Mileage Matters: Data-Driven Classification for Smarter Vehicle Sales

NORQUEST COLLEGE

Members:
Inderjeet Singh
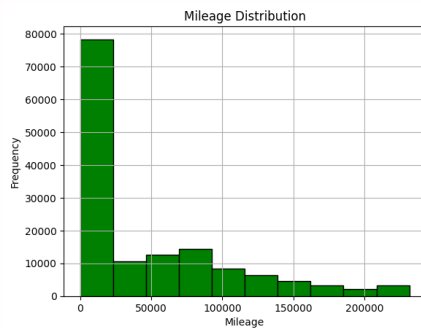Aditya Mehta
Harvir Kaur
Ishmeet Singh

Instructor:
Md Mahbub Mishu, PhD

Contact
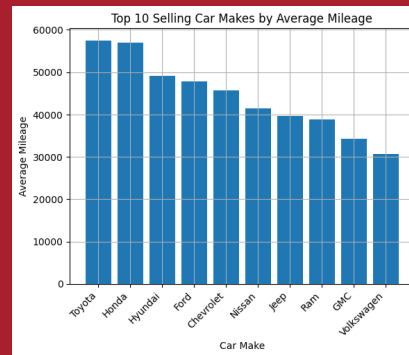Email: mdmahbub.mishu@norquest.ca
Phone: +1 825−439−3680

Introduction

We analyzed **145,114 vehicle records** to understand mileage-based trends and their impact on vehicle valuation. Dealerships struggle with **pricing and inventory management** due to varying mileage effects on resale value. Through **Exploratory Data Analysis (EDA)**, we found a **strong negative correlation (-0.6) between mileage and price**, confirming that higher mileage reduces vehicle worth. Most vehicles fall within **0 – 25,000 km**, while luxury cars retain value better despite high mileage. Missing data was handled via imputation. These insights will aid in developing a **multi-class classification model** to help dealerships optimize **pricing, marketing, and inventory strategies**, ensuring better customer targeting and decision-making.
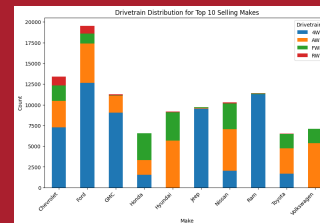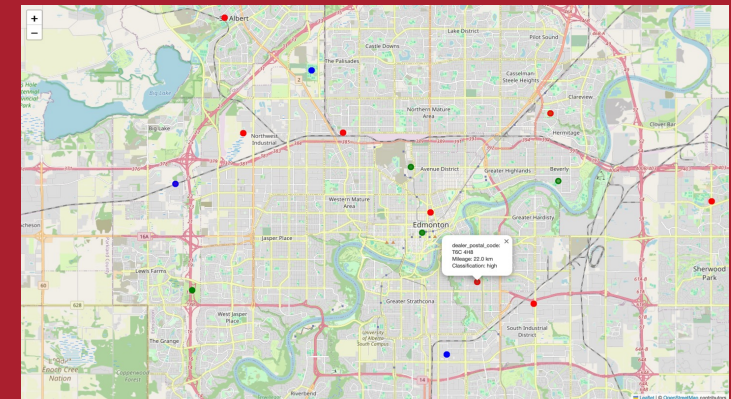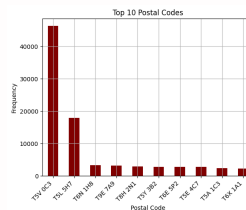
Make and drivetrain count

Mileage classification by Postal Codes

Mileage Distribution
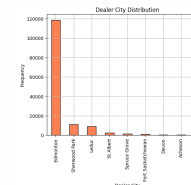

Top 10 listing by average Mileage
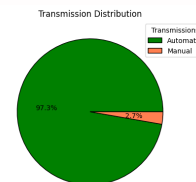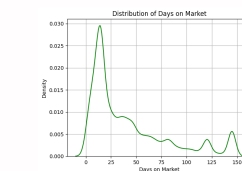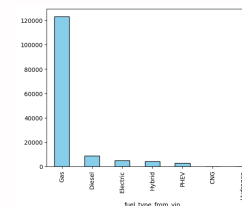

Stock type VS Listing type

Top 10 regions

Univariate Analysis
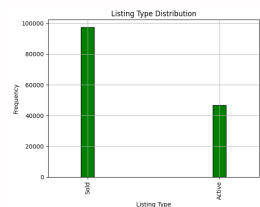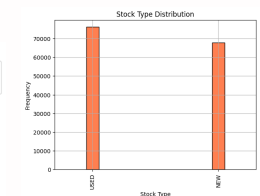
CONCLUSION

Our **EDA on 145,114 vehicle records** reveals that **mileage negatively impacts price (-0.6 correlation)**, with most vehicles in the **0-100,000 km range**. **Used vehicles** are mainly from **2015-2023**, while **new inventory is dominated by 2024 models**. The market is led by **Ford, Chevrolet, and GMC**, with **AWD and 4WD being the most common drivetrains**. Gas-powered vehicles dominate, electric and hybrids remain limited, and **97.3% of vehicles have automatic transmissions.** Edmonton has the **highest dealership concentration** with rapid inventory turnover. These insights will aid in developing a **multi-class classification model** for optimized **pricing, inventory management, and customer targeting**