



1 Descrição do Projeto

Durante a evolução da pandemia da COVID-19 vários institutos uniram as forças para disponibilizar os dados a respeito de sua evolução. Um destes sites é o *Our World in Data*¹ que atualiza os dados diariamente por meio de arquivos CSV, XLSX e JSON. Os arquivos são equivalentes e possuem 59 variáveis (atributos). Por simplicidade organizaram os dados todos em um único arquivo, então obviamente há uma grande quantidade de dados repetidos indicando que não estão normalizados.

O projeto de vocês será de fazer uma engenharia reversa neste arquivo de dados por meio das seguintes tarefas:

1. Normalização dos dados produzindo o esquema relacional equivalente
2. Criação dos diagramas do esquema relacional
3. Implementação do esquema relacional no Banco de Dados *MySQL*
4. Carga dos dados no bancos de dados
5. Implementação de consultas

2 “O que” deve ser entregue

1. **Modelo Lógico** na forma do diagrama representando o esquema Relacional que resultou da normalização. Utilize a notação *crow's foot* para os relacionamentos e a notação **IDEF1X** para as entidades (Retângulo com 3 compartimentos) (sugestão de software: *MySQL Workbench*)
2. **Modelo Físico** na forma do *script* de criação em SQL para o Banco de Dados. O *script* deve ter em sua primeira parte a criação de todas as tabelas, respectivos atributos e restrições, entretanto as restrições de integridade referencial (chaves estrangeiras) devem ser definidas após a criação de **todas** as tabelas.
3. **Script de Inserção** dos dados para as tabelas. O script de inserção será executado logo após o script de criação do banco de dados.
4. **Consultas em SQL** requisitadas

2.1 Consultas SQL

As seguintes consultas devem ser implementadas:

1. Quantos países estão na base de dados?
2. Qual o total de casos para o mundo no dia 14/04/2021?
3. Quais foram os 10 países com mais casos confirmados no mês de Março/2021 (ordem decrescente)?
4. Liste os 10 países com maior e os 10 com menor expectativa de vida.
5. Liste os continentes contendo o total de casos de cada um em 2021
6. Liste os países da América do Sul e inclua as informações do total de pessoas totalmente vacinadas em 2021, em valores absolutos e percentual. Ordene o resultado em ordem decrescente pelo percentual de vacinados.

¹<https://ourworldindata.org/>

7. Liste os países informando o grau de restrições² que foram aplicadas à população (há um atributo com este índice (*stringency*) que engloba várias medidas como fechamento de escolas, proibição de viagens, ...) e o total de novos casos por milhão de habitantes confirmados para o mês de Março/2021.
8. Quais países não possuem informação de pacientes na UTI para o mês de Fevereiro/2021
9. Qual foi o dia com a maior quantidade de novos casos registrados de COVID-19 no Brasil?
10. Qual foi o dia com a maior quantidade de mortes confirmadas por COVID-19 no mundo?

3 “Como” deve ser feita a entrega

- A entrega deve ser feita por meio de um arquivo compactado (.zip). O nome do arquivo deve ser composto pela matrícula dos integrantes do grupo, em ordem crescente, separada por um hífen (-) (ex. aaaa-bbbb-cccc.zip). O arquivo compactado deve conter:
 - Um arquivo em formato pdf com nome “diagrama.pdf” contendo o diagrama do esquema relacional resultante do processo de normalização.
 - Um arquivo contendo o *script* de criação do banco de dados em formato de texto puro com o nome “create.sql”
 - Um arquivo contendo o *script* de carga (inserção) do banco de dados em formato de texto puro com o nome “load.sql”
 - Um arquivo contendo o *script* de consultas ao banco de dados em formato de texto puro com o nome “query.sql”
- Formate seu trabalho para que fique legível. Coloque os diagramas na orientação do papel (retrato ou paisagem) que privilegie uma boa visualização. Tenha cuidado com o layout do seu diagrama.

4 Critérios de Avaliação

A pontuação será assim distribuída:

- 20% - Normalização
- 20% - Modelo Lógico (Relacional)
- 20% - Modelo Físico
- 20% - *Script* de Inserção dos dados
- 20% - Consultas SQL

A entrevista tem importância fundamental no processo de avaliação do projeto. Caso fique evidente desconhecimento dos integrantes do grupo ou todo o grupo em relação ao conteúdo entregue poderão ocorrer penalizações na nota final.

5 Cronograma para o Projeto

- 10/05/2021 - Entrega do trabalho pelo PVA.net.
- 11, 18/05/2021 - Início das entrevistas com os grupos

²Os países sem este valor informado devem ser descartados do resultado. Utilize a média do valor para o mês.