

Trabalho Final

MC536

Cavaleiros da Query Estruturada

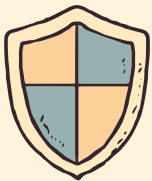
Willian Takayuki Ozako - 207267

Frederico Meletti Rappa - 216047

João Alberto Moreira Seródio - 218548



Sumário



Introdução

- Motivação
- Objetivos



Etapa 3

- Modelo Conceitual
- Modelo Lógico
- Tratamento dos Dados
- Queries

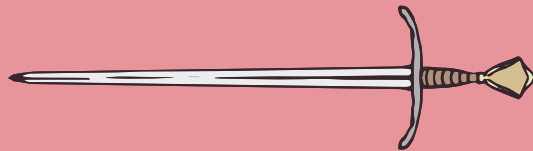


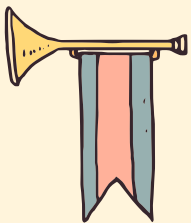
Etapa 4

- Modelo de Grafos
- Queries
- Visualização

01

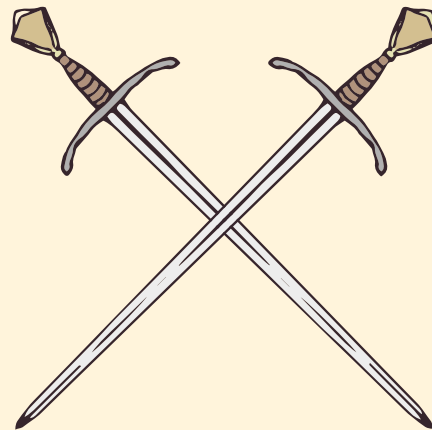
Introdução





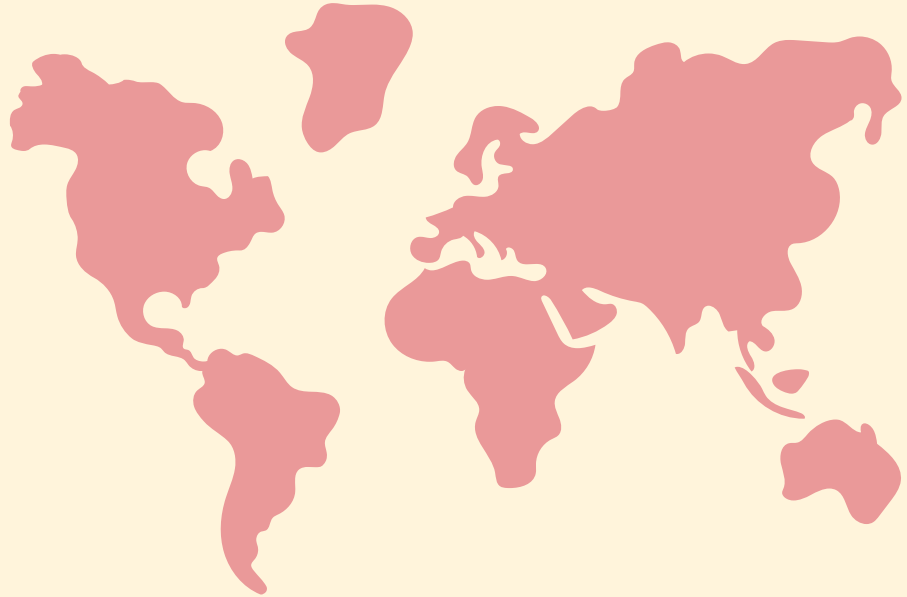
Motivação

Visualizar possíveis correlações entre o número de casos de DSTs, aspectos socioeconômicos e políticas públicas, a fim de determinar fatores chave no controle de novos casos.



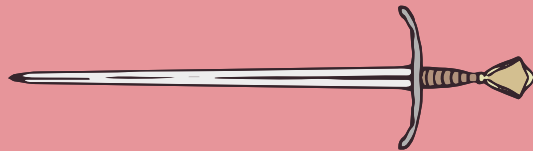
Análise de DSTs e aspectos socioeconômicos

- Foco em casos de infecção por HIV
- Busca por correlação entre quantidade de novas infecções, aspectos socioeconômicos no país (e.g. grau de escolaridade, renda) e políticas públicas



02

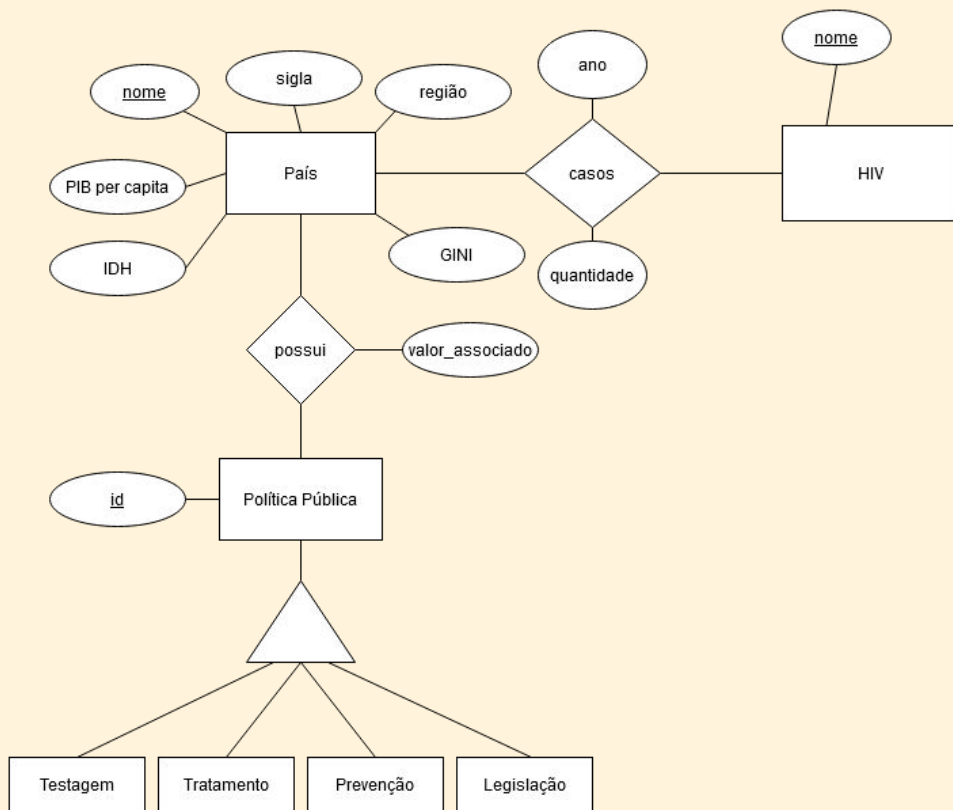
Etapa 03



Análise com Modelo Relacional

(SQL + Python)

Modelo Conceitual



Modelo Lógico Relacional

País							
nome	sigla	região	wb_class	GINI (x 100)	IDH (x 1000)	PIB per Capita	População
Algeria	DZA	Africa	Upper middle income	353	736	14610.0	40400000

Política		
id	política	tipo
P1	Country adopted 2015 WHO guidelines ...	law

Infecções por HIV				
país (ou GLOBAL)	ano	mínimo	nominal	máximo
ARG	2000	5200.00	6200.00	7400.00

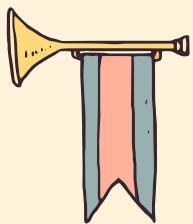
PolíticaPaís				
país	política	éNumérico	valor	valorNumérico
Afghanistan	P1	False	Yes, partially	

Testagem				
país	Política AT	Política BT	...	Política ATZ
Afghanistan	Yes	Yes		No

Tratamento				
país	Política ATr	Política BTr	...	Política ATrZ
Afghanistan	Yes	Yes		Countrywide

Prevenção				
país	Política AP	Política BP	...	Política APZ
Afghanistan	Yes	Yes		Yes

Legislação				
país	Política AL	Política BL	...	Política ALZ
Afghanistan	Yes, patially	No		Countrywide



Tratamento dos Dados

Utilizou-se Jupyter Notebooks em Python para confeccionar as tabelas utilizadas no modelo lógico tabular.

Os dados das fontes de dados utilizadas são tratados e então inseridos em um DataFrame, que após completado o processo é exportado como CSV.

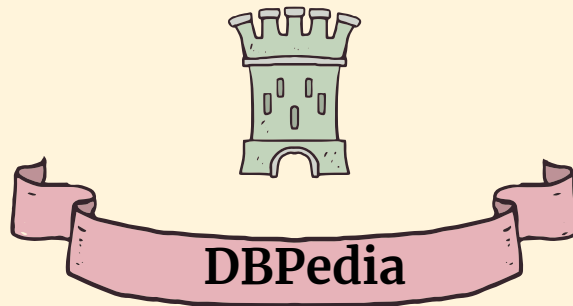


Tabela de Países



Lista de Países (170)

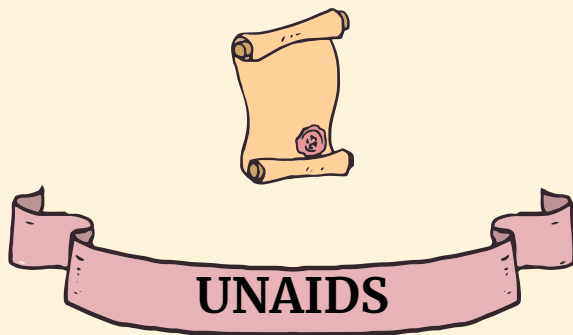
- Nome
- Label
- Região
- Classificação World Bank



Utilizando a lista de países,
adicionamos os seguintes campos:

- Gini (124)
- PIB per Capita PPP (138)
- IDH (137)
- População (116)

Tabelas de Políticas



Divisão das Políticas em 4 tipos:

- Testagem
- Tratamento
- Legislação
- Prevenção



Criação de tabelas que facilitam queries:

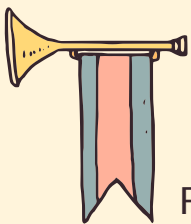
- Políticas (id, política, tipo)
- País-Políticas (país, políticaID, éNumérico, valor, valorNumérico)

Tabela de Infecções



HIV_0000000026 GHO

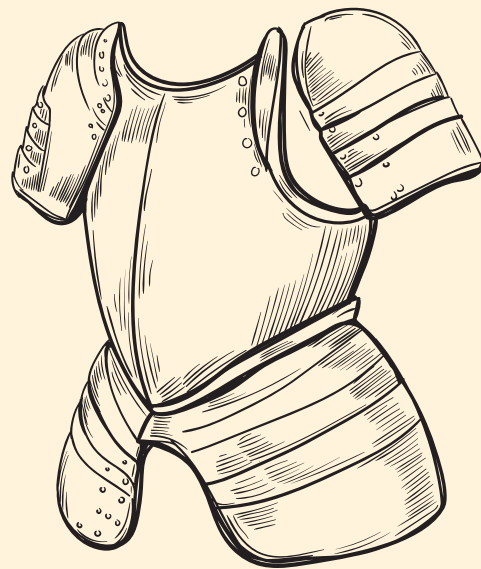
- Região/País
- Ano
- Nominal
- Mínimo
- Máximo



Queries

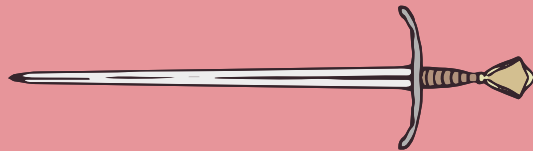
Foram realizadas 7 queries buscando as seguintes análises:

- Qtd de Políticas x Renda QR
- Região x Novas Infecções QR
- Socioeconômico x Novas Infecções QRR2
- Tratamento x Renda QR
- Distribuição Preservativos x Região QR
- Preservativos por Habitante QR
- Estratégia x Região QR



03

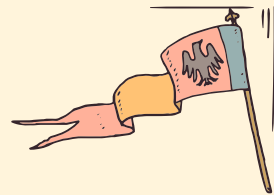
Etapa 04



Análise com Modelo de Grafos

(Neo4j + Cytoscape)

Problema: como agrupar diferentes países e diferentes DSTs?



Como agrupar diferentes países?

- Modelo de grafos
- Critérios socioeconômicos?
- Políticas públicas semelhantes?

Como agrupar diferentes DSTs?

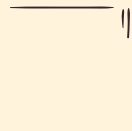
- Agentes causadores?

Ferramentas utilizadas:

- Banco de dados - Neo4j
- Algoritmo de agrupamento: Louvain
- Visualização: Cytoscape



Agrupando países com dados socioeconômicos

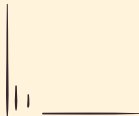


Que combinação utilizar?

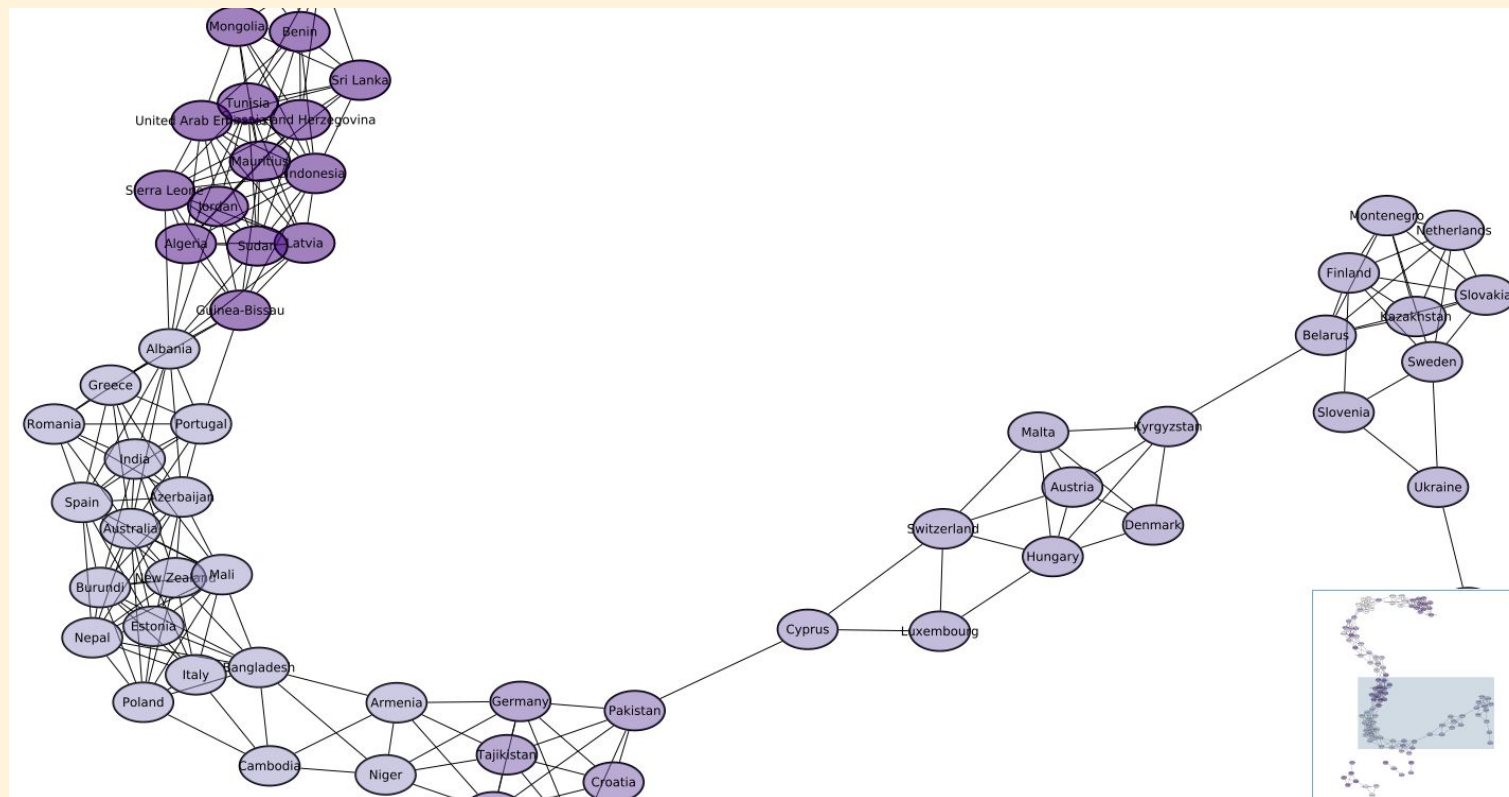
- PIB per Capita PPP
- IDH
- Gini
- IDH + Gini
- IDH + Gini + PIB per Capita

Algoritmo:

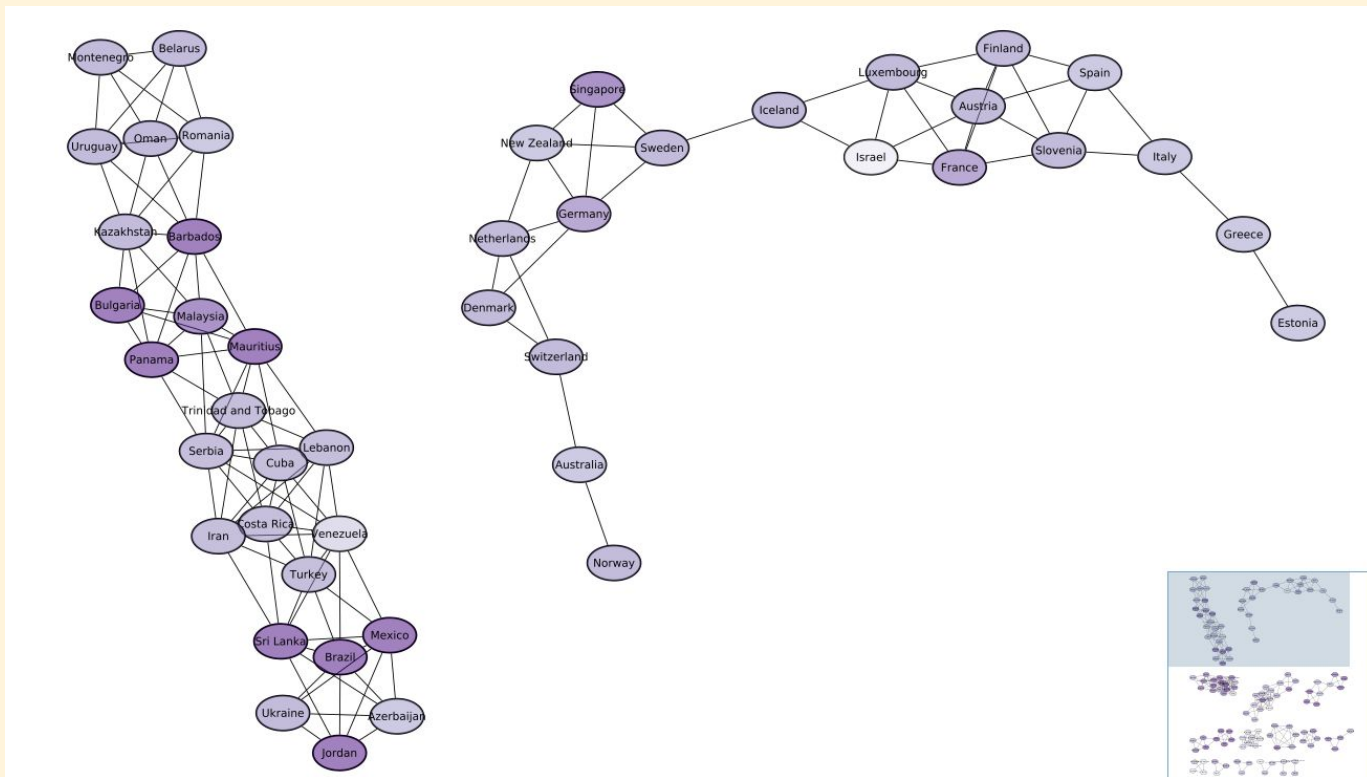
- Se dois países tem um dado socioeconômico semelhante, cria uma aresta entre eles e quanto maior a proximidade do valor, maior o peso
- Aplicar Louvain com e sem peso
- Verificar manualmente se o agrupamento faz sentido

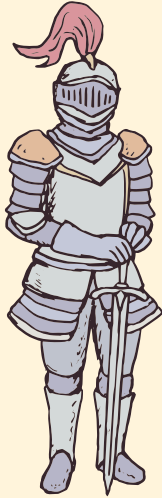


Gini + IDH com Peso - grafo de Gini



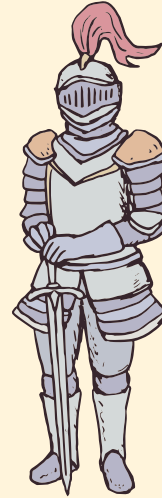
Gini + IDH com Peso - grafo do IDH





Obrigado!

Willian Takayuki Ozako - 207267
Frederico Meletti Rappa - 216047
João Alberto Moreira Seródio - 218548



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Please keep this slide for attribution.

