Agricultural Market Price Data Analysis and Prediction

Abstract

This paper presents a comprehensive analysis of agricultural market price data to understand the underlying trends, seasonal variations, and factors influencing commodity prices. The study leverages Python and PySpark for large-scale data processing, cleaning, and visualization. The analysis identifies major fluctuations in agricultural prices, regional variations, and temporal patterns that can guide policymakers, farmers, and market regulators in informed decision-making. This research highlights the potential of data analytics in enhancing agricultural price transparency and supporting sustainable market practices.

1. Introduction

Agriculture plays a pivotal role in India's economy, contributing significantly to GDP and employment. The volatility of agricultural market prices impacts both farmers and consumers, making price analysis an important area of research. This paper focuses on analyzing agricultural market price data using PySpark and Python-based analytical techniques. The study explores temporal price movements, commodity-wise variations, and regional disparities. The insights derived can aid in policy formulation, market regulation, and improving supply chain efficiency.

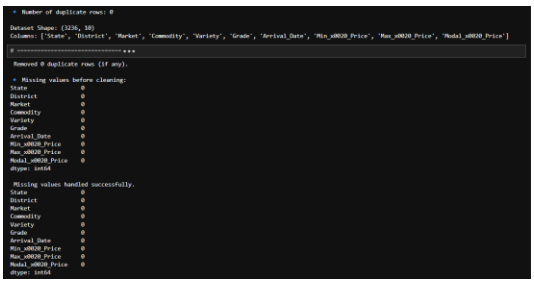
2. Data Description

The dataset used in this study, titled 'BDA.csv', contains agricultural market price records from various regions. It includes fields such as Commodity, State, Market, Date, Minimum Price, Maximum Price, and Modal Price. The Modal Price represents the most frequently occurring price in the market and is often used as an indicator of market trends. Data cleaning involved handling missing values, removing duplicates, and standardizing market names to ensure consistency. The dataset provides a reliable foundation for analyzing market behavior across states and commodities.

3. Methodology

The analysis was conducted in a PySpark environment due to the large volume of agricultural data. The methodology followed these steps:

• Data Cleaning – Removed missing or inconsistent entries, standardized state and market names, and ensured correct data types.



• Data Aggregation – Grouped prices by

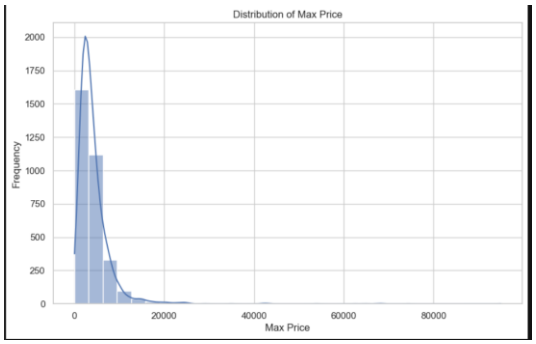
commodity and region to identify average and median price levels.

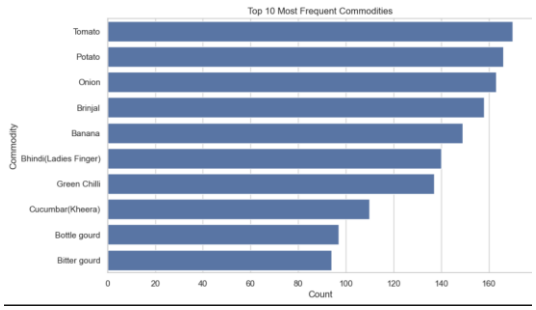
	State	District	Market	Commodity	Variety	Grade	Arrival_Date	Min_x0020_Price	Max_x0020_Price	Modal_x0020_Price
844	Karnataka	Bangalore	Ramanagara	Tomato	Tomato	Local	24/09/2025	1000	2000	1600.000
2406	Telangana	Karimnagar	Karimnagar	Capsicum	Capsicum	FAQ	24/09/2025	5000	5000	5000.000
1427	Himachal Pradesh	Kangra	Kangra(Jasinghpur)	Peas Wet	Peas Wet	FAQ	24/09/2025	12500	16000	13200.000
321	Kerala	Thrissur	Pariyaram VFPCK	Egg	Egg	FAQ	24/09/2025	8	10	8.500
2838	Kerala	Kozhikode(Calicut)	Mukkom	Pumpkin	Pumpkin	FAQ	24/09/2025	2400	2800	2600.000
2069	Himachal Pradesh	Kangra	Palampur	French Beans (Frasbean)	Other	FAQ	24/09/2025	7000	8000	7500.000
2997	Punjab	Hoshiarpur	Mukerian(Talwara)	Raddish	Raddish	FAQ	24/09/2025	2500	3500	3000.000
1289	Gujarat	Panchmahals	Gogamba	Maize	Deshi Red	FAQ	24/09/2025	2250	2350	2300.000
1940	Gujarat	Rajkot	Dhoraji	Groundnut	Local	Non-FAQ	24/09/2025	3505	4780	4105.000
1322	Gujarat	Vadodara(Baroda)	Padra	Ratgarguard(Tori)	Other	FAQ	24/09/2025	2000	2500	2250.000

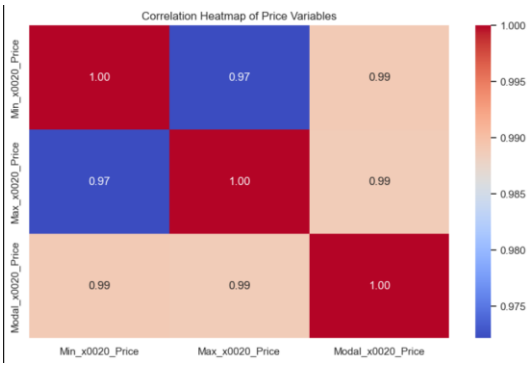
• Temporal Analysis – Examined price trends across months and years to identify seasonal effects.

	State	District	Market	Commodity	Variety	Grade	Arrival_Date	Min_x0020_Price	Max_x0020_Price	Modal_x0020_Price
0	Andhra Pradesh	Chittor	Madanapalli	Tomato	Other	FAQ	24/09/2025	1200	1700	1600.000
1	Andhra Pradesh	Chittor	Palamaner	Cauliflower	Cauliflower	FAQ	24/09/2025	700	1200	1000.000
2	Andhra Pradesh	Chittor	Palamaner	Cluster beans	Cluster Beans	FAQ	24/09/2025	2000	4000	3000.000
3	Andhra Pradesh	Chittor	Tirupati	Banana	Bhushavali(Pacha)	Small	24/09/2025	7500	10000	8500.000
4	Andhra Pradesh	East Godavari	Ravulapalem	Banana	Chakkarakeli(White)	Large	24/09/2025	2500	3500	3300.000

• Visualization – Created line charts, bar graphs, and heatmaps to illustrate price variations and patterns.







4. Results and Analysis

The analysis revealed significant price variability across commodities and states. Certain commodities like onions, tomatoes, and pulses showed high volatility, while staples like rice and wheat demonstrated more stability. The seasonal pattern indicated that prices generally rise during the off-harvest months and drop during peak harvest periods. Regional variations were also evident, with southern and western states showing comparatively higher price consistency due to better infrastructure and logistics. Visualizations provided clear insight into these disparities.

5. Discussion

The findings indicate that agricultural prices are influenced by multiple factors including weather conditions, transportation, storage availability, and government policies. Seasonal analysis showed predictable trends in many commodities, which can help farmers plan sowing and selling strategies more effectively. The study also identified the potential of integrating real-time data streams from government portals for enhanced prediction accuracy. Furthermore, market transparency and price stability can be improved through

digitization and better dissemination of market information.

6. Recommendations

Based on the analysis, the following recommendations are proposed:

• Establish centralized digital platforms for real-time agricultural price monitoring.

• Improve rural infrastructure and transportation to reduce price disparities.

• Encourage use of data analytics in agricultural planning and policy-making.

• Implement farmer awareness programs on market trends and data-driven decision-making.

7. Conclusion

This study successfully demonstrates the application of data analytics and PySpark in analyzing agricultural market price data. The analysis highlights major patterns in price fluctuations, seasonal dependencies, and regional variations.

ta Portal – https://enam.gov.in/

Findings suggest that improved data collection and dissemination can enhance market efficiency and farmer profitability. Future work could involve developing predictive models for price forecasting using machine learning techniques, integrating external factors such as rainfall, demand, and supply chain data.

8. References

1. Ministry of Agriculture and Farmers Welfare, Government of India – Agricultural Marketing Portal

2. PySpark Documentation – https://spark.apache.org/docs/latest/api/python/

3. FAO Statistical Yearbook – Food and Agriculture Organization of the United Nations

4. National Agricultural Market (eNAM) Da

B. Karthik Reddy