State-Level Petrol Price Behaviour in India: Trends and Forecasts

[1]Sayesha Poddar

[1]Modern School Vasant Vihar

Author Note

[1]Sayesha Poddar

Email ID: sayeshapoddar@gmail.com

Github Repository: https://github.com/sayeshapoddar/petrol-price-analysis

Abstract

This project studies how petrol prices changed across 15 major Indian states from 2017 to 2022. It uses monthly data from a public Kaggle dataset and turns it into a clean time series for each state. Post data preparation, the project explores long-term trends, monthly changes, and differences between states. It also compares the period before and after 2020 to see how the COVID-19 shock affected fuel prices.

The analysis showed that some states consistently had higher petrol prices than others, mainly because of differences in taxes and transport costs. Rajasthan, Telangana, and Madhya Pradesh stood out for both higher average prices and higher volatility. It then discusses simple forecasting models using rolling averages and Holt's linear trend method to predict prices for 2023. These forecasts showed slow and steady changes instead of sudden jumps.

Overall, this project aims to help understand how real economic data behaves, how to clean and analyse time series, and how simple forecasting methods can give useful insights about price movements.

1. Introduction

Petrol plays a major role in daily life in India. It affects how much households spend on travel, how businesses manage transport costs, and how states collect revenue through fuel taxes. Because petrol is an essential good, changes in its price can influence inflation and the overall cost of living. This makes fuel prices an important topic for both consumers and policymakers. Even small increases can add up quickly, especially for families that depend on daily commuting or for industries that rely on road transport.

One thing that stands out about India is that petrol prices are not the same across the country. Each state sets its own tax rates, mainly through Value Added Tax, which creates noticeable differences from place to place. Transport costs, distance from refineries, and local policies also play a role. As a result, two neighbouring states can have very different petrol prices even when global crude oil prices stay the same. Understanding these differences helps explain how state decisions and local conditions shape what people pay at the pump.

The period from 2017 to 2022 is especially useful to study because it includes both steady, normal years and a major global shock. Before 2020, petrol prices moved in response to typical market and policy factors. After the COVID-19 outbreak, the situation changed sharply. Global oil demand collapsed during lockdowns and then rose quickly as economies reopened. This created unusual price movements that affected all states, but not always in the same way.

Comparing the pre-COVID and post-COVID periods helps show how sensitive different states are to global and domestic changes.

This project was built around a few simple questions. How much do petrol prices differ across Indian states? Which states are consistently more expensive, and which ones are more stable? How much did prices change after 2020, and did all states see the same shift? Finally, can basic forecasting methods give a rough idea of what prices might look like in the near future?

By answering these questions, the project aims to show how real economic data can be analysed in a clear and practical way, while revealing patterns that connect state policies, market conditions, and price behaviour over time.

2. Data and Methods

2.1 Dataset Description

This project uses a publicly available dataset from Kaggle that contains monthly petrol prices for 15 major Indian states between 2017 and 2022. The dataset is structured in a wide format, where each column represents a state and each row represents a month. The states included are Delhi, Maharashtra, Tamil Nadu, Karnataka, Kerala, Rajasthan, Punjab, Haryana, Telangana, Assam, Gujarat, West Bengal, Bihar, Uttar Pradesh, and Madhya Pradesh. The time range covers May 2017 to December 2022, which gives 67 months of data in total.

To prepare the dataset for analysis, the first step was to convert the date values from strings (such as “2017_Aug”) into proper datetime objects. This allowed the data to be sorted chronologically and treated as a time series. After checking for missing values, a few gaps were found within certain state columns. Because petrol price movements tend to be gradual rather than random, simple interpolation was used to fill these missing values. This keeps the dataset complete without adding unrealistic jumps. After cleaning, the data was reshaped into both long and wide formats so it could be used easily for plotting and forecasting.

2.2 Exploratory Analysis

The exploratory analysis focused on understanding how petrol prices behaved across states and over time. It started by plotting each state’s price trend from 2017 to 2022. These line charts helped show which states were consistently high-priced, which ones were lower, and how prices moved during different periods. To smooth out short-term fluctuations and highlight general direction, rolling averages were added, especially the 3-month and 6-month moving averages. These made long-term trends easier to see.

To understand differences across states, the average price and volatility (standard deviation) was calculated for each state. This showed that states like Madhya Pradesh, Telangana, and Rajasthan had both higher average prices and larger swings over time. In contrast, Delhi and Gujarat remained relatively stable and lower-priced. These differences connect directly to state-level tax choices and transport-related costs.

2.3 Forecasting Approach

The forecasting part of the project used simple, interpretable models to predict petrol prices for 2023.

The first method was a **naive baseline**, where the forecast simply repeats the last observed value. This provides a starting point to compare other models against. The second method used **rolling averages**, especially the 6-month average. Rolling averages smooth the data and reflect the idea that recent trends are more important than older values. This often gives more reasonable short-term forecasts than the naive method.

The third and main method was **Holt's linear trend model**, which is an extension of simple exponential smoothing. Holt's method was chosen because it captures both the level and the direction of the data without being too complex. It allows the model to follow gradual upward or downward trends instead of staying flat. This is helpful for petrol prices, which rarely stay constant. Holt's method produced believable 12-month forecasts that showed slow upward movement or stabilization depending on the state's recent trend.

Together, these methods offered a clear and manageable way to understand how future prices might behave while keeping the analysis simple and realistic.

3. Key Findings

3.1 Long-run price ranking

Looking at the full 2017–2022 period, clear and consistent patterns appeared across states. Madhya Pradesh, Telangana, and Rajasthan stayed at the higher end of the price range almost the entire time. These states also showed higher averages in both the raw data and in the rolling mean calculations. On the other hand, Delhi and Gujarat remained among the lowest, with smoother and more stable price paths. These long-run gaps suggest that state-level tax policies and location-based transport costs play a major role in shaping petrol prices.

3.2 Volatility

The analysis showed that some states experienced stronger month-to-month changes than others. Rajasthan and Telangana had the largest fluctuations, which was visible both in their standard deviation values and in their line charts. These states reacted more sharply to national or global price movements. In contrast, Delhi remained comparatively stable, likely because of more predictable tax rates and its position as a major distribution point.

3.3 The 2020 break

A clear and noticeable shift occurred around March 2020. Prices for all states showed a jump or change in pattern soon after the COVID-19 outbreak. The early months of the pandemic brought unusual drops in global crude prices, followed by a fast recovery as economies reopened. This resulted in a sharp rise in petrol prices across India. The “pre-2020” and “post-2020” comparison showed that every state moved to a higher price level after this point, even though the size of the jump differed by state.

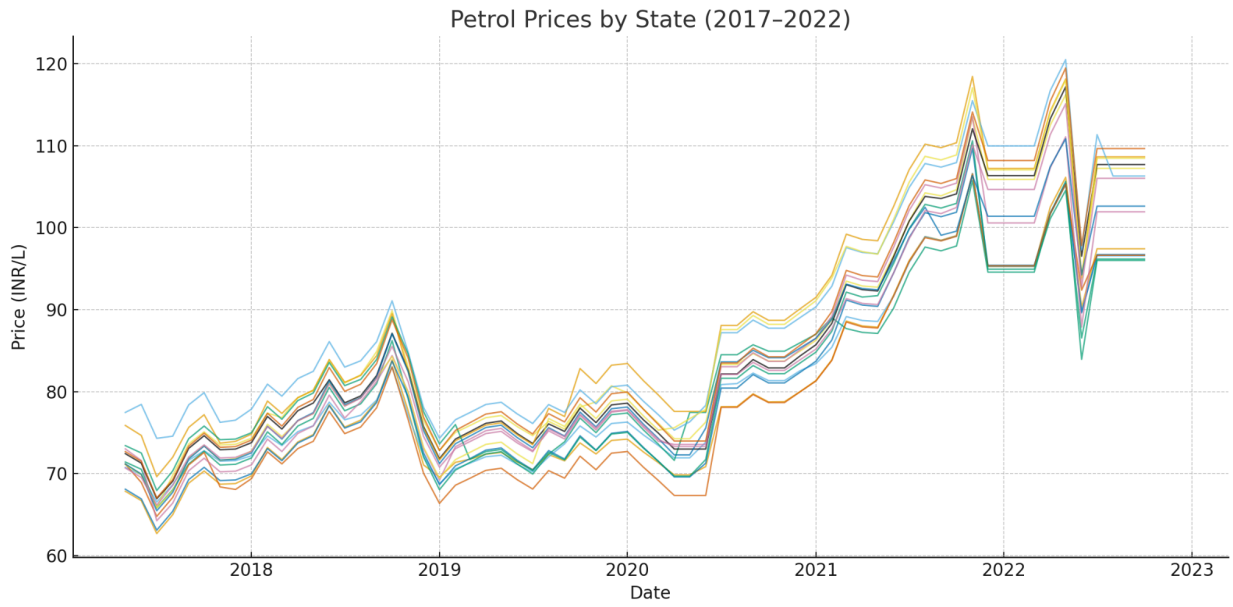
3.4 Forecast results

The simple forecasting models, especially Holt’s linear trend method, suggested that petrol prices in 2023 would likely show moderate upward movement or remain stable rather than experience large spikes. High-cost states are expected to stay high, keeping the same overall ranking seen in earlier years. No extreme jumps or sudden drops were predicted by any of the models. This lines up with the recent behaviour of the series, where prices rise slowly over time and react more to long-term factors than short-term shocks.

4. Visual Summary

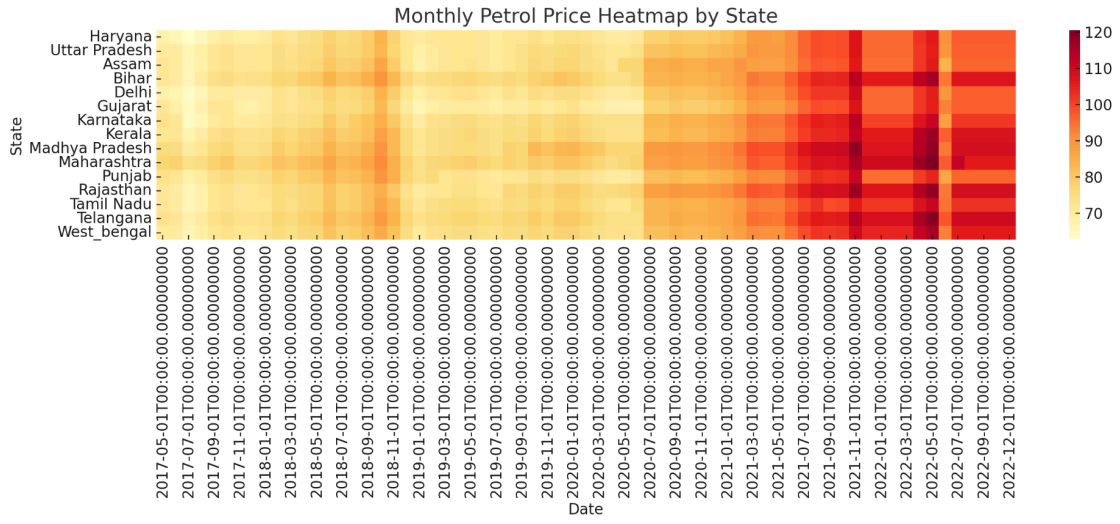
The figures below highlight the main patterns seen in the data and support the findings discussed earlier.

4.1State-wise line plot (2017–2022)



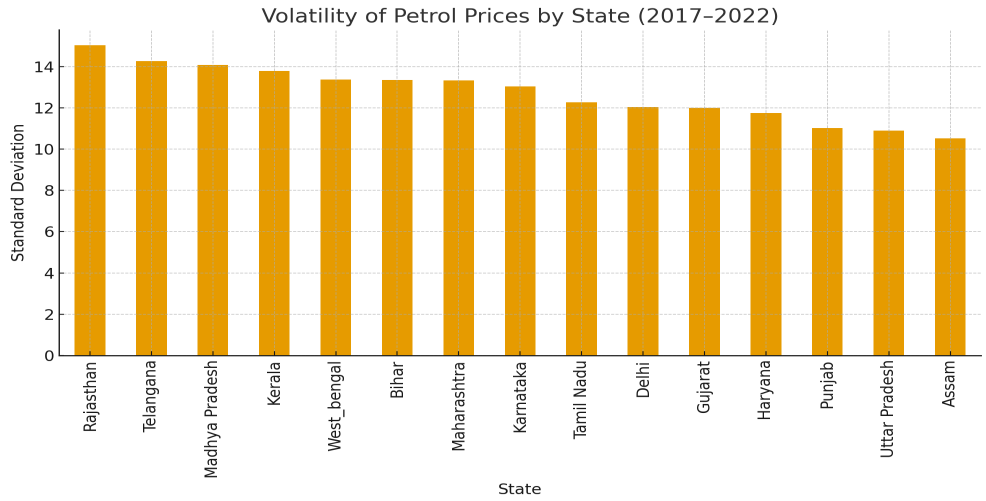
This chart shows how petrol prices moved in each state from 2017 to 2022. The lines make it easy to see long-run differences between states, as well as the strong upward movement after 2020. States such as Madhya Pradesh and Telangana stay near the top of the range, while Delhi and Gujarat form the lower end.

4.2 Monthly price heatmap



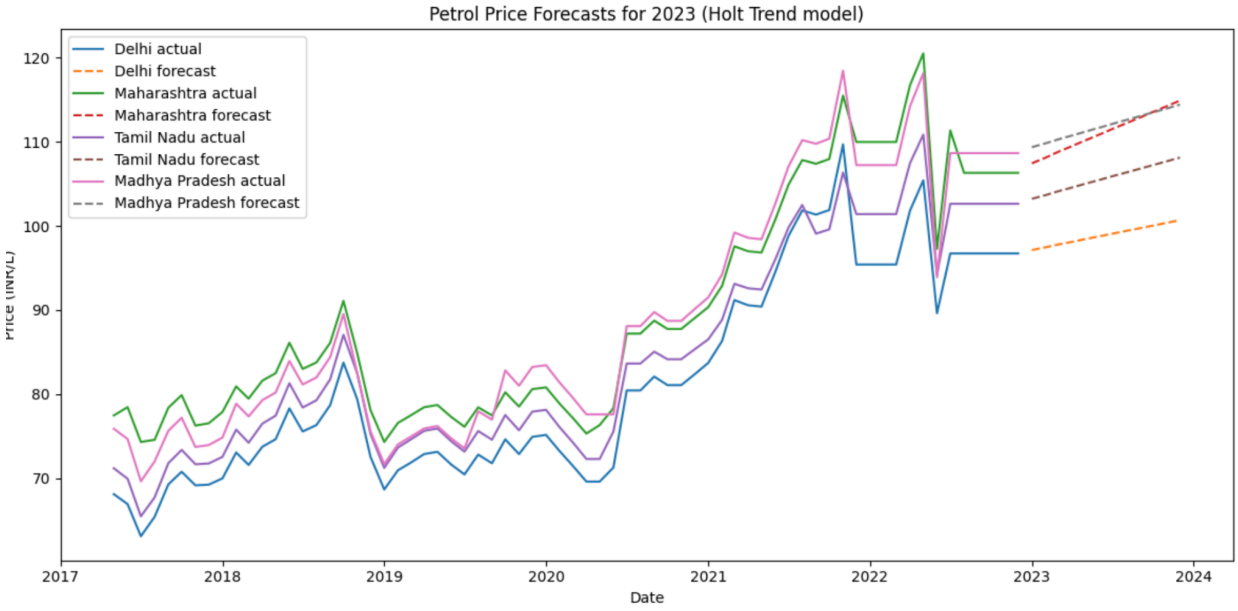
The heatmap gives a compact view of how prices changed month by month across all states. Darker shades appear more often in later years, reflecting rising prices. Certain states consistently show darker bands, matching their higher long-run averages.

4.3 Volatility ranking



A bar chart or ranking table helps compare how much prices fluctuated in each state. Rajasthan and Telangana stand out with the highest volatility values, while Delhi and Gujarat remain more stable. This supports the idea that some states adjust more sharply to market or tax changes.

4.4 Forecast plot (actual vs forecast for selected states)



This chart shows the real data up to 2022 and the predicted values for 2023 using Holt's linear trend method. The forecasts follow the direction of each state's recent pattern without large jumps. The trend lines remain smooth and realistic, showing either gentle increases or steady behaviour depending on the state.

5. Discussion

The results of this project show that petrol prices in India are shaped by a mix of national market forces and state-level policy choices. The reason some states remain more expensive than others becomes clear when looking at how fuel taxes work. States set their own Value Added Tax rates on petrol, and these taxes make up a large share of the final price. States like Rajasthan, Madhya Pradesh, and Telangana have higher VAT levels, which leads to consistently higher petrol prices over the entire period. Transport costs also matter. States farther from coastal refineries or major distribution points tend to face additional logistical expenses, which can push prices up even when global crude oil prices stay stable.

Volatility differences also point to how state policies interact with broader market movements. Rajasthan and Telangana show the highest month-to-month fluctuations. These states adjust their tax components more frequently or respond more directly to changes in global crude prices. Delhi shows the opposite pattern. As a national hub with stable tax policies and strong supply chains, its prices change more slowly and predictably. This makes Delhi a useful reference point for comparing how other states behave.

The forecasting part of the project helped connect recent trends to short-term expectations. Simple models, especially Holt's linear trend method, gave a realistic sense of how prices might move in the following year. These models do not aim to predict sudden shocks, but they are helpful for understanding the general direction of prices based on recent patterns. In this case, the forecasts suggested modest increases or stable movement rather than any extreme jumps. This matches the smoother trend seen toward the end of 2022, when petrol prices levelled off after earlier volatility.

Overall, the discussion shows that petrol pricing is not random. It reflects predictable economic structure, policy choices, and the way different states respond to market signals. Simple forecasting adds another layer of understanding by showing how past patterns shape short-term expectations.

6. Conclusion

This project gave me a practical understanding of how real economic data behaves and how it can be analysed using clear and manageable methods. Working with monthly petrol prices across 15 Indian states showed me how important it is to clean and structure data properly before looking for patterns. Converting the dataset into a long time series format, filling missing values, and adding rolling averages all helped make the analysis more reliable. This project also showed how different visual tools, such as line plots and heatmaps, can highlight trends that would be difficult to notice by looking at raw numbers alone.

Beyond the technical steps, the findings helped explain the economic forces behind fuel pricing in India. The long-run differences between states reflected VAT structures, transport costs, and local market conditions. The contrast between high-volatility states and more stable ones showed how policy choices shape how quickly prices respond to shocks. The clear shift after 2020 reinforced how global events, like the COVID-19 pandemic, can influence domestic prices even when state-level policies remain the same.

Using simple forecasting methods was another important part of the project. Instead of relying on complex models, rolling averages and Holt's linear trend was used to produce short-term predictions for 2023. These methods were easy to interpret and still provided meaningful insights. The forecasts pointed toward gradual increases or stability rather than drastic changes, which matched the recent pattern of petrol prices in the dataset. This experience showed me how forecasting can help form expectations about basic economic variables without requiring very advanced tools.

There are several directions this project could grow in. Adding diesel prices or LPG data would allow for comparisons across different fuels. Studying price elasticity of demand could show how sensitive consumers are to fuel costs. Finally, connecting the data with state-level tax changes or policy announcements could help explain why certain shifts happened at specific times. Overall, this project built a strong foundation for exploring more detailed economic questions in the future.