EDA Report – Dynamic Pricing (April–June 2018)

■ 1. Data Loading & Cleaning

File Used: Cleaned_Classes_April_June_final.csv

Key Columns: ActivitySiteID, ActivityDescription, BookingStartTime, MaxBookees, Number_Booked, Price_INR, BookingEndDateTime, DynamicPrice
Data Shape: 3289 records and 8 columns
Data Types: 4 numeric fields (float64), 4 categorical/object fields
Processing Steps: Parsed datetime, extracted Month/Hour/Weekday, standardized numeric fields.

■ 2. Descriptive Statistics

Metric	Value
Mean MaxBookees	~0.0
Mean Number_Booked	~0.0
Mean Price (INR)	~0.0
Mean DynamicPrice	-0.021864
Std MaxBookees	~1.0
Std Number_Booked	~1.0
Std Price_INR	~1.0
Std DynamicPrice	~1.19
Min/Max Values	Range ~-2 to ~2.5 (standardized)

■ 3. Key Visualizations (Generated)

• Booking Trend Over Time: Interactive time series showing total bookings.
• Monthly Booking Distribution: Boxplot showing variation by month.
• High-Demand Classes Distribution: Histogram of most booked classes.
• Correlation Heatmap:
Number_Booked ↔ Price_INR = -0.18
Number_Booked ↔ MaxBookees = 0.44
Price_INR ↔ MaxBookees = -0.09

■ 4. Outliers & Anomalies

• Outliers identified in Number_Booked and Price_INR using boxplot + IQR.
• Number of outlier records detected: ~17
• Recommendation: Inspect rows for unrealistic capacities or duplicates.

■ 5. Top Classes (By Average Attendance)

Activity Description	Avg Attendance	Mean Price	Count
Body Combat 11-12pm	4.52	-0.44	1
Line Dance 10.45-11.45am	3.65	-0.44	13
Body Attack 10-11am	3.57	0.18	1
Zumba Gold 12.45-1.45pm	2.42	0.18	13
Body Balance 12-1pm	2.10	0.82	1

Body Pump 5.45-6.45pm	1.96	0.18	13
Aqua Fit 11.15-12.00	1.72	-1.08	10
Aquafit 6.00-7.00pm	1.70	-1.08	13

■ 6. Notes & Next Steps

• Segment price elasticity models by class type for precision.
• Add external factors (holidays, weather) to improve forecasts.
• Deploy lightweight monitoring system to validate live dynamic pricing adjustments.