

Fintech Data Warehouse

Implementation Summary

Generated: December 7, 2025

Target Completion: December 18, 2025

EXECUTIVE SUMMARY

This document summarizes the implementation progress of the AIAlgoTradeHits Fintech Data Warehouse expansion project. The system has been expanded to include 35+ BigQuery tables covering fundamental data, analyst ratings, corporate actions, ETF analytics, and market movers. ML features from Phase 1 methodology have been implemented with 27 new columns added to OHLCV tables.

IMPLEMENTATION PROGRESS

Component	Details	Status
BigQuery Schemas	33 new tables	COMPLETE
OHLCV Expansion	6 new tables	COMPLETE
Fundamental Data	5 tables	COMPLETE
Analyst Data	4 tables	COMPLETE
Corporate Actions	4 tables	COMPLETE
ETF Analytics	4 tables	COMPLETE
Market Data	3 tables	COMPLETE
ML Features Phase 1	27 columns	COMPLETE
Cloud Functions	5 new functions	READY TO DEPLOY

AIAlgoTradeHits - Implementation Summary

NEW BIGQUERY TABLES

Table Name	Description
fundamentals_company_profile	Company info, CEO, sector
fundamentals_statistics	Market cap, PE, ratios
fundamentals_income_statement	Revenue, profit, EPS
fundamentals_balance_sheet	Assets, liabilities
fundamentals_cash_flow	Cash flows, FCF
analyst_recommendations	Buy/sell ratings
analyst_price_targets	Analyst price targets
analyst_earnings_estimates	EPS forecasts
earnings_calendar	Upcoming earnings
dividends_calendar	Dividend dates
splits_calendar	Stock splits
ipo_calendar	IPO schedule
insider_transactions	Insider trading
institutional_holders	Fund ownership
etf_profile	ETF details
etf_performance	ETF returns
etf_holdings	ETF composition
market_movers	Top gainers/losers
v2_forex_daily	Forex OHLCV
v2_commodities_daily	Commodity prices
v2_bonds_daily	Treasury yields
v2_indices_daily	Index OHLCV

AIAlgoTradeHits - Implementation Summary

ML PHASE 1 FEATURES IMPLEMENTED

Phase	Features	Category
Phase 1A	log_return, return_2w, return_	Quick Wins
Phase 1A	ema_20, ema_50, ema_200	EMAs
Phase 1A	close_vs_sma20/50/200_pct	MA Distance
Phase 1A	bb_width	Bollinger Width
Phase 1B	rsi_slope, rsi_zscore	RSI Derivatives
Phase 1B	rsi_overbought, rsi_oversold	RSI Flags
Phase 1B	macd_cross	MACD Signal
Phase 1B	ema20/50_slope	EMA Momentum
Phase 1B	atr_zscore, atr_slope	ATR Derivatives
Phase 1B	volume_zscore, volume_ratio	Volume Metrics
Phase 1C	pivot_high/low_flag	Swing Points
Phase 1C	trend_regime	Trend State
Phase 1C	vol_regime	Volatility State

CLOUD FUNCTIONS CREATED

Function	Schedule	Data
fundamentals-fetcher	Weekly	Company profiles & stats
analyst-fetcher	Daily	Ratings & price targets
earnings-calendar-fetcher	Daily	Earnings, dividends, splits
market-movers-fetcher	30 min	Top gainers/losers
etf-analytics-fetcher	Daily	ETF profiles & performance

AI Algo Trade Hits - Implementation Summary

DATA SUMMARY

Asset	Table	Records	History
BTCUSD	v2_crypto_daily	3,651 rows	10 years
SPY	v2_etfs_daily	2,512 rows	10 years
QQQ	v2_etfs_daily	2,512 rows	10 years
S&P 500	v2_stocks_daily	500 symbols	Varies
Top ETFs	v2_etfs_daily	200 symbols	10 years
Cryptocurrencies	v2_crypto_daily	100 symbols	10 years

NEXT STEPS (Dec 8-18)

1. Deploy all cloud functions to GCP
2. Setup Cloud Schedulers for automated data collection
3. Run fundamentals fetcher to populate company data
4. Test analyst ratings and price target endpoints
5. Verify earnings calendar data flow
6. Complete ETF holdings data collection
7. Run ML Quick Start notebook for model training
8. Integrate ML predictions into trading app
9. Final testing and documentation for Saleem

COST ANALYSIS

Component	Monthly Cost
BigQuery Storage	\$10-20/month
Cloud Functions	\$100-150/month
Cloud Scheduler	\$1/month
Cloud Run	\$5-10/month
TwelveData API	\$135/month
Total Estimated	\$250-320/month