DOI: https://doi.org/10.30525/2256-0742/2025-11-5-126-138

MODELLING OPERATIONAL ARCHETYPES FOR CORPORATE FLEET ELECTRIFICATION

Aivars Rubenis[1], Jelena Tonova[2], Vadims Morozovs[3]

Abstract. The decarbonisation of corporate vehicle fleets is a central challenge in achieving Europe's climate neutrality targets under the European Green Deal. Although corporate vehicles constitute only a share of the total fleet, they account for a disproportionate fraction of total mileage and associated CO_2 emissions. Despite fiscal incentives and regulatory support, the adoption of battery electric vehicles (BEVs) within the corporate sector remains significantly below private uptake, primarily due to uncertainty about operational feasibility and charging constraints. This study presents a data-driven framework for assessing fleet electrification potential based on empirical driving data and simulation-based modelling. Using vehicle usage records from the carmonitor.eu telematics platform, the research identifies four distinct operational archetypes within corporate fleets, differentiated by travel intensity, trip fragmentation, and temporal driving structure. These archetypes are derived through a clustering methodology employing standardised behavioural indicators, principal component analysis (PCA), and k-means segmentation, validated by silhouette and Davies–Bouldin indices. Results demonstrate pronounced heterogeneity in fleet operation, with daily driving distances, trip frequency, and vehicle availability varying substantially across clusters. Scenario-based modelling reveals that electrification feasibility depends not only on total mileage but also on temporal driving regularity and charging opportunity windows. Vehicles with predictable daily cycles and long overnight parking are found to be most suitable for early electrification, while high-mileage or irregular-use vehicles require access to fast-charging infrastructure and larger battery capacities. The study concludes that segmenting corporate fleets by operational archetype provides a robust analytical foundation for transition planning, enabling tailored recommendations for vehicle selection, charging infrastructure, and total cost of ownership optimisation. By linking empirical usage data with simulation and scenario modelling, the paper contributes a replicable methodological approach for evidence-based fleet decarbonisation strategies across Europe.

Keywords: Electric Vehicles (EV), Fleet electrification, Data-driven decision-making, Corporate sustainability and mobility, Car Fleet Operational Archetypes.

JEL Classification: Q01, O18

Nomenclature	
AC	Alternating Current
AIC	Akaike Information Criterion
BEV	Battery Electric Vehicle
BIC	Bayesian Information Criterion
CO_2	Carbon Dioxide
CV	Coefficient of Variation
DC	Direct Current
EM	Expectation–Maximisation
EU	European Union
EEA	European Environment Agency
EV	Electric Vehicle
GHG	Greenhouse Gases
GMM	Gaussian Mixture Model
ICE	Internal Combustion Engine
IQR	Interquartile Range
JEL	Journal of Economic Literature
PCA	Principal Component Analysis
SD	Standard Deviation
TCO	Total Cost of Ownership

[1] Latvia University of Life Sciences and Technologies, Latvia *(corresponding author)*
E-mail: aivars.rubenis@ivorygroup.eu
ORCID: https://orcid.org/0000-0001-8765-7790
[2] Vidzeme University of Applied Sciences, Latvia
ISMA University of Applied Sciences, Latvia
ORCID: https://orcid.org/0009-0002-3770-387X
[3] Vozorom SIA, Latvia

1. Introduction

Efforts to decarbonise transport have accelerated as governments and industries seek to align with global climate targets. In Europe, the European Climate Law establishes a legal commitment to achieve climate neutrality by 2050 as part of the European Green Deal (Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 Establishing the Framework for Achieving Climate Neutrality and Amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law'), 2021), identifying road transport as a key sector requiring rapid emission reductions. While most sectors have reduced greenhouse gas (GHG) emissions over recent decades, transport remains an outlier – responsible for about 25% of total EU emissions in 2022 and still rising (European Environment Agency, 2024).

Corporate vehicles, despite benefiting from fiscal incentives such as tax exemptions in many EU Member States (Transport & Environment, n.d.), have contributed relatively little to emission reduction efforts. They represent around 60% of all new car registrations in the EU (Héliot, & Ferrara, 2025) and, because they are driven roughly twice as much as private cars, account for 74% of new vehicle CO_2 emissions (Cornelis & Antich, 2023). Yet, battery-electric vehicle (BEV) uptake in this segment remains modest. In 2024, BEVs made up only 16.3% of new corporate registrations in Germany versus 25.6% among private users; in France the figures were 12.0% and 22.1%, respectively; and in Denmark, 26.1% compared to 53.1% (Antich, Arnau Oliver, 2024). The European Commission notes that electrifying these high-mileage fleets could achieve substantial emission reductions within a short timeframe (Decarbonise Corporate Fleets, 2025).

Corporate fleets are increasingly recognised as pivotal to accelerating the EV transition. Their shorter replacement cycles supply the second-hand market with affordable used EVs, thus expanding access for private consumers. Estimates suggest that full corporate fleet electrification could add nearly seven million used EVs to Europe's market by 2035 (Cornelis & Antich, 2025).

Previous research on private vehicle electrification has explored financial factors such as vehicle cost (Coffman et al., 2017), charging infrastructure availability (Haustein et al., 2021), and incentives – including in Latvia (Rubenis et al., 2019) – as well as non-financial influences such as social norms, environmental awareness, and aesthetics (Krishnan & Koshy, 2021), without going into details of various subsections of the fleets and suitability of EVs for all of those.

This article aims to facilitate the transition of corporate fleets from internal combustion engine (ICE) vehicles to EVs, particularly looking at how to evaluate the suitability of an EV as a direct replacement for an existing ICE vehicle. This concept builds on the framework introduced in "The Road to Zero-Emission Fleets: The Role of Data-Driven Decision-Making," (Rubenis et al., 2025) which outlined the methodological basis for integrating empirical fleet monitoring, simulation modelling, and decision-support analytics into corporate fleet transition planning.

Even though for corporate fleets adoption is primarily determined by economic feasibility, typically assessed through Total Cost of Ownership (TCO) models (Al-Alawi & Bradley, 2013) and further influenced by fiscal measures and incentives (Di Foggia, 2021), first of all, understanding how these corporate vehicles are used, is essential for designing effective electrification strategies.

Real-world fleet operations are inherently heterogeneous, reflecting a wide range of trip patterns, travel intensities, and temporal behaviours that cannot be adequately captured by aggregated averages or single performance indicators. To address this complexity, we have employed clustering techniques to identify groups of vehicles with similar operational characteristics, or archetypes, based on empirical driving data. This data-driven segmentation enables the differentiation of fleet users according to their mobility behaviour – such as daily distance, trip frequency, and schedule regularity – and provides a structured foundation for evaluating electrification suitability. By modelling these operational *archetypes*, it becomes possible to align vehicle selection, charging strategies, and economic assessments with the actual patterns of use observed in the field, rather than relying on generalized assumptions or static usage profiles.

2. Methods

2.1. Driving Simulation

Dataset

The study relies on operational logs exported from the client's existing telematics platform carmonitor.eu. This European vehicle data repository provides comprehensive information on fleet composition and vehicle utilisation patterns. Data are provided as per-vehicle, semicolon-delimited CSV files and contain a time-ordered sequence of contiguous usage "events", where an event is defined as a period during which the vehicle remains in a single operational state. Two states are used in the raw export: driving for intervals when the vehicle is in motion and parked for intervals when the vehicle is stationary.

Each record contains six fields: Car_ID, Start, End, Event, Distance KM, and Average Speed. Start and End are timestamp strings indicating the beginning and end of the event interval as recorded by the platform. The timestamps are supplied without explicit timezone metadata by the export; for analysis, we treat them as local clock time as provided by the fleet operator and keep them consistent across all vehicles. The Event is a categorical label taking values, driving or parked. Distance KM records the total distance attributed by the platform to the event interval in kilometres; by convention, this field is populated for driving intervals and left blank in the parked state. Average Speed is the event-level average in kilometres per hour, likewise typically populated for driving intervals and absent for parked intervals.

The fundamental unit in the raw data is an interval, not an instantaneous point sample. That is, each row describes a span [Start, End) during which the operational state is assumed constant, and for which aggregate metrics (distance and average speed, when applicable) are reported. Intervals may vary in length from minutes to hours, and consecutive intervals are expected to tile the observation timeline for each vehicle with minimal gaps or overlaps; however, as with many operational datasets, occasional gaps (no state reported) or overlaps (partly redundant intervals) can occur due to connectivity, device resets, or post-processing. The export uses a consistent decimal and unit convention (kilometres, kilometres per hour).

As received, parked intervals contain no distance or speed information by design, and a small proportion of driving intervals may also have missing values in one of those fields; these are handled during preprocessing and imputation, which are described in the next subsection.

2.2. Clustering Methodology: General Approach

The objective of the clustering analysis is to identify groups of vehicles with comparable patterns of daily use. Vehicles that are used in a similar way – for example, driven mainly in the morning and evening with long overnight parking, or used continuously throughout the working day – can be expected to have similar charging opportunities and energy profiles. Clustering therefore provides an analytical foundation for the subsequent simulation and economic modelling of the fleet.

Clustering in this study is used to identify groups of vehicles with similar patterns of daily use, forming behavioural archetypes that serve as a foundation for fleet electrification analysis. The process begins by converting detailed driving and parking records into daily indicators such as distance, trip frequency, and time-of-day activity. Each vehicle's long-term behaviour is then summarised through aggregated statistics, standardised to ensure comparability. Using these behavioural profiles, vehicles are grouped with the k-means algorithm, and the optimal number of clusters is selected based on internal validity measures. The resulting clusters represent distinct usage types – such as commuter, service, or low-utilisation vehicles – providing a structured basis for subsequent modelling of energy demand and transition scenarios.

2.3. Feature Construction and Aggregation

Data representation

The transformation of raw vehicle event logs into consistent numerical features is a critical step before clustering. The aim is to represent each vehicle's daily behaviour in a way that is both comparable across vehicles and robust to differences in trip frequency or observation period length. This process consists of two levels: (a) constructing day-level indicators from individual driving and parking events, and (b) aggregating these daily measures into long-term vehicle-level descriptors.

Let $\mathcal{V} = 1,\ldots,n$ denote the set of vehicles in the dataset, and let $\mathcal{D}_i = 1,\ldots,T_i$ represent the set of observation days for vehicle i.

Construction of day-level features

The event log for each vehicle i contains consecutive records $\left(s_{ij}, e_{ij}, \xi_{ij}, d_{ij}, v_{ij}\right)$, where s_{ij} and e_{ij} denote the start and end timestamps of interval j, $\xi_{ij} \in \text{driving}, \text{parked}$ is the event type, d_{ij} is the distance travelled (km) if the event is "driving", and v_{ij} is the average speed (km·h^{-1}).

The first step transforms this sequence into a set of daily statistics representing how each vehicle is used during each day t.

For every vehicle i and day $t \in \mathcal{D}_i$, we compute a vector of p features

$$\mathrm{x}_{\mathrm{it}} = [\text{total\_distance}_{\mathrm{it}}, \text{driving\_time}_{\mathrm{it}}, \text{parking\_time}_{\mathrm{it}}, \text{avg\_speed}_{\mathrm{it}}, \text{morning\_share}_{\mathrm{it}}, \text{evening\_share}_{\mathrm{it}}, \text{night\_parking}_{\mathrm{it}},]^{\top} \quad (1)$$

These quantities are derived as follows:

– Total daily distance is the sum of all driving intervals' distances d_{ij} within the day.

– Number of trips counts transitions from parked to driving states.

– Driving and parking times are computed as the total durations of corresponding event types within the 24-hour window.

– Average speed is the distance-weighted mean of v_{ij} over all driving intervals.

– Morning and evening shares represent the proportion of total daily distance occurring between 6–10 a.m. and 4–8 p.m., respectively, capturing commuting intensity.

– Night parking hours measure the total duration of parking between 10 p.m. and 6 a.m., which approximates overnight availability for charging.

Collectively, these indicators summarise when, how far, and how intensively each vehicle is used during a typical day. They form the basis for comparing operational roles between vehicles.

Aggregation to vehicle-level profiles

Because each vehicle is observed over many days, the daily indicators $\{x_{it}\}_{t \in \mathcal{D}_i}$ are summarised into a *vehicle-level fingerprint* using robust statistical operators that are resistant to outliers and missing data.

Formally, the aggregation operator $\Phi(\cdot)$ maps the set of daily observations for vehicle i into a fixed-length feature vector:

$$z_i = \Phi(\{x_{it}\}_{t \in \mathcal{D}_i}) =$$

$$\left[median(x_{i1}), \quad IQR(x_{i1}), \quad p95(x_{i1}), \quad \ldots\right]^{\top}. \quad (2)$$

For each daily indicator we compute three summary measures: the *median* (typical value), the *interquartile range* (variability), and the *95th percentile* (upper limit of observed intensity). These statistics capture not only the average behaviour but also the stability and extremes of usage. For example, a vehicle with a low median distance but a high 95th percentile likely performs occasional long trips despite mostly short travel.

The aggregation yields the matrix

$$Z = [z_1^{\top}, z_2^{\top}, \ldots, z_n^{\top}]^{\top} \in \mathbb{R}^{n \times q}, \quad (3)$$

where n is the number of vehicles and q the number of aggregated features per vehicle. This matrix is the direct input for clustering.

Normalisation and scaling

Since the features z_i are expressed in different units (kilometres, hours, shares), we standardise them before comparing vehicles.

For each feature j, let μ_j and s_j denote robust measures of central tendency and dispersion (typically the median and median absolute deviation), and standardise the data as

$$\tilde{Z}_{ij} = \frac{Z_{ij} - \mu_j}{s_j}.$$

To limit the influence of extreme values, we apply a winsorisation operator $W_{\tau}(\cdot)$ that caps the highest and lowest quantiles (e.g. at 2.5 % and 97.5 %). The resulting matrix $\overline{Z} = W_{\tau}(\tilde{Z})$ provides a robust, scale-free basis for distance calculations.

If several features are highly correlated, we apply a principal component analysis (PCA) transformation. PCA projects the data onto a lower-dimensional orthogonal space $Y = \tilde{Z}U$, where $U \in \mathbb{R}^{q \times r}$ contains the first r principal component loadings capturing a target proportion (typically 90 %) of total variance. The clustering is then performed in this reduced feature space.

(d) Outputs of the preprocessing pipeline

The result of this multi-stage preprocessing is a dataset of vehicle-level behaviour profiles $Y = [Y_1, \ldots, Y_n]^{\top}$, each representing a single vehicle by its typical usage statistics

3. Clustering Algorithm

Clustering aims to partition the vehicles into K groups such that vehicles within the same cluster have similar use patterns and those in different clusters are dissimilar.

The most straightforward and widely used approach is the k-means algorithm, which minimises the total within-cluster variance:

$$\min_{\{m_k\}_{k=1}^{K}, \{C_k\}_{k=1}^{K}} J_K = \sum_{k=1}^{K} \sum_{i \in C_k} \| Y_i - m_k \|^2, \quad (4)$$

where m_k is the centroid (mean profile) of cluster k, and C_k is the set of vehicles assigned to it.

In practice, the optimisation is performed iteratively: vehicles are first assigned to the nearest centroid, centroids are recalculated as the average of the assigned members, and the process repeats until assignments stabilise.

Because k-means can converge to local minima, it is initialised multiple times with random seeds, and the configuration with the smallest objective J_K is retained.

As a robustness check, a Gaussian Mixture Model (GMM) can also be fitted, which allows for probabilistic rather than hard assignments. In that case, the likelihood

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \log\left(\sum_{k=1}^{K} \pi_k \, \mathcal{N}(Y_i \mu_k, \Sigma_k)\right) \quad (5)$$

is maximised using the Expectation–Maximisation (EM) algorithm, yielding posterior probabilities γ_{ik} that vehicle i belongs to cluster k.

Determining the number of clusters

The optimal number of clusters K^* is selected by evaluating a range of candidate values using several complementary indices.

For k-means, we compute the silhouette coefficient $S(K)$, which measures how similar each observation is to its own cluster compared with other clusters, and the Davies–Bouldin index $DB(K)$, which penalises overlapping clusters.

The selected number of clusters K^* is the smallest value that provides both high silhouette scores and low DB, corresponding to a distinct yet parsimonious segmentation.

For GMMs, we additionally use information criteria such as the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC).

To test the stability of the clustering solution, we apply a resampling approach by repeating the clustering on bootstrapped subsets of the data and measuring the Adjusted Rand Index between solutions. The final K^* balances internal fit, interpretability, and stability.

Cluster assignment and interpretation

Once the cluster structure is defined, each vehicle is assigned to the nearest centroid:

$$c_i = \arg\min_k \| Y_i - m_k \|, i = 1, \ldots, n, \quad (6)$$

and, when applicable, its confidence is measured by the relative distance margin or by the posterior probability $\max_k \gamma_{ik}$ from the GMM.

Cluster characteristics are then examined in the original feature space. For each cluster k, we compute the median and variability of key indicators (e.g. daily distance, number of trips, parking hours). These statistics form the cluster profile

$$\acute{z}_k = \text{median}\{z_i : i \in C_k\}, \quad (7)$$

which serves as a basis for interpretation. Descriptive tags such as "commuter-type vehicle", "urban operative vehicle", or "low-utilisation pool car" are assigned using heuristic rules based on these median values and known operational patterns.

Vehicles whose profiles lie far from any cluster centroid – identified by Euclidean distances – are flagged as potential outliers rather than forced into a cluster.

4. Fleet Descriptive Statistics

4.1 Overview of Fleet-Level Descriptive Statistics

The descriptive statistics summarise the fundamental operational characteristics of the analysed corporate vehicle fleet and establish the empirical context for subsequent segmentation.

Overall, the fleet demonstrates moderate daily utilisation with substantial variation among vehicles. The median of vehicle-level daily distances was 48.0 km (SD = 10.6), with the central 50% of vehicles ranging between 41 and 54.5 km. Nevertheless, higher-intensity users were clearly present: the 95th percentile of daily distance reached 158.3 km (SD = 56.2), and extreme cases exceeded 290 km. These values indicate that while most vehicles follow moderate, stable use patterns, a smaller subset engages in significantly more demanding operational cycles, likely reflecting field service or inter-urban travel.

The frequency of daily trips was consistent with light-duty commercial or commuter-type operation. The median number of trips per day was 2.0 (M = 2.44, SD = 0.69), confirming that most vehicles perform one outbound and one return trip per day. The mean number of trips per day was 3.35 (SD = 0.52), showing that some vehicles undertake additional intermediate journeys. The upper quartile reached approximately 3.6 trips per day, with occasional peaks above four, which may correspond to multi-stop service routes.

The duration and intensity of daily operation were limited relative to the total time vehicles spent inactive. Median driving time was 1.57 hours

Table 1
Descriptive fleet statistics

Indicator	Mean	SD	Median	IQR (P25–P75)	Range (min–max)
Median daily distance (km)	48.01	10.56	48	41.0–54.5	26.5–73.0
95th percentile daily distance (km)	158.33	56.18	139.3	115.3–198.3	73.3–297.2
Median trips per day	2	0.69	2	2.0–3.0	2.0–4.0
Mean trips per day	3.35	0.52	3.37	2.97–3.65	2.41–4.90
Median driving time (h)	1.57	0.3	1.57	1.34–1.75	0.84–2.39
Median parked time (h)	22.43	0.3	22.43	22.25–22.66	21.61–23.16
Drive–park ratio	0.06	0.01	0.07	0.06–0.07	0.03–0.10
Night parking duration (h)	16.26	1.19	15.91	15.40–16.74	14.36–20.18
Midday parking duration (h)	3.92	0.69	4.1	3.82–4.37	2.02–4.85
Median trip distance (km)	15.8	3.68	16.23	13.56–18.40	7.0–23.5
95th percentile trip distance (km)	23	5.29	23.5	19.82–26.66	10.25–33.22
Median trip duration (h)	0.54	0.13	0.56	0.51–0.64	0.24–0.76
95th percentile trip duration (h)	0.73	0.16	0.77	0.66–0.83	0.34–0.96
Coefficient of variation (daily km)	0.8	0.13	0.8	0.70–0.89	0.51–1.13
Start-time entropy (bits)	3.76	0.08	3.77	3.71–3.81	3.59–3.98
Share of weekend distance (%)	24.9	4.86	24.2	21.7–27.9	13.9–38.5

(SD = 0.30), while median parked time reached 22.43 hours (SD = 0.30), producing a mean drive–park ratio of 0.06 (SD = 0.01). This strong imbalance between motion and idleness demonstrates that the vehicles are available for charging during a large portion of each day, especially overnight. Indeed, vehicles were parked for an average of 16.26 hours per night (SD = 1.19), supplemented by roughly 3.9 hours of midday parking (SD = 0.69).

Trip-level indicators further characterise the short-distance nature of daily operations. The median trip length averaged 15.8 km (SD = 3.7), while the 95th percentile of trip distance was 23.0 km (SD = 5.3). Trip durations were correspondingly brief, with median trip times of 0.54 hours (≈32 minutes) and a 95th percentile of 0.73 hours (≈44 minutes). These results suggest that most travel occurs within a local or regional radius well within the typical range of contemporary battery-electric vehicles.

Temporal regularity measures highlight moderate variability in usage. The coefficient of variation in daily distance averaged 0.80 (SD = 0.13), indicating that many vehicles alternate between light and heavy usage days. The mean start-time entropy of 3.76 bits (SD = 0.08) suggests that departure times are partly predictable – typical of routine operations – yet still allow flexibility across weekdays. Weekend use was limited, comprising roughly 25% of total distance (SD = 0.05), confirming that the fleet's activity is predominantly weekday-based.

Taken together, these descriptive statistics reveal a heterogeneous operational structure encompassing both predictable, low-intensity users and irregular, high-mileage vehicles. The coexistence of such distinct behavioural profiles indicates that the fleet cannot be adequately represented by a single operational model. This heterogeneity provides the empirical foundation for the clustering analysis that follows, which classifies vehicles into operational archetypes to inform electrification suitability assessments and charging strategy development.

4.2 Determination of the Optimal Number of Clusters

To identify the most appropriate number of behavioural clusters, a sensitivity analysis was conducted for $K = 2$ to $K = 10$ using two complementary diagnostics: the *within-cluster sum of squares* (inertia) and the *mean silhouette coefficient*. The results are summarised in Figure 2.

The inertia values decreased monotonically from 1862 at $K = 2$ to 643 at $K = 10$, reflecting the expected reduction in within-cluster variance as more clusters are added. The *elbow method* therefore focuses on the point where the rate of improvement begins to flatten, indicating diminishing returns from additional partitions. In this dataset, a pronounced inflection was visible between $K = 2$ and $K = 4$, after which the curve gradually levelled off, suggesting that four clusters capture most of the structure present in the data.

The silhouette coefficients exhibited a similar pattern. The highest mean silhouette value (0.48) was obtained for the two-cluster solution, indicating strong separation at the coarsest partition. However, this configuration merged clearly distinct behavioural groups and was therefore judged overly simplistic. The silhouette value declined to 0.37 for $K = 3$ and to 0.24 for $K = 4$, with

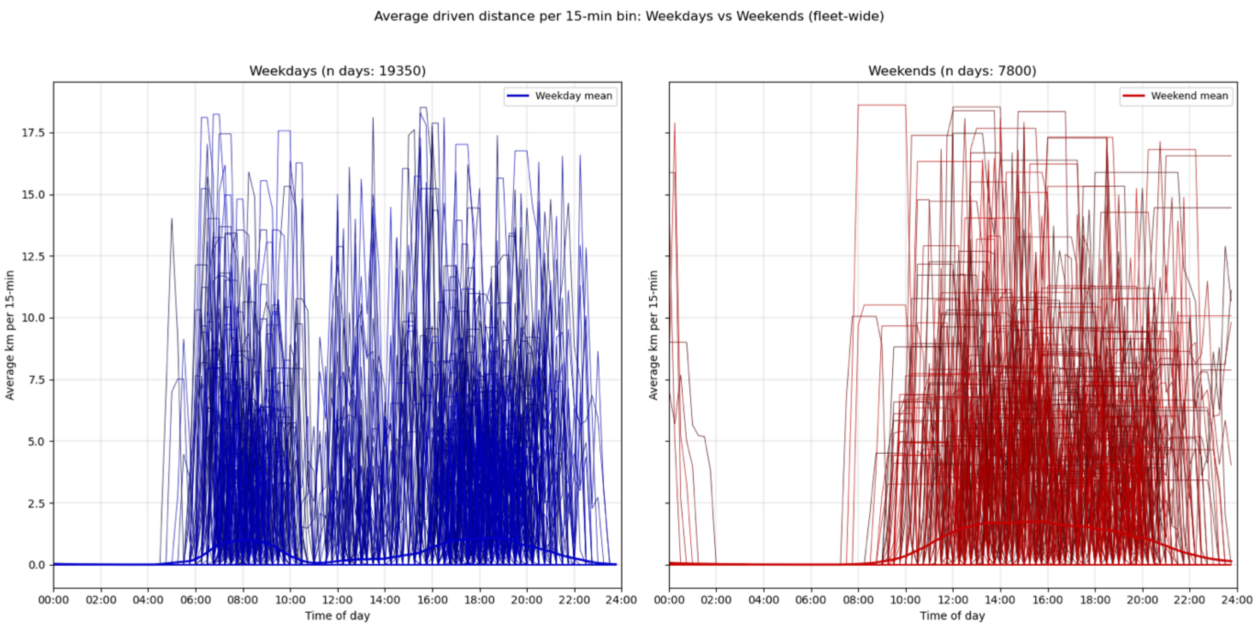

Figure 1. Fleet-wide Average Distance Driven in 15-Minute Intervals

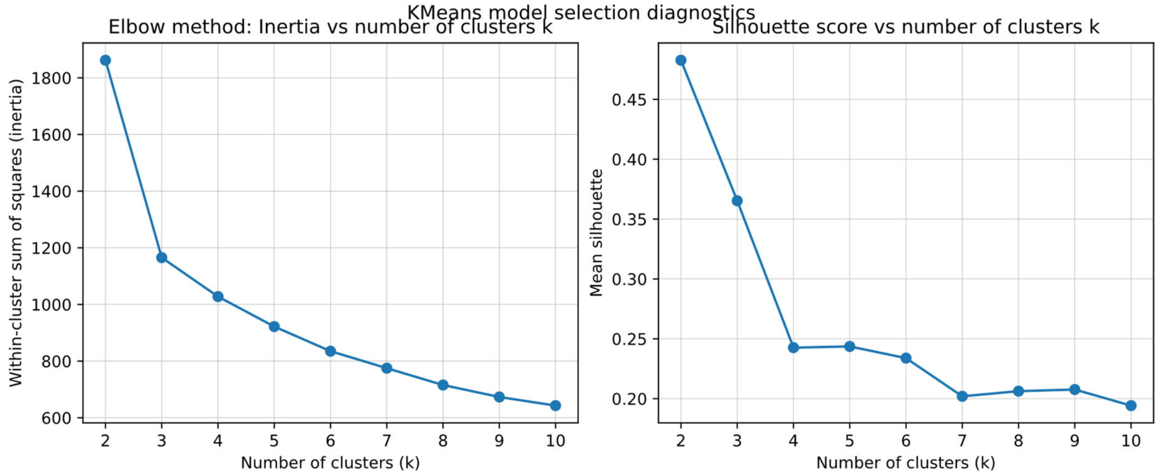

Figure 2. Cluster selection diagnostics

only marginal changes thereafter ($0.20 \le S(K) \le 0.24$ for $K = 5 - 10$). This flattening indicates that increasing the number of clusters beyond four yields no substantive improvement in internal cohesion.

Taken together, the elbow shape of the inertia curve and the stabilisation of the silhouette score beyond $K = 4$ both indicate that a four-cluster solution offers the best compromise between explanatory power and parsimony. This configuration provides sufficient granularity to distinguish major operational patterns while maintaining clear interpretability. Consequently, $K = 4$ was selected as the optimal number of clusters for subsequent analysis and interpretation.

4.3 The Fleet

The clustering solution separates vehicles primarily along two orthogonal dimensions of use: the intensity of daily distance accumulation and the fragmentation of that distance into trips. 4 clusters were observed.

Cluster description

Cluster 0: high-mileage, high-variability users

Vehicles assigned to Cluster 0 exhibit the highest intensity of daily distance, reflected by strong positive loadings for the 95th percentile of daily kilometres, the standard deviation of daily kilometres, and the mean and median daily kilometres. In the trips-per-day versus distance plane these vehicles occupy the high-distance region, with trip counts that can vary but are less diagnostic than the sheer volume of kilometres. This archetype is consistent with long-range service, field operations, or inter-urban logistics in which routing is dynamic and peaks of very long travel days occur with meaningful frequency.

Cluster 1: low-utilization, schedule-regular commuters

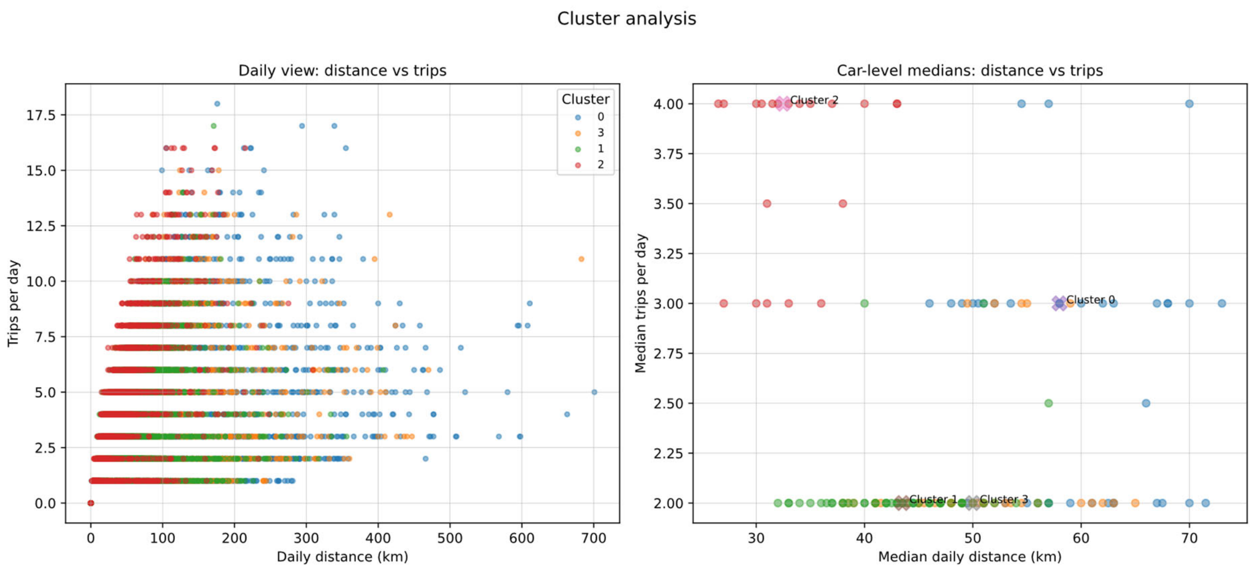

Figure 3. Vechicle Use Cluster Analysis

Cluster 1 sits at the opposite end of the utilization spectrum. The salient features all load negatively: coefficient of variation and standard deviation of daily kilometres, and the number of trips per day are all below fleet averages. Start-time entropy is also lower, indicating a more regimented schedule with departures concentrated in narrow time windows. The share of weekend kilometres is depressed, which reinforces the interpretation of routine weekday use. In the two-dimensional plot, these vehicles populate the low-distance, low-trips region.

<u>Cluster 2: multi-stop urban duty with short trip lengths</u>

Cluster 2 is distinguished by simultaneous signals of high fragmentation and short trip characteristics. Median trips per day load strongly positive, while median and 95th percentile trip distances and durations load strongly negative. Mean mid-day parking time and the 95th percentile of mid-day maximum parking duration are both notably below average, indicating that vehicles remain active rather than stationary during business hours. In contrast, night-time parking accumulates above-average hours, consistent with depot dwell or home garaging. Plotted against trips per day and distance, these vehicles gravitate to the high-trips, low-distance corner.

<u>Cluster 3: low-frequency, long-duration trips with mid-day dwell</u>

Vehicles in Cluster 3 display above-average trip durations at both the median and the upper tail, accompanied by higher median driving hours and a modestly elevated mid-day parking total. The median number of trips per day is below average, and median parked hours are somewhat reduced, suggesting days organized around one or a few longer journeys interspersed with a notable mid-day stop. In the visualization these vehicles appear at lower trip counts with moderate distances that are achieved through longer, less fragmented drives.

Comparative Interpretation

Comparing all four clusters, two dimensions emerge as dominant in defining fleet heterogeneity: daily driving intensity and temporal dispersion. Cluster 0 occupies the high-intensity, long-duration extreme; Cluster 1 represents low-intensity, time-bound commuting; Cluster 2 captures high fragmentation within a concentrated workday; and Cluster 3 combines moderate distance with extended driving windows. Weekday and weekend contrasts are most pronounced in Clusters 0 and 2, where weekend activity exceeds weekday levels, while Cluster 1 remains strongly weekday-oriented.

The average 15-minute distance profiles reveal distinct operational signatures for each cluster, reflecting the diversity of use patterns within the analysed corporate fleet. The graphs display the mean distance travelled across 15-minute intervals throughout the day, separately for weekdays (blue) and weekends (red), with shaded areas indicating the 10th–90th percentile spread. These results highlight significant variation in both daily intensity and temporal structure of vehicle use across the four clusters.

Cluster 0 is characterised by the most intensive and prolonged daily activity profile. Weekday driving begins gradually around 06:30, peaks between 07:30 and 09:00, and remains consistently high throughout business hours, with sustained movement continuing well into the evening. Weekend activity is even broader, spanning nearly the entire daytime period and reaching higher intensity levels than on weekdays. The extended operational window and weak midday trough suggest irregular routing or service-oriented use with minimal downtime. This pattern corresponds to long-distance, high-mileage vehicles that require substantial daily energy input and access to fast-charging options.

In contrast, Cluster 1 displays a highly structured and predictable daily pattern typical of commuter or short-duty vehicles. Two distinct weekday peaks are visible: one in the morning between 07:00–09:00 and another around 16:00–18:00, with little activity outside these windows. The rest of the day is dominated by vehicle idling or parking. Weekend driving is minimal and occurs primarily in the early afternoon. The low amplitude of the distance curve confirms that these vehicles accumulate relatively small daily mileage. Their predictable operation and long stationary periods make them well suited for overnight charging, with no operational dependence on public infrastructure.

Cluster 2 shows a markedly different temporal structure, with the highest concentration of activity during mid-day hours. Weekday driving intensity increases around 09:00, peaks between 11:00 and 15:00, and declines gradually toward evening. Weekend driving follows a similar pattern, though at slightly higher intensity and extended duration. This cluster likely represents service or delivery vehicles with numerous short trips concentrated in business hours and reduced movement during morning and evening commute periods. Their operational predictability and daytime concentration make them ideal candidates for depot-based overnight charging, supplemented by limited opportunity charging if required for longer shifts.

Cluster 3 demonstrates an intermediate pattern combining elements of commuting and long-route travel. Weekday activity shows two subdued peaks – one in the morning and one in the afternoon – but unlike Cluster 1, driving continues through mid-day with moderate intensity. The weekend profile is broader, covering most daylight hours with steady movement. These vehicles likely undertake fewer but longer trips, possibly with defined mid-day tasks or longer-distance assignments. Their operational rhythm suggests moderate energy

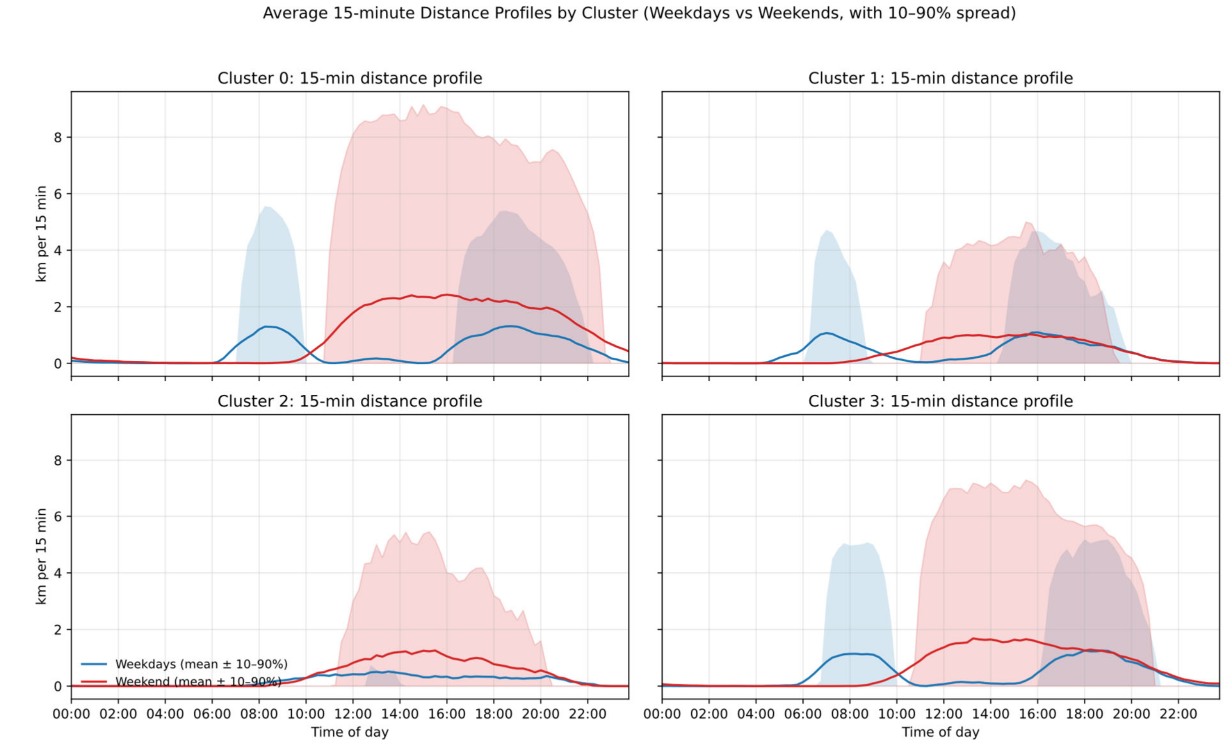

Figure 4. Average 15 minute Distance Profiles by Cluster

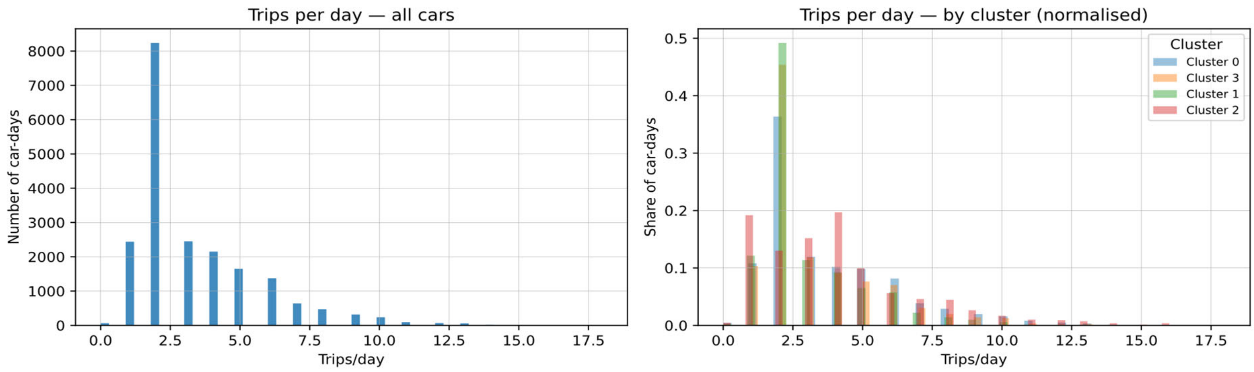

Figure 5. Vechicle Trips per Day

requirements but limited opportunities for daytime recharging, making scheduled mid-day or destination charging particularly relevant.

4.4 The Driving Patterns

Vehicle trip distribution

The trip distribution plot in Figure 5 reveals clear differentiation between clusters in both the number and length of trips. Vehicles in Cluster 1 exhibit a sharply peaked distribution centred on one to two daily trips, consistent with regular commuter or short-shift use. The low spread confirms that these vehicles operate in a stable, highly predictable pattern with limited day-to-day variation. In contrast, Cluster 2 shows a broader distribution with a modal range of three to five trips per day, characteristic of multi-stop duty cycles or service operations where vehicles make repeated short journeys between locations. Cluster 0 displays a right-skewed distribution, combining relatively few but often long trips, aligning with long-distance or inter-urban usage patterns that contribute disproportionately to total mileage. Finally, Cluster 3 sits between these extremes, reflecting moderate trip counts with mixed trip lengths and more variable day-to-day utilisation.

The two graphs depicting vehicle trip distance distribution (Figure 6) and average parking duration during the night (Figure 7) provide further empirical evidence of the heterogeneity identified in the descriptive statistics and clustering analysis.

The average distance per car figure reinforces this segmentation. Cluster-level averages diverge

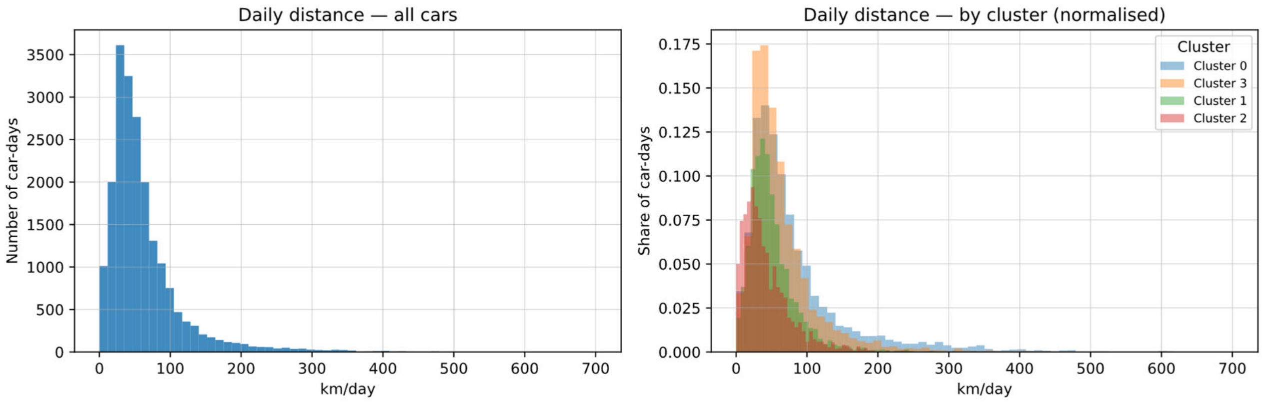

Figure 6. Distance Driven per Day

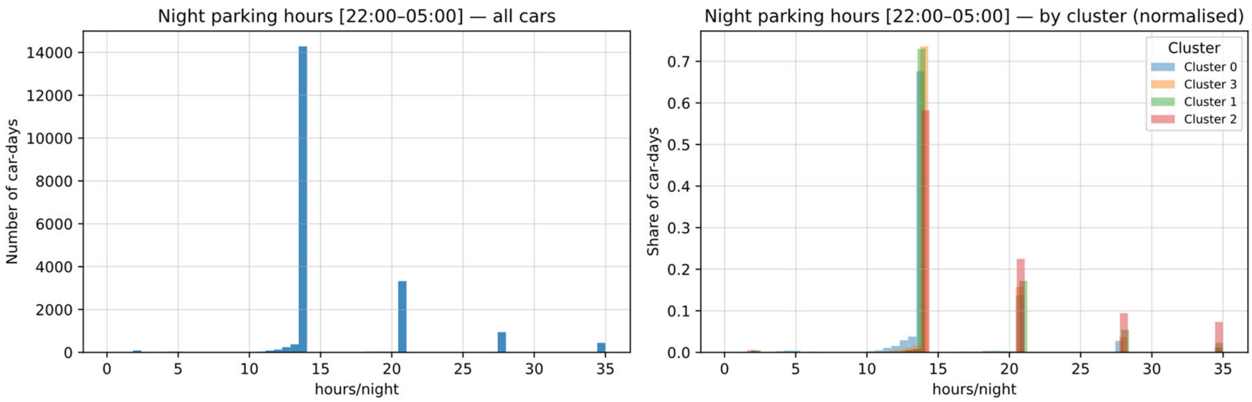

Figure 7. Parking Hours per Night

substantially, with Cluster 0 vehicles showing the highest mean daily distance, typically exceeding 100 km, while Clusters 1 and 2 average less than half of that. Cluster 3 occupies an intermediate position but demonstrates a wider internal spread, suggesting that while most vehicles remain within moderate daily ranges, a subset occasionally undertakes longer journeys. These differences in both total distance and trip fragmentation confirm that the fleet encompasses several distinct operational regimes rather than a single unified pattern of use.

Weekday usage analysis

Analysing the average driving data by individual cars, it is notable, that most of the vehicles follow a distinct trip distribution, which we have labelled managerial vehicles.

The Figure 8 presents the average driving distance by time of day, separated by day of the week, showing a clear

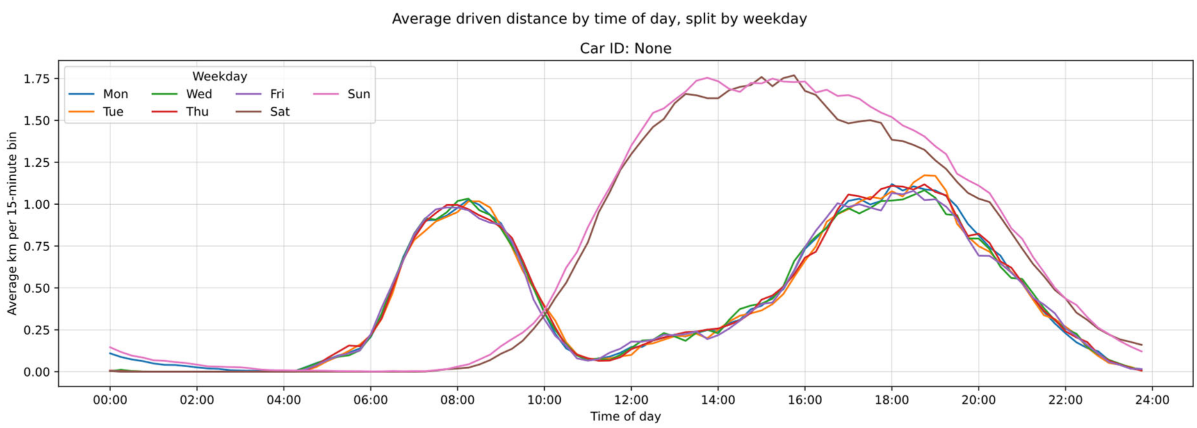

Figure 8. Average Distance Driven by Days of the Week

bimodal pattern typical of daily commuting or service-related vehicle use. Two distinct peaks are visible: the first in the morning between 07:00 and 09:00, corresponding to departure trips, and the second in the afternoon between 15:30 and 18:30, linked to return or end-of-day travel. Between these peaks, driving activity remains low, indicating that most vehicles are stationary or engaged in longer single trips during mid-day hours.

The weekday profiles (Monday–Thursday) are almost identical, reflecting regular and predictable operation patterns. In contrast, Friday and Saturday show elevated afternoon and evening activity, with Saturday maintaining high levels until around 21:00, suggesting more flexible or leisure-related driving behaviour. Night-time use across all days is minimal, confirming that the fleet operates mainly within daylight hours.

From an operational perspective, the mid-day trough between 10:00 and 14:00 represents a natural opportunity window for vehicle charging, while the extended evening activity on weekends indicates the need for more adaptive charging strategies. Overall, the temporal distribution underscores a consistent daily rhythm of use, combining high predictability during workdays with greater variability toward the weekend – an important input for modelling time-dependent charging demand and energy optimization strategies.

5. Discussion

From a methodological perspective, the coherence of feature rankings within each group supports the validity of the clustering solution. In particular, the alignment between the scatterplot positioning and the signed z-scores of trip-count, duration, distance, and temporal-entropy features indicates that the clusters are not artifacts of the algorithm but reflect underlying behavioural regimes. This strengthens the case for using cluster membership as a stratification variable in subsequent techno-economic analysis, including battery sizing, charger power selection, and charging-window allocation in smart-charging scenarios.

In the broader analytical context, these results strengthen the argument for data-driven fleet segmentation. The pronounced variation in trip frequency and travel intensity indicates that electrification strategies must be tailored to the specific operational profile of each cluster. Vehicles with low trip counts and consistent daily distances can be readily electrified with smaller battery packs and depot-based overnight charging, whereas those in high-trip or high-mileage clusters will require larger-capacity vehicles and flexible access to daytime charging infrastructure. Thus, the trip distribution patterns provide a quantitative link between observed behaviour and the differentiated transition pathways modelled in later sections of the study.

From an electrification standpoint, Cluster 0 presents the greatest requirement for either large usable battery capacity or reliable access to high-power public DC charging, given the combination of high daily energy demand and temporal unpredictability. Operational planning for this segment should emphasize range-adequate model selection and route-integrated charging strategies.

For Cluster 1 electrification suitability is correspondingly high. Modest energy requirements and strong schedule predictability support overnight depot or home charging without the need for opportunistic fast charging. This group can typically be transitioned with smaller-capacity battery electric models at minimal operational risk, delivering early cost and emissions benefits.

Cluster 2 is a canonical stop-and-go urban service pattern – e.g., delivery rounds, service calls, or intra-city distribution – where daily energy demand is moderate but distributed across many short drive segments with limited mid-day idle windows. Such cycles are well aligned with battery-electric operation provided that overnight charging is reliable; in dense urban contexts, limited dwell at mid-day reduces the practicality of low-power opportunity charging, so either predictable depot returns or occasional DC top-ups may be warranted during peak days. Regenerative braking benefits are likely to be material in this segment due to frequent decelerations.

For Cluster 3 electrification implications are nuanced. While daily kilometre totals are not as extreme as in Cluster 0, the continuous nature of driving episodes and the presence of a single mid-day dwell create a natural anchor for scheduled AC or DC top-ups, for example at a known destination or customer site. Right-sizing battery capacity to cover the dominant continuous leg while planning a medium-power charge during the mid-day stop typically yields robust operational feasibility.

6. Conclusion

The findings of this study provide a empirical understanding of corporate fleet usage and its implications for vehicle electrification. Through clustering analysis, four distinct operational archetypes were identified, each characterized by unique travel intensity, trip fragmentation, and temporal driving patterns. These include high-mileage vehicles with irregular schedules, low-utilisation commuters with stable routines, multi-stop urban-duty vehicles with short repetitive trips, and low-frequency users combining longer drives with mid-day dwell periods. The identification of these patterns confirms that corporate fleets are far from homogeneous and that electrification strategies must be differentiated according to operational behaviour.

The analysis of average daily driving profiles reinforces these distinctions. Weekday usage exhibits a pronounced bimodal distribution, with morning and afternoon peaks reflecting conventional working hours, while weekend activity extends later into the evening, indicating greater behavioural flexibility. This temporal structure highlights predictable windows of vehicle availability – particularly the mid-day trough and overnight hours – that are well suited for controlled or scheduled charging. Conversely, fleets operating with higher temporal variability, particularly on weekends or across regional routes, will require more flexible or distributed charging infrastructure, including access to fast chargers during operational hours.

These empirical insights suggest that fleet electrification should not be approached through a single financial or technical criterion such as purchase cost or nominal range. Instead, transition planning must account for the specific duty cycles that determine energy demand, charging availability, and operational risk. High-mileage, irregular-use vehicles will depend on fast-charging networks and robust route planning, whereas low-distance, schedule-stable fleets can achieve immediate cost and emissions benefits with smaller, depot-charged BEVs. Urban, multi-stop vehicles appear particularly well suited for early electrification, benefiting from short trip distances, high regenerative braking potential, and overnight depot access.

From a methodological standpoint, integrating empirical usage data with simulation and scenario-based modelling offers a powerful means to evaluate electrification feasibility under dynamic conditions. This approach enables the assessment of total cost of ownership and emissions performance under realistic operational constraints, supporting evidence-based transition planning.

The next phase of this research will extend these analyses by combining the empirical datasets provided by carmonitor.eu with economic simulation to quantify cost structures for each of the clusters, the corresponding infrastructure requirements, and emission outcomes across alternative fleet transition pathways.

Acknowledgement

This paper has been published within the research project "Software Development for the Transition to Environmentally Friendly, Zero-Emission Vehicles in Company Transport Fleets" carried out under the grant program funded by the Recovery and Resilience Facility and Central Finance and Contracting Agency of Latvia as part of the research program "Competency Centre for Digital Products" Project number: 2.2.1.3.i.0/2/24/A/CFLA/001.

Author's Contributions

A.R. conducted the main research and analysis. J.T. provided methodological guidance and critically reviewed the manuscript. V.M. was responsible for data acquisition, dataset preparation, and securing financial support for the research.

References:

Al-Alawi, B. M., & Bradley, T. H. (2013). Total cost of ownership, payback, and consumer preference modeling of plug-in hybrid electric vehicles. *Applied Energy*, 103, 488–506. DOI: https://doi.org/10.1016/j.apenergy.2012.10.009

Antich, Arnau Oliver (2024). *Unveiling Europe's Corporate Car Problem*. Transport & Environment.

Coffman, M., Bernstein, Paul, & Wee, S. (2017). Electric vehicles revisited: A review of factors that affect adoption. *Transport Reviews,* 37(1), 79–93. DOI: https://doi.org/10.1080/01441647.2016.1217282

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Decarbonise Corporate Fleets, No. COM/2025/96, European Commission (2025). Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025DC0096

Cornelis, S., & Antich, A. O. (2023, May). *The corporate cars problem and what the EU can do about it.* Transport & Environment. Available at: https://www.transportenvironment.org/uploads/files/Briefing-on-Fleets-Regulation-3.pdf

Cornelis, S., & Antich, A. O. (2025). Corporate fleets and the EU Automotive Industrial Action Plan: How the European Commission can boost demand for made-in-Europe EVs [Briefing]. Available at: https://www.transportenvironment.org/articles/fleets-electrification-law-could-deliver-2-million-ev-sales-half-what-carmakers-need-to-meet-their-2030-emissions-targets-2#:~:text=presiding%20over,car%20on%20the%20used%20market

Di Foggia, G. (2021). Drivers and challenges of electric vehicles integration in corporate fleet: An empirical survey. *Research in Transportation Business & Management,* 41, 100627. DOI: https://doi.org/10.1016/j.rtbm.2021.100627

European Environment Agency (2024). *Sustainability of Europe's mobility systems* (Web Report No. 01/2024). European Environment Agency. Available at: https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems

Haustein, S., Jensen, A. F., & Cherchi, E. (2021). Battery electric vehicle adoption in Denmark and Sweden: Recent changes, related factors and policy implications. *Energy Policy,* 149, 112096. DOI: https://doi.org/10.1016/j.enpol.2020.112096

Héliot, R., & Ferrara, G. (2025). *Recommendations for the Upcoming EU Corporate Fleets Legislation for Cars and Vans.* Available at: https://www.emobilityeurope.org/wp-content/uploads/2025/09/Recommendations-for-the-Upcoming-EU-Corporate-Fleets-Legislation-for-Cars-and-Vans.pdf

Krishnan, V. V., & Koshy, B. I. (2021). Evaluating the factors influencing purchase intention of electric vehicles in households owning conventional vehicles. *Case Studies on Transport Policy*, 9(3), 1122–1129. DOI: https://doi.org/10.1016/j.cstp.2021.05.013

Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 Establishing the Framework for Achieving Climate Neutrality and Amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law'), No. 2021/1119, European Parliament (2021).

Rubenis, A., Laizāns, A., & Zvirbule, A. (2019). Evaluation of a Necessity for Subsidies for Electric Vehicle Purchases in Latvia: 2013–2017. In I. Kabashkin, I. Yatskiv (Jackiva), & O. Prentkovskis (Eds.), *Lecture Notes in Networks and Systems* (pp. 637–646). Springer. DOI: https://doi.org/10.1007/978-3-030-12450-2_61

Rubenis, A., Tonova, J., & Morozovs, V. (2025). The Road to Zero-Emission Fleets: The Role of Data-Driven Decision-Making. *Environmental and Climate Technologies,* 29(1), 600–610. DOI: https://doi.org/10.2478/rtuect-2025-0041

Transport & Environment. (n.d.). *Car taxation: Good Tax Guide.* T&E. Retrieved September 12, 2025, from https://www.transportenvironment.org/topics/cars/good-tax-guide

Received on: 13th of September, 2025
Accepted on: 30th of October, 2025
Published on: 13th of November, 2025