# Diagrams and Design

## Fact Table:

| Column Name    | Data Type | Description                              |
|----------------|-----------|--|
| transaction_id | INT (PK)  | Unique identifier for each transaction.  |
| stock_id       | INT (FK)  | Foreign key to the Stock dimension table |
| date_id        | INT (FK)  | Foreign key to the Date dimension table  |
| open_price     | DECIMAL   | Opening price of the stock               |
| close_price    | DECIMAL   | Closing price of the stock.              |
| high_price     | DECIMAL   | Highest price of the stock.              |
| low_price      | DECIMAL   | Lowest price of the stock                |
| volume         | INT       | Number of shares traded                  |
| rsi            | DECIMAL   | Relative Strength Index                  |
| sma            | DECIMAL   | Simple Moving Average                    |
| bollinger_band | DECIMAL   | Bollinger Band value                     |

### **Dimension Tables:**

**Stock Dimension Table:** This table contains information about individual stocks (companies).

| Column Name  | Data Type | Description                                       |
|--------------|-----------|---|
| stock_id     | INT (PK)  | Unique identifier for each stock                  |
| stock_symbol | VARCHAR   | Stock ticker symbol (e.g., AAPL, TSLA)            |
| company_name | VARCHAR   | Name of the company (e.g., Apple Inc.)            |
| sector       | VARCHAR   | Sector the company belongs to (e.g., Technology)  |
| currency     | VARCHAR   | Currency in which the stock is traded (e.g., USD) |
| exchange     | VARCHAR   | Stock exchange (e.g.,<br>NASDAQ, NYSE)            |

**Date Dimension Table:** It contains a record for every day in your data's time range, with attributes like day, week, month, quarter, and year

| Column Name | Data Type | Description                       |
|-------------|-----------|-----------------------------------|
| date_id     | INT (PK)  | Unique identifier for each date.  |
| full_date   | DATE      | The full date (YYYY-MM-DD)        |
| day_of_week | VARCHAR   | Day of the week (e.g.,<br>Monday) |
| month       | VARCHAR   | Month (e.g., January)             |
| quarter     | VARCHAR   | Quarter (e.g., Q1, Q2, Q3, Q4)    |
| year        | INT       | Year (e.g., 2023)                 |
| is_holiday  | BOOLEAN   | Whether the date is a holiday     |

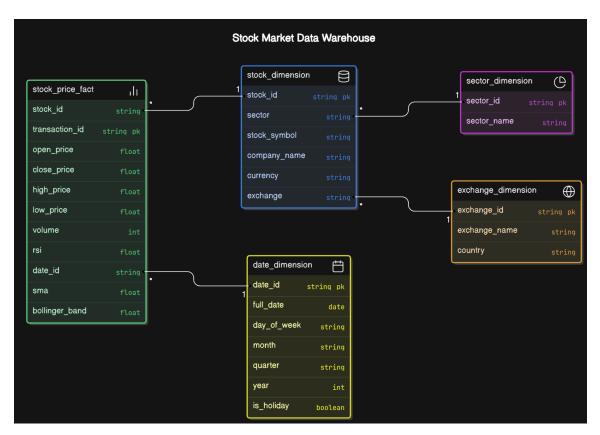
**Sector Dimension Table**: This table <u>categorizes</u> the companies by their business sectors (e.g., Technology, Healthcare).

| Column Name | Data Type | Description                           |
|-------------|-----------|---------------------------------------|
| sector_id   | INT (PK)  | Unique identifier for each sector.    |
| sector_name | VARCHAR   | Name of the sector (e.g., Technology) |

**Exchange Dimension Table**: This table contains details about stock exchanges (e.g., NASDAQ, NYSE).

| Column Name   | Data Type | Description                               |
|---------------|-----------|---|
| exchange_id   | INT (PK)  | Unique identifier for each exchange       |
| exchange_name | VARCHAR   | Name of the stock exchange (e.g., NASDAQ) |
| country       | VARCHAR   | Country of the stock exchange             |

## ERD Diagrams:



This Stock Price Data Warehouse design captures essential stock trading data by organizing it into a fact table and several dimension tables in a star schema.

#### ETL Outline:

#### Data Extraction

- Goal: Fetch raw financial data from Yahoo Finance for selected stock tickers.
- Method: Use the yfinance library to pull data such as stock prices, historical data, financial statements, and dividends.

#### Data transformation

- . Goal: Clean, normalize, and reformat the data for consistency.
- Steps: Handle missing values (e.g., fill in gaps or drop incomplete rows).
  - Normalize data types (e.g., date formats, numerical precision).

#### Data Loading

- Goal: Load the transformed data into a data warehouse for future analysis.
  - . Steps: Establish a connection to the SQL database.
    - Create tables to store the data.