06.12.2024

# GG Taxi Analysis

**DS116 DATA VISUALIZATION**

**Final project**

**PRESENTED BY**

Artsvik Avetisyan
Ani Harutyunyan
Gayane Hovsepyan

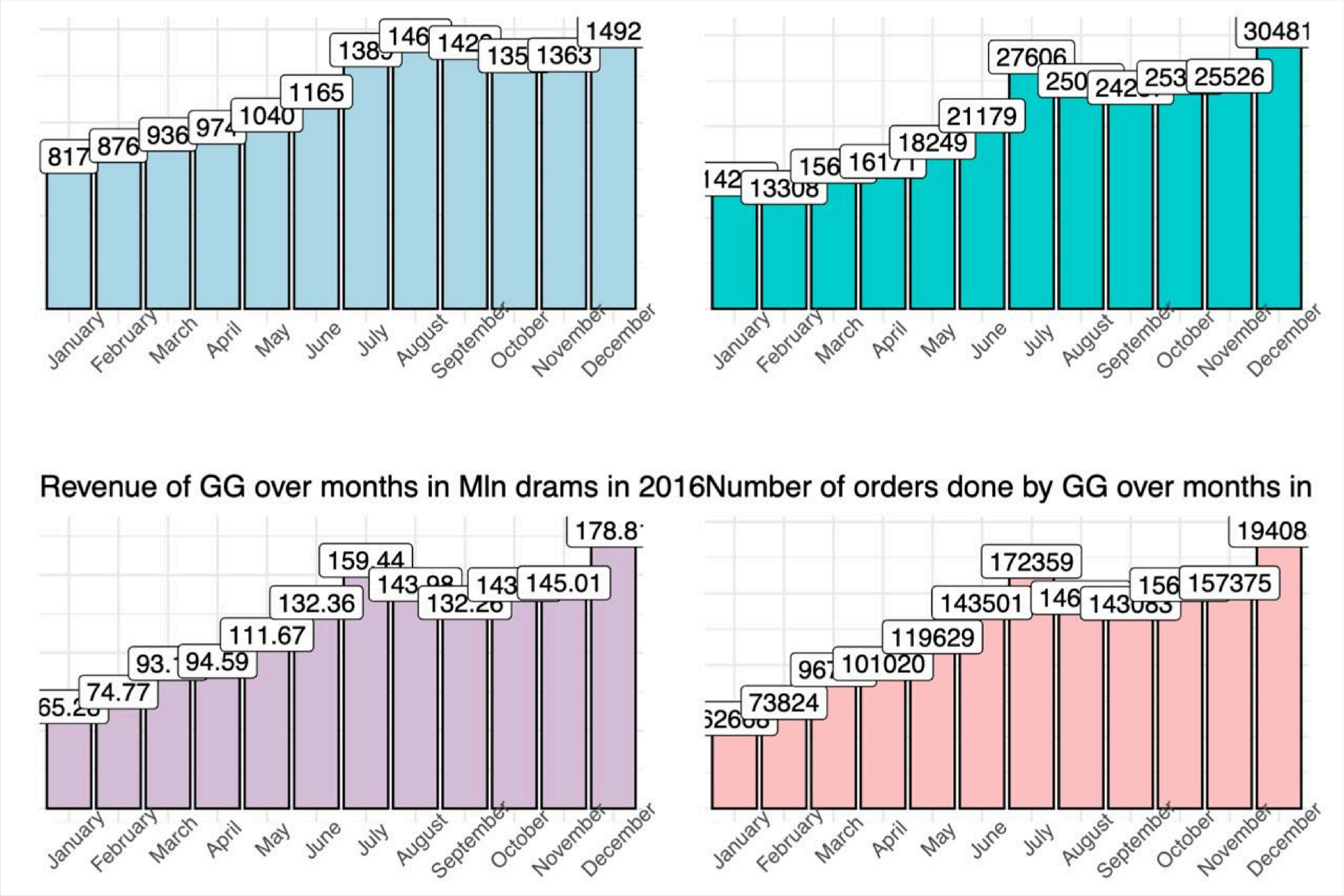
# Introduction

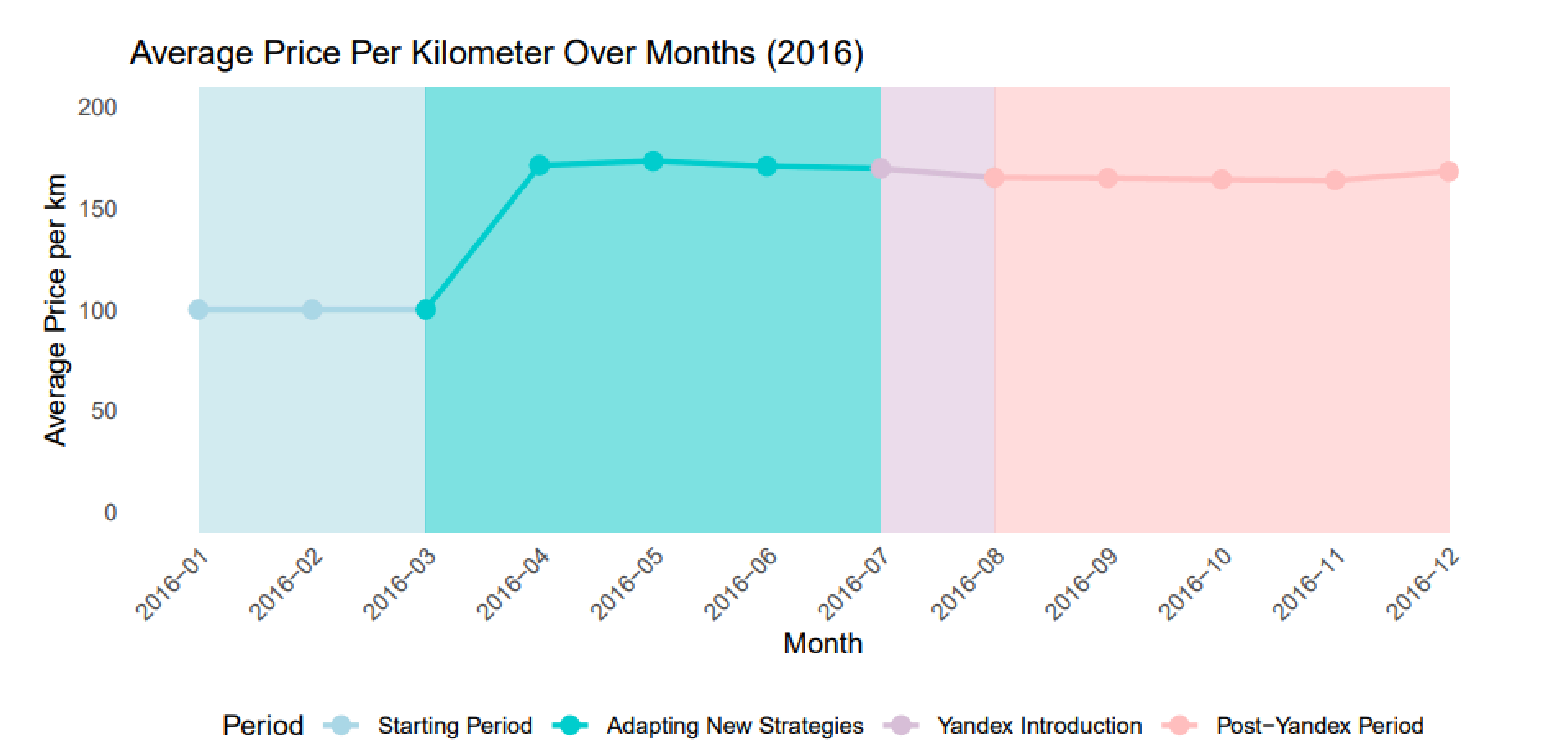
This project report analyzes data from GG Taxi to uncover patterns and trends that could help improve their service. We focused on how factors like time of day, weather conditions, and the start and end locations of rides influence service usage. This focused analysis not only highlights specific challenges and opportunities for GG Taxi but also contributes to the broader discussion on optimizing urban taxi services using data-driven strategies.

# General Findings

## Yandex Taxi

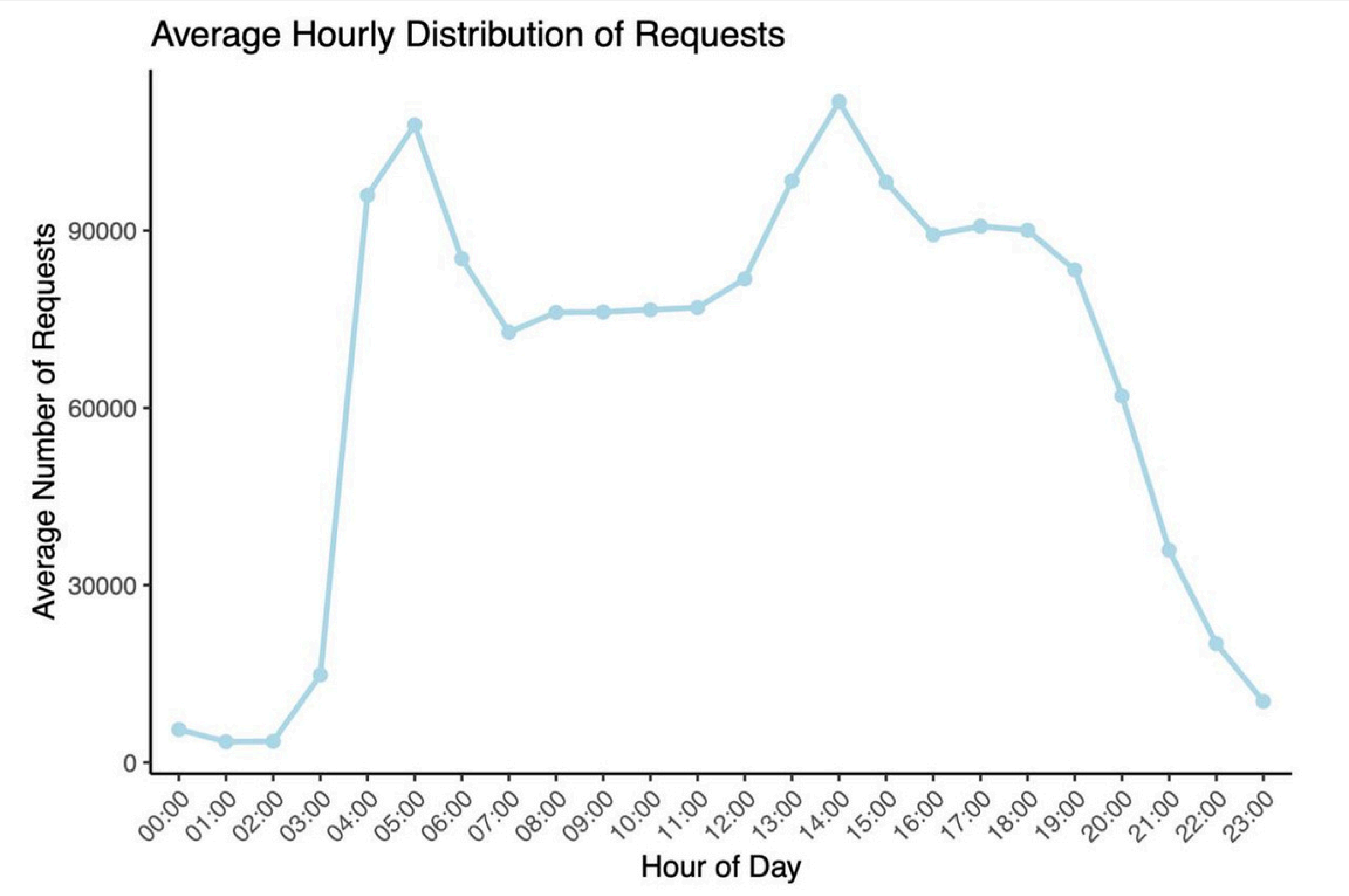
he graphs show that GG Taxi experienced significant peaks across all performance metrics — drivers, users, revenue, and completed orders — in July 2016, likely due to strategic efforts to counter the launch of Yandex Taxi in Yerevan. This rapid growth suggests aggressive marketing campaigns and driver recruitment initiatives to maintain customer loyalty and attract new users in response to heightened competition.



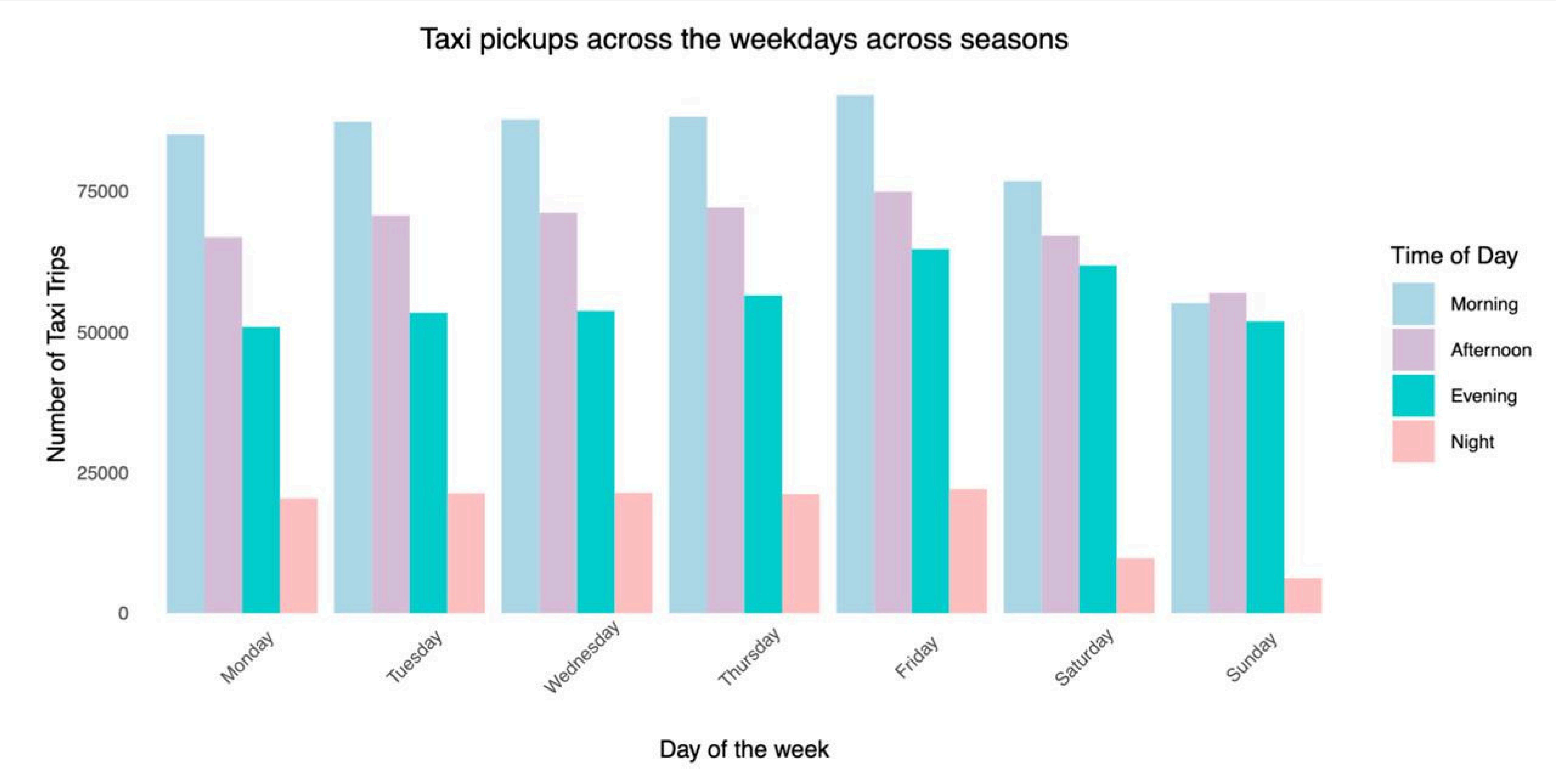



The graph illustrates GG Taxi's pricing policy during its early operations. Initially, prices remained constant at 100 AMD, reflecting traditional Armenian practices, but experienced a significant rise after March, possibly due to economic factors or competitive strategies. Following the entry of Yandex Taxi into the Armenian market in July, GG reduced its prices

# Daily and Hourly Trends



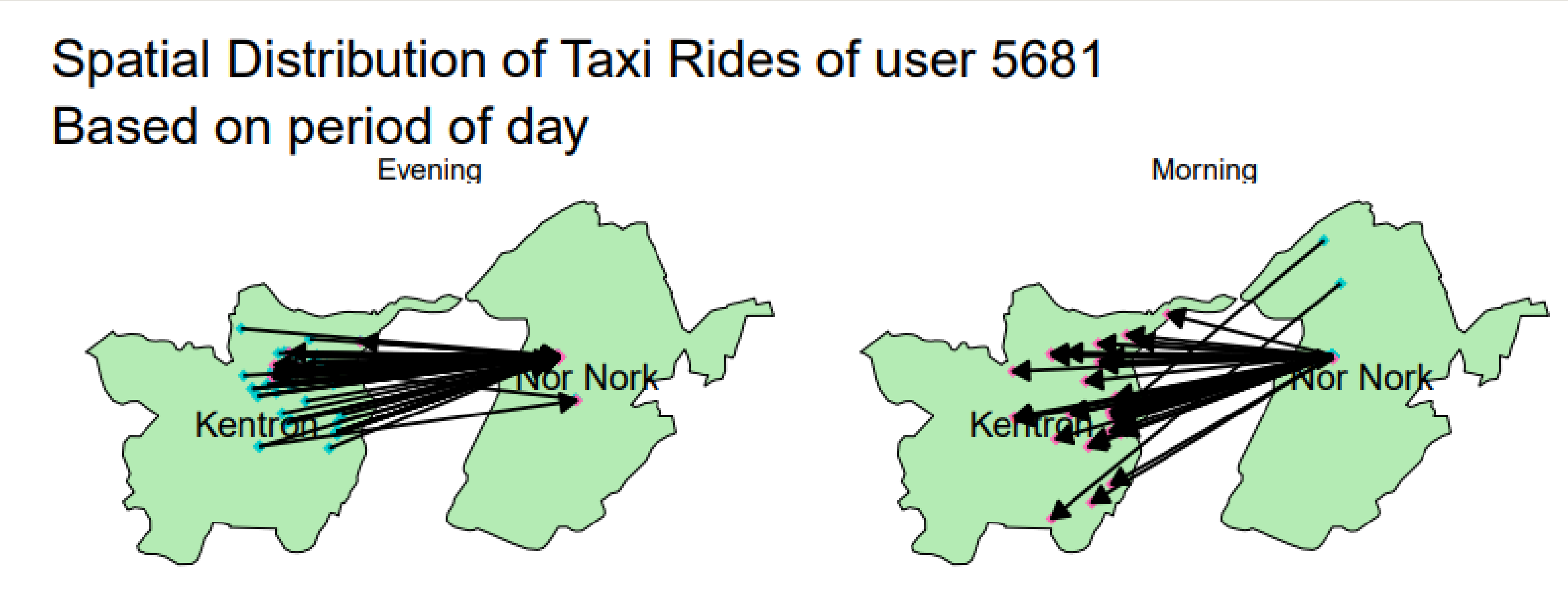

The graph depicts the hourly distribution of taxi requests in Yerevan, showing a sharp increase in demand during the early hours, peaking at 14:00, likely due to lunch-hour activities and school dismissals. After this mid-afternoon peak, demand steadily declines into the evening, dropping significantly after 20:00 as workdays end and late-night travel decreases.




This graph shows the distribution of taxi demand across days and times of the week. Weekday mornings consistently have the highest demand, likely driven by work and school commutes, while night trips are the lowest throughout the week. On weekends, demand patterns shift slightly, with nights seeing even fewer trips, reflecting reduced late-night travel activity.

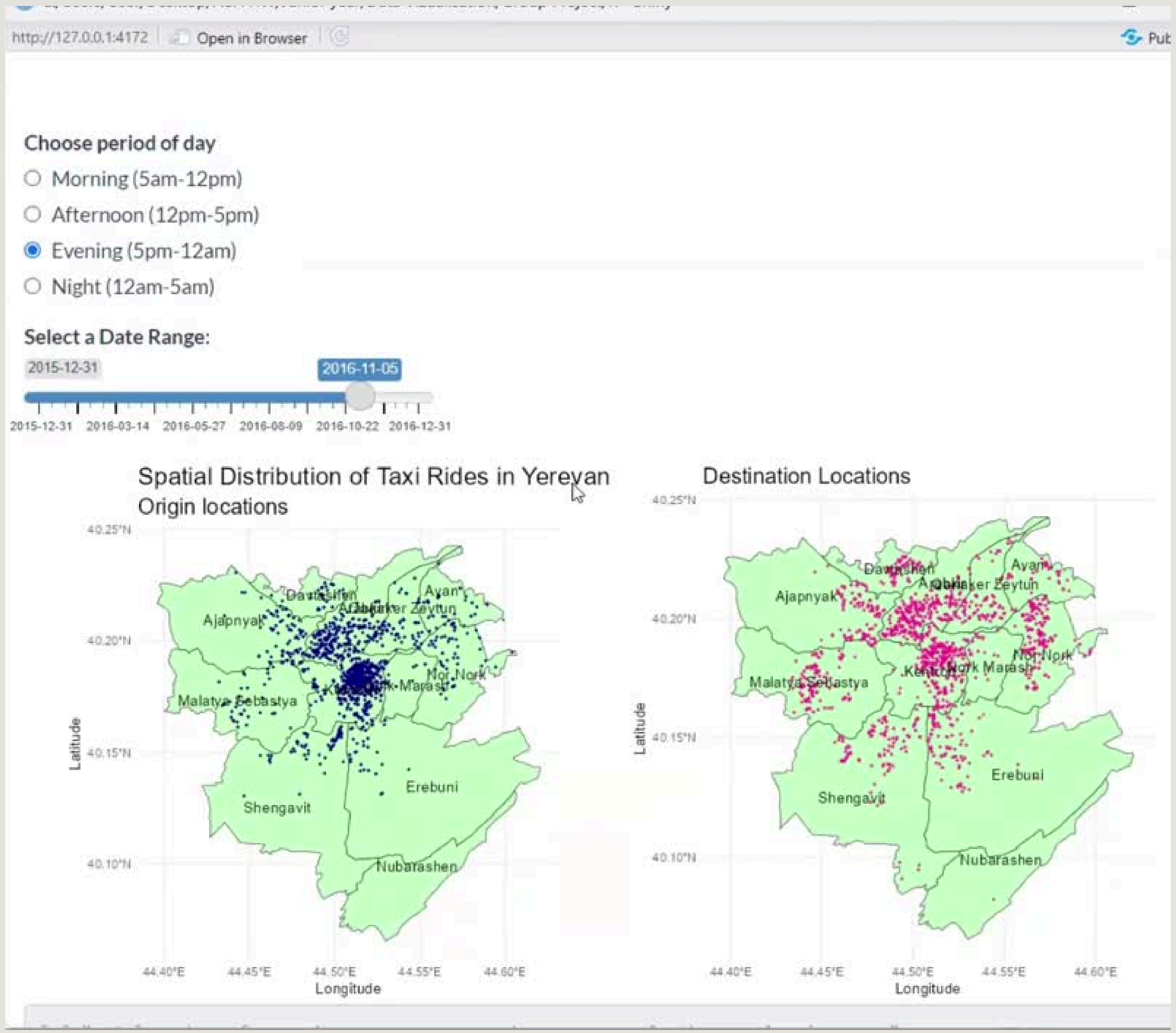
# Hypothesis

Kentron serves as a major hub for taxi activity.
Most taxi rides in Yerevan that originate from residential areas, with the city center (Kentron) being the primary destination in the morning, in the evening these rides predominantly return passengers to residential areas, suggesting a pattern of commuting behavior where the most common route is Home to Kentron and Kentron to Home.

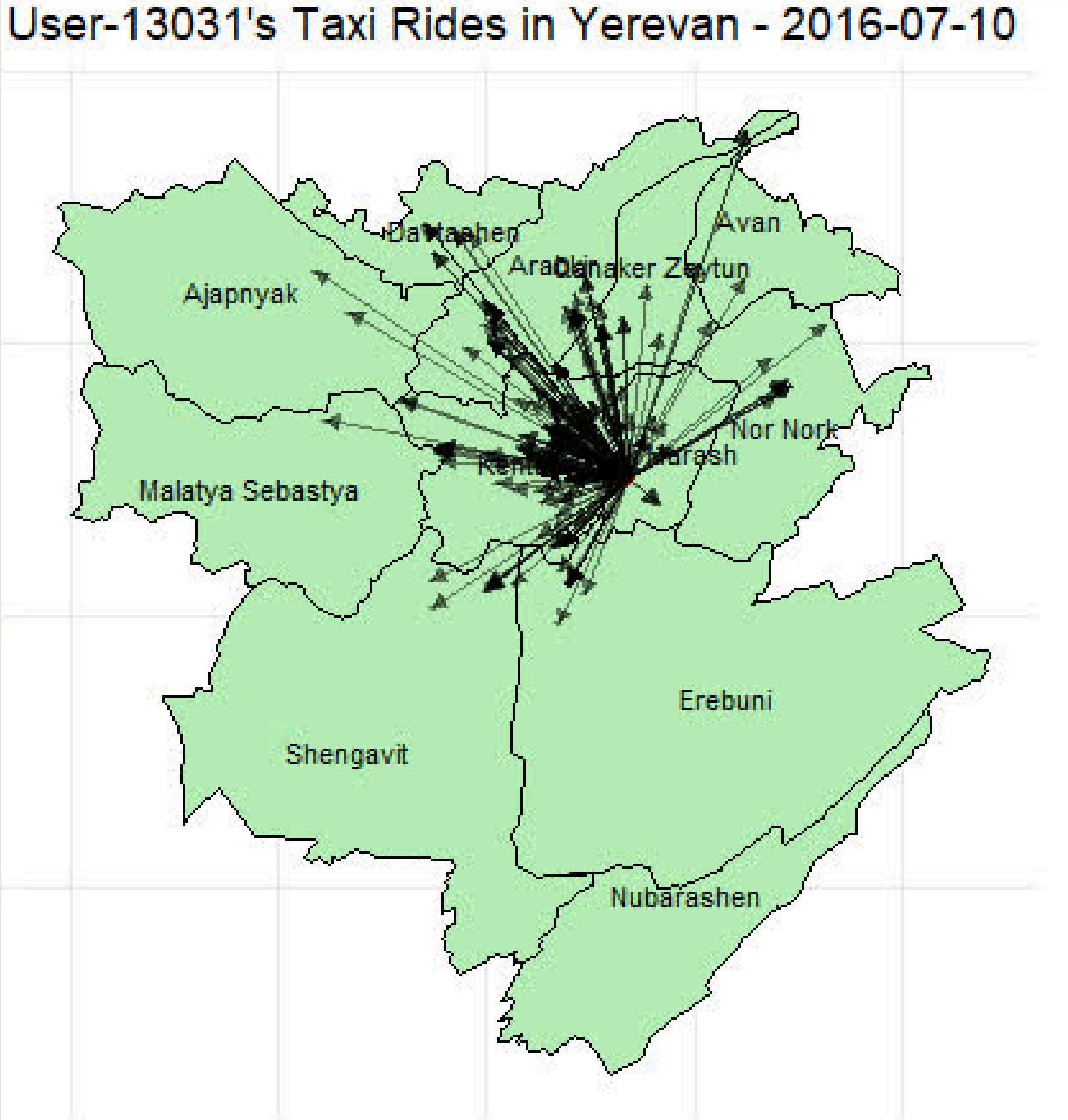



**Case Study: User 5681 shows a clear pattern where the user consistently travels from Nor Nork to Kentron in the morning and returns in the evening. The data reveals that all rides in Nor Nork cluster around the user's home, with most rides following the same home-to-Kentron and Kentron-to-home route.**

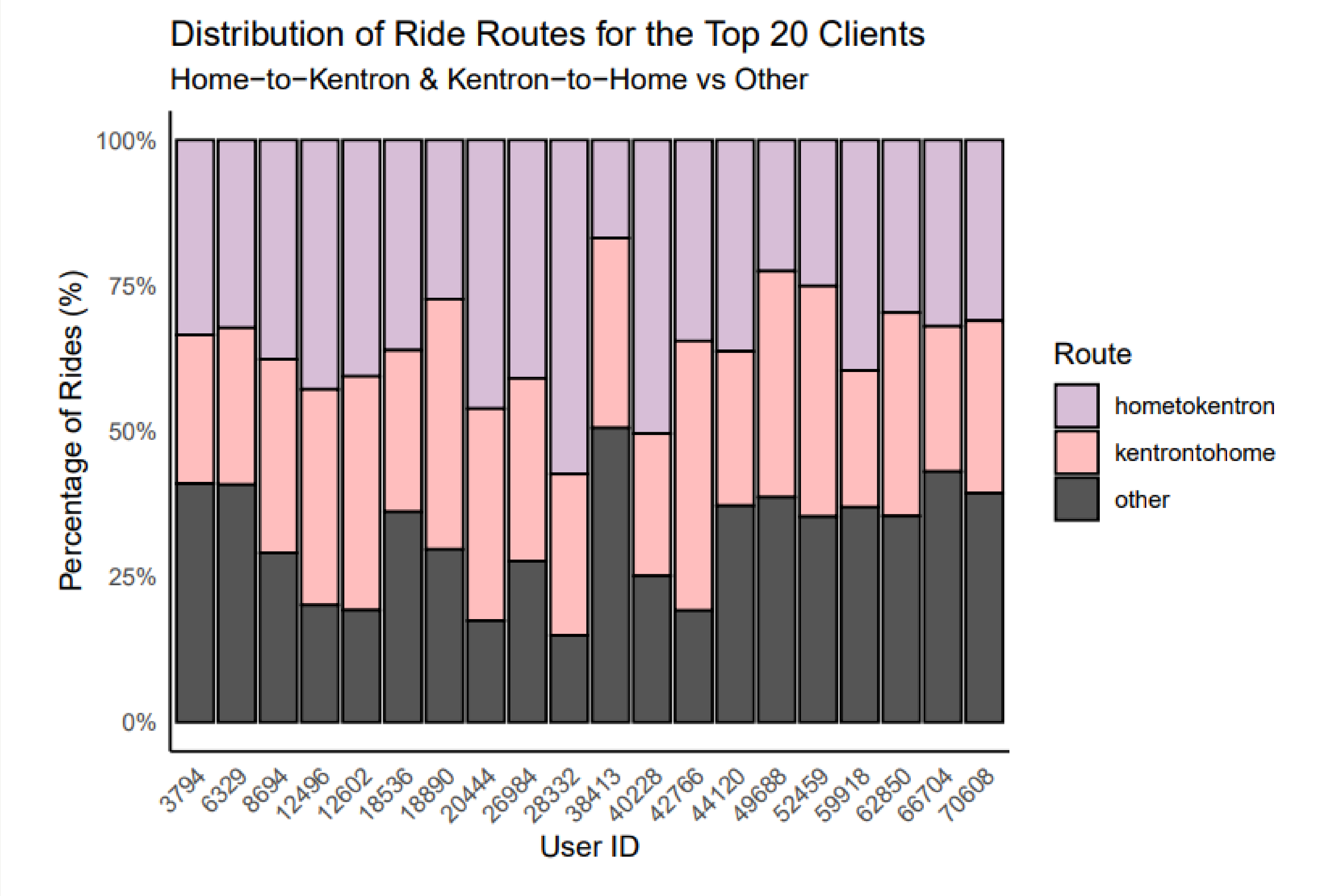
**An interactive UI visualization highlights Kentron's centrality in taxi movements, showing that most morning rides originate from other districts to Kentron, while evening rides reverse this pattern. The accompanying data table confirms that taxis converge in Kentron during the day and redistribute to other districts in the evening.**



# The Outlier: A Mystery User's Taxi Journey

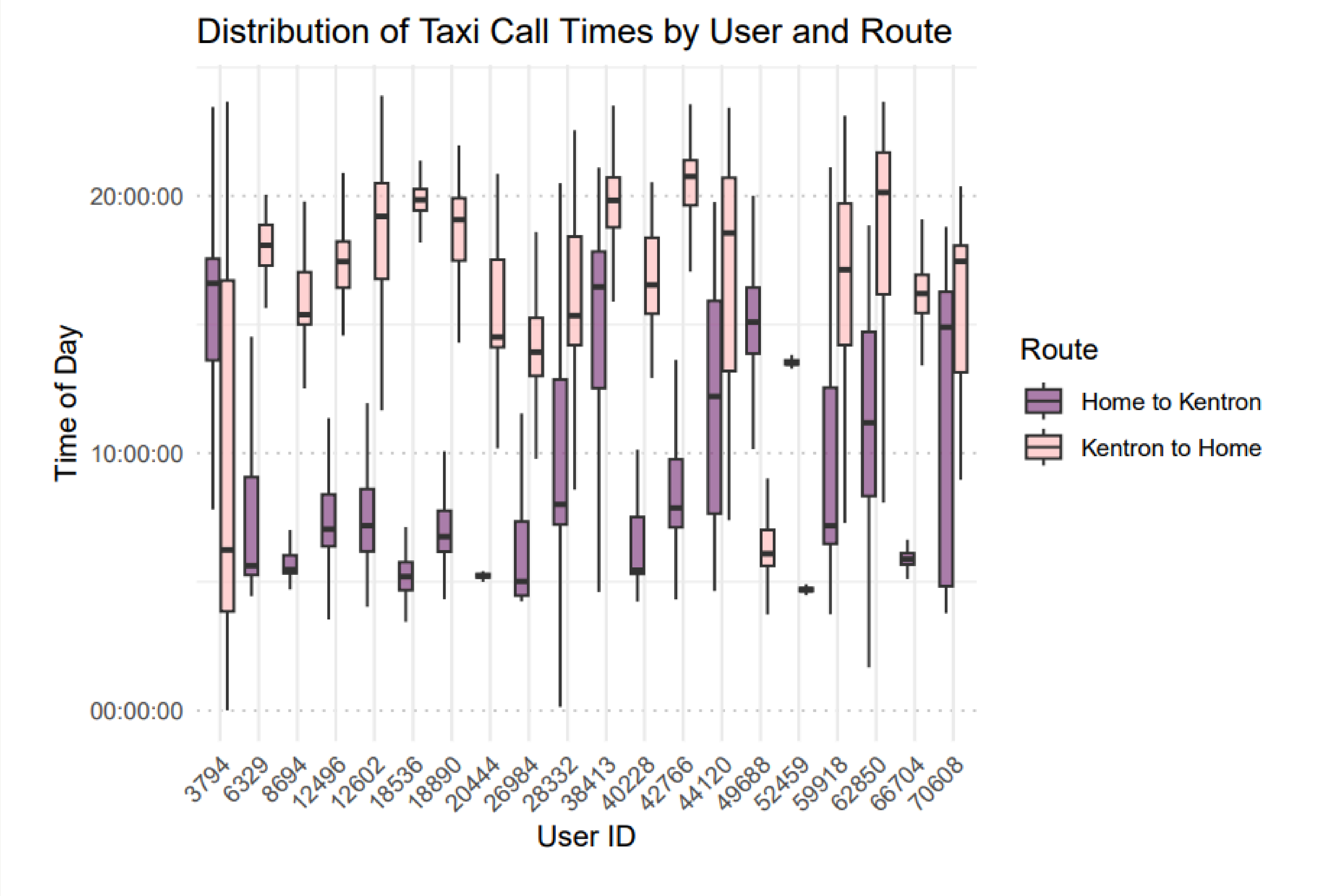



**A notable outlier was a user with 1,502 taxi orders in a year, averaging 4 rides daily, without any recurring destinations or return trips. This unique pattern suggested the user was likely using taxis for delivery activities, evidenced by diverse, non-repeating locations and supported by an animated visualization of their ride history.**

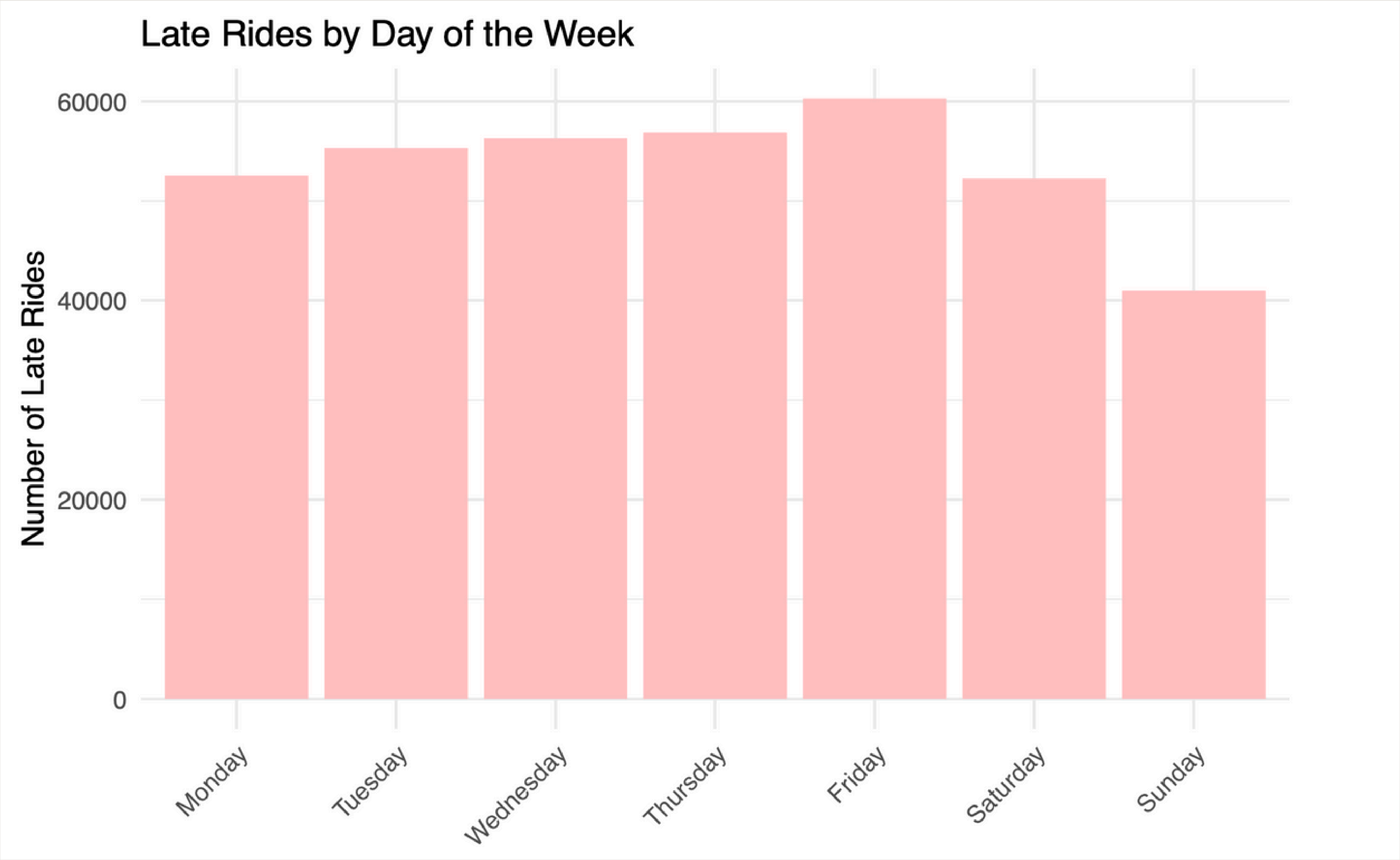



**The charts confirm that the majority of taxi rides are taken for the same purpose: users travel from their home to Kentron in the morning and return home in the evening. This pattern is evident in both the absolute values and the percentage breakdowns shown in the charts.**

**Boxplot visualizations show narrow ranges for morning rides (Home-to-Kentron), indicating consistent taxi usage around 7–8 AM, while evening rides (Kentron-to-Home) display wider variability, with most trips occurring between 5–7 PM.**



# Orders with Lateness Across Weekdays



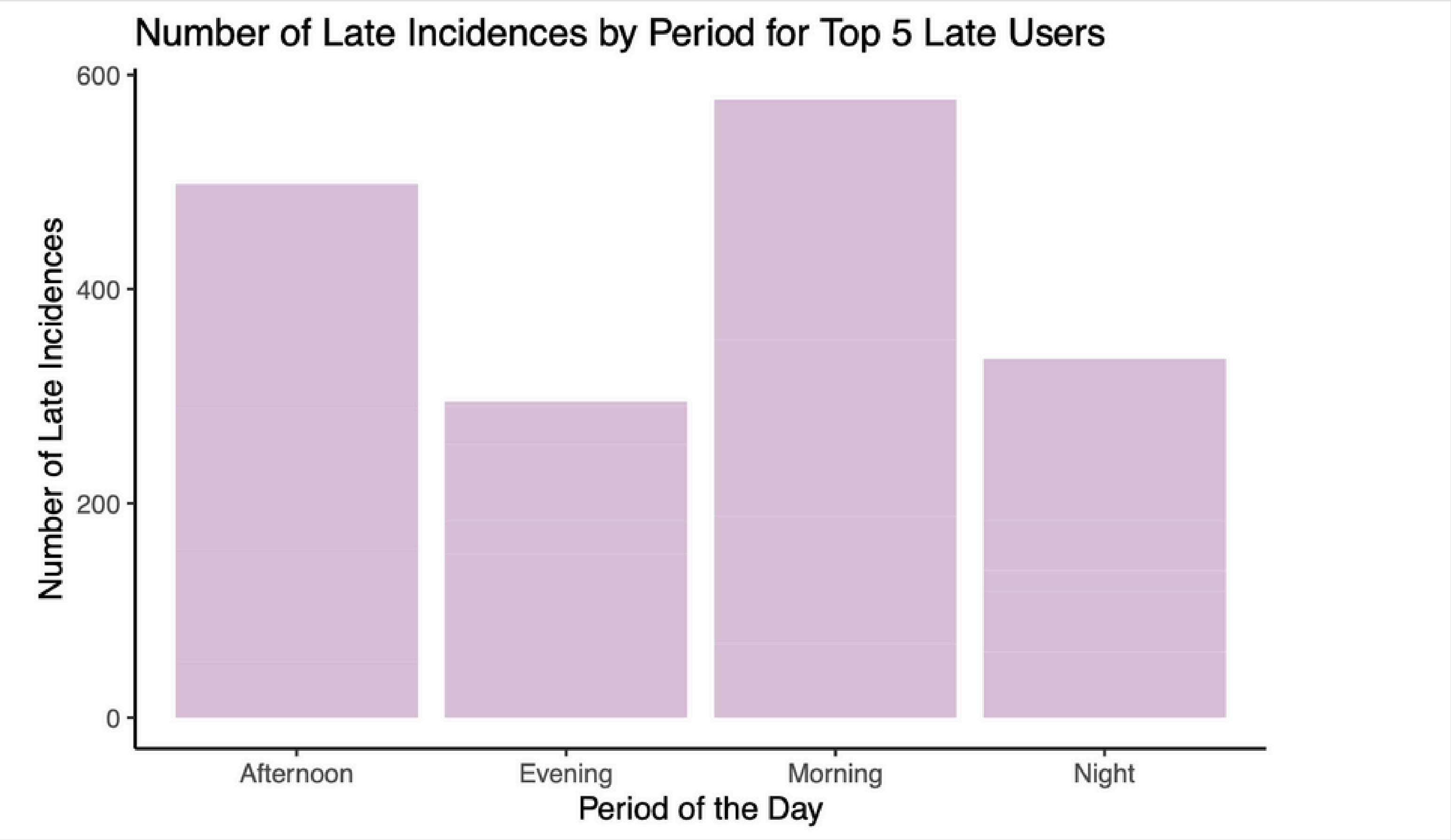

***HYPOTHESIS: "USERS TEND TO BE LATER AS THE WEEK PROGRESSES DUE TO CUMULATIVE FATIGUE, WITH A DROP IN LATE RIDES DURING THE WEEKEND DUE TO FEWER COMMITMENTS."***

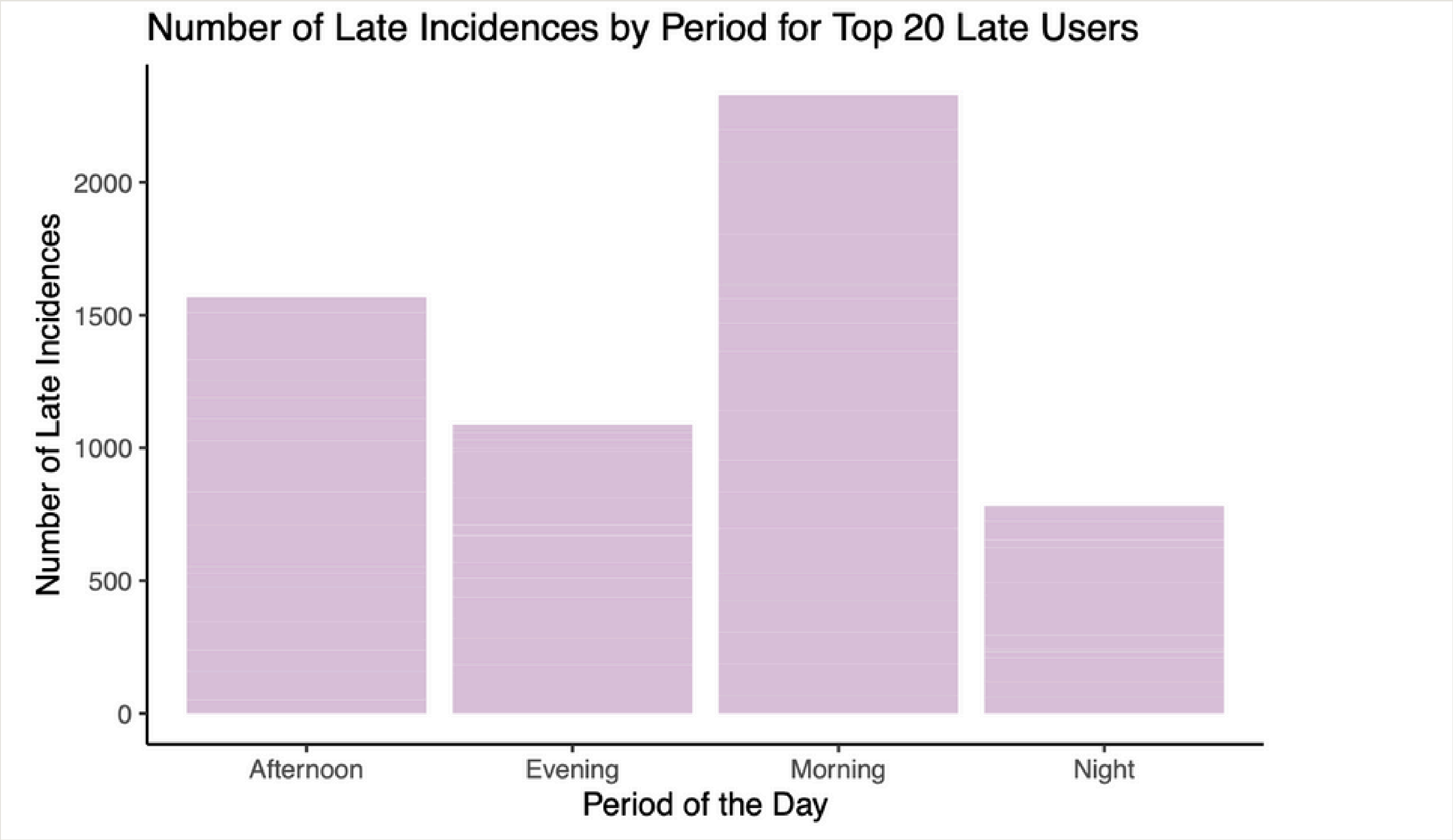
The hypothesis is accepted, with the analysis showing a clear pattern in late rides throughout the week. Late rides increase gradually from Monday to Friday, likely due to escalating work pressures, traffic, and fatigue. However, there is a sharp decline in late rides over the weekend, suggesting fewer work obligations and a more relaxed schedule. These insights highlight the importance of adjusting service delivery and resource allocation based on weekly patterns to improve punctuality and meet user needs.

# Number of Late Incidences by Period

***HYPOTHESIS: "LATENESS VARIES SIGNIFICANTLY ACROSS DIFFERENT TIMES OF THE DAY, ESPECIALLY IN THE MORNING, POTENTIALLY INFLUENCED BY THE MORNING RUSH OR OVERSLEEPING."***

The hypothesis is accepted. The analysis of late incidences by period, illustrated through two graphs representing the top 5 and top 20 late users, reveals significant daily timing patterns. The morning period consistently records the highest frequency of late rides across both user groups, reflecting challenges such as the morning rush or oversleeping. Among the top 5 late users, afternoon incidences are slightly lower than in the morning but still significant, while night incidences are higher than those in the evening. For the top 20 late users, the pattern is almost similar, with morning delays dominating, followed by the afternoon, while night and evening incidences change their places compared to the top 20 late users. This trend highlights the morning as a critical period for addressing lateness, driven by predictable factors like traffic congestion and individual preparation routines. Proposed interventions to mitigate lateness during this peak period include real-time traffic updates, flexible work or school start times, and encouraging habits that promote punctuality. Such strategies could address the unique challenges of the morning rush and improve timeliness across user groups.

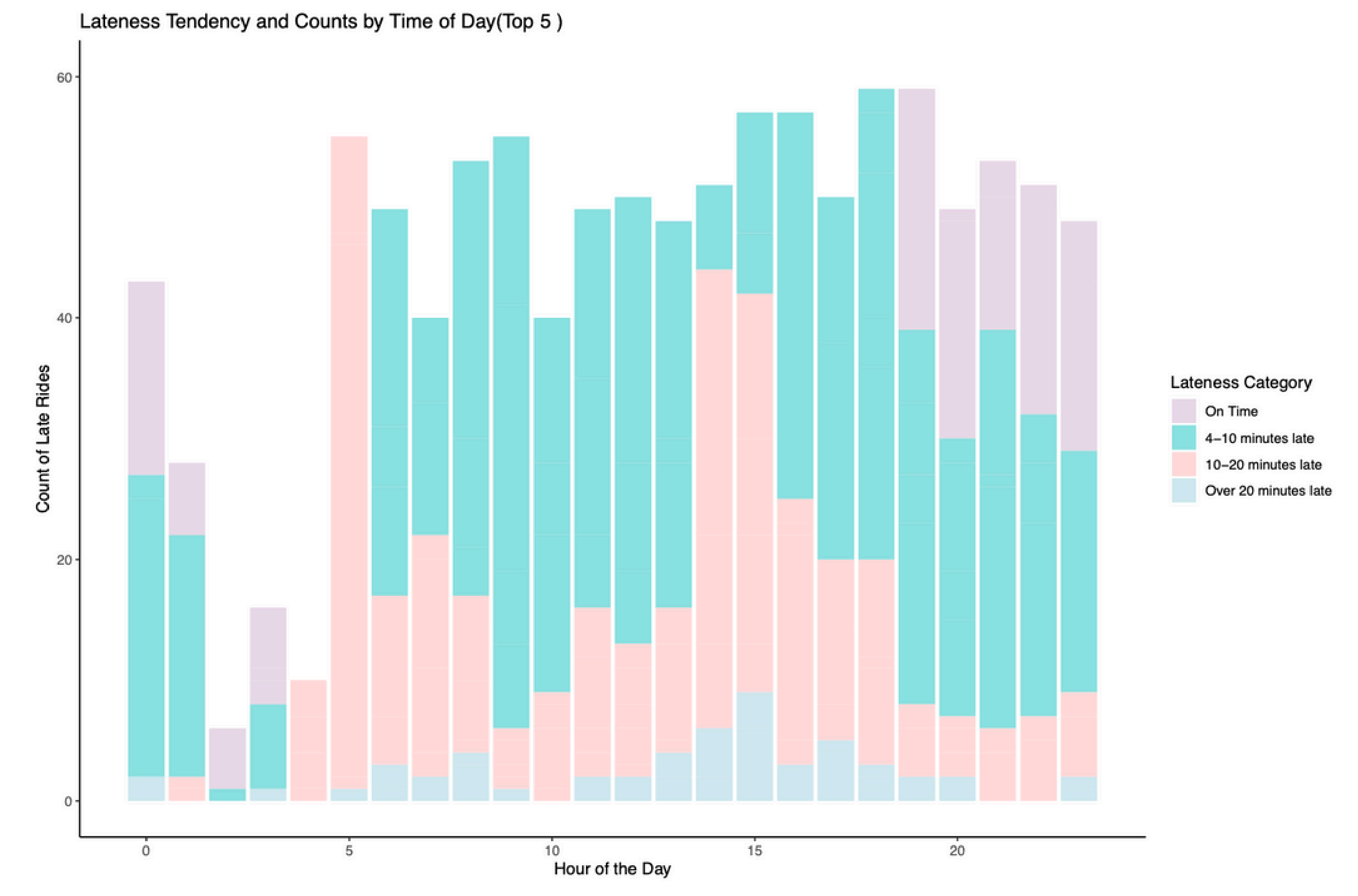


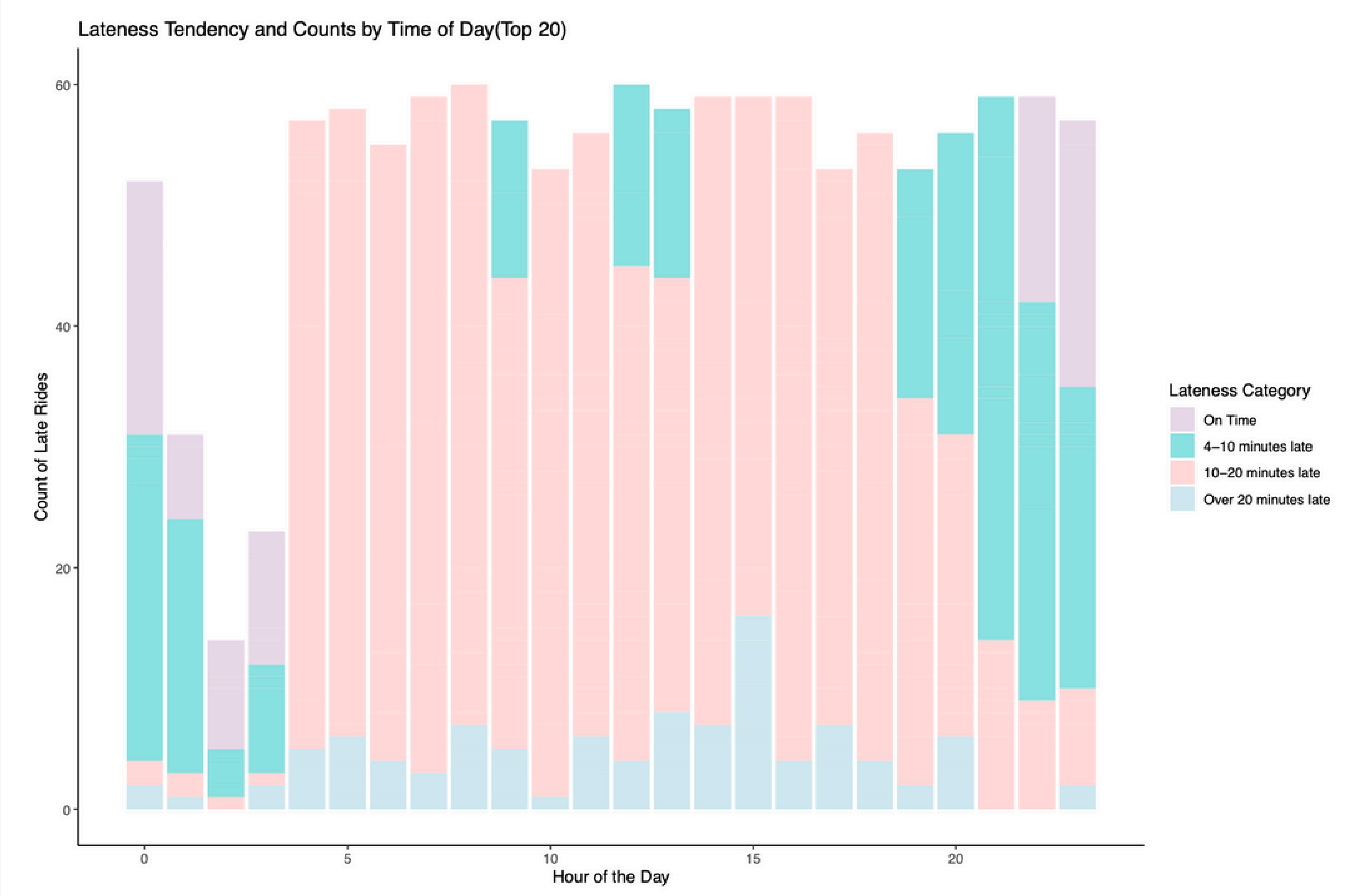



# Lateness Tendency

***HYPOTHESIS: "TOP 20 LATE USERS EXHIBIT MORE CONSISTENT LATENESS PATTERNS THROUGHOUT THE DAY COMPARED TO THE TOP 5 LATE USERS, WHOSE LATENESS PATTERNS VARY SIGNIFICANTLY DEPENDING ON THE TIME OF DAY."***

The hypothesis is accepted, with the analysis revealing distinct lateness patterns across the top 5 and top 20 late users. The top 20 users show consistent lateness, with most late rides falling within the 10-20 minute range, indicating habitual lateness less affected by external factors. In contrast, the top 5 users display more variability, with noticeable peaks at certain times, suggesting their lateness is influenced by external factors like work schedules and traffic. This comparison supports the hypothesis, highlighting that the top 20 users have more consistent lateness, while the top 5 users show greater fluctuations due to external influences.

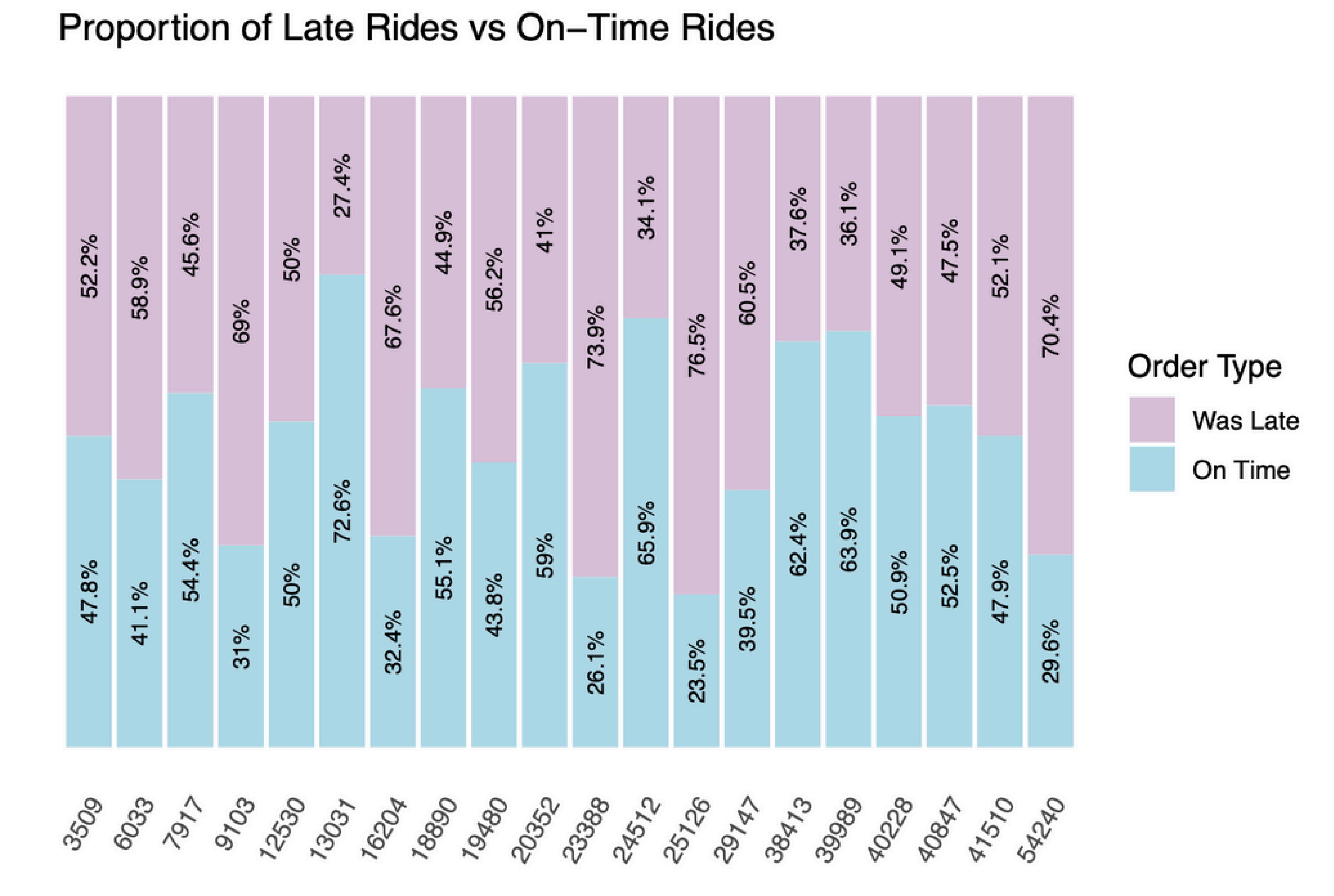


# Lateness as a Habit

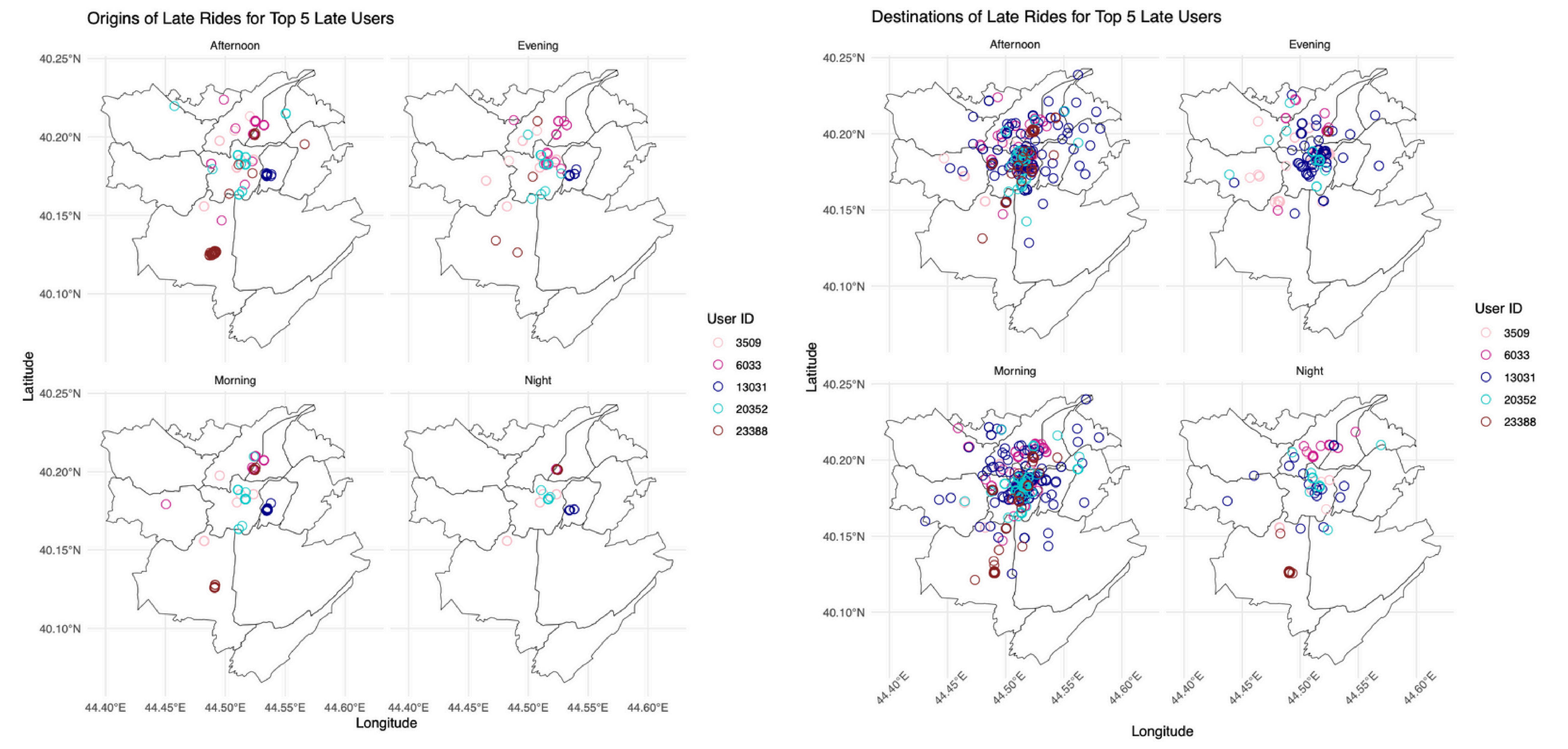
***HYPOTHESIS: "PEOPLE WHO ARE LATE MORE THAN 200 TIMES HAVE A HABITUAL TENDENCY TO BE LATE."***

12 out of the top 20 users showed a higher proportion of late rides, suggesting a consistent pattern of lateness. This indicates that punctuality is influenced by individual habits, while a smaller group shows sporadic lateness due to external factors. This distinction reinforces the hypothesis that habitual lateness is a defining trait for many users.



# Lateness as a Habit

***HYPOTHESIS: "IN THE MORNING, INDIVIDUALS CONSISTENTLY START FROM THE SAME LOCATION, WHILE THEIR DESTINATIONS VARY."***




The hypothesis is accepted. In the morning, users consistently depart from the same location, likely their homes, suggesting a fixed starting routine. However, their destinations are highly variable, indicating they don't commute to a single place but instead engage in various activities. This contrast highlights how unpredictable routines may contribute to delays, as varied schedules and destinations make timing more challenging, reinforcing the idea that their lateness is influenced by dynamic daily routines rather than external factors tied to a fixed destination.

Taxi

GG ~~Plot~~