**Critical Analysis of Government Vegetable Mandi Dataset**

**To:** Dr.B.Jogeswara Rao

**From:** P. Charan Sai

**RollNo:** 2211CS010459

**Date:** 22/10/2025

**Subject:** Analysis of Government Vegetable Mandi Dataset

**Dataset Description**

This report presents an analysis of the Government Vegetable Mandi (Market) dataset. The data provides a cross-sectional snapshot of agricultural commodity prices from various wholesale markets (Mandis) across India on a single day. The primary domain is agricultural economics and market analysis, aimed at understanding price structures, market activity, and regional variations in trade.

**Structure and Scope:**

* **Records:** The dataset contains **1,370 clean and complete records**, with no missing values found in the key Commodity or Modal\_x0020\_Price columns.
* **Key Columns:** The data includes State, District, Market, Commodity, Variety, Grade, Min\_x0020\_Price (Minimum Price), Max\_x0020\_Price (Maximum Price), and Modal\_x0020\_Price (Most Frequent Price).
* **Geographic Coverage:** The dataset spans **19 states**, **121 districts**, and **178 distinct markets**.
* **Commodity Diversity:** It covers **119 unique commodities** across **147 different varieties**.
* **Data Limitation:** A crucial aspect of this dataset is that all entries are recorded for a single date: **30/09/2025**. This makes it snapshot in time and does not allow for time-series analysis of price trends.

**Observed Insights & Hidden Facts**

The analysis revealed several key patterns and non-obvious facts about the Indian agricultural market structure.

**1. Uneven Geographic Distribution of Market Activity** There is a significant disparity in the number of market entries reported from different states. **Kerala** dominates the dataset with 279 entries, suggesting a high level of market activity or more comprehensive data reporting. In contrast, states like **Andhra Pradesh** (1 entry) and **Bihar** (2 entries) are vastly underrepresented. This highlights a regional imbalance in market participation or data collection.

**2. Significant Price Variation Across States and Commodities** Average modal prices show extreme variation based on both geography and commodity type.

* **By State:** **Nagaland** reports the highest average price across its commodities at ₹7,305.88, while **Bihar** reports the lowest at ₹2,300.00. This indicates that regional economic factors and local supply-demand dynamics heavily influence pricing.
  + **By Commodity:** High-value commodities such as **Black pepper** (average price ₹52,000) and **Arecanut** (₹48,954) command premium prices. In contrast, **Potato**, the most frequently listed commodity in the dataset (79 times), has a much lower average price, reflecting its status as a staple crop with high volume and lower unit cost.

**3. Price Volatility Hotspots** Price stability varies considerably by state. **Karnataka** exhibits the highest price variation, with a standard deviation of ₹9,756.14. This suggests a more volatile and less predictable market environment compared to other states like Maharashtra (₹2,259.01), which shows greater price stability. This insight is crucial for risk assessment by traders and farmers.

**4. Hidden Facts and Data Nuances**

* **Single-Day Data:** The analysis of the Arrival\_Date confirmed that all data points are from September 30, 2025. This fact makes it impossible to analyze price trends, seasonality, or market behavior over time.
* **Data-Driven Stability:** The commodity with the most "stable" price was found to be **Thondekai**, which showed no price variation. However, this is because it appears only once in the dataset, making its stability a statistical artifact rather than a market insight.

**Recommendations**

Based on the findings, the following actionable recommendations are proposed:

**For Policymakers:**

1. **Strengthen Market Reporting:** The wide disparity in market entries suggests that data reporting mechanisms in states like Andhra Pradesh and Bihar should be investigated and improved. Standardizing and enhancing data collection across all states would provide a more balanced and accurate national picture of agricultural trade.
2. **Address Price Volatility:** For states with high price variation like Karnataka, government intervention could focus on price stabilization mechanisms. This could include improving supply chain logistics, investing in modern storage facilities to manage supply shocks, and ensuring the effective implementation of Minimum Support Prices (MSPs) for volatile commodities.

**For Traders and Farmers:**

1. **Leverage Arbitrage Opportunities:** The significant price differences between states (e.g., Nagaland vs. Bihar) present clear opportunities for interstate trade and arbitrage. Farmers in low-price regions could form cooperatives to explore selling their produce in high-price markets to maximize revenue.
2. **Strategic Crop Selection:** The data highlights high-value commodities like Black pepper and Arecanut. Farmers in suitable agro-climatic zones could consider diversifying into these crops for higher profitability, while also balancing the production of high-volume staples like potatoes.

**For Future Analysis:**

1. A graph of states with red and green bars

   AI-generated content may be incorrect.**Implement Time-Series Data Collection:** The most critical recommendation is to expand data collection to cover daily or weekly entries. A longitudinal dataset would unlock powerful predictive analytics, enabling the forecasting of price trends, understanding seasonality, and building more sophisticated market models.

**Fig 1:** This bar chart illustrates the states with the highest and lowest number of market entries in the dataset, showing that Kerala and Uttar Pradesh have the most entries, while Telangana and Andhra Pradesh have the fewest.

A graph of a chart

AI-generated content may be incorrect.

**Fig 2:** This bar chart shows the top 5 commodities by average price, indicating that Black Pepper has the highest average price, followed by Arecanut (Betelnut/Supari), Coconut Oil, Cumin Seed (Jeera), and Mustard Oil.

A graph with orange bars

AI-generated content may be incorrect.

**Fig 3:** This bar chart indicates that commodities such as Mint, Rajgira, and Onion Green have the lowest average prices, showing minimal price variation in low-cost goods.

A graph of red bars

AI-generated content may be incorrect.

A graph of the states with the highest price variation

AI-generated content may be incorrect.**Fig 4:** This graph highlights that commodities like Black Pepper and Coconut Oil in Kerala and Karnataka have the highest recorded prices across states.

**Fig 5:** This chart shows that Karnataka has the highest price fluctuation among states, followed by Nagaland and Kerala, indicating significant regional price variability.

**A colorful pie chart with text

AI-generated content may be incorrect.Fig 6:** This pie chart visualizes the distribution of the top 10 states from the mandi dataset. It highlights that Kerala is the most significant contributor, representing the largest percentage of market entries among the top states.

**Conclusion:**

The analysis of the Vegetable Mandi (Market) dataset highlights the diversity and structure of agricultural trade across India. The data, consisting of 1,370 clean and complete records, spans 19 states, 121 districts, and 178 markets, showcasing a rich variety of 119 commodities and 147 varieties.

The findings reveal that Kerala has the highest representation, while Potato emerges as the most traded commodity. Price variables (Min, Max, and Modal Prices) show noticeable variation across states and commodities, reflecting regional market dynamics and commodity demand differences.

Overall, the dataset offers strong potential for agricultural market analytics, price trend prediction, and regional performance evaluation. It serves as a valuable resource for policymakers, traders, and data analysts to understand market patterns, support decision-making, and improve efficiency in India’s agricultural marketing system.