Futures & Options Database Design (NSE/BSE/MCX)

Overview

This project designs and implements a normalized relational database to store and analyze high-volume Futures & Options (F&O) trading data from Indian exchanges. The schema is designed using real-world NSE F&O data from Kaggle and is scalable to support additional exchanges such as BSE and MCX.

Dataset size: 2.5M+ rows (sample used 100k locally for execution).

Database Design

The schema follows Third Normal Form (3NF) to eliminate redundancy and support scalability.

Core Tables

exchange: Stores exchange metadata (NSE, BSE, MCX)
instrument: Stores tradable symbols and instrument type (FUT/OPT)
expiry: Stores contract-level details (expiry date, strike price, option type)
trades: Central fact table storing daily OHLC prices, volume, and open interest

A star schema was intentionally avoided to reduce duplication and improve ingestion performance for time-series trading data.

Data Ingestion

Raw CSV data is first loaded into a staging table and then normalized into dimension and fact tables.
Due to local Windows environment constraints, a representative sample was used for execution. The schema and queries are designed to scale to the full dataset and beyond (10M+ rows).

Indexing & Optimization

Indexes were created on:
trade_date (time-series queries)
instrument_id
exchange_id

BRIN indexing on trade_date is recommended for very large datasets.

EXPLAIN ANALYZE confirms index usage for time-based queries with millisecond-level execution.

Analytics Queries

The project includes:
Top symbols by Open Interest (OI) change
7-day rolling volatility using window functions
Cross-exchange settlement price comparison
Option chain summary by expiry and strike
Optimized max-volume query for last 30 days

Scalability

The design supports:
Multi-exchange ingestion
Partitioning by date or exchange
High-volume analytical workloads