Processamento Analítico de Dados

Trabalho - Segunda Parte Prof.^a Dr.^a Cristina Dutra de Aguiar PAE João Pedro de Carvalho Castro

Alunos:

•	Dalton Hiroshi Sato	11275172
•	Lucas Yuji Matubara	10734432
	Sávio Duarte Fontes	10737251

Introdução

Com dados provenientes do SUS, e de estações meteorológicas, o grupo buscou encontrar uma possível correlação entre o clima (umidade) e a ocorrência de doenças respiratórias (e seus sintomas).

Com uso das ferramentas Python 3 (versão 3.9.12 e 3.10.5), Pandas, NumPy, PostgreSQL e pgAdmin 4, os dados foram filtrados, organizados, armazenados e preenchidos na base de dados.

Alterações entre a primeira e segunda parte

Os campos sobre a vacinação não foram preenchidos corretamente pelos usuários do sistema, e apesar da existência de um dicionário de dados e de uma documentação sobre as funcionalidades de cada campo, não havia um padrão no preenchimento dessas colunas, gerando uma grande confusão aos seus valores e que demandariam grande esforço e tempo para tentar filtrá-las adequadamente.

Por essa razão, as dimensões relativas à vacinação foram incorporadas na dimensão paciente.

Alterações entre a primeira e segunda parte

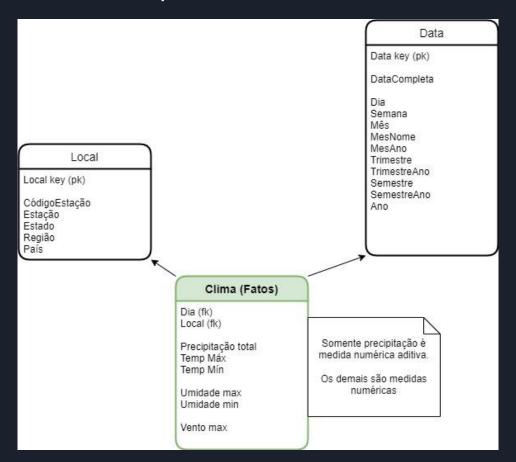
Sobre a dimensão em comum Local, não havia consistentemente uma maneira de relacionar ambos os esquemas estrelas de Clima e de Saúde. Isso ocorre em função de diferentes referências que ambas as fontes de dados continham: o Clima possuía estações de medição que cobriam diversas cidades, e cidades que possuíam diversas estações.

Além disso, Saúde não fazia referência à sua cidade, e perante a existência de mais de 5000 municípios no Brasil, outra tarefa extremamente trabalhosa surgiu.

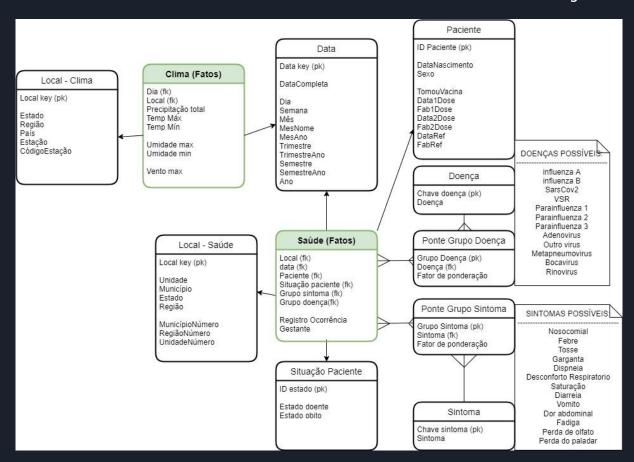
DOENCAS POSSÍVEIS: Local influenza A Local key (pk) influenza B SarsCov2 Unidade Municipio Parainfluenza 1 Estado Parainfluenza 2 Região Parainfluenza 3 Adenovirus MunicipioNumero Doenca Outro virus RegiãoNumero Metapneumovirus Chave doença (pk) UnidadeNumero Bocavirus Doenca Rinovirus Situação Paciente ID estado (pk) Ponte Grupo Doença Saúde (Fatos) Estado doente Grupo Doença (pk) Estado obito Doença (fk) Local (fk) Fator de ponderação data (fk) Paciente (fk) Situação paciente (fk) Grupo sintoma (fk) Ponte Grupo Sintoma Grupo doença(fk) Grupo Sintoma (pk) Registro Ocorrência Sintoma (fk) Gestante Fator de ponderação Data Data key (pk) DataCompleta Sintoma Dia Semana Chave sintoma (pk) Mês Sintoma Paciente MesNome MesAno SINTOMAS POSSÍVEIS ID Paciente (pk) Trimestre TrimestreAno DataNascimento Nosocomial Semestre Sexo Febre SemestreAno Tosse Ano TomouVacina Garganta Dispneia Data1Dose Fab1 Desconforto Respiratorio Data2Dose Saturação Fab2 Diarreia DataRef Vomito FabRef Dor abdominal Fadiga Perda de olfato Perda do paladar

Esquemas estrelas - Saúde

Esquemas estrelas



Constelação de fatos



Consultas

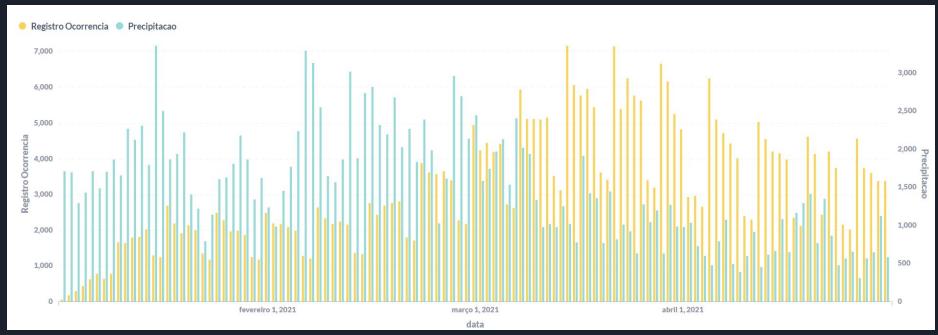
- Slice and Dice: Verificar entradas mensais em hospital que aconteceram na cidade de São Paulo no ano de 2021
- Drill-Across: Verificar o número de infectados (qualquer vírus) e de precipitação diários durante os 4 primeiros meses de 2021;
- Roll-Up: Verificar precipitação total por mês por região.
- **Drill-Down**: Verificar por semana quantos foram os gestantes ao longo do tempo
- **Pivot**: Verificar número de contágios por semana por estado, trocando as perspectivas de tempo e local.

Consultas - Slice and Dice



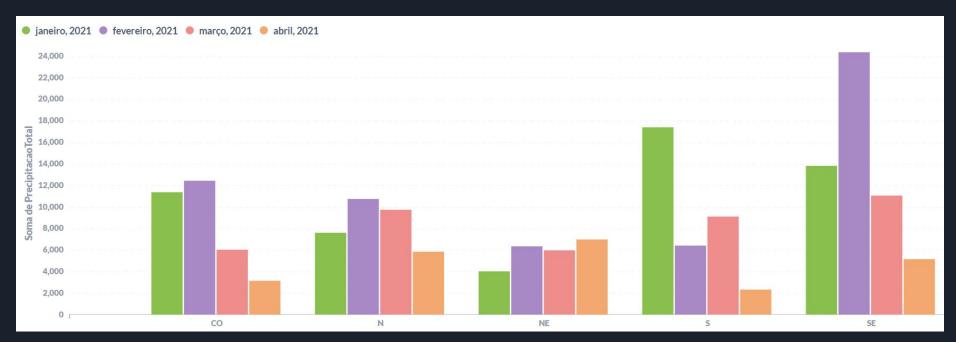
Variação de contágios mensais em São Paulo capital

Consultas - Drill-Across



Variação de precipitação e de infectados ao longo dos 4 primeiros meses de 2021.

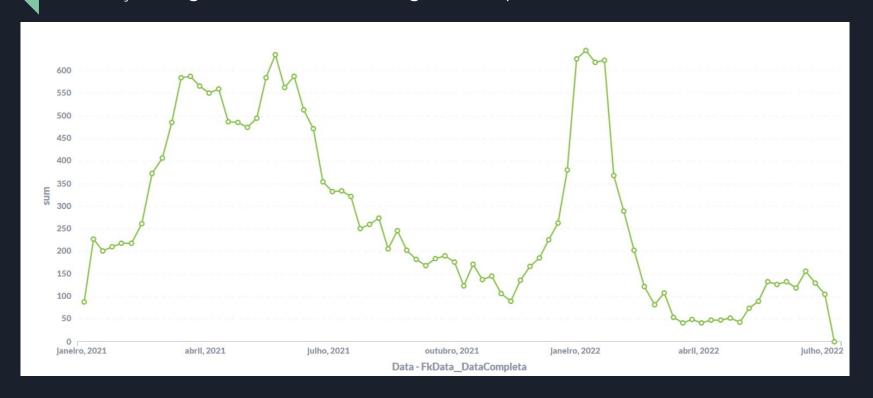
Consultas - Roll-up



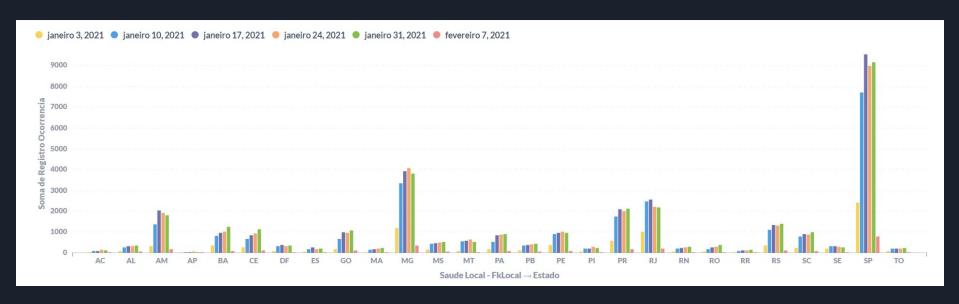
Precipitação total em cada região entre Janeiro e Abril de 2021

Consultas - Drill-down

Variação de gestantes diários ao longo do tempo



Consultas - Pivot



Variação de contágios nas 6 primeiras semanas de 2021, por estado

Referências bibliográficas

Datasus. "SRAG 2021 e 2022 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19 -

OPENDATASUS." *OPENDATASUS*, 2022, https://opendatasus.saude.gov.br/dataset/srag-2021-e-2022. Accessed 18 May 2022.

INMET. "Climate Weather Surface of Brazil - Hourly." Kaggle, outubro 2021,

https://www.kaggle.com/datasets/PROPPG-PPG/hourly-weather-surface-brazil-southeast-region. Accessed 18 May 2022.

Slides apresentados em aula