**Title:** Optimising Car Sharing Profitability with a Regional Pricing Strategy

**Deliverable**: A usable pricing tool to help us model the effects of different pricing options on profitability

**Background:** Car share is price sensitive and seasonal as many B2C journeys are for social/discretionary use such as shopping or leisure trips. We have a significant amount of trip data on how and when our vehicles are used, plus a clear idea on our costs to provide the service both fixed and variable, but we have traditionally charged one fixed price across all locations even though demand and local economies vary enormously.

**Why is this piece of work required:** We want to look at how best to use our data on **usage and demand patterns** to better inform our pricing strategy. Ideally, we would want to develop a straightforward pricing tool where we could input all costs – including variables such as fuel/electricity – and output options for hourly and daily pricing based on locations, demand, seasonal variations or potentially time of day (e.g. cheaper overnight hire).

We would want this to give us a clearer idea of profitability in locations and the potential outcomes if we adopted different pricing options and what the impact of this might have on utilisation.

**Input:**

1. Booking lengths (booking duration)
2. No of booking
3. Miles travels (total number of miles travelled region-wise)
4. Mileage (Total mileage region-wise)

**Vital URL:**

1. [Dynamic pricing through data science](https://www.youtube.com/watch?v=cKPaIsOQslo)

**Problem Statement:**

* I’m looking to develop region specific pricing tool for four vehicle types – City, Everyday, Family and Van. I’m excited to predict hourly and daily rates based on the historical data.
* Historical data has information about – vehicle information, bookings details, location, rates and tariffs.
* This seems a regression problem.

**Predictive Modelling and Feature Engineering**

**Key Components of the Dataset**

1. **Vehicle Information**: `vehicle\_description`, `vehicle\_type`, `Fuel Type`, `Size Category` — These features can help you determine the pricing based on the type of vehicle, which may vary in cost due to size, fuel efficiency, or luxury status.

2. **Location Information**: `location\_description`, `location` — This is crucial for dynamic pricing as the demand and pricing strategy might differ by location due to varying levels of traffic, popularity, or economic factors.

3. **Booking Details**:

* `booking\_start`, `booking\_end`, `booking\_duration`, `booking\_actual\_start`, `booking\_actual\_end`, `booking\_actual\_duration` — These temporal features are essential for understanding when demand is highest and how long vehicles are typically used.
* `booking\_mileage`, `booking\_estimated\_cost`, `booking\_actual\_cost\_total` — These could inform you about the usage patterns and costs associated with different bookings.

4. **Rates and Tariffs**: `booking\_rates\_hours`, `booking\_rates\_24hours`, `hourly\_rate`, `daily\_rate` — Directly related to your target variables. Analyzing how these rates change in response to other variables in your dataset can help establish a pricing model.

5. **Seasonality and Time Variability**: `booking\_created\_at`, `booking\_cancelled\_at` — These features can help you capture seasonal effects and how they influence booking patterns.

6. **Demand Indicators**: Count of bookings per time slot or location, derived from `booking\_id` and timestamps — Useful for understanding and predicting peak demand periods.

**Suggested Steps for Building Your Model**

1. Feature Engineering:

- Extract time components like hour of the day, day of the week, month, and possibly the year from `booking\_start` and `booking\_end`.

- Create new features such as holiday flags or special event periods which might influence demand.

- Compute derived variables like average daily bookings per location, average duration, and cancellation rates.

2. **Exploratory Data Analysis (EDA)**:

- Analyze the distribution and trends of hourly and daily rates.

- Investigate patterns based on vehicle type, location, and time to identify potential high-demand triggers.

3. **Model Selection**:

- Given the dataset includes time-series data, consider models capable of capturing temporal dynamics like XGBoost, LightGBM, or even deep learning models such as LSTM if the sequence of bookings is important.

- Prophet or ARIMA could be useful to model and forecast prices based on detected seasonal trends and cycles.

4. **Model Training**:

- Split your data into training and test sets, considering the chronological order to avoid look-ahead bias.

- Employ cross-validation, particularly time-series cross-validation.

5. **Model Evaluation and Tuning**:

- Evaluate model performance using appropriate metrics (e.g., RMSE for regression tasks).

- Adjust hyperparameters and perhaps ensemble several models to improve accuracy.

6. **Deployment**:

- Implement the model in a production environment where it can update pricing in real-time or near-real-time based on the incoming booking data and other contextual information.

- Set up a monitoring system to track the performance of your pricing model and adjust as needed.