

Documentação do Data Lake Open Brewery

Visão Geral do Projeto

Este projeto estabelece um Data Lake moderno utilizando a Arquitetura Medalhão (Bronze, Silver, Gold). O objetivo é ingerir, limpar e analisar dados globais de cervejarias da API Open Brewery DB para gerar insights de negócios.

Arquitetura de Dados

1. Camada Bronze (Raw)

- **Processo**: Ingestão automatizada da Open Brewery DB via API REST.
- **Armazenamento**: Formato JSON original preservando o estado bruto.

2. Camada Silver (Limpa)

- **Processo**: Limpeza, deduplicação e padronização dos dados.
- **Armazenamento**: Formato Parquet otimizado, particionado por `brewery_type`.

3. Camada Gold (Analítica)

- **Processo**: Agregações complexas e cálculo de métricas de negócio.
- **Qualidade**: Verificações automatizadas garantem a integridade dos dados.

Insights Analíticos

1. Score de Maturidade Digital

- **Definição**: Porcentagem de cervejarias em um estado que possuem site e telefone registrados.
- **Insight**: Pontuações altas indicam um mercado digitalmente avançado e acessível.

2. Índice de Diversidade Regional

- **Definição**: Quantidade de tipos únicos de cervejaria presentes no estado.
- **Insight**: Alta variedade indica um mercado heterogêneo e maduro.

3. Especialização de Mercado (Contagem Micro)

- **Definição**: Total de cervejarias vs. cervejarias "micro" por estado.
- **Insight**: Destaca centros de produção artesanal.

4. Data Trust Score (Confiança de Dados)

- **Definição**: Média de completude de campos críticos (Endereço, Telefone e Website).
- **Insight**: Indicador de confiabilidade da base de dados comercial.

5. Hubs Geográficos

- **Definição**: Cidades com a maior concentração de cervejarias.

Open Brewery Data Lake - Relatório Estratégico

- **Insight**: Identifica pontos estratégicos e competitivos.

6. Cervejarias por Tipo e Estado

- **Definição**: Detalhamento da contagem de cervejarias por categoria (micro, nano, brewpub, etc.) para cada estado.
- **Insight**: Útil para entender o perfil industrial de cada região.

7. Distribuição Global por País e Tipo

- **Definição**: Comparação dos tipos de cervejaria entre diferentes países.
- **Insight**: Revela tendências internacionais na estrutura do mercado cervejeiro.

8. Cobertura Geográfica por Estado

- **Definição**: Mergulho profundo na distribuição de cervejarias entre as cidades dentro dos estados.
- **Insight**: Essencial para planejamento logístico e penetração de mercado regional.

Detailed Data Tables (Top Results)

Breweries By Type And State

brewery_type	state_province	brewery_count
micro	ACT	7
micro	Alabama	26
closed	Alabama	15
brewpub	Alabama	5
planning	Alabama	3
taproom	Alabama	3
micro	Alaska	28
brewpub	Alaska	18
closed	Alaska	7
taproom	Alaska	4

Breweries By Country And Type

country	brewery_type	brewery_count
Australia	micro	491
Australia	large	22
Australia	regional	1
Austria	large	10
Austria	bar	2
Austria	nano	2
Austria	brewpub	1
Canada	micro	54
Canada	brewpub	44
Canada	closed	12

Top Cities By Brewery Count

city	state_province	country	brewery_count
Denver	Colorado	United States	92
San Diego	California	United States	91
Portland	Oregon	United States	85
Seattle	Washington	United States	85
Chicago	Illinois	United States	64
Austin	Texas	United States	49
Houston	Texas	United States	40
San Francisco	California	United States	39
Minneapolis	Minnesota	United States	39
Cincinnati	Ohio	United States	34

Geo Coverage By State

country	state_province	city	brewery_count
Australia	ACT	Fyshwick	4

Open Brewery Data Lake - Relatório Estratégico

country	state_province	city	brewery_count
Australia	ACT	Braddon	2
Australia	ACT	Gungahlin	1
Australia	NSW	Marrickville	12
Australia	NSW	Brookvale	6
Australia	NSW	Alexandria	4
Australia	NSW	Botany	4
Australia	NSW	South Murwillumbah	4
Australia	NSW	Armidale	3
Australia	NSW	Newtown	3

Digital Maturity

state_province	total_breweries	digitally_ready_count	maturity_score
ACT	7	7	100.0
Argyll	4	4	100.0
Chungcheongnamdo	3	3	100.0
Clare	1	1	100.0
Carlow	1	1	100.0
Busan	9	9	100.0
Bute	1	1	100.0
Bouche du Rhône	3	3	100.0
Bolzano	4	4	100.0
Kerry	3	3	100.0

Regional Diversity

state_province	unique_brewery_types
Washington	11
Wisconsin	9
Maryland	8
Texas	8
Oregon	8
Colorado	8
California	8
Missouri	8
Tennessee	8
Ohio	8

Market Specialization

state_province	total_brewery_count	micro_brewery_count
California	919	466
Washington	498	245
Colorado	449	228
New York	419	226
Michigan	375	143
Texas	352	196
Pennsylvania	345	129
Florida	312	190

Open Brewery Data Lake - Relatório Estratégico

state_province	total_brewery_count	micro_brewery_count
North Carolina	311	185
Ohio	303	163

Data Trust Score

state_province	trust_score
ACT	100.0
Argyll	100.0
Chungcheongnamdo	100.0
Clare	100.0
Carlow	100.0
Busan	100.0
Bute	100.0
Bouche du Rhône	100.0
Bolzano	100.0
Kerry	100.0