Humanitarian Risk Analytics

An Integrated Dashboard Project Using SQL, Python, ML, and Power BI

Saurabh sagar
August 2025

Turning data into direction for vulnerable communities

Executive Summary

Synthesizing Poverty-Weighted Food Price Insights for Humanitarian Decision-Making

This project analyzes poverty-weighted food prices across Indian markets to identify regions facing disproportionate economic stress. It integrates SQL cleaning, Python modeling, and Power BI dashboarding to deliver actionable insights for humanitarian response.

Key Findings

- Bihar shows highest price burden
- Rice and wheat are most volatile
- Cluster 1 reveals compounded vulnerability
- Outliers flag hidden stress zones

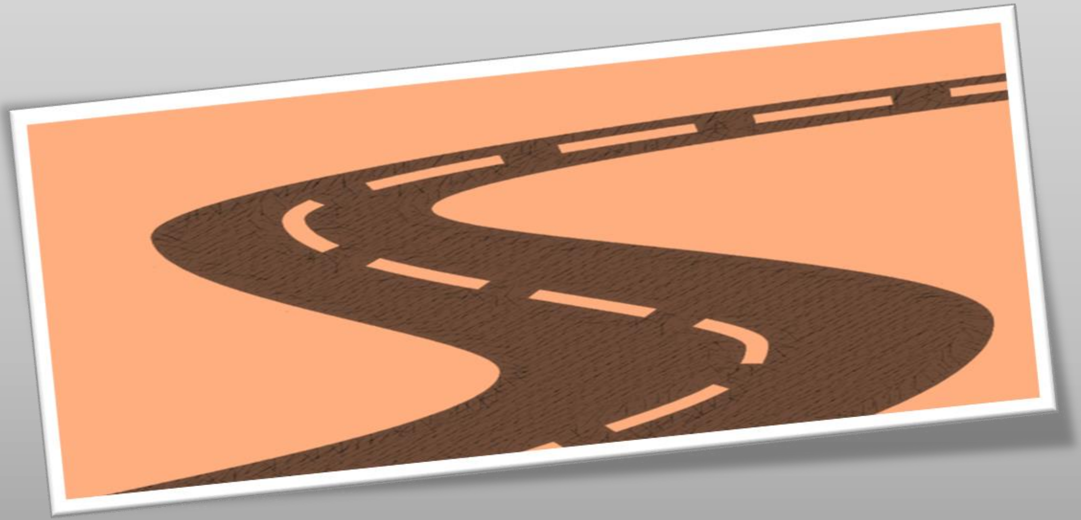
Impact

- Policy targeting based on spatial and commodity-specific risk
- Scalable reporting for humanitarian agencies and state governments
- A replicable model for risk analytics in other regions or sectors

Project Overview

From Raw Data to Humanitarian Insight

This project analyzes poverty-weighted food prices across India using a structured, multi-phase workflow. From SQL-based data cleaning to machine learning clustering and Power BI dashboarding, each phase builds toward actionable insights for humanitarian decision-making.



Phases:

- **Project Setup** – Folder structure, reproducibility, documentation
- **Data Cleaning & SQL Setup** – Normalize, filter, and join datasets
- **Exploratory Analysis** – Price trends, poverty correlations
- **ML Modelling** – Regression, clustering, SHAP interpretation
- **Dashboarding** – Interactive visuals, slicers, annotations
- **AI-Augmented Reporting** – Executive summaries, policy framing
- **Final Report** – PowerPoint deck with visuals and insights

Folder Structure and Workflow Setup

Organized for Reproducibility and Collaboration

This structure supports modular development across SQL, Python, and Power BI. Each phase is documented and version-controlled, enabling reproducibility and collaboration. The Report/ folder houses the final PowerPoint deck, while README.md and CHANGELOG.md track project evolution.

```
Humanitarian_Risk_Analytics_India/
├── data/                # Raw and cleaned datasets
├── sql/                 # SQL scripts
├── python/              # Jupyter notebooks
├── Power_BI/            # Power BI pdf or .pbix file
├── Screenshots/         # Screenshots of dashboard (png)
├── report/              # Final PowerPoint deck (PPTX and PDF)
├── README.md            # Project summary
```

Folder Structure

- **Project Setup** – Folder structure, reproducibility, documentation
- **Data Cleaning & SQL Setup** – Normalize, filter, and join datasets
- **Exploratory Analysis** – Price trends, poverty correlations
- **ML Modelling** – Regression, clustering, SHAP interpretation
- **Dashboarding** – Interactive visuals, slicers, annotations
- **AI-Augmented Reporting** – Executive summaries, policy framing
- **Final Report** – PowerPoint deck with visuals and insights

SQL Phase: Cleaning & Integration

Structuring Food Price and Poverty Data for Analysis

This project analyzes poverty-weighted food prices across India using a structured, multi-phase workflow. From SQL-based data cleaning to machine learning clustering and Power BI dashboarding, each phase builds toward actionable insights for humanitarian decision-making.

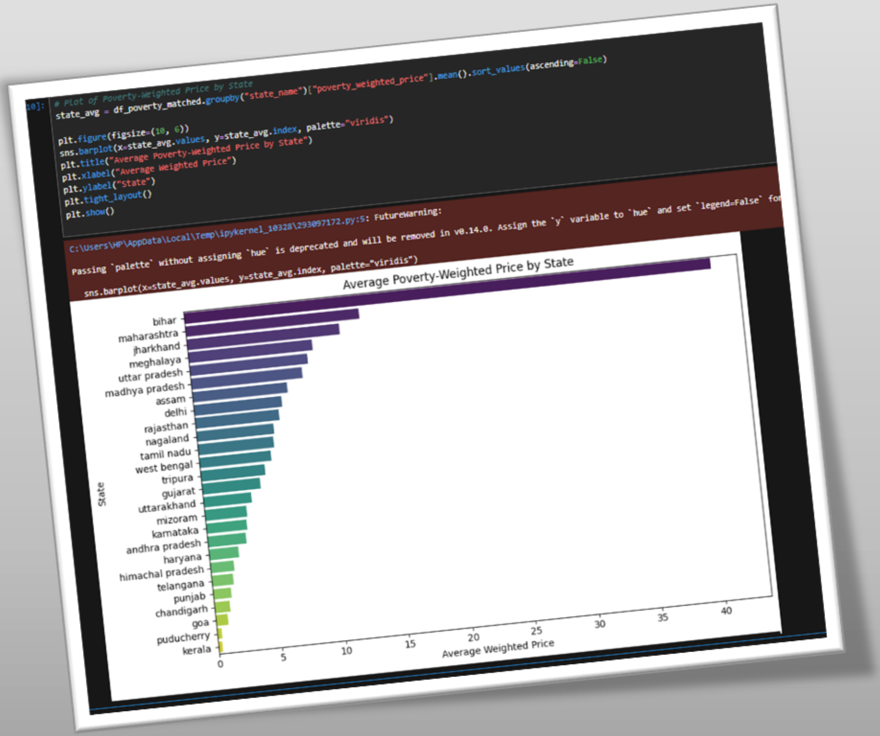
The SQL phase focused on transforming raw food price and poverty datasets into a clean, analysis-ready format. Key steps included:

- Dropping rows with missing or zero prices
- Normalizing **units** (e.g., per kg/liter)
- Standardizing **date** formats
- Filtering to recent years **(2015–2025)**
- Cleaning poverty indicators and aligning time periods
- Joining datasets on **provider_admin1_name** and **reference_period_start**

EDA: Price Trends & Poverty Correlations

Early Insights from Python Notebooks

Using Python (pandas, matplotlib, seaborn), we explored food price trends across states and their relationship to poverty indicators



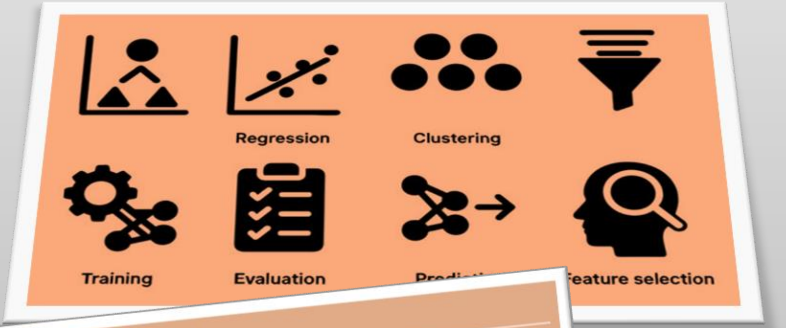
Key insights include:

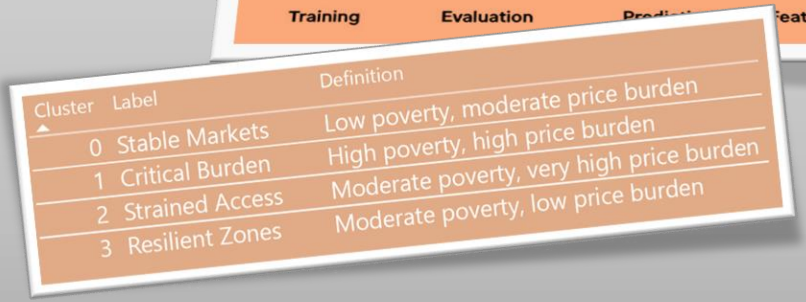
- Bihar shows consistently high poverty-weighted food prices
- Volatility is highest in rice, wheat, and sugar
- Positive correlation between MPI and price-to-poverty ratio
- Certain states (e.g., Jharkhand, Odisha) show price spikes despite moderate poverty

ML Modeling: Vulnerability Clusters & Risk Drivers

Using KMeans and SHAP to Surface Humanitarian Stress

Using Python (pandas, matplotlib, seaborn), we explored food price trends across states and their relationship to poverty indicators





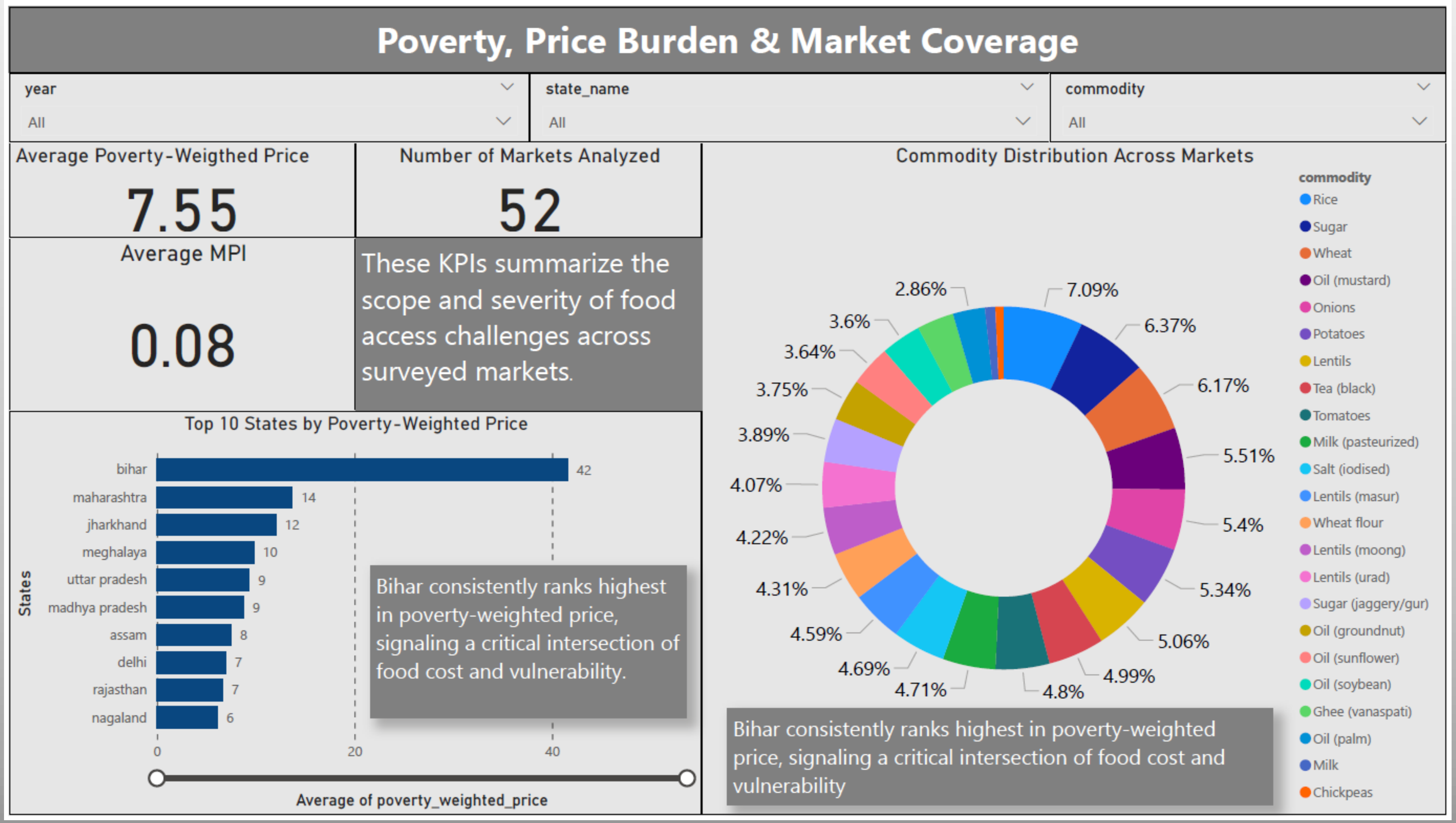
Cluster	Label	Definition
0	Stable Markets	Low poverty, moderate price burden
1	Critical Burden	High poverty, high price burden
2	Strained Access	Moderate poverty, very high price burden
3	Resilient Zones	Moderate poverty, low price burden

Machine learning was used to segment markets and explain price burden dynamics:

- **KMeans Clustering** grouped markets into 4 vulnerability profiles based on MPI, poverty ratio, and price burden indicators
- **Cluster 1 ("Critical Burden")** revealed markets with deep poverty and high food prices
- **SHAP Analysis** explained which features most influenced price-to-poverty ratio—MPI and volatility were top drivers

Dashboard Overview: National Trends & KPI Highlights

Poverty-Weighted Food Prices Across India



Key Insights from Overview Dashboard

Poverty-Weighted Food Prices Across India

Key Findings:

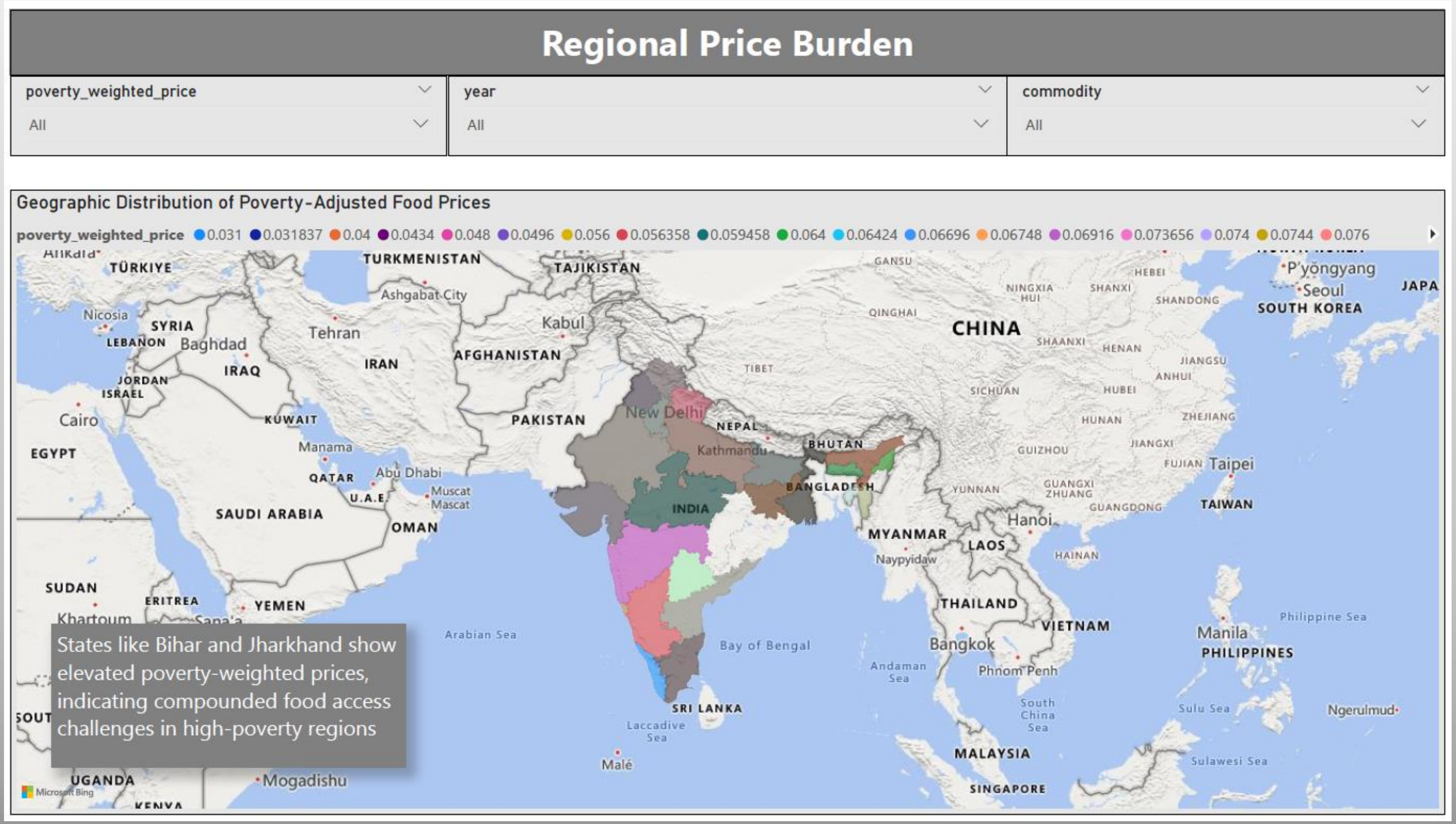
- Bihar shows the highest poverty-weighted price burden across all states
- Rice and wheat dominate consumption but show price volatility in vulnerable regions
- KPI cards highlight average MPI, price-to-poverty ratio, and top 5 vulnerable states
- Pie chart reveals commodity distribution, with rice and wheat comprising the majority

Implications & Suggested Actions:

- Prioritize Bihar and Jharkhand for price stabilization programs
- Monitor rice and wheat markets for volatility spikes
- Consider expanding poverty data collection in underrepresented regions
- Use dashboard KPIs for monthly humanitarian risk tracking

Dashboard Preview: Vulnerability Map

Mapping Poverty-Weighted Price Burden Across India



Key Insights from Vulnerability Map

Mapping Poverty-Weighted Price Burden Across India

Key Findings:

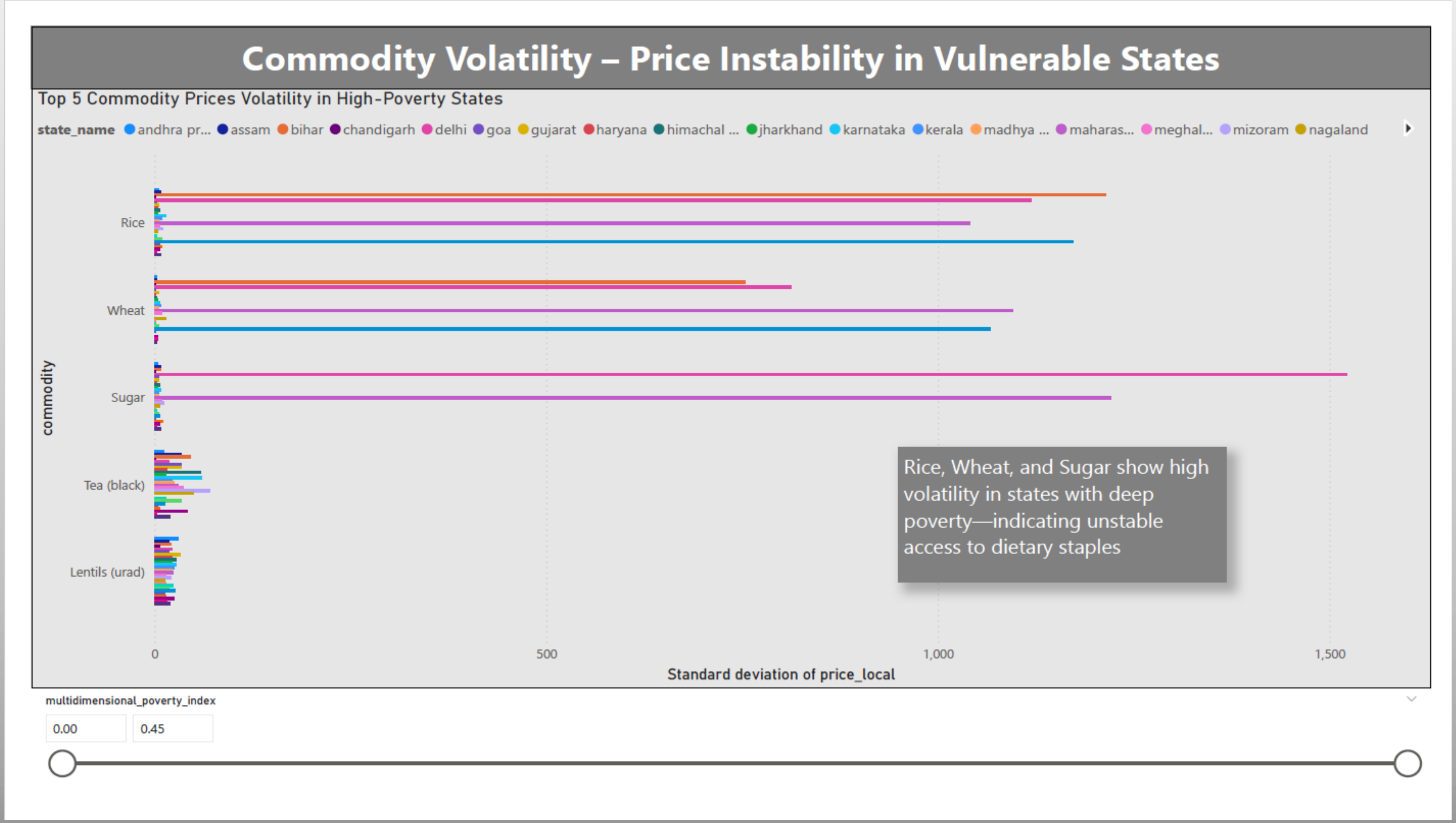
- Bihar, Jharkhand, and Odisha show **consistently high poverty-weighted prices**
- Southern states (e.g., Kerala, Tamil Nadu) show **lower burden despite high prices**
- Spatial clustering reveals **regional disparities in affordability**
- Commodity filters show **rice and wheat** as dominant drivers in high-risk zones

Implications & Suggested Actions:

- Prioritize **eastern states** for price stabilization and poverty relief
- Use spatial insights to **target food distribution and subsidy programs**
- Monitor **commodity-specific stress** in high-burden regions
- Consider integrating **district-level poverty** data for finer granularity

Dashboard Preview: Volatility Chart

Identifying Price Instability in Vulnerable Markets



Key Insights from Volatility Chart

Identifying Price Instability in Vulnerable Markets

Key Findings:

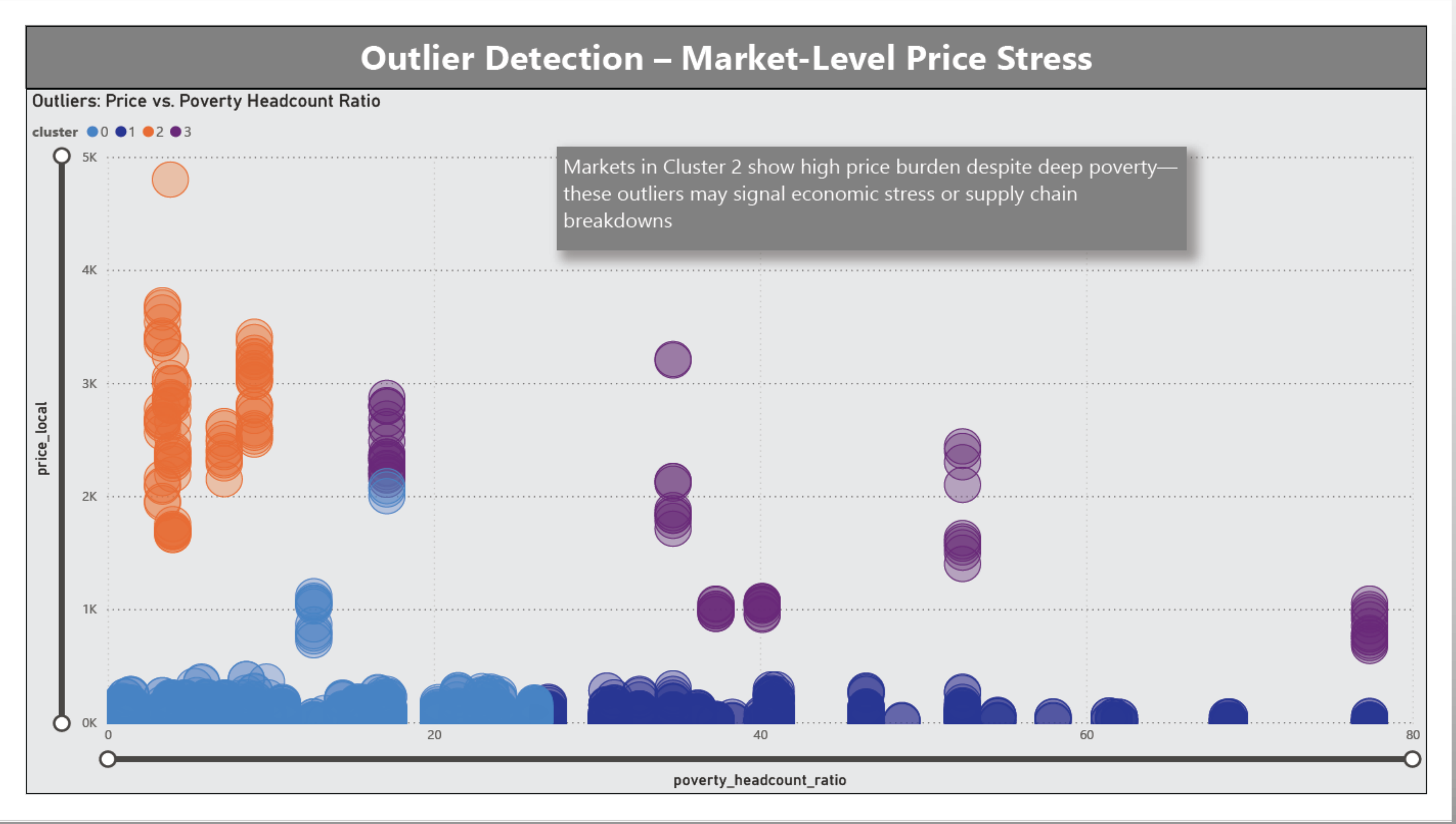
- **Rice, wheat, and sugar** show the highest volatility across multiple states
- **Bihar and Jharkhand** experience frequent price spikes in core staples
- Volatility is **not uniform**—some states show stable prices despite poverty
- **Seasonal patterns** may influence volatility in certain commodities

Implications & Suggested Actions:

- Prioritize **price monitoring systems** for volatile staples in high-poverty regions
- Consider **buffer stock policies** or subsidies for rice and wheat
- Use volatility insights to **time interventions** (e.g., pre-harvest or lean season)
- Recommend **commodity-specific dashboards** for deeper tracking

Dashboard Preview: Outlier Detection

Identifying Deviations in Poverty-Price Relationships



Key Insights from Outlier Detection

Identifying Deviations in Poverty-Price Relationships

Key Findings:

- Several markets show **high price-to-poverty ratios despite moderate MPI**
- Outliers are concentrated in **Bihar, Uttar Pradesh, and Chhattisgarh**
- Some commodities (e.g., sugar, pulses) show **unexpected spikes** in specific regions
- These outliers may reflect **supply chain disruptions or local inflation shocks**

Implications & Suggested Actions:

- Flagged markets should be **prioritized for field validation and monitoring**
- Consider deploying **targeted subsidies or price caps** in outlier zones
- Use outlier detection to refine **early warning systems** for humanitarian stress
- Recommend integrating **real-time market data** to improve responsiveness

Dashboard Preview: Vulnerability Clusters

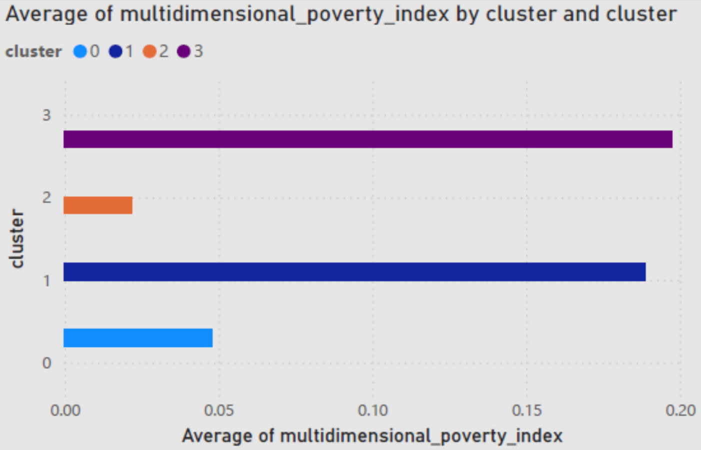
Clustering Markets by Poverty & Price Burden

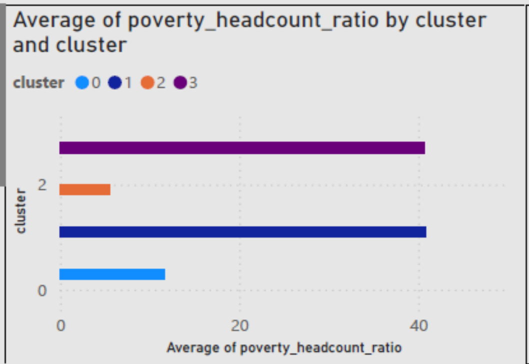
Cluster Profiles: Poverty & Price Burden Segmentation

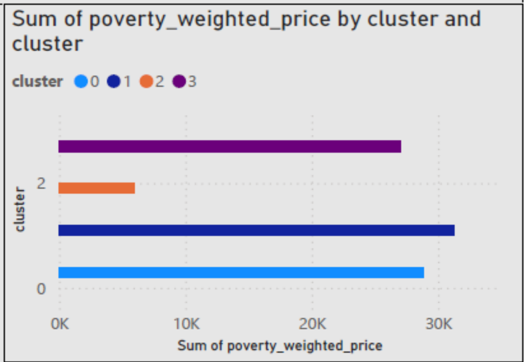
Cluster Summary: Poverty & Price Profiles

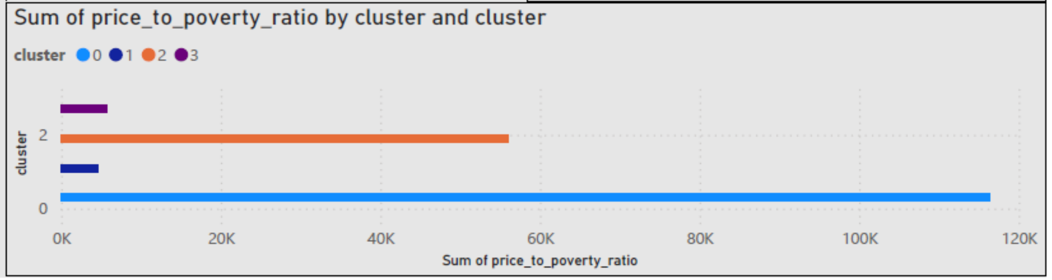
cluster	Average of multidimensional_poverty_index	Average of poverty_headcount_ratio	Average of poverty_weighted_price	Average of price_to_poverty_ratio	Count of market	Count of state_name
0	0.05	11.69	3.21	12.97	48	23
1	0.19	40.85	9.94	1.47	24	15
2	0.02	5.56	59.12	560.30	3	3
3	0.20	40.73	287.30	60.66	3	3
Total	**0.08**	**19.30**	**7.55**	**14.85**	**52**	**26**

Cluster	Label	Definition
0	Stable Markets	Low poverty, moderate price burden
1	Critical Burden	High poverty, high price burden
2	Strained Access	Moderate poverty, very high price burden
3	Resilient Zones	Moderate poverty, low price burden









Key Insights from Vulnerability Clusters

Clustering Markets by Poverty & Price Burden

Key Findings:

- **Cluster 1 (“Critical Burden”)** includes markets with high poverty and elevated food prices
- **Cluster 2 (“Strained Access”)** shows very high price burden despite moderate poverty
- **Cluster 0 (“Stable Markets”)** reflects low poverty and manageable prices
- Clustering reveals **distinct risk profiles** across regions—enabling targeted response

Implications & Suggested Actions:

- Use cluster labels to **prioritize humanitarian interventions**
- Recommend **price stabilization** in Cluster 1 zones
- Monitor Cluster 2 for **emerging stress despite moderate poverty**
- Share cluster definitions with stakeholders for **policy alignment and targeting**

AI-Augmented Summary

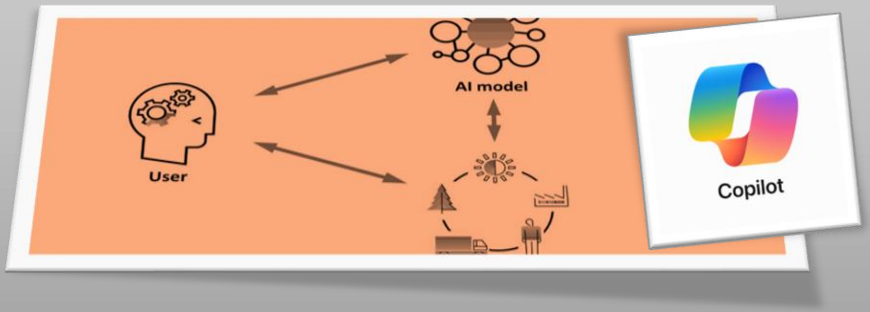
Synthesizing Insights for Decision-Makers

Copilot-generated insights from dashboard analysis reveal:

- **Cluster 1 markets** face compounded risk—deep poverty and elevated food prices
- **Volatility in core staples** (rice, wheat) amplifies humanitarian stress
- **Outlier markets** may signal emerging crises or data gaps
- Spatial and commodity filters enable **targeted monitoring and response**

This dashboard enables:

- **Prioritization of high-risk zones** for policy intervention
- **Commodity-specific planning** for food security programs
- **Scalable reporting** for humanitarian agencies and state governments
- A replicable model for **risk analytics in other regions**



Policy Recommendations

Targeted Actions Based on Dashboard Insights

Strategic Recommendations

- **Stabilize prices** in Cluster 1 markets through targeted subsidies
- **Monitor rice and wheat volatility** in high-poverty states like Bihar and Jharkhand
- Expand **poverty data collection** in underrepresented regions for better targeting
- Deploy **early warning systems** using outlier detection and volatility tracking
- Use dashboard KPIs for **monthly humanitarian risk reporting**

Implementation Suggestions

- Partner with **state governments** for localized interventions
- Integrate dashboard into **existing food security programs**
- Share cluster definitions with **NGOs and policy think tanks**
- **Recommend district-level dashboards** for granular monitoring
- Encourage **open data sharing** to improve transparency and collaboration

Closing & Next Steps

Scaling Impact Beyond This Dashboard

Summary Statement

This dashboard project integrates food price and poverty data to surface humanitarian risk across India. Through SQL, Python, ML modeling, and Power BI, it delivers actionable insights for policy, monitoring, and intervention.

Next Steps

- Expand to district-level granularity for finer targeting
- Integrate real-time market feeds for dynamic updates
- Collaborate with state agencies and NGOs for deployment
- Package dashboard as a public-facing tool or GitHub repo
- Explore replication in other countries or sectors (e.g., health, education)

Thank You

For Your Time, Attention, and Commitment to Data-Driven Impact

Saurabh sagar
August 2025
GitHub: [lfsaurabh/Humanitarian-risk-analysis-india]

Turning data into direction for vulnerable communities