Guilherme Paleari

Plano de Análise de Dados – Análise com SQL

Etapa 1 - Definição do Objetivo

O objetivo deste trabalho é extrair informações relacionado ao Covid-19 durante o contexto da pandemia ao redor do mundo. As 10 perguntas abaixo deverão ser respondidas:

**# 1- Qual a média de mortos dos 5 primeiros países que apresentaram mais mortes por Covid-19?**

**1° United States com 274.030,92**

**2° Brasil com 182.642,16**

**3° Índia com 110.945,26**

**4° Mexico com 93.108,06**

**5° Peru com 77.217,53**

**# 2- Qual data teve a maior proporção de mortes em relação ao total de casos no Brasil?**

**Na data de 12-05-2020 com 178.214 casos, 12.461 mortes e gerando um percentual de 6.99%.**

**# 3- Qual a proporção média entre o total de casos e a população das 5 primeiras localidades?**

**1° Andorra com 7,6%**

**2° Montenegro com 6,23%**

**3° San Marino com 6,17%**

**4° Czechia com 5,66%**

**5° Bahrain com 5,11%**

**Obs.: Pode parecer estranho os resultados, mas esses países ficaram com a proporção maior por terem uma população menor, sendo assim, se teve muitos casos de Covid-19, consequentemente a proporção ficou alta e claramente os resultados estão coerentes.**

**# 4- Considerando o maior valor do total de casos, quais os 3 países com a maior taxa de infecção em relação à população?**

**1° Andorra com 13.918 infectados e 18,01% em relação a população**

**2° Seychelles com 15.857 infectados e 16,12% em relação a população**

**3° Montenegro com 100.327 infectados e 15,97% em relação a população**

**# 5- Quais os continentes com o maior número de mortes?**

**1° North América com 605.526**

**2° South América com 524.417**

**3° Asia com 402.005**

**4° Europa com 135.637**

**5° África com 61.840**

**6° Oceania com 910**

# As questões 6, 7, 8, 9 e 10 são questões com análises mais complexas e as consultas estão em Scripts no arquivo **Consultas\_Covid19.sql.**

**# 6- Qual o número de novos vacinados e a média móvel de novos vacinados ao longo do tempo por localidade? Considerando apenas os dados da América do Sul.**

**# 7- Qual o número de novos vacinados e o total de novos vacinados ao longo do tempo por continente? Considerando apenas os dados da América do Sul.**

**# 8- Qual o número de novos vacinados e o total de novos vacinados ao longo do tempo por continente? Considerando apenas os dados da América do Sul.**

**# 9- Qual o percentual da população com pelo menos 1 dose da vacina ao longo do tempo? Considerando apenas os dados do Brasil.**

**# 10- Durante o mês de Maio/2021 o percentual de vacinados com pelo menos uma dose aumentou ou diminuiu no Brasil?**

Etapa 2 - Coleta e Extração de Dados

Os dados estão disponíveis em formato CSV e foram extraídos da fonte conforme indicado no arquivo **CarregandoDadoseCriandoColunas\_Covid19.sql**. Os dados foram carregados via linha de comando no SGBD MySQL.

Etapa 3 - Escolha das Ferramentas

Para este trabalho usarei Linguagem SQL uma vez que os dados estão estruturados e disponíveis em um banco de dados relacional.

Etapa 4 - Aplicação das Técnicas de Análise

Para responder às perguntas de negócio usei a função window (quando for necessário cálculo ao longo do tempo) e calculei média e desvio padrão para explicar a variabilidade dos dados. Utilizei CTE para filtrar consultas com o resultado da função WIndow e utilizei também cálculos de média móvel.

Etapa 5 - Entrega dos Resultados

Conclusão:

Embora os primeiros casos tenham ocorrido na Asia, a América do Norte foi o continente que mais sofreu com casos de Covid-19 seguido da América do Sul.

Os EUA foi o recordista por apresentar uma média de 274.030,92 mortes por Covid-19, entretanto deu início a sua vacinação em 21-12-2020, já outros países demoraram para começarem a se vacinar e foram aumentando seu número de casos de Covid-19.

O Brasil é um exemplo disso, estartou sua vacinação no dia 17-01-2021 mas devido ao atraso das vacinas e poucas doses, sofreu com o aumento de casos e número de mortes.