Mandi Price Analysis Report

Mandi Price Analysis Report

1. Dataset Description

- **Source:** Agricultural Mandi dataset (mandi.csv), loaded and processed using **PySpark** and **Pandas**.

- **Size:** 1370 records across multiple states in India.

- **Features:**

    - **State, District, Market** – geographical details of mandis.

    - **Commodity, Variety, Grade** – type of crop/vegetable and its classification.

    - **Arrival_Date** – trading/arrival date.

    - **Min Price, Max Price, Modal Price** – recorded prices for commodities.

- **Price Range:**

    - Minimum recorded price: **₹0**

    - Maximum recorded price: **₹65,300**

    - Modal (most common) price average: ~**₹3,764.**

2. Operations Performed

- **Data Cleaning:** Removed nulls, ensured numeric conversion of price columns

- **Descriptive Statistics:** Used .describe() for count, mean, min, max of prices

- **Exploratory Analysis:**

- o Computed mandi counts per state and visualized via bar chart.

- o Grouped commodities to analyze **price variability (standard deviation)**.

- o **Visualization:** Used **Matplotlib** for state-wise mandi counts and commodity distribution.

- o **Commodity Insights:** Checked for commodities with most stable and most volatile prices.

3. Key Insights

1. **Geographic Spread:**

   - o States like **Gujarat** and **Andhra Pradesh** have a higher concentration of mandis.

   - o Smaller states/UTs (e.g., **Chandigarh**) have fewer mandis.

2. **Commodity Pricing:**

   - o Prices vary widely, with some crops (e.g., **Cotton**) having very high market values.

   - o Certain vegetables show **high volatility**, while some (e.g., Thondekai) showed little recorded variance.

3. **Outliers:**

   - o Presence of extreme max prices (₹65,000+) compared to average (~₹3,700) suggests either luxury commodities (like Cotton, Yam) or data anomalies.

4. **Seasonal Patterns:**

   - o Dataset is timestamped (30/09/2025 entries visible), indicating daily mandi price tracking. More time-series analysis could reveal seasonal trends.

4. Recommendations

1. **Price Monitoring:**
   Deploy automated alerts for sudden spikes or drops in commodity prices to support farmers and traders.

2. **Commodity Stabilization:**
   Identify crops with highest price volatility and consider introducing **minimum support prices (MSP)** or storage interventions.

3. **Data Quality Improvement:**

   o Fix anomalies like **0 values** in Min/Max price.

   o Ensure consistent formatting of Variety and Grade.

4. **Future Analysis:**

   o Perform **time-series forecasting** to predict seasonal price changes.

   o Conduct **state-level comparisons** to inform agricultural policy.

   o Use clustering to group commodities with similar price behaviors.