**Objective:**  
Efficiently allocate drivers in areas where ride demand is likely to be high by analyzing historical ride request data, clustering similar demand patterns, and forecasting future demand using machine learning.

1. **Model Building for Demand Forecasting**

**What i Did:**

* Chose **cnt (ride count)** as the target.
* Selected relevant features from the cleaned & clustered data.
* Split data into train and validation sets using train\_test\_split.
* Applied two models:
  + **Linear Regression**
  + **Random Forest Regressor**
* Evaluated using **Mean Absolute Error (MAE)**

**Why:**

* **Forecasting future ride demand** helps pre-position drivers efficiently.
* Random Forest outperformed Linear Regression slightly:
  + Linear Regression MAE: 59.88
  + Random Forest MAE: 59.33
* This validates the use of **tree-based models** for capturing non-linear relationships

1. **Model Evaluation & Selection**

**What i Did:**

* Compared MAE scores to assess model accuracy.

**Why:**

* Evaluation ensures the selected model **generalizes well** to unseen data.
* MAE is an intuitive metric — lower means more accurate predictions.
* Random Forest was chosen for its **slightly better performance** and robustness.

**✅ Final Insights: Model Evaluation Summary**

Here’s a clear interpretation of your results:

| **Model** | **MAE (Mean Absolute Error)** |
| --- | --- |
| **Linear Regression** | **59.88** |
| **Random Forest** | **59.33** |

**🔍 Interpretation:**

* **Random Forest Regressor performed slightly better** with a lower MAE (59.33) compared to Linear Regression (59.88).
* While the improvement is **modest**, Random Forest still **offers better predictive power**, especially for **non-linear** relationships.
* If **real-time predictions and interpretability** are key, Linear Regression could still be valuable.
* But for **accuracy-focused tasks like demand forecasting**, **Random Forest** is your best bet.