

GST Sales Data Analysis Report

Objective

The objective of this analysis is to examine the GST Sales Data across different states, identify key trends, and generate insights related to revenue contributions, industry performance, and monthly sales patterns. The assignment tasks include data cleaning, exploratory data analysis (EDA), visualization, and reporting.

1. Understanding the Dataset

The dataset consists of the following key variables:

- **State:** The state where the transaction occurred.
 - **Sales Amount:** The total sales recorded.
 - **Tax Amount:** The GST collected.
 - **Date of Transaction:** The date when the transaction was recorded.
-

2. Data Cleaning & Preprocessing

To ensure accurate analysis, the following data preprocessing steps were performed:

- **Handled Missing Values:** Removed or imputed missing sales and tax amount values.
 - **Standardized Date Formats:** Converted all dates into a uniform format (YYYY-MM-DD).
 - **Removed Duplicates:** Eliminated redundant entries.
 - **Standardized State Names:** Applied title casing and removed extra characters.
-

3. Exploratory Data Analysis (EDA)

3.1 State-wise GST Revenue Contribution

- **Top States by Sales Contribution:**
 - **Haryana:** ₹59.49M (16.44%)
 - **Karnataka:** ₹45.65M (12.62%)

- **Maharashtra:** ₹38.92M (10.75%)
 - **Madhya Pradesh & West Bengal:** Significant contributors.
- **Lowest Contributors:**
 - **Andaman & Nicobar Islands, Ladakh, and Other Countries** contributed less than 0.001% each.

3.2 Monthly GST Sales Trends

- **Sales Trends Over Time:**
 - Gradual increase from **August 2024 to October 2024**.
 - **November 2024 saw an extraordinary spike (₹359.83M)**, likely due to seasonal trends, business events, or product launches.

3.3 Industry-wise GST Contributions

- The highest GST-contributing industries include HSN/SAC codes:
 - **8012320.0, 8021200.0, 8041020.0, 8129090.0, and 20081920.0.**
- The lowest contributors include **10064000.0, 11061000.0, and 39199020.0.**

3.4 Outlier Detection & Removal

- **IQR & Z-Score Methods:** Identified and removed extreme outliers in the **Total Document Value (OP)**.
- **Impact of Outliers:** The cleaned data exhibited more stable and realistic sales trends.

3.5 Refund & Negative Transactions Analysis

- Certain customers (GSTINs) had significantly high refunds, potentially indicating:
 - **High return rates.**
 - **Fraudulent refund claims.**
 - **Operational errors in reporting.**

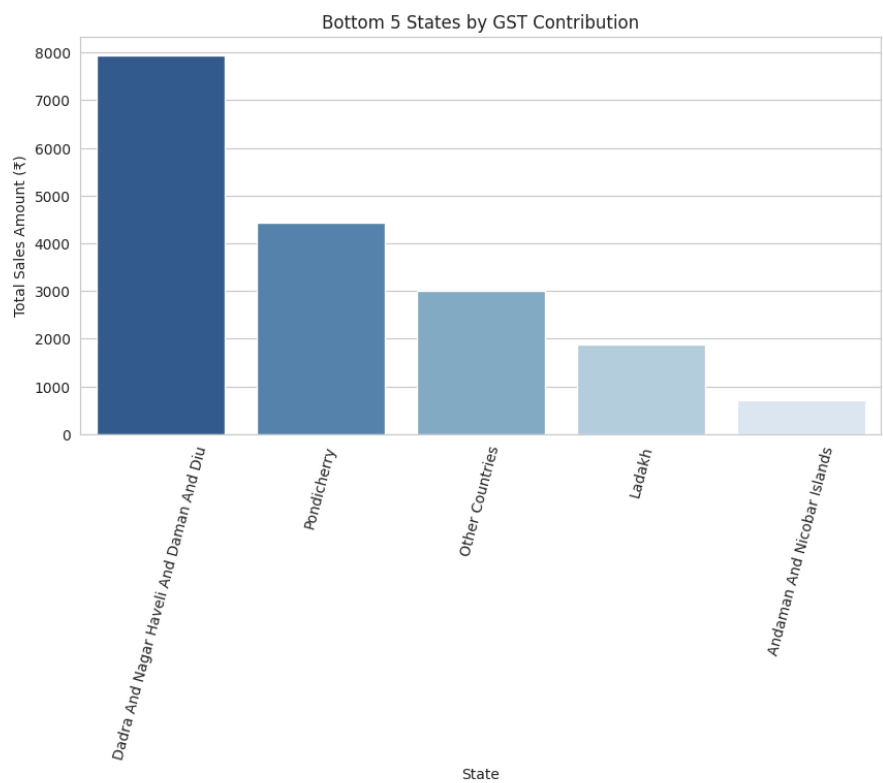
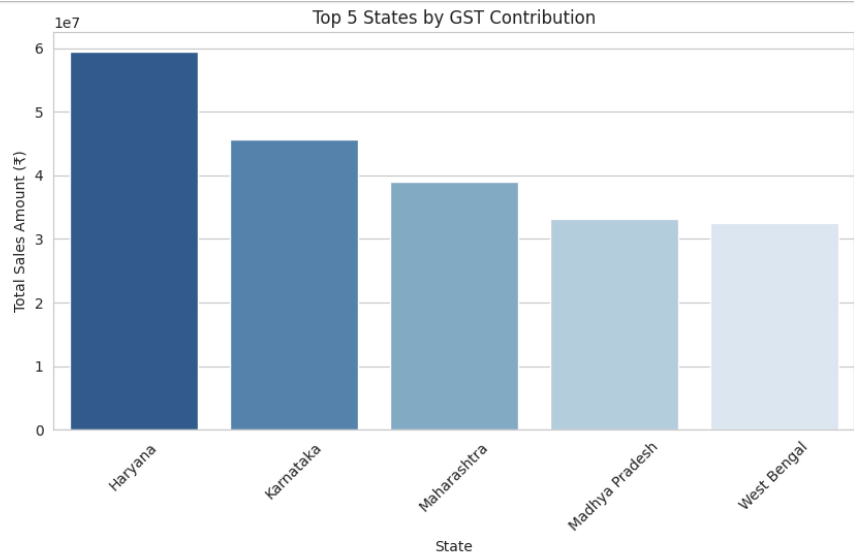
3.6 High-Value Transactions & Customer Impact

- A few GSTINs (e.g., **06AAACZ8867B1Z4, 06AABCF8078M1Z5**) contributed disproportionately to high-value transactions.
- Requires further scrutiny for potential fraud or high-volume business activity.

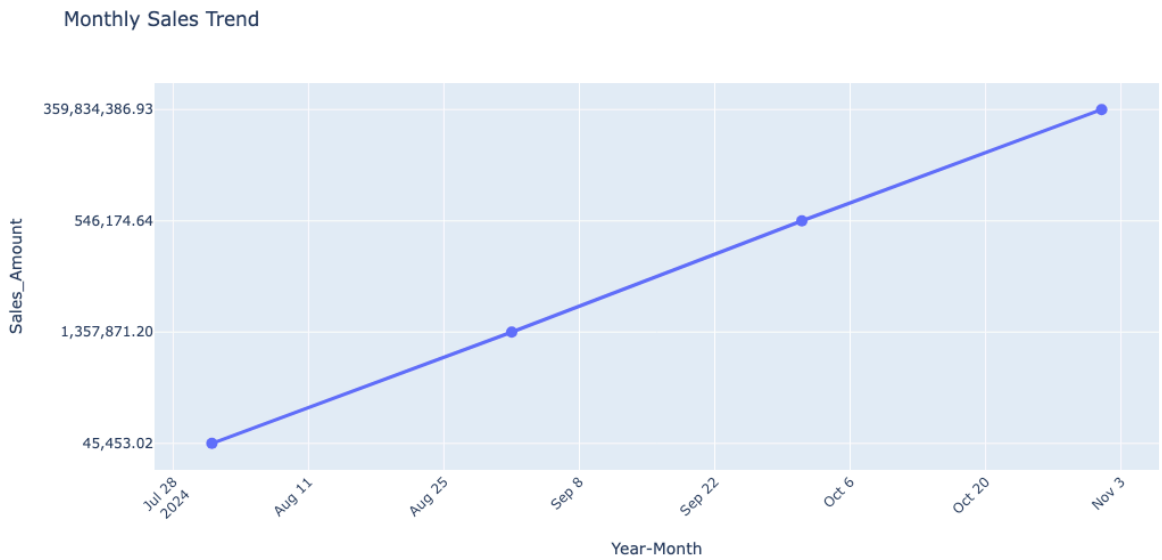
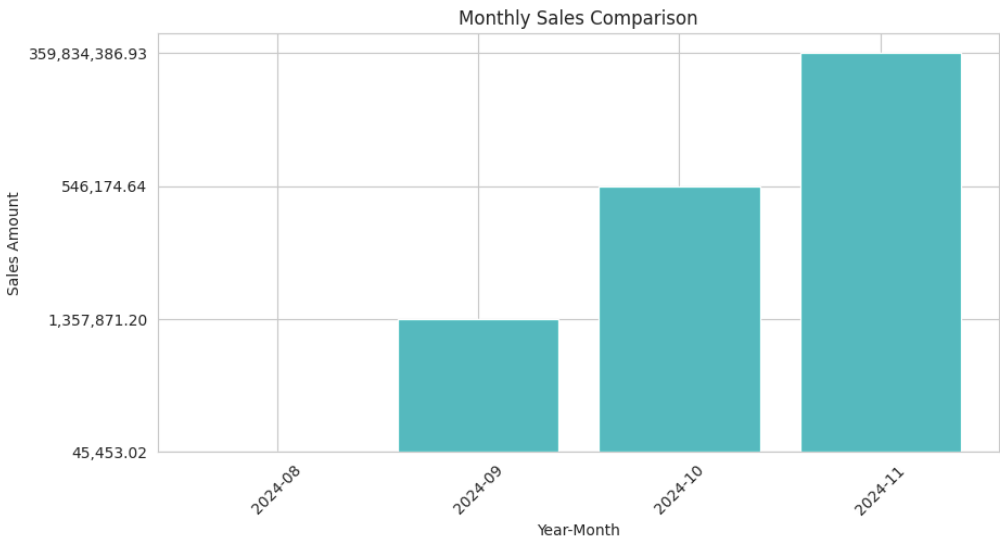
4. Visualization & Reporting

Key Visualizations:

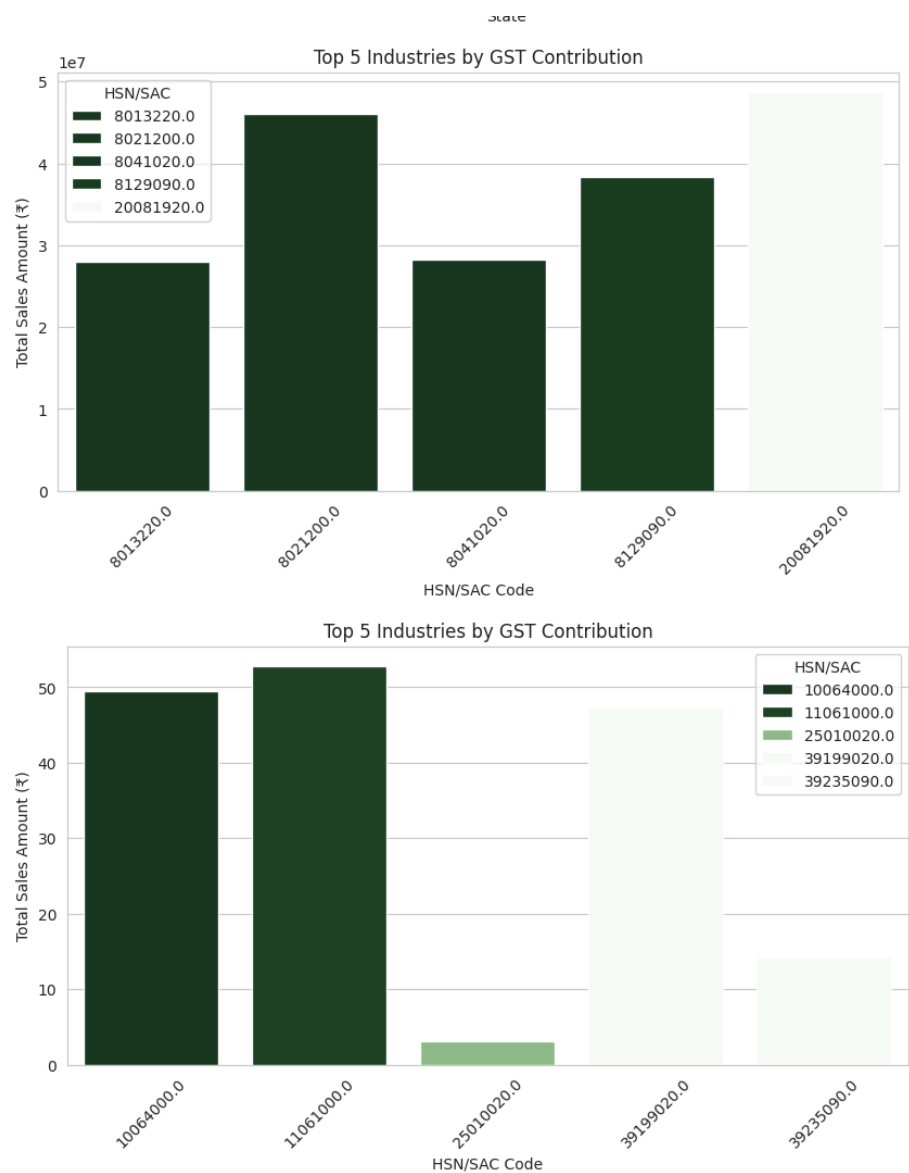
- **State-wise Sales Contribution** (Bar Chart with a red dashed line for the average contribution).



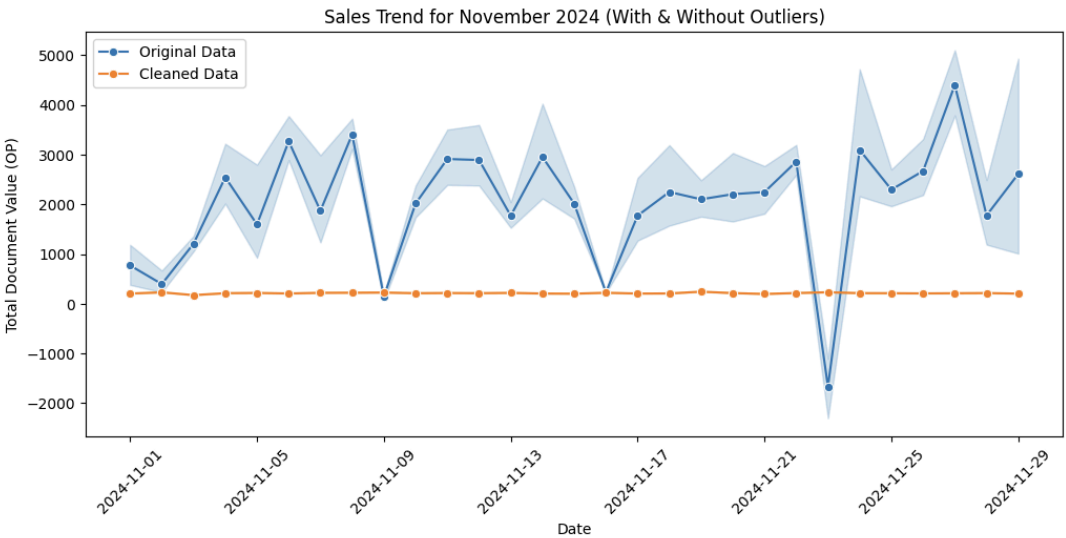
- **Monthly Sales Trends** (Line graph showing a sharp increase in November 2024).



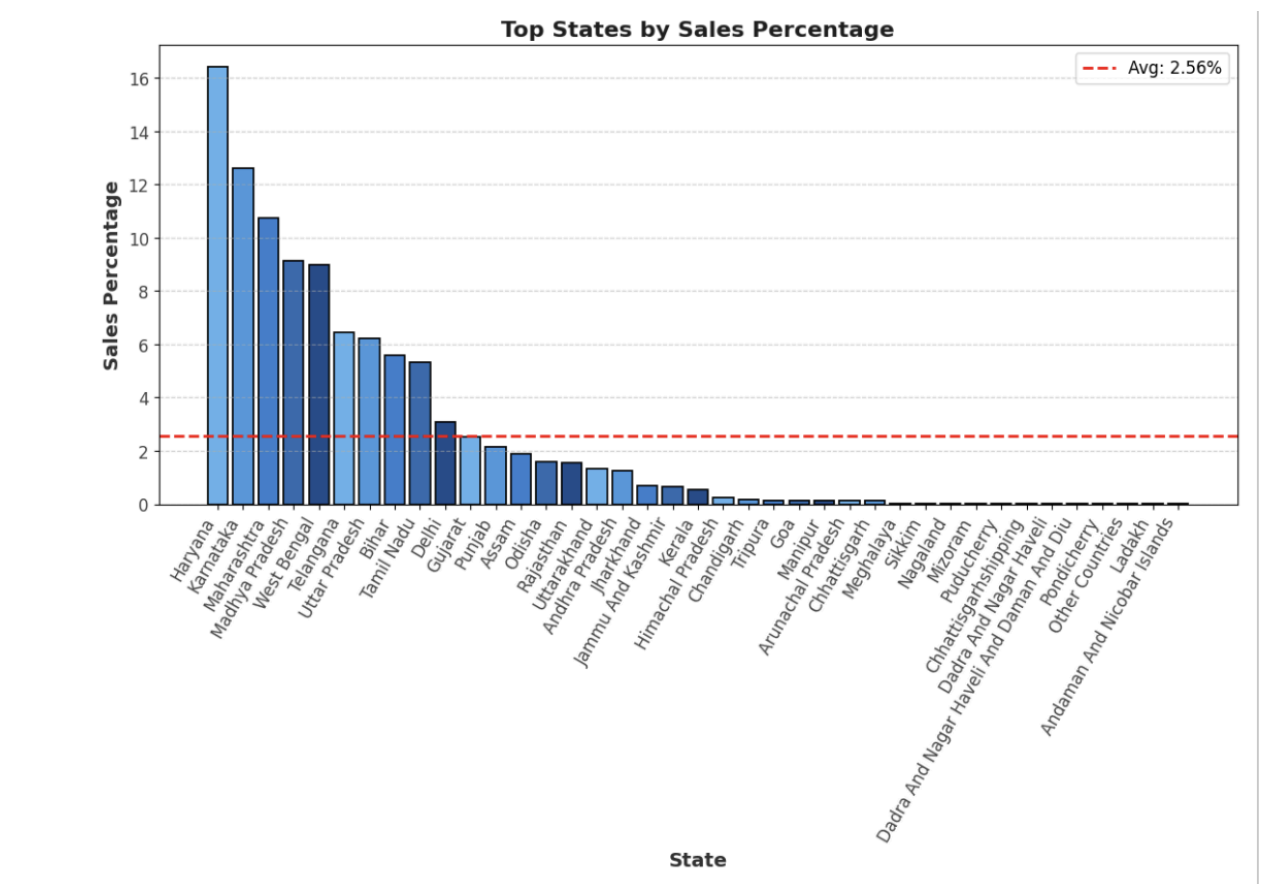
- **Industry-wise GST Contribution** (Bar Chart comparing different HSN/SAC codes).



- **Outlier Comparison (Boxplot)** (Before and after cleaning).



- **Top & Bottom States by GST Contribution** (Bar Charts).



5. Business Implications & Recommendations

- **Focus on Key Revenue States:** Marketing and expansion should target **Haryana, Karnataka, and Maharashtra**.
- **Prepare for Seasonal Sales Spikes:** The **November trend suggests a peak sales season**—businesses should optimize promotions accordingly.
- **Monitor High-Value Transactions & Refunds:** High refunds and large transactions from select customers require closer scrutiny.
- **Explore Market Potential in Low-Contributing States:** Smaller states (e.g., **Tripura, Arunachal Pradesh, Mizoram**) may offer **growth opportunities** with targeted marketing.

Conclusion

This analysis provided an in-depth look into the GST sales trends, identifying top revenue states, key industries, and monthly patterns. The results can be used to drive business decisions, optimize tax planning, and target new market opportunities. Further deep dives into customer behavior and transaction patterns will enhance strategic planning.