






BCG X Challenge 2023

Descritivo de bases

SÃO PAULO / 30 DE SETEMBRO



Bases de dados (1 / 4)

	1	2	3
	Registros Administrativos	Censo	INMET
 Descrição	<ul style="list-style-type: none">Indic. de desenvolvimento humano dos municípios	<ul style="list-style-type: none">Indicadores do censo no nível município	<ul style="list-style-type: none">Dados climáticos de estações meteorológicas
 Principais dados	<ul style="list-style-type: none">SaúdeEconomiaEducaçãoInfraestrutura	<ul style="list-style-type: none">Indicadores para múltiplas áreas averiguadas no censo<ul style="list-style-type: none">- Saúde- Educação- Economia	<ul style="list-style-type: none">Dados meteorológicos de precipitação, temperatura, umidade, vento, etc.Informações sobre estações meteorológicas
 Granularidade	<ul style="list-style-type: none">Município	<ul style="list-style-type: none">Município	<ul style="list-style-type: none">Estação
 Frequência	<ul style="list-style-type: none">Anual (2012-2017)	<ul style="list-style-type: none">Pontual (1999, 2000, 2010)	<ul style="list-style-type: none">Data-Hora (2020-2022)





Bases de dados (2/4)

	4	5	6
	Exportações	Produção Agrícola Municipal	Transportes
 Descrição	<ul style="list-style-type: none">• Informações sobre exportações brasileiras	<ul style="list-style-type: none">• Índice de produção agrícola de diversos produtos	<ul style="list-style-type: none">• Aproximação de custo de transportes
 Principais dados	<ul style="list-style-type: none">• Produto exportado• Município de origem• País de destino• Valor de remessa• Peso exportado	<ul style="list-style-type: none">• Área colhida• Área plantada• Rendimento médio	<ul style="list-style-type: none">• Custo de transporte entre municípios e porto mais próximo
 Granularidade	<ul style="list-style-type: none">• Município	<ul style="list-style-type: none">• Município	<ul style="list-style-type: none">• Município
 Frequência	<ul style="list-style-type: none">• Mensal (2010-2022)	<ul style="list-style-type: none">• Anual (2003-2022)	<ul style="list-style-type: none">• Pontual

Bases de dados (3/4)

	7	8	9
	ANA	Lista de municípios do semiárido	De-para de municípios
 Descrição	<ul style="list-style-type: none">Qualidade da água e área irrigada	<ul style="list-style-type: none">Lista dos municípios da região semiárida	<ul style="list-style-type: none">De-para entre código do município e nome / UF
 Principais dados	<ul style="list-style-type: none">Índice de qualidade da águaÁrea irrigada por culturaÁrea adicional irrigável	<ul style="list-style-type: none">MunicípioUF	<ul style="list-style-type: none">Código municípioNome municípioUFLatitude/longitude
 Granularidade	<ul style="list-style-type: none">Município	<ul style="list-style-type: none">Município	<ul style="list-style-type: none">Município
 Frequência	<ul style="list-style-type: none">Anual (2019-2021)	<ul style="list-style-type: none">Pontual	<ul style="list-style-type: none">Pontual

Bases de dados (4/4)

		10
	QEdu	
	Descrição	<ul style="list-style-type: none">• Informações sobre escolas e matrículas
	Principais dados	<ul style="list-style-type: none">• Número de escolas• Características da infraestrutura escolar• Matrículas e docentes• Características do corpo docente
	Granularidade	<ul style="list-style-type: none">• Município
	Frequência	<ul style="list-style-type: none">• Anual (2019-2020)

1. Registros Administrativos



Nome do arquivo

dados_registros_administrativos_[total/cor/sexo]_2012_2017.xlsx



Dicionário

Mesmo arquivo, aba "Legenda"



Fonte

Atlas do Desenvolvimento Humano
<http://www.atlasbrasil.org.br/acer vo/biblioteca>



Descrição

Base com indicadores de desenvolvimento humano dos municípios construídos pelo Atlas do Desenvolvimento Humano a partir de registros administrativos (com quebras por sexo e cor)



Principais variáveis

Indicadores do município para múltiplas áreas

- Saúde (ex: percentual de internações)
- Educação (ex: taxa de evasão, taxa de matrícula)
- Economia (ex: rendimentos)
- Infraestrutura (ex: acesso à saneamento básico, água potável)



Observações importantes

- Notas metodológicas estão disponíveis no site do Atlas do Desenvolvimento Humano

2. Censo



Nome do arquivo

Censo_municipal_estadual_nacional.xlsx
Censo_municipal_estadual_nacional_quebras_cor_sexo_domicilio.xlsx



Dicionário

No mesmo arquivo, aba "Siglas"



Fonte

Atlas do Desenvolvimento Humano
<http://www.atlasbrasil.org.br/acer vo/biblioteca>



Descrição

Base com resultados do Censo realizado pelo IBGE, com indicadores no nível municipal, estadual e nacional, além de quebras por sexo, cor e local do domicílio (rural vs. urbano)



Principais variáveis

Indicadores para múltiplas áreas averiguadas no censo

- Saúde (ex: expectativa de vida, mortalidade)
- Educação (ex: taxa de analfabetismo, taxa de atendimento escolar)
- Economia (ex: renda domiciliar)



Observações importantes

- Censo realizado em 1990, 2000 e 2010
- Municípios do semiárido não possuem quebras por cor/sexo

3. INMET



Nome do arquivo

INMET_[2020/2021/2022].csv



Dicionário

Nome de colunas autoexplicativo



Fonte

Instituto Nacional de Meteorologia
<https://portal.inmet.gov.br/>

Banco de Dados
<https://bdmep.inmet.gov.br/>



Descrição

Base com dados meteorológicos das estações automáticas do INMET, com informações a nível data-hora para cada estação, separados por ano



Principais variáveis

Indicadores meteorológicos coletados pelas estações de forma automática

- Dados meteorológicos (ex: precipitação, temperatura, umidade, vento)
- Informações da estação (ex: região, UF, latitude, longitude, altitude)



Observações importantes

- Possível gerar tabelas de dados adicionais para estações convencionais, seguindo as instruções disponíveis em <https://bdmep.inmet.gov.br/>
- Para junção dos dados com municípios, é possível utilizar lat/lon como chave em um GeoDataFrame (geopandas)

4. COMEX



Nome do arquivo

EXP_[2020/2021/2022]_MUN.csv



Dicionário

SH4: códigos de grupos de produtos
KG_LIQUIDO: Quantidade exportada (KG)
VL_FOB: valor de fobização (R\$)



Fonte

COMEX STAT
<http://comexstat.mdic.gov.br/pt/home>



Descrição

Bases com dados de exportação a nível de grupo de produtos (SH4), município e ano/mês. De para de produtos e municípios estão disponíveis em sh4_agropecuária.xls e na base de lat_lon municípios.



Principais variáveis

- Quantidade exportada (em KG)
- Valor de fobização (Custo em R\$)



Observações importantes

- Possível gerar tabelas de dados adicionais em <http://comexstat.mdic.gov.br/pt/municipio>

5. PAM



Nome do arquivo

pam_area_colhida.csv
pam_area_plantada.csv
pam_rendimento_medio.csv
pam_valor_producao.csv



Dicionário

Nome de colunas autoexplicativo



Fonte

IBGE - Produção Agrícola Municipal
<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html>



Descrição

Bases com dados de produção agrícola, com informações a nível de município e produtos, separados por ano (20,21 e 22). Cada base tem informações de variáveis diferentes. Bases filtradas para produtos relevantes para Amigos do Bem e municípios do semiárido



Principais variáveis

- Área colhida
- Área plantada
- Rendimento médio
- Valor da produção



Observações importantes

- Possível gerar tabelas de dados adicionais para outras variáveis, produtos, anos e municípios, seguindo as instruções disponíveis em <https://sidra.ibge.gov.br/tabela/5457>

6. Transportes



Nome do arquivo

transport_cost.xlsx



Dicionário

Nome de colunas autoexplicativo



Fonte

Construída internamente pelo BCG apenas para uso durante o desafio do BCG X Challenge 2023



Descrição

Bases com custos de transportes aproximados entre municípios e os portos mais próximos. Custo de referência para uma viagem entre o município e o porto levando um volume de 50m³.



Principais variáveis

- Município (de origem)
- Porto (de destino)
- Custo (R\$)



Observações importantes

Custos aproximados de transporte rodoviário na região.

7. ANA



Nome do arquivo

Indicadores de Qualidade da Agua.csv
Atlas Irrigacao - Area Atual e
Potencial.xlsx



Dicionário

MIN, MED, MAX: Índice de qualidade da
água mínimo/médio/máximo no ano



Fonte

ANA
<https://dadosabertos.ana.gov.br/>



Descrição

Bases com informações sobre qualidade da água nos municípios e registros de
área irrigada por cultura e potencial adicional irrigável



Principais variáveis

- Valores mínimos, médios e máximos de qualidade da água
- Estação e corpo d'água, com latitude e longitude
- Área irrigada por tipologia de cultura
- Área adicional irrigável



Observações importantes

- Dados adicionais disponíveis no Atlas de Irrigação e [Atlas Águas](#)
- Para junção dos dados com municípios, é possível utilizar lat/lon como chave em um GeoDataFrame (geopandas)

8. QEdu



Nome do arquivo

Dados_QEdu_Analitico_[2019/2020].csv



Dicionário

Nome de colunas autoexplicativo



Fonte

QEdu
<https://analitico.qedu.org.br/>



Descrição

Base com informações educacionais sobre escolas, número de matrículas e corpo docente, por município



Principais variáveis

- Total de escolas, matrículas e docentes
- Números totais em quebras específicas (e.g., escolas sem abastecimento de água)
- Características dos docentes (e.g., docentes no Ensino Fundamental)



Observações importantes

- Possível gerar base com outras quebras no site do QEdu