# Predictive Analysis Using Multiple Linear Regression

Manasi Odassery

B94

1. Introduction

This document synthesizes the findings from two separate predictive modeling projects. The first project explores factors influencing profitability in startups, and the second examines determinants of pricing for used Toyota Corolla cars. By applying multiple linear regression, these analyses illuminate the underlying patterns and provide a basis for enhanced decision-making.

2. Objective

- Startups: To determine how different types of expenditures (R&D, Administration, Marketing) influence startup profitability.
- Toyota Corolla: To identify key vehicle attributes that affect the market price of Toyota Corolla models.

3. Methodology

- Data Collection: Data for the startups was sourced from a compiled database of startup financials. The Toyota Corolla dataset was obtained from an automotive resale platform detailing various car attributes.
- Data Preparation: Both datasets underwent cleaning, transformation, and normalization processes to prepare for modeling. Categorical variables were encoded, and numerical outliers were addressed.
- Model Building: Multiple linear regression models were developed for each dataset using Python's sklearn library.
- Validation: Models were validated using a train-test split approach to ensure they accurately predict unseen data.

4. Analysis I: 50 Startups

- Data Overview: Consisted of financial details from 50 startups, focusing on three main expenditures.
- EDA Insights: Highlighted significant correlations between R&D spending and higher profits.

- Model Results: Achieved an $R^2$ of 0.93, indicating excellent model performance.
- Business Impact: Suggested increased allocation towards R&D for enhanced profitability.

5. Analysis II: Toyota Corolla

- Data Overview: Included attributes like age, mileage, horsepower of 1,436 Toyota Corolla cars.
- EDA Insights: Showed that newer, less-used cars with higher horsepower tend to have higher prices.
- Model Results: Produced an $R^2$ of 0.85, confirming strong predictive capability.
- Business Impact: Recommendations were made for dealerships to focus on newer, high-performance models in their sales strategy.

6. Combined Strategic Recommendations

Strategic insights from both analyses advocate for data-driven decision-making to optimize budget allocation in startups and inventory management in car dealerships, enhancing overall profitability and customer satisfaction.

7. Conclusion

The regression analyses have provided valuable insights into factors driving profitability and pricing in distinct sectors. These findings can guide more informed strategic planning and operational adjustments.