

CarDekho Dataset - Preprocessing & Exploratory Data Analysis

Prepared for: Assignment Submission

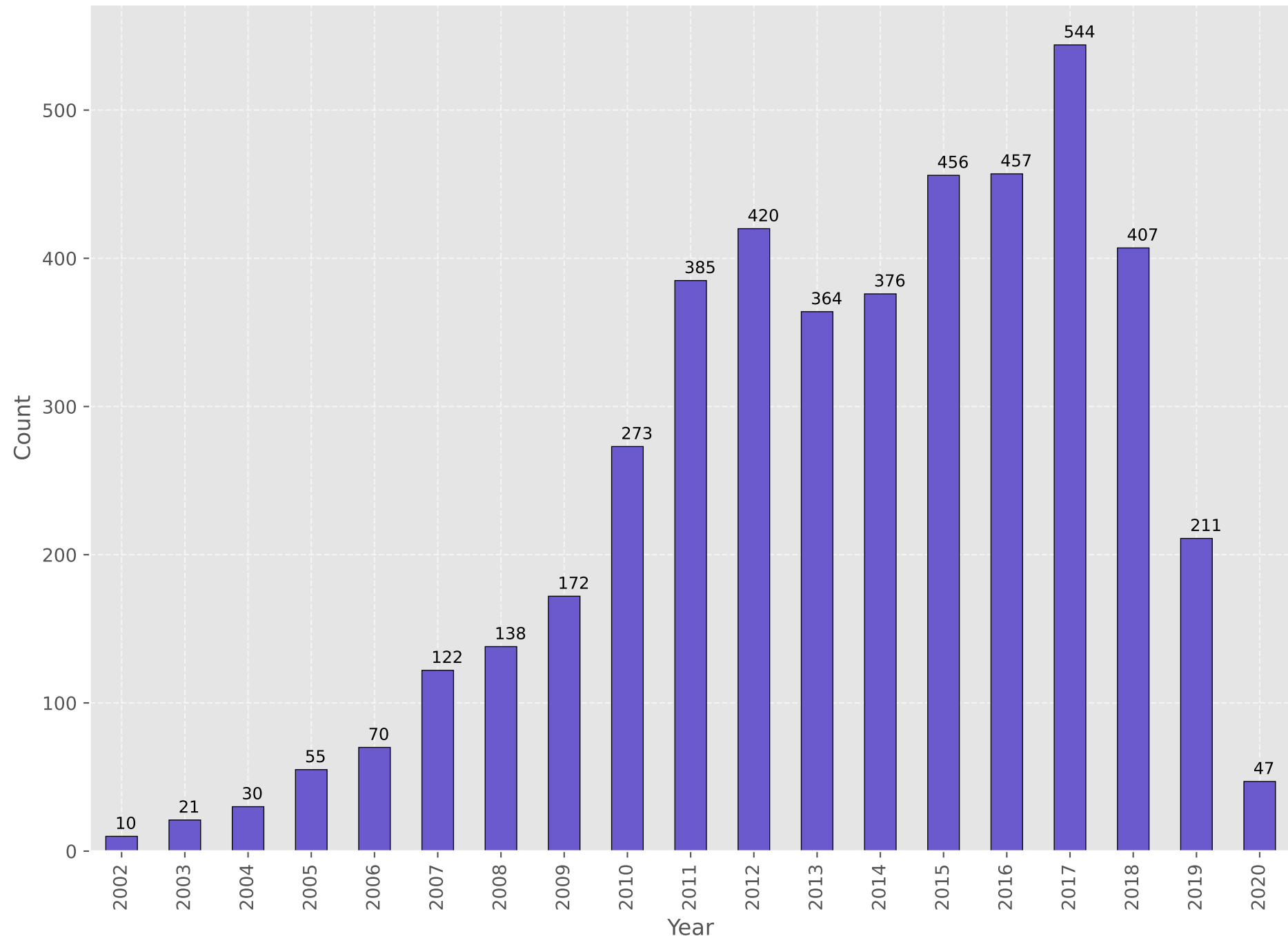
Contents:

1. Data Overview & Preprocessing Summary
2. Analysis 1: Year-wise Distribution of Cars
3. Analysis 2: Average Selling Price by Fuel Type
4. Analysis 3: Engine Size vs Selling Price
5. Analysis 4: Seller Type vs Average Selling Price
6. Analysis 5: Transmission Type Impact on Price
7. Correlation Heatmap

Preprocessing Steps:

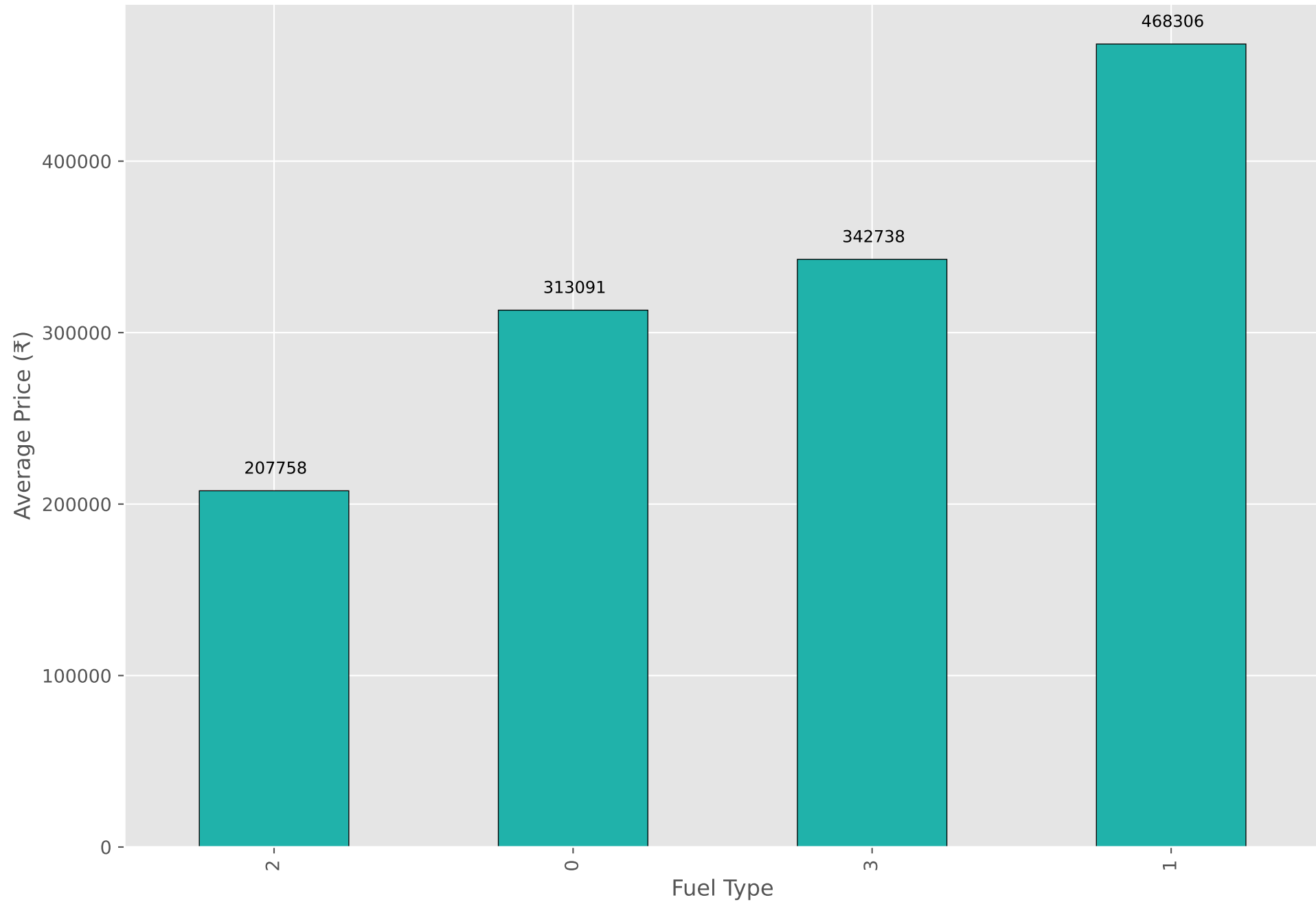
- Missing value handling (Mean/Mode)
- Label Encoding for categorical fields
- Outlier removal using IQR
- Feature Scaling using StandardScaler

Analysis 1: Number of Cars by Manufacturing Year



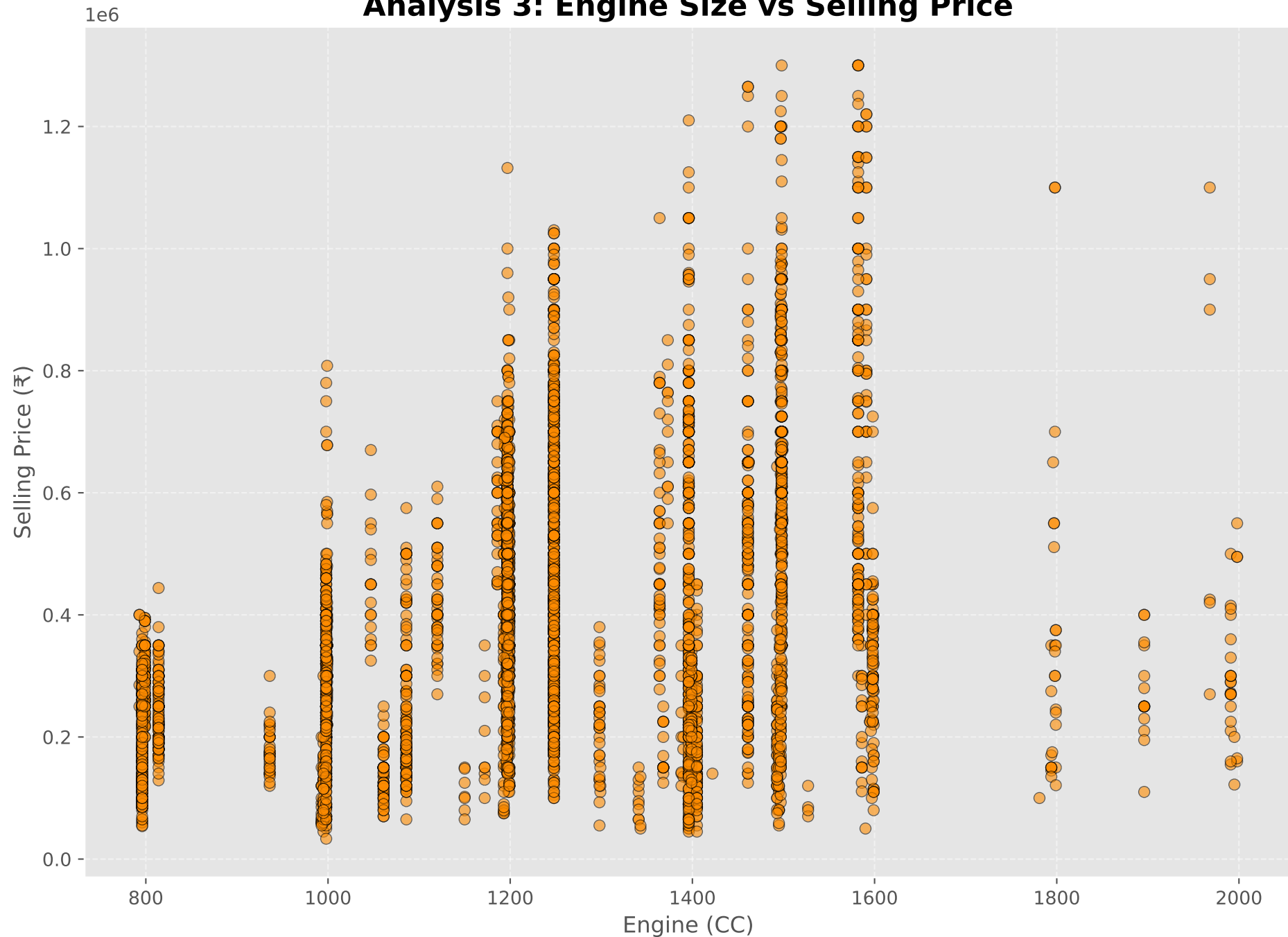
Insight: Most cars are from 2010-2017, indicating stronger resale market for mid-aged cars.

Analysis 2: Average Selling Price by Fuel Type



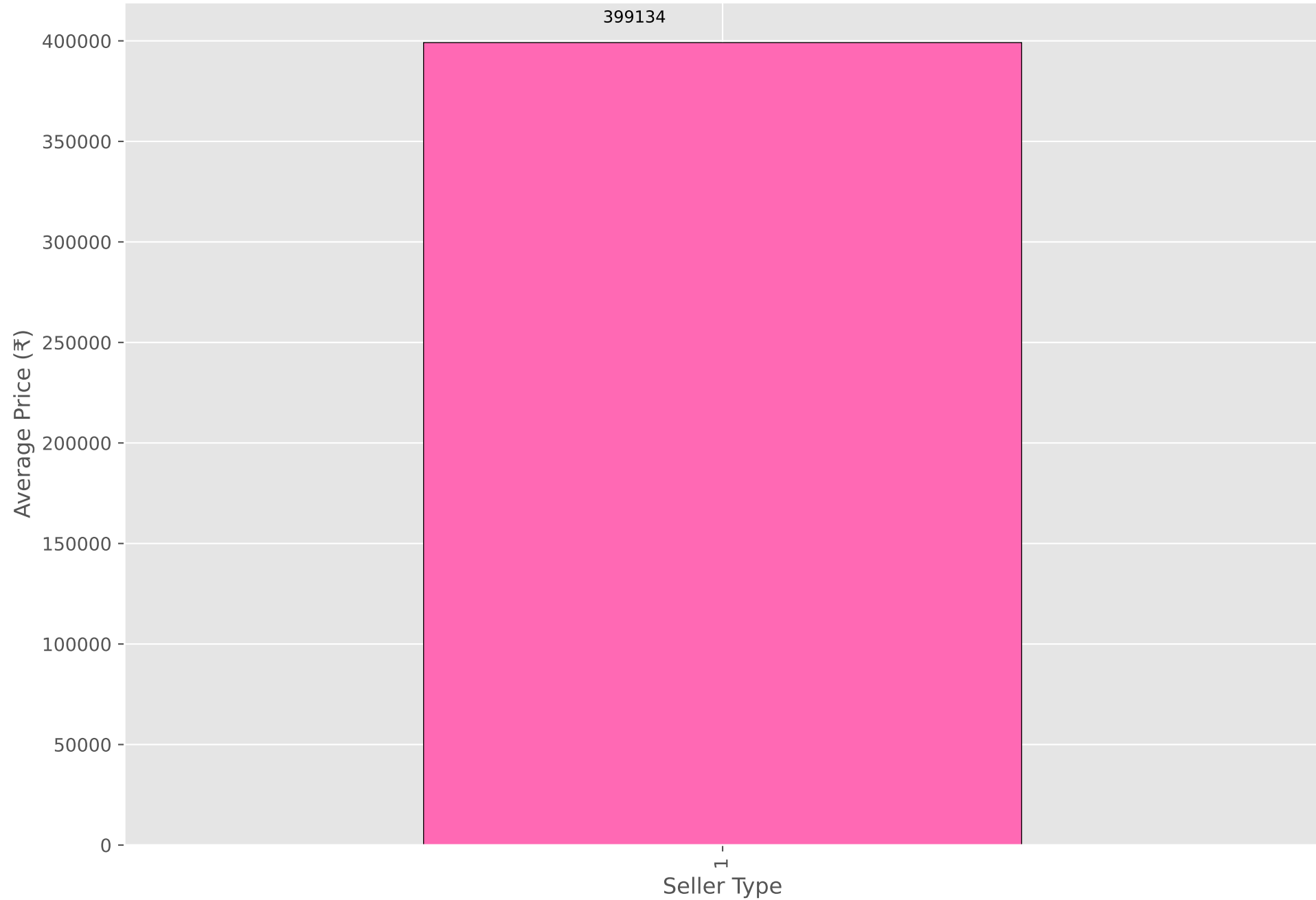
Insight: Diesel cars show higher resale prices compared to Petrol or CNG cars.

Analysis 3: Engine Size vs Selling Price



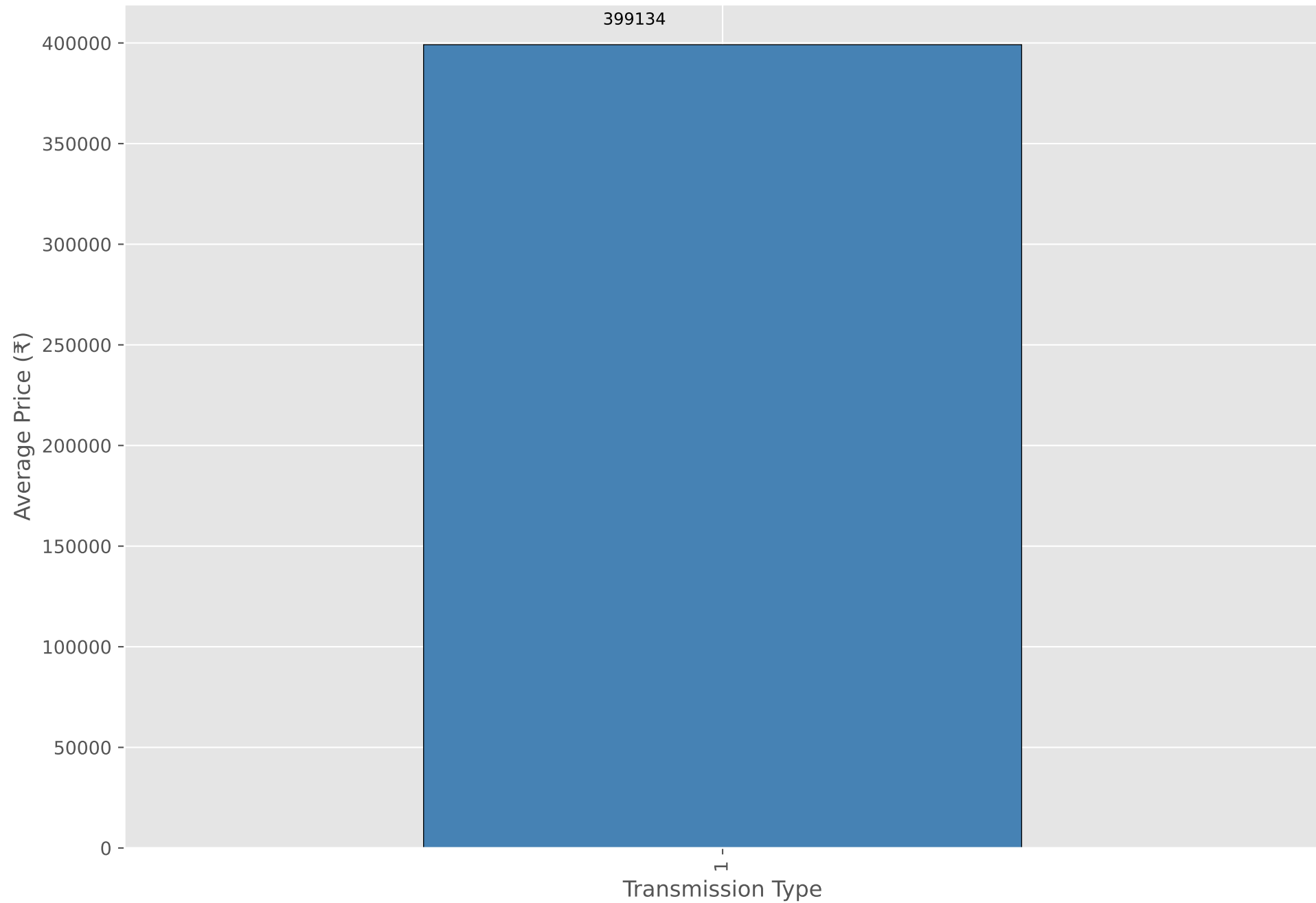
Insight: Correlation = 0.347. Larger engines tend to command higher prices.

Analysis 4: Average Selling Price by Seller Type



Insight: Dealers list higher prices due to added service and warranty value.

Analysis 5: Average Price by Transmission Type



Insight: Automatic cars attract higher resale prices due to comfort demand.

Correlation Heatmap of Numerical Features

