

Dataset de índices econômicos mundiais

Membro	RA
Leonardo Rener de Oliveira	201270
César Guedes Carneiro	261031
Matheus Silva de Deus	241882

Grupo
Invest Bank - INVB3

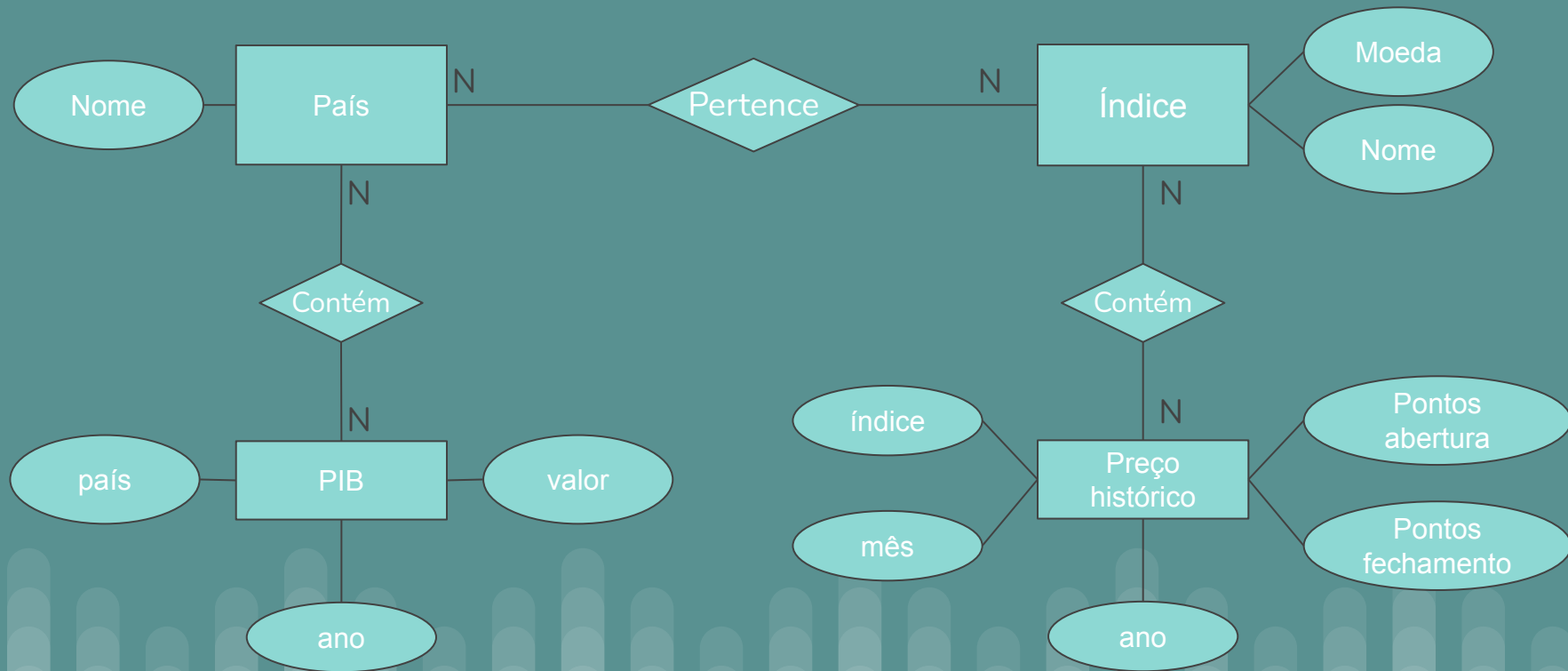
Motivação e Contexto Gerador

O tema do nosso dataset é a evolução histórica de índices econômicos em conjunto com o PIB do país.

No contexto atual existem diversos datasets com tema econômico, entretanto são em sua maioria pagos ou de difícil acesso.

Por isso nossa motivação está em criar uma forma simples e gratuita de estudar os principais índices de investimento econômicos do mundo e suas relações com a economia do país de origem. Buscamos realizar essa comparação através da relação Bolsa x PIB (Produto Interno Bruto), o PIB representa a soma de todos os bens e serviços finais produzidos numa determinada região, durante um período determinado.

Modelo Conceitual



Modelos Lógicos

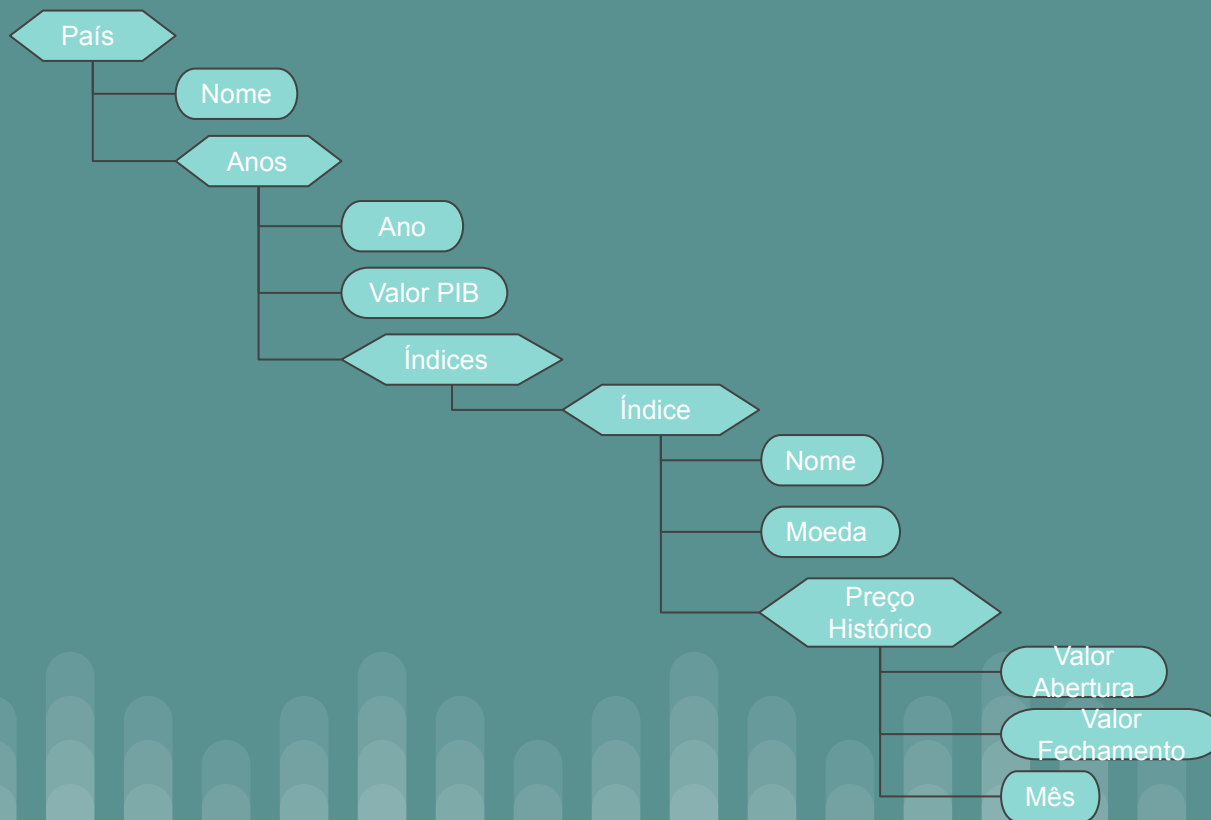
- Modelo Relacional
- Modelo de Documentos

Modelo Relacional

- País(nome)
- PIB(país, ano, valor)
CHE: país para País
- Índice_Econômico(nome, moeda)
- Pertence(país, índice)
CHE: país para País e índice para Índice_Econômico
- Histórico_de_preços(índice, ano, mês, pontosAbertura, pontosFechamento)
CHE: índice para Índice_Econômico

Chaves primárias são representadas da forma chave-primária

Modelo de Documentos



Fontes de Dados Utilizadas

- Investing.com : <https://br.investing.com/>
- IMF : <https://data.imf.org/regular.aspx?key=63122827>

Utilizamos operações de extração, integração, tratamento e transformação dos dados.

Extração dos Dados

A extração dos dados foi feita em sua maior parte através de um API scraper, através de uma biblioteca do Python chamada Investpy e scripts implementados por nós. Utilizamos tal scraper para capturar os países, moeda, nome de índices e suas pontuações históricas ao longo dos anos.

Além disso, foi também feita uma extração manual de dados, para a coleta da planilha do GDP dos países listados no dataset.

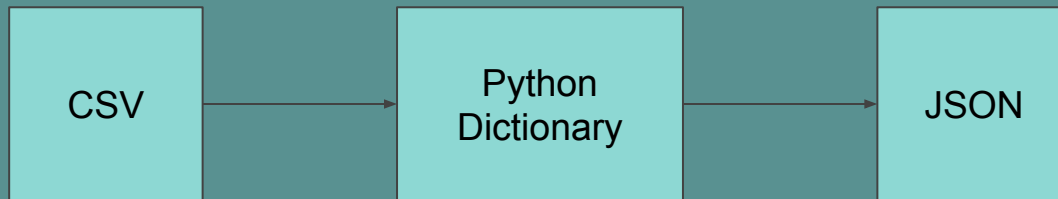
Integração de Dados

Realizamos a integração dos dados da Investing com a International Monetary Fund (IMF), através do relacionando dos índices financeiros com o PIB dos países.

Tratamento dos Dados

Foi necessário um tratamento dos dados de PIB recebidos da International Monetary Fund (IMF), isso porque alguns países não possuíam o valor do PIB durante todo o período apurado (2015 a 2020). Realizamos a substituição desses dados faltante que continham o valor “...” para o valor null, assim tornando o campo PIB somente de valores numéricos ou nulos.

Transformação dos Dados



Perguntas que podem ser respondidas com o dataset

- Qual a média histórica (2015-2020) para o PIB de todos os países ?
- Quais países, obtiveram desempenho maior do que a média histórica do PIB, no ano de 2019?
- Qual a previsão do GDP de um país com base nos valores de abertura e fechamento do índice econômico do mês de um determinado mês?
- Qual foi o índice com maior variação em sua pontuação histórica, no período de 1 mês, de 2015 a 2020? E a qual país tal índice pertence?

Query

- Qual a média histórica (2015-2020) para o PIB de todos os países ?

MongoDB
(python)

```
a = countries.find(
    {},
    {"years.GDP":1,"_id":0}
)
rep = 0
som = 0
for i in a:
    for j in i['years']:
        if(j):
            if isinstance(j['GDP'],float):
                som += j['GDP']
            rep += 1
som = som/rep
print(som)
```

Query

- Quais países, obtiveram desempenho maior do que a média histórica do PIB, no ano de 2019?

MongoDB
(python)

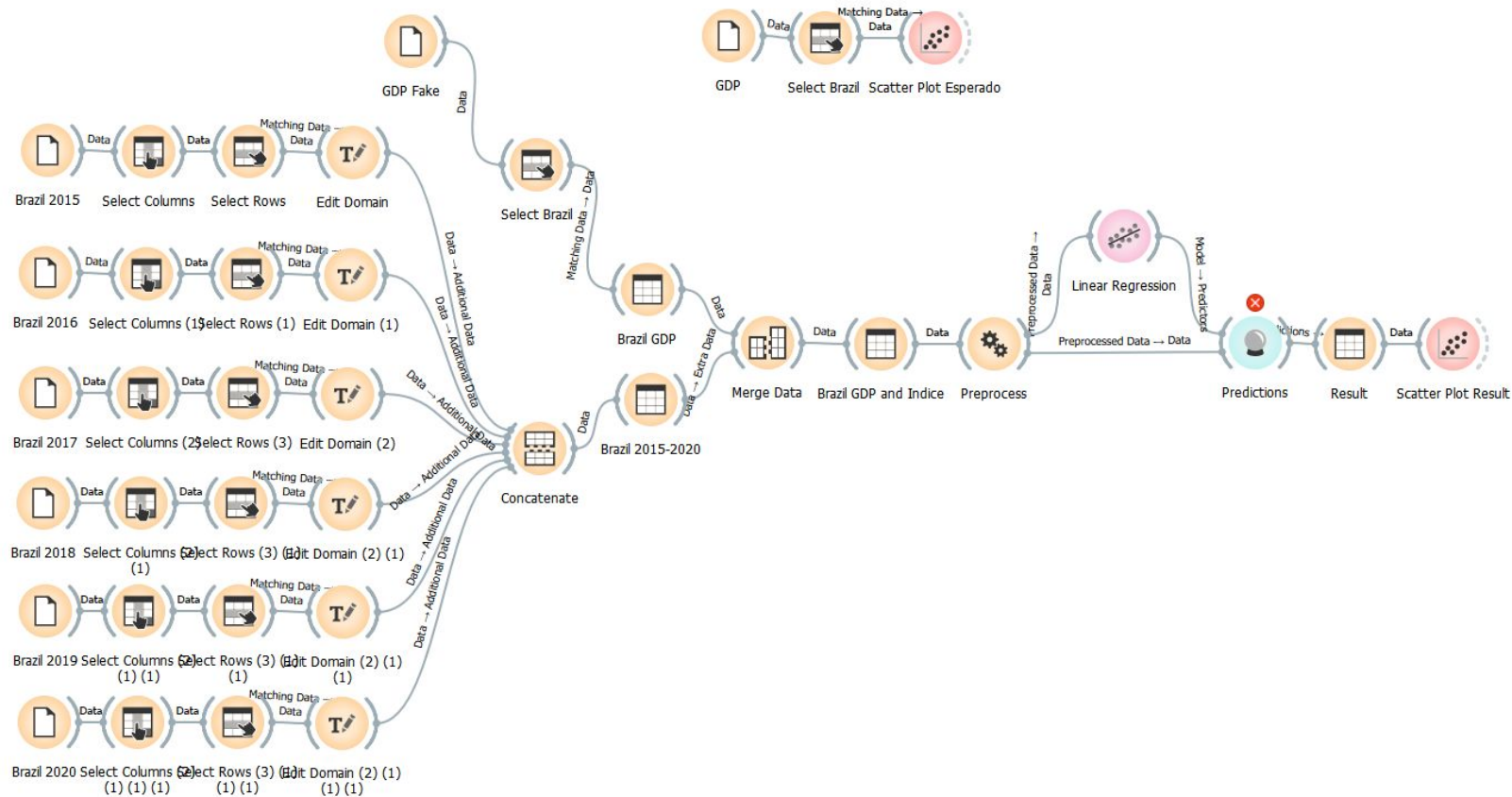
```
b = countries.find(
    { "years": { "$elemMatch":
        { "year": 2019,
          "GDP": { "$gt":
            244682543.798},
        }
    }
},
    {'name.$': 1, "_id": 0}
)

for i in b:
    print(i)
```

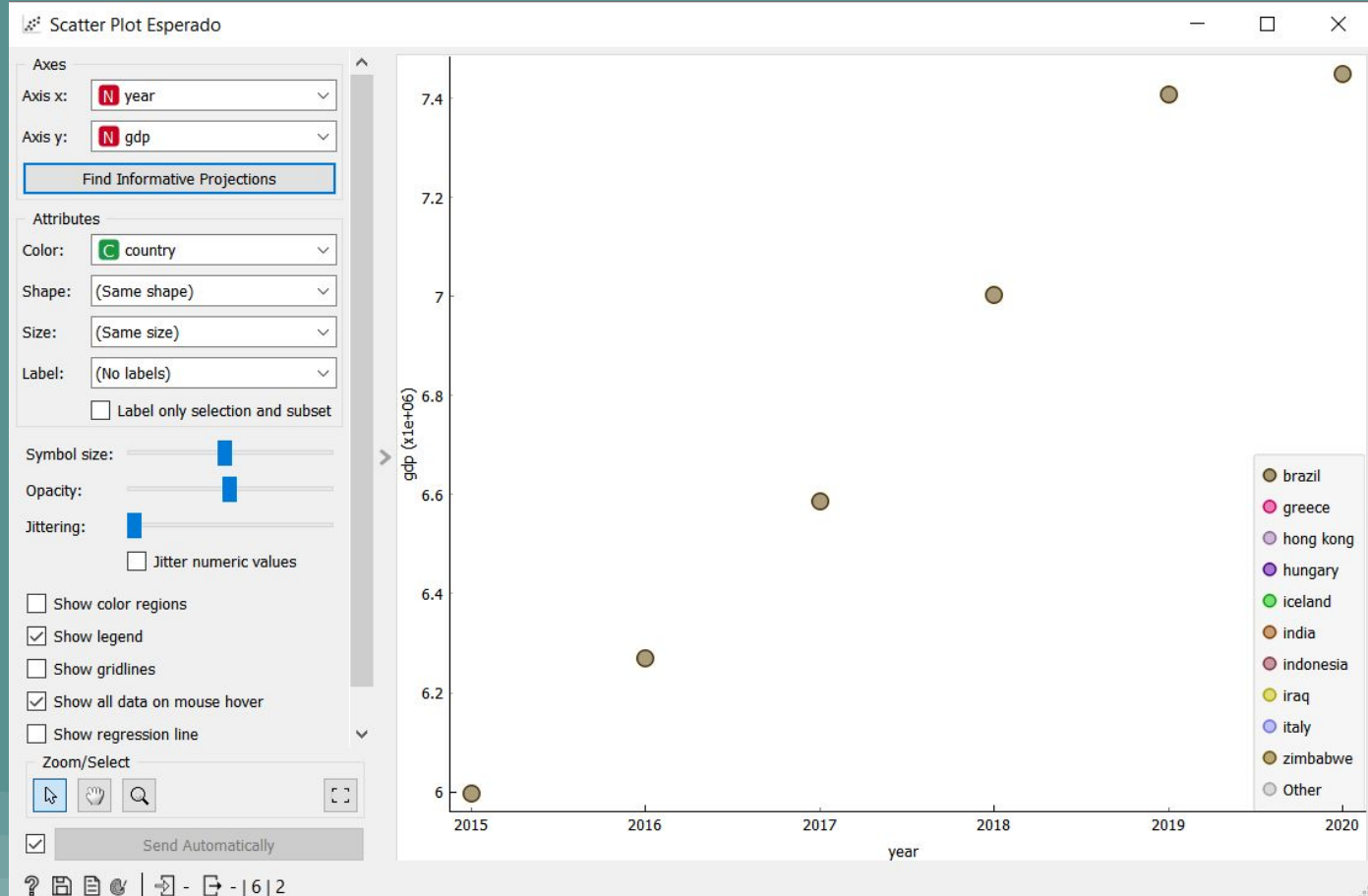
Query

- Qual a previsão do GDP de um país com base nos valores de abertura e fechamento do índice econômico do mês de um determinado mês?

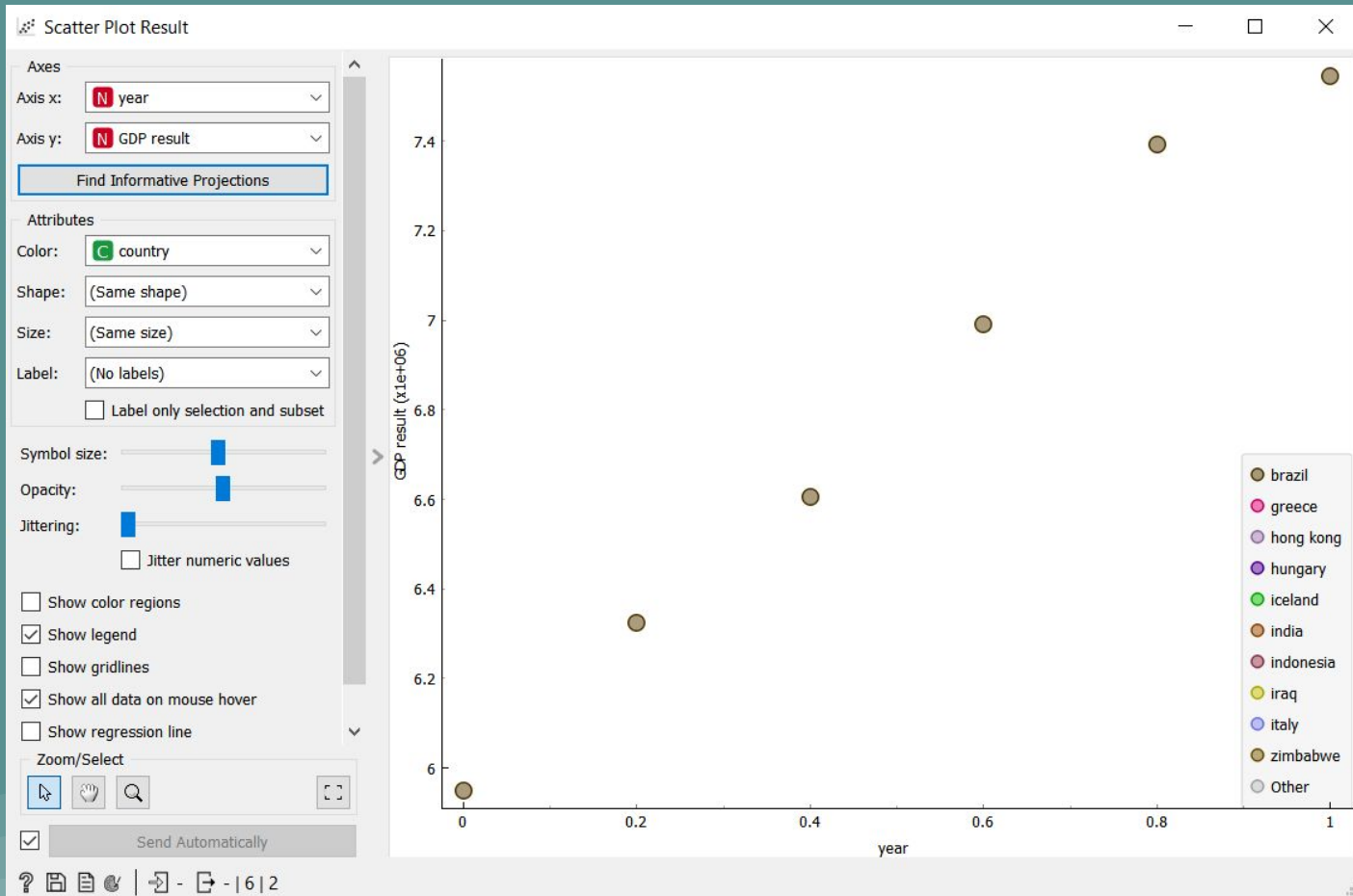
WorkFlow



Resultado Esperado



Resultado Esperado



Query

- Qual foi o índice com maior variação em sua pontuação histórica, no período de 1 mês, de 2015 a 2020? E a qual país tal índice pertence?

Database queries

Qual foi o índice com maior variação em sua pontuação histórica, no período de 1 mês, de 2015 a 2020? E a qual país tal índice pertence?

In [25]: %sql

```
SELECT
    indFC.Country, histD.Indice_name, histD.Year, histD.Month, ABS(histD.Close - histD.Open) AS variation
FROM
    HISTORICAL_DATA AS histD, INDICE_FROM_COUNTRY AS indFC
WHERE variation = (
    SELECT
        MAX(ABS(histD.Close - histD.Open))
    FROM
        HISTORICAL_DATA AS histD
) AND indFC.Indice_name = histD.Indice_name
```

* sqlite:///temp_db.db
Done.

Out[25]:

Country	Indice_name	Year	Month	variation
jamaica	JSE Market	2020	3	52811.0

Datasets

Os datasets está disponível em:

<https://github.com/Matheuss1/Trabalho-BD>

Obrigado!