Medical Data Analysis Report

Generated on: 2025-01-10 06:29:01

Analyzed File: indian_medicine_data.csv

Executive Summary

The provided medical dataset comprises 253973 records with 9 features, including 'id', 'name', 'price(■)', 'Is_discontinued', 'manufacturer_name', 'type', 'pack_size_label', 'short_composition1', and 'short_composition2'. Essential statistics indicate that while 'id' has a mean of 126987, 'price(■)' has a significantly higher mean of 270.53, suggesting a skewed distribution. Missing values are present in 'short_composition2' for 141802 records, which should be addressed for comprehensive analysis.

Dataset Overview

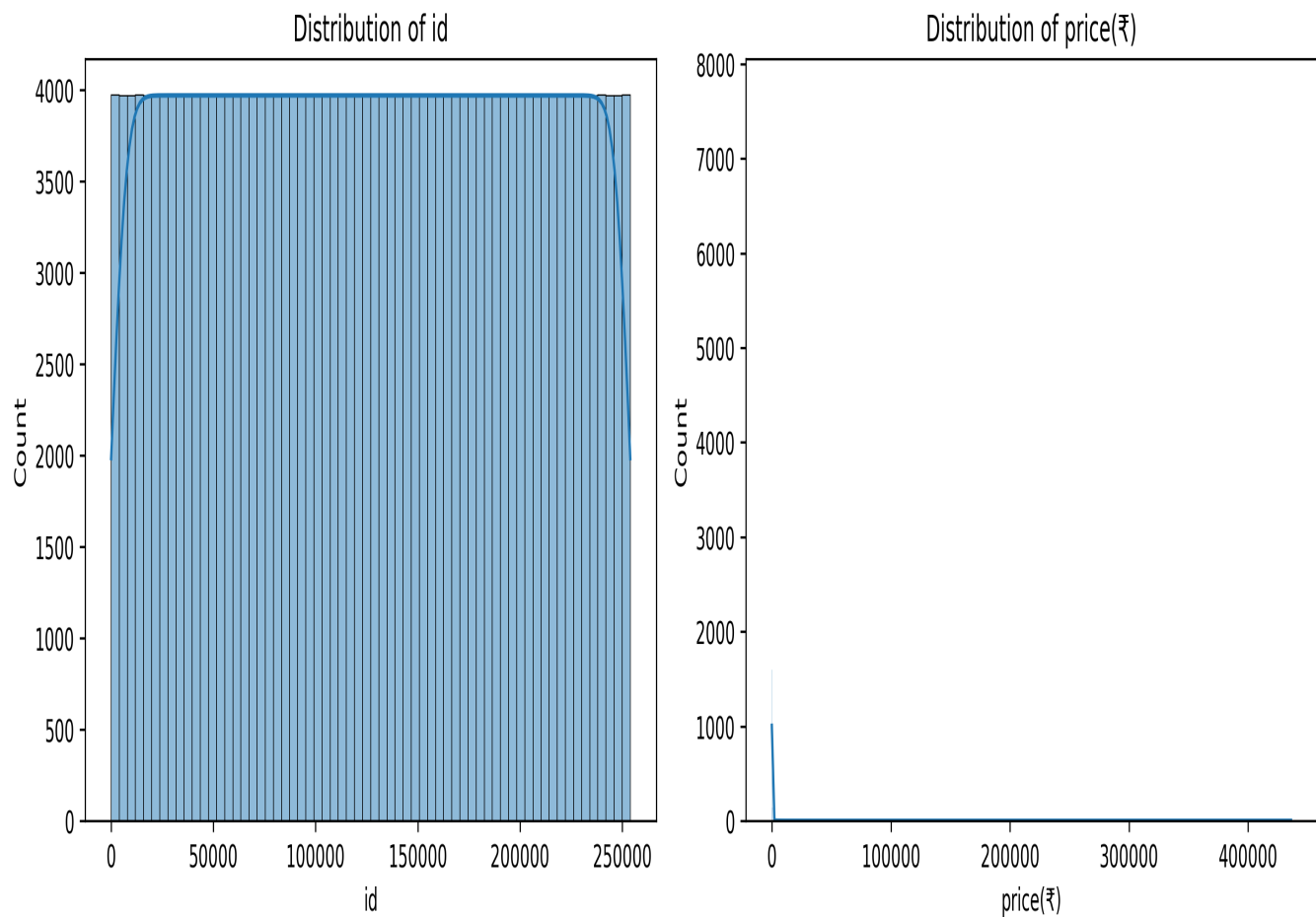| Total Records	253973
Total Features	9
Numeric Features	2
Categorical Features	6
Missing Values	141802

Key Insights

• The average price of medications in the dataset is ■270.53, but there is substantial variation, with a maximum price of ■436000.

• The most common type of medication is 'Injection', followed by 'Tablet' and 'Syrup'.

• The top manufacturers include 'Cipla', 'Sun Pharmaceutical Industries', and 'Abbott Healthcare Pvt. Ltd.'

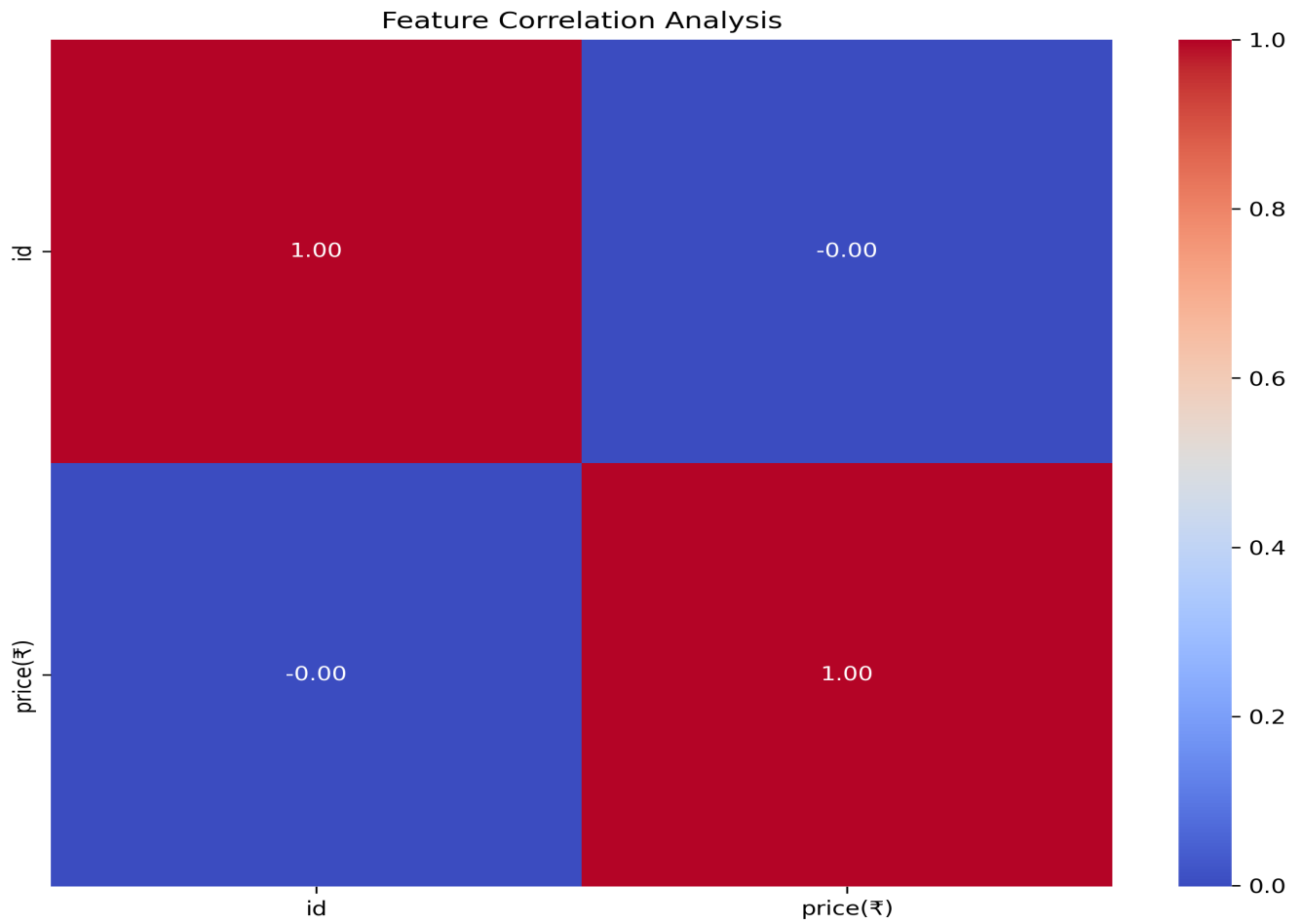• The most common pack size label is '10', followed by '5' and '20'.

• The most frequent short composition is 'Paracetamol' for 'short_composition1' and 'Ibuprofen' for 'short_composition2', highlighting the prevalence of pain relievers in the dataset.
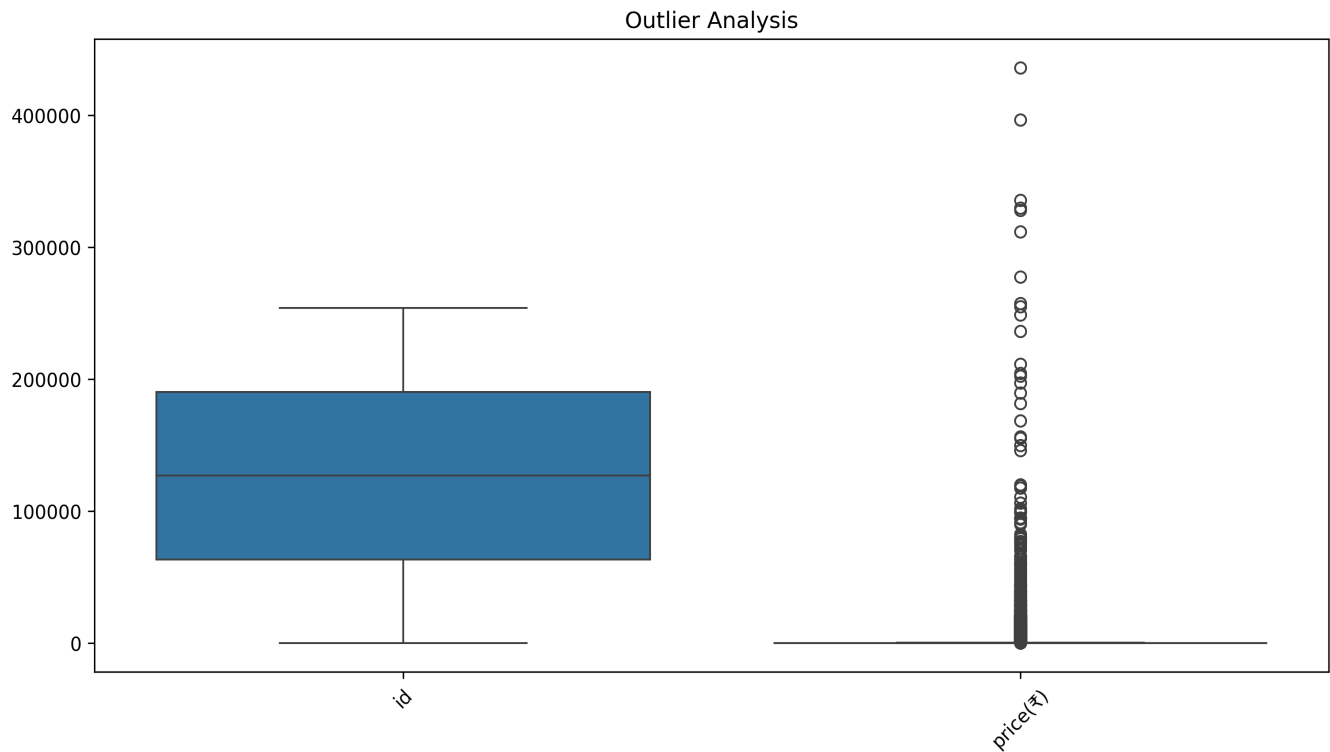
Data Visualizations

Feature Distributions

Correlation Analysis

Feature Correlation Analysis

Outlier Analysis
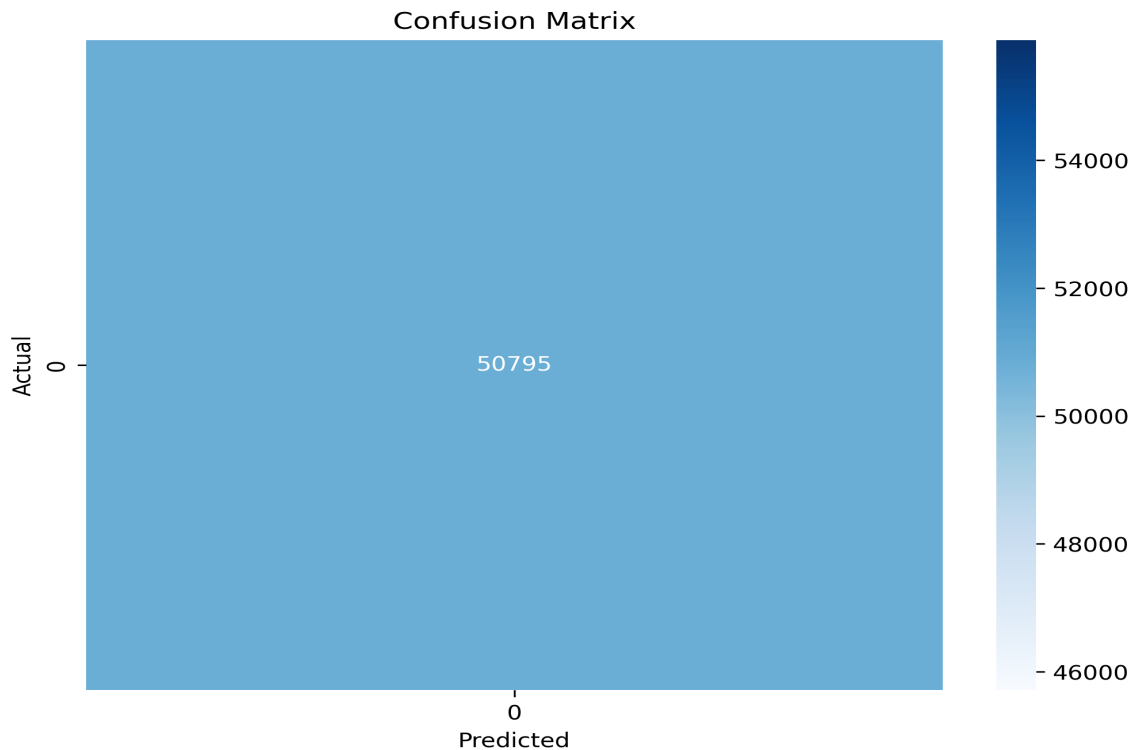
Outlier Analysis

Prediction Analysis

Model Accuracy: 100.00%

Classification Report

precision recall f1-score support 0 1.00 1.00 1.00 50795 accuracy 1.00 50795 macro avg 1.00 1.00 1.00 50795 weighted avg 1.00 1.00 1.00 50795

Confusion Matrix

Confusion Matrix

Recommendations

• Impute missing values in 'short_composition2' using appropriate statistical methods to minimize bias and enhance data integrity.

• Explore the reasons behind the wide price range and advocate for policies that promote affordability and equitable access to essential medications.

• Monitor market concentration and encourage competition to increase patient choice and drive innovation in the pharmaceutical industry.

• Promote rational use of pain relievers through public health campaigns and provider education to address potential overuse and misuse.

• Regularly review and update the dataset to ensure it remains current and informative for research and decision-making.