



Data Analysis em Saúde no Brasil

Cleber Martins, Guilherme Schneider
e Naiome Alves



Fonte dos Dados



- Dados Demográficos: População Estimada, Área Territorial, Escolarização, IDH, Mortalidade Infantil e PIB per Capta.
- Filtrado por: Local: município; Período: Ano (2021)

tabnet

DATA SUS
DEPARTAMENTO DE INFORMÁTICA DO SUS

- geração de resultados em tabela;
- importação para formado csv;

Fonte dos Dados

- Dados da Rede Assistencial: Tipos de Estabelecimentos e Recursos Humanos (Médicos e Áreas Correlatas: Assistência Social, Fisioterapia, Nutrição)
- Filtrado por: Local: município; Período: Ano (2021)

GE
estatística

tabnet

DATASUS
DEPARTAMENTO DE INFORMÁTICA DO SUS

- geração de resultados em tabela;
- importação para formado csv;

Tratamento dos Dados

Limpeza da Base

- Missing Values: 1.404 Base IBGE
- Substituição de "," por "."
- Padronização do Código do Município
- Seleção de Colunas
- Combinação de base de dados
- Excluir cabeçalho e rodapé de informações
- Tratamento de caracteres não-numerais

Pré-processamento

- Codificar arquivos para Latin-1
(Ç, ~, etc)
- Eliminar Missings
- Criação de Dataframe



Análise Descritiva e Exploratória

- Feature Engineering
- Plotagem de Gráficos
- Análise describe (Min, Max, Desvio padrão)

- Dataframe para criação de novas variáveis
- Comparação entre variáveis

Legendas (Attribute Information)

CD_MUN - Código do Município

AR_TER_km2 - Área Territorial

POP_EST - População Estimada

DENS_DEM - Densidade Demográfica

ESC - Escolaridade

IDHM - Índice de Desenvolvimento Humano

MORT_INF - Mortalidade Infantil

PIB_P_CAP - Produto Interno Bruto Per Capta

MD_CIR_G - Médico Cirurgião-Geral

MD_CLIN - Médico Clínico

MD_GIN_OBS - Médico Ginecologista-Obstetra

MD_FAM - Médico de Família e Comunidade

MD_PED - Médico Pediatra

MD_GERI - Médico Geriatria

MD_URO - Médico Urologista

AS_SOC - Assistente Social

ENF - Enfermeiro

FISIO - Fisioterapeuta

FONO - Fonoaudiólogo

NUTRI - Nutricionista

ODONTO - Odontologista

PSICO - Psicólogo

PSIQ - Psiquiatra

POSTO - Posto de Saúde

CS_UBS - Centro de Saúde / Unidade Básica de Saúde

POLICL - Policlínica

HOSP_GER - Hospital Geral

HOSP_ESP - Hospital Especializado

PRT_SOC - Pronto-Socorro Geral

PRT_ESP - Pronto-Socorro Especializado

df.describe

```
[34] df.describe()
```

	AR_TER_km2	POP_EST	DENS_DEM	ESC	IDHM	MORT_INF	PIB_P_CAP	MD_CIR_G	MD_CLIN	MD_GIN_OBS	...	ODONTO	PSICO	PSIQ	POSTO	CS_UBS	POLICL	HOSP_GER	HOSP_ESP	PRT_SOC	PRT_ESP
count	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	...	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000	5570.000
mean	242470.196	38297.601	108.246	97.138	591.770	19.030	23513.942	1.833	25.862	4.192	...	16.085	13.481	1.699	1.583	7.100	1.876	0.971	0.180	0.055	0.014
std	258410.737	224288.153	571.882	2.865	210.379	10.862	24238.463	20.493	187.253	43.857	...	97.454	85.028	20.265	3.120	13.388	14.963	3.418	1.958	0.384	0.174
min	26.100	771.000	0.130	51.800	0.450	1.960	4788.180	0.000	0.000	0.000	...	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
25%	3715.945	5454.000	11.570	96.600	579.000	11.242	9880.370	0.000	1.000	0.000	...	1.000	1.000	0.000	0.000	2.000	0.000	0.000	0.000	0.000	0.000
50%	170743.000	11732.000	24.400	97.800	651.000	18.085	17433.840	0.000	3.000	0.000	...	2.000	2.000	0.000	0.000	4.000	0.000	1.000	0.000	0.000	0.000
75%	379095.000	25764.750	51.835	98.500	712.000	24.365	28729.907	0.000	10.000	1.000	...	7.000	8.000	1.000	2.000	8.000	1.000	1.000	0.000	0.000	0.000
max	999517.000	12396372.000	13024.560	100.000	862.000	181.820	583171.850	1067.000	9833.000	2355.000	...	4384.000	4102.000	1200.000	78.000	522.000	581.000	157.000	68.000	15.000	6.000

8 rows x 29 columns

A decorative graphic in the bottom right corner featuring several overlapping hexagons in shades of teal and green, set against a light gray background.





