

# Análise da COVID-19 com SQL

---

Este projeto tem como objetivo realizar uma análise exploratória dos dados da pandemia da COVID-19 utilizando SQL Server. Através de queries bem estruturadas, investigamos o impacto do vírus ao longo do tempo em diferentes países e continentes, avaliando taxas de mortalidade, índices de vacinação e níveis de infecção populacional.

## Perguntas de Negócio (Insights)

- 1. Qual é a taxa de mortalidade por casos em Portugal?
- 2. Qual a proporção de pessoas infectadas em relação à população total de cada país?
- 3. Quais países apresentam a maior taxa de infecção populacional?
- 4. Quais continentes apresentaram maior número total de mortes?
- 5. Qual a taxa de mortalidade global por novos casos?
- 6. Qual a percentagem da população vacinada por país?
- 7. Como evolui a vacinação ao longo do tempo em cada país?

## Fonte dos Dados

Os dados utilizados neste projeto provêm do repositório da Our World in Data (<https://ourworldindata.org/coronavirus-source-data>), divididos em duas tabelas principais:

- CovidDeaths
- CovidVaccinations

## Ferramentas Utilizadas

- SQL Server
- Janela de consultas (Query Editor)
- Funções agregadas, CTEs, subqueries, tabelas temporárias e views

## Principais Análises e Código

### 1. Taxa de Mortes por Casos em Portugal

```
SELECT  
location,  
date,  
COALESCE(total_cases, '-') total_cases,  
COALESCE(total_deaths, '-') total_deaths,
```

```
ROUND((CAST(total_deaths AS FLOAT)/total_cases) * 100, 2)
DeathPercentage
FROM portfolio_project_covid..CovidDeaths
WHERE location LIKE '%Portugal%' AND continent IS NOT NULL
ORDER BY 1,2;
```

---

> Resultado: Em alguns momentos, a taxa de mortalidade em Portugal ultrapassou os 3%.

## 2. Casos por População (Mundo)

```
SELECT location,
       date,
       population,
       COALESCE(total_cases, '-') total_cases,
       ROUND((CAST(total_cases AS FLOAT)/population) * 100, 2)
CasePercentage
FROM portfolio_project_covid..CovidDeaths;
```

---

> Resultado: A percentagem varia drasticamente por país — alguns ultrapassam 20% da população infetada.

## 3. Países com Maior Percentagem de Infetados

```
SELECT location,
       population,
       MAX(total_cases) HighestInfectionCount,
       ROUND(MAX((total_cases/population)) * 100, 2)
PercentPopulationInfected
FROM portfolio_project_covid..CovidDeaths
GROUP BY location, population
ORDER BY 4 DESC;
```

---

> Destaque: Países europeus e da América Latina lideram o ranking de infeções proporcionais.

## 4. Total de Mortes por Continente

```
SELECT continent, MAX(CAST(total_deaths AS INT)) TotalDeathCount
FROM portfolio_project_covid..CovidDeaths
WHERE continent IS NOT NULL
GROUP BY continent
ORDER BY 2 DESC;
```

---

> Resultado: A Europa e a América lideram em números absolutos de mortes.

## 5. Taxa de Mortalidade Global

```
SELECT SUM(new_cases) total_cases,  
       SUM(CAST(new_deaths AS INT)) total_deaths,  
       ROUND(SUM(CAST(new_deaths AS INT)) / SUM(New_Cases) * 100,  
2) DeathPercentage  
FROM portfolio_project_covid..CovidDeaths  
WHERE continent IS NOT NULL;
```

---

> Resultado: A taxa de mortalidade global está em torno de 2%.

## 6. Evolução da Vacinação Mundial

```
SELECT dea.continent,  
       dea.location,  
       dea.date,  
       dea.population,  
       vac.new_vaccinations,  
       SUM(CONVERT(BIGINT, vac.new_vaccinations)) OVER  
(PARTITION BY dea.Location ORDER BY dea.location, dea.date)  
RollingPeopleVaccinated  
FROM portfolio_project_covid..CovidDeaths dea  
JOIN portfolio_project_covid..CovidVaccinations vac  
  ON dea.location = vac.location AND dea.date = vac.date  
WHERE dea.continent IS NOT NULL  
ORDER BY 2,3;
```

---

> Resultado: Podemos visualizar a tendência crescente da vacinação, especialmente a partir de 2021.

## 7. Percentagem da População Vacinada com CTE

```
WITH Pop_vs_Vac (continent, location, date, population,  
new_vaccinations, RollingPeopleVaccinated)  
AS (  
  SELECT dea.continent,  
         dea.location,  
         dea.date,  
         dea.population,  
         vac.new_vaccinations,  
         SUM(CONVERT(BIGINT, vac.new_vaccinations)) OVER  
(PARTITION BY dea.Location ORDER BY dea.location, dea.date)  
RollingPeopleVaccinated
```

```

FROM portfolio_project_covid..CovidDeaths dea
JOIN portfolio_project_covid..CovidVaccinations vac
  ON dea.location = vac.location AND dea.date = vac.date
WHERE dea.continent IS NOT NULL
)
SELECT *, ROUND((RollingPeopleVaccinated / Population) * 100, 2)
PercentagePeopleperPopulation
FROM Pop_vs_Vac;

```

---

> Resultado: Permite visualizar a percentagem de vacinados acumulada ao longo do tempo por país.

## 8. Percentagem da População Vacinada com Tabela Temporária

```

DROP TABLE IF EXISTS #PercentPopulationVaccinated;
CREATE TABLE #PercentPopulationVaccinated (
  continent NVARCHAR(255),
  location NVARCHAR(255),
  date DATETIME,
  population NUMERIC,
  new_vaccinations NUMERIC,
  RollingPeopleVaccinated NUMERIC
);

INSERT INTO #PercentPopulationVaccinated
SELECT dea.continent,
  dea.location,
  dea.date,
  dea.population,
  vac.new_vaccinations,
  SUM(CONVERT(INT,vac.new_vaccinations)) OVER (PARTITION BY
  dea.location ORDER BY dea.location, dea.date)
FROM portfolio_project_covid..CovidDeaths dea
JOIN portfolio_project_covid..CovidVaccinations vac
  ON dea.location = vac.location AND dea.date = vac.date;

SELECT *, (RollingPeopleVaccinated / Population) * 100 AS
VaccinationPercentage
FROM #PercentPopulationVaccinated;

```

---

> Resultado: Alternativa usando temp table para calcular a percentagem vacinada.

## 9. Criação de View com Percentagem Vacinada

```
CREATE VIEW PercentPopulationVaccinated AS  
SELECT dea.continent,  
       dea.location,  
       dea.date,  
       dea.population,  
       vac.new_vaccinations,  
       SUM(CONVERT(INT, vac.new_vaccinations)) OVER (PARTITION  
BY dea.location ORDER BY dea.location, dea.date)  
RollingPeopleVaccinated  
FROM portfolio_project_covid..CovidDeaths dea  
JOIN portfolio_project_covid..CovidVaccinations vac  
ON dea.location = vac.location AND dea.date = vac.date  
WHERE dea.continent IS NOT NULL;
```

---

> Resultado: View reutilizável para análises futuras sobre a percentagem da população vacinada.

### Conclusão

Este projeto demonstra a capacidade de explorar grandes volumes de dados com SQL, criando insights relevantes sobre a pandemia da COVID-19. Utilizamos técnicas avançadas como CTEs, tabelas temporárias, funções de janela (window functions) e views, para realizar cálculos eficientes e reutilizáveis.

### Autor

Desenvolvido por Joelson Baptista

Projeto pessoal para portfólio na área de Data Analytics