EXPLORATORY DATA ANALYSIS OF CHARGING EVENTS

EV Charging Behaviours - Data Science intern role

BY

Srinivasan Gomadam Ramesh

Master's in Business Analytics

The University of Texas at Dallas

Contents

Exploratory data analysis of charging events

Introduction:

The dataset consists of charging events data of 16 charging locations having energy meter readings & duration of charging, having 277 instances of charging events, the dataset has the following columns

1. **Start Time**: Timestamp When the charging event started.
2. **Meter Start (Wh)**: Meter reading at the start of charging.
3. **Meter End (Wh)**: Meter reading at the end of charging.
4. **Meter Total (Wh)**: Total energy consumed during the charging event.
5. **Total Duration (s)**: Total duration of the charging event in seconds.
6. **Charger_name**: Name of the charging locations (data type: object/string).

From the dataset variables, I could assume to find insights such as charging trends and pattern on the time frame across all the chargers, charging capacity of various chargers and instances, type of charging behavior like overnight charging, rapid charging etc.

Executive Summary:

- Charger_1 has multiple outlets and varied start meter readings suggest a high-traffic, multi-charger station.
- Weekly Charging Trends: Mondays peak in charging events, hinting at a weekly preparation cycle for EV users.
- Power Rate Variance: Charger 11 exhibits the capability for rapid charging with its high-power rate spikes above 100 KW.
- Energy Usage Patterns: The average energy consumption is around 9.66 KWh per session, aligning with residential charging type.
- Charging Time Diversity: Chargers 1 and 16 are predominantly used for brief durations, indicating fast or daytime charging preferences.
- Anomalies in Duration: Extended charging durations over 25 hours raise flags for potential data errors or non-EV charging scenarios

Data Cleaning & Preparation

Observations

Date time inconsistencies: Based on the observations from the csv file dataset, we could find lot of inconsistencies in the data such as non-sequential Start times, as the 'Start Time' entries are not in chronological order. This might suggest that the data has not been sorted by the start time, or the charging sessions could be from different charging stations that are not synchronized.

Charging locations are independent: The charging locations (charger name) has no relation/correlation among other charging locations, each charger name can be considered as unique dataset. Still we could find the date time inconsistency as the events Where not in chronological order but the meter readings were **cumulative/consecutive** based on Meter start and End readings are almost consecutive and so the corresponding index Where in sequential order, having discontinuity in the meter readings as well.

Zero Entries/ missing values: The columns total duration(s) and Meter total (Wh) has many of the entries Where 0, there were discrepancies in these records, both charging events having no charge transfer and 0 duration charge time so these records having no significance, similarly many records

having 0-meter total (zero energy transfer) but the charging duration were logged and vice versa. This clearly signifies the data has some anomalies properly might be several reasons one such might be potential malfunctions in the logger, date time reset/ maintenance in the chargers, may be due to delayed logging of the data to server

Cleaning

Based on the dataset consistencies above 3 reasons these records are dropped:

- '0' entries of total duration(s)
- '0' entries of Meter total (Wh)

tried sorting of the records in ascending based on the Start time instance Which seems to be inconsistent so I considered using the index and the meter start and meter End as reference, not the Start time column to understand the sequence of pattern Which seemed to be logical and consistent.

Data Preparation

As per the understanding of the dataset pattern from visual exploration, though the meter readings Where consistent cumulative but the values were not continuous this might be due to some inaccuracies or resolution-based issue due to loggers & generally the energy consumption are measured in units (KWH) so the Meter Total (Wh) is transformed to KWH in new column. Tried transforming/changing the date time discrepancies but there was not any pattern to identify the correct date values.

Total duration(s): data column is transformed to a new column in hour format Total Duration (hr) Which provided practical understanding.

Start Time: This column is converted to date/time format for time series analysis and for data transformations.

End Time: Based on the duration and start time the end date/time of charging event in calculated added to a new column.

Day of the week: based on the start time date values is calculated and added to find the pattern/insights

Hour of the day: It is also calculated based on the start time values to understand the pattern charging time frame of the users

Charger_ID: This column is added to better visualize the trend and sort the chargers in order.

Power Rate (KW): This column provides the rate of charging, which is an average value energy transfer over entire charging duration.

Descriptive statistics

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 167 entries, 12 to 276
Data columns (total 13 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Start Time           167 non-null    datetime64[ns]
 1   Meter Start (Wh)     167 non-null    int64
 2   Meter End(Wh)        167 non-null    float64
 3   Meter Total(Wh)      167 non-null    float64
 4   Total Duration (s)   167 non-null    int64
 5   Charger_name         167 non-null    object
 6   Meter Total(KWH)     167 non-null    float64
 7   Total_duration(hrs)  167 non-null    float64
 8   End Time             167 non-null    datetime64[ns]
 9   Day of the Week      167 non-null    object
 10  Hour of Day          167 non-null    int64
 11  Charger_ID           167 non-null    int64
 12  Power Rate (KW)      167 non-null    float64
dtypes: datetime64[ns](2), float64(5), int64(4), object(2)
memory usage: 18.3+ KB
```

Figure 1 Info of the datatype of each Columns

A total of 167 non-null instances are filtered after data cleaning, capturing a wide array of information from starting and ending meter readings to the total energy consumed and the rate at Which power is transferred during each event. This section outlines the central tendencies and variations of the main features.

	Meter Start (Wh)	Meter End(Wh)	Meter Total(Wh)	Total Duration (s)	Meter Total(KWH)	Total_duration(hrs)	Hour of Day	Charger_ID	Power Rate (KW)
count	1.670000e+02	1.670000e+02	167.000000	1.670000e+02	167.000000	167.000000	167.000000	167.000000	167.000000
mean	3.216183e+05	3.312812e+05	9662.899281	1.199537e+05	9.662899	33.320484	10.622754	5.898204	2.577300
std	3.477542e+05	3.461375e+05	14569.940599	3.464249e+05	14.569941	96.229139	3.606582	4.050338	7.581268
min	0.000000e+00	5.659070e+03	9.770000	1.200000e+01	0.009770	0.003333	3.000000	1.000000	0.000022
25%	7.363000e+04	8.107181e+04	1205.520000	5.131500e+03	1.205520	1.425417	8.000000	3.000000	0.114345
50%	1.787700e+05	1.788596e+05	4698.860000	3.263200e+04	4.698860	9.064444	10.000000	4.000000	0.739817
75%	4.907780e+05	5.140272e+05	12356.640000	8.806850e+04	12.356640	24.463472	12.000000	8.000000	2.383610
max	1.193443e+06	1.197233e+06	126350.920000	3.020411e+06	126.350920	839.003056	22.000000	16.000000	88.494808

Figure 2 Summary statistics of each column

Energy Consumption

The total meter readings have an average of approximately 9662 Wh, with the lowest being at 0 Wh and the highest at around 126350 Wh, indicating a wide range of initial conditions.

Total (KWh): When converted to kilowatt-hours, the average consumption stands at around 9.66 KWh, aligning with typical residential charging patterns.

Charging Duration

Total Duration (s): The duration of charging events averages close to 1,966 seconds (or roughly 32.7 minutes), with some sessions lasting a mere second and others extending to over 12,365 seconds.

Total Duration (hrs): When translated into hours, the average duration is approximately 0.55 hours, further highlighting the variability of charging sessions.

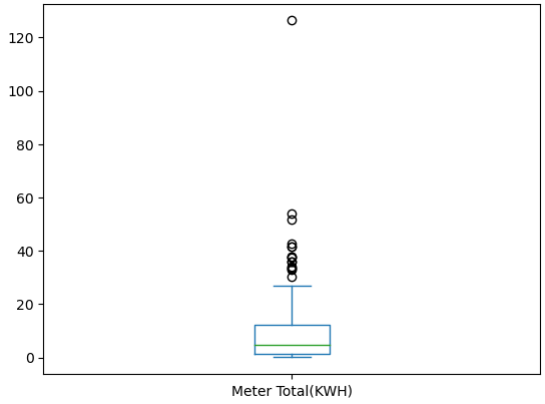
Charging Rates and Times:

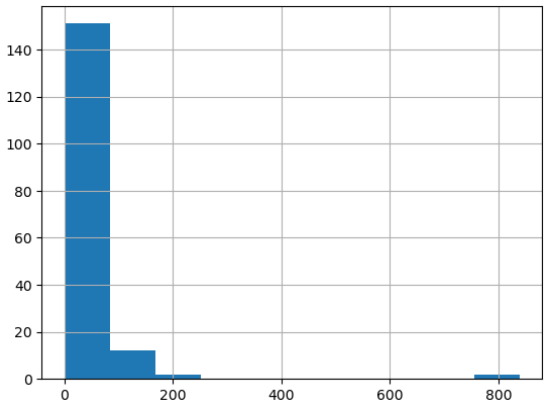
Power Rate (KW): The rate at Which power is transferred during a charging event averages at 2.57 KW, showcasing the average power rating of the of charging sessions

Visual Analysis:

Meter Total (KWh) Boxplot: The boxplot reveals a concentration of energy consumption within a lower quartile, with outliers suggesting that occasional charging sessions consume substantially more energy.

Total duration (hrs) Histogram: The histogram further emphasizes the skewness towards lower charging durations Which could be indicative of the predominance of short, quick charging sessions or the capacity of the vehicles being charged.

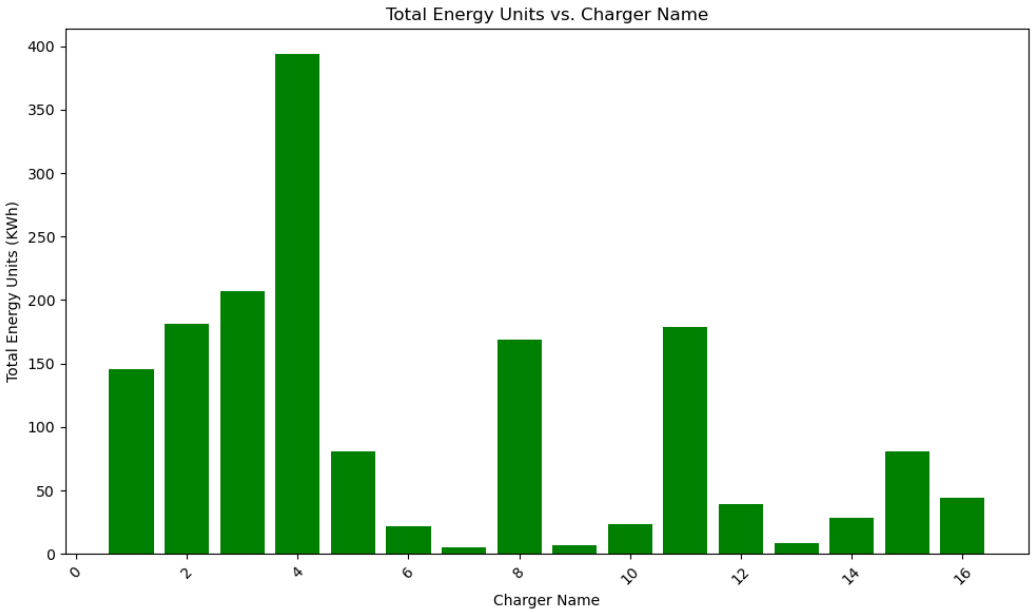




Data Visualization & Insights

Total Energy Units vs. Charger Name

The bar chart depicts the total energy consumption attributed to each charger, providing insight into the utilization patterns of the charging stations. The x-axis represents the charger names or ID, While the y-axis quantifies the total energy units in kilowatt-hours (KWh) consumed by electric vehicles including all sessions at each station.

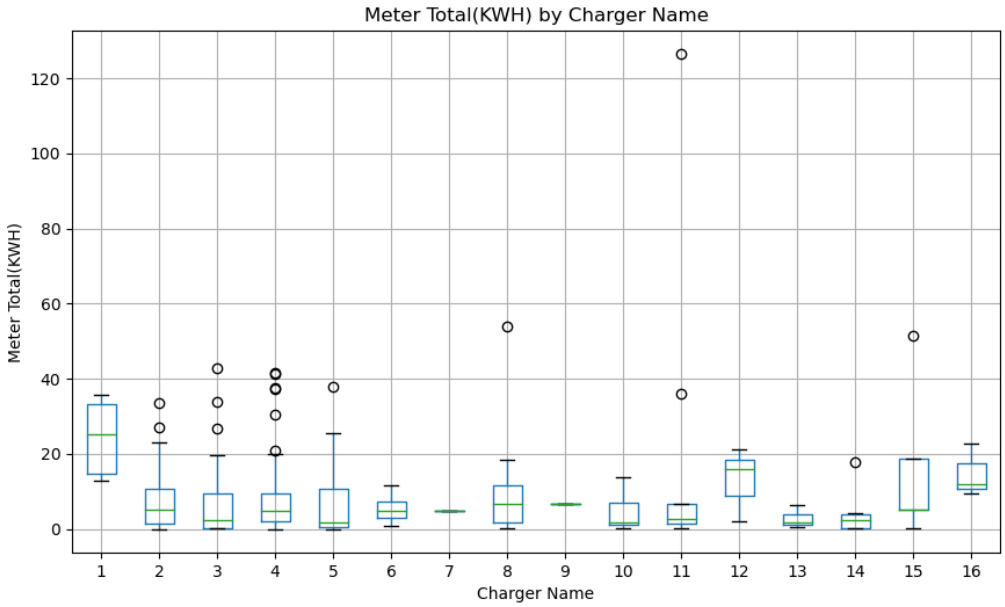


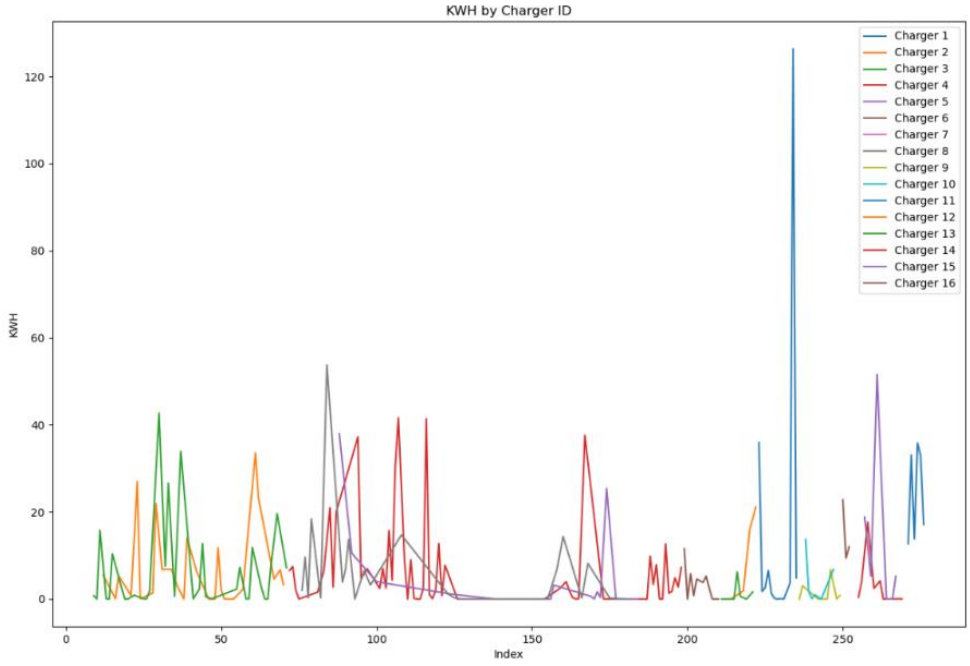
We can infer that the charger 4 seems to be frequently & might be placed in location for high demand of EV chargers based on the available data points. Similarly, charger locations of **1,2,3,8 & 11** have similar pattern for the demand.

Energy Consumption Distribution by Charger

The boxplot provides a detailed visualization of the distribution of energy consumption across the 16 charging stations identified by charger ID on the x-axis vs energy in kilowatt-hours (KWh) distribution. Most chargers show a relatively narrow IQR, which suggests a consistent pattern of energy consumption among different charging sessions for the same charger. The median value in the plot signifies the capacity/ types of vehicles charged or charging pattern like overnight or fast charging.

Sessions having relatively high units of charging seems to have higher storage capacity of 120KWH or can be inferred as heavy-duty vehicle. On average the meter total signifies a load like mid-size EV and some higher segment EV.



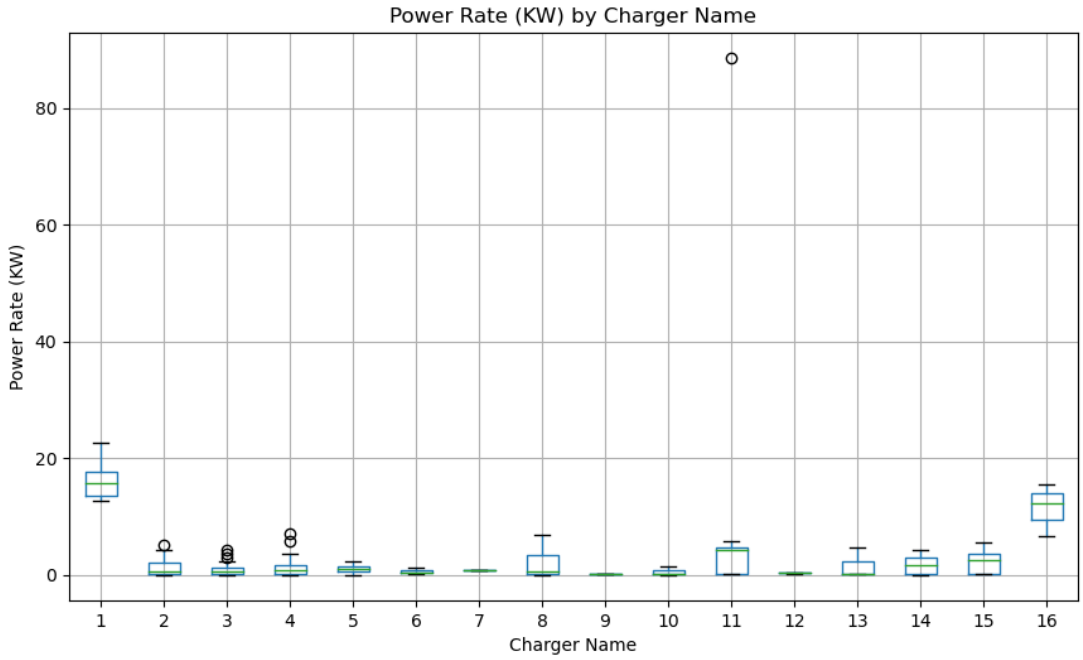


This line graph represents the charging event in sequential way with index as reference as the date time is not consistent. The trend signifies many chargers have similar charging demand compared to few outliers like charger 11 and 5 having diverse set of demand.

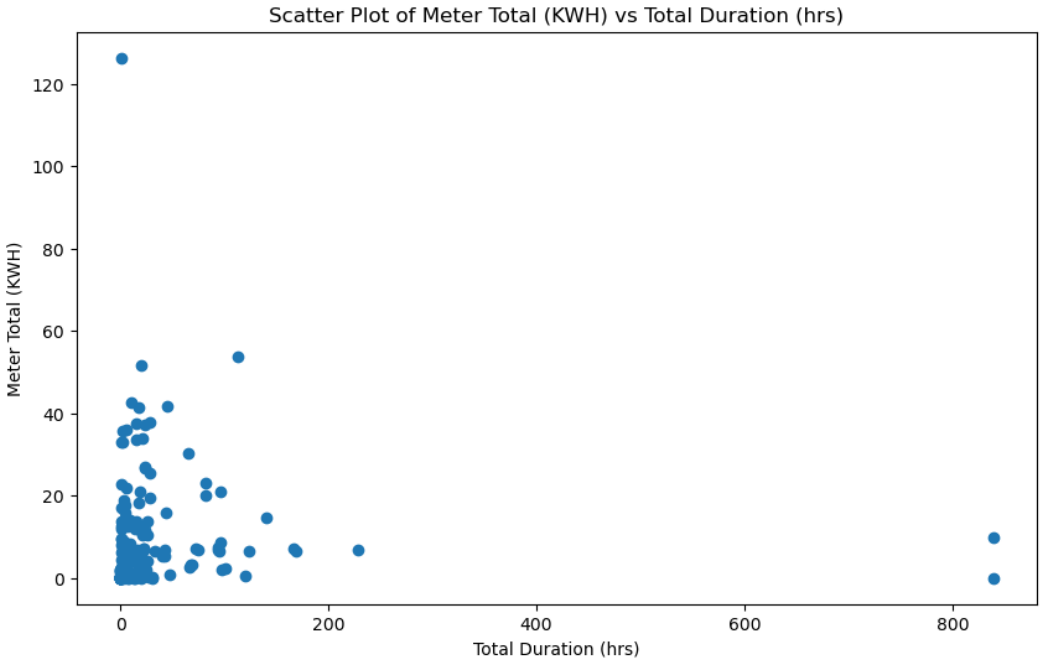
Power Rate (KW) by Charger Name

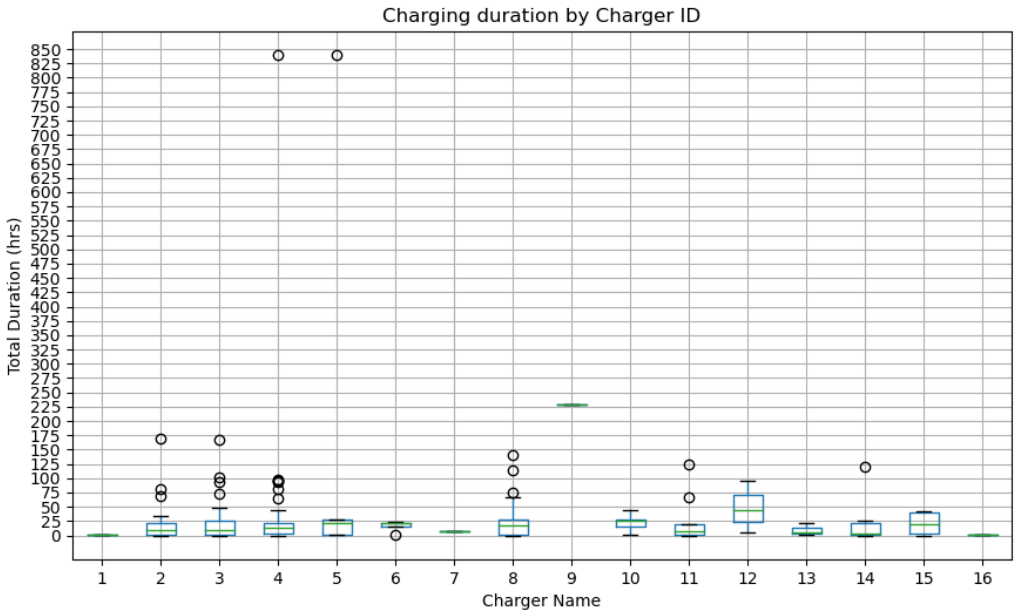
The boxplot portrays the range of power rates for each charger, with chargers identified on the x-axis and the corresponding power rate in kilowatts (KW) on the y-axis. The power rate, indicating the speed at Which chargers can deliver energy, shows variability across chargers.

Chargers are consistently delivering power within a specific range, indicated by the compactness of the boxes. This value of power rate is average value based on each session of charge duration. We can conclude that if the data is reliable, this we can confirm the max power rating of the chargers, charger station 11 has max rating of around 100+KW range, and charging station 1 has around 30KW+, and all remaining chargers might be having around 20KW range Which is based on the data. The charger 11 might be having rapid charging capability. The median power rates across chargers are relatively consistent, with slight variations. This could suggest common charger capabilities or a common set of charging behaviours among electric vehicle user.



Charging Duration by Charger ID





Presented here is a boxplot detailing the total duration of charging sessions for each charger, labeled by their respective IDs on the x-axis, against the total time in minutes on the y-axis. After detailed viewing of the plot, charger location 1 & 16 has least duration consistently having few comparatively fast or day charging behavior. These outlier like 850hrs of charging means error in data or may be the charger been connected for long in location even after the charging is completed might be due to user behavior. There are several instances of charging sessions with more than 25hrs of charging this might infer that the load/device may not be EV can be different or the charger location may be determined by the home charging setup, where drivers could charge overnight and leave their vehicle until they return.

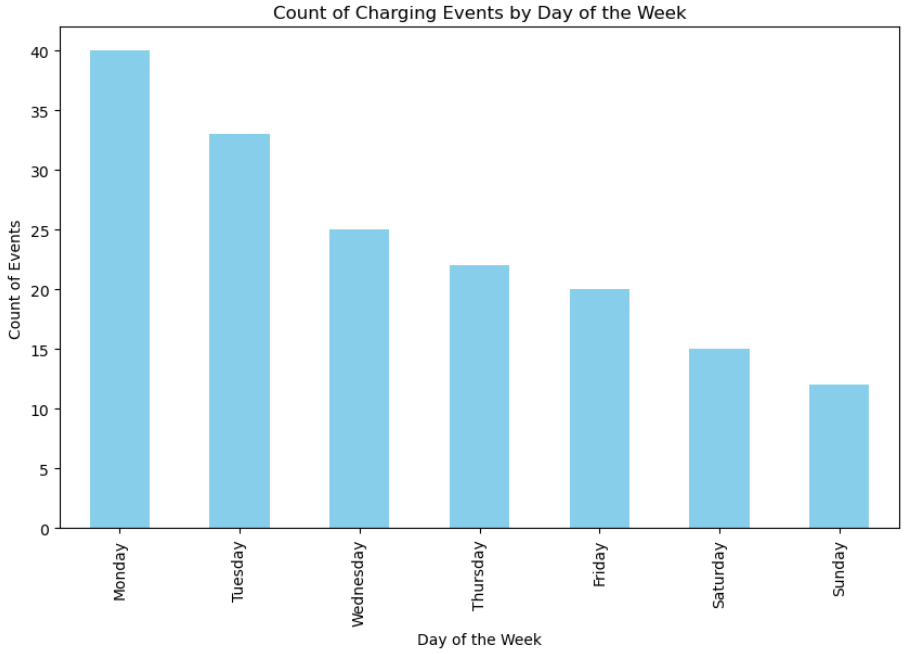
Count of Charging Events by Day of the Week

The bar chart demonstrates the number of charging events recorded for each day of the week. The x-axis categorizes the days from Monday to Sunday, While the y-axis indicates the count of events.

Monday shows the highest number of charging events, which might indicate a pattern Where users charge their vehicles in preparation for the worKWeek.

There is a noticeable decline in charging events as the week progresses, with a slight dip midweek on Wednesday. This could reflect a trend Where users are less likely to charge their vehicles during the middle of the week or that the need for recharging decreases after an initial start-of-week charge.

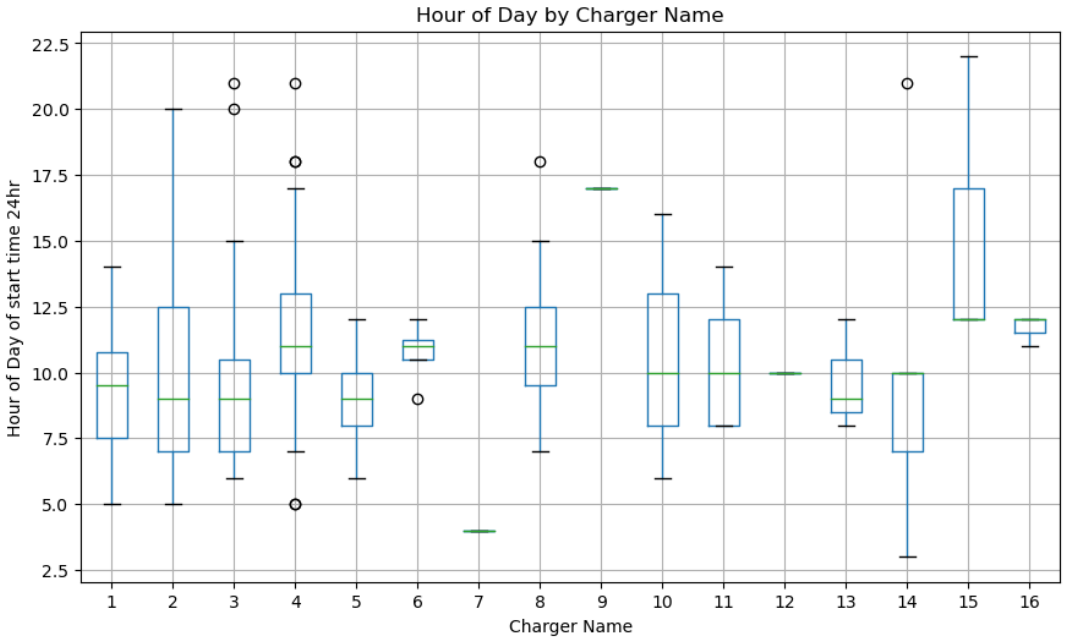
The weekends, particularly Sunday, show a significant decrease in charging events. This decline could be due to a variety of factors, such as reduced commuting or users' travel patterns, suggesting that less charging is required on days off.



Hour of Day by Charger Name

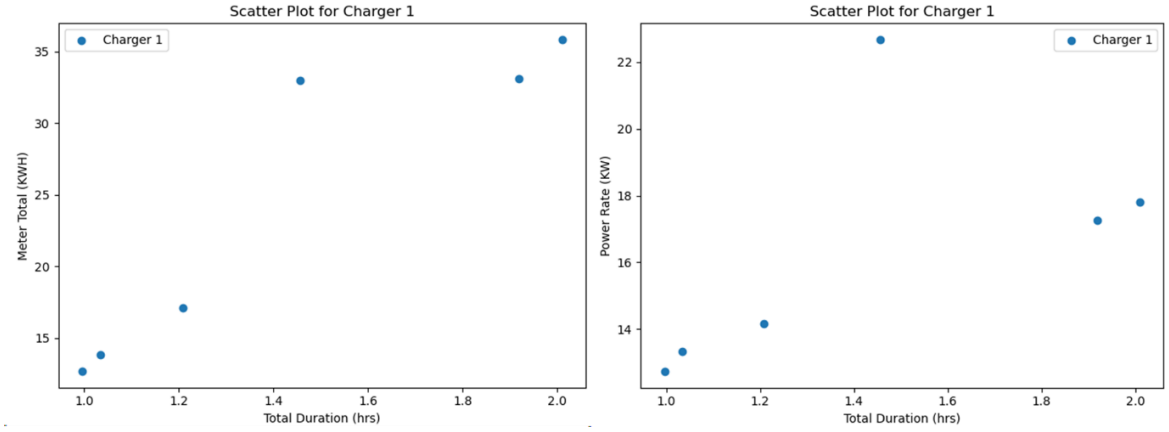
The boxplot displays the spread and median of charging times throughout the day for each charger, named or numbered along the x-axis. The y-axis represents hours of the day, ranging from the early morning to late night.

There is a large variance in the charging type/pattern, since the data for each location is less these box plots are skewed, the charger 11 for instance has overnight charging behavior as the start time of the charging is around evening. The charging locations 2 & 4 has wide spectrum of range for the starting time.



Charger 1:

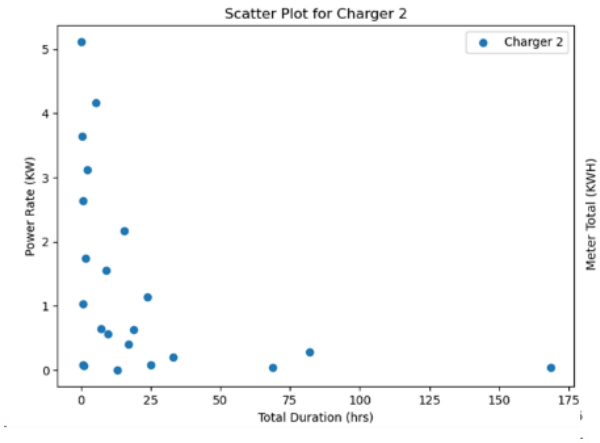
Comparing both Power rate & meter total for the charging instances of the charging location1, the charge duration is typically less than 2 hours and thus power rating is around 25KW range, these can be a commercial charging location spot apart from residential/home-based chargers, comparing the time of charging also. Exclusively based on the meter reading for the charger 1 we could find start meter reading were not cumulative means the location has multiple charging outlets

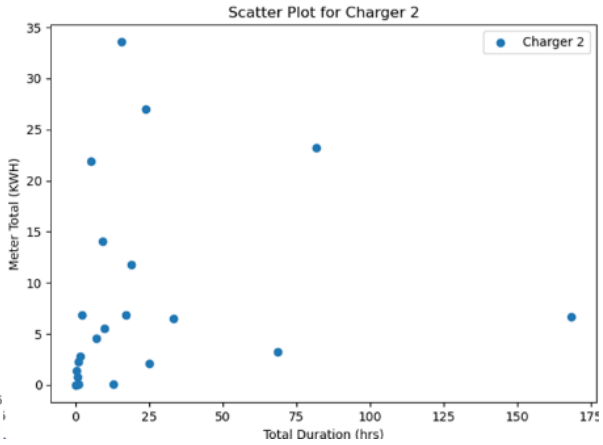


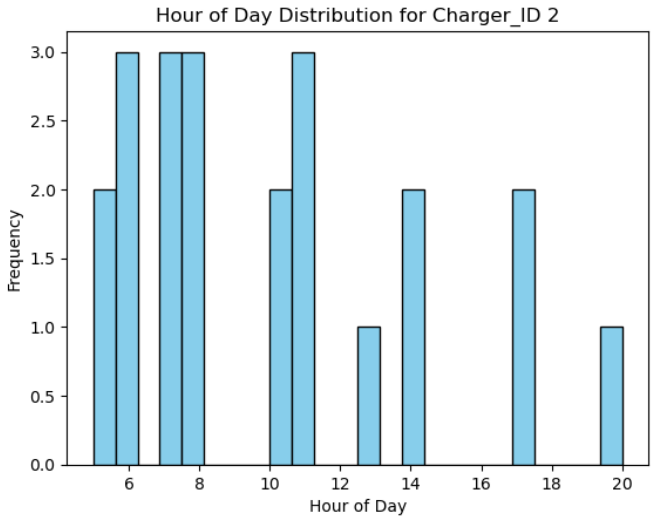
	Start Time	Meter Start (Wh)	Meter End(Wh)	Meter Total(Wh)	Total Duration (s)	Charger_name	Meter Total(KWH)	Total_duration(hrs)	End Time	Day of the Week	Hour of Day	Charger_ID	Power Rate (KW)
271	2019-05-09 10:55:00	0	12695.83	12695.83	3587	charger_1	12.69583	0.996389	2019-05-09 11:54:47	Thursday	10	1	12.741842
272	2019-09-09 05:47:00	0	33101.51	33101.51	6906	charger_1	33.10151	1.918333	2019-09-09 07:42:06	Monday	5	1	17.255348
273	2019-11-09 14:05:00	0	13807.38	13807.38	3726	charger_1	13.80738	1.035000	2019-11-09 15:07:06	Saturday	14	1	13.340464
274	2019-12-09 11:05:00	0	35804.92	35804.92	7234	charger_1	35.80492	2.009444	2019-12-09 13:05:34	Monday	11	1	17.818318
275	2019-09-16 07:17:00	0	32996.70	32996.70	5240	charger_1	32.99670	1.455556	2019-09-16 08:44:20	Monday	7	1	22.669489
276	2019-09-16 09:33:00	0	17109.95	17109.95	4350	charger_1	17.10995	1.208333	2019-09-16 10:45:30	Monday	9	1	14.159959

Charger 2:

Comparing both Power rate & Meter total for the charging instances of the charging location2, Majority of the instances Where short duration, typically in early morning charging, very less instances of over night charging, then typical charging location power rate is 35KW+

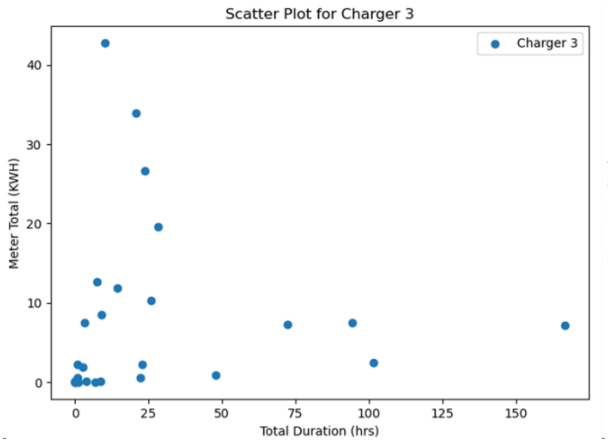


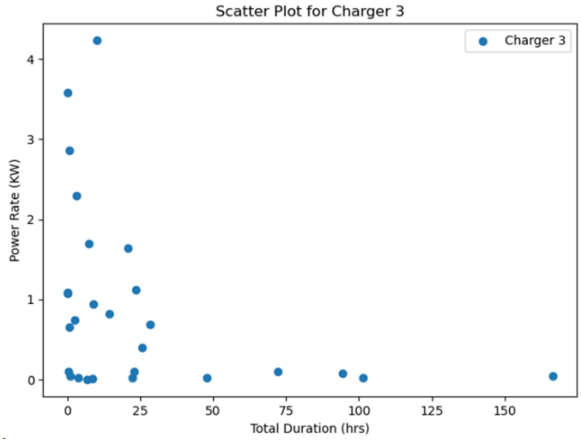


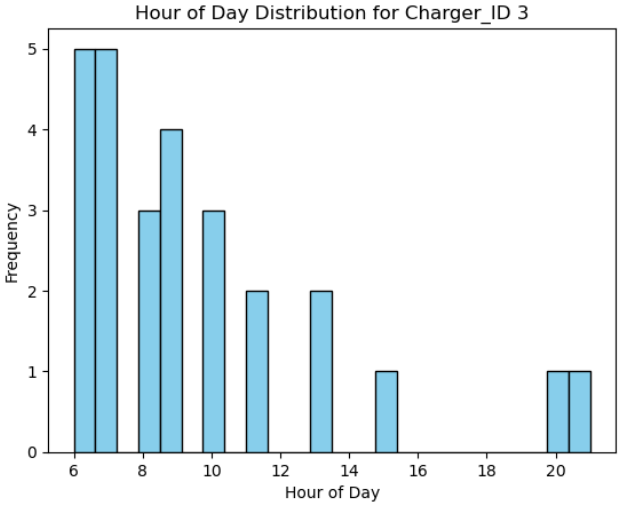


Charger 3:

Comparing both Power rate & meter total for the charging instances of the charging location3, the typical load is of mid-range EV less than 40KWh capacity or 3 Wheelers and majority are long duration charging & low power rating charger of around 5KW ~ 10KW. Typically start time of charging in morning

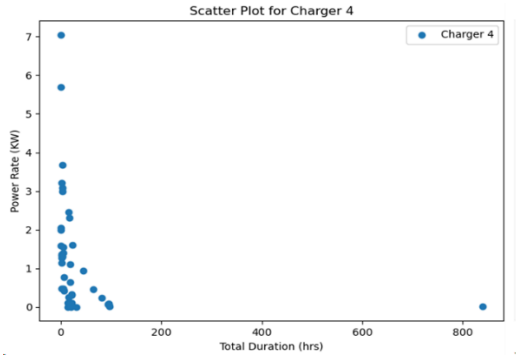


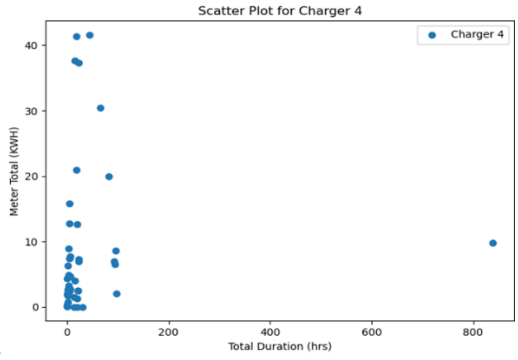


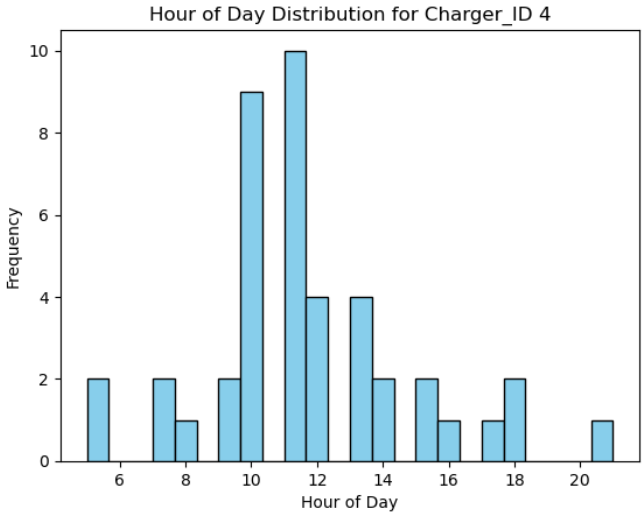


Charger 4:

Comparing both Power rate & meter total for the charging instances of the charging location 4, based on the charging duration it can be considered as slow or overnight charging behavior pattern.

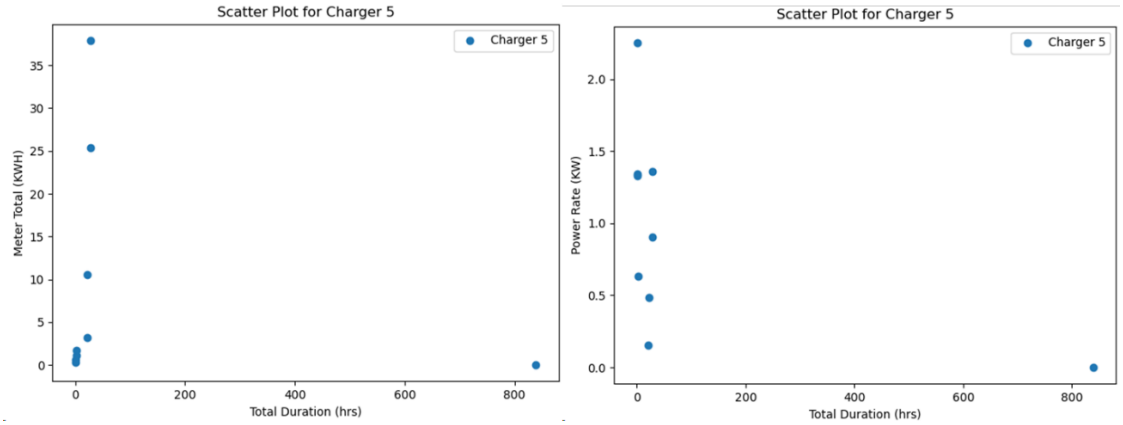






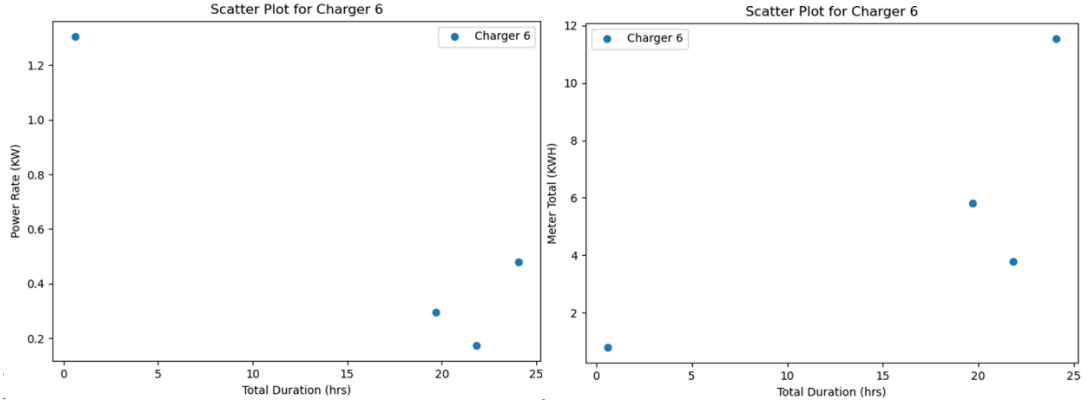
Charger 5:

Comparing both Power rate & meter total for the charging instances of the charging location5, we can infer the load demand for the charger is of average 25KWh range, may be mid-size or low range EV



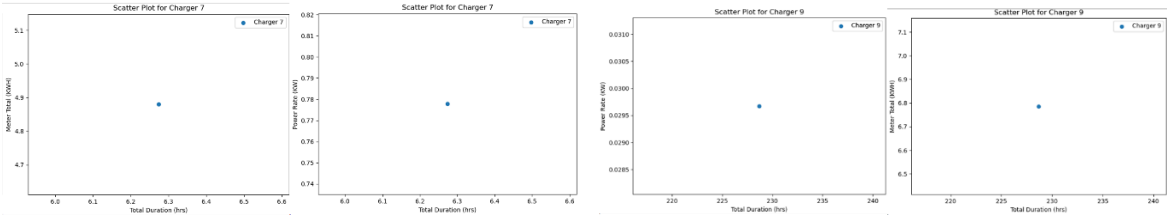
Charger 6:

Comparing both Power rate & meter total for the charging instances of the charging location 6, the load requirement of around 5KWh capacity may be an EV of 2-Wheeler or any household application.



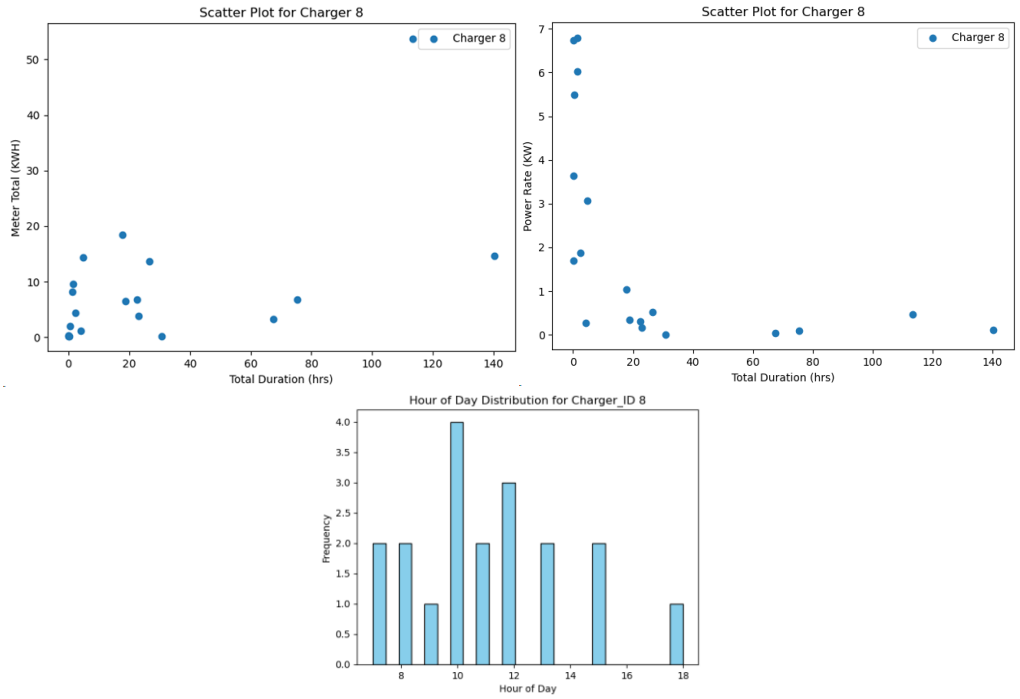
Charger 7 & Charger 9:

Based on the data at hand, it is not possible to draw any conclusion



Charger 8:

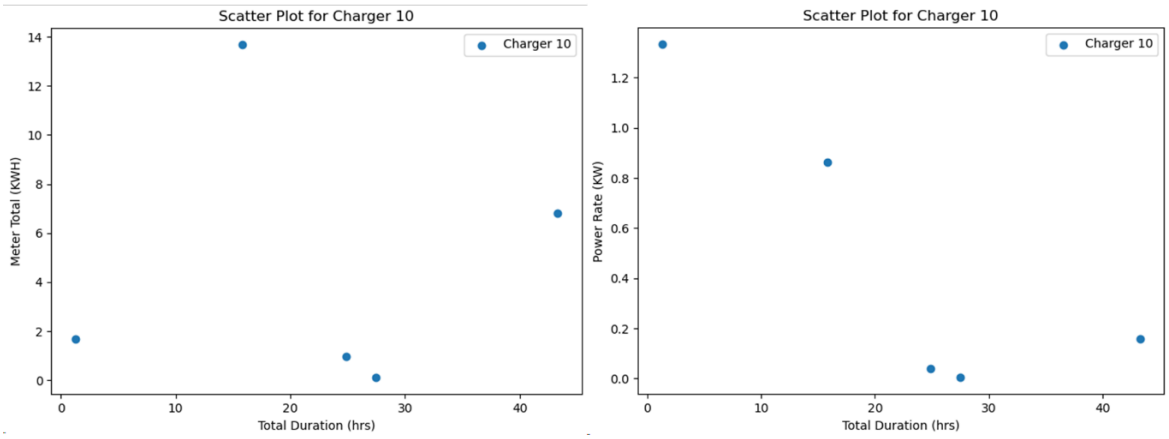
Comparing both Power rate & meter total for the charging instances of the charging location 8 ,



The load range is around 20KWh and majority of the instances are slow/overnight charging.

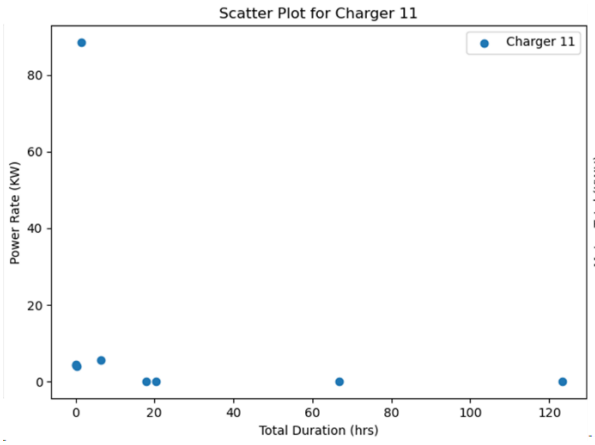
Charger 10:

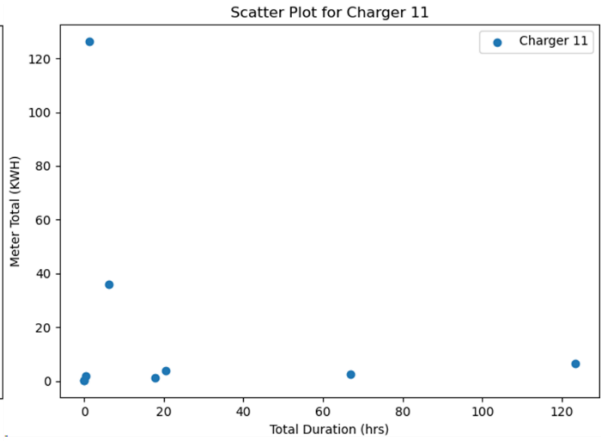
Comparing both Power rate & meter total for the charging instances of the charging location10, the pattern appears to be overnight/slow charging behavior with load range of 15KWh.



Charger 11:

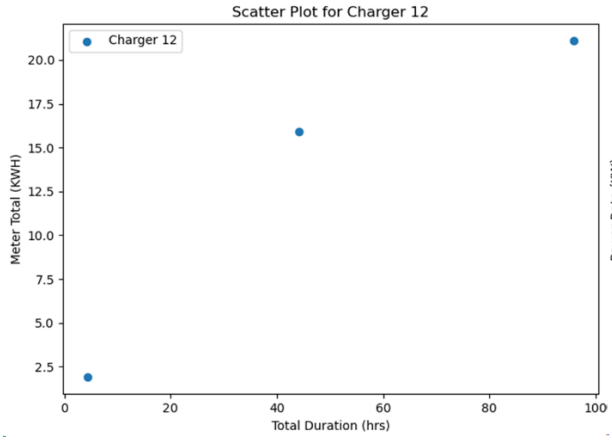
Comparing both Power rate & meter total for the charging instances of the charging location11, the pattern has both rapid and overnight charging behaviors, If the charging duration is accurate, we can verify Whether charger 11 has a power rating within the 120KW range.

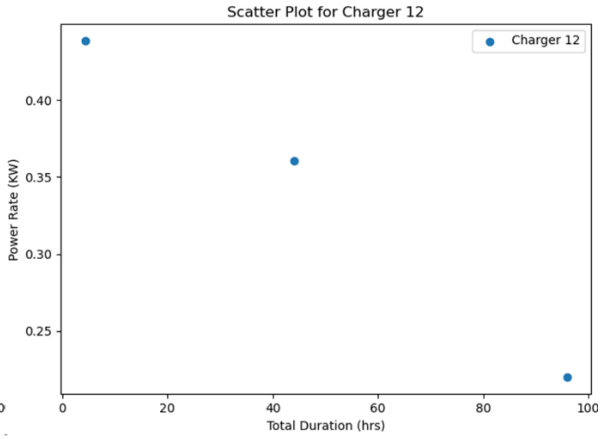




Charger 12:

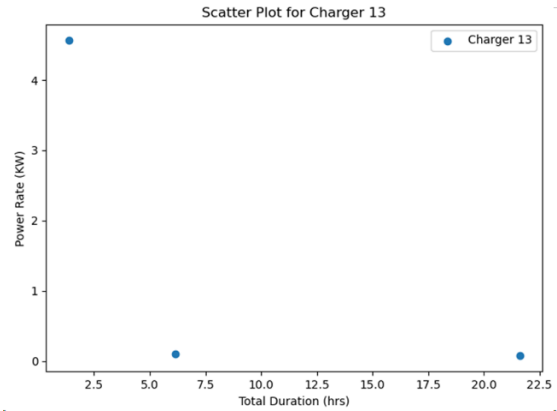
Comparing both Power rate & meter total for the charging instances of the charging location12, typically having long duration charging.

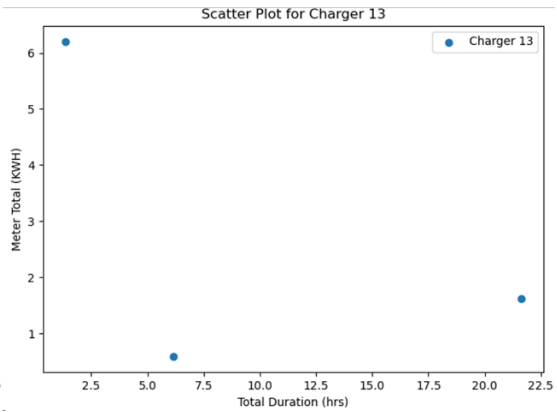




Charger 13:

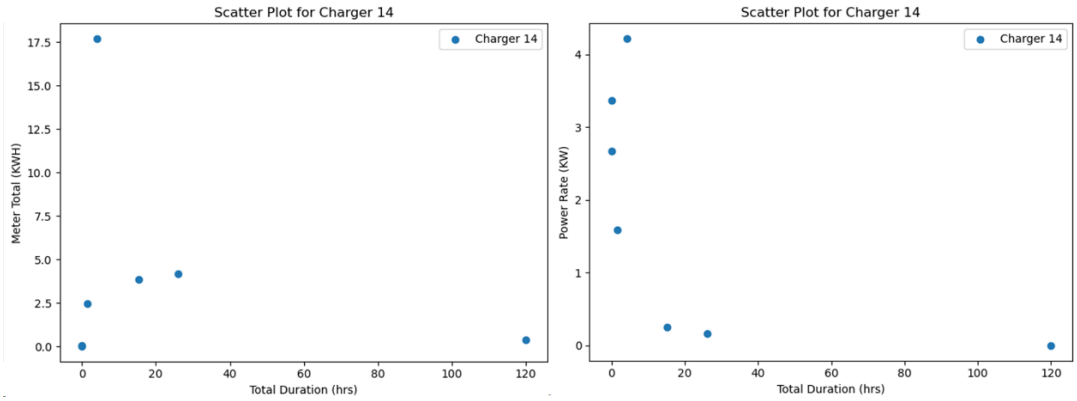
Comparing both Power rate & meter total for the charging instances of the charging location13. The scatter plots do not indicate a strong correlation between charging duration and power rate. The data points are sparse and do not form a clear pattern, suggesting that for Charger 13, the duration of charging does not consistently affect the rate at Which power is delivered





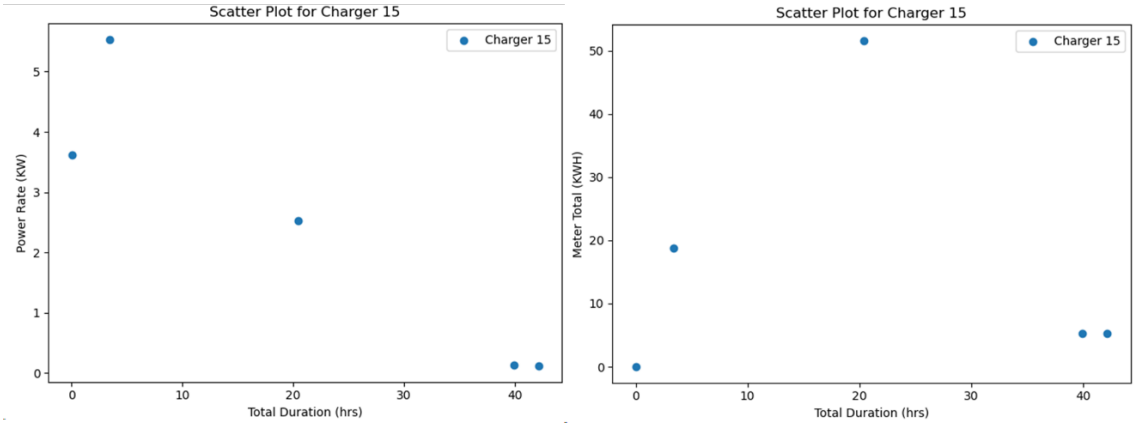
Charger 14:

Comparing both Power rate & meter total for the charging instances of the charging location14 having common patter 10hr + Which can be slow/ overnight based on total unit consumed



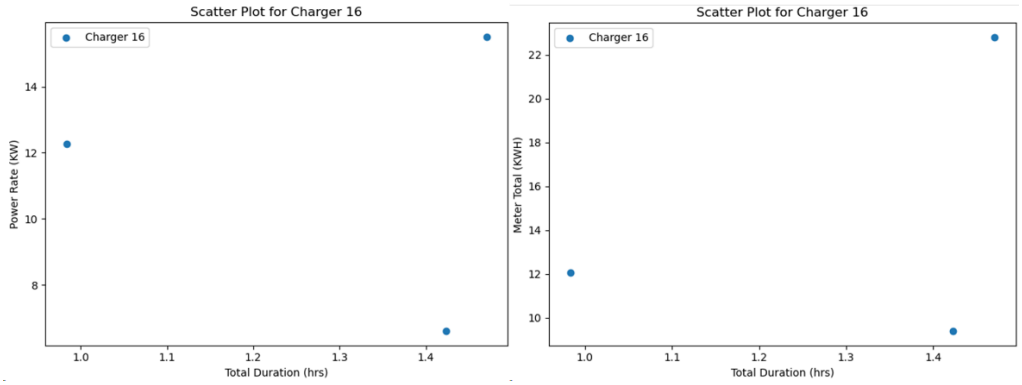
Charger 15:

Comparing both Power rate & meter total for the charging instances of the charging location15, the max charging load capacity were around 50KWh representing a midsize EV.



Charger 16:

Comparing both Power rate & meter total for the charging instances of the charging location16 we can infer with charging duration of around 2 hrs and total units charging pattern of rapid charging behavior, with power rating of 25KW+.



Final Insights

1. Charging Patterns: Monday exhibits the highest charging activity, suggesting a start-of-week charging trend, while weekend charging, especially on Sunday, decreases significantly, possibly due to less commuting the charging location may be commercial nonresidential charger.
2. Charging Duration and Energy Consumption: The duration of charging events averages around 33hrs minutes and median of 9 hrs overall & 75% of the instances within 24 hrs, but there is significant variability. Thus, most energy consumption is concentrated within a lower quartile, indicating short, quick charging sessions are common, although there are outliers indicating occasional higher consumption, possibly due to longer sessions or larger capacity vehicles.
3. Charger Utilization: Charger 4 displays the highest energy throughput, implying high utilization possibly due to its location or charger capabilities. In contrast, chargers like 1 and 16 have shorter average charging durations, indicating fast or day charging behaviors. Charging location 1 might have multiple chargers.
4. Time-of-Day Usage: The charging times vary significantly across chargers, with some indicating overnight charging patterns and others showing a broader range of operational hours.
5. Power Rate: The average power rate is 2.57 KW across charging sessions, with Charger 11 potentially offering rapid charging capabilities, evidenced by its high-power rate outliers. Majority of the charging have slow charging rate based on the units consumed and total duration.
6. Variability Across Chargers: Chargers show different patterns in terms of energy consumption, charging duration, and power rate, indicating a mix of charging behaviors and different type of vehicles capacities/load, average of 10KWh capacity & max of 125KWh.
7. Data Anomalies: Instances of unusually long charging durations, such as those over 25 hours, suggest possible data errors, non-EV loads, or varied user behaviors Where vehicles remain connected after charging is complete.
8. Charger Performance: A lack of strong correlation between charging duration and power rate for certain chargers suggests that the duration does not consistently affect the power delivery rate.

Conclusion

The exploratory data analysis of the charging events dataset has revealed nuanced behaviors in EV charging across various stations and times. It is evident that charging event patterns are influenced by a variety of factors, including the day of the week, the charger's location, and the time of day. The insights suggest opportunities for understanding requirement of charger type & charging pattern based on the user behavior for targeted infrastructure improvements, and data-driven operational decisions. Chargers with consistently high or rapid charging rates could be further optimized to serve the growing demand for quick charging options. The time pattern is not ideally captured with the time trend due to anomalies in the dataset, but due to charging capacity user prefer overnight charging/long hours charging optimized for their requirement. In conclusion, this analysis offers a foundational understanding of current charging trends and provides a data-driven approach to enhance the EV charging infrastructure.