

Rencana Persiapan Visualisasi

1. Kesiapan Data

Dataset Brazil Stock Market atau data pasar saham Brazil telah melalui beberapa tahapan, yakni:

1. ETL Pipeline

- a. Ekstraksi data dari file .csv.
- b. Transformasi dilakukan sesuai kebutuhan, seperti penyesuaian nama hari dan bulan serta menambahkan atribut historis, seperti *startedAt*, *endedAt*, *isActive*.
- c. Data dimuat ke dalam *Data Warehouse* DuckDB sebagai penyimpanan akhir untuk *query* dan visualisasi.

2. Pemeriksaan Kualitas Data

- a. Pemeriksaan nilai NULL untuk kolom penting.
- b. Validasi bahwa nilai numerik (*openValueStock*, *closeValueStock* dan lain-lain) tidak negatif.
- c. Deteksi duplikat berdasarkan kombinasi kunci.
- d. Pemeriksaan *foreign key* *keyCompany* agar valid dan konsisten antar tabel.

2. Rekomendasi Alat Visualisasi

Berikut adalah alat visualisasi yang direkomendasikan:

1. Metabase

- Gratis dan *open-source*.
- Bisa dihubungkan langsung ke DuckDB.
- Mendukung query SQL dan drag-and-drop.

2. Power BI / Tableau

- Tools *enterprise* dengan banyak fitur.
- Mudah dipresentasikan ke stakeholder bisnis.

3. Contoh Query

Berdasarkan dokumen kebutuhan bisnis, dataset pasar saham Brazil dapat digunakan untuk membantu investor memahami tren, risiko, dan peluang pasar saham Brazil. Beberapa analisis yang dilakukan berikut query-nya dapat meliputi:

- Mengetahui pola harga saham per-sektor dengan mengecek rata-rata harga penutupan per sektor tiap bulan/tahun

```
SELECT
  dt.yearTime,
  dt.monthTime,
```

```

dc.sectorCompany,
AVG(fs.closeValueStock) AS avg_close_price
FROM factStocks fs
JOIN dimTime dt ON fs.keyTime = dt.keyTime
JOIN dimCompany dc ON fs.keyCompany = dc.keyCompany
GROUP BY dt.yearTime, dt.monthTime, dc.sectorCompany
ORDER BY dt.yearTime, dt.monthTime, dc.sectorCompany;

```

- Mendapatkan informasi korelasi volume perdagangan terhadap harga saham (harga penutupan)

```

SELECT
dc.sectorCompany,
CORR(fs.quantityStock, fs.closeValueStock) AS volume_price_corr
FROM factStocks fs
JOIN dimCompany dc ON fs.keyCompany = dc.keyCompany
GROUP BY dc.sectorCompany;

```

- Mengetahui performa saham dengan melihat saham yang memiliki return tertinggi (selisih harian harga saham pembukaan ke penutupan)

```

SELECT
dc.stockCodeCompany,
dc.nameCompany,
AVG((fs.closeValueStock - fs.openValueStock) / fs.openValueStock) * 100 AS
avg_daily_return_percent
FROM factStocks fs
JOIN dimCompany dc ON fs.keyCompany = dc.keyCompany
GROUP BY dc.stockCodeCompany, dc.nameCompany
ORDER BY avg_daily_return_percent DESC
LIMIT 10;

```

- Mengetahui dampak peristiwa ekonomi terhadap harga saham, (pada contoh ini diambil peristiwa COVID)

```

SELECT
CASE WHEN dt.datetime < '2020-03-01' THEN 'Before COVID' ELSE 'After
COVID' END AS period,
AVG(fs.closeValueStock) AS avg_close_price
FROM factStocks fs
JOIN dimTime dt ON fs.keyTime = dt.keyTime
GROUP BY period;

```

- Mengetahui Anomali Perdagangan dengan mencari data yang bervolume lebih tinggi dibanding biasanya (misalnya di atas 95th percentile)

```
WITH volume_stats AS (  
  SELECT  
    fs.keyCompany,  
    PERCENTILE_CONT(0.95) WITHIN GROUP (ORDER BY  
fs.quantityStock) AS vol_95  
  FROM factStocks fs  
  GROUP BY fs.keyCompany  
)  
SELECT  
  dc.stockCodeCompany,  
  dt.datetime,  
  fs.quantityStock,  
  fs.closeValueStock  
FROM factStocks fs  
JOIN dimTime dt ON fs.keyTime = dt.keyTime  
JOIN dimCompany dc ON fs.keyCompany = dc.keyCompany  
JOIN volume_stats vs ON fs.keyCompany = vs.keyCompany  
WHERE fs.quantityStock > vs.vol_95  
ORDER BY fs.quantityStock DESC  
LIMIT 20;
```