SECTOR: AGRICULTURE
Data Visualization & Analytics — Capstone

DATASET COVERAGE
2018 – 2026

CAPSTONE PROJECT PRESENTATION

Commodity-Wise Mandi Price Analysis

Farmer Decision Support System

"Identifying the best market and time to sell for maximum farmer returns"

PROJECT TEAM

Group 4
SECTION B

Khushi	2401010225	Kshitiz	2401010242
Abhinav	2401010017	Abhijeet	2401010014
Anshika	2401010080	Nishant	2401010302

FACULTY MENTORS
Satyaki Sir
Ayushi Ma'am

PROJECT OVERVIEW

Context & Problem Statement

Sector Context

Farmers face high price variability across mandis and seasons. Selling decisions are often based on limited, fragmented, or delayed information rather than data-driven market intelligence.

Key Decision-Makers

- Individual Farmers
- Farmer Producer Organizations (FPOs) / Co-ops
- APMC Market Committees & Agri-Advisory Teams

CORE PROBLEM STATEMENT

*"Farmers lack clarity on **when and where** to sell their crops due to price variations across mandis and seasons, leading to **suboptimal selling decisions** and reduced profits."*

PROJECT OBJECTIVE

Maximize Farmer Returns

Identify which **markets** and **time periods** offer the best prices to guide smarter, data-driven selling decisions, moving from intuition-based to insight-based agriculture.

Core Business Question: How can we predict optimal selling windows and locations to minimize distress sales?

METHODOLOGY

Data Engineering Pipeline

DATASET SCOPE
6,000 Rows • 18 Columns • 2018–2026

SOURCE

Raw Dataset

Agmarknet / Kaggle Mirror

- Mixed Date Formats
- Missing Modal Prices
- Inconsistent Text Case

PROCESS

Cleaning & Std.

Google Sheets

- ✓ Date Standardization
- ✓ Imputation (Mid-Range)
- ✓ Outlier Validation

TRANSFORM

Feature Eng.

Derived Metrics

- ⚡ Price Spread & PPI
- ⚡ Relative Market Adv. (RMA)
- ⚡ Profit Opportunity Flag

SINK

Dashboard

Decision Output

- Market Intelligence
- Seasonal Trends
- Actionable Alerts

DATA DICTIONARY & CLEANING LOG

Column Name	Data Type	Description	Action Taken
Arrival_Date	Date	Date of commodity arrival	✓ Standardized
Modal_Price	Numeric	Most frequent traded price (₹/q)	🔧 Imputed
State / Market	Categorical	Geographic location identifiers	✓ Trimmed
Price_Spread	Derived	Max_Price − Min_Price	+ Created
RMA_Index	Derived	Modal − State_Avg	+ Created

Dataset Coverage

8 STATES	**8** MARKETS
8 COMMODITIES	**7** YEARS

⚠ Critical Fixes

Fixed **150+ missing values** in Modal Price using mid-range estimation. Standardized 8 variations of state names to ensure accurate regional aggregation.

MEASUREMENT STRATEGY

KPI & Metrics Framework

STRATEGIC GOAL
Maximize Farmer Returns

Modal Price

`Most Frequent Traded Price (₹/q)`

The primary revenue proxy per mandi and time period. Represents the market clearing price for the majority of volume.

FARMER DECISION RELEVANCE

Price Spread

`Max_Price - Min_Price`

Volatility indicator showing the range of traded prices. High spread indicates quality variation or market inefficiency.

RISK & AGGREGATION SIGNAL

Price Position Index

`(Modal - Min) / (Max - Min)`

Normalized score (0-1) indicating where the modal price sits within the daily range. Higher is better for sellers.

NEGOTIATION STRENGTH

Relative Market Advantage (RMA)

`Modal - State Average` *Benchmarks local market vs state baseline*

Measures the structural price premium of a specific mandi compared to the state average. Positive RMA indicates a consistently better-paying market regardless of general price trends.

LOCATION CHOICE Guides "Where to Sell" decisions

Profit Opportunity Indicator

`High / Normal Class` *Composite Signal*

Binary classification based on thresholds (e.g., Top 25% Modal Price + Positive RMA + PPI > 0.6). Identifies optimal selling windows.

ACTION TRIGGER Guides "When to Sell" decisions

EXPLORATORY ANALYSIS

Key Insights from EDA

ANALYSIS SCOPE
6,000 Records • 8 States

TIMING

Peak Selling Month

May

Highest average Modal Price observed in May, suggesting optimal pre-monsoon liquidation.

SEASONALITY

Highest Paying Season

Rabi

Rabi crops consistently command better prices compared to Kharif or Summer harvests.

TOP MARKET

Lucknow Mandi

₹3,110/q

Consistently highest average modal price, serving as a benchmark for northern region sales.

MACRO

Yearly Price Trend

Stable

Long-term trend (2018-2026) shows relative stability, indicating reliable baseline planning.

STRUCTURE

RMA Leaders

Top structural advantage

- Mysuru
- Jaipur
- Kolkata

AVG VOLATILITY

₹490/q

High spread indicates significant arbitrage potential through aggregation.

AVG PPI

0.50

Mid-range position suggests balanced market power; room for negotiation.

OPPORTUNITY

High Profit Share

Maharashtra	42.8%
Uttar Pradesh	42.3%
Rajasthan	41.6%

ADVANCED ANALYTICS

Price Position & Structural Advantage

ANALYSIS FOCUS
PPI • RMA • Spread

NEGOTIATION POWER

Price Position Index (PPI)

`(Modal Price – Min Price) / (Max Price – Min Price)`

WHAT IT MEASURES

Where the traded price sits within the day's range. A PPI of 1.0 means selling at the day's maximum price.

INTERPRETATION

- **> 0.7:** Strong seller negotiation power
- **< 0.3:** Buyers dominating; distress sales

STRUCTURAL EDGE

Relative Market Advantage (RMA)

`Mandi Modal Price – State Average Price`

WHAT IT MEASURES

The inherent price premium of a specific market compared to its regional benchmark. Identifies "structurally better" markets.

INTERPRETATION

- Positive (+): **Market consistently pays above state avg**
- Negative (-): **Market underperforms regional peers**

VOLATILITY & RISK

Price Spread Analysis

`Max Price – Min Price (Daily Range)`

WHAT IT MEASURES

The degree of intraday price volatility. High spread indicates quality differentiation or information asymmetry.

INTERPRETATION

- High Spread: **Opportunity for sorting/grading uplift**
- Low Spread: **Commodity market; volume driven**

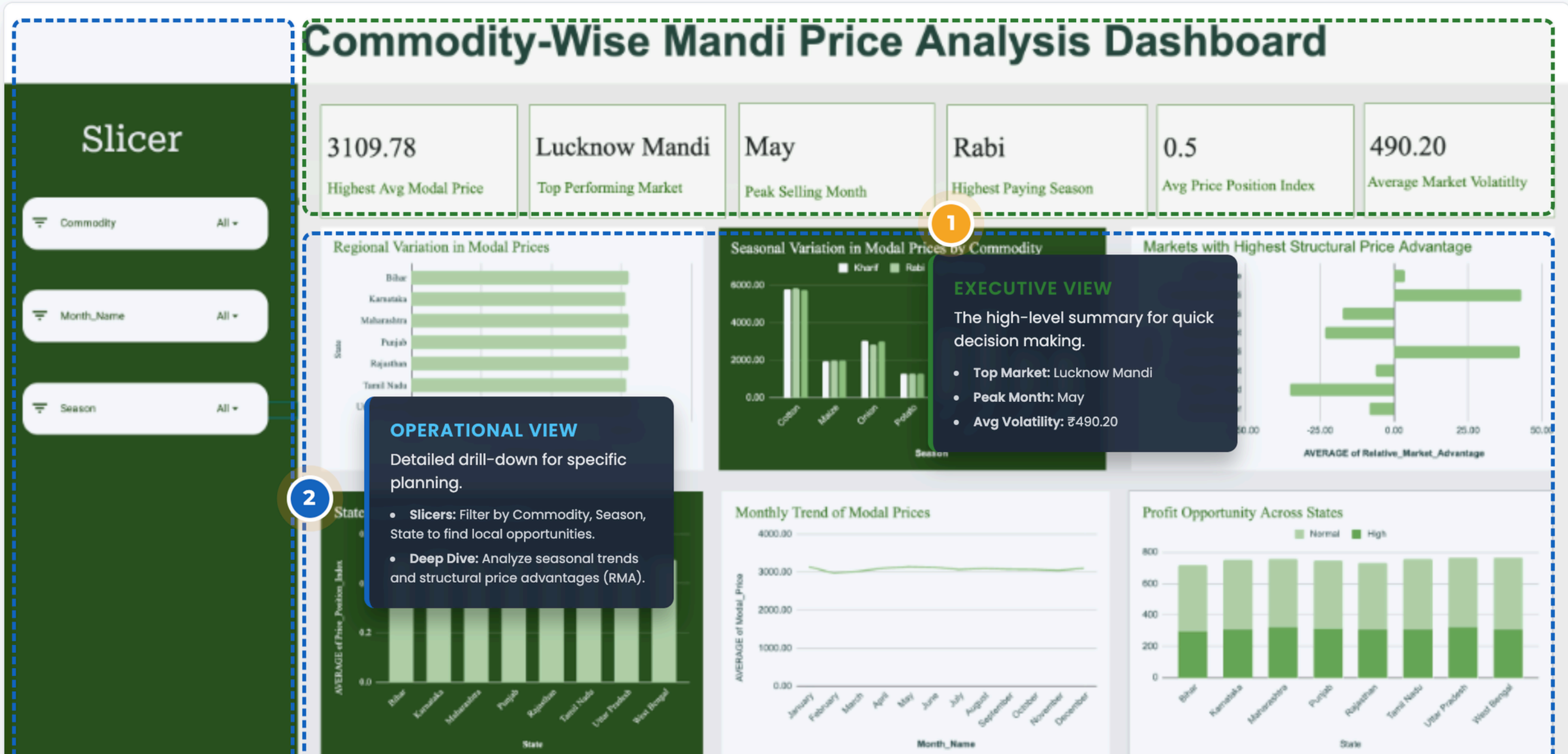
STRATEGIC SYNTHESIS
Combine RMA for location choice + PPI for timing execution.

TARGET PPI	RMA STRATEGY	VOL. CAPTURE
> 0.60	Positive Only	Aggregation

TOOL WALKTHROUGH

Farmer Decision Dashboard

VISUALIZATION
Executive & Operational Views



ACTION PLAN

Insight-Driven Recommendations

SEASONAL STRATEGY

01

Align with Summer Peak

Shift major sales volume to the **April–May** window to capture the "Summer Premium" when post-harvest supply tightens.

↑ Data: April Peak ₹2,166/qtl

02

Avoid Winter Trough

Withhold non-perishable stock during **Nov–Jan** to bypass the annual low-price window caused by market glut.

↓ Data: Nov Low ~₹2,078/qtl

GEOGRAPHIC ARBITRAGE

03

Route to Maharashtra

Direct logistics to markets in **Maharashtra** which consistently outperform national averages due to superior infrastructure.

Data: MH Avg ₹2,175/qtl

04

Target Lucknow Mandi

Prioritize **Lucknow Mandi** as a key node, offering premium realization opportunities above the state baseline.

Data: Top Performing Node

DATA-DRIVEN TACTICS

05

Manage Price Risk

Leverage **Volatility Data** to identify risk; avoid selling during high-spread periods or exploit arbitrage in volatile crops.

Data: Avg Spread ₹448/qtl

06

Dynamic Allocation

Use **Commodity Slicers** to shift focus to high-value crops like Soyabean vs staples based on real-time dashboard trends.

Data: Soyabean > Staples

BUSINESS VALUE

Data-Driven Impact Assessment

DATASET FINDINGS
Measurable ROI

COST

Strategic Market Selection

Shifting sales to markets with positive **RMA (Relative Market Advantage)** directly improves net realization.

✓ **Impact: Capture state-level price premiums (e.g., Mysuru, Jaipur)**

TIME

Seasonal Targeting

Aligning sales with peak pricing windows identified in EDA (e.g., **May** and **Rabi season**).

✓ **Impact: Maximize revenue per cycle by avoiding low-price months**

EFFICIENCY

Volatility Management

Leveraging **PPI** and **Price Spread** analysis to execute trades during optimal intraday windows.

✓ **Impact: Reduce distress sales; capture arbitrage from ₹490 spread**

Why Approve This Solution?

Based on analysis of 6,000 records

01 **Proven Negotiation Room**

Dataset shows average **PPI of 0.50**, confirming significant room (0.50 margin) to negotiate towards max prices.

02 **Clear Arbitrage Value**

Average market volatility (Spread) of **₹490.20/q** validates the need for timed aggregation strategies.

03 **Regional Scalability**

Identified **40%+ high-opportunity share** in key states (Maharashtra, UP, Rajasthan), ensuring broad applicability.

BOTTOM LINE

Data confirms actionable variability in both time and location, justifying decision support intervention.

PROJECT CONCLUSION

Limitations & Next Steps

FUTURE OUTLOOK
Scope & Roadmap

Data Gaps & Constraints

Current project limitations and assumptions

Restricted Scope

Analysis limited to 8 specific states/markets. Regional generalizations may not apply pan-India.

Data Quality (Imputation)

Some `Modal_Price` values were imputed using mid-range calculations where official records were missing.

Quality Nuances

"Variety" field captures some grades, but granular quality parameters (moisture, foreign matter) are absent.

Cost Model Exclusions

Net realization estimates do not strictly account for variable transport, loading, and storage costs.

Future Roadmap

Strategic enhancements and analytical expansion

Expand Coverage

Scale data ingestion to include all major Tier-1 and Tier-2 mandis across India for comprehensive analysis. This will reduce regional bias and improve model generalizability across different agricultural zones.

Integrate Costs

Embed logic for **Net Realization Price**: (Modal Price - Transport Cost/km - Handling). This provides a truer picture of farmer profitability beyond just the top-line selling price.

Feature Engineering

Incorporate "Daily Arrival Volume" and "Rainfall Deviation" as predictive features for price volatility. Correlating weather anomalies with price spikes will enhance the predictive power of the decision framework.