Toyota Sales Dataset — Full Data Analysis Report

Executive Summary

This report presents a comprehensive data analysis of the Toyota sales dataset using Python (Pandas, NumPy, Matplotlib/Seaborn) and SQL-style analytical logic. The objective is to uncover pricing drivers, customer preferences, depreciation behavior, and actionable business recommendations to support inventory planning, pricing strategy, and profitability.

Key Findings (High-Level):

- Vehicle **price is primarily driven by year (age), mileage, fuel type, and transmission**.
- **Hybrid and automatic vehicles** consistently command premium prices.
- Vehicles aged **3–6 years with moderate mileage** offer the best balance between demand and margin.
- Certain Toyota models demonstrate **strong resale value stability**, while others underperform.

Dataset Overview

Dataset Description

The dataset contains listings of Toyota vehicles with attributes related to pricing, technical specifications, and usage history.

Key Columns:

- `model` – Toyota vehicle model
- `year` – Manufacturing year
- `price` – Sale price
- `mileage` – Distance driven
- `fuelType` – Petrol, Diesel, Hybrid
- `transmission` – Manual or Automatic
- `engineSize` – Engine displacement (litres)
- `mpg` – Fuel efficiency
- `tax` – Road tax

BUSINESS INSIGHTS & RECOMMENDATIONS

Mileage & Depreciation Insights

🔍 Insights

- Price declines steadily as mileage increases.
- The sharpest price drop occurs after **~60,000–80,000 miles**.
- Some models retain value better at higher mileage.

�🗸 Recommendations

- Target inventory below **60,000 miles** for maximum resale value.
- Aggressively price high-mileage cars to reduce holding time.
- Promote models that show strong value retention even at higher mileage.

Age (Year) Insights

🔍 Insights

- Newer vehicles command significantly higher prices.
- Depreciation accelerates after **5–6 years**.
- Very old vehicles stabilize at a low price range.

✓ Recommendations

- Best resale window: **3–6 year-old cars**.
- Avoid overstocking cars older than 8–10 years unless deeply discounted.
- Emphasize "low age + low mileage" bundles in listings.

Fuel Type Insights

🔍 Insights

- **Hybrid cars have the highest average prices**.
- Petrol cars dominate volume.
- Diesel cars show lower prices and declining attractiveness.
- Hybrid listings increase in newer model years.

✅ Recommendations

- Increase acquisition of **hybrid Toyota models**.
- Reduce diesel inventory exposure.
- Market hybrids as **fuel-efficient and future-proof** options.

Transmission Insights

🔍 Insights

- **Automatic cars are priced higher** than manual.
- Manual cars dominate older and budget segments.
- Automatic cars depreciate slower.

✅ Recommendations

- Prioritize automatic transmission vehicles for higher margins.
- Price manual vehicles competitively to move inventory faster.
- Clearly highlight transmission type in listings.

Engine & MPG Insights

🔍 Insights

- Larger engines correlate with higher prices.
- High MPG cars maintain better resale value.
- Mid-range engines (1.5–2.0L) offer the best balance.

✅ Recommendations

- Focus on **1.5–2.0L engine vehicles**.
- Promote high-MPG cars as low-running-cost options.
- Avoid large engines unless positioned as premium.

Executive Summary

📍 Key Insights

- Price is mainly driven by **year, mileage, fuel type, transmission**.
- Hybrids and automatics command premium pricing.
- Mid-age, mid-mileage cars perform best.
- Certain Toyota models consistently outperform others.

🎯 Strategic Recommendations

- Optimize inventory around **3–6 year-old automatic hybrids**.
- Reduce diesel and high-mileage stock.

Conclusion

This analysis demonstrates how Python-driven analytics can uncover powerful insights from automotive sales data. By aligning inventory and pricing strategies with data-backed insights, dealerships and analysts can significantly improve profitability, turnover, and decision-making accuracy.