Politecnico di Torino

Department of Electronics and Telecommunications

Department of Telecommunications, Electronics and Physics

ICT for SMART MOBILITY

Laboratory Report

Laboratory 1

Academic Year 2025/2026

Student:

Lorenzo Cavallaro - s346742
Giovanni Grieco - s346012
Negar Abdi - s343747

13th November 2025

Table of Contents

1 Step 1 – Preliminary Data Analysis

1.1 Overview of the datasets

Table 1 shows the number of documents present in each collection for the Car2Go and Enjoy car-sharing services.

Table 1: Number of documents in each collection for Car2Go and Enjoy.

Car-Sharing Service	**Collection**	**Number of Documents**
Car2Go	ActiveBookings	8,743
Car2Go	ActiveParkings	4,790
Car2Go	PermanentBookings	28,180,508
Car2Go	PermanentParkings	28,312,676
Enjoy	enjoy_ActiveBookings	0
Enjoy	enjoy_ActiveParkings	0
Enjoy	enjoy_PermanentBookings	6,653,472
Enjoy	enjoy_PermanentParkings	6,689,979

The nearly equal number of records in *PermanentParkings* and *PermanentBookings* is due to the one-to-one relationship between bookings and subsequent parking events. Therefore, analyses mainly focus on bookings to avoid redundancy. The small excess of parking records likely results from data inconsistencies or system errors.

Then, a cities coverage analysis has been performed and highlighted in Table 2, showing that Car2Go operate in various international cities, while Enjoy is used only in Italy.

Table 2: Cities included in each dataset.

Car-Sharing Service	**Cities**
Car2Go	Amsterdam, Austin, Berlin, Calgary, Columbus, Denver, Firenze, Frankfurt, Hamburg, Madrid, Milano, Montreal, Munchen, New York City, Portland, Rheinland, Roma, San Diego, Seattle, Stuttgart, Torino, Toronto, Twin Cities, Vancouver, Washington DC, Wien
Enjoy	Bologna, Catania, Firenze, Milano, Roma, Torino

Subsequently, Table 3 presents the results obtained from a temporal coverage study of data collections. Start and end times are showed in GMT +1, corresponding to the Turin timezone. It is important to note that the *ActiveParkings* and *ActiveBookings* collections do not include a *final_time* field, as these entries are still active. Moreover, an analysis on the data demonstrates that time has been stored according the GMT time, instead the date according to the local timezone of the city.

Table 3: Start and end times for each dataset collection.

Collection	**init_time**	**final_time**
ActiveBookings	2017-12-11 14:45:38	Not Available
ActiveParkings	2018-01-24 06:56:00	Not Available
PermanentBookings	2016-12-13 18:38:23	2018-01-31 14:13:03
PermanentParkings	2016-12-13 18:37:38	2018-01-31 14:13:03
enjoy_PermanentBookings	2017-05-05 16:06:21	2019-06-10 18:20:35
enjoy_PermanentParkings	2017-05-05 16:05:36	2019-06-10 18:20:35
enjoy_ActiveBookings	None	None
enjoy_ActiveParkings	None	None

1.2 Torino, Seattle, Madrid data analysis

An initial estimation of the total number of distinct license plates in the rental data was performed, revealing **Torino - Car2Go**: 609, **Torino – Enjoy**: 1,900, **Seattle – Car2Go**: 1,473, and **Madrid – Car2Go**: 475. The same analysis was then carried out for the week from 1st December to 7th December 2017, showing **Torino – Car2Go**: 411, **Torino – Enjoy**: 279, **Seattle – Car2Go**: 647, and **Madrid – Car2Go**: 405.

Additionally, a complete analysis of the number of rentals during the month of December was carried out for each city, and the results are shown in Table 4.

Table 4: Number of bookings recorded in December 2017.

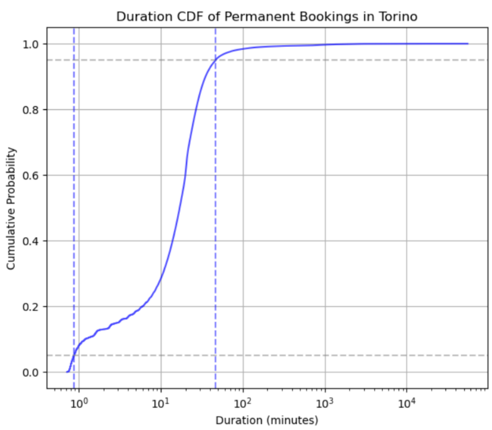
City	Car2Go Bookings (Dec 2017)	Enjoy Bookings (Dec 2017)
Torino	95,586	63,192
Seattle	83,077	0
Madrid	171,057	0

Finally, a study of alternative modes of transport for each city shows that the recording system is active only in Turin, with a total value of 568,402, while in the other cities the count is zero.

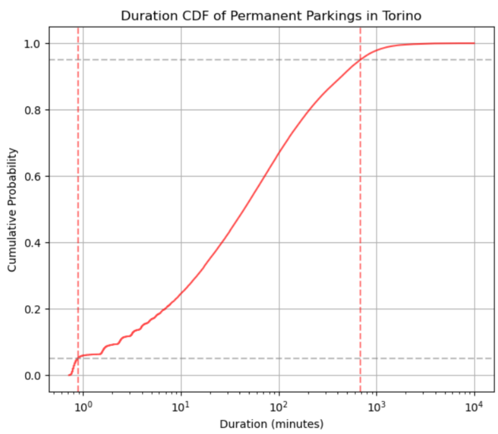
2 Step 2 – Car-Sharing Usage Characterization

2.1 Temporal and Statistical Analysis of Rental and Parking Usage

In this section, the analysis considers each city of the group, over the period between **November 1st, 2017** and **January 31st, 2018**.



(a) CDF for Booking durations.

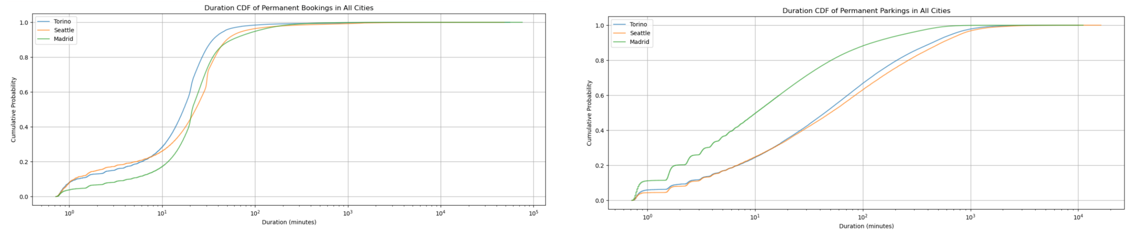


(b) CDF for Parking durations.

Figure 1: Cumulative Distribution Functions (CDFs) for durations in Turin.

As shown in Figure 1, the cumulative distributions of booking and parking durations display distinct behaviors. In the parking CDF (Figure 1b), the cumulative share of events grows gradually with duration, indicating that longer parking periods are relatively common. Conversely, in the booking CDF (Figure 1a), the curve rises steeply at shorter durations, meaning that most bookings are completed within a few minutes. Moreover, the 5th and 95th percentiles are highlighted in the plots, as they can be used to reasonably identify and remove outliers.

The comparison of CDFs across different cities is shown in Figure 2, where Figure 2a illustrates the trends in booking durations and Figure 2b shows those for parking durations.



(a) CDF for Booking durations across all cities.

(b) CDF for Parking durations across all cities.

Figure 2: Comparison of CDFs of booking and parking durations across all cities.

The data were then grouped by day to compute daily statistics, including the mean, standard deviation, median, and percentiles. Figure 3 shows the daily mean duration, highlighting peak rental periods during holidays, such as Christmas in Seattle and Turin, and Thanksgiving (23 November 2017) in Seattle.

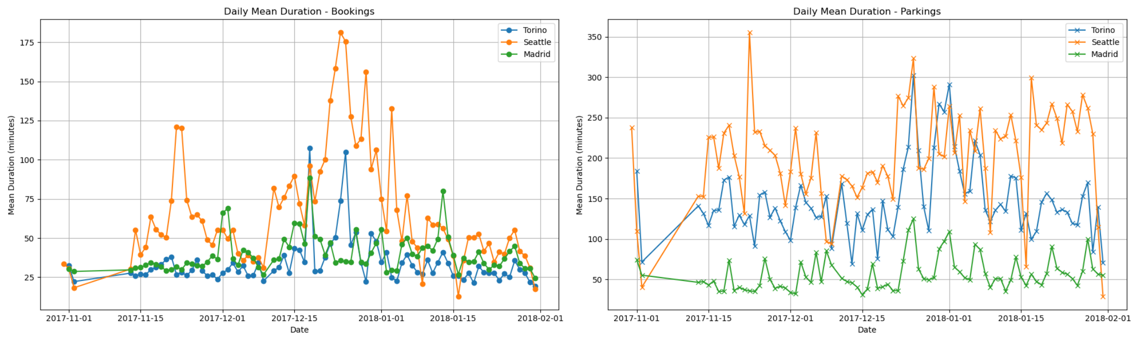
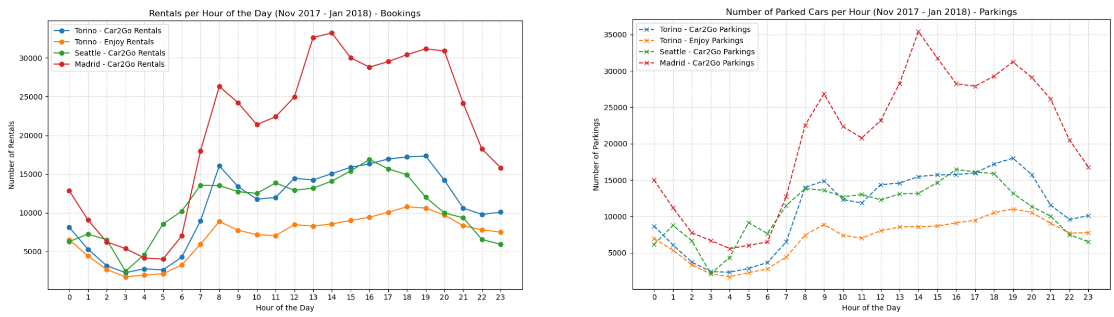


Figure 3: Daily Mean Durations for both parkings and bookings

Following the same approach, the data were aggregated by hour of the day to count the number of rentals occurring in each hour, producing an hourly usage trend for each service in each city. Figure 4 highlights peak utilization hours at around 8 a.m., 1–2 p.m., and 7 p.m. Madrid is the city with the highest overall service usage.



(a) Hourly distribution of bookings.

(b) Hourly distribution of parkings.

Figure 4: Hourly usage trends for bookings and parkings across cities, highlighting peak hours.

2.2 Outlier filtering

To remove outliers, the initial experiment used the 5th and 95th percentiles as thresholds. However, this resulted in extremely low and high cutoff values that were considered unrealistic for the system.

The implemented solution uses static thresholds, common to all cities, determined based on real-world reasonability and the statistical properties of the data. Specifically, bookings lasting less than **2 minutes** or more than **60 minutes** and parkings lasting less than **1 minute** or more than **400 minutes** were filtered out.

This decision was guided by analyzing histograms of the duration distributions (Figure 5) and by examining the percentage of data removed using the chosen thresholds (Table **??**).

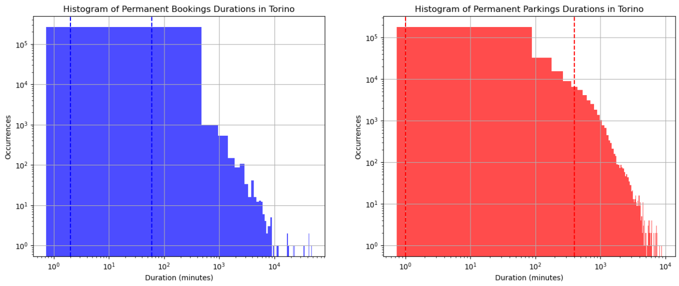


Figure 5: Histograms of booking and parking durations across all cities, used to identify reasonable thresholds for outlier filtering.

City	Bookings (%)			Parkings (%)		
	In range	**Below range**	**Above range**	**In range**	**Below range**	**Above range**
Torino	84.04	12.32	3.64	81.38	8.57	10.06
Seattle	77.19	15.08	7.73	80.93	4.30	14.77
Madrid	84.50	5.11	10.39	86.60	10.65	2.75

Table 5: Percentage of booking and parking durations within, below, and above the defined thresholds for outlier filtering.

Filtering outliers provides a clearer view of typical booking activity by removing extreme values that could distort the analysis. The resulting filtered CDFs (Figure 6) are smoother, highlighting ordinary usage patterns across Turin, Seattle, and Madrid. Most bookings occur within a 15–30 minute window, with similar trends across cities. In contrast, the parking CDF rises sharply in the 1–50 minute interval, demonstrating greater variability in parking durations compared to bookings. Moreover, to demonstrates the difference between outliers filtered data and not filtered data, Figure 7 illustrates this comparison in the city of Turin.

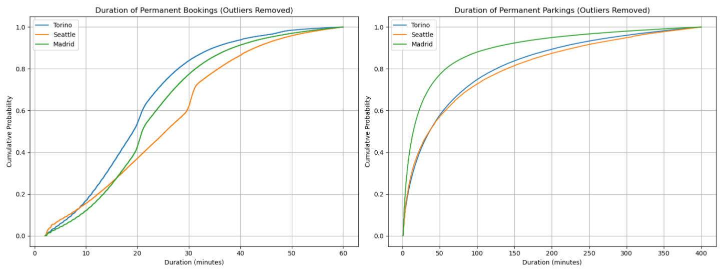


Figure 6: Filtered Cumulative Distribution Functions (CDFs) for booking (left) and parking (right) durations.

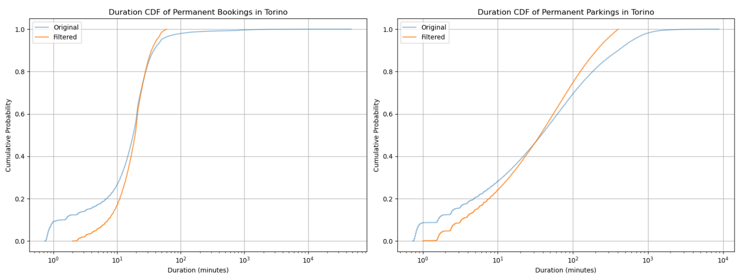


Figure 7: Comparison between Filtered CDFs and Not Filtered CDFs for booking (left) and parking (right) durations.

An additional analysis was performed to assess how daily statistics are affected by the outlier filtering process. Figure 8 shows the daily median along with the 25th and 75th percentiles for unfiltered bookings (left) and parkings (right), highlighting the influence of extreme values. Figure 9 presents the same statistics after filtering, showing smoother trends and reduced variability, which better reflect typical daily usage patterns.

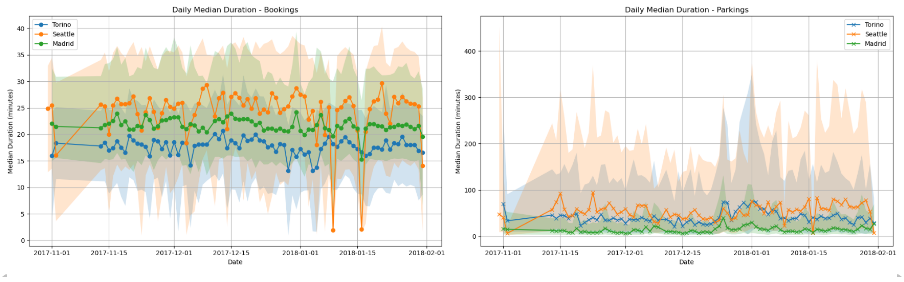


Figure 8: Daily Median, 25th and 75th Percentiles for not filtered bookings (left) and parkings (right)

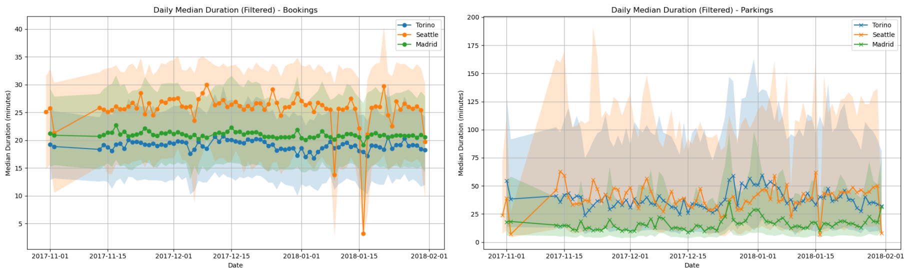


Figure 9: Daily Median, 25th and 75th Percentiles for filtered bookings (left) and parkings (right)

2.3 Effect of Public Transport Duration on Rental Activity

From Figure 10, the number of rentals appears highest when the public transport duration is around 15 minutes. The trend shows that the number of bookings increases rapidly up to this peak and then gradually decreases as public transport durations become longer. This pattern indicates that users are most likely to choose the rental service when public transport trips take approximately 15 minutes, while both very short and very long public transport durations correspond to fewer bookings.

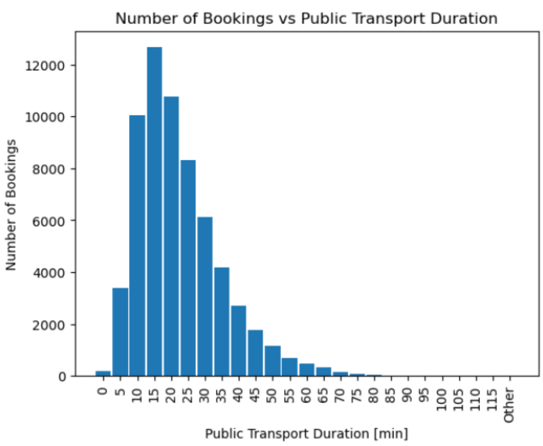


Figure 10: Number of bookings as a function of public transport duration bins.

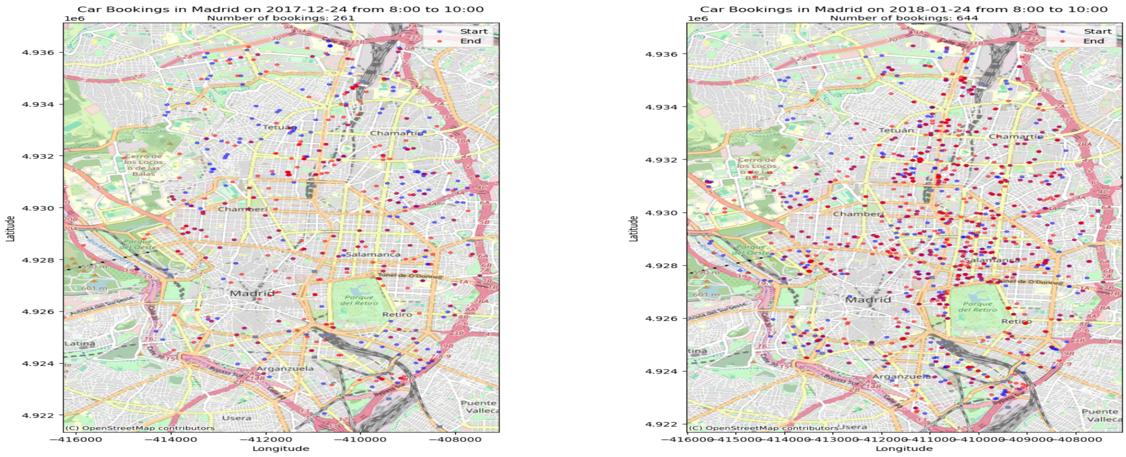
2.4 Spatial analysis of parking origins and destinations

To analyze the spatial distribution of cars in Madrid at the times of rental pickup and drop-off, the origin and destination positions of cars were plotted for different hours of the day to examine how rental patterns vary over time and across different days.

Figure 11 presents a comparison of rental activity for two mornings: December 24, 2017 (Christmas holiday) and January 24, 2018 (regular weekday). Each subfigure shows both the pickup (origin) and drop-off (destination) locations of rentals. On the holiday morning, there were 261 bookings, with pickups dispersed across residential and leisure areas and drop-offs concentrated in shopping or tourist zones. On the regular weekday morning, there were 644 bookings, with pickups primarily in residential neighborhoods and drop-offs clustered in business districts, reflecting typical commuting patterns.

Figure 12 complements this analysis by showing the area-based distribution of start and end points using polygons to indicate the probability density of rentals within each area. This provides a clearer view of the zones with the highest likelihood of rentals beginning or ending, allowing for a more detailed understanding of spatial demand patterns across the city.

Together, these visualizations highlight how rental demand in Madrid is both spatially and temporally dependent, and how special days such as holidays can significantly influence patterns. Insights from these analyses can inform fleet redistribution strategies to improve service availability in high-demand areas during different times and events.



(a) December 24, 2017 — 8:00–10:00

(b) January 24, 2018 — 8:00–10:00

Figure 11: Comparison of rental activity at different dates.

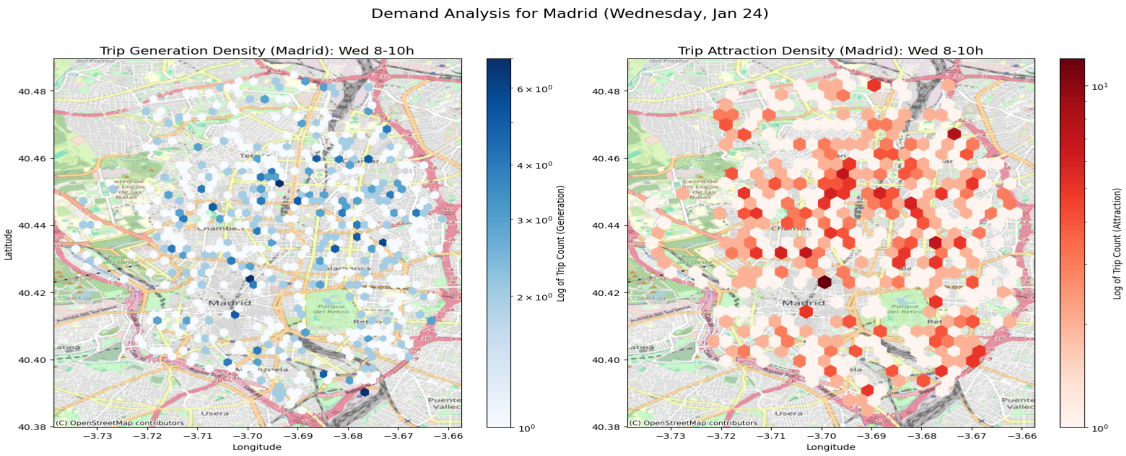


Figure 12: Area-based distribution of rental start (left) and end points (right).