

NYC Taxi Demand Analysis

Exploratory Data Analysis Report

Dataset Period: July 01, 2014 - January 31, 2015

Total Observations: 10,320

Frequency: 30-minute intervals

Key Statistics

Average trips per 30-min: 15,138

Maximum trips (30-min): 39,197

Minimum trips (30-min): 8

Standard deviation: 6,939

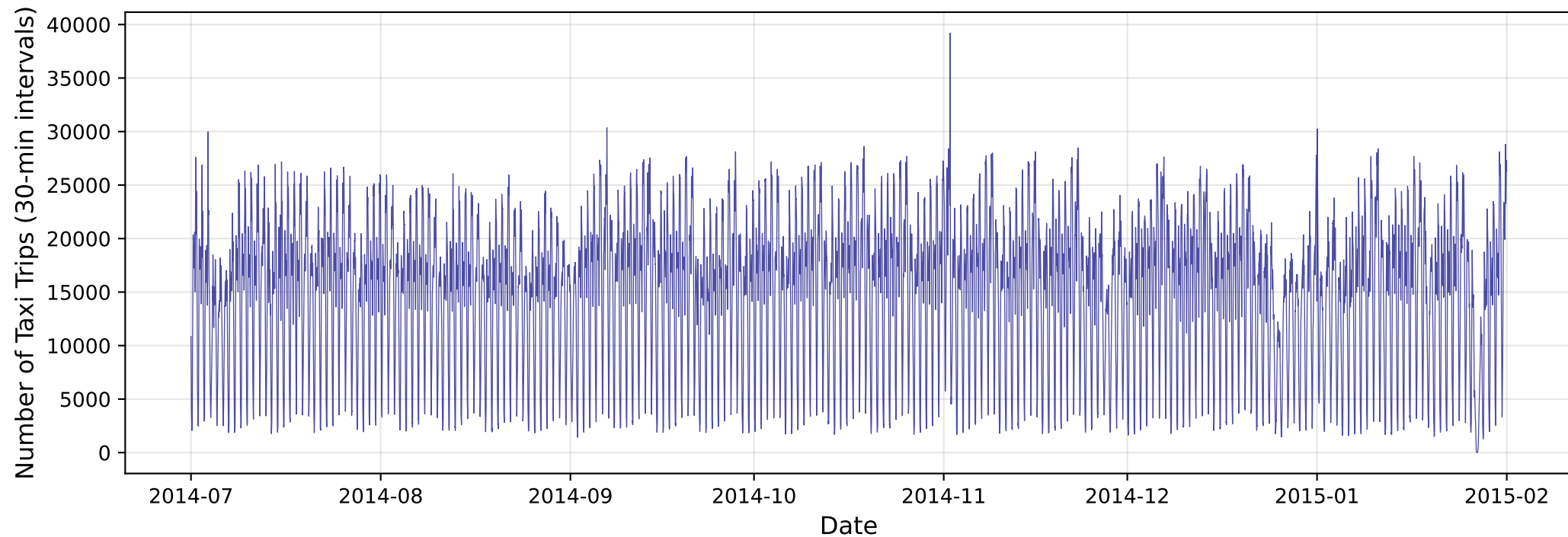
Data Quality:

- Missing values: 0
- Duplicate timestamps: 0
- Negative values: 0

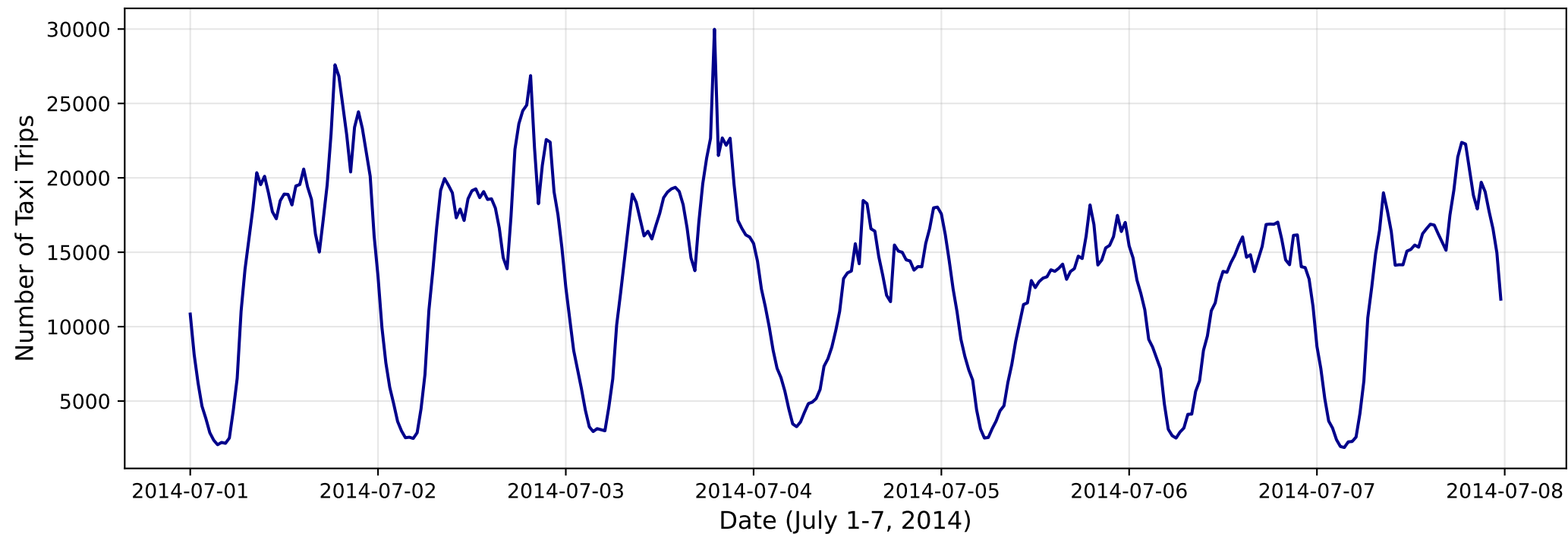
Generated on: September 12, 2025 at 10:37 PM

Time Series Analysis Project

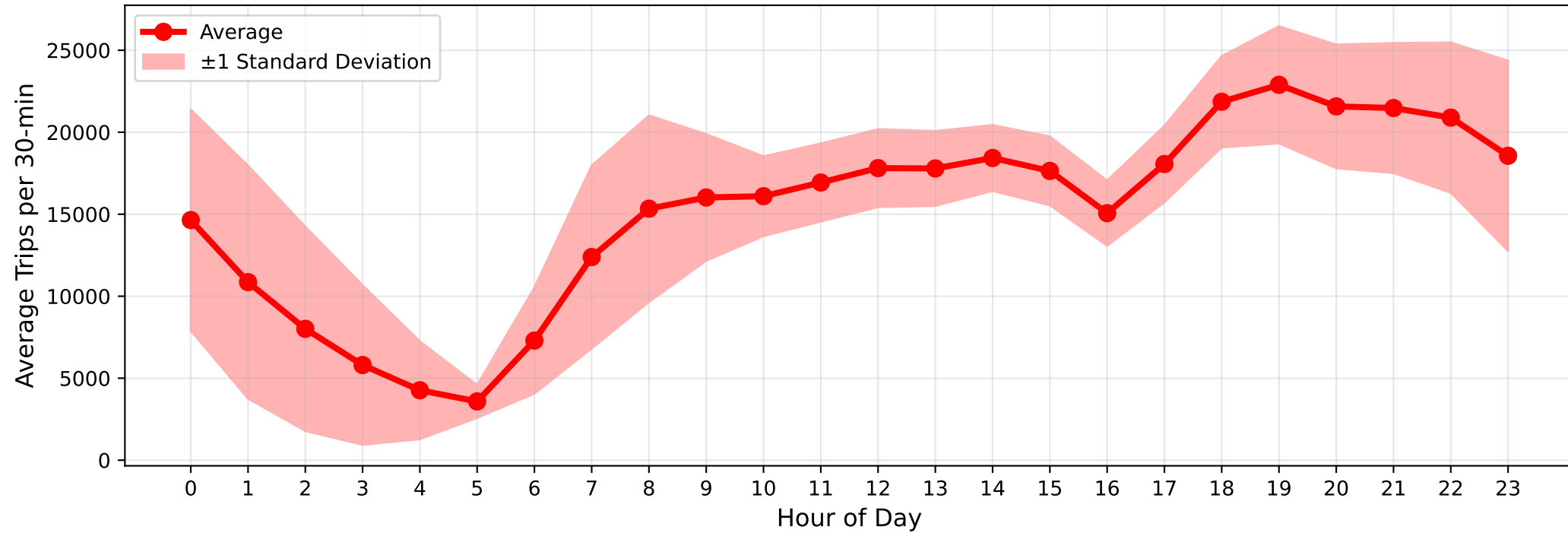
NYC Taxi Trips Over Time (Complete Dataset)



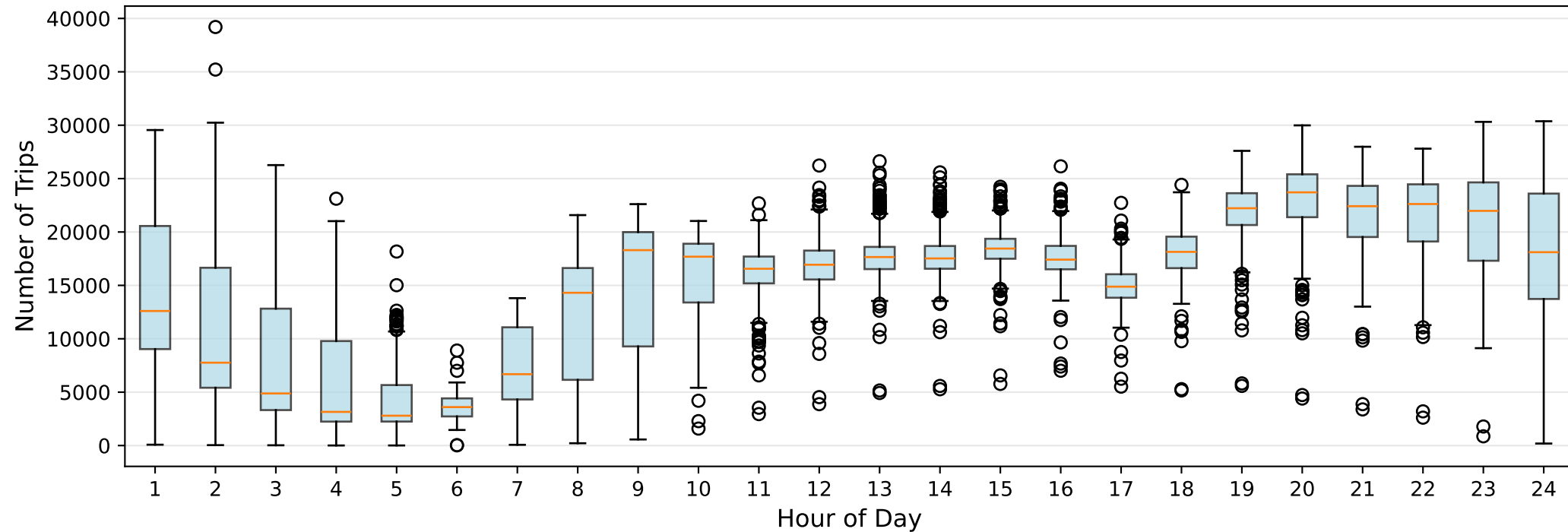
First Week Detail - Daily Patterns Visible



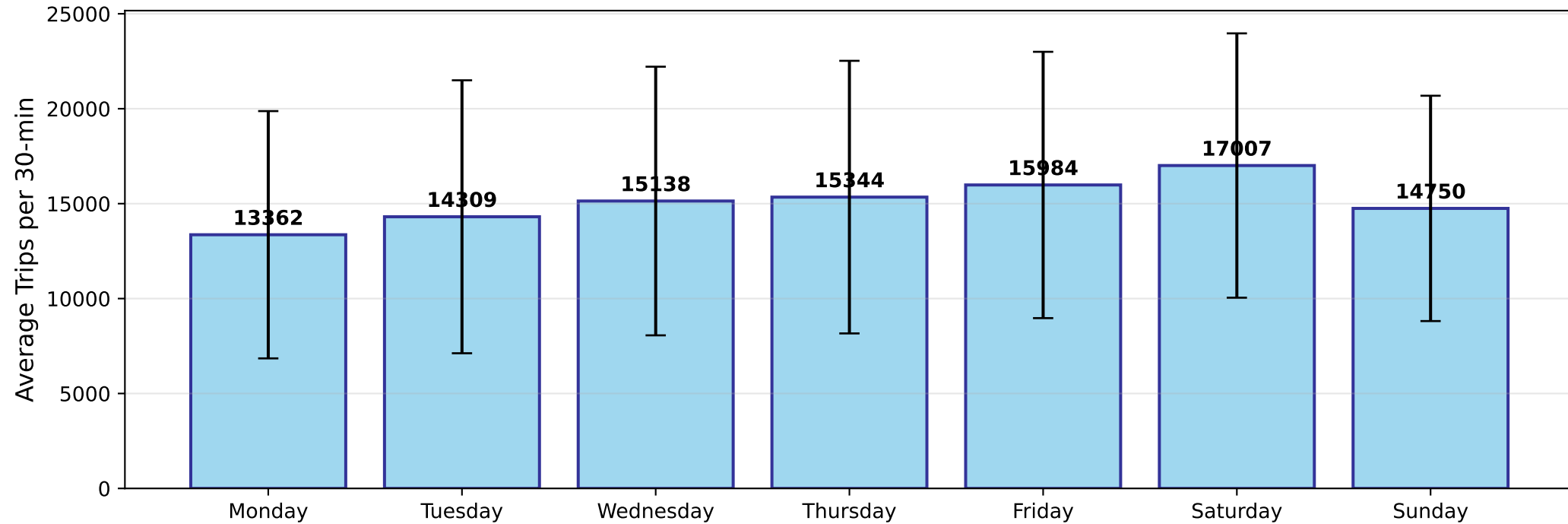
Average Taxi Trips by Hour of Day



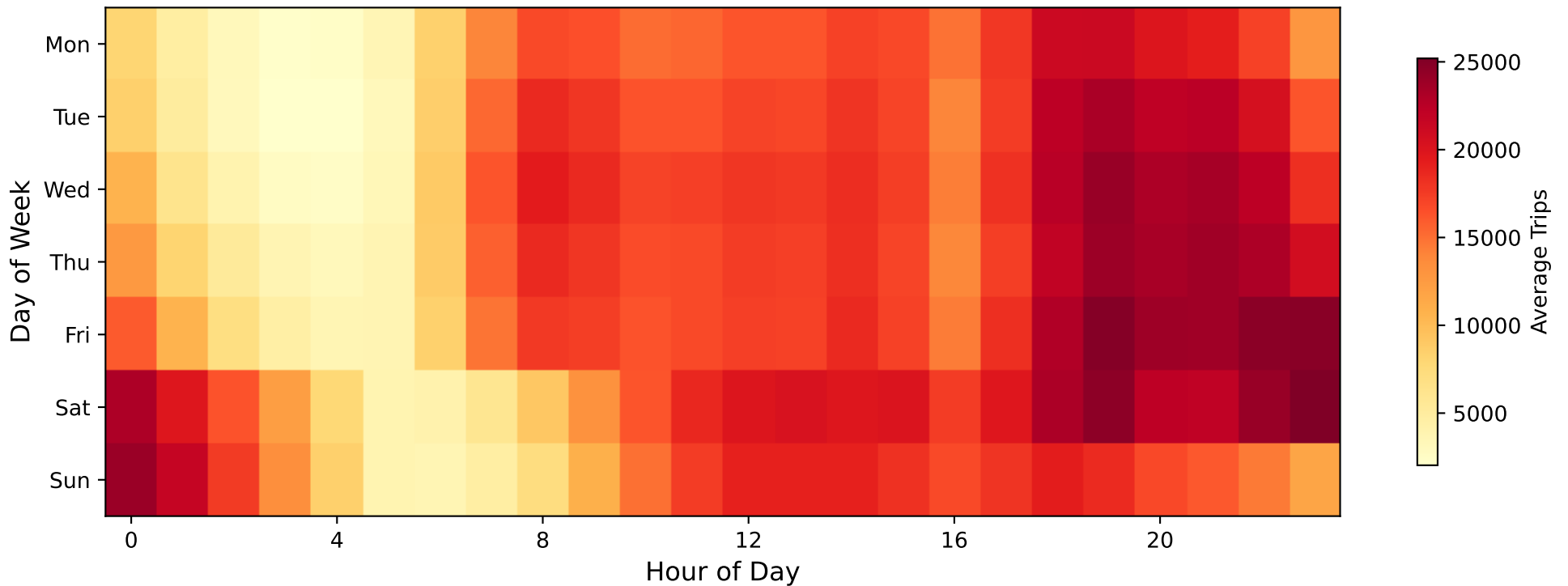
Distribution of Trips by Hour (Boxplot)



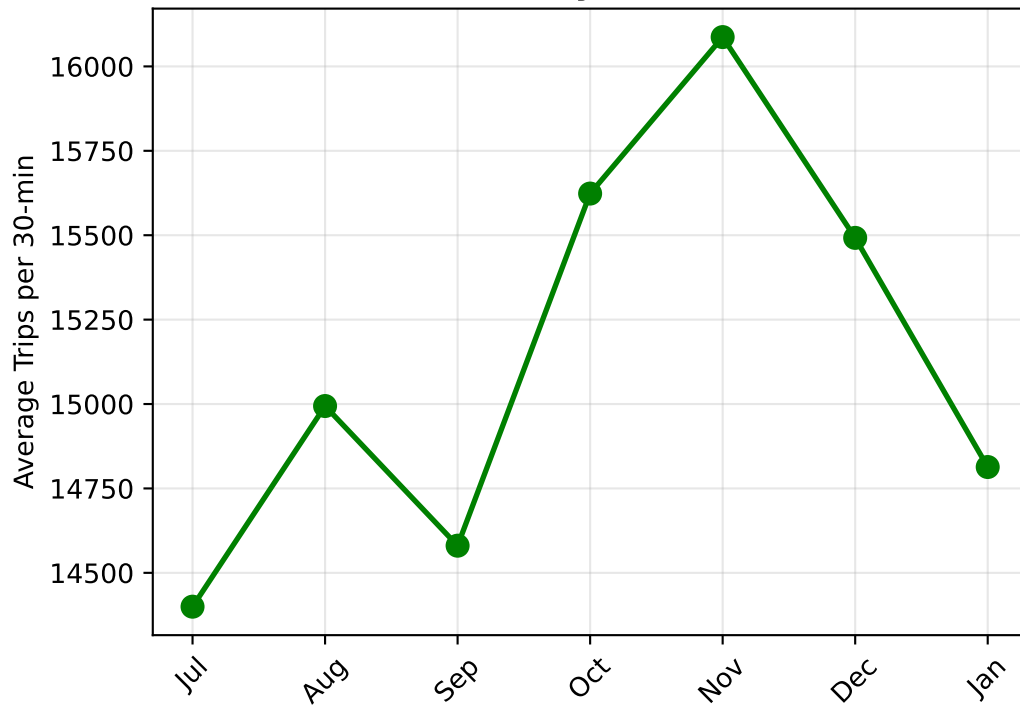
Average Taxi Trips by Day of Week



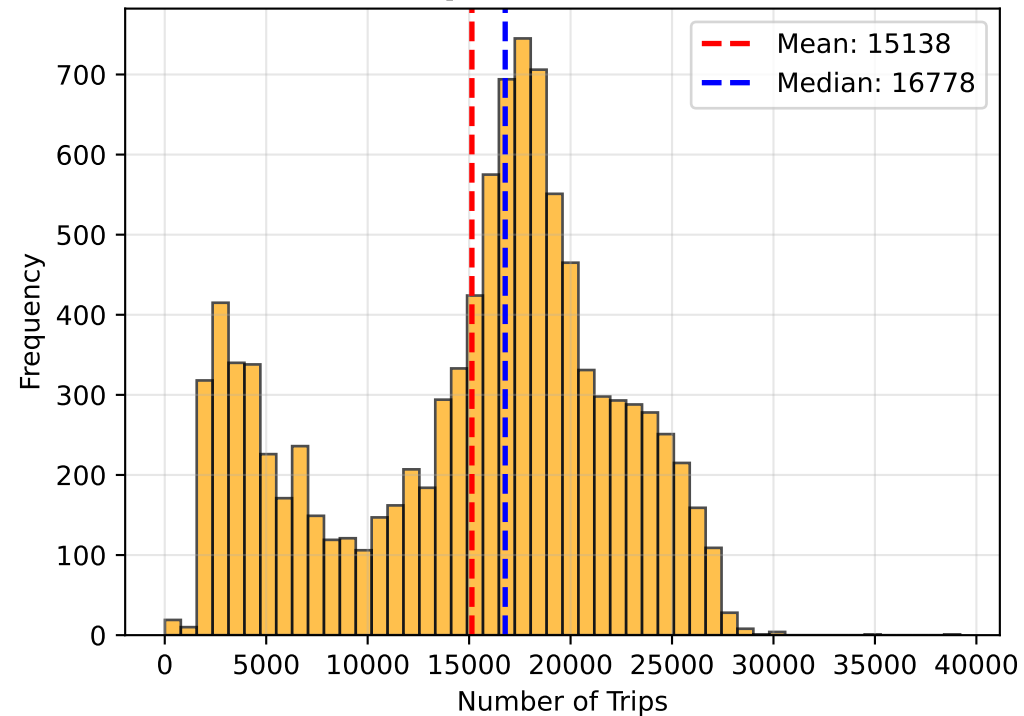
Demand Heatmap: Hour vs Day of Week



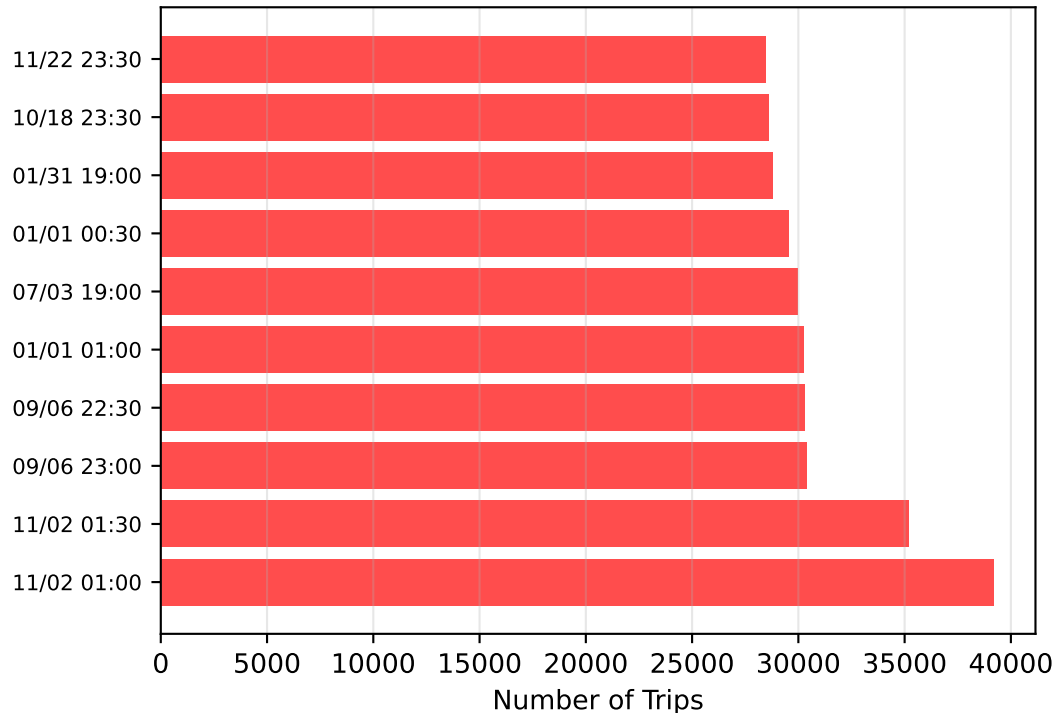
Monthly Trends



Trip Distribution



Top 10 Peak Periods



Summary Statistics

Statistic	Value
Total Trips	156,219,716
Average per 30min	15,138
Peak Period	November 02, 2014 01:00
Peak Trips	39,197
Coefficient of Variation	0.46
Weekday Average	14837
Weekend Average	15897

Key Insights from Exploratory Data Analysis

TEMPORAL PATTERNS DISCOVERED:

□ Daily Patterns:

- Peak Hour: 7:00 PM (22,892 avg trips)
- Rush Hours: 6-8 PM consistently highest demand
- Quiet Period: 3-6 AM (lowest demand)
- Rush Hour Ratio: 6.4x higher than minimum

□ Weekly Patterns:

- Busiest Day: Saturday (17,007 avg trips)
- Quietest Day: Monday (13,362 avg trips)
- Weekend Effect: 7.1% higher than weekdays
- Business Days: Steady increase Thu-Fri

□ Seasonal Trends:

- Peak Month: October (16,087 avg trips)
- Holiday Spikes: New Year's Eve/Day highest peaks
- Seasonal Variation: 11.7% difference

✂ Demand Characteristics:

- High Variability: $CV = 0.46$
- Strong Predictability: Clear daily and weekly cycles
- Peak Demand: 39,197 trips on November 02, 2014
- Business Impact: Predictable patterns enable optimization

□ MODELING IMPLICATIONS:

- Strong seasonality suggests seasonal models (SARIMA, Holt-Winters)
- Time-of-day features critical for ML models
- Lag features will be highly predictive
- Weekend/weekday effects should be captured
- Holiday effects need special handling