**EE4211 - DATA SCIENCE FOR IOT**

National University of Singapore

College of Design and Engineering

# Group Project
## Dynamic Car Park Pricing Model

## Submitted by:


**Group-4**
**Arumugam Vijayakumar Bhavya - A02389393J**
**Li Tian - A0293441E**
**Ng Yong Qian - A0238989Y**
**Ahmad Farhan Bin Saiuflddin - A0254515H**

# 1. Introduction and Problem Description

**Urban Challenge**: Cities experience significant issues with car park congestion, especially during peak hours. This leads to inefficiencies, user frustration, and increased emissions. Traditional static pricing models don't adapt to real-time demand changes, causing underutilization during off-peak times and overcrowding during peak periods.

**Objective**: This report presents a data-driven dynamic pricing model for carparks aimed at optimizing space usage, enhancing revenue, and improving user experience by predicting demand and adjusting prices accordingly.

**Scope and Importance**: Implementing this system can alleviate urban congestion, balance carpark demand, and contribute to sustainable city initiatives. The approach aligns with smart city goals and supports better urban planning and traffic management.

# 2. Experimental Setup

**Assumptions**:

- Historical data trends, including occupancy, temperature, and rainfall, serve as reliable predictors of future parking demand.
- External variables, particularly weather conditions and holiday periods, exert significant influence on parking utilization patterns.

**Dataset Description**:

- **Carpark Data**: Contains timestamps, total lots, and available spaces.
- **Weather Data**: Includes hourly temperature and rainfall measurements. (two different dataset)
- **Holiday Data**: Denotes holiday dates, collected to understand their impact on parking demand.

**Preprocessing Steps**:

- **Timestamp Parsing**: Converted raw timestamps to datetime objects for feature engineering.
- **Feature Engineering**:
    - **Time-Based Features**: Extracted features such as the day of the week and hour of the day, and combine them to capture interactions which might influence parking patterns.
    - **Lag Variables**: This feature serves as temporal memory in our model. For instance, if we implement a 3-hour lag, our model considers parking availability from 3 hours ago to predict current demand. This captures important temporal dependencies if a parking facility reaches capacity during morning rush hour, this often indicates sustained high occupancy for subsequent hours. And we've carefully selected multiple lag periods to capture both short-term and long-term patterns.
    - **Rolling Averages**: Rather than using instantaneous temperature or rainfall measurements, we calculate averages over specific time windows. Applied rolling averages to temperature and rainfall data to smooth out short-term fluctuations, providing a clearer view of trends and patterns.
    - **Holiday Flags**: Added binary features to mark holidays.

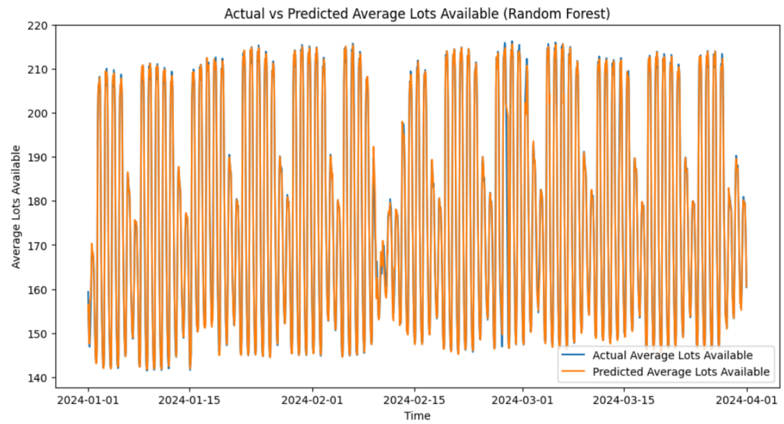
**Model**:

- **Algorithm**: Selected the Random Forest Regressor due to its robustness against overfitting and its ability to handle complex interactions between features. This model is well-suited for capturing non-linear relationships in the data.
- **Training and Tuning**:
    - **Training Period**: January to March 2024.
    - **Testing Period**: April to June 2024.
    - **Hyperparameter Tuning**: Conducted using GridSearchCV and TimeSeriesSplit for time series cross-validation.

**Performance Metrics**:

- **Mean Squared Error (MSE)** and **R-squared ($R^2$)** for both training and validation phases to assess model accuracy.

*Prediction on validation set*




# 3. Results and Analysis

**Model Performance**:

- **Training Results**:
  - MSE: 3.9798
  - R-squared: 0.9930
- **Validation Results**:
  - MSE: 5.6228
  - R-squared: 0.9897
- **Test Results**:
  - MSE: 7.7460
  - R-squared: 0.9861

**Key Predictive Factors**: The feature importances highlight which factors contributed most to the accuracy of our model predictions:

1. **Top Features:**
   - **Lagged Occupancy (Lag 1):** This feature, with an importance score of **0.784**, was by far the most significant. It reflects the strong dependence of current parking availability on recent trends.

    - **Hour:** Scored **0.185**, indicating that time of day is a critical predictor due to daily demand cycles.
2. **Moderately Important Features:**
    - **Lagged Occupancy (Lag 2 & Lag 3):** These provide additional historical context, contributing scores of **0.0043** and **0.0104**, respectively.
    - **Day of Week:** While less impactful overall (**0.0101**), it helped capture weekly patterns like weekend vs. weekday differences.
3. **Less Influential Features:**
    - **Rainfall and Temperature:** Both had very small importance scores (~**0.0009**). This suggests weather effects are minimal for parking demand.
    - **Holidays:** Contributed an insignificant **0.00001**, likely because the dataset contained few holiday instances.

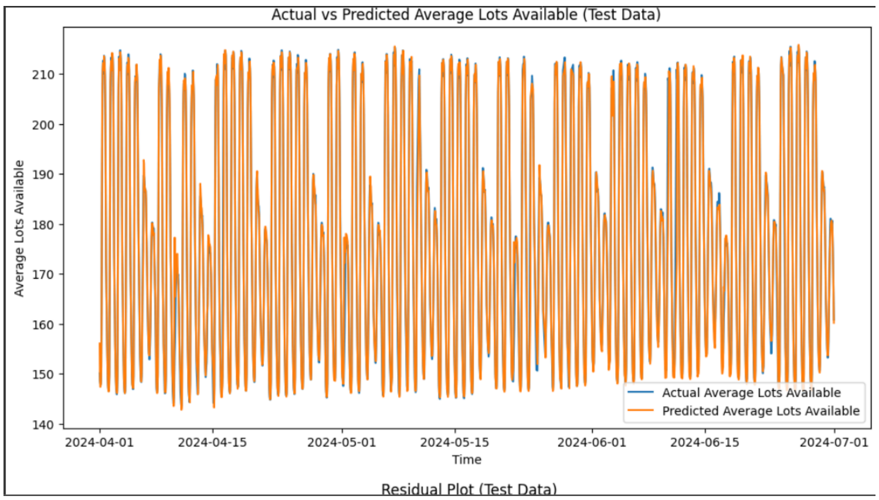
- The dominance of lagged occupancy underscores that short-term trends are key for forecasting.
- While external factors like weather and holidays provide some context, they had little impact on the model compared to time-based features.

```
Feature importances:
hour: 0.18944578575078103
day_of_week: 0.01142427657273982
hour_day_interaction: 0.002289293592890342
avg_lots_available_lag1: 0.7799829547657675
avg_lots_available_lag2: 0.003622748887148263
avg_lots_available_lag3: 0.011099805710945322
rainfall: 0.0001979625938676763
temperature: 0.0005940927815642026
rolling_temperature: 0.0006589688500043803
rolling_rainfall: 0.0006770985072700197
is_holiday: 7.011987021480182e-06
```

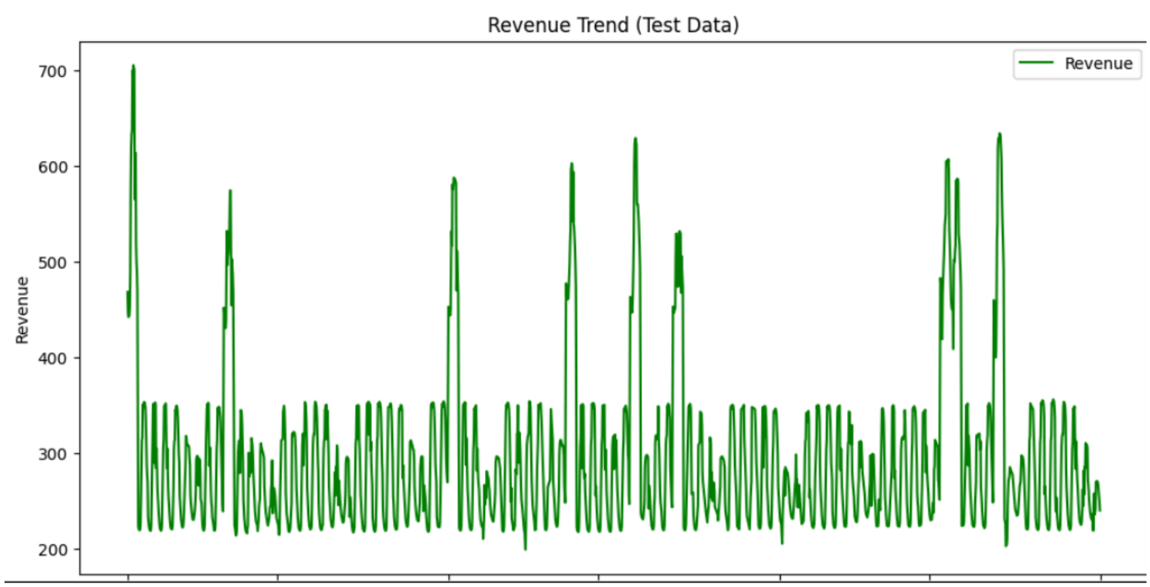
**Visualization of Results**:

- **Actual vs. Predicted Occupancy**: Graphs illustrating the model's accuracy over time, showing close tracking between actual and predicted values.

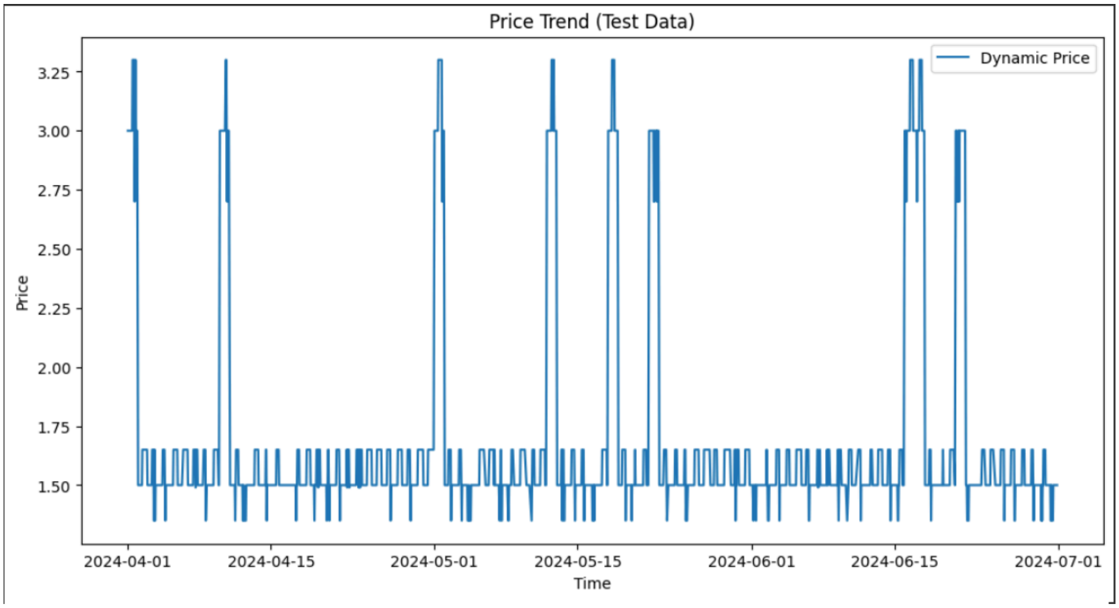
*Prediction on test set*



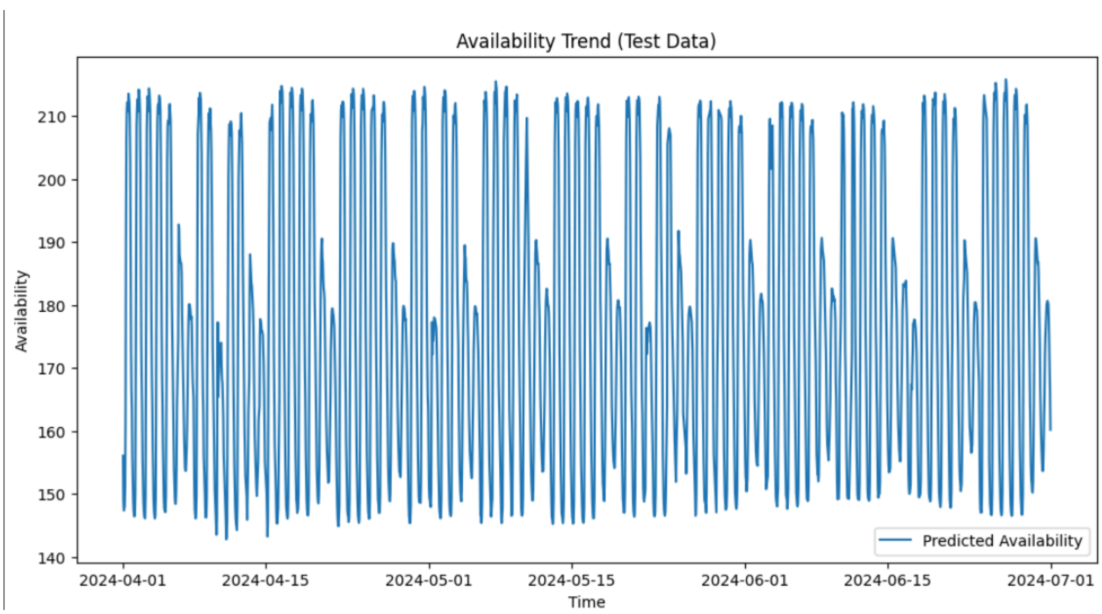

- **Price Adjustments and Revenue Analysis**:
  - The dynamic pricing model raised prices during high demand periods and lowered them during low demand times, achieving an estimated total revenue of SGD 573,050.89 during the test period.




Revenue Trend: The graph shows significant spikes during high-demand periods, where prices were raised to manage congestion and boost revenue.




Price Trend: Conversely, prices dropped during low-demand periods to encourage parking, smoothing utilization across the day.




These adjustments helped increase revenue while improving user satisfaction by maintaining parking availability. For instance, in off-peak hours, price reductions led to higher occupancy than a static pricing model would have achieved

# 4. Reflections and Conclusions

**Key Insights**:

- **Dynamic Adjustment**: The model effectively adjusted parking rates in real-time to optimize space utilization and maximize revenue. By responding to fluctuating demand during peak and off-peak hours, the model ensured a balanced distribution of vehicles, reducing the likelihood of overcrowding during high-demand periods and underutilization when demand was low.
- **Weather and Holiday Effects**: Although the influence of weather conditions and holidays was relatively minor compared to other factors, these external variables provided valuable insights into seasonal and temporal variations in parking behavior. Understanding these patterns allows for more nuanced pricing strategies that account for occasional spikes or drops in demand, enhancing the model's overall responsiveness.
- **User Experience**: Implementing a dynamic pricing system significantly improved the user experience by making parking availability more predictable. Drivers experienced fewer instances of frustration caused by unavailable parking spaces, as the system proactively managed demand. This predictability not only enhances user satisfaction but also encourages more consistent use of parking facilities.

**Practical Benefits**:

- **Reduced Congestion**: By evenly distributing parking demand throughout the day, the model effectively alleviated congestion during peak hours. This reduction in traffic bottlenecks not only makes the parking process smoother but also contributes to overall urban traffic flow improvements, mitigating the ripple effects of parking congestion on surrounding roadways.
- **Increased Revenue**: The ability to adjust prices based on real-time demand allowed parking operators to capitalize on high-demand periods by increasing rates, thereby boosting revenue. Conversely, lowering prices during low-demand times attracted more users, ensuring higher occupancy rates and maximizing profitability across different times of the day.
- **Sustainability**: Minimizing the time drivers spend searching for available parking spaces leads to lower vehicle emissions, contributing to cleaner air and reduced environmental impact. Additionally, improved traffic flow

resulting from reduced congestion supports broader sustainability goals by decreasing the carbon footprint associated with urban transportation.
- **Societal Impact**: The dynamic pricing model plays a crucial role in promoting urban sustainability by addressing both emissions and traffic congestion. Real-time applications that guide users to affordable and available parking spaces enhance smart city initiatives, fostering more efficient and user-friendly urban navigation. This not only benefits individual drivers but also supports the collective well-being of the community by creating a more organized and environmentally conscious urban environment.

**Key Takeaway:**

- **Data Integration**: Combining different data sources (e.g., real-time event data) could further enhance the accuracy and effectiveness of the pricing model. Integrating these additional data points allows for more precise predictions and adjustments, ensuring the model remains adaptable to a wide range of influencing factors.
- **Feature Engineering**: The creation of lag variables and rolling window features was pivotal in capturing time-dependent trends and autocorrelations within the data. These engineered features enabled the model to recognize and anticipate patterns in parking demand, leading to more accurate and reliable predictions.
- **Model Scalability**: The dynamic pricing approach demonstrated scalability, making it applicable to various parking facilities across different locations. By tailoring strategies to location-specific patterns and demand characteristics, the model can be effectively implemented in diverse urban settings, accommodating the unique needs and behaviors of each area.

**Future Work**:

- **Integration with Navigation Apps**: Linking the dynamic pricing model with popular navigation applications can provide users with real-time suggestions for optimal parking options based on their preferences and current data. This integration would enhance user convenience, allowing drivers to make informed decisions seamlessly within their existing navigation tools.
- **Enhanced Real-Time Updates**: Incorporating live feeds for traffic conditions, nearby events, and other real-time data sources can improve the model's responsiveness. By continuously updating the system with the latest

information, the pricing model can adjust more swiftly to sudden changes in demand, ensuring sustained efficiency and user satisfaction.

- **User-Focused Incentives**: Developing loyalty programs or offering discounts during low-demand periods can encourage increased engagement and utilization of parking facilities. These incentives not only attract more users during typically slower times but also foster customer loyalty, creating a more stable and predictable revenue stream for operators.