicas mais comumente utilizadas, a do *Excel*, serão apresentadas algumas funções desse *software*. A utilização da função que encontramos na planilha eletrônica depende da base de dados já estar organizada. A partir da aba funções, estatística e média, pode-se indicar qual o intervalo para o qual a média deve ser calculada. A Figura 7 apresenta o *layout* da planilha para o analista.

Figura 7 — Exemplo para o cálculo da média na planilha eletrônica Excel

The image shows the Microsoft Excel interface with the 'MÉDIA' (AVERAGE) function dialog box open. The dialog box displays the formula bar with the text '=742,6281533' and the range 'Núm1: 742,6281533'. Below the dialog box, a data table is visible with columns labeled H through Q. The table contains numerical data for various locations, such as 'ACEGUÁ', 'AGUA SANTA', 'AGUDO', etc.

	H	I	J	K	L	M	N	O	P	Q
4	4300034	ACEGUÁ	378,29	633,63	1,36	17,02			378,29	633,63
5	4300059	AGUA SANTA	283,57	1491,17	0,88	23,67			283,57	1491,17
6	4300109	AGUDO	427,84	861,06	1,93	20,40			427,84	861,06
7	4300209	ARABICÁ	319,56	925,70	3,26	36,92			319,56	925,70
8	4300307	ALEGREIM	195,15	961,62	2,36	13,45			195,15	961,62
9	4300406	ALEGRETE	459,15	705,29	6,52	86,55			459,15	705,29
10	4300455	ALEGRIA	210,31	692,76	2,46	21,69			210,31	692,76
11	4300471	ALHURRANTE TAMANDARÉ DO SUL	232,90	760,16	0,84	18,51			232,90	760,16
12	4300505	ALPESTRE	133,79	625,11	0,90	16,78			133,79	625,11
13	4300554	ALTO ALEGRE	354,04	880,85	2,58	27,86			354,04	880,85
14	4300577	ALTO FELIZ	313,51	802,20	1,82	20,83			313,51	802,20
15	4300600	ALVORADA	334,82	600,48	0,51	99,88			334,82	600,48
16	4300603	AMARAL FERRADOR	280,83	499,77	0,27	12,32			280,83	499,77
17	4300646	AMETISTA DO SUL	193,00	611,20	2,46	16,48			193,00	611,20
18	4300661	ANDRÉ DA ROCHA	337,21	792,19	4,37	27,98			337,21	792,19
19	4300700	ANTA GORDA	403,50	810,41	1,43	19,56			403,50	810,41
20	4300800	ANTÔNIO PRADO	623,98	1004,85	2,99	57,51			623,98	1004,85
21	4300885	ARAMBARÉ	195,28	599,72	0,49	51,50			195,28	599,72
22	4300887	ARARICÁ	303,12	609,69	0,37	46,64			303,12	609,69
23	4300900	ARATIBA	286,24	1171,70	2,80	28,06			286,24	1171,70
24	4301000	ARROIO DO MEIO	446,16	850,06	3,26	63,70			446,16	850,06
25	4301005	ARROIO DO SAL	427,95	800,40	2,03	94,36			427,95	800,40
26	4301007	ARROIO DO PADRE	192,55	883,18	0,23	4,87			192,55	883,18
27	4301010	ARROIO DOS RATOS	272,54	623,54	1,15	89,73			272,54	623,54
28	4301020	ARROIO DO TIGRE	246,84	654,51	2,38	30,62			246,84	654,51
29	4301030	ARROIO GRANDE	324,09	528,13	1,87	74,37			324,09	528,13
30	4301400	ARVOREZINHA	377,66	675,18	2,40	42,84			377,66	675,18
31	4301500	AUGUSTO PESTANA	419,01	1016,34	2,49	34,72			419,01	1016,34
32	4301555	AURORA	223,55	686,10	0,83	21,38			223,55	686,10
33	4301600	BASE	450,22	795,56	8,49	86,07			450,22	795,56
34	4301603	BALNEÁRIO PINHAL	338,55	662,85	4,59	75,56			338,55	662,85
35	4301604	BALNEÁRIO PINHAL	728,16	847,78	7,77	70,95			728,16	847,78

Fonte: Elaborado pela autora.

Uma dica interessante é clicar imediatamente na célula vazia abaixo da coluna de que se quer calcular a média ou à direita da linha. A planilha vai apresentar um intervalo definido, que pode ser alterado caso haja algum equívoco em relação aos dados que serão avaliados na média. Ao clicar no “OK” representado na Figura 7, a célula será preenchida com o valor resultado da função, ou seja, a média numérica do conjunto de variáveis solicitado. A Figura 8 apresenta o resultado do exemplo. Note que a célula que estava vazia é preenchida com o valor da média, mas, na linha de fórmula, estão especificados a função média e o intervalo de análise. Eventuais ajustes podem ser feitos no local onde a fórmula está indicada (marcada com a seta). Como o resultado está associado

ao intervalo, a alteração de qualquer valor no intervalo refletirá no resultado apresentado para a função. Essa propriedade, comum a qualquer função que trata um intervalo na planilha eletrônica, economiza tempo ao analista, que não necessita, a cada alteração na base de dados, reorganizar a fórmula da função.

Figura 8 — Exemplo de resultado da função média na planilha eletrônica Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
469	432225	4322251	TUPANDI	285,91	808,39	0,00	27,18		4322251	TUPANDI	285,91	808,39	0,00	27,18	285,91	808,39	
470	432230	4322301	TUPARENDI	266,06	768,25	4,24	44,02		4322301	TUPARENDI	266,06	768,25	4,24	44,02	266,06	768,25	
471	432232	4322327	TURUÇU	223,60	566,71	0,00	31,66		4322327	TURUÇU	223,60	566,71	0,00	31,66	223,60	566,71	
472	432234	4322343	LIBHIRETAMA	176,62	590,04	1,19	6,18		4322343	LIBHIRETAMA	176,62	590,04	1,19	6,18	176,62	590,04	
473	432235	4322350	UNIÃO DA SERRA	244,72	1068,08	0,80	8,94		4322350	UNIÃO DA SERRA	244,72	1068,08	0,80	8,94	244,72	1068,08	
474	432237	4322376	UNISTALDA	173,46	392,38	0,00	24,06		4322376	UNISTALDA	173,46	392,38	0,00	24,06	173,46	392,38	
475	432240	4322400	URUGUAIANA	465,95	714,26	7,14	92,25		4322400	URUGUAIANA	465,95	714,26	7,14	92,25	465,95	714,26	
476	432250	4322509	VACARIA	410,45	797,86	7,03	86,82		4322509	VACARIA	410,45	797,86	7,03	86,82	410,45	797,86	
477	432252	4322525	VALE VERDE	362,81	509,24	0,70	14,37		4322525	VALE VERDE	362,81	509,24	0,70	14,37	362,81	509,24	
478	432253	4322533	VALE DO SOL	252,93	553,23	0,83	9,12		4322533	VALE DO SOL	252,93	553,23	0,83	9,12	252,93	553,23	
479	432254	4322541	VALE REAL	293,55	795,41	0,00	53,51		4322541	VALE REAL	293,55	795,41	0,00	53,51	293,55	795,41	
480	432255	4322558	VANINI	287,03	780,56	3,77	36,75		4322558	VANINI	287,03	780,56	3,77	36,75	287,03	780,56	
481	432260	4322606	VENÂNCIO AIRES	340,02	779,48	3,72	48,26		4322606	VENÂNCIO AIRES	340,02	779,48	3,72	48,26	340,02	779,48	
482	432270	4322707	VERA CRUZ	335,01	796,95	2,15	39,87		4322707	VERA CRUZ	335,01	796,95	2,15	39,87	335,01	796,95	
483	432280	4322806	VERANÓPOLIS	544,35	1141,67	4,53	68,79		4322806	VERANÓPOLIS	544,35	1141,67	4,53	68,79	544,35	1141,67	
484	432285	4322855	VESPASIANO CORREIA	279,85	895,32	0,96	9,68		4322855	VESPASIANO CORREIA	279,85	895,32	0,96	9,68	279,85	895,32	
485	432290	4322905	VIADUTOS	324,93	683,03	2,36	29,27		4322905	VIADUTOS	324,93	683,03	2,36	29,27	324,93	683,03	
486	432300	4323002	VIAMO	393,26	707,05	2,39	92,79		4323002	VIAMO	393,26	707,05	2,39	92,79	393,26	707,05	
487	432310	4323101	VICENTE DUTRA	169,45	475,13	1,14	29,29		4323101	VICENTE DUTRA	169,45	475,13	1,14	29,29	169,45	475,13	
488	432320	4323200	VICTOR GRAEFF	328,36	589,66	4,63	27,09		4323200	VICTOR GRAEFF	328,36	589,66	4,63	27,09	328,36	589,66	
489	432330	4323309	VILA FLORES	554,45	728,71	2,94	26,19		4323309	VILA FLORES	554,45	728,71	2,94	26,19	554,45	728,71	
490	432335	4323358	VILA LÂNGARO	216,57	902,87	1,39	8,35		4323358	VILA LÂNGARO	216,57	902,87	1,39	8,35	216,57	902,87	
491	432340	4323408	VILA MARIA	384,38	1101,27	2,61	26,29		4323408	VILA MARIA	384,38	1101,27	2,61	26,29	384,38	1101,27	
492	432345	4323457	VILA NOVA DO SUL	178,94	442,26	2,00	31,73		4323457	VILA NOVA DO SUL	178,94	442,26	2,00	31,73	178,94	442,26	
493	432350	4323501	VIA ALBURE	394,20	900,87	2,46	22,44		4323501	VIA ALBURE	394,20	900,87	2,46	22,44	394,20	900,87	
494	432360	4323606	VISTA ALEGRE DO PRATA	409,66	1231,80	0,51	15,70		4323606	VISTA ALEGRE DO PRATA	409,66	1231,80	0,51	15,70	409,66	1231,80	
495	432370	4323705	VISTA GAÚCHA	267,22	990,13	3,23	13,82		4323705	VISTA GAÚCHA	267,22	990,13	3,23	13,82	267,22	990,13	
496	432375	4323754	VITÓRIA DAS MISSÕES	240,51	482,44	0,34	10,97		4323754	VITÓRIA DAS MISSÕES	240,51	482,44	0,34	10,97	240,51	482,44	
497	432377	4323770	WESTFALIA	485,80	1136,67	2,38	16,28		4323770	WESTFALIA	485,80	1136,67	2,38	16,28	485,80	1136,67	
498	432380	4323803	XANGRI-LÁ	479,67	798,62	1,43	98,68		4323803	XANGRI-LÁ	479,67	798,62	1,43	98,68	479,67	798,62	
499									430001	JAGUÁ DO SUL	#N/D	#N/D	#N/D	#N/D	#N/D	#N/D	
500									430002	JAGUÁ DOS PATOS	#N/D	#N/D	#N/D	#N/D	#N/D	#N/D	

Fonte: Elaborado pela autora.

Na mesma lista de funções estatísticas, é possível selecionar as funções de mínimos e máximos, gerando assim novos valores para complementação da análise. A importância da média para a análise de um grupo de variáveis é a obtenção de um parâmetro referente a um conjunto específico de dados. Na Figura 8, a média apresentada refere-se à renda domiciliar *per capita* média dos municípios gaúchos. Assim, para o Estado do Rio Grande do Sul, analisando o conjunto dos 497 municípios, a renda domiciliar *per capita* média é de R\$ 742,83. Essa informação é útil para compararmos com algum município ou região específica do Estado. A partir dessa comparação é possível identificar se um município está em condição de renda melhor ou pior do que a média do Estado.

Além de permitir a identificação da média, é possível identificar, por meio da função máximo, o município de maior renda domiciliar *per capita*. O processo de identificação, na base de dados, do município com maior valor para essa variável no Estado gaúcho se dá por meio da função CTRL+L (localizar). O caminho é o mesmo para a identificação da variável de valor mínimo.

Essas funções e esses recursos das planilhas eletrônicas permitem a manipulação rápida de uma série de informações, números e dados que facilitam o trabalho dos analistas e pesquisadores. Para além desses recursos, é possível ainda espacializar (organizando a formatação de mapas) essas informações. Isso será abordado de forma introdutória na próxima unidade.

Além de calcular e permitir a análise das médias, máximos e mínimos de uma base, a função classificar da planilha eletrônica permite que os valores sejam organizados do maior para o menor ou do menor para o maior a partir da identificação da variável de referência. É possível também colocar as variáveis categóricas ou de texto em ordem alfabética (como, por exemplo, no caso dos municípios).

No caso de classificação das variáveis, é preciso algum cuidado com a determinação da variável que realizará a classificação. Dessa forma, é preciso selecionar toda a base de dados e indicar qual coluna deve ser a norma — a forma de classificação — ou os dados serão misturados. Isso significa que a base de dados deve ser colocada em ordem no seu intervalo completo, com todas as linhas e colunas, ou teremos vários dados embaralhados. Cuidado! No caso de isso acontecer, sempre é possível clicar na seta desfazer ao lado do ícone de salvamento das alterações no canto esquerdo da tela.

Nessa mesma barra de instrumentos da classificação de variáveis, diversos outros elementos facilitam a análise de dados para os pesquisadores. Estão disponíveis ferramentas de teste de hipóteses e de outros tipos de análise de dados. Nessa análise de dados mais completa, que está disponível para a instalação no *Excel*, podem-se fazer, por exemplo, cálculos de correlação entre variáveis, geração de números aleatórios, a apresentação de toda a estatística descritiva de um conjunto de dados, análise de covariância, análises de regressão, etc.

4.2 ANÁLISE DA MATRIZ DE CORRELAÇÃO E A COMPOSIÇÃO DE INDICADORES E ÍNDICES

Em determinadas situações, pode ser útil ao analista identificar a relação entre duas ou mais variáveis em uma base de dados extensa. Nesse caso, um dos instrumentos oferecidos pelas planilhas eletrônicas é a matriz de correlação. A relação de correlação indica como ocorre a variação entre um par de variáveis de análise. Considerando o conjunto dos dados, a planilha gera uma matriz em que cada variável é analisada em função de outra.

A análise de correlação permite identificar como as variáveis variam juntas numa base de dados. Apenas para o entendimento desses resultados, convém salientar que, nesse tipo de análise, o resultado numérico varia de -1 a 1. Valores próximos de -1 ou 1 indicam que as variáveis do par apresentam forte correlação, ou seja, se alteram de forma inversa (-1)¹ ou de forma conjunta (1)² ao longo da base: quanto mais próximo de 0, mais fraca a correlação. Na base a seguir, pode-se identificar que as variáveis 10 e 23 possuem uma correlação igual a -0,28, com uma integração maior do que aquelas que apresentam valores menores do que estes. O sinal do resultado indica que, quando uma variável aumenta, a outra diminui, e vice-versa. Além da identificação da dinâmica entre as variáveis, é preciso que o analista interprete os números relacionados aos resultados obtidos, verificando aquilo que era esperado e buscando explicações para as relações que contradizem o senso comum. A Figura 9 apresenta o exemplo da matriz de correlação organizada para este material.

1 Quando uma cresce, a outra imediatamente diminui na mesma proporção.

2 Quando uma cresce, a outra cresce também; e quando uma diminui a outra também diminui.

mente composto, com sentido e objetivo. Existem testes para verificar se ele pode ser utilizado, sendo considerado nesse caso como estatisticamente significativo. Um exemplo de indicador é a quantidade de médicos por mil habitantes em um município; a unidade de medida do indicador decorre das variáveis que o compõem diretamente.

O índice apresenta um nível mais complexo de agregação; é formado a partir da reunião de vários indicadores, que são representados por um número apenas. É uma unidade de medida em si. É o que ocorre com o Índice de Desenvolvimento Humano, por exemplo (já discutido neste volume). O resultado do índice permite apenas identificar a relação entre seus componentes e não uma unidade; a comparabilidade decorre do cálculo do mesmo índice (com o mesmo tipo de variáveis) para situações diversas. As variáveis no índice só fazem sentido quando olhadas de forma agregada e não separadamente. A vantagem é o fato de ter a possibilidade de ponderar e analisar variáveis em distintas unidades de contagem em distintos momentos históricos.

4.3 ALGUNS ELEMENTOS GERAIS

A possibilidade da inserção de fórmulas e a organização dos elementos combinados, como verificamos, ajudam e facilitam grandemente o trabalho do analista para o estabelecimento das relações apresentadas por um conjunto de dados. Entretanto, é sempre importante ponderar que é preciso que os resultados tenham repercussão na realidade. Algumas análises só fazem sentido com determinadas funções estatísticas. Por exemplo, faz sentido comentar sobre a média de renda *per capita* de um município ou estado, mas será que faz sentido comentar um somatório das rendas? Tudo depende daquilo que o arcabouço teórico quer apresentar ao leitor e da situação que está representada por aquele valor, seja ele uma média ou um índice. Algumas variáveis com alto valor podem representar uma sociedade com uma boa qualidade de vida (por exemplo, expectativas de anos de vida). Por outro lado, outras variáveis com valor alto podem representar exatamente o contrário (por exemplo, analfabetismo). Assim, a análise e a leitura do trabalho exigem atenção quanto a essa construção.

Questões para a reflexão:

- 1) Verifique, na tabela configurada na unidade anterior, quais são as médias, moda, medianas, máximos e mínimos para cada uma das variáveis. Qual dos municípios tem a maior população? E o maior percentual da população vivendo em área rural?
- 2) Baixe, na base FEEDADOS, as variáveis PIB *per capita* e expectativa de vida ao nascer dos municípios gaúchos. Os municípios com valores mínimos e máximos entre as variáveis são os mesmos? Como você explicaria essa diferença?

APRESENTANDO OS DADOS ORGANIZADOS

A possibilidade da busca e análise dos dados permite que uma série de informações seja disponibilizada e reconhecida. A pergunta que tínhamos no início do trabalho agora, após a pesquisa, tem uma resposta. Sabemos qual a dinâmica do fenômeno analisado e agora é preciso divulgar, aos interessados, os resultados obtidos sobre determinados temas. Dessa forma, o exercício a que se dedica esta unidade é o de apresentar aos leitores as informações, seja através da escrita, das figuras ou de uma apresentação oral.

As figuras são elementos importantes da apresentação e divulgação de trabalhos, pois permitem aos leitores e à audiência uma maior aproximação dos temas em discussão. Pense na sua realidade: é mais fácil ou mais difícil se lembrar de informações que são acompanhadas, por exemplo, de uma imagem? É possível imaginar que exista diferença no nível de atenção e até na lembrança da fala ou da escrita se há a presença de alguma figura reforçando o que é dito ou escrito? Tendo um elemento gráfico, não fica mais fácil lembrar e realizar as relações necessárias para a interpretação? No meu caso, e no de muitas pessoas, a resposta é sim. Por isso, é importante, sempre que possível, sintetizar e traduzir as informações em recursos gráficos.

Esta última seção tem como objetivo apresentar elementos que ajudam na concepção e na execução de figuras que facilitam o entendimento e a análise dos dados reunidos. Essas informações passam a ser organizadas de forma prática e rápida, comunicando aos leitores os resultados da nossa pesquisa. É preciso estar atento ao fato de que a figura precisa sempre melhorar o entendimento do contexto em análise, devendo ser capaz de resumir o que foi discutido anteriormente em forma de textos ou de alguma apresentação oral.

Assim como existem formas diferentes de organizar os dados, como existem interpretações mais ou menos adequadas em relação à utilização das fórmulas e dos indicadores, pode-se considerar também que existem figuras adequadas a cada tipo de informação. Por exemplo, há diferença, no que diz respeito às fontes dos dados, para a configuração de quadros ou de tabelas.

Esta unidade trará, além desta breve introdução, a apresentação de uma série de possibilidades para a apresentação das informações, todas disponíveis nas planilhas eletrônicas, onde organizamos os dados com a orientação da unidade anterior. As figuras mais comumente utilizadas são quadro, tabela, gráfico e figuras. Além dos elementos gráficos, serão comentados aspectos referentes à apresentação escrita do trabalho e a alguns cuidados da apresentação oral.

5.1 ELEMENTOS GRÁFICOS MAIS UTILIZADOS

Considerando essa função mais geral de facilitar a interpretação dos dados e informações que respondem ao nosso problema de pesquisa, pode-se imaginar objetivos específicos para cada um dos elementos visuais que se pretende utilizar em determinado texto. Sempre que formos montar alguma figura, é bom ter em mente qual seu objetivo específico. Ela pode estar ali para comparar informações em relação a algum elemento; pode ter a função de indicar algum fluxo; pode estar tentando estabelecer alguma hierarquia ou proporção. Podemos ter figuras que indicam a localização, que apresentam uma cronologia de determinado fato. Ao longo do trabalho, as figuras devem ser numeradas em ordem crescente, conforme vão aparecendo.

Se estamos analisando, por exemplo, processos produtivos, as figuras podem ser utilizadas para mapear e identificar esses processos. Podemos utilizar algumas figuras com setas para indicar as relações e a causalidade dos fenômenos, entre tantos outros objetivos. O importante é identificar o objetivo para, após a configuração do elemento gráfico, podermos verificar se esse objetivo foi realmente alcançado.

Em trabalhos com um grande volume de texto e informações (como monografias, dissertações e teses) é exigida a configuração de uma lista de figuras no início do trabalho.

5.1.1 QUADROS

Os quadros reúnem informações geralmente mais qualitativas. Organizados pelo autor do trabalho, sempre devem apresentar a fonte das informações. Em relação à forma, pode-se destacar que o quadro possui os lados delimitados por uma borda contínua.

No quadro, é possível reunir informações tanto de trabalhos de outros autores como do próprio autor do trabalho, baseado em outros autores. Normalmente, esse tipo de figura tem como função apresentar algum tipo de panorama geral que precisa ser comparado. Com os quadros é possível classificar, ordenar e hierarquizar as informações que foram analisadas anteriormente. Neles, podem ser apresentados tanto dados qualitativos quanto dados quantitativos.

A Figura 10 apresenta um exemplo de quadro; note que ele ajuda o leitor a separar as funções de cada ente da federação em relação a um programa específico.

Figura 10 — Exemplo de quadro — classificação¹

Quadro 1 — Distribuições de atribuições em relação ao PBF nas três esferas de governo.

Governo Federal	Governo Estadual	Governo Municipal
Disponibiliza o Cadastro Único	Promove o registro documental em seu território	Identifica as famílias socialmente vulneráveis
Transfere o benefício financeiro para as famílias	Fomenta a intersetorialidade	Procede ao cadastramento das famílias
Coordena e normatiza a operacionalização do PBF	Acompanha indicadores, apoiando a Gestão Municipal	Executa a gestão do benefício

Fonte: Coordenação-Geral de Apoio à Gestão Descentralizada — Departamento de Operação/SENARC/MDS (2012).

¹ KÜHN; TONETTO, 2017, p. 91.

5.1.2 TABELAS

As tabelas são um dos recursos gráficos mais utilizados. Assim como os quadros, dispõem dados e informações em linhas e colunas, permitindo determinado nível de agregação e síntese de informações mais abrangentes. Na maioria das vezes, a tabela apresenta dados quantitativos que são resultado da pesquisa e da análise realizadas pelo autor do trabalho.

Diferentemente dos quadros, as tabelas apresentam os lados direito e esquerdo sem uma delimitação (são “abertas”), mantendo delimitações em cima e embaixo. Na parte inferior da figura, é sempre necessário apresentar a origem da fonte dos dados trabalhados. São elementos gráficos bastante comuns na apresentação dos resultados das pesquisas.

A disposição das informações entre linhas e colunas deve permitir ao leitor a interpretação do dado indicado pelas coordenadas. Na Figura 11, por exemplo, é possível identificar (1) que a maior privação está relacionada às condições sanitárias, (2) que houve retrocesso em relação à alfabetização e (3) que o maior avanço se refere à renda.

Figura 11 — Exemplo de tabela²

Tabela xx — Percentual de privações nos indicadores de pobreza em 2000 e 2010

PRIVAÇÕES	2000	2010	DIF %
Alfabetização	1,46%	2,05%	-0,59
Anos de estudo	47,58%	40,30%	7,28
Condição de ocupação do imóvel	18,60%	18,58%	0,02
Energia elétrica	1,47%	0,22%	1,25
Itens domiciliares	2,05%	0,43%	1,62
Abastecimento de água	24,67%	20,24%	4,43
Tipo de escoadouro das instalações sanitárias	75,04%	65,38%	9,66
Destino do lixo	20,01%	13,62%	6,39
Número de banheiros no domicílio	6,42%	1,39%	5,03
Renda domiciliar	16,53%	8,63%	7,90
Emprego	33,31%	29,43%	3,88

Fonte: Elaborado pela autora, baseado nos microdados do IBGE.

É preciso atentar que as informações devem sempre ser precedidas de uma análise do autor, servindo a figura para reforço ou reorganização de informações apresentadas de outra forma, oral ou escrita.

5.1.3 GRÁFICOS

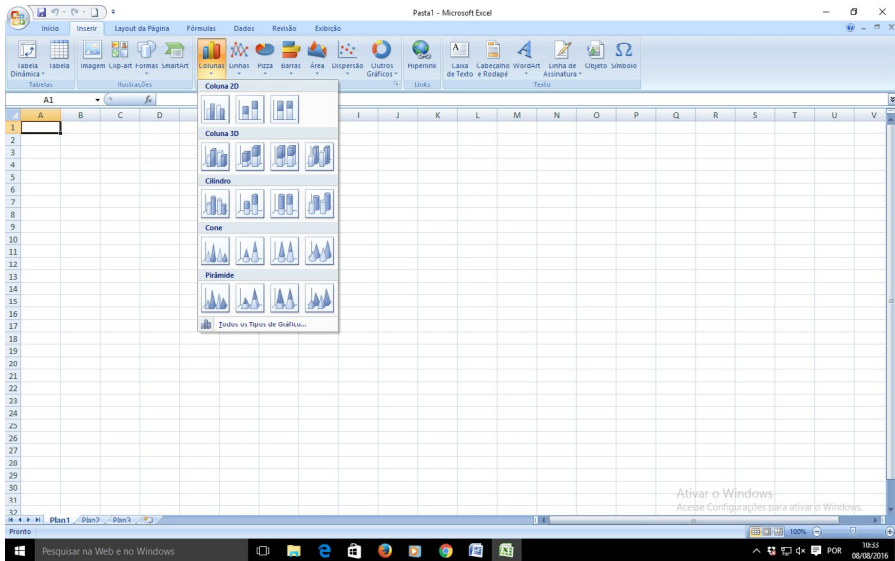
Os gráficos permitem uma visualização instantânea das relações apresentadas; por exemplo: quem é maior, quem é menor e quem cresce mais ou menos. As informações quantitativas são transformadas em informação visual que permite uma interpretação mais “amigável” das relações. Entretanto, é preciso estar atento ao que se quer apresentar para evitar que o gráfico dificulte ainda mais, ao invés de auxiliar, na leitura das informações. As planilhas eletrônicas, bem como os *softwares* estatísticos, apresentam a funcionalidade no seu pacote

² MARIN; VIEIRA; KÜHN, 2017.

de funcionalidades. Normalmente, o próprio ícone do programa que leva à formatação do gráfico está representado por um gráfico de barras.

Como exemplo, a Figura 12 apresenta como a ferramenta que permite a configuração de Gráficos aparece no *software* de planilha eletrônica *Excel*. Na aba “Inserir”, inicia-se o detalhamento para a configuração de figuras gráficas a partir de uma base de dados já existente.

**Figura 12 — Exemplo de acesso aos tipos de gráficos
— Gráfico para o *Excel***



Fonte: Elaborado pela autora.

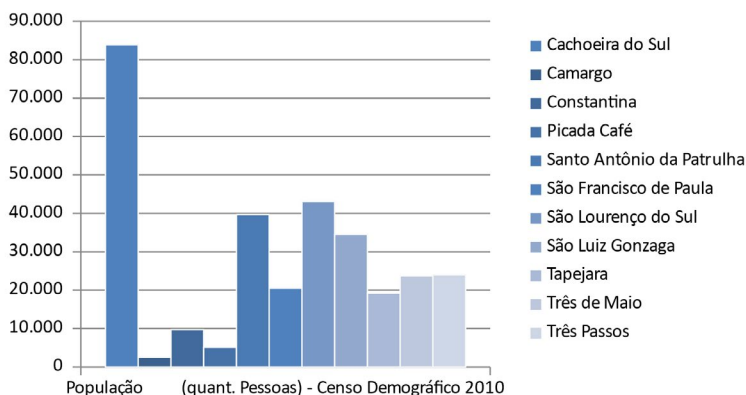
Observe que as versões atuais do *software* já separam na própria barra os tipos de gráficos. Ao clicar em cada um dos tipos, é permitida a visualização das opções em relação à formatação do gráfico. A Figura 13 apresenta esse exemplo para o gráfico de colunas. A formatação do gráfico se dá de modo bem intuitivo, sendo possível testar vários tipos de gráficos, alterando cores, formato de fontes e legendas. É sempre importante ter em mente o objetivo do gráfico e verificar se, de fato, o formato escolhido facilita a interpretação dos dados e das informações disponíveis.

A partir da organização das tabelas nas planilhas eletrônicas, o gráfico se configura automaticamente, necessitando, na maioria dos casos, apenas de um posterior ajuste de forma. Para a familiarização com o instrumento, busque uma base de dados simples e procure fazer os mais diferentes tipos de gráficos.

Normalmente, a realização dos gráficos nas planilhas eletrônicas permite que o desenho seja colocado em nova aba da planilha ou ao lado da própria tabela. Cada um deve testar as duas formas e escolher aquela que melhor se adapta a sua forma de trabalho. Uma vez organizada a base e gerado o gráfico, pode-se realizar uma série de modificações que envolvem a formatação da fonte (tamanho, tipo, cor, etc.), a especificação de linhas de gráfico, especificação das cores que devem ser utilizadas no gráfico, introdução de uma linha de tendência para o conjunto de dados em análise. Pode-se configurar também o local e o formato da legenda que será utilizada. Em relação aos gráficos, é sempre importante lembrar-se de inserir a indicação da representação dos eixos, sob pena de o leitor ou de quem está assistindo à apresentação não compreender a relação que está sendo estabelecida (perdendo assim a principal função da organização do gráfico). Na figura gerada, com um clique no botão direito do *mouse* sobre os eixos, edita-se esse elemento. Existem diversos modelos de gráficos, mas as características de cada um permitem identificar aquele mais adequado ao objetivo que se pretende com a visualização dos dados.

O gráfico de colunas apresentado permite a comparação entre itens. Podemos verificar o comportamento de determinada variável em condições diferentes (como, no caso exemplificado, a população total em diversos municípios). O gráfico de colunas apresenta um recurso 3D que permite a inserção de outro plano de análise (por exemplo, mais uma variável ao contexto). Os marcadores em 3D podem ser utilizados em forma de cilindro ou de cone.

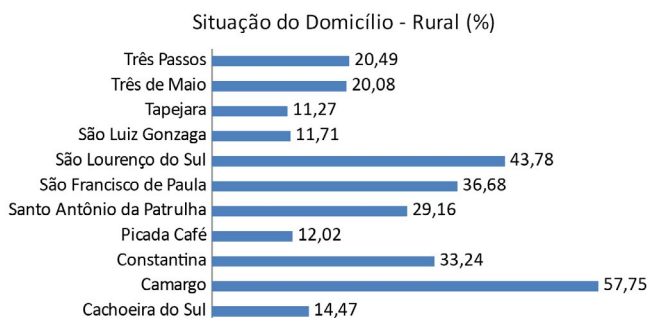
Figura 13 — Exemplo: gráfico de colunas



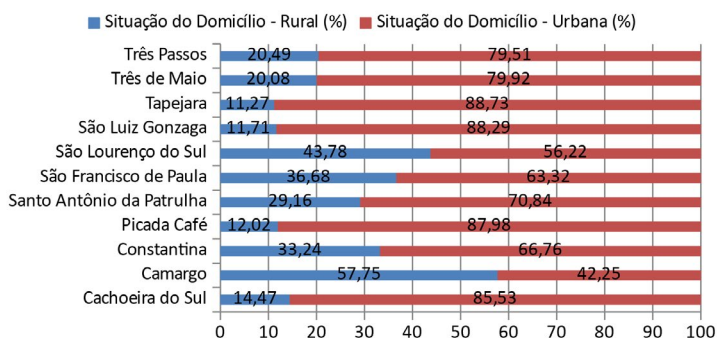
Fonte: IBGE, Censo 2010 Elaborado pela autora.

O gráfico de barras é análogo ao gráfico de colunas. Cumpre as funções apresentando um elemento gráfico que identifica os movimentos da variável na horizontal. Permite a comparação de elementos em relação a determinada variável, bem como permite o estabelecimento e a apresentação de padrões em relação à série analisada. Também é um dos tipos de gráficos bastante utilizados, assim como o de colunas. O gráfico de barras permite também a configuração empilhada, em que é possível apresentar uma comparação em relação aos totais e às barras comparadas. A Figura 14 (a, b, c) apresenta exemplos de gráfico de barras.

Figura 14 (a) — Exemplo: gráfico de barras simples

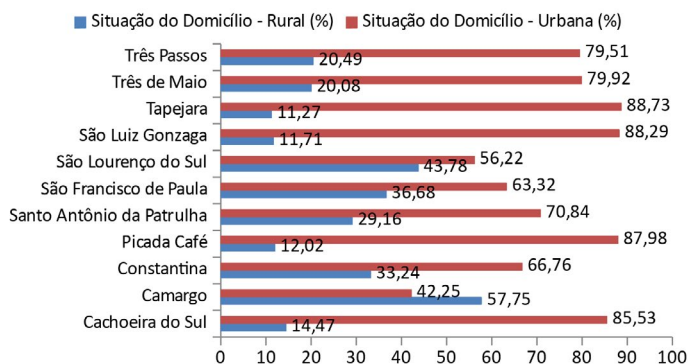


Fonte: IBGE, Censo 2010 Elaborado pela autora.



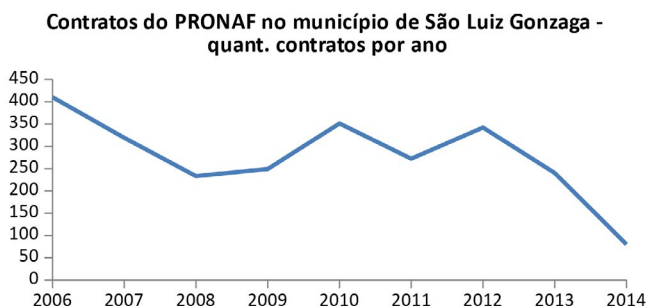
Fonte: IBGE, Censo 2010 Elaborado pela autora.

Figura 14 (c) — Exemplo: gráfico de barras comparadas



Fonte: IBGE, Censo 2010 Elaborado pela autora.

O gráfico de linhas é utilizado para mostrar evoluções ou tendências dos dados para intervalos e categorias comparáveis ou ao longo do tempo. É preciso identificar a análise a ser feita e os dados que serão utilizados, uma vez que corremos o risco de ter um emaranhado de linhas que não ajudam na interpretação do contexto que se quer apresentar. A Figura 15 apresenta um exemplo de gráfico de linhas.

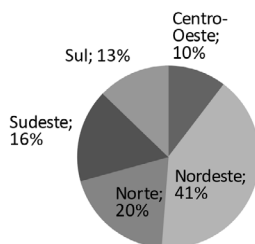
Figura 15 — Exemplo: gráfico de linhas

Fonte: BACEN. Elaborado pela autora.

Um dos tipos mais comuns e de mais fácil interpretação, quando bem utilizado, é o gráfico do tipo pizza. Ele serve para apresentar o tamanho proporcional de determinado atributo em relação a um determinado total. Refere-se sempre à distribuição de uma variável em determinado momento do tempo. Não apresenta a evolução dos dados, mas representa muito bem uma fotografia das situações. A figura pode apresentar os percentuais escritos na própria “pizza”, o que prescinde da utilização de uma nova figura de tabela, por exemplo. A Figura 16 apresenta um exemplo de gráfico de pizza.

Figura 16 — Exemplo: gráfico de pizza

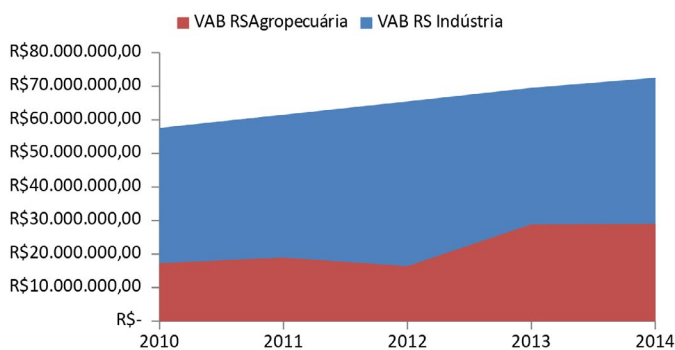
Quantidade de Territórios Rurais Brasileiros por macrorregião



Fonte: Elaborado pela autora com base em dados do Ministério do Desenvolvimento Agrário.

O gráfico de área torna possível a análise de mudanças, ao longo do tempo, de um conjunto de variáveis que podem ser comparadas. As observações são apresentadas de forma sobreposta e permitem a comparação de variáveis ao longo do tempo. É importante destacar que a utilização dessa representação não é adequada quando na base de dados existem números negativos, o que dificulta a interpretação. No caso de valores negativos, talvez o mais indicado seja usar o gráfico de linhas. A Figura 17 apresenta o gráfico de área.

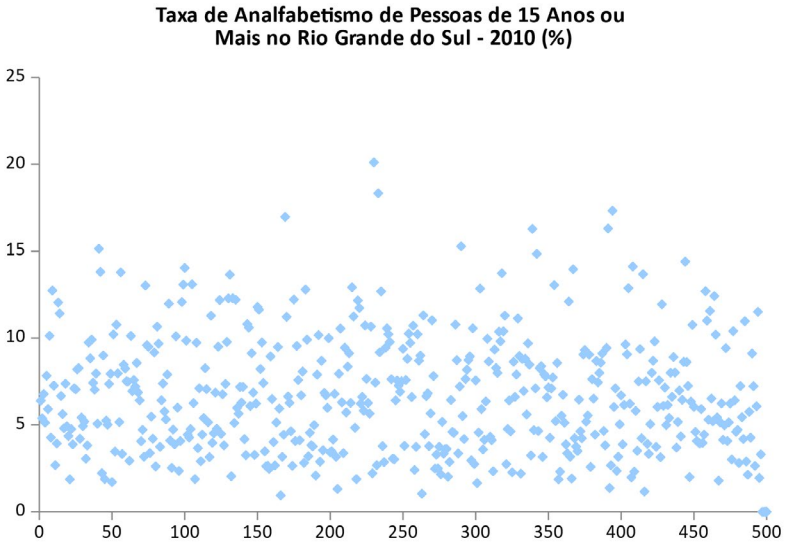
Figura 17 — Exemplo: gráfico de área



Fonte: FEEDADOS Elaborado pela autora.

O gráfico de dispersão serve para a identificação de pontos de relação entre variáveis no espaço. A partir da análise desses pontos, é possível perceber a tendência de comportamento conjunto. A série de dados é analisada conjuntamente. A Figura 18 apresenta um exemplo do gráfico de dispersão.

Figura 18 — Exemplo: gráfico de dispersão



Fonte: FEEDADOS Elaborado pela autora.

O gráfico de ações é formatado a partir de uma variável em que é possível identificar mínimos e máximos em unidades comparáveis. A principal utilização é na comparação diária de preços mínimos e máximos de ações no mercado financeiro. A Figura 19 representa um gráfico de ações.

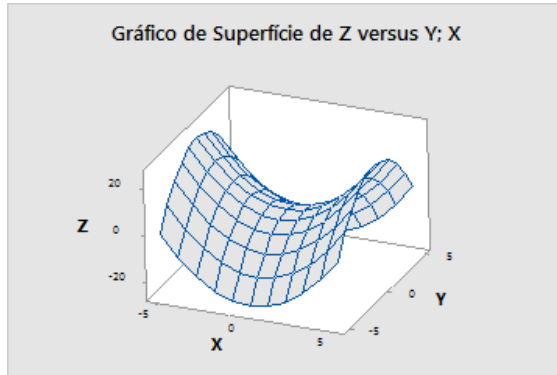
Figura 19 — Exemplo: gráfico de ações



Fonte: CRUDDEN research group, dados de 2012.

O gráfico de superfície permite a combinação tridimensional de variáveis, apresentando um conjunto de vetores (x,y,z) a partir de valores identificados em tabelas. A Figura 20 exemplifica a configuração de um gráfico de superfície.

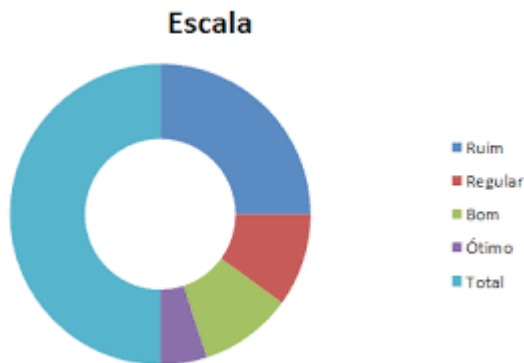
Figura 20 — Exemplo: gráfico de superfície



Fonte: Minitab 18.

O gráfico de rosca permite a avaliação da frequência de um atributo em relação ao total de observações. A Figura 21 representa o exemplo de um gráfico de rosca.

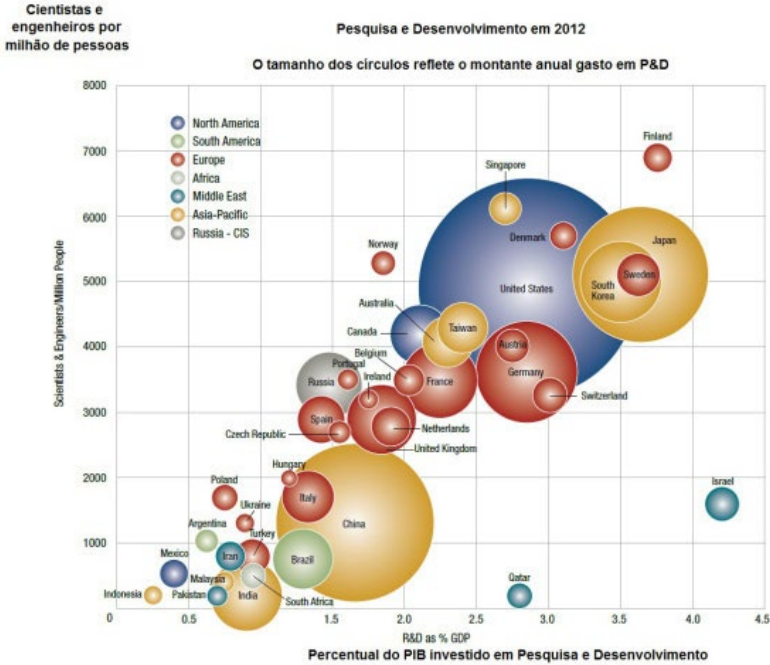
Figura 21 — Exemplo: gráfico de rosca



Fonte: ONTRAINING, 2012.

O gráfico de bolhas é um tipo de gráfico de dispersão capaz de apresentar a intensidade de relacionamento das variáveis com um terceiro elemento identificado. A Figura 22 apresenta um exemplo de gráfico de bolhas.

Figura 22 — Exemplo: gráfico de bolhas



Fonte: Segredo da estatística, 2014.

O biograma, gráfico de radar ou estrela, é outra forma de representação de variáveis. Essa forma gráfica de reunião de informações é bastante utilizada para a representação de índices. Permite a agregação de valores a partir de várias séries de dados. Dessa forma, viabiliza, além de uma análise mais ampla e geral, a comparação entre cada dimensão identificada. Os índices são elementos estatísticos que permitem a percepção de um conjunto de variáveis em um só indicador. Um exemplo famoso de índice é o Índice de Desenvolvimento Humano (IDH). Esse índice reúne aspectos como saúde, educação e renda. Nesse caso, cada uma das três dimensões em análise é composta por mais de uma variável, resultando, de acordo com a ponderação indicada, em um novo número que classifica a realidade analisada. Existem diversos índices que ajudam os pesquisadores a interpretar a realidade social a partir de uma interação

mais complexa. Considerando a reunião das dimensões, os gráficos mais comumente utilizados não permitem essa representação gráfica mais complexa; nesse sentido, é utilizado então o biograma, que permite essa visualização conforme a figura a seguir. O exemplo apresentado aqui foi retirado do relatório de análise do Índice de Condições de Vida (ICV) calculado para os Territórios Rurais brasileiros.

Figura 23 – Exemplo de Biograma (radar) – ICV



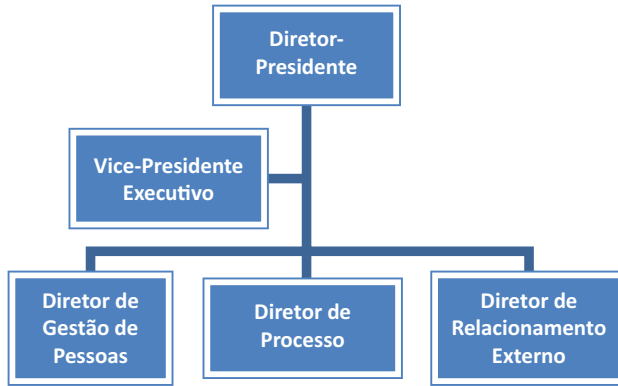
5.1.4 ORGANOGRAMAS E FLUXOGRAMAS

Dois elementos importantes que podem ser utilizados tanto para a visualização de concepções apresentadas ao longo de textos como para organizar ideias para as apresentações visuais são os organogramas e os fluxogramas. Ambos estão disponíveis nos editores de textos mais comumente utilizados e nos programas que ajudam na preparação de apresentações visuais — por meio de slides, por exemplo. Quando na barra de ferramentas do *software* for possível inserir algum desenho, certamente haverá a opção para a composição desses dois tipos de figuras.

O organograma permite a apresentação de relações e encadeamentos, diretos e indiretos. É um elemento gráfico importante para a apresentação contextual, assim como os gráficos. Através dele podem ser apresentados elemen-

tos como estruturas de empresas e cooperativas, árvores genealógicas, entre outros elementos. A Figura 24 apresenta um exemplo simples de organograma.

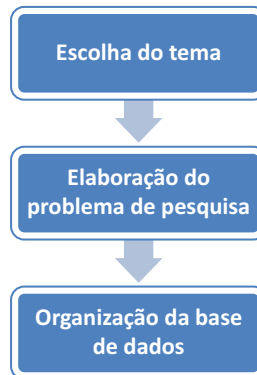
Figura 24 — Exemplo de organograma



Fonte: Elaborado pela autora.

O fluxograma permite o estabelecimento de relações de causa e consequência, bem como apresenta, de forma gráfica, a sequência de um processo e suas alternativas de desdobramento. Qualquer processo com começo e fim pode ser representado a partir de um fluxograma. A Figura 25 apresenta um exemplo de fluxograma.

Figura 25 — Exemplo de fluxograma



Fonte: Elaborado pela autora.

Figura 27 — Exemplo de mapa II



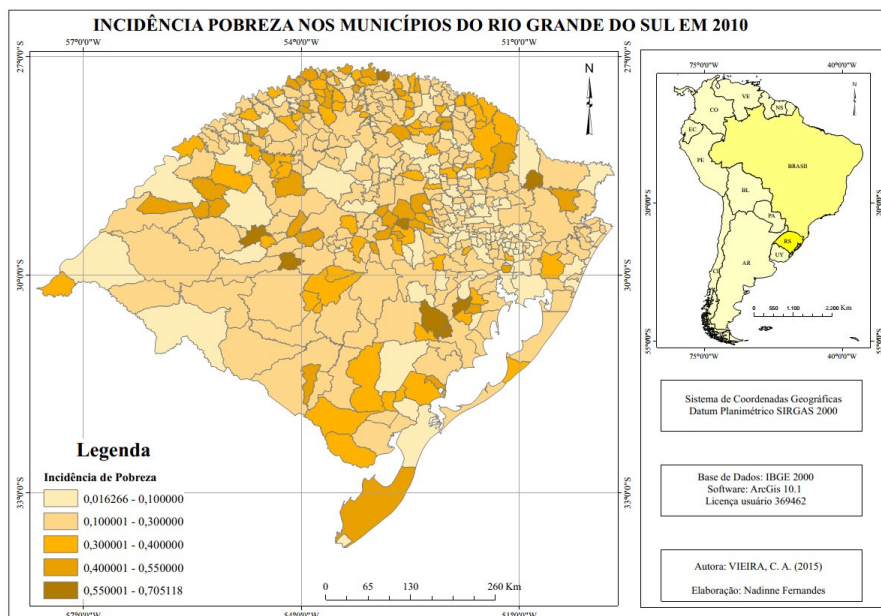
Fonte: Wikipedia.

Pode-se perceber que o primeiro mapa não evidencia o contexto do município no âmbito mais amplo da escala territorial, enquanto que no segundo isso é permitido. Se a função do mapa é localizar o leitor que não conhece o município (que talvez nem chegue a visitá-lo algum dia), o segundo mapa é o mais adequado. O mapa do *Google Maps* normalmente é adequado não para a identificação de um contexto mais amplo, mas para aquela pessoa que está ou vai até o município e precisará localizar um lugar específico ou realizar um deslocamento dentro do município. Veja que, claramente, os dois mapas têm funções específicas e a sua utilização nos trabalhos exige que o pesquisador ou aluno saiba o objetivo da figura no contexto em que vai se inserir. Perceba que é mais fácil compreender a localização do município no segundo mapa. Os dois mapas tratam do mesmo município; você sabe qual é?³

Outra questão importante em relação aos mapas, assim como nos gráficos, é a atenção à necessidade ou não de cores na legenda, bem como a percepção de ter uma legenda clara daquilo que se quer evidenciar com a apresentação do mapa. A Figura 28 apresenta um exemplo de mapa configurado adequadamente que permite a sua interpretação visual.

³ Trata-se do município de Três Passos, no Rio Grande do Sul.

Figura 28 — Exemplo de mapa de intensidade para variável específica Incidência de Pobreza em 2010



Fonte: IBGE Elaborado para Marin, Vieira e Kühn (2017).

Algumas das bases de dados já comentadas permitem a organização de mapas a partir da seleção de agregados territoriais e da seleção das variáveis de interesse. A FEE e o IPEA são algumas dessas instituições.

5.2 OUTROS ELEMENTOS IMPORTANTES NA APRESENTAÇÃO DE TRABALHOS DE PESQUISA

5.2.1 A ESCRITA

Além da organização dos elementos gráficos, é muito importante, e nunca é demais, atentar à organização da escrita dos trabalhos. Esta breve seção busca chamar a atenção para alguns elementos que, se ignorados, por vezes, acabam prejudicando o reconhecimento do esforço de produção intelectual pelo leitor. O potencial do trabalho de pesquisa pode ser prejudicado tanto por uma apresentação oral ruim quanto por uma apresentação escrita confusa.

Primeiramente, convém evidenciar que não é papel do autor do trabalho supor que o leitor já tenha algum conhecimento sobre o tema. A condução do leitor pelo caminho do tema que está sendo tratado, normalmente, exige que se passe da percepção teórica utilizada ao objeto específico em estudo, passando pela explicação da metodologia de análise a ser realizada e pela composição das variáveis, dos indicadores e dos índices que forem utilizados.

Listemos outras considerações importantes e que podem ajudar na organização das ideias: por uma questão de clareza, é importante evitar a utilização de gírias; da mesma forma, devem-se evitar frases longas demais (em uma frase muito longa a ideia pode ficar confusa e o leitor poderá não entender; podemos imaginar que cada frase deve ter no máximo um verbo principal e jamais será capaz de abrigar mais de uma referência); evitar a repetição da mesma palavra na mesma frase ou no mesmo parágrafo (geralmente, uma segunda leitura do material pronto evidencia claramente essas repetições); evitar os pleonasmos;⁴ evitar a utilização de adjetivos ao longo do texto (os adjetivos — feio/bonito; errado/certo — representam um julgamento de valor que não deve permear a realização do trabalho).

Uma atitude importante em relação à organização escrita de um trabalho de pesquisa é sempre reler o que foi escrito. Essa nova leitura permite sempre a identificação de algo que precisa ser corrigido ou que pode ser melhorado. Dessa forma, é fundamental que se evite deixar a realização do trabalho para os últimos dias, buscando sempre organizar, após o término da primeira versão, um período para a releitura antes do encaminhamento do trabalho final.

5.2.2 A APRESENTAÇÃO

Além da organização de um trabalho escrito, não raro, há estímulos (se não obrigação, como no caso dos trabalhos de final de graduação) para a apresentação para um grupo da pesquisa desenvolvida. Esse é um tema complexo, pois a diversidade humana faz com que ocorram diferenciados comportamentos em relação a essa etapa de finalização da pesquisa. Isso é evidente pelo fato de que algumas pessoas têm facilidade na organização e na apresentação

⁴ Alguns pleonasmos famosos: subir para cima; elo de ligação; a última versão definitiva; gritar bem alto; exceder em muito; criação nova; amanhecer o dia; de sua livre escolha; certeza absoluta; em duas metades iguais; há anos atrás; vereador da cidade; surpresa inesperada; escolha opcional; planejar antecipadamente, etc.

desses momentos e que outras têm extrema dificuldade. De qualquer forma, é importante que o pesquisador consiga conversar e sustentar uma apresentação oral sobre seu objeto de pesquisa. Obviamente, as pessoas têm rituais e formas diferentes de se organizar; esse breves comentários destinam-se àqueles alunos que ainda não estão familiarizados com essa atividade.

Em linhas gerais, pode-se salientar que a apresentação de um trabalho ou resultado de pesquisa é uma etapa importante da formação profissional e acadêmica. Como há um costume de que o processo de formação se dê com alguém falando e outros ouvindo, a passagem de ouvinte a orador pode ser um pouco desconfortável. Algumas pessoas acham importante realizar “treinos” de apresentação. Para fazer uma tentativa como essa, convide alguém da sua confiança e que não saiba sobre o tema. Dessa forma, a pessoa poderá questionar os aspectos que ficaram confusos ao longo da apresentação.

Tenha o cuidado de não falar nem muito rápido (o que pode levar à confusão de termos), nem muito devagar (o que pode levar à desatenção dos ouvintes). Assim como na escrita, tente evitar vícios de linguagem e gírias. Quanto ao visual, tome o cuidado para manter a postura adequada; muitas vezes quem não fica confortável na condição de orador tende a permanecer em um canto da sala como se estivesse querendo fugir dali (mesmo que esteja, mantenha-se no centro da sala e lembre-se sempre de que você realizou o trabalho: ninguém sabe mais sobre ele do que você).

Apresente figuras (quadro, tabelas e mapas) que facilitem o entendimento e a visualização das relações que você considerou mais importantes e significativas nessa pesquisa ou tema. Entretanto, é preciso tomar cuidado com apresentações em que o instrumental audiovisual acaba servindo apenas para a apresentação de textos. Utilizando uma apresentação de slides, cuide com a cor dos fundos e das fontes; pondere sobre a necessidade de animação na passagem dos slides (bem como a utilização de recursos sonoros). Lembre-se: quem ouve está interessado no tema que está sendo apresentado e não exatamente na forma de apresentação.

Convém atentar ainda que é sempre conveniente, na medida do possível, conhecer o lugar antes da apresentação. Assim, você poderá ficar mais seguro quanto ao manuseio da aparelhagem audiovisual, conhecendo seus recursos e limitações. De qualquer forma, sempre pense em uma forma alternativa de apresentação, pois imprevistos podem acontecer e nem por isso você deve deixar de apresentar seu trabalho/sua pesquisa.

Questões para refletir:

- 1) Considere a tabela a seguir:

TABELA — População Brasileira — Local da residência

Residência	População
Rural	29.830.007
Urbana	160.925.792

Fonte: IBGE, Censo 2010.

- a) Qual a menor forma de representar as informações graficamente? Gráfico de coluna, de linha, pizza? Por quê? O que poderíamos dizer sobre essas informações (Qual a população total? Qual o percentual de moradores na área rural?)
- 2) Considere as seguintes informações sobre o Produto Interno Bruto (PIB), obtidas no site do IBGE:

PIB brasileiro em 2010 — R\$ 3.885.847;

PIB brasileiro em 2011 — R\$ 4.376.382;

PIB brasileiro em 2012 — R\$ 4.814.760;

PIB brasileiro em 2013 — R\$ 5.331.619;

PIB brasileiro em 2014 — R\$ 5.778.953;

PIB brasileiro em 2015 — R\$ 6.000.570;

PIB brasileiro em 2016 — R\$ 6.266.895.

Pense em que tipo de informação é necessário para escrever algo que transforme esses dados em conhecimento. Pode-se dizer que o PIB cresceu apenas utilizando essa variável? Qual seria o tipo de figura mais adequado para representar graficamente o que ocorreu com o PIB brasileiro? Faz sentido montar um gráfico de pizza para representar essa informação? Por quê?

- 3) Releia algum texto seu. O que você acha que pode ser melhorado? Como as pessoas entenderiam melhor aquilo que você está querendo dizer?
- 4) Agora imagine que esse texto será apresentado. Como você faria essa apresentação? De quanto tempo precisaria? Que recursos você utilizaria? É possível pensar em algo além da apresentação de slides?

O trabalho de pesquisa envolve a busca por dados e informações. Num mundo em que o acesso às informações não é mais uma barreira para grande parte da população, resta-nos a tentativa de organizar e separar, dentro de um grande “mar” de opções, elementos que de fato possam nos ajudar e contribuir para a nossa formação acadêmica e para a compreensão das relações sociais. O objetivo deste breve material foi o de reunir alguns elementos que possam ajudar tanto na seleção das informações quanto na organização desses dados para que os objetivos de pesquisa de cada um possam ser alcançados.

Certamente esse não é um tema restrito. Pelo contrário, é amplo e a cada momento pode-se dizer que a quantidade de informações disponível aumenta, de modo que, muitas vezes, o essencial é ter claro o objetivo da pesquisa para sair em busca de dados e informações que possam ajudar nessa compreensão. A evolução e as transformações do acesso aos meios de informação e comunicação podem fazer com que os exemplos aqui apresentados possam estar em breve superados. Entretanto, espera-se que a motivação para a organização das informações possa estar presente em cada trabalho a ser realizado após a leitura desse material (e, principalmente, o convencimento da necessidade dessa organização).

Uma base de dados bem organizada permite a continuidade e a retomada de trabalhos pelo mesmo pesquisador ou por outros interessados. Considerando a Ciência como um ambiente cumulativo de conhecimento, essa organização torna-se algo fundamental.

Por outro lado, espera-se também que o processo interpretativo de dados possa ter ficado mais próximo daqueles interessados em iniciar um processo de pesquisa e que este tenha sido – quem sabe – o primeiro de muitos outros textos sobre etapas e procedimentos para a realização de pesquisa. Que seja um dos primeiros passos!

Boa sorte e bom trabalho a todos nós que nos aventuramos na tentativa de fazer Ciência, compreendendo e revelando mais e mais das relações que mantêm nossas trajetórias na sociedade em que nos inserimos.

- ALVES, Rubem. **Filosofia da Ciência**. São Paulo: Brasiliense, 1981.
- BACHELARD, Gaston. **A filosofia do não**. In: _____. **A filosofia do não; O novo espírito científico; A poética do espaço**. Seleção de textos de José Américo Motta Pessanha. Tradução de Joaquim José Moura Ramos et al. São Paulo: Abril Cultural, 1978. p. 16-87. (Os Pensadores).
- COSTA, Sérgio Francisco. **Introdução ilustrada à Estatística**. 2. ed. São Paulo: Harbra, 1992.
- FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário da Língua Portuguesa**. 2. ed. Rio de Janeiro: Nova Fronteira, 1986.
- GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. (Org.). **Métodos de Pesquisa**. Porto Alegre: Ed. da UFRGS, 2009.
- GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- GOLDENBERG, Mirian. **A arte de pesquisar**. Rio de Janeiro: Record, 1997.
- HAN, Jiawei; KAMBER, Micheline. **Data Mining: Concepts and Techniques**. 2nd ed. San Francisco: Elsevier, 2006.
- KÜHN, Daniela Dias; TONETTO, Elci da Silva. O Programa Bolsa Família e os indicadores sociais de combate à pobreza no Rio Grande do Sul: um olhar multidimensional. **Desenvolvimento em questão**, Ijuí, v. 15, n. 39, p. 86-111, abr./jun. 2017.
- KUHN, Thomas Samuel. **A estrutura das revoluções científicas**. 5. ed. Tradução de Beatriz Vianna Boeira e Nelson Boeira. São Paulo: Perspectiva, 1998.
- LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Metodologia científica**. 5. ed. São Paulo: Atlas, 1999.
- MARIN, Solange Regina; VIEIRA, Carine de Almeida; KÜHN, Daniela Dias. Método Alkire-Foster: uma aplicação para a medição de pobreza multidimensional no Rio Grande do Sul (2000-2010). Artigo a ser publicado na revista **Planejamento e políticas públicas**, Brasília, IPEA, 2017.
- MINAYO, Maria Cecília. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 3. ed. São Paulo: Hucitec-Abrasco, 1994.

MORIN, Edgar. **Ciência com consciência**. 8. ed. rev. e mod. pelo autor. Tradução de Maria D. Alexandre e Maria Alice Sampaio Dória. Rio de Janeiro: Bertrand Brasil, 2005.

PENHA, Eli Alves. **A criação do IBGE no contexto da centralização política do Estado Novo**. Rio de Janeiro: IBGE/CCDI, 1993. (Documentos para Disseminação, Memória Institucional n. 4).

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. Índice de Desenvolvimento Humano Municipal. Brasília, DF: PNUD, IPEA, FJP, 2013. (Atlas do Desenvolvimento Humano no Brasil 2013).

ROSINI, Alessandro Marco; PALMISANO, Angelo. **Administração de sistemas de informação e a gestão do conhecimento**. São Paulo: Thomson Learning, 2003.

SANTOS, Antonio Raimundo dos. **Metodologia científica**: a construção do conhecimento. Rio de Janeiro: DP&A, 1999.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015.

TRIOLA, Mario F. **Introdução à Estatística**. Rio de Janeiro: LTC – Livros Técnicos e Científicos, 1999.

Este livro foi composto na tipologia Lapidary333 BT, em corpo 12
e impresso no papel Offset 75 g/m² na Gráfica da UFRGS

Editora da UFRGS • Ramiro Barcelos, 2500 – Porto Alegre, RS – 90035-003 – Fone/fax
(51) 3308-5645 – editora@ufrgs.br – www.editora.ufrgs.br • Direção: Alex Niche Teixeira •
Editoração: Luciane Delani (coordenadora), Clarissa Felkl Prevedello, Cláudio Marzo da Silva,
Cristina Thumé Pacheco e Lucas Ferreira de Andrade • Administração: Aline Vasconcelos da
Silveira, Cláudio Oliveira Rios, Fernanda Kautzmann, Gabriela Campagna de Azevedo, Getúlio
Ferreira de Almeida, Heloísa Polese Machado, Janer Bittencourt, Jaqueline Trombin e Laerte
Balbinot Dias • Apoio: Luciane Figueiredo