

# ACH0021

## Tratamento e Análise de Dados / Informações

Aula 3: Manipulação de dados

Alexandre Ribeiro Leichsenring  
alexandre.leichsenring@usp.br



# Organização

## 1 SPSS

- Visão Geral
- Estatísticas descritivas

## 2 Tipologia de variáveis

# Organização

- 1 SPSS
  - Visão Geral
  - Estatísticas descritivas

- 2 Tipologia de variáveis

# Visão Geral do SPSS

**SPSS:** Statistical Package for the Social Sciences

O SPSS é um pacote para análise e tratamento estatístico de dados. Nosso intuito será o de aplicar os conceitos vistos no curso e estudar os resultados fornecidos pelo programa.

*Sugestão:* Visitar os links e manuais disponibilizados na Internet e utilizar o Help do aplicativo.

*Atenção:* As figuras do SPSS apresentadas nas notas de aula não necessariamente correspondem a versão utilizada no laboratório.

# Referências

## Links

- Site oficial no Brasil. Permite baixar gratuitamente uma versão de avaliação.  
<http://www-03.ibm.com/software/products/pt/spss-statistics>

## Referências bibliográficas



Pestana, M. H., & Gageiro, J. N.

*Análise de dados para Ciências Sociais - A Complementaridade do SPSS*,  
Silabo, Lisboa, 2014.



Field, A. P

*Descobrimo a estatística usando o SPSS*,  
Porto Alegre Artmed 2009.



IBM Corporation,

IBM SPSS Statistics 20 Brief Guide,  
[Clique aqui para baixar.](#)



Levesque, R.

*Programming and Data Management for IBM SPSS Statistics 20: A Guide for  
IBM SPSS Statistics and SAS Users*,  
IBM CORP. 2011.



# Introdução

- Vamos introduzir os aspectos básicos e demonstrar uma seção típica de utilização do SPSS
- Vamos abrir um arquivo de dados e gerar medidas estatísticas básicas
- Vamos aprender algumas técnicas de manipulação de dados

# Abrindo um arquivo

Para abrir um arquivo:

- 1 Arquivo ► Abrir ► Dados
- 2 Selecione o arquivo: “Inclusão-Produtiva.sav” ([clique aqui para baixar](#))
- 3 ► Abrir

Por padrão, são abertos arquivos com extensão “.sav”

# Janela Visualização da Variável (Variable View)

Permite criar variáveis ou modificar definições existentes.

Inclusao-Produtiva.sav [Conjunto\_de\_dados1] - IBM SPSS Statistics Editor de dados

Arquivo Editar Visualizar Dados Transformar Analisar Marketing direto Gráficos Utilitários Janela Ajuda

	Nome	Tipo	Largura	Decimais	Rótulo	Valores	Ausente	Colunas	Alinhar
1	QUEST	Númerico	5	0	Número do Questionário	Nenhum	Nenhum	7	Direito
2	CD_IBGE	Númerico	7	0	Código IBGE do Município	{2100709, ...	Nenhum	9	Direito
3	SG_UF	Sequência	2	0	Sigla da UF	Nenhum	Nenhum	7	Esquerdo
4	CD_SEXO	Sequência	1	0	Código do sexo	{F, Feminin...	Nenhum	12	Esquerdo
5	IDADE	Númerico	5	0	Idade da pessoa (em 01/jul/2010)	Nenhum	Nenhum	5	Direito
6	CD_TIPO_...	Númerico	4	0	Código do tipo de localidade	{1, Urbana}	Nenhum	18	Direito
7	q5	Númerico	1	0	Q5. Estado civil	{1, Solteiro...	Nenhum	10	Direito
8	q10_a	Númerico	2	0	Q10. Faz quanto tempo que está morando, sem ...	{0, 00}...	Nenhum	5	Direito
9	q12	Númerico	1	0	Q12. Situação do domicílio:	{1, Próprio...	Nenhum	10	Direito
10	q13	Númerico	2	0	Q13. Qual é o material predominante na constru...	{1, Tijolo/Al...	Nenhum	12	Direito
11	q14	Númerico	2	0	Q14. Qual é a forma de abastecimento de água ...	{1, Rede p...	Nenhum	16	Direito
12	q15	Númerico	2	0	Q15. A água usada para beber é:	{1, Filtrada...	Nenhum	12	Direito
13	q16	Númerico	2	0	Q16. Qual é a forma de iluminação utilizada no ...	{1, Relógio ...	Nenhum	16	Direito
14	q17	Númerico	2	0	Q17. De que forma é feito o escoamento do ba...	{1, Rede p...	Nenhum	15	Direito
15	q18	Númerico	2	0	Q18. O lixo do seu domicílio:	{1, É coleta...	Nenhum	9	Direito
16	q19_1	Númerico	2	0	Q19. Quantas pessoas moram nesse domicílio, ...	{0, Nenuh...	Nenhum	5	Direito
17	q19_2	Númerico	2	0	Q19. Quantas pessoas moram nesse domicílio, ...	{0, Nenuh...	Nenhum	5	Direito
18	q19_3	Númerico	2	0	Q19. Quantas pessoas moram nesse domicílio, ...	{0, Nenuh...	Nenhum	5	Direito
19	q19_4	Númerico	2	0	Q19. Quantas pessoas moram nesse domicílio, ...	{0, Nenuh...	Nenhum	5	Direito

Visualização de dados **Visualização da variável**

IBM SPSS Statistics O processador está pronto



# Janela Visualização da Variável (*Variable View*)

**Descrição do Banco de Dados:** Um bom banco de dados contém uma descrição detalhada das suas variáveis.

- Tipo: Define o tipo de entrada no campo da variável. Os mais comuns são *Numérico* para números e *Sequência* para letras;
- Rótulo: Descrição do significado da variável;
- Valores: Descreve os significados de cada resposta possível;
- Ausente: Define as respostas inválidas ou não respondidas;
- Medir: Define o tipo de escala usada: Nominal, Ordinal ou Escalar (numérico).

# Janela Visualização de Dados (*Data View*)

Mostra o conteúdo de uma base de dados permitindo inserir ou alterar dados.

Inclusao-Produtiva.sav [Conjunto\_de\_dados1] - IBM SPSS Statistics Editor de dados

Arquivo Editar Visualizar Dados Transformar Analisar Marketing direto Gráficos Utilitários Janela Ajuda

Visível: 34 de 34 variáveis

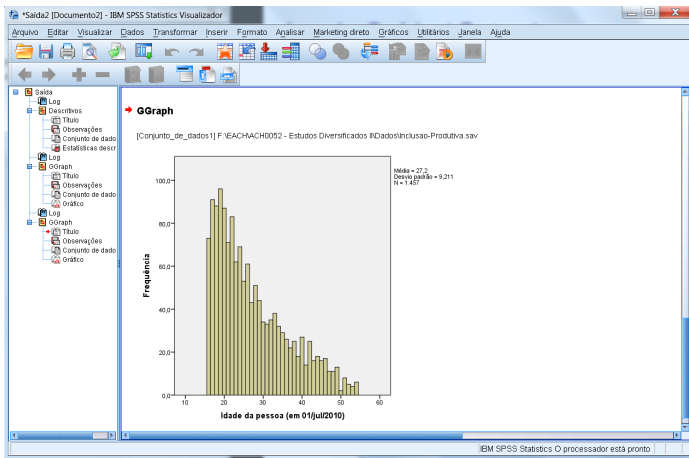
	QUEST	CD_IBGE	SG_UF	CD_SEXO	IDADE	CD_TIPO_LOCALIDADE	q5	q10_a	q12	q13
1	12308	Bacabeira MA		Masculino	22	Rural	Solteiro(a)	23	Próprio	Tijolo/Alvenaria
2	12249	Bacabeira MA		Feminino	48	Rural	Viúvo(a)	48	Próprio	Tijolo/Alvenaria
3	12164	Bacabeira MA		Masculino	28	Rural	Solteiro(a)	29	Próprio	Taipa não revest...
4	12148	Bacabeira MA		Masculino	30	Rural	Solteiro(a)	30	Próprio	Tijolo/Alvenaria
5	12145	Bacabeira MA		Masculino	18	Rural	Solteiro(a)	19	Cedido	Tijolo/Alvenaria
6	12034	Bacabeira MA		Feminino	46	Urbana	Amasiado	06	Próprio	Tijolo/Alvenaria
7	10119	São Luís MA		Masculino	25	Urbana	Amasiado	26	Próprio	Taipa revestida
8	10442	São Luís MA		Feminino	41	Urbana	Amasiado	42	Próprio	Tijolo/Alvenaria
9	12010	Bacabeira MA		Feminino	38	Rural	Solteiro(a)	08	Próprio	Tijolo/Alvenaria
10	10321	São Luís MA		Masculino	46	Urbana	Amasiado	30	Próprio	Tijolo/Alvenaria
11	12071	Bacabeira MA		Feminino	49	Rural	Casado(a)	50	Próprio	Tijolo/Alvenaria
12	11086	São José d... MA		Feminino	46	Urbana	Viúvo(a)	10	Próprio	Tijolo/Alvenaria
13	10108	São Luís MA		Feminino	44	Urbana	Casado(a)	30	Próprio	Tijolo/Alvenaria
14	18014	Santa Rita MA		Feminino	29	Rural	Solteiro(a)	08	Próprio	Tijolo/Alvenaria
15	12177	Bacabeira MA		Masculino	33	Rural	Viúvo(a)	07	Próprio	Tijolo/Alvenaria
16	10296	São Luís MA		Masculino	31	Urbana	Solteiro(a)	31	Próprio	Tijolo/Alvenaria
17	15016	Itapecuru M... MA		Feminino	38	Urbana	Casado(a)	11	Próprio	Tijolo/Alvenaria
18	13002	Anajatuba MA		Feminino	42	Urbana	Amasiado	42	Próprio	Tijolo/Alvenaria

Visualização de dados Visualização da variável

IBM SPSS Statistics O processador está pronto

# Janela Saída (*Output*)

**Viewer ou Output:** A janela de Output mostra automaticamente os resultados automaticamente sempre um procedimento de análise é executado.



# Tipos de escalas no SPSS

Para selecionar um tipo de escala para determinada variável, no menu:

Dados ► Definir Propriedades da Variável.

Ou diretamente em Medir na janela de variáveis.

IBM SPSS Statistics Editor de dados

Arquivo Editar Visualizar **Dados** Transformar Análise Marketing direto Gráficos Utilitários Janela Ajuda

Nome Tipo Largura Decimais Rótulo Valores Ausente Colunas Alinhar M

### Definir propriedades da variável

Lista de variável verificada

Não r.	Medida	Função	Variável
			QUEST
			CD IBGE
			SG UF
			CD SEXO
			IDADE
			CD TIPO LOCALIDADE
			q5
			q10 a
			q12
			q13
			q14
			q15
			q16
			q17
			q18
			q18 1
			q18 2
			q18 3
			q18 4

Variável atual: LOCALIDADE Rótulo: Código do tipo de localidade

Nível de medida: Normi Sugerir

Função: Escala Ordinal Nominal

Valores não rotulados: Insira ou edite rótulos na grelha. Você pode inserir valores adicionais na parte inferior.

Alterado	Ausente	Contagem	Valor	Rótulo
1		871	1	Urbana
2		586	2	Rural
3				

Casos verificados: 1457  
Limite da lista de valor: 200

Copiar propriedades: De outra variável... Para outras variáveis... Rótulos automáticos

OK Colar Redefinir Cancelar Ajuda

Visualização de dados Visualização da variável

IBM SPSS Statistics O processador está pronto

# Organização

- 1 SPSS
  - Visão Geral
  - Estatísticas descritivas

- 2 Tipologia de variáveis

# Tabela de frequências

É possível gerar tabelas de frequência facilmente, através do menu:

Analisar ► Estatísticas descritivas ► Frequências

The screenshot shows the IBM SPSS Statistics interface. The 'Analisar' menu is open, and the 'Estatísticas descritivas' sub-menu is selected. Within this sub-menu, the 'Frequências...' option is highlighted. The background shows a data table with columns for 'Nome', 'Tipo', 'Largura', 'res', 'Ausente', 'Colunas', 'Alinhar', and 'M'. The table contains 24 rows of data, including variables like QUEST, CD\_IBGE, SG\_UF, CD\_SEXO, IDADE, CD\_TIPO, q5, q10\_a, q12, q13, q14, q15, q16, q17, q18, q19\_1, q19\_2, q19\_3, q19\_4, q19\_5, q19\_6, q19\_7, q20, q21, and q22.

# Tabela de frequências

Selecione uma ou mais variáveis e clique em OK

\*Inclusao-Produtiva.sav [Conjunto\_de\_dados1] - IBM SPSS Statistics Editor de dados

Arquivo Editar Visualizar Dados Transformar Analisar Marketing direto Gráficos Utilitários Janela Ajuda

	Nome	Tipo	Largura	Decimais	Rótulo	Valores	Ausente	Colunas	Alinhar
1	QUEST	Numérico	5	0	Número do Questionário	Nenhum	Nenhum	7	Direito
2	CD_IBGE	Numérico	7	0	Código IBGE do Município	{2100709, ...	Nenhum	9	Direito
3	SG_UF	Sequência	2	0	Sigla da UF	Nenhum	Nenhum	7	Esquerdo
4	CD_SEXO	Sequência	1	0	Código do sexo	{F, Feminin...		12	Esquerdo
5	IDADE	Numérico	5	0	Idade da pessoa (em 01/jul/2010)	Nenhum	Nenhum	5	Direito
6	CD_TIPO_...	Numérico	4	0	Código do tipo de localidade	{1, Urbana}...	Nenhum	18	Direito
7	q5	Numérico	1	0	Q5. Estado civil	{1, Solteiro...	Nenhum	10	Direito
8	q10_a	Numérico	2	0	Q10. Faz quanto tempo que está morando, sem ...	{0, 00}...	Nenhum	5	Direito
9	q12	Numérico	1	0	Q12. Situação do domicílio	{1, Próprio	Nenhum	10	Direito
10	q13	Numérico	2	0	Q13. Qual é o r...				Direito
11	q14	Numérico	2	0	Q14. Qual é a f...				Direito
12	q15	Numérico	2	0	Q15. A água us...				Direito
13	q16	Numérico	2	0	Q16. Qual é a f...				Direito
14	q17	Numérico	2	0	Q17. De que fo...				Direito
15	q18	Numérico	2	0	Q18. O lixo do s...				Direito
16	q19_1	Numérico	2	0	Q19. Quantas p...				Direito
17	q19_2	Numérico	2	0	Q19. Quantas p...				Direito
18	q19_3	Numérico	2	0	Q19. Quantas p...				Direito
19	q19_4	Numérico	2	0	Q19. Quantas p...				Direito
20	q19_5	Numérico	2	0	Q19. Quantas p...				Direito
21	q19_6	Numérico	2	0	Q19. Quantas p...				Direito
22	q19_7	Numérico	2	0	Q19. Quantas pessoas moram nesse domicílio, ...	{0, Nenuh...	Nenhum	5	Direito
23	q20	Numérico	2	0	Q20. Quantidade de mulheres grávidas	{0, Nenuh...	Nenhum	4	Direito
24	q21	Numérico	2	0	Q21. Quantidade de mães amamentando	{0, Nenuh...	Nenhum	4	Direito

**Frequências**

Variável(is):

- Número do Q...
- Código IBGE ...
- Sigla da UF [S...
- Código do sex...**
- Idade da pess...
- Código do tipo...
- Q5. Estado ci...

Exibir tabelas de frequência

OK Colar Redefinir Cancelar Ajuda

Estadísticas...  
Gráficos...  
Formato...  
Bootstrap...

# Tabela de frequências

Veja o resultado na janela de Saída.

```
FREQUENCIES VARIABLES=CD_SEXO  
/ORDER=ANALYSIS.
```

## → Frequências

[Conjunto\_de\_dados1] F:\EACH\ACH0052 - Estudos Diversificados IIDados\Inclusao-Produtiva.sav

### Estadísticas

Código do sexo

N	Válido	1457
	Ausente	0

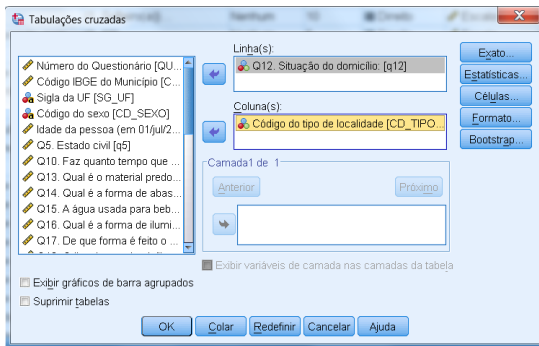
### Código do sexo

		Frequência	Porcentual	Porcentagem válida	Porcentagem acumulativa
Válido	Feminino	968	66,4	66,4	66,4
	Masculino	489	33,6	33,6	100,0
	Total	1457	100,0	100,0	



# Tabulações cruzadas (*Crosstabs*)

- São tabelas de frequência envolvendo mais de uma variável (preferencialmente qualitativas).
- Analisar ► Estatísticas descritivas ► Tabela de referência cruzada
- Selecione variáveis para Linha e para Coluna



- Clique em OK.

# Tabulações cruzadas (*Crosstabs*)

- Saída:

➔ **Tabulações cruzadas**

[Conjunto\_de\_dados1] F:\EACH\ACH0052 - Estudos Diversificados I\Dados\Inclusao-Produtiva.sav

**Resumo do processamento de caso**

	Casos					
	Válido		Ausente		Total	
	N	Porcentagem	N	Porcentagem	N	Porcentagem
Q12. Situação do domicílio: * Código do tipo de localidade	1457	100,0%	0	0,0%	1457	100,0%

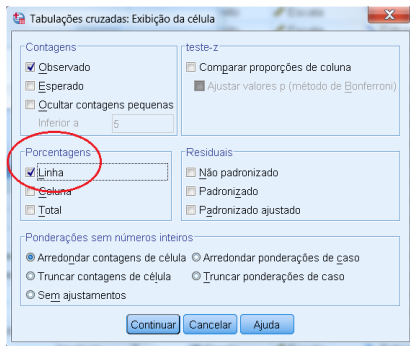
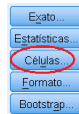
**Q12. Situação do domicílio: \* Código do tipo de localidade Tabulação cruzada**

Contagem

		Código do tipo de localidade		Total
		Urbana	Rural	
Q12. Situação do domicílio:	Próprio	719	516	1235
	Alugado	58	10	68
	Arrendado	0	1	1
	Cedido	59	29	88
	Invasão	33	30	63
	Financiado	2	0	2
Total		871	586	1457

# Tabulações cruzadas (*Crosstabs*)

- Você pode solicitar a tabela com porcentagens.
- Após selecionar as variáveis, clique em Células
- Escolha entre as possíveis formas de se obter a porcentagem (na Linha, por exemplo)
- OK



# Tabulações cruzadas (*Crosstabs*)

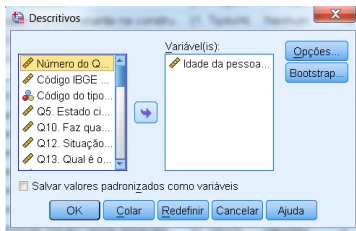
- Saída:

Q12. Situação do domicílio: \* Código do tipo de localidade Tabulação cruzada

			Código do tipo de localidade		Total
			Urbana	Rural	
Q12. Situação do domicílio:	Próprio	Contagem	719	516	1235
		% dentro de Q12. Situação do domicílio:	58,2%	41,8%	100,0%
	Alugado	Contagem	58	10	68
		% dentro de Q12. Situação do domicílio:	85,3%	14,7%	100,0%
	Arrendado	Contagem	0	1	1
		% dentro de Q12. Situação do domicílio:	0,0%	100,0%	100,0%
	Cedido	Contagem	59	29	88
% dentro de Q12. Situação do domicílio:		67,0%	33,0%	100,0%	
Invasão	Contagem	33	30	63	
	% dentro de Q12. Situação do domicílio:	52,4%	47,6%	100,0%	
Financiado	Contagem	2	0	2	
	% dentro de Q12. Situação do domicílio:	100,0%	0,0%	100,0%	
Total	Contagem	871	586	1457	
	% dentro de Q12. Situação do domicílio:	59,8%	40,2%	100,0%	

# Estatísticas descritivas

- Para variáveis quantitativas (escalares), pode-se facilmente gerar um conjunto de *Estatísticas descritivas*, isto é, um conjunto de medidas resumo.
- No menu: Analisar ► Estatísticas descritivas ► Descritivos.
- Selecione a(s) variável(is) desejada(s), e clique em OK



- Saída:

## ► Descritivos

[Conjunto\_de\_dados1] F:\EACH\ACH0052 - Estudos Diversificados I\Dados\Inclusao-Produtiva.sav

**Estatísticas descritivas**

	N	Mínimo	Máximo	Média	Desvio padrão
Idade da pessoa (em 01/jul/2010)	1457	16	54	27,20	9,211
N válido (de lista)	1457				

# Organização

- 1 SPSS
  - Visão Geral
  - Estatísticas descritivas

- 2 Tipologia de variáveis

# Dados e variáveis

## Dados

---

- Informação numérica coletada sobre produtos, serviços, processos, indivíduos ou itens
- Variam por natureza, não havendo dois objetos ou indivíduos idênticos.
- Características de interesse são chamadas de *variáveis*

## Variáveis

---

- Variáveis são características observáveis em cada elemento pesquisado.
- Sua medição permite atingir os objetivos específicos.
- Para cada elemento pesquisado, em dado momento, cada variável pode assumir **apenas** um valor: pesquisa deve ser planejada de maneira a garantir tal condição.

# Dados e variáveis

## Variáveis

---

**Qualitativas** Dados apresentam como possíveis resultados uma qualidade ou atributo do indivíduo pesquisado, geralmente classificados em categorias

**Quantitativas** Resultam de medidas numéricas tiradas sobre objetos ou pessoas



# Dados e variáveis

## Variáveis Qualitativas

---

Apresentam como possíveis resultados uma qualidade ou atributo do indivíduo pesquisado. Exemplo:

- a) Sexo
- b) Cor dos olhos
- c) Classe social
- d) Escolaridade

## Níveis de mensuração

**Nominais** Dados nominais são *rotulados*;  
Permitem *classificar e categorizar*

**Ordinais** Dados ordinais são *graduados*  
Permitem *ordenar*

# Dados e variáveis

## Variáveis Quantitativas

---

- São dados numéricos
- Não informa apenas ordenação mas indica também precisamente distância entre valores
- Permitem cálculos

Apresentam como possíveis resultados números resultantes de uma contagem ou mensuração;

- Peso
- Altura
- Distância
- Rendimento

## Tipos

**Discretas** Determinam quantidades inteiras

**Contínuas** Resultam geralmente de medições

# Estatísticas Descritivas por tipo de variável

## Qualitativas

- Tabelas de frequência  
Analyze ► Descriptive Statistics ► Frequencies
- Tabelas de contingência  
Analyze ► Descriptive Statistics ► Crosstabs

## Quantitativas

- Estatísticas descritivas  
Analyze ► Descriptive Statistics ► Descriptives