

Stock Market Conflict Impact Analysis: Full Workflow Documentation

Project Overview

This project automates the extraction, transformation, and visualization of stock market data to analyze how global conflicts influence major companies across the Banking, Defense, and Energy sectors.

We use:

- **Docker + Apache Airflow** for automation and scheduling
- **Yahoo Finance (via ``)** for data extraction
- **Power BI** for data modeling and interactive dashboarding

Technologies Used

Tool / Tech	Purpose
Docker	Containerized Airflow environment
Apache Airflow	Pipeline orchestration and scheduling
Python + yfinance	Fetching stock market data from Yahoo
Pandas	Data transformation (moving averages, returns)
Power BI	Business Intelligence visualization platform

Step-by-Step Breakdown

1. Setup Apache Airflow with Docker

- Created a folder `airflow_stock_project` with `dags/`, `logs/`, `plugins/`
- Pulled official Airflow `docker-compose.yaml`
- Created `.env` file with:

```
AIRFLOW_UID=50000
```

- Ran:

```
docker compose up airflow-init
docker compose up
```

- Accessed Airflow at <http://localhost:8080>

2. Define ETL Pipeline in Airflow DAG

- DAG: `stock_etl_cof`
- Schedules a daily pipeline
- For each ticker:
 - Downloads 1-year historical stock data
 - Calculates:
 - 20-day Moving Average (MA20)
 - 50-day Moving Average (MA50)
 - 5-day Volatility
 - 3-day Return
 - Saves two sets of CSVs:
 - Full data per ticker (e.g., `COF_data.csv`)
 - ± 5 -day windows around event dates (e.g., `COF_Iran_Israel_2025.csv`)

Conflict Windows Tracked

The DAG filters stock data around these key conflict events:

Event Name	Date
Iran-Israel Conflict Begins	2025-06-05
India-Pakistan Sinddor	2025-05-12
Israel-Gaza Escalation	2025-06-18

Each event generates filtered CSVs per company.

Data Model in Power BI

- All event-window CSVs are **appended** into one dataset: `War-conflict-stockdata`
- Added **DAX columns**:
 - `CompanyName` using SWITCH based on ticker
 - `Sector` based on predefined mapping

- EventName derived from file name or lookup

Visualizations in Power BI

Filters/Slicers:

- Conflict/War (event-level filter)
- Sector (Banking, Defense, Energy)

Visuals Created:

Visual Title	Description
Stock Price Trend During Conflict Timeline	Line chart of Close by Date per company during selected conflict
Latest Market Close vs Conflict Period Close	Clustered bar chart: main bar = current Close, highlight = conflict Close
20/50-Day Moving Averages Table	Table showing MA20 and MA50 at the end of selected conflict
Latest Close by Company	Clustered bar of current close price for selected companies
Slicers	For EventName, Sector, and optionally CompanyName

Refresh Strategy

- Airflow DAG runs daily via Docker
- New CSVs are written into event_windows/
- Power BI reads from folder-based source (automatically appends new files)
- Manual refresh in Power BI: **Home > Refresh**

Summary

This project showcases a full **ETL pipeline** that:

- Automates data ingestion from Yahoo Finance
- Tracks market reaction to geopolitical conflict
- Uses Airflow for daily orchestration
- Visualizes dynamic, filterable insights in Power BI

It highlights stock market volatility, trends, and sector-specific responses to real-world events.

