

# AQI INDICATORS

**Banco de dados em MySQL / API em Python / Dashboard em Power BI**

Contexto: O indicador AQI, ou Índice de Qualidade do Ar, é uma medida utilizada para avaliar a qualidade do ar em uma determinada região. Ele considera diversos poluentes atmosféricos e atribui a eles uma classificação de acordo com os níveis de concentração. O AQI fornece uma visão geral da qualidade do ar e ajuda a informar a população sobre possíveis impactos na saúde e na qualidade de vida.

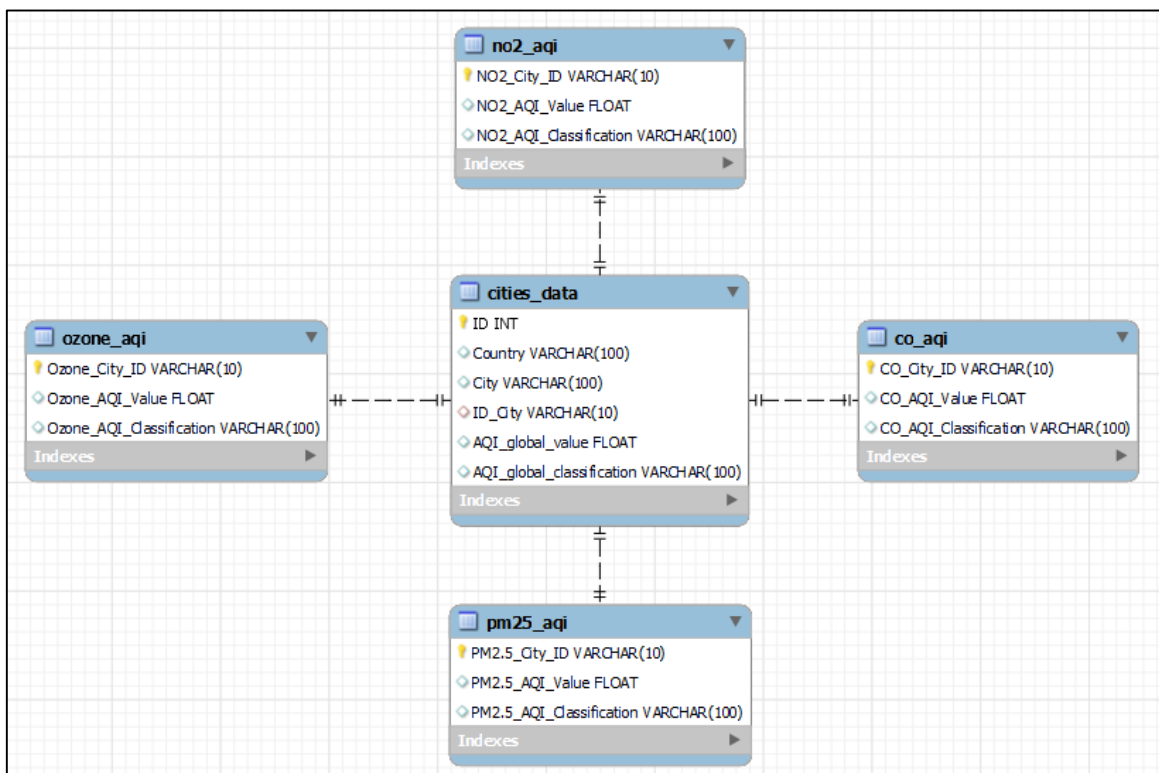
## 1. CRIAÇÃO DO BANCO DE DADOS EM MySQL

Primeiramente, foi construído um banco de dados relacionais no MySQL. O banco de dados possui 5 tabelas, nas quais a tabela mãe contém as colunas de países, cidades, AQI global e classificação do AQI global, além de um ID para cada país, o qual possibilita o relacionamento da tabela mãe com as tabelas filhas. A coluna de ID dos países foi construída por meio de algoritmo em linguagem python, o qual será apresentado nos tópicos seguintes.

Cada tabela filha apresenta informações de AQI respectiva a cada poluente analisado do *dataset* a ser importado posteriormente: CO, NO<sub>2</sub>, Ozônio e PM 2.5.

O Diagrama Entidade-Relacionamento (DER) do banco de dados elaborado pode ser visto na Figura 1.

Figura 1: DER do banco de dados elaborado neste projeto.



## 2. CRIAÇÃO DE API PARA ALIMENTAÇÃO DO BANCO DE DADOS

Para conectar o *dataset* de informações de AQI com o banco de dados em MySQL e realizar a alimentação de dados, elaborou-se uma API em linguagem Python. Esse algoritmo possui basicamente 4 etapas, as quais são:

1. Conexão do Python ao MySQL;
2. Leitura do *dataset* em CSV com as informações a serem importadas;
3. Criação de ID para as cidades, facilitando o relacionamento entre tabelas;
4. Inserção dos dados nas suas respectivas tabelas.

O *dataset* importado nesse estudo pode ser consultado em: <https://www.kaggle.com/datasets/hasibalmuzdadid/global-air-pollution-dataset>.

O algoritmo elaborado pode ser visto abaixo:

```
# IMPORTANDO BIBLIOTECAS

import MySQLdb as MS
import pandas as pd
import numpy as np

# CONECTANDO O PYTHON AO MYSQL

host = "localhost"
user = "root"
password = "*****"
db = "aqi_countries"
port = 3306

con = MS.connect (host, user, password, db, port)
c = con.cursor(MS.cursors.DictCursor)

# LENDO O DATASET

x = pd.read_csv(r'C:\Users\carlo\Desktop\DATA
SCIENCE\AQI_Countries_1.csv')
dados = pd.DataFrame(x)

# CRIANDO IDs PARA AS CIDADES

dados.insert(2, "ID_City", np.arange(len(dados["City"])), True)

# MECANISMO DE INSERÇÃO DE DADOS
```

```

def insert_data(values, table):

    global c, con

    query = "INSERT INTO " + table
    query = query + " VALUES " + ",".join("(" + v + ")" for v in
values])

    c.execute(query)
    con.commit()

# INSERINDO DADOS NAS TABELAS

for linha in range (len(dados['City'])):

    values1 = ["DEFAULT, '{0}', '{1}', '{2}', {3},
'{4}'".format(dados['Country'][linha], dados['City'][linha],
dados['ID_City'][linha], dados['AQI Value'][linha], dados['AQI
Category'][linha])]
    insert_data(values1, "cities_data")

    values2 = ["'{0}', {1}, '{2}'".format(dados['ID_City'][linha],
dados['CO AQI Value'][linha], dados['CO AQI Category'][linha])]
    insert_data(values2, "co_aqi")

    values3 = ["'{0}', {1}, '{2}'".format(dados['ID_City'][linha],
dados['NO2 AQI Value'][linha], dados['NO2 AQI Category'][linha])]
    insert_data(values3, "no2_aqi")

    values4 = ["'{0}', {1}, '{2}'".format(dados['ID_City'][linha],
dados['Ozone AQI Value'][linha], dados['Ozone AQI
Category'][linha])]
    insert_data(values4, "ozone_aqi")

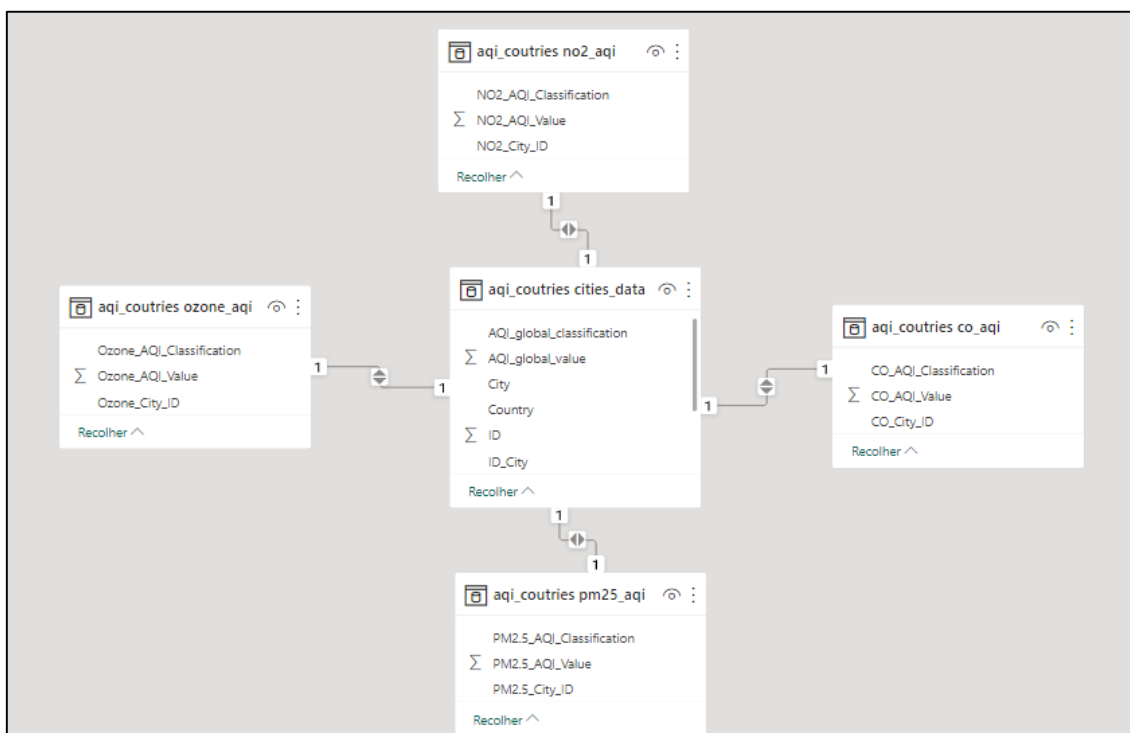
    values5 = ["'{0}', {1}, '{2}'".format(dados['ID_City'][linha],
dados['PM2.5 AQI Value'][linha], dados['PM2.5 AQI
Category'][linha])]
    insert_data(values5, "pm25_aqi")

```

### 3. ELABORAÇÃO DO DASHBOARD NO POWER BI

Com o banco de dados devidamente construído e alimentado, importou-se o mesmo no Microsoft Power BI. As relações entre as tabelas criadas no MySQL e as formatações das colunas foram corretamente interpretadas no Power BI (Figura 2), dispensando-se, assim, quaisquer edições de tabelas dentro desse software. Com isso, construiu-se apenas o *dashboard* propriamente dito.

Figura 2: Relações entre tabelas no Power BI, as quais foram mantidas do MySQL.



No *dashboard* (Figura 3), foram criadas as seguintes visualizações:

- Mapa com cidades com valores mais críticos de AQI;
- Gráfico de colunas 100% empilhadas com a proporção entre classificações de AQI (de cidades) para países com mais cidades analisadas;
- Gráfico de barras com os piores AQI médios por países, juntamente com gráficos de linha indicando os valores de AQI médio por poluente;
- Indicador KPI para os valores de AQI (Limite estabelecido de 100 indica a saída de um valor de KPI moderado para insalubre);
- Cartão com a classificação de acordo com valor de AQI.

Com estes artifícios estatísticos elencados anteriormente, pode-se observar o problema sob diversas perspectivas, tais como:

➤ Onde há maiores concentrações de cidades com níveis mais críticos de AQI?

“Observa-se, no mapa, maiores concentrações de cidades críticas na Índia e em suas proximidades, no Leste da China, no Leste da França, no Oeste da África e no Sul do México.”

- Nos países com mais cidades estudadas, em quais há maiores proporções de cidades com valores críticos de AQI?

“Dentre os países com mais cidades registradas, observa-se no gráfico um maior teor de cidades críticas na Índia, na China e no México.”

- Considerando os valores médios de AQI, qual país se encontra em situação mais crítica?

“Observa-se, pelo gráfico, uma situação mais crítica na Coreia do Sul. Tal criticidade não é observada nos outros dois questionamentos levantados acima pelo fato do volume de cidades analisadas nesse país ser bem menor que em outros já mencionados.”

- Quais poluentes são mais preponderantes para a elevação de AQI nos países mais críticos?

“Segundo o gráfico, em todos há uma grande incidência de poluição por PM 2.5. Entretanto, nos Emirados Árabes Unidos, no Kuwait e no Catar, há grandes teores de poluição por Ozônio.”

Figura 3: *Dashboard* de indicadores de AQI construído no Power BI.

