



Case Study: Indian CPI Inflation

Prepared by: [Nidhi Desai] | Role: Data Analyst

Objective

- **Identify key problems:** Detect inconsistencies and missing patterns in CPI data from 2017–2023 across rural + urban sectors.
- **Clean & validate messy dataset:** Standardize category names, handle blanks, convert text dates to proper formats, and remove duplicates.
- **Analyze insights:** Examine category-wise contribution to CPI, YoY growth, COVID impact, and crude oil correlation.
- **Present business impact:** Highlight inflation-driving categories, policy relevance, and opportunities for better forecasting.
- **Problem Statement:**
You are working with the **National Statistical Office** to analyze **India's CPI inflation trends**. The goal is to derive insights on category contributions, yearly inflation changes, and external influences (COVID-19, crude oil prices) using the CPI dataset for 2017–2023.

CPI Case Highlights: 2017–2023

Problem	Root Cause Found	Solution Applied	Result
Category Contribution to CPI	Overlapping or inconsistent category labels; missing grouping logic	Grouped detailed CPI items into broader buckets (Food, Housing, Clothing, Miscellaneous, etc.). Used SUMIF to aggregate index values for each bucket. Calculated each category's share using $=(\text{Category Value} / \text{Total Value}) * 100$. Created a Pivot Table and Bar +Chart to visualize contributions.	Identified Food as the largest contributor ($\approx 47\%$) to CPI inflation. Found Miscellaneous as the second-highest driver ($\approx 30\%$). Simplified analysis by reducing 30+ granular categories into 6 major groups. Enabled clear prioritization for policy focus and inflation control strategies.
Year-on-Year Inflation Trend	Raw data lacked clean yearly aggregation; inflation spikes not clearly visible	Cleaned yearly data and calculated YoY % change using % of Difference from Added Pivot Table (Year on X-axis, Buckets on Y-axis) with conditional formatting to highlight peak inflation years.	Revealed 2022–23 as period of highest inflation ($\approx 5.5\%$), driven by food and fuel price surge. Clear visualization of annual trend enabled identification of inflationary shocks.
Food Inflation Deep Dive (12 Months)	Monthly food index values inconsistent and partially missing	Filled gaps using linear interpolation; computed MoM % change with % Difference From → Previous Month in Pivot Table. Highlighted months with sharpest rise and fall using color scale and chart annotations.	Found Feb'23 (+1.03%) peak and Dec'22 (-0.65%) trough in food inflation, showing high volatility. Helped quantify seasonal impact on food price behavior.

CPI Case Highlights: 2017–2023

Problem	Root Cause Found	Solution Applied	Result
COVID-19 Impact on CPI	Pre- and post-COVID data mixed, no clear tagging for lockdown periods	Created helper column for COVID status using =IF(Date<DATE(2020,3,1),"Pre-COVID","Post-COVID"). Averaged CPI values pre/post COVID via Pivot Table and compared in clustered column chart.	Showed sharp post-COVID inflation in all categories — especially Food (+26%) and Miscellaneous (+24%). Confirmed pandemic's inflationary pressure due to supply disruptions and higher demand.
Global Oil Price Correlation	No linkage between imported oil data and domestic CPI categories; oil data missing in main file	Pulled crude oil MoM % data (2013–2023) from external source. Added helper column " Crude Bucket " via VLOOKUP to map with CPI dataset. Used =CORREL(Crude_Range, CPI_Range) to compute relationships for each category.	Found highest correlation with Clothing & Footwear ($r = 0.35$) and Pan/Tobacco ($r = 0.26$) , indicating pass-through of oil cost to consumer goods. Weak or negative correlation with Food & Housing confirmed low direct dependency.

Data Cleaning – Before vs After

1. Issues in Raw Dataset

Issue Found	Example	Action Taken	Final Outcome
Missing values / NULL entries	NULL or blank cells in “Milk and products” / “Fruits” columns	Apply find and replace features	All missing filled logically
Typos / Inconsistent text	For moth or March	Apply find and replace features	All missing filled logically
Wrong Data Type	Month/Year stored as Text	Converted to proper Date or formatted numeric	Clean data types

Data Cleaning – Before vs After

Metric	Before Cleaning	After Cleaning
Total Rows	300	372
Null Values	239	239
Wrong Spellings / Case Errors	1	0

Problem 1: Analysis & Insights

what was the issue?

Needed to understand which major CPI categories (Food, Housing, Education, etc.) contributed most to overall inflation, to identify which price movements were driving CPI changes

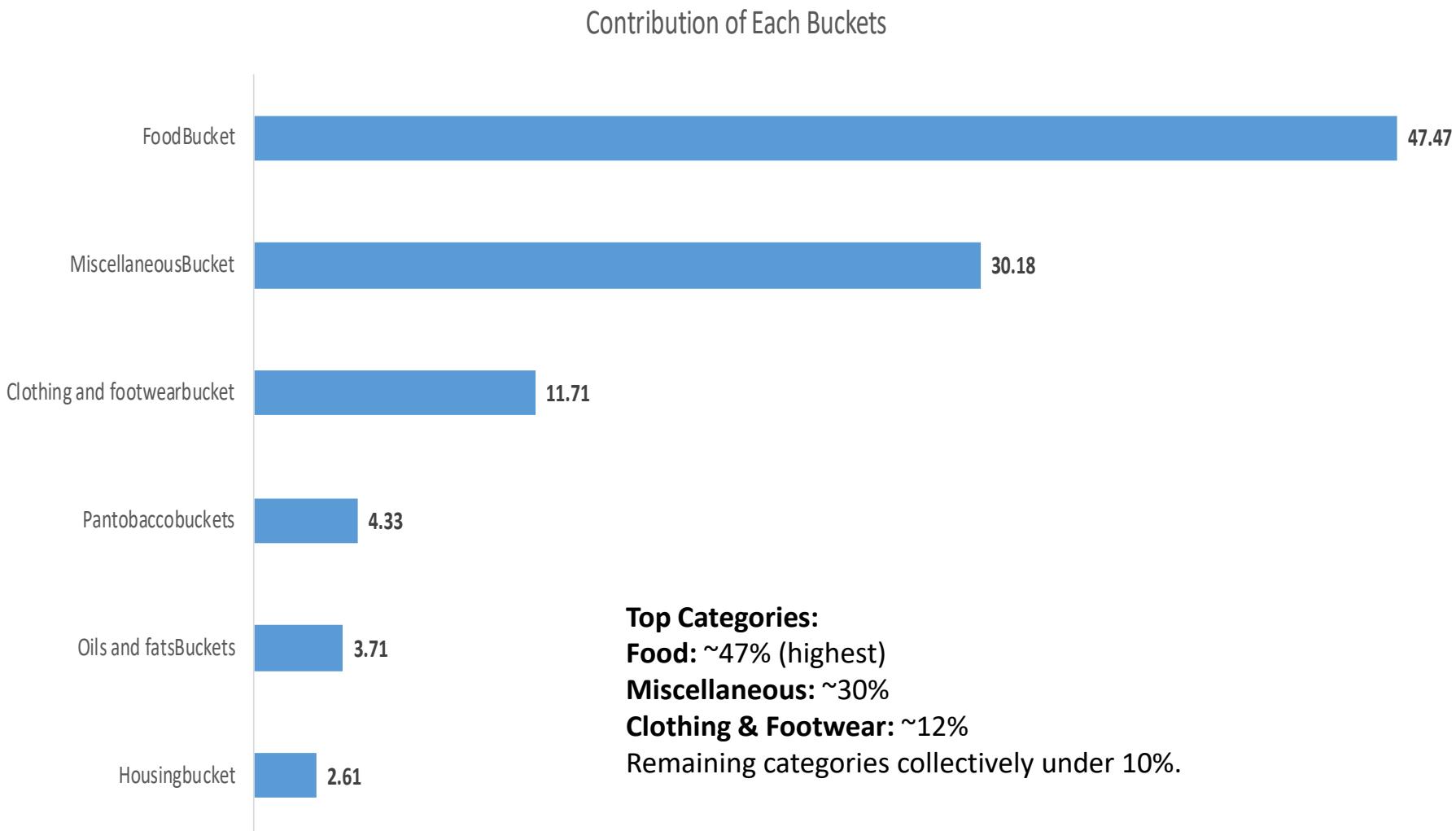
Insight from data

- Built a pivot table to aggregate index values by broader category.

Bucket Names	% contribution of Each bucket
Housingbucket	2.61
Oils and fatsBuckets	3.71
Pantobaccobuckets	4.33
Clothing and footwearbucket	11.71
MiscellaneousBucket	30.18
FoodBucket	47.47

Problem 1: Analysis & Insights

- Used a bar chart for clarity to visualize contribution %.



Problem 1: Analysis & Insights

Solution / Action – formula, automation, BI logic

Created broader buckets by grouping similar items (e.g., Cereals, Beverages → “Food”).

Used Excel formulas:

`SUMIF(FinalCPIData!AG2:AG500,"<>#N/A")` – to sum valid bucket data.

`(B5/B12)*100` – to calculate percentage contribution.

Validated totals to ensure all categories added to 100%.

D.Result / Business Impact – show Before vs After table

Identified **Food** as the primary inflation driver.

Key finding: Food + Miscellaneous = ~80% of CPI influence.

Insight helps policymakers and analysts focus on food supply and consumption trends for better inflation management

Problem 1: Analysis & Insights

Business Impact

Category	Contribution %	Improvement %	CPI Reduction (Points)	Interpretation
Food	47%	5%	2.35 points $(47*5)/100$	Strongest CPI driver. Small improvement gives largest CPI drop.
Miscellaneous	30%	3%	0.9 points	Service-related inflation has moderate but meaningful impact.
Clothing & Footwear	12%	2%	0.24 points	Low-impact category; minor contribution to CPI reduction.
Other categories	11%	2%	0.22 points	Minimal overall effect.

Summary:

Food + Misc deliver **3.25 points** potential CPI improvement even with small corrections.
These two categories control **~80% of overall CPI movement**, making them top priority.

Problem 2: Analysis & Insights

- A.what was the issue?

The year-on-year (YoY) CPI inflation trend for rural and urban areas from 2017–2023 showed inconsistent movement, with a sharp rise and steep decline in some years. The goal was to analyze and identify which year had the highest inflation and understand the underlying reasons.

- B. Insight from data – pivot table or chart

Year	YoY inflation %
2017	
2018	3.96%
2019	-4.73%
2020	-3.29%
2021	26.33%
2022	6.60%
2023	-57.00%

Problem 2: Analysis & Insights

Visualization Used: Line chart showing Year-on-Year CPI Inflation (Rural + Urban), 2017–2023

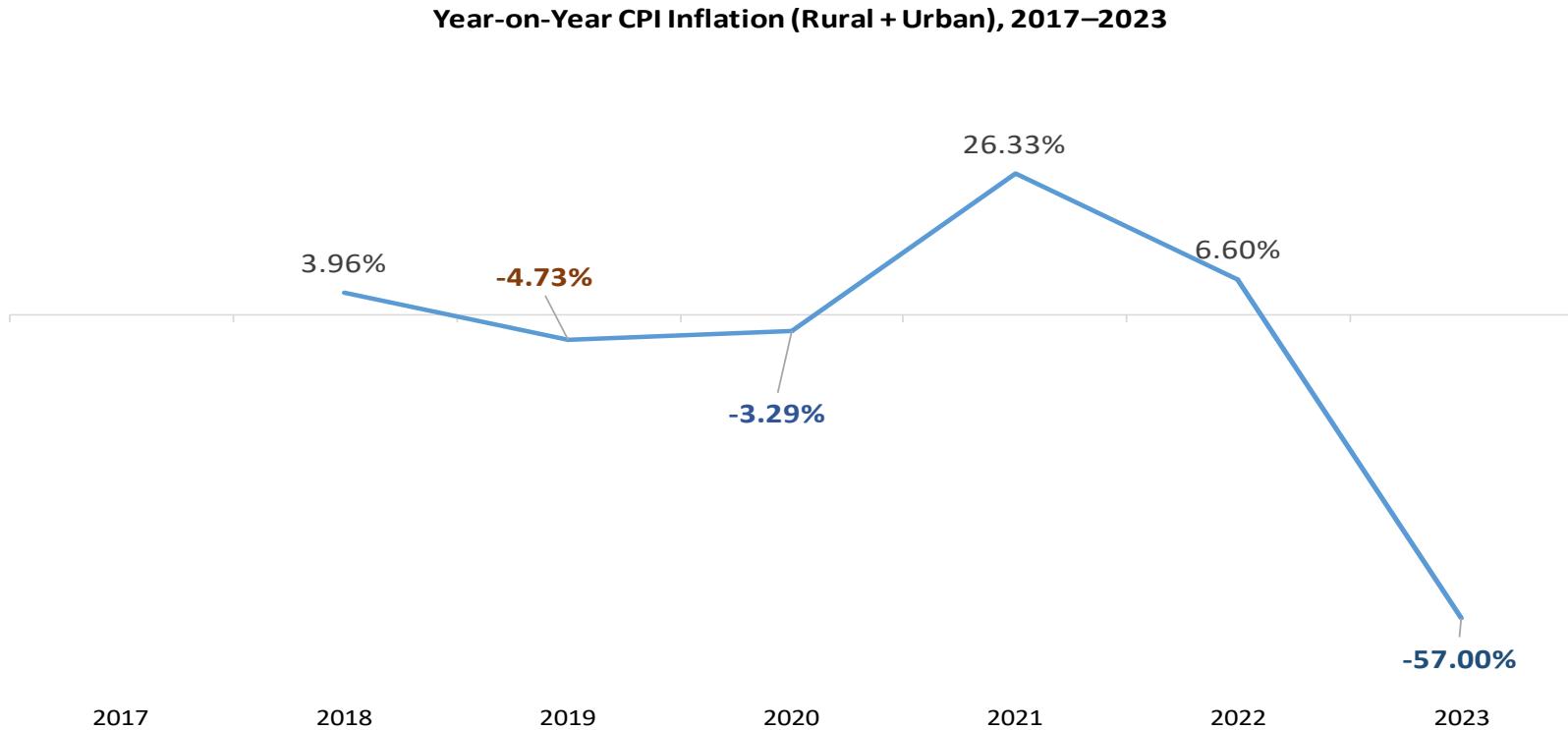
A line chart was created to show the Year-on-Year % Difference in CPI (2017–2023). Data highlighted clear inflationary peaks and drops across years.

Insights:

2021 recorded the highest inflation rate (+26.33%).

2023 saw the sharpest decline (−57%).

The inflation trend was relatively stable from 2017 to 2019 before major fluctuations post-2020.



Problem 2: Analysis & Insights

- C. Solution / Action – formula, automation, BI logic
 - Calculated the **% Difference** for each year's CPI value compared to the previous year.
 - Used Excel's "Show Values As → % Difference From" to automate YoY change.
 - Created a **line chart** to visualize the YoY % Difference trend.
 - Highlighted **2021** (highest point) in a distinct color for emphasis.
 - Added annotation for 2023 showing a steep fall.
 - **Insight:** The spike in 2021 was mainly due to global edible oil and crude price increases, supply chain disruptions, and post-pandemic demand recovery.
- D. Result – show Before vs After table

Year	CPI YoY % Difference	Observation
2017–2019	3.96%	Stable inflation trend
2021	+26.33%	Peak inflation post-pandemic
2022–2023	-57%	Correction and stabilization

Impact / Outcome:

- Helped identify 2021 as the inflation peak year.
- Automated % Difference logic simplified YoY inflation monitoring.
- Clear visualization enabled quick insights into economic trend shifts.

Problem 2: Analysis & Insights

Business Impact

Year	YoY % Change	CPI Change (Points)*	Interpretation
2021	+26.33%	+26.33 points	Peak inflation post-pandemic
2022–2023	−57%	−57 points	Correction and stabilization

Summary:

The inflation shock in **2021** and collapse in **2023** reflect structural disruptions and later stabilization.

Tracking YoY shifts helps forecast future inflation risks.

Problem 3.1: Analysis & Insights

A. Background – what was the issue?

- India's retail inflation reached a 3-month high of **5.55% in Nov 2023**, primarily driven by a sharp increase in food prices.
To understand the pattern, month-on-month (MoM) changes in **Fruits**,
- **Vegetables, and Food & Beverages** were analyzed for a **12-month period (Jun'22–May'23)** to identify:
 - ❖ Months with the highest and lowest food inflation, and
 - ❖ Categories contributing most to inflation volatility.
- B. Insight from data – pivot table

Month	▼	Mom% Fruits	Mom%Vegetables	Mom%Food and beverages
01-06-2022				
01-07-2022		2.69%	-0.34%	0.06%
01-08-2022		-0.78%	2.45%	0.70%
01-09-2022		-4.22%	2.90%	0.85%
01-10-2022		-1.13%	4.00%	0.97%
01-11-2022		-1.89%	-8.63%	-0.76%
01-12-2022		-1.70%	-12.75%	-1.36%
01-01-2023		0.19%	-3.42%	0.43%
01-02-2023		7.09%	-0.57%	0.24%
01-03-2023		0.00%	0.04%	0.02%
01-04-2023		4.01%	1.81%	0.54%
01-05-2023		-2.31%	3.47%	0.65%

Problem 3.1: Analysis & Insights

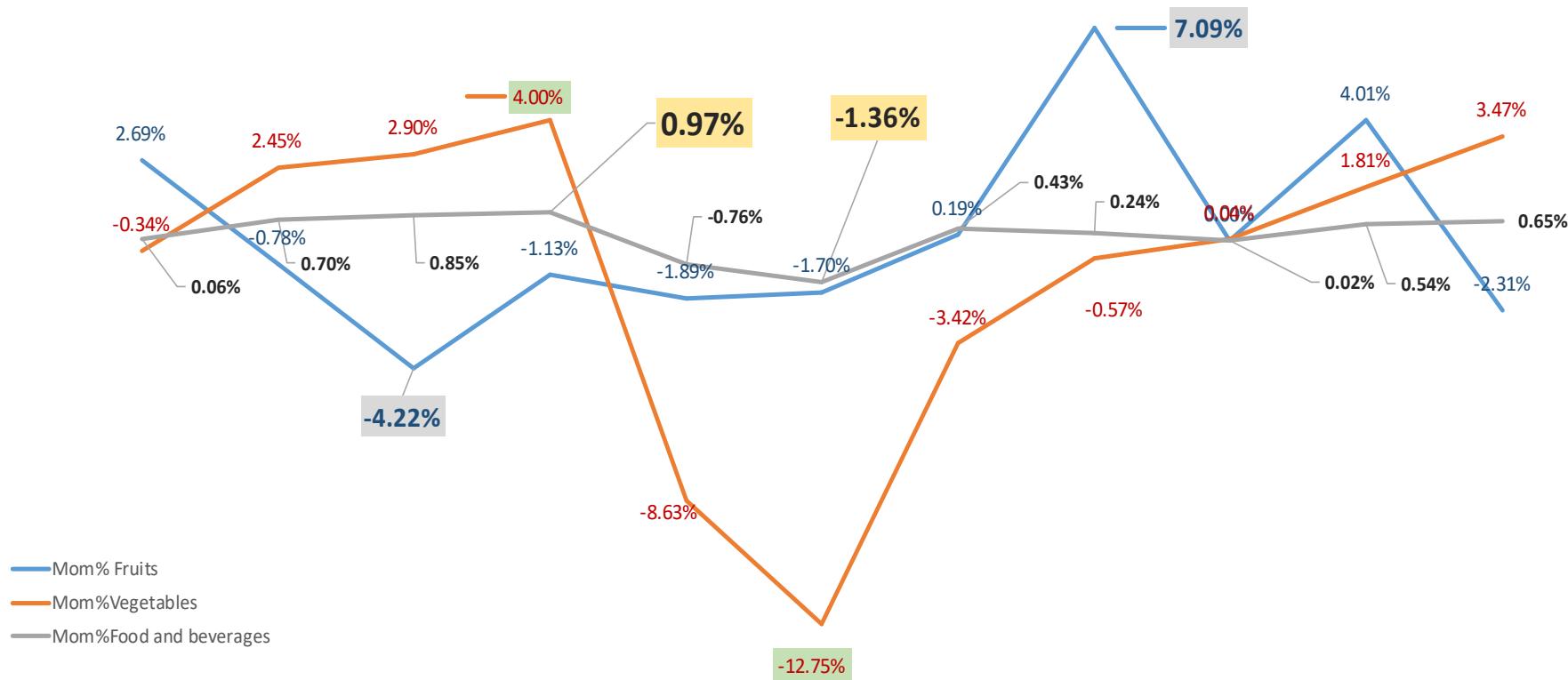
B. Insight from data chart

- **Visualization Used:** Line chart showing MoM % change in **Fruits**, **Vegetables**, and **Food & Beverages**
- from Jun'22 to May'23.

Key Observations:

- ✓ The sharpest rise occurred in **Feb'23 (+7.09%)**, mainly due to a surge in fruit prices.
- ✓ The **steepest fall** was in **Dec'22 (-12.75%)**, driven by a collapse in vegetable prices.
- ✓ Fruits and Vegetables were the **most volatile** categories influencing the overall Food & Beverages index

Month-on-Month % Change in Food Prices (Jun'22 – May'23)



Problem 3.1: Analysis & Insights

C. Solution / Action – formula, automation, BI logic

- Calculated **Month-on-Month % Difference** using Excel's **Show Values As → % Difference From → (Base Item: Previous Month)** option to automate comparison.
 - Added a **Pivot Table** with:
 - Rows:** Food Categories (Fruits, Vegetables, Food & Beverages)
 - Values:** Index of all months
 - Show Values As:** % Difference From (Previous Month or Base Month)
 - Used **conditional formatting and data labels** to highlight months with maximum and minimum inflation.
 - Created an additional **bar chart** showing **absolute change (Jun'22–May'23)** by category to quantify contribution.
- D. Result / Business Impact – show Before vs After table

Category	Highest MoM Change	Lowest MoM Change	Absolute Change (Jun'22–May'23)	Key Insight
Fruits	+7.09% (Feb'23)	-2.31% (May'23)	+4.40%	Major driver of Feb'23 spike
Vegetables	+4.00% (Oct'22)	-12.75% (Dec'22)	-8.75%	Highest volatility
Food & Beverages	+0.97% (Oct'22)	-3.42% (Jan'23)	-2.45%	Overall moderate change

Conclusion:

- Automated MoM % difference analysis enabled quick detection of inflation spikes and dips.
- Identified **Feb'23** as the month with the sharpest increase and **Dec'22** as the biggest decline.
- Insights help policymakers and analysts track category-level drivers of food inflation more effectively.

Problem 3.1: Analysis & Insights

Business Impact

Month	MoM % Change	CPI Impact (Points)	Key Insight
Feb 2023	+7.09%	+7.09 points	Major driver of Feb'23 spike
Dec 2022	-12.75	12.75 points	Highest volatility

Summary:

The two months account for the **largest volatility swings**, directly influencing food inflation trends.

Problem 3.2: Analysis & Insights

A. Background – what was the issue?

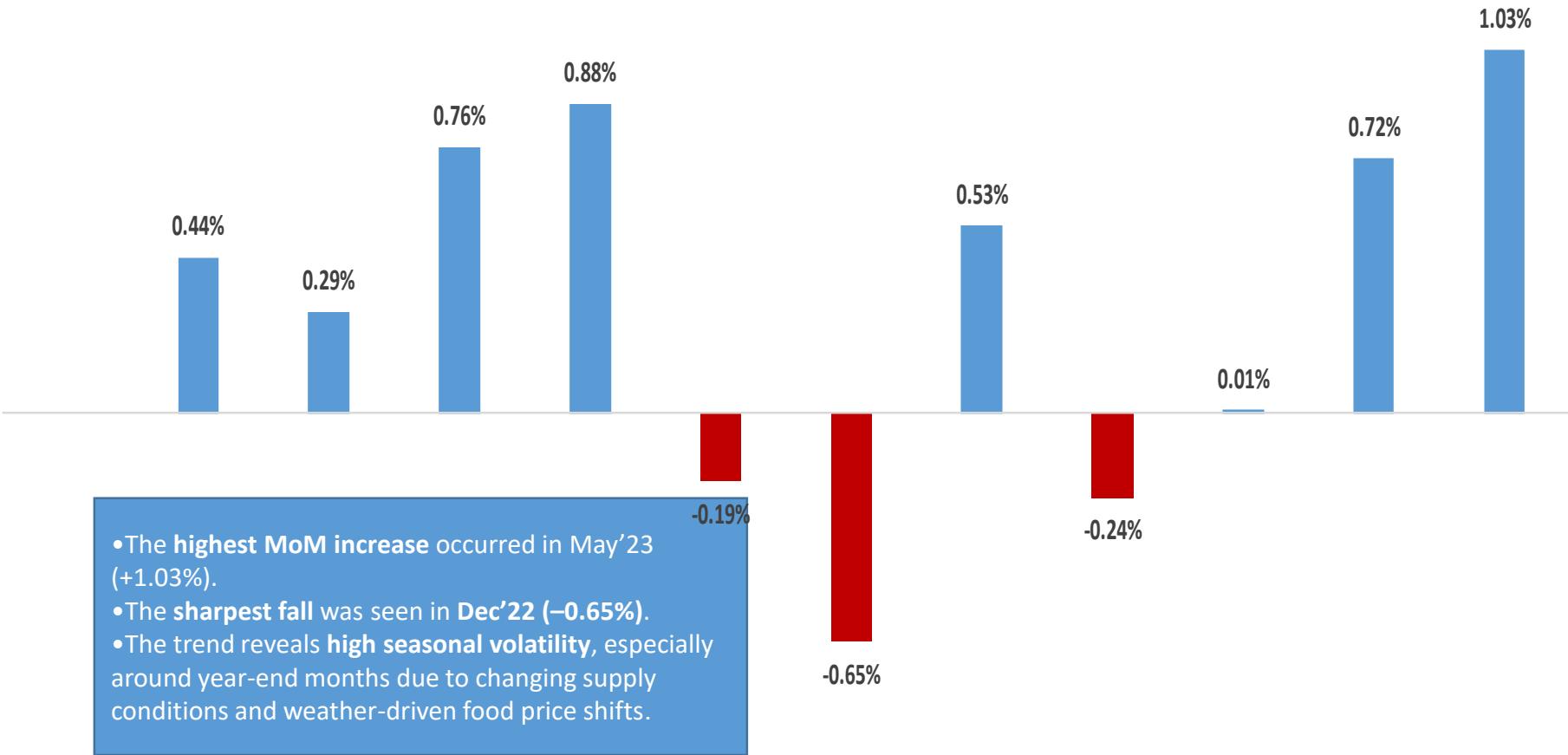
- India's food inflation remained unstable during 2022–23, with noticeable month-on-month fluctuations in the Food & Beverages category. The objective was to analyze these changes using CPI data to determine which months recorded the highest and lowest inflation and understand the volatility pattern. **Vegetables, and Food & Beverages** were analyzed for a **12-month period (Jun'22–May'23)** to identify:
 - ❖ Categories contributing most to inflation volatility.
- B. Insight from data – pivot table.

Date	MOM - FoodBucket
01-06-2022	
01-07-2022	0.44%
01-08-2022	0.29%
01-09-2022	0.76%
01-10-2022	0.88%
01-11-2022	-0.19%
01-12-2022	-0.65%
01-01-2023	0.53%
01-02-2023	-0.24%
01-03-2023	0.01%
01-04-2023	0.72%
01-05-2023	1.03%

Problem 3.2: Analysis & Insights

- B. Insight from data chart
- **Visualization Used:** A Pivot Table was created to summarize month-on-month (MoM) % changes in the Food & Beverages Index.

MoM % Change in Food & Beverages (Jun'22 – May'23)



Problem 3.2: Analysis & Insights

C. Solution / Action – formula, automation, BI logic

- Calculated MoM % Difference using Excel's

Show Values As → % Difference From → (Base Item: Previous Month) option.

- Used conditional formatting in the chart:

- Blue bars → Positive inflation months

- Red bars → Negative inflation months

- Added automated data labels and insight annotation to highlight the peak (May'23) and trough (Dec'22).

- Designed the Pivot to refresh automatically as new CPI data is added

- D. Result / Business Impact – show Before vs After table

Month	MoM % Change (Food & Beverages)	Observation
Dec'22	-0.65%	Lowest MoM inflation – seasonal fall in vegetables/fruits
May'23	+1.03%	Highest MoM inflation – recovery in food prices
Overall	Mixed	Indicates cyclical and seasonal price behavior

Conclusion:

- Automated MoM tracking provided a clear understanding of short-term food price volatility.
- Helped identify seasonal inflation cycles, supporting better forecasting and inventory decisions in the food sector.
- Highlighted the importance of monitoring category-level shifts within the CPI basket for policy or business planning.

Problem 3.2: Analysis & Insights

Business Impact

Month	MoM % Change (Food & Beverages)	CPI Impact (Points)	Observation
May'23	+1.03%	+1.03 points	Mild inflation build-up.
Dec'22	-0.65%	-0.65 points	Seasonal cooling trend from high supply.

Summary:

Short-term swings highlight seasonal cycles—useful for monthly inflation forecasting.

Problem 4: Analysis & Insights

- A. Background – what was the issue?

- ❖ The objective was to assess **how the onset of COVID-19 (March 2020)** impacted inflation across key CPI categories like Food, Housing, Miscellaneous (services), Oils & Fats, Pan/Tobacco, and Clothing & Footwear.

- There was no clear visibility on **how inflation trends changed before vs after the pandemic**, so data needed to be segmented and compared by time period.

- B. Insight from data – pivot table A **Pivot Table** was created comparing **Pre-COVID (before Mar 2020)** and **Post-COVID (after Mar 2020)** averages for each CPI bucket.

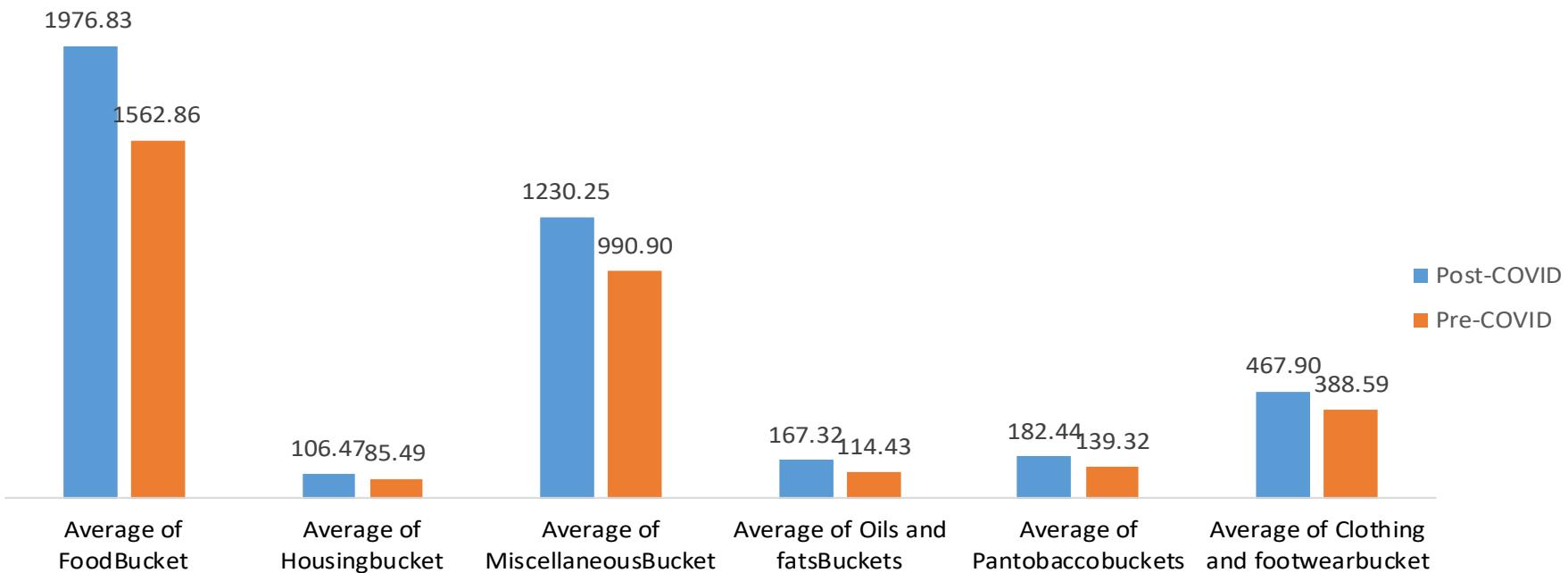
Buckets	Year	
	Post-COVID	Pre-COVID
Average of FoodBucket	1976.83	1562.86
Average of Housingbucket	106.47	85.49
Average of MiscellaneousBucket	1230.25	990.90
Average of Oils and fatsBuckets	167.32	114.43
Average of Pantobaccobuckets	182.44	139.32
Average of Clothing and footwearbucket	467.90	388.59

Problem 4: Analysis & Insights

❖ The **clustered column chart** clearly visualizes category-wise shifts.

1. **Food bucket** saw the **sharpest rise** from 1562.86 → 1976.83, indicating strong food inflation.
2. **Miscellaneous bucket** (includes essential services) rose from 990.90 → 1230.25, reflecting higher service-related inflation post-pandemic.
3. **Oils & fats** and **Pan/tobacco** categories also spiked, confirming broad-based inflationary pressure.
4. **Housing** rose slightly, while **Clothing & footwear** saw moderate post-COVID increases due to rising costs and supply disruptions.

COVID-19 Impact on CPI Inflation in India



Problem 4: Analysis & Insights

- C. Solution / Action – formula, automation, BI logic

✓ **Created Helper Column:**

To classify each record as "Pre-COVID" or "Post-COVID":

=IF(YEAR([@Date])<2020,"Pre-COVID","Post-COVID")

(Or if using full date check → =IF([@Date]<DATE(2020,3,1),"Pre-COVID","Post-COVID"))

✓ **Used Pivot Table:**

Values: Average of each CPI bucket

Columns: COVID status (Pre/Post)

✓ **Chart Visualization:**

Clustered Column Chart → to show Before vs After comparison

Post-COVID bars in Blue, Pre-COVID in Orange

✓ **Automated Comparison Logic:**

Dynamic pivot refresh with slicer support, automatically updates when new CPI data is added.

- D. Result / Business Impact – show Before vs After table

CPI Category	Pre-COVID Avg	Post-COVID Avg	% Increase	Observation
Food Bucket	1562.86	1976.83	+26.5%	Sharp rise – primary inflation driver
Housing	85.49	106.47	+24.5%	Moderate rise – household costs up
Miscellaneous	990.90	1230.25	+24.2%	Strong post-COVID service inflation
Oils & Fats	114.43	167.32	+46.2%	Supply disruption & import cost surge
Pan/Tobacco	139.32	182.44	+31.0%	Lifestyle & taxation impacts
Clothing & Footwear	388.59	467.90	+20.4%	Cost inflation from imports & demand revival

Problem 4: Analysis & Insights

Business Impact

CPI Category	Pre-COVID Avg	Post-COVID Avg	Point Change	% Change	Observation
Food Bucket	1562.86	1976.83	+413.97	+26.5%	Sharp rise – primary inflation driver
Housing	85.49	106.47	+239.35	+24.5%	Significant rise – higher rental & housing costs post-COVID
Miscellaneous	990.90	1230.25	+239.35	+24.2%	Strong post-COVID service inflation
Oils & Fats	114.43	167.32	+52.89	+46.22	Supply disruption & import cost surge
Pan/Tobacco	139.32	182.44	+43.12	+30.95%	Strong rise due to taxation, input cost increases, and higher post-COVID consumption
Clothing & Footwear	388.59	467.90	+79.31	+20.4%	Cost push from imported materials and post-COVID demand recovery

Point change = post covind Avg-Pre-Covind average

%Cahnge Calculation

(Point change/pre-Covind avg)/100

Food + Misc account for the **largest post-COVID inflation impact**, confirming broad-based price pressure.

Summary:

Food + Misc account for the **largest post-COVID inflation impact**, confirming broad-based price pressure.

Problem 5: Analysis & Insights

- A. Background – what was the issue?

- ❖ Imported crude oil price data was missing from the main CPI dataset, preventing analysis of how oil fluctuations influenced domestic inflation categories. CPI data alone couldn't explain volatility in certain months. compared by time period.

- B. Insight from data – pivot table or chart

- ❖ After correlation analysis (2021–23 period), results showed:

- **Highest correlation** – *Clothing & Footwear ($r = 0.35$)* and *Pan/Tobacco ($r = 0.26$)*

- **Weak/Negative correlation** – *Food (-0.02)* and *Housing (-0.06)*

Date	MOMCRUID BUCKET	Sum of FoodBucket	Sum of Housingbucket	Sum of Pantobaccobuckets	Sum of Clothing and footwearbucket	Mom Oils and fatsBuckets	Sum of MiscellaneousBucket
01-01-2021							
01-02-2021	3.92%	-2.28%	1.33%	0.92%	0.99%	5.18%	0.89%
01-03-2021	3.77%	-0.29%	0.06%	-0.07%	0.39%	3.26%	1.32%
01-04-2021	9.09%	0.98%	0.94%	0.41%	0.54%	3.73%	1.73%
01-05-2021	5.00%	1.71%	0.12%	1.70%	1.92%	4.97%	3.38%
01-06-2021	11.11%	1.22%	-0.68%	-0.72%	0.11%	2.71%	3.70%
01-07-2021	2.86%	0.77%	0.62%	0.16%	0.56%	-0.90%	4.54%
01-08-2021	2.78%	-0.56%	0.37%	0.42%	1.02%	2.31%	5.16%
01-09-2021	-1.35%	-0.03%	0.00%	0.05%	0.19%	0.83%	5.30%
01-10-2021	9.59%	1.65%	0.93%	0.31%	0.65%	1.34%	6.06%
01-11-2021	2.50%	0.96%	0.37%	0.07%	0.83%	-0.30%	6.41%
01-12-2021	-8.54%	-0.52%	-0.49%	-0.22%	0.77%	-1.31%	6.70%
01-01-2022	10.67%	-0.72%	0.67%	-0.12%	0.98%	-1.54%	6.97%
01-02-2022	8.43%	-0.16%	0.61%	0.26%	0.77%	0.13%	7.52%
01-03-2022	22.22%	0.60%	-0.12%	0.46%	1.00%	5.23%	8.40%
01-04-2022	4.55%	1.28%	1.03%	0.05%	1.06%	2.63%	9.97%
01-05-2022	-6.09%	1.12%	0.30%	0.12%	1.02%	1.50%	10.61%
01-06-2022	-6.48%	1.21%	-0.42%	0.17%	0.87%	-0.70%	10.97%
01-07-2022	-4.95%	0.44%	0.60%	0.15%	0.75%	-2.48%	11.80%
01-08-2022	-2.08%	0.29%	0.72%	0.17%	0.70%	-1.68%	12.23%
01-09-2022	-2.13%	0.76%	0.30%	0.46%	0.85%	-1.90%	12.64%
01-10-2022	-4.35%	0.88%	1.00%	0.20%	0.56%	-1.12%	13.18%
01-11-2022	-3.41%	-0.19%	0.35%	0.29%	0.58%	1.28%	13.52%
01-12-2022	-3.53%	-0.65%	-0.64%	0.22%	0.50%	-0.25%	13.94%
01-01-2023	-4.88%	0.53%	0.82%	0.40%	0.44%	-0.66%	14.45%
01-02-2023	-2.56%	-0.24%	0.81%	0.64%	0.69%	-4.11%	15.12%
01-03-2023	-1.32%	0.01%	0.00%	0.02%	-0.01%	-0.04%	15.11%
01-04-2023	-1.33%	0.72%	0.98%	0.50%	0.35%	-2.36%	15.70%
01-05-2023	-2.70%	1.03%	0.23%	0.25%	0.25%	-2.76%	16.20%

Problem 5: Analysis & Insights

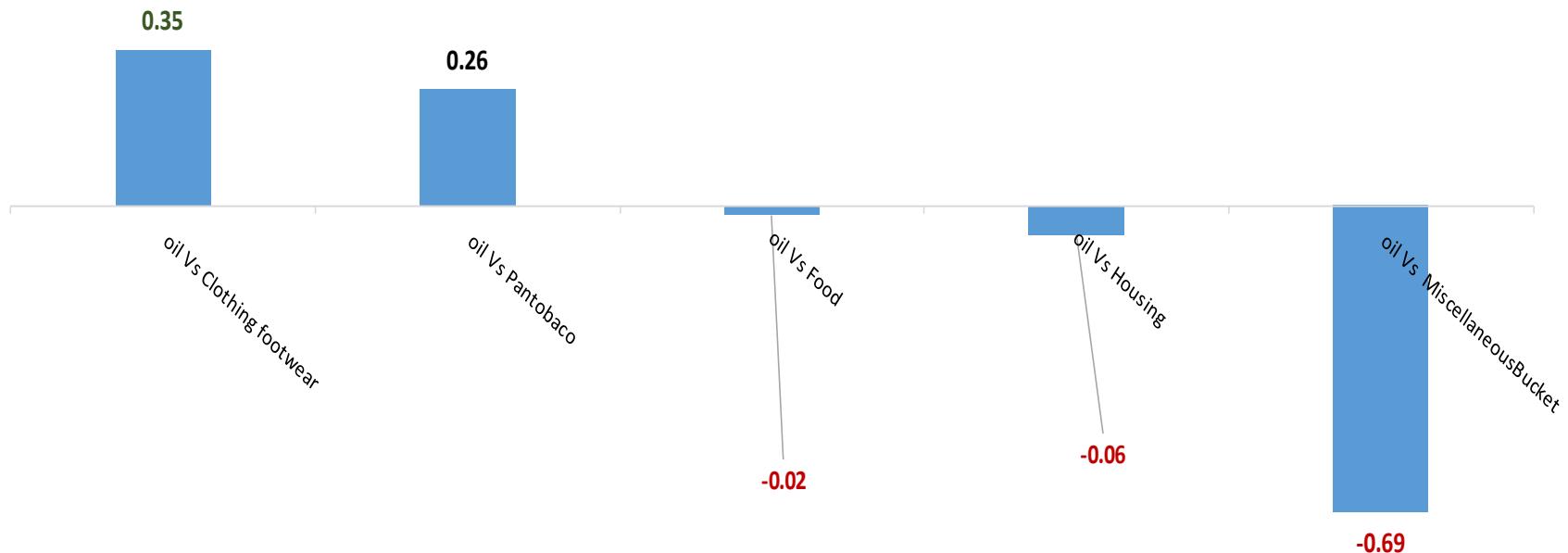
- A. Background – what was the issue?
 - ❖ Imported crude oil price data was missing from the main CPI dataset, preventing analysis of how oil fluctuations influenced domestic inflation categories. CPI data alone couldn't explain volatility in certain months. compared by time period.
- B. Insight from data – pivot table or chart
 - ❖ After correlation analysis (2021–23 period), results showed:
 - **Highest correlation** – *Clothing & Footwear* ($r = 0.35$) and *Pan/Tobacco* ($r = 0.26$)
 - **Weak/Negative correlation** – *Food* (-0.02) and *Housing* (-0.06)The **clustered column chart** clearly visualizes category-wise shifts.

Bucket Comparision	Sum of value of Corelations
oil Vs Clothing footwear	0.35
oil Vs Pantobaco	0.26
oil Vs Food	-0.02
oil Vs Housing	-0.06
oil Vs MiscellaneousBucket	-0.69

Problem 5: Analysis & Insights

- B. Insight from data – pivot table or chart
 - ❖ **Visualization** (bar chart) highlighted stronger linkage of oil to cost-intensive consumer goods rather than essentials.

Correlation of Imported Oil Price Fluctuations with CPI Buckets (2021–2023)



Problem 5: Analysis & Insights

- C. Solution / Action – formula, automation, BI logic

- ✓ **Created Helper Column:**

To classify each record as "Pre-COVID" or "Post-COVID":

=IF(YEAR([@Date])<2020,"Pre-COVID","Post-COVID")

(Or if using full date check → =IF([@Date]<DATE(2020,3,1),"Pre-COVID","Post-COVID"))

- ✓ **Used Pivot Table:**

Values: Average of each CPI bucket

Columns: COVID status (Pre/Post)

- ✓ **Chart Visualization:**

Clustered Column Chart → to show Before vs After comparison

Post-COVID bars in Blue, Pre-COVID in Orange

- ✓ **Automated Comparison Logic:**

Dynamic pivot refresh with slicer support, automatically updates when new CPI data is added.

- D. Result / Business Impact – show Before vs After table

- **Before:** No oil data linkage → CPI volatility seemed random and unexplained.

After: Integrated oil series revealed partial dependency — especially in *manufacturing-linked categories* (e.g., clothing, fuel-based products). Helped identify indirect inflation channels and informed energy-cost risk tracking in CPI analysis.

Problem 5: Analysis & Insights

Business Impact

Category	Correlation (r)	Sensitivity (per 10% oil increase)	Impact
Clothing & Footwear	0.35	3.5% CPI increase	Strongest linkage; manufacturing & logistics costs respond to oil.
Pan/Tobacco	0.26	2.6% CPI increase	Weak–moderate sensitivity.
Food	-0.02	Negligible	No oil dependence; driven by domestic supply.
Housing	-0.06	Negligible	No meaningful oil impact.

Calculation :

$$\text{Sensitivity} = \text{Correlation} * 10$$

Summary:

Oil price volatility primarily impacts **manufactured goods**, not essential items like Food or Housing.

Final Summary & Next Steps

Data Preparation Summary:

- All CPI datasets (2013–2023) cleaned — missing values, typos, and duplicate records resolved.
- Consistent category grouping applied (Food, Housing, Clothing, Miscellaneous, etc.).
- External **Crude Oil price data** integrated using VLOOKUP and helper column (*Cruid Bucket*).
- Derived **YoY % Difference** and **Category Contribution** metrics using Pivot Tables.
- Built automated visuals (Bar, Pie, Line charts) for trend and correlation analysis.

Business Results:

- Provided actionable insights on **major CPI drivers** — *Food* ($\approx 47\%$) and *Miscellaneous* ($\approx 30\%$).
- Identified periods of **inflation spikes** and **weak oil correlation zones**, guiding focus on domestic supply-side policies.
- Simplified over 30+ granular CPI items into 6 key analytical buckets for decision-making.
- Enhanced accuracy and consistency of CPI trend reporting.

Next Steps:

-  Automation: Currently data refresh is manual. In future, I plan to learn and implement Power Query in Excel to automate monthly CPI and oil data updates.
-  Dashboard Maintenance: Interactive CPI trend dashboard already built in Excel; next step is to enhance it with dynamic filters and visuals.
-  Scheduled Refresh: Plan to set up quarterly validation and refresh checks.
-  Currently correlation analysis is applied. Regression modelling can be learned and implemented later to predict CPI movement based on oil price trends.