Mirror Market

A commodity market intelligence platform that monitors global agricultural markets -- soybeans, coffee, palm oil, corn, wheat, sugar, cotton, and livestock -- pulling data from 11 free sources across 27 countries into a single database with professional-grade analysis, a daily briefing, and an interactive Streamlit dashboard.

What It Does

Mirror Market runs a data pipeline that collects, cleans, validates, and stores market data from 11 independent sources. The analysis engine then processes everything into a daily briefing: technical signals (MACD, Bollinger Bands, RSI divergence), crush spreads, forward curve structure, export sales demand, cross-market correlations, seasonal patterns, and a "Market Drivers" narrative that connects data across sources to surface insights no single section shows alone. An interactive Streamlit dashboard provides 7 pages of charts and analysis.

Data Sources (All FREE)

Layer 1 -- Commodity Futures Prices

Source: Yahoo Finance (CME/CBOT/ICE) | **Frequency**: Daily (~15 min delay)

Ticker	Contract	What It Tells You
ZS=F	Soybeans	Benchmark global soybean price. Most-traded ag future. Driven by US/Brazil planting, Chinese demand, and weather.
ZL=F	Soybean Oil	Used in cooking oil and biodiesel. Competes with palm oil. Rising biofuel mandates push this up.
ZM=F	Soybean Meal	Animal feed ingredient. ~65% of a crushed soybean's value. Tight protein meal supply lifts this.
KC=F	Coffee (Arabica)	Premium coffee bean. Brazil is #1 producer. Weather in Minas Gerais moves this market.
ZC=F	Corn	Largest US crop. THE #1 driver of soybean acreage -- when corn is more profitable, farmers plant less soy.
ZW=F	Wheat	Competes for acreage. Food inflation proxy. Global supply disruptions (Black Sea, drought) ripple into all grains.
SB=F	Sugar	Competes with ethanol for processing capacity. Affects biofuel demand dynamics (soybean oil vs ethanol).
CT=F	Cotton	Competes for acreage in US South, Brazil, and India. A cotton rally can pull area away from soybeans.
LE=F	Live Cattle	Beef herd expansion = more soybean meal demand. Livestock cycles are the demand side of soybeans.
HE=F	Lean Hogs	Hog cycle drives meal consumption globally. China's hog herd is the world's largest meal consumer.

How it's used: Price lines in the briefing, all technical indicators (MA, RSI, MACD, Bollinger, volatility), crush spread calculation, seasonal comparison, correlation matrix, signal detection. Market Drivers section uses corn/soy price ratio to flag acreage competition and livestock prices to flag feed demand shifts.

Layer 2 -- USDA Crop Fundamentals

Source: USDA NASS QuickStats API | **Frequency**: Annual (updated ~January) | **Requires**: `USDA_API_KEY`

Fetches US soybean production, yield (bushels/acre), and area harvested.

How it's used: Briefing shows year-over-year changes -- e.g. "US soybean production: 4,165M bu (+3.2% YoY)". Helps answer: is the US growing more or less soybeans than last year?

Layer 2b -- USDA Crop Progress & Condition

Source: USDA NASS QuickStats API | **Frequency**: Weekly (during growing season) | **Requires**: `USDA_API_KEY`

The most price-moving weekly report for US crops. Fetches for soybeans and corn: - **Progress**: % planted, emerged, blooming, setting pods, dropping leaves, harvested - **Condition**: % rated excellent, good, fair, poor, very poor

A drop in good/excellent % = potential yield loss = price rally. This is the data that moves markets intraweek.

How it's used: Briefing shows latest condition ratings and week-over-week changes. A 3+ point drop in good/excellent ratings is a strong bullish signal.

Layer 3 -- Economic Context (FRED)

Source: Federal Reserve Economic Data | **Frequency**: Daily/Monthly | **Requires**: `FRED_API_KEY`

Series	What It Tells You
US Dollar Index (DTWEXBGS)	Strong dollar = commodities get more expensive for foreign buyers = downward price pressure. Weak dollar = tailwind for commodities.
CPI (CPIAUCSL)	Inflation backdrop. High CPI can drive commodity demand as an inflation hedge.
Fed Funds Rate (FEDFUNDS)	Interest rate environment. Rising rates strengthen the dollar (headwind for commodities) and increase storage costs.
Treasury 2Y (DGS2)	Short-term yield. Moves with Fed expectations.
Treasury 10Y (DGS10)	Long-term yield. The 2Y/10Y spread is the classic recession signal.
Treasury 30Y (DGS30)	Ultra-long yield. Reflects inflation expectations decades out.
Ethanol PPI (WPU06140341)	Producer price index for ethanol. Tracks biofuel cost -- soybean oil competes with ethanol for blend mandates.

How it's used: Briefing shows latest values with directional commentary (e.g. "Dollar Index: 104.52 (up 0.3% -- headwind for commodities)"). The yield curve section shows the 2Y/10Y spread -- when it inverts (goes negative), it signals recession risk and potential commodity demand destruction. Market Drivers section flags dollar moves >0.5% as a cross-market signal.

Layer 4 -- COT Positioning (Commitment of Traders)

Source: CFTC via cot_reports library | **Frequency**: Weekly (published Fridays, data from prior Tuesday)

Tracks positioning for 10 commodities: Soybeans, Soybean Oil, Soybean Meal, Coffee, Corn, Wheat, Sugar, Cotton, Live Cattle, and Lean Hogs.

Shows how different trader groups are positioned in each commodity: - **Commercials** (hedgers): Farmers, processors, exporters. They trade to manage business risk. When they are heavily short, they expect prices to drop. - **Non-commercials** (speculators): Hedge funds, managed money. They trade for profit. Extreme spec positions often mark turning points.

How it's used: Briefing shows net positions for each group. Market Drivers section flags "crowded trades" -- when spec positioning is extreme AND RSI confirms overbought/oversold, reversal risk is elevated.

Layer 5 -- Weather Data

Source: Open-Meteo API | **Frequency**: Daily + 7-day forecast

Monitors 20 growing regions across 6 continents:

Region	Why It Matters
US Midwest (Iowa)	**Heart of the US soybean belt. Summer heat/drought during pod-fill (Jul-Aug) can slash yields.**
US Illinois	**#1 US soybean state by production.**
Brazil Mato Grosso	**Brazil's #1 soybean state (~30% of production). Dry conditions during planting (Oct-Nov) delay the crop.**
Brazil Parana	**Brazil's #2 soybean state. Frost risk during June-July (southern hemisphere winter).**
Brazil Minas Gerais	**Coffee capital of Brazil. #1 Arabica state. Frost or drought here moves global coffee prices.**
Brazil Bahia	**Cacao + coffee region. Northeast Brazil growing area.**
Argentina Pampas	**#3 soybean exporter. La Nina brings drought here; El Nino brings floods.**
Argentina Cordoba	**#2 Argentina soybean province.**
Paraguay Chaco	**#4 global soybean exporter. Expanding soy frontier.**
Colombia Coffee Region	**#3 Arabica producer. Too much rain during harvest = quality issues.**

Region	Why It Matters
Ethiopia Sidama	**#1 Africa coffee producer. Birthplace of Arabica.**
Ivory Coast	**#1 cocoa producer (cross-reference for tropical agriculture conditions).**
Vietnam Central Highlands	**#2 global Robusta producer. Drought during dry season (Jan-Apr) damages trees.**
Indonesia Riau (Sumatra)	**#1 palm oil belt. El Nino = drought = lower yields.**
Malaysia Sabah (Borneo)	**#2 palm oil state. Flooding during monsoon season disrupts harvest.**
India Madhya Pradesh	**India's soybean capital. Monsoon timing drives the entire crop.**
India Maharashtra	**#2 India soybean state. Late monsoon onset = delayed planting.**
Thailand Surat Thani	**#3 global palm oil producer.**
China Heilongjiang	**China's domestic soybean belt. Non-GMO soybeans for food use.**

How it's used: Briefing flags heavy rain (>20mm), extreme heat (>38C), and dry conditions (<1mm). Market Drivers section connects weather alerts with rising prices to identify "weather premiums building."

Layer 6 -- Global Supply & Demand (PSD)

Source: USDA Foreign Agricultural Service (bulk CSV) | **Frequency**: Monthly

Production, imports, exports, crush, beginning stocks, ending stocks, domestic consumption, total supply, and total distribution for **8 commodities** across **27 countries**.

Commodities: Soybeans, Soybean Oil, Soybean Meal, Palm Oil, Coffee, Corn, Wheat, Cotton

Countries: United States, Brazil, Argentina, Paraguay, Uruguay, Bolivia, Colombia, Mexico, China, India, Indonesia, Malaysia, Thailand, Vietnam, Japan, South Korea, Pakistan, Bangladesh, European Union, Ethiopia, Nigeria, South Africa, Ivory Coast, Tanzania, Uganda, Kenya, Australia

How it's used: Briefing highlights the numbers that move markets -- Brazil soybean production, China soybean imports, US production, Indonesia palm oil output. Year-over-year changes in these figures drive long-term price trends. The global view across 27 countries prevents blind spots -- e.g. a surge in Indian soybean imports or a drop in Argentine production shows up here before it hits prices.

Layer 7 -- Currency Exchange Rates

Source: Yahoo Finance | **Frequency**: Daily

Pair	Why It Matters
BRL/USD	**Brazilian Real. Brazil exports ~50% of global soybeans. Weak BRL = Brazilian farmers get more Reais per dollar = incentivized to sell = more supply hitting world markets = price pressure. This is the single most important currency for soybeans.**
ARS/USD	**Argentine Peso. Argentina is #3 soybean exporter. Chronic devaluation here affects export pacing.**
COP/USD	**Colombian Peso. #3 Arabica coffee producer.**
PYG/USD	**Paraguayan Guarani. #4 global soybean exporter.**
CNY/USD	**Chinese Yuan. China imports ~60% of globally traded soybeans. Yuan moves affect their buying power.**
IDR/USD	**Indonesian Rupiah. #1 palm oil producer.**
MYR/USD	**Malaysian Ringgit. #2 palm oil producer.**
VND/USD	**Vietnamese Dong. #2 Robusta coffee producer.**
INR/USD	**Indian Rupee. Major soybean and palm oil consumer/processor.**
THB/USD	**Thai Baht. #3 global palm oil producer.**
ETB/USD	**Ethiopian Birr. #1 Africa coffee producer. Birr volatility affects Ethiopian export competitiveness.**

How it's used: Briefing shows exchange rates with trade impact commentary (e.g. "Real weakening -- Brazil exports cheaper"). Correlation analysis measures how tightly currencies and commodity prices move together. Market Drivers flags when BRL moves >1% in a week.

Layer 8 -- World Bank Monthly Prices

Source: World Bank Pink Sheet (Excel) | **Frequency**: Monthly

Provides benchmark prices for commodities not covered by daily CBOT futures: - **Robusta Coffee** -- no daily free source, this is the best free data - **Palm Oil** -- CPO (crude palm oil) benchmark - Also covers Soybeans, Soybean Oil, Soybean Meal (cross-reference with CBOT)

How it's used: Briefing shows month-over-month percentage changes. Useful for longer-term trend analysis on Robusta and Palm Oil.

Layer 9 -- DCE Chinese Futures

Source: AKShare (Dalian Commodity Exchange) | **Frequency**: Daily

Contract	What It Tells You
DCE Soybean (A0)	**Chinese domestic soybean price. China is the world's largest soybean importer (~100M MT/year). When DCE prices rise relative to CBOT, it signals strong Chinese demand.**
DCE Soybean Meal (M0)	**Chinese meal price. Hog herd expansion = more meal demand = higher prices.**
DCE Soybean Oil (Y0)	**Chinese cooking oil price. Government reserve releases can cap upside.**
DCE Palm Oil (P0)	**Chinese palm oil price. Competes directly with soybean oil.**
DCE Corn (C0)	**Chinese corn price. China feed demand indicator -- corn and soybean meal are both animal feed.**

How it's used: Briefing shows DCE prices in CNY alongside CBOT prices in USD, so you can see the China-vs-US price gap. A widening gap suggests Chinese import demand is heating up.

Layer 10 -- USDA Export Sales

Source: USDA FAS OpenData API (ESR) | **Frequency**: Weekly (Thursdays) | **Requires**: `FAS_API_KEY`

Weekly export sales data -- the #1 indicator of demand pace. Every grain trader checks this every Thursday.

Data Point	What It Tells You
Net Sales	**New export sales minus cancellations. Rising = strong demand.**
Weekly Exports	**Actual shipments that week. Pace matters -- behind USDA projections = bearish.**
Accumulated Exports	**Season-to-date total. Compare to USDA forecast to gauge demand pace.**
Outstanding Sales	**Sold but not yet shipped. Large outstanding = shipping bottleneck or basis play.**
Top Buyers	**Who is buying -- China's share of soybean sales is the key demand signal.**

Commodities tracked: Soybeans, Soybean Oil, Soybean Meal, Corn, Wheat, Cotton.

How it's used: Briefing shows weekly net sales and top 3 buyer destinations per commodity. Market Drivers flags when China accounts for >30% of weekly soybean purchases (strong demand signal).

Layer 11 -- Forward Curves

Source: Yahoo Finance (individual contract months) | **Frequency**: Daily

The forward curve shows the price of each upcoming delivery month -- revealing market structure:

Structure	What It Means	Price Pattern
Contango	**Adequate supply, storage costs priced in**	**Future > Spot (upward-sloping)**
Backwardation	**Tight supply, strong immediate demand**	**Spot > Future (downward-sloping)**

Constructs tickers programmatically (e.g. `ZSN26.CBT` = Soybeans Jul 2026) and fetches the latest close for each contract month.

Commodities tracked: Soybeans, Soybean Oil, Soybean Meal, Corn, Wheat, Coffee, Sugar, Cotton, Live Cattle, Lean Hogs.

How it's used: Briefing shows the term structure (contango/backwardation) per commodity with spread percentages. Market Drivers flags backwardation (tight supply signal) and steep contango (>5%). Dashboard provides a visual forward curve chart for each commodity.

Global Coverage

Countries Tracked (27 via PSD)

Region	Countries
Americas	United States, Brazil, Argentina, Paraguay, Uruguay, Bolivia, Colombia, Mexico
Asia	China, India, Indonesia, Malaysia, Thailand, Vietnam, Japan, South Korea, Pakistan, Bangladesh
Europe	European Union
Africa	Ethiopia, Nigeria, South Africa, Ivory Coast, Tanzania, Uganda, Kenya
Oceania	Australia

Commodities by Region

Commodity	Key Producers Tracked	Key Importers Tracked
Soybeans	US, Brazil, Argentina, Paraguay, Uruguay, India, China (Heilongjiang)	China, EU, Japan, South Korea, Indonesia
Coffee	Brazil, Vietnam, Colombia, Ethiopia, India, Indonesia	EU, US, Japan
Palm Oil	Indonesia, Malaysia, Thailand	China, India, EU, Pakistan, Bangladesh
Corn	US, Brazil, Argentina	China, Japan, South Korea, Mexico, EU
Wheat	US, EU, Australia, Argentina, India	China, Indonesia, Nigeria, Brazil
Cotton	US, India, Brazil, Australia	China, Bangladesh, Vietnam, Pakistan
Livestock (feed demand)	US (cattle + hogs)	-- (tracked as demand signal for soybean meal)

Analysis Features

Technical Indicators (`analysis/technical.py`)

- **Moving Averages** (20/50/200-day SMA) -- trend direction at three timeframes
- **RSI** (14-day, Wilder smoothing) -- overbought/oversold momentum
- **MACD** (12/26/9) -- momentum shifts and trend strength
- **Bollinger Bands** (20-day, 2 std) -- volatility compression and breakout detection
- **Historical Volatility** (20-day and 60-day, annualised) -- how much the price is swinging
- **Price Changes** -- daily and weekly percentage moves

Trading Signals (`analysis/signals.py`)

- **20/50 MA crossover** -- short-term golden/death cross (severity: warning)
- **50/200 MA crossover** -- major trend shift, the "big" golden/death cross (severity: alert)
- **RSI extremes** -- overbought (>70) or oversold (<30)
- **RSI divergence** -- price makes new high but RSI doesn't (bearish) or vice versa. This is the most reliable RSI signal.
- **MACD crossover** -- momentum turning up or down
- **Bollinger Band squeeze** -- volatility at 120-day low, breakout imminent
- **Volume spike** -- today's volume >2x the 20-day average

Crush Spread (`analysis/spreads.py`)

Soybean processing margin: `(Oil price x 11) + (Meal price x 2.2) - Bean price`. Positive = profitable to crush. Widening spread = processors buying more beans = price support.

Forward Curve Analysis (`analysis/forward_curve.py`)

- **Contango/backwardation detection** -- classifies term structure based on sequential price changes
- **Curve slope** -- average price change per month across the curve
- **Calendar spreads** -- price difference between any two contract months
- Summary includes structure type, front/back prices, spread, and market implication

Correlations (`analysis/correlations.py`)

- Cross-commodity matrix (e.g. Soybeans vs Soybean Meal: +0.77 strong positive)
- Commodity-vs-currency (e.g. Soybeans vs BRL/USD: how tightly are they linked?)
- Rolling correlation (how the relationship changes over time)

Seasonal Patterns (`analysis/seasonal.py`)

Compares current price to its historical average for this calendar month. Shows "Above seasonal (+5.2%)" or "Below seasonal (-3.1%)". Soybeans typically peak Jun-Jul (weather uncertainty) and dip at harvest (Oct-Nov).

Yield Curve Analysis

Uses the 2Y/10Y Treasury spread from FRED data. When the spread is negative (inverted), it signals recession risk -- historically the most reliable recession predictor. This matters for commodities because recession = demand destruction = bearish pressure.

Market Drivers Narrative (`analysis/briefing.py`)

Connects dots across data sources -- the part no single section shows: - **BRL + exports**: Weak Real = cheaper Brazilian soy/coffee on world markets - **COT + RSI crowding**: Specs heavily long AND RSI overbought = reversal risk - **Weather + price**: Active weather alerts in growing regions AND prices rising = weather premium building - **Dollar strength**: Strong dollar = headwind for all USD-denominated commodities - **Corn/soy acreage competition**: Corn/soy price ratio signals which crop farmers will plant next season. High ratio = less soybean acreage = bullish soybeans. - **Livestock demand**: Rising cattle/hog prices = expanding herds = more soybean meal demand = price support - **Export sales pace**: China accounting for >30% of weekly soybean purchases signals strong demand - **Forward curve structure**: Backwardation signals tight supply; steep contango signals adequate supply

Data Freshness Tracking

The pipeline records when each layer last succeeded. The briefing shows warnings at the top if any layer is more than 7 days stale (e.g. "WARNING: USDA data is 45 days old").

Data Validation (`processing/cleaner.py`)

Sanity checks run during cleaning: - Flags daily price moves >10% (possible data corruption or extreme event) - Flags zero/negative volume (data gap) - Warnings only -- doesn't block the pipeline

Interactive Dashboard

Run with `streamlit run app/dashboard.py`. 7 pages of visual analysis:

Page	What It Shows
Overview	Full text briefing + data freshness status table
Price Charts	Candlestick chart per commodity with MA/Bollinger overlays, RSI subplot, MACD subplot, volume bars. Commodity dropdown selector.
Forward Curve	Line chart: x=contract month, y=price. Visual contango/backwardation detection with metrics.
Crush Spread	Time series with profitability shading (green = profitable, red = negative margin).
COT Positioning	Grouped bar chart: commercial vs speculator net positions. Time series detail view.
Weather	Color-coded table of 20 regions -- red for extreme heat, blue for heavy rain, yellow for dry.
Correlations	Plotly heatmap of cross-commodity correlation matrix with values.

All pages reuse existing `read_*()` and analysis functions. Data updates when you re-run `python main.py`.

What's Missing (Optional Paid Upgrades)

Missing Data	Why	Upgrade Cost	Service
Daily Robusta Coffee prices	Not available on free platforms	~$10-50/month	Commodities-API.com (symbol: ROBUSTA)
Daily Palm Oil prices	Not available on free platforms	~$10-50/month	Commodities-API.com (symbol: CPO)
ICE certified coffee stocks	Requires exchange subscription	Varies	ICE Data Services
Real-time prices	Exchange licensing fees	$500+/month	Barchart cmdtyView

Note: The free World Bank monthly data for Robusta and Palm Oil is sufficient for trend analysis and seasonal studies. Daily prices would add granularity for shorter-term trading decisions.

Configurable Thresholds (`config.py`)

These can be tuned without touching analysis code:

Threshold	Default	What It Controls
RSI_OVERBOUGHT	70	RSI level that flags overbought
RSI_OVERSOLD	30	RSI level that flags oversold
VOLUME_SPIKE_MULTIPLIER	2.0	Multiple of 20-day avg volume to flag as unusual
WEATHER_HEAVY_RAIN_MM	20	Precipitation threshold for heavy rain alert
WEATHER_EXTREME_HEAT_C	38	Temperature threshold for crop stress alert
WEATHER_DRY_THRESHOLD_MM	1	Below this = "dry conditions" alert
FRESHNESS_WARNING_DAYS	7	Days before stale data warning appears

How to Run

```
# Set API keys (one-time, optional -- 8 of 11 layers work without them)
export USDA_API_KEY="your-key-here"
export FRED_API_KEY="your-key-here"
export FAS_API_KEY="your-key-here"

# Install dependencies
pip install -r requirements.txt

# Run the pipeline (fetches all 11 layers, cleans, validates, stores)
python main.py

# Generate daily briefing
python -c "from analysis.briefing import generate_briefing; print(generate_briefing())"

# Launch interactive dashboard
streamlit run app/dashboard.py
```

Tech Stack

- **Language**: Python 3.10+
- **Database**: SQLite (local, no server needed)
- **Data Libraries**: pandas, yfinance, requests, AKShare, cot_reports
- **Dashboard**: Streamlit + Plotly (interactive charts)
- **Analysis**: Pure Python/pandas (no paid analytics tools)

Project Structure

```
Mirror_Market/
config.py                          # Tickers, API keys, URLs, thresholds
main.py                            # Pipeline orchestrator
data/
fetchers/
yfinance_fetcher.py        # Layers 1 + 7 (prices + currencies)
usda_fetcher.py            # Layers 2 + 2b (fundamentals + crop progress)
fred_fetcher.py            # Layer 3
cot_fetcher.py             # Layer 4
weather_fetcher.py         # Layer 5
psd_fetcher.py             # Layer 6
worldbank_fetcher.py       # Layer 8
akshare_fetcher.py         # Layer 9
export_sales_fetcher.py    # Layer 10 (USDA FAS export sales)
forward_curve_fetcher.py   # Layer 11 (individual contract months)
storage/
mirror_market.db           # SQLite database (gitignored)
processing/
cleaner.py                     # Data cleaning + validation
combiner.py                    # SQLite storage + freshness tracking
analysis/
technical.py                   # SMA, RSI, MACD, Bollinger, volatility
signals.py                     # Signal detection (crossovers, divergence, squeeze)
spreads.py                     # Crush spread calculation
correlations.py                # Cross-market correlation analysis
seasonal.py                    # Seasonal pattern comparison
forward_curve.py               # Forward curve analysis (contango/backwardation)
briefing.py                    # Daily briefing generator (all sections + Market Drivers)
```

```
app/
dashboard.py                    # Interactive Streamlit + Plotly dashboard (7 pages)
```