

Dados Bitcoin

Klayvert G. A. R. de Carvalho

Fonte dos Dados: <https://mempool.space/pt/>

Aplicação OLAP: <https://tinyurl.com/mrs9adxh>

Github: <https://github.com/KlayRodrigs/projeto-sad>

ETAPA 1 - PLANEJAMENTO

1. Contextualização

Valores do Bitcoin - análise financeira e de performance

2. Escopo/objetivo do Data Mart

Analisar as principais variáveis econômicas e temporais que compõem o Bitcoin ao longo de 15 anos.

3. Arquitetura Tecnológica

Fonte dos Dados



ETL



Aplicação OLAP



4. Processo

- Planejamento;
- Levantamento;
- Modelagem dimensional;
- ETL;
- Projeto físico do Data Mart
- Aplicação OLAP.

4. Abordagem

Foi utilizado a *bottom-up*, visto que o foco principal seria construir um data mart pequenos e em algum outro momento centralizá-lo em um data warehouse. A modelagem usada foi o esquema estrela (*star-schema*) que permite realizar consultas de forma mais performática.

5. Usuários

Pessoas, empresas ou órgãos governamentais com interesses em criptoativos.
Ex: investidores, receita-federal e entusiastas.

ETAPA 2 - LEVANTAMENTO DE NECESSIDADES

6. Consultas de apoio à decisão

- Maiores máximas e menores mínimas de cada ano;
- Preço histórico total;
- Variações em espaços de Halving, 4 em 4 anos;
- Somatório dos meses;

7. Indicadores de análise

- Volume financeiro em dólares movimentado.

ETAPA 3 - MODELAGEM

9. Modelo Relacional

dim_date		fact_price_daily	
date_key	int	date_key	int NN
full_date	date NN	open_price	decimal(38,10) NN
year	int NN	high_price	decimal(38,10) NN
month	int NN	low_price	decimal(38,10) NN
day	int NN	close_price	decimal(38,10) NN
		volume	decimal(38,10) NN
		load_ts	datetime NN
		source_ref	varchar(200)



dbdiagram.io

10. Modelo Dimensional

- A. Área de Negócios
Finanças
- B. Processo
Levantamento dos valores do criptoativo Bitcoin
- C. Granularidade
Ano x Mês x Dia

10. Modelo Dimensional

- D. Atributos e Hierarquia das Dimensões
dim_date (dia > mês > ano > full_date)

10. Modelo Dimensional

E. Estimativa de espaço (para 5 anos)

Tamanho atual do banco:

Total por linha (fact_price_daily) = $4 + 100 + 8 + 52 + 20 = 184$ bytes/linha

Dias em 5 anos = $5 \times 365 = 1825$ dias

Resultado = 0.32MB em 5 anos (1 linha/dia)

Exemplo = 100 instrumentos/dia

1. linhas_5anos = $1825 \times 100 = 182.500$

2. bytes_total = $182.500 \times 184 = 33.580.000$ bytes

3. MB = $33.580.000 / 1.048.576 \approx 32,02$ MB

Resultado: $\approx 32,02$ MB (5 anos, 100 linhas/dia)

ETAPA 4 - PROJETO FÍSICO DO DB

11. Modelo Relacional do Data Mart (físico)

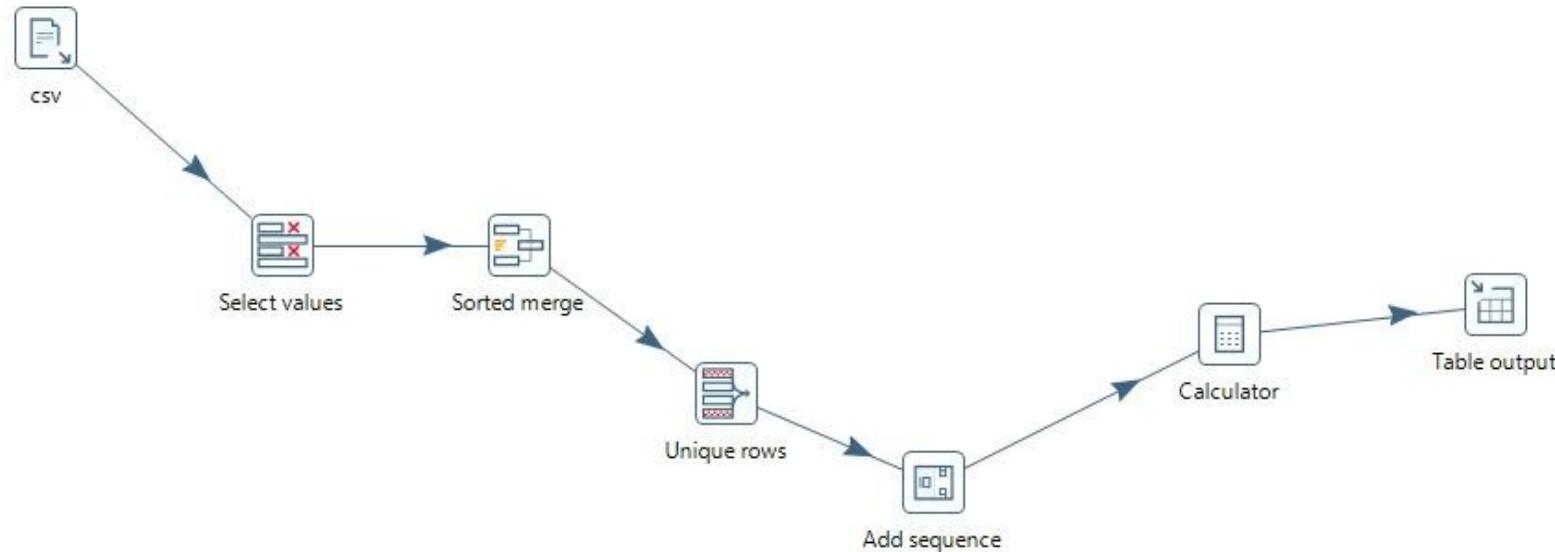
```
CREATE TABLE IF NOT EXISTS public.fact_price_daily
(
    price bigint,
    "Open" bigint,
    high numeric(11, 2),
    low numeric(11, 2),
    change character varying(100) COLLATE pg_catalog."default",
    sk_tempo double precision
);
END;
```

11. Modelo Relacional do Data Mart (físico)

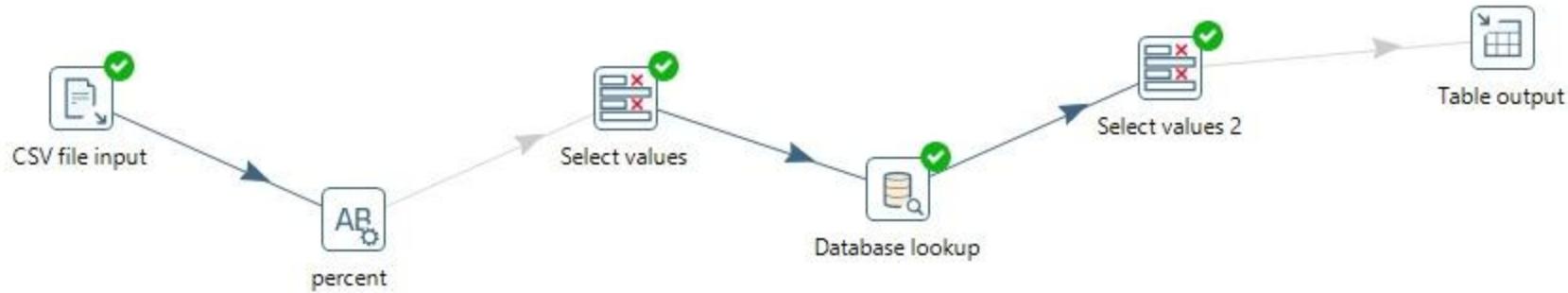
```
CREATE TABLE IF NOT EXISTS public.dim_time
(
    id bigserial NOT NULL,
    fulldate timestamp without time zone,
    sk_tempo double precision,
    year double precision,
    mounth double precision,
    day double precision
);
END;
```

ETAPA 5 - ETL

12. Plano de Carga da Dimensão Tempo

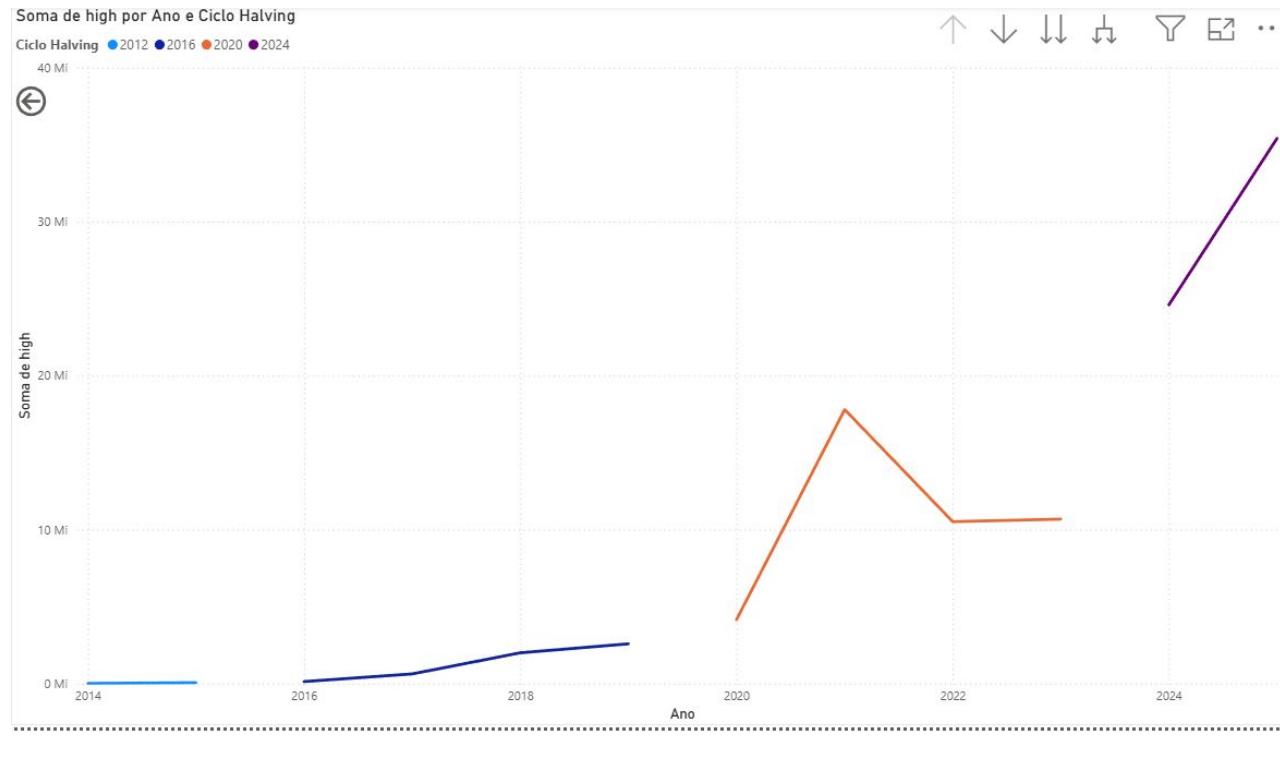


12. Plano de Carga da Fato

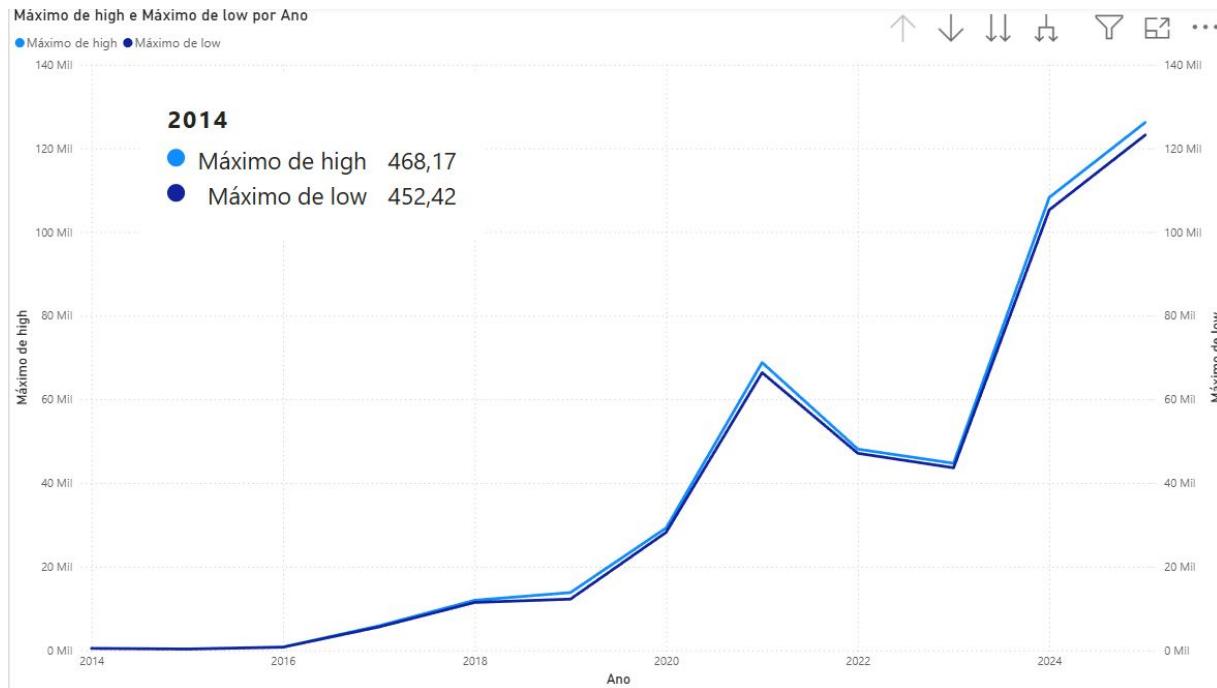


ETAPA 6 - APLICAÇÃO OLAP E PAINEL DE BORDO

13. Consulta OLAP 1 (Variação 4 em 4 anos)



13. Consulta OLAP 2 (Máximas e mínimas)



13. Consulta OLAP 3 (Preço total)



13. Consulta OLAP 4 (Informações totais)

Ano	Mês	Dia	price
2015	agosto	24	210
2015	agosto	18	211
2015	abril	14	219
2015	abril	26	219
2015	agosto	25	222
2015	abril	17	223
2015	abril	18	223
2015	abril	19	223
2015	junho	1	223
2015	junho	7	223
2015	abril	15	224
2015	junho	4	224
2015	abril	13	225
2015	abril	20	225

Referências

1. Dados fontes (abertos):
<https://mempool.space/>