

CS-732 Data Visualisation (A-1)

Keshav Chandak
IMT2021003
keshav.chandak@iiitb.ac.in

Sunny Kaushik
IMT2021007
sunny.kaushik@iiitb.ac.in

Samarjeet Wankhade
IMT2021013
samarjeet.wankhade@iiitb.ac.in

I. DATASET

The [dataset](#) encompassed Consumer Price Index (CPI) values for Indian states spanning from 2011 to 2018, with monthly data available for each state categorized into rural, urban, and combined (rural+urban) segments. Consumer Price Index (CPI) quantifies the average changes in prices that consumers pay for a standard basket of goods and services over a specific period, reflecting inflation or deflation in a country.

II. TASKS

We have undertaken three distinct visualization tasks to gain insights from the data-set:

T1: To Study CPI of different regions of India using line charts.

T2: Using cartographic Visualisation and compare trends across CPI variants.

T3: To analyse the given data-set using 2 Dimensional racing visualisation and draw generalised inferences from it.

III. VISUALISATION-1

A. Data Preprocessing

To effectively visualize the changes occurring over months and years in rural and urban sectors within each state, the data initially combined for both sectors was segmented into distinct rural and urban dataset. This division facilitated clearer and more precise visualizations for each sector's trends and variations.

B. Description

In this visualization, line charts are utilized to compare trends in rural, urban, and combined rural-urban Consumer Price Index (CPI) data. Additionally, it allows for a detailed analysis of CPIs across various regions of India, including the Northern, Southern, Eastern, Western, North Eastern, and Central regions.

- Northern Region: Himachal Pradesh, Punjab, Uttarakhand, Uttar Pradesh, Haryana, Jammu and Kashmir
- Southern Region: Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Telangana
- Eastern Region: Bihar, Orissa, Jharkhand, West Bengal
- Western Region: Daman and Diu, Dadar and Nagar Haveli, Maharashtra, Gujarat, Goa, Rajasthan
- North Eastern Region: Assam, Sikkim, Nagaland, Meghalaya, Manipur, Mizoram, Tripura, Arunachal Pradesh
- Central Region: Madhya Pradesh and Chhattisgarh

C. Visualisation - Rural

The most common observation in each of the graphs plotted was that the CPI for each state of India fell for the transition from the year 2014-2015 due to a number of factors :

- Falling global oil prices: The price of crude oil fell by more than 50 percent in 2014, which had a significant impact on the CPI in India. Oil is a major input cost for many industries, and its decline led to lower prices for goods and services across the board.
- Depreciating rupee: The Indian rupee depreciated by more than 10 percent against the US dollar in 2014, which also helped to lower the CPI. A weaker rupee makes imported goods more expensive, which can put downward pressure on prices.
- Good monsoon: The monsoon season was good in 2014, which helped to keep food prices in check. Food is a major component of the CPI, so a good monsoon can have a significant impact on the overall index.

The CPI started to increase again in 2016 due to the following factors:

- Rising global oil prices: The price of crude oil started to recover in 2016, which had a negative impact on the CPI.
- Weakening rupee: The Indian rupee continued to weaken against the US dollar in 2016, which also put upward pressure on prices.
- Adverse weather conditions: The monsoon season was not as good in 2016, which led to higher food prices.

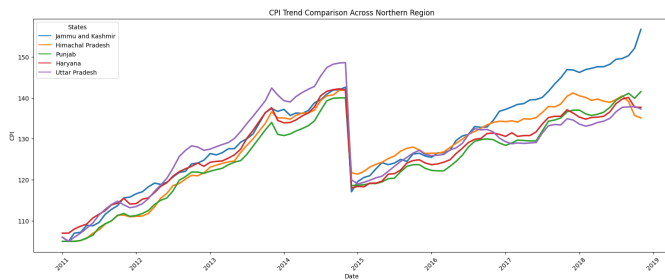


Fig. 1. CPI trend for rural part of Northern Region.

The above plot is comparing the trends of the variations in CPI in rural sector for the Northern states. There are two main highlights for this line chart:

- The CPI for the state of Uttar Pradesh showed higher values as compared the other states, which is because the population of the state has significant rural population and the other major factor is the investment of the government in the rural development.
- The plot for the state of Jammu and Kashmir shot up during 2017-2019 and this was because of the Kashmir floods which resulted in increase in the prices of the essential commodities.

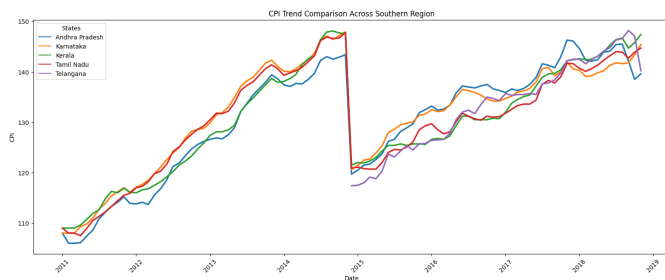


Fig. 2. CPI trend for rural part of Southern Region.

The main highlight of the Southern region was that the average CPI was comparatively greater as compared to the Northern region.

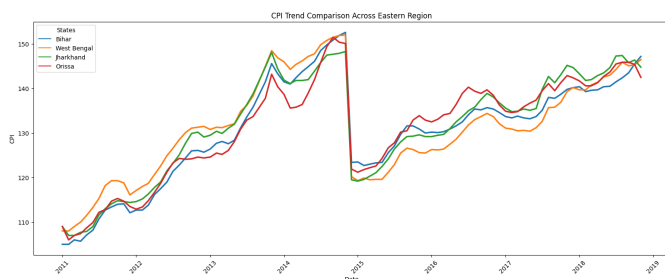


Fig. 3. CPI trend for rural part of Eastern Region.

The Eastern region was well balanced as per the plot with the average CPI higher than that of Southern region.

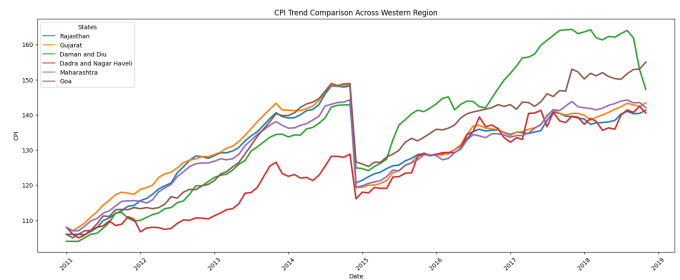


Fig. 4. CPI trend for rural part of Western Region.

The Western region plot has two major highlights:

- The plot for Dadra and Nagar Haveli didn't show a dip during the years 2014-2015.
- The plot for Daman and Diu during the years 2017-2019 showed an exceptional behaviour.

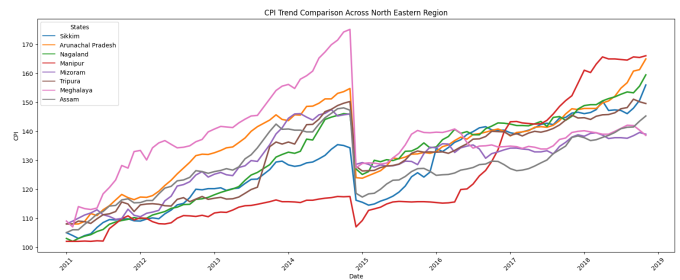


Fig. 5. CPI trend for rural part of North Eastern Region.

The main highlights for the North Eastern plot are:

- The plot for Meghalaya went to the peak among all the states during 2015.
- The plot for Manipur showed a sideways trend through 2011-2016.

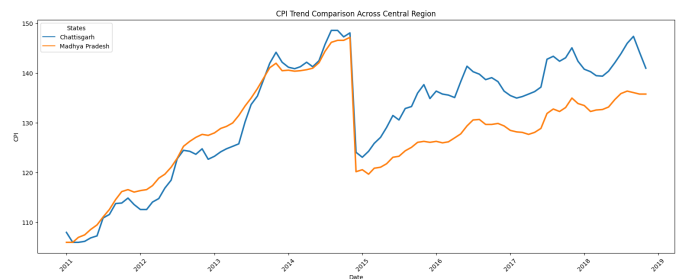


Fig. 6. CPI trend for rural part of Central Region.

The plot for the Central region had two states, Madhya Pradesh and Chhattisgarh.

D. Visualisation - Urban

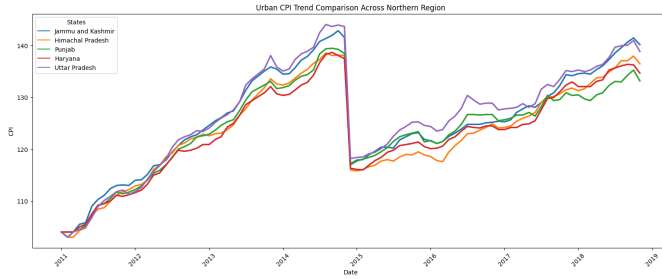


Fig. 7. CPI trend for urban part of Northern Region.

The CPI trend in the urban sector for the northern states was similar for all the states however, the CPI was lesser in the urban sector as compared to the rural sector.

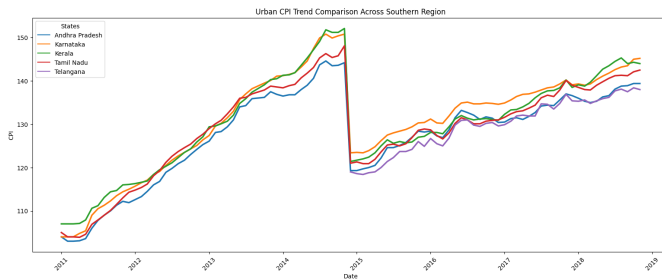


Fig. 8. CPI trend for urban part of Southern Region.

The CPI trend in the urban sector for the southern states was similar for all the states, however the CPI was lesser in the rural sector as compared to the urban sector. Is this because of the variation in the population of rural and urban regions in the two respective regions?

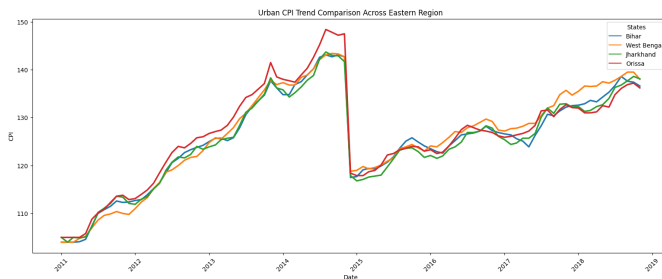


Fig. 9. CPI trend for urban part of Eastern Region.

The CPI trend in the eastern region was almost similar across all the states.

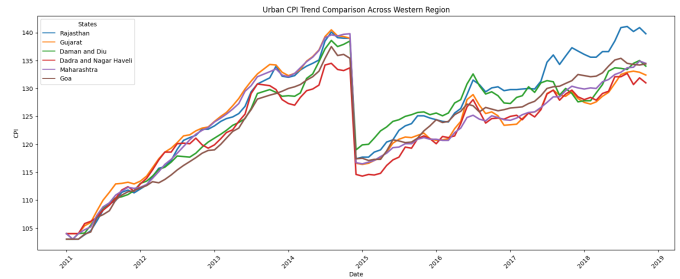


Fig. 10. CPI trend for urban part of Western Region.

The CPI trend of the Western region in urban sector was reached a lesser peak value as compared to the rural sector. Here, again the question arises, is it due to the population difference in the rural and urban sectors.

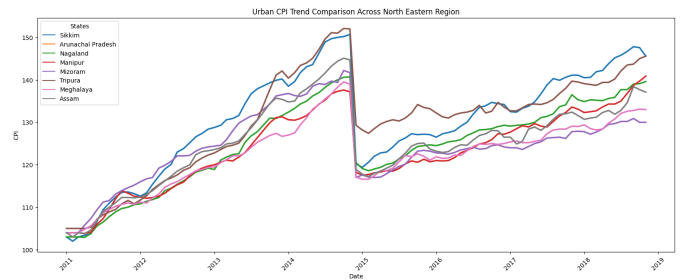


Fig. 11. CPI trend for urban part of North Eastern Region.

In the North Eastern region, the trend for a greater CPI in rural sector than urban sector was relevant.

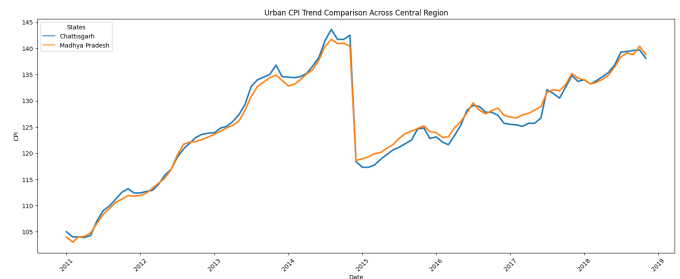


Fig. 12. CPI trend for urban part of Central Region.

The trend for the central region was very similar in the states of Madhya Pradesh and Chhattisgarh.

Now, as we have looked on the different regions of India statewise in both the rural and urban sector. Let's compare the collective regionwise mean of CPI in the two sectors.

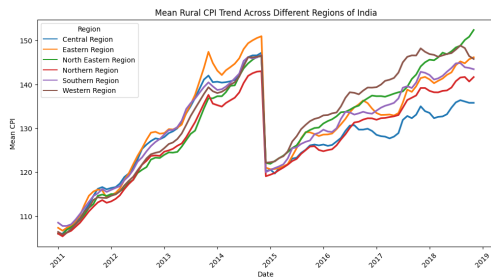


Fig. 13. CPI trend for different regions in the rural sector.

Till the year 2015 the eastern region dominated with the highest CPI, from 2015-2018 it was taken over by the states from the western region.

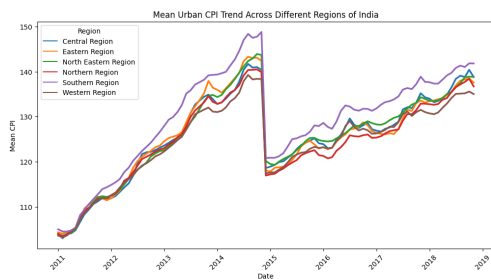


Fig. 14. CPI trend for different regions in the urban sector.

The highlight for this plot was the mean CPI for the Southern region which showed an exceptionally high values as compared to the other regions in the urban sector.

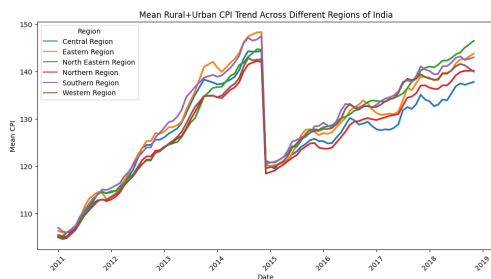


Fig. 15. CPI trend for different regions in the rural and urban sector.

The above plot shows the CPI trend for the different regions of India considering the both rural and urban sector.

E. Inferences

- In India, when taking into account all the states, the Consumer Price Index (CPI) values were higher in the rural sector compared to the urban sector. This trend can be attributed to the fact that approximately 65 percent of the Indian population resides in rural areas, according to the 2019 data.
- In the northern states of India, the Consumer Price Index (CPI) was higher in the rural sector, whereas in the southern region, the opposite trend was observed. This disparity can be attributed to the varying rural population proportions in these respective regions.
- The Southern region of India exhibited a similar trend in both sectors, with the highest CPI recorded in the urban sector across all regions.
- The data for the region of Telangana commences in 2015, as it was formerly a constituent of Andhra Pradesh (though it was formed in 2014 we had data from 2015).
- Meghalaya consistently maintained higher CPI values in the rural sector compared to all other states, reaching its peak in 2015.
- During the transition from 2014 to 2015, while the CPI in all states experienced a decline, Manipur stood out by maintaining its CPI at an all-time low without displaying a significant dip in 2015.

IV. VISUALISATION-2

A. Data Preprocessing

In this visualization, we are trying to use cartographic map to depict average CPI for each sector over the years. In order to put the information in cartographic map, we have preprocessed average for each state over the years categorically based on rural, urban and rural+urban sector. Also, in order to use tableau for visualization, we created average rural CPI, average urban CPI and average rural+urban CPI as new columns. This helps us to easily create the visualization by simple drag-and-drop.

B. Description

In this visualization, cartographic maps are utilized to compare trends in rural, urban, and combined rural-urban Consumer Price Index (CPI) data. These maps will visually represent the CPI for different regions, allowing viewers to see how the cost of living varies between rural and urban areas. By comparing the CPI values across regions, we can identify which areas have a higher or lower cost of living. The data given to us did not contain CPI for urban parts of Arunachal Pradesh. Also, the data for Ladakh is not available as it has recently been established as Union Territory of India, and the dataset given to us contains data from 2011-2018. Through this task, we will try to reason out the average CPI for various states based on various internal and external aspects of the state.

C. Visualisation - Rural

We can see from the visualization that Meghalaya and Arunachal Pradesh have the highest average rural CPI among the other states, while their neighbouring states Assam and Manipur have lower average rural CPI. This might be due to the fact that:

- **Absence of Urban CPI Data for Arunachal Pradesh:** The lack of urban CPI data for the state of Arunachal Pradesh can be indicative of the fact that it is a predominantly rural and hilly state with limited urbanization. The state's remote and challenging terrain makes it less attractive for urban development and investment.
- **Agricultural Reliance in Arunachal Pradesh and Meghalaya:** The economies of Arunachal Pradesh and Meghalaya are primarily agrarian, with a heavy reliance on agriculture and allied activities. Rural areas often face unique economic challenges, such as limited access to markets and infrastructure, which can contribute to higher inflation in essential goods and services.
- **Diversified Economies in Assam and Manipur:** Assam and Manipur may have more diversified economies with a mix of agriculture, industry, and services. This economic diversification can contribute to more stable and lower inflation rates compared to states heavily reliant on agriculture.

In the context of lower rural Consumer Price Index (CPI) in northern states like Punjab, Uttar Pradesh, Himachal Pradesh,

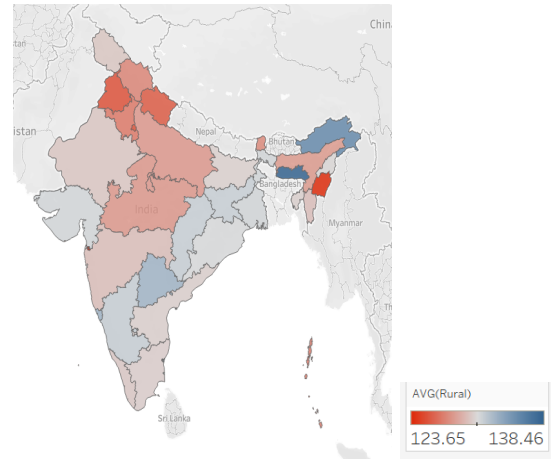


Fig. 16. I-Cartographic map showing average rural CPI for years 2011-2018, and II-showing legend for the same.

Uttarakhand, and Haryana for the years 2011 to 2018, several factors contribute to this phenomenon:

- **Agricultural Dominance and High Productivity:** These states have a strong agricultural sector with high productivity due to favorable soil, climate, and modern farming practices, resulting in a steady supply of affordable food.
- **Lower Population Density:** The lower population density in rural areas of these states reduces demand-driven inflationary pressures.
- **Government Support and Subsidies:** State governments provide significant support to farmers through subsidies and agricultural schemes, contributing to price stability.

D. Visualisation - Urban

In the period from 2011 to 2018, several factors contributed to the higher average urban Consumer Price Index (CPI) in southern states such as Kerala, Tamil Nadu, Karnataka, and Telangana. Here are the four key reasons for the elevated urban CPI during this period:

- **Urbanization and Population Density:** The significant urbanization witnessed in these southern states during this period led to higher population density in urban areas. This urban influx increased the demand for goods and services, contributing to price inflation.
- **Economic Growth and Employment Opportunities:** Robust economic growth and industrialization in these states attracted individuals from rural areas in search of employment opportunities. The resulting urban migration drove up demand for housing, food, and services, pushing prices higher.

During the period from 2011 to 2018, several factors played a significant role in contributing to the lower average urban Consumer Price Index (CPI) in states like Gujarat and Maharashtra. These factors reflect the states' economic diversity, infrastructure development, and effective governance. Here are the four key reasons for the reduced urban CPI during this period:

- **Industrial and Economic Diversity:** Gujarat and Maharashtra have diversified economies that encompass agriculture, industry, and services. The presence of multiple economic sectors contributes to price stability and lower inflation rates in urban areas.
- **Manufacturing and Industrial Base:** Both states have robust manufacturing and industrial bases, attracting a skilled workforce and increasing production and the availability of goods at competitive prices in urban markets.
- **Infrastructure Development:** Substantial investments in infrastructure development, including transportation networks, enhance supply chain efficiency, reduce transportation costs, and improve overall living conditions, contributing to a lower cost of living.
- **Access to Ports and Markets:** Proximity to major ports like Mundra Port(Gujarat), Jawaharlal Nehru Port(Maharashtra) facilitates the efficient movement of goods, ensuring a consistent supply and competitive pricing for urban consumers in these states.

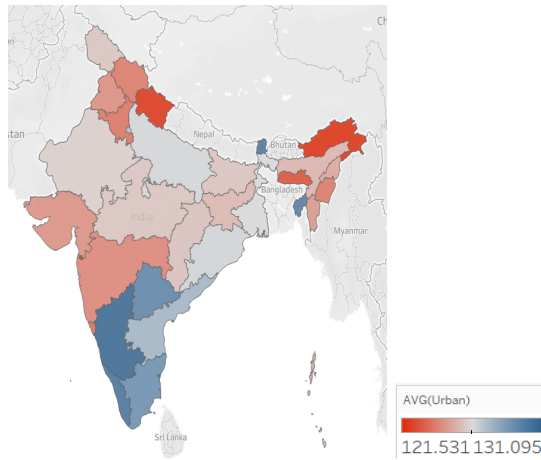


Fig. 17. I-Cardographic map showing average urban CPI for years 2011-2018, and II-showing legend for the same.

E. Visualisation - Rural+Urban

During the period from 2011 to 2018, certain regions, including Dadra and Nagar Haveli, Lakshadweep, and Manipur, exhibited a lower average rural+urban Consumer Price Index (CPI). Here are the key reasons for the reduced rural+urban CPI during this period:

- **Geographical Isolation:** Geographical isolation of these regions makes it challenging for goods to reach them. Limited accessibility can lead to lower demand and lower prices.
- **Government Subsidies and Support:** Government subsidies and support for essential commodities and services can help control price inflation, especially in regions where the cost of living is generally lower.
- **Lower Population Density:** The lower population density in these regions means less demand-driven inflationary

pressure, as there are fewer consumers competing for goods and services.

- **Influence of Tourism:** While tourism may contribute to the economy in places like Lakshadweep, it can also lead to controlled pricing in an effort to attract tourists and maintain affordability.

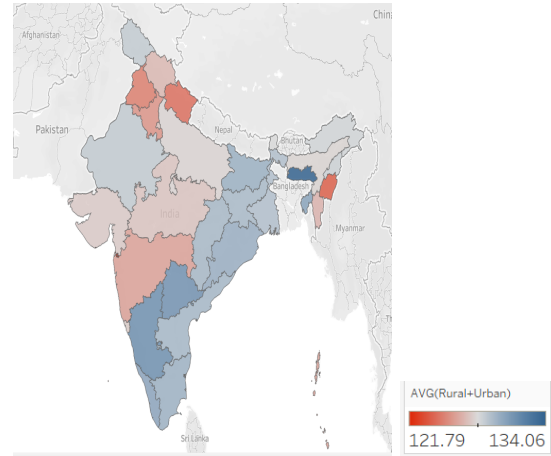


Fig. 18. I-Cardographic map showing average rural+urban CPI for years 2011-2018, and II-showing legend for the same.

F. Inferences

Based on the analysis of Consumer Price Index (CPI) data and related factors for different regions in India, the following inferences can be deduced:

- **Geographical isolation and limited urbanization** can lead to reduced accessibility and lower demand for goods and services, resulting in lower CPI levels. This is especially noticeable in geographically isolated regions like Lakshadweep. Also, if those places have support of tourism like Andaman and Nicobar Islands and Lakshadweep, their CPIs tend to be lower as they maintain affordability to attract more tourists.
- **Government subsidies and support for essential commodities and services** play a significant role in controlling price inflation. Areas where governments provide effective subsidies tend to have lower CPI levels.
- **Regions with lower population density** generally experience less demand-driven inflationary pressure, contributing to lower CPI levels. States with sparser populations, such as Uttarakhand and Assam, tend to have lower CPI, whereas states like Karnataka, Tamil Nadu and Kerala have high average CPI due to high population density. Also, the prime occupation of the majority of the population also impacts the CPIs.
- **The influence of tourism** can impact CPI, with tourism contributing to the economy but also affecting pricing strategies. In tourist destinations like parts of Maharashtra, controlled pricing is used to attract tourists and maintain affordability, leading to lower CPI levels.

V. VISUALISATION-3

A. Data Preprocessing

On top of the data processed in visualization-1, we have transformed the data in to create <State, Month Year, CPI-rural, CPI-urban, CPI-combined> tuple in order to create 2 dimensional racing Visualization. Note that in order to find generalised inferences we had to remove the data which was not available or partially available(i.e. Telangana and Arunachal Pradesh). This was all done using excel and it's built in functionality Power Query editor.

B. Description

In this visualization, we have used racing scatter plots and racing bar charts to analyse collective nature and rural and urban trends in 3 Dimensional Visualizations. Please refer to respective video provided with this document for the visualisations used in this Task.

C. Scatter plot Races: Rural, Urban, Rural and Urban-Combined

First of all, we will try to plot the three graphs for CPI-rural, CPI-urban, and CPI-combined on the y-axis and States on the x-axis, animated over time frames where each frame shows the respective value for a month of the said year.

- [Video reference for rural](#),
- [Video reference for urban](#),
- [Video reference for Rural and Urban-Combined](#)

Observations:

- Collective movement of all states: All states exhibit a tendency to follow a similar pattern or trend in their behaviour or actions.
- Dip in 2015 did not hamper this collective nature: Despite a decline or setback in 2015, the overall tendency of states to move together or share common behavior remains unaffected. The collective nature, in this case, remains strong and resilient despite a temporary downturn.
- These behavior is seen in all CPI variants.

Conclusions:

- Consistency in Behavior: The first observation indicates that there is a consistent and collective movement among all states in India. This suggests that states tend to act in a similar manner or follow similar trends in various situations or contexts. This could be due to shared cultural, economic, or political factors that influence their behavior.
- Resilience to Setbacks: The second observation highlights the resilience of this collective nature. Even when there was a dip or setback in 2015, the overall tendency of states to move together or exhibit common behavior remained strong. This implies short-term fluctuations may not significantly disrupt the overarching pattern of similarity in state behavior.

D. Bar Charts Race

In trends following from previous observations it can be noted that the CPI changes across state have a relation. This indicates the use of Racing bar charts to confirm the collective nature of CPI. X-axis contains Urban CPI, Y-axis has States and these all are spanned over time and sorted according CPI value from Highest to lowest.

- [Video reference for Rural](#)
- [Video reference for Urban](#)
- [Video reference for Rural and Urban-Combined](#)

Observations:

- Almost all states have similar CPI at a given time: Most states have comparable Consumer Price Index (CPI) values at a specific point in time, indicating a degree of uniformity or similarity in their economic conditions.
- CPI among various states seems to be a close competition: CPI values among different states are closely competitive or seem to have minimal variation, indicating a balanced economic environment.
- There also seems to be a restoring power (or force) whenever a state lags behind or leads in CPI: When a state falls behind or takes the lead in terms of CPI, there is a tendency for it to bounce back and regain its position. There is a restorative force at play, ensuring a certain equilibrium in CPI among states.

Conclusions:

- CPI Uniformity: The first observation suggests that, over time, most states maintain similar levels of economic development, reflected in their comparable Consumer Price Index (CPI) values. This hints a general economic stability or consistent factors influencing price levels across states.
- Competitive CPI Values: The second observation highlights that CPI values across states exhibit a narrow range, indicating a healthy competition among states to maintain their economic positions and manage CPI effectively.
- Restorative Mechanism: The third observation identifies a "restorative force" in state economic dynamics, ensuring states return to equilibrium if they fall behind or surge ahead in CPI. This suggests the presence of economic policies, market forces, or government interventions that mitigate extreme CPI disparities and promote stability.
- Collective CPI Trends: The first and second conclusion recognizes the collective nature of CPI movements across all states, attributed to interdependence through trade. This ensures that market principles like Supply and Demand, Surplus Trade, which influence CPI remain balanced.
- Restorative Force Through Trade: The third conclusion highlights how trade plays a role in the restorative force, as states with low CPIs export surplus products to others, helping stabilize prices. Government interventions, such as subsidies, also contribute to this equilibrium.

E. Scatter plot Race for Rural and Urban

Following from previous observations it is hinted that the CPI changes across state have a relation. This indicates the use of Racing scatter plot to see and confirm there collective nature. Y-axis contains Urban CPI, X-axis has Rural CPI and these all are spanned over time. This can be further inferred from trend lines plotted in this visualisation.

- [Video reference for Rural and Urban-Combined](#)
- [Trend Lines](#)

Observations:

- The CPI urban and rural has a Linear incremental behavior. Consumer Price Index (CPI) for both urban and rural areas shows a consistent and gradual increase over time, following a linear pattern.
- All states appear to have concentrated initially but as time progresses they appear to be more scattered. Initially, states had similar or closely clustered CPI values, but over time, these values have become more spread out or dispersed, indicating greater variability among states.
- Diverging trend lines: The trend Lines in this graph is diverging as well which indicates that over time the Some states are focusing more on Rural and Some States are focusing more on Urban populations.

Conclusions:

- The observed behavior of CPI in Indian states can be thought of as a combination of a general trend and noise. The general trend represents the gradual increase in CPI values over time, while the noise represents deviations from this trend. As time progresses, it appears that these noise deviations tend to increase. Which implies that the collective nature may not be an inherent thing. But, similar to how Maximum Likelihood Estimation (MLE) in artificial intelligence can be used to model and then predict the future values for now.
- Building on the previous point, this approach of considering a general trend and noise can potentially make it easier to predict the future behavior of states in India. By understanding the underlying trend and recognizing the noise patterns, policymakers and analysts can better anticipate how CPI might evolve in different states. This predictive capability can aid in making informed decisions related to inflation management and economic policies.
- It's important to note that despite the increasing variability and noise in CPI values among states, there is still a "restoring force" at play. This implies that while states may deviate from the general trend in the short term, they are likely to be influenced by economic factors or policies that bring them back towards the overall trend over time. In other words, no single state is likely to deviate from the general trend for an extended period, as economic forces tend to balance out regional disparities.
- Collective nature may not exist for Urban CPI and Rural CPI individually for all states in India. But from previous two visualisation it does exist for Rural + Urban CPI.

VI. REFERENCES

- 1) [Statista - Rural Population by State and Union Territory in India](#)
- 2) [MacroTrends - India Rural Population](#)
- 3) [The Indian Express - Rural Pinches More in High Inflation States](#)
- 4) [Quartz - Why Is Inflation Higher in Rural India Than in Urban?](#)
- 5) [Data.gov.in - State-level Consumer Price Index \(Rural/Urban\)](#)
- 6) [Urbanisation in Arunachal Pradesh](#)
- 7) [CEIC Data - Consumer Price Index Tamil Nadu Miscellaneous](#)
- 8) [The Indian Express - On Inflation: Gujarat Ahead of National Avg, Himachal Fared Best of All States in Sept](#)

VII. TOOLS USED

- 1) [Tableau](#)
- 2) [Python](#)
- 3) [Excel](#)
- 4) [Google Colab](#)
- 5) [Flourish](#)
- 6) [Latex](#)
- 7) [SQL](#)

VIII. AUTHOR CONTRIBUTIONS

- 1) Sunny Kaushik(IMT2021007): Contribution to T1 (CPI Trends via Line Charts), primarily focused on visualizing CPI trends across distinct regions of India using line charts. The author gathered and processed the CPI data, then designed and created the line charts to effectively represent these trends.
- 2) Keshav Chandak(IMT2021003): Contribution to T2 (Cartographic Visualization of CPI Trends), specialized in cartographic visualization techniques. Author's main contribution was to employ these methods to visually compare trends in CPI variants across different regions, providing a geographical perspective to the data.
- 3) Samarjeet Wankhade(IMT2021013): Contribution to T3 (2D Racing Visualization and Analysis), specialized in two-dimensional racing visualization, a unique approach for data analysis. The author creatively applied this technique to analyze the given dataset, drawing valuable insights and generalizations from the data through an innovative visualization style.