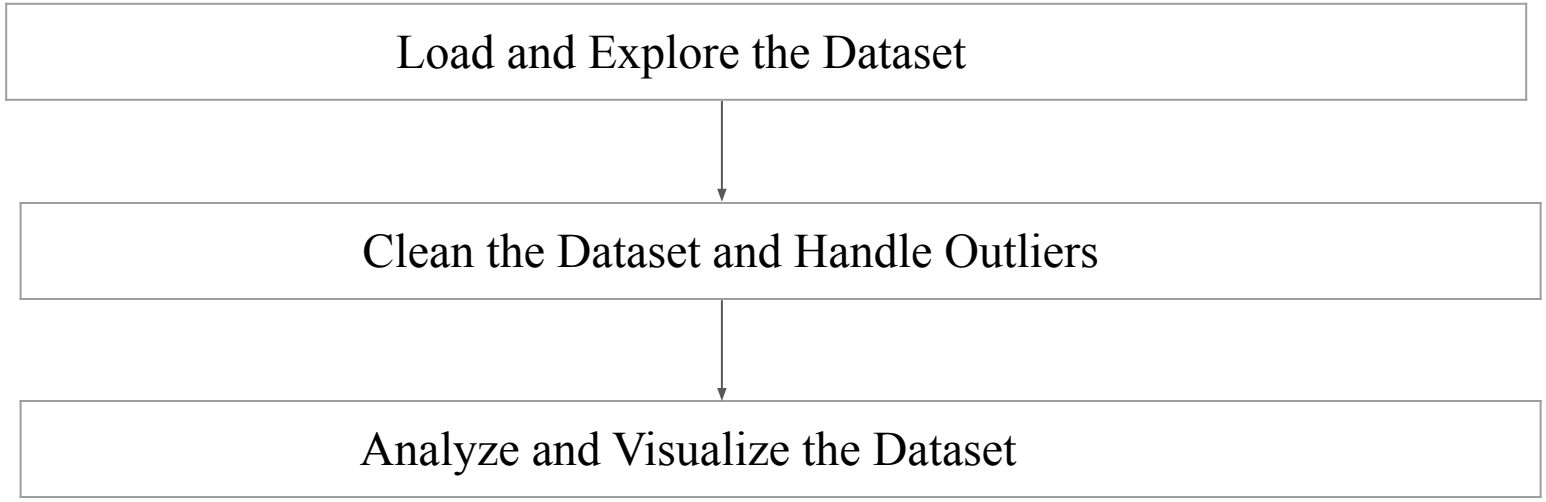
ELECTRIC VEHICLE ADOPTION AND CHARGING BEHAVIOUR ANALYSIS

.

ARUNIMA K
DA/DS
Date: 06/11/2024

Workflow


Load and Explore the Dataset
Clean the Dataset and Handle Outliers
Analyze and Visualize the Dataset

INTRODUCTION

- This analysis seeks to explore electric vehicle adoption and charging behavior patterns, offering insights to enhance EV infrastructure and guide strategic planning.

- The dataset provides details on charging sessions and user demographics, which will be analyzed through data understanding, preprocessing, and visualization to reveal meaningful trends.

DATA UNDERSTANDING

No:of Rows :1320

No:of Columns:20

	User ID	Vehicle Model	Battery Capacity (kWh)	Charging Station ID	Charging Station Location	Charging Start Time	Charging End Time	Energy Consumed (kWh)	Charging Duration (hours)	Charging Rate (kW)	Charging Cost (USD)	Time of Day	Day of Week	State of Charge (Start %)	State of Charge (End %)	Distance Driven (since last charge) (km)	Temperature (°C)	Vehicle Age (years)	Charger Type	User Type
0	User_1	BMW i3	108.463007	Station_391	Houston	2024-01-01 00:00:00	2024-01-01 00:39:00	60.712346	0.591363	36.389181	13.087717	Evening	Tuesday	29.371576	86.119962	293.602111	27.947953	2.0	DC Fast Charger	Commuter
1	User_2	Hyundai Kona	100.000000	Station_428	San Francisco	2024-01-01 01:00:00	2024-01-01 03:01:00	12.339275	3.133652	30.677735	21.128448	Morning	Monday	10.115778	84.664344	112.112804	14.311026	3.0	Level 1	Casual Driver
2	User_3	Chevy Bolt	75.000000	Station_181	San Francisco	2024-01-01 02:00:00	2024-01-01 04:48:00	19.128876	2.452653	27.513593	35.667270	Morning	Thursday	6.854604	69.917615	71.799253	21.002002	2.0	Level 2	Commuter
3	User_4	Hyundai Kona	50.000000	Station_327	Houston	2024-01-01 03:00:00	2024-01-01 06:42:00	79.457824	1.266431	32.882870	13.036239	Evening	Saturday	83.120003	99.624328	199.577785	38.316313	1.0	Level 1	Long-Distance Traveler
4	User_5	Hyundai Kona	50.000000	Station_108	Los Angeles	2024-01-01 04:00:00	2024-01-01 05:46:00	19.629104	2.019765	10.215712	10.161471	Morning	Saturday	54.258950	63.743786	203.661847	-7.834199	1.0	Level 1	Long-Distance Traveler

Summary Statistics

	User ID	Vehicle Model	Battery Capacity (kWh)	Charging Station ID	Charging Station Location	Charging Start Time	Charging End Time	Energy Consumed (kWh)	Charging Duration (hours)	Charging Rate (kW)	Charging Cost	Time of Day	Day of Week	State of Charge (Start %)	State of Charge (End %)	Last charge distance	Temperature (°C)	Vehicle Age (years)	Charger Type	User Type
count	1320	1320	1320.000000	1320	1320	1320	1320	1320.000000	1320.000000	1320.000000	1320.000000	1320	1320	1320.000000	1320.000000	1320.000000	1320.000000	1320.000000	1320	1320
unique	1320	5	NaN	462	5	1320	1309	NaN	NaN	NaN	NaN	4	7	NaN	NaN	NaN	NaN	NaN	3	3
top	User_1	Tesla Model 3	NaN	Station_108	Los Angeles	2024-01-01 00:00:00	2024-01-04 16:26:00	NaN	NaN	NaN	NaN	Evening	Saturday	NaN	NaN	NaN	NaN	NaN	Level 1	Commuter
freq	1	280	NaN	9	297	1	2	NaN	NaN	NaN	NaN	362	205	NaN	NaN	NaN	NaN	NaN	459	476
mean	NaN	NaN	74.534692	NaN	NaN	NaN	NaN	42.642894	2.269377	25.963003	22.551352	NaN	NaN	49.130012	75.141590	153.596788	15.263591	3.612843	NaN	NaN
std	NaN	NaN	20.626914	NaN	NaN	NaN	NaN	21.843792	1.061037	13.656279	10.751494	NaN	NaN	24.074134	17.080580	83.825619	14.831216	2.309824	NaN	NaN
min	NaN	NaN	1.532807	NaN	NaN	NaN	NaN	0.045772	0.095314	1.472549	0.234317	NaN	NaN	2.325959	7.604224	0.862361	-10.724770	0.000000	NaN	NaN
25%	NaN	NaN	62.000000	NaN	NaN	NaN	NaN	24.868084	1.397623	14.431294	13.368141	NaN	NaN	27.786903	62.053266	81.931571	2.800664	2.000000	NaN	NaN
50%	NaN	NaN	75.000000	NaN	NaN	NaN	NaN	42.642894	2.258136	25.963003	22.076360	NaN	NaN	48.241771	75.682496	153.596788	14.630846	4.000000	NaN	NaN
75%	NaN	NaN	85.000000	NaN	NaN	NaN	NaN	60.545496	3.112806	36.955838	31.646044	NaN	NaN	69.277921	88.201370	221.962330	27.981810	6.000000	NaN	NaN
max	NaN	NaN	193.003074	NaN	NaN	NaN	NaN	152.238758	7.635145	97.342255	69.407743	NaN	NaN	152.489761	177.708666	398.364775	73.169588	11.688592	NaN	NaN

Missing Values

User ID	0
Vehicle Model	0
Battery Capacity (kWh)	0
Charging Station ID	0
Charging Station Location	0
Charging Start Time	0
Charging End Time	0
Energy Consumed (kWh)	66
Charging Duration (hours)	0
Charging Rate (kW)	66
Charging Cost (USD)	0
Time of Day	0
Day of Week	0
State of Charge (Start %)	0
State of Charge (End %)	0
Distance Driven (since last charge) (km)	66
Temperature (°C)	0
Vehicle Age (years)	0
Charger Type	0

Missing Values Handling

- Data imputation using the median

```
df['Energy Consumed (kWh)'].fillna(df['Energy Consumed (kWh)'].median(), inplace=True)
df['Charging Rate (kW)'].fillna(df['Charging Rate (kW)'].median(), inplace=True)
df['Distance Driven (since last charge) (km)'].fillna(df['Distance Driven (since last charge) (km)'].median(), inplace=True)
```

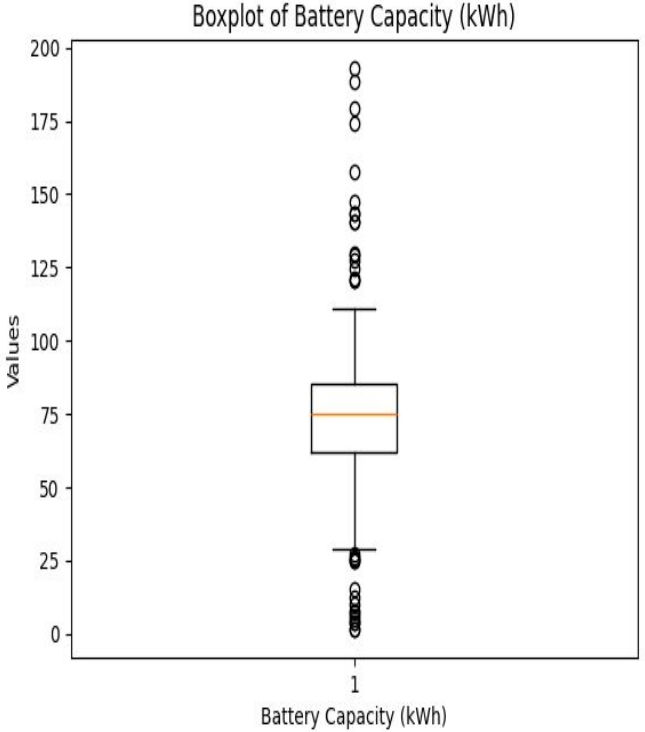
Selecting Relevant Columns

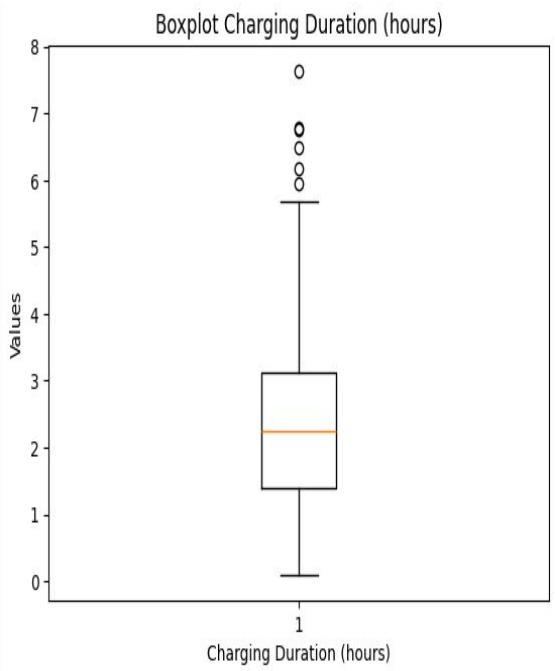
```
df=df[['User ID',
       #'Vehicle Model',
       'Battery Capacity (kWh)',
        #'Charging Station ID',
       #'Charging Station Location',
        'Charging Start Time',
       'Charging End Time', 'Energy Consumed (kWh)',
        'Charging Duration (hours)',
       'Charging Rate (kW)',
        'Charging Cost (USD)',
       'Time of Day', 'Day of Week',
        'State of Charge (Start %)',
       'State of Charge (End %)',
        'Distance Driven (since last charge) (km)',
       #'Temperature (°C)',
        'Vehicle Age (years)',
       'Charger Type',
       'User Type'].copy()]
df.head()
```

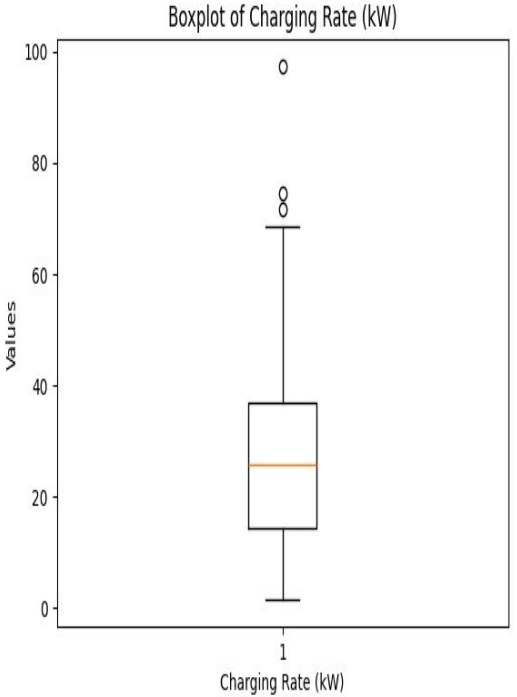
Shape of Dataset

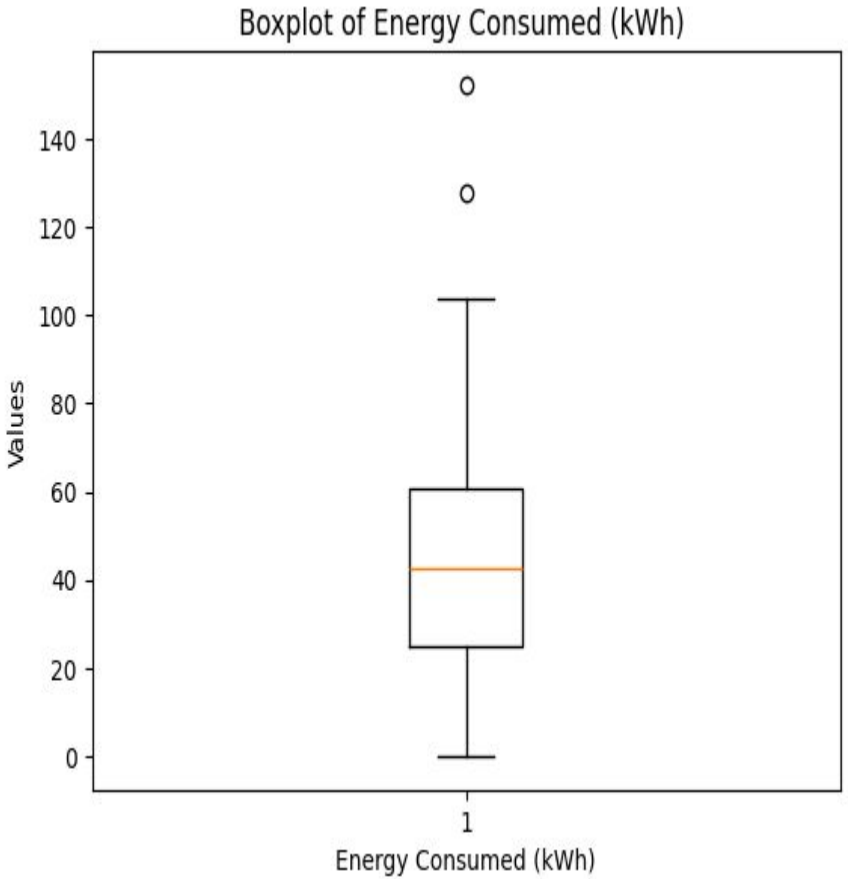
shape	Before	After
No:of Rows	1320	1320
No:of Columns	20	16

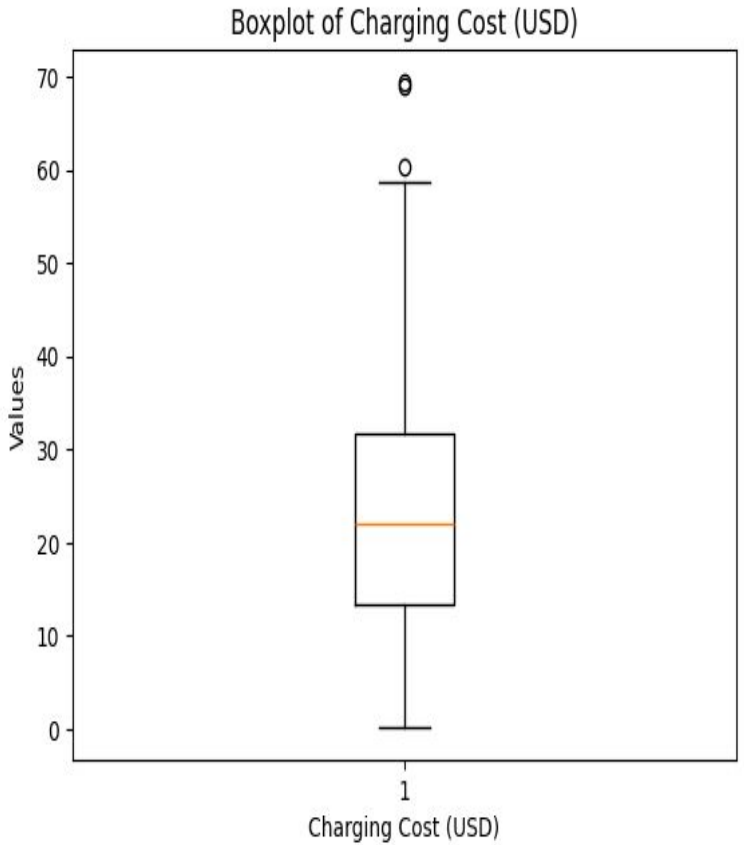
Outliers


Boxplot of Battery Capacity (kWh)
Values
200
175
150
125
100
75
50
25
0
1
Battery Capacity (kWh)


Boxplot Charging Duration (hours)
Values
8
7
6
5
4
3
2
1
0
1
Charging Duration (hours)


Boxplot of Charging Rate (kW)
Values
100
80
60
40
20
0
1
Charging Rate (kW)


Boxplot of Energy Consumed (kWh)
Values
0
20
40
60
80
100
120
140
1
Energy Consumed (kWh)


Boxplot of Charging Cost (USD)
Values
0
10
20
30
40
50
60
70
1
Charging Cost (USD)

IQR

1.Calculate the IQR:

$$IQR = Q3 - Q1$$

2.**Determine the outlier boundaries**:

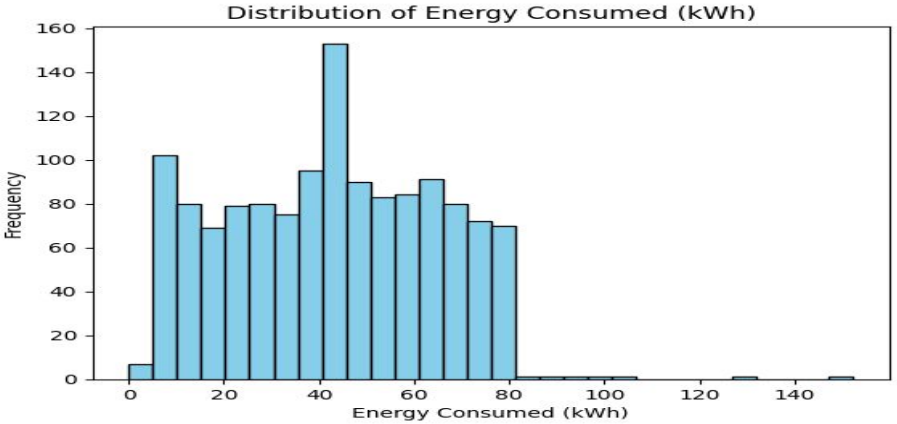
$$\text{Lower Bound} = Q1 - 1.5 \times IQR$$

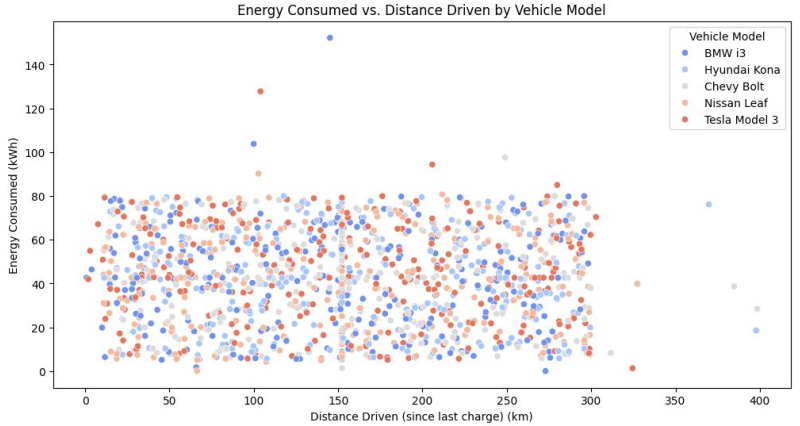
$$\text{Upper Bound} = Q3 + 1.5 \times IQR$$

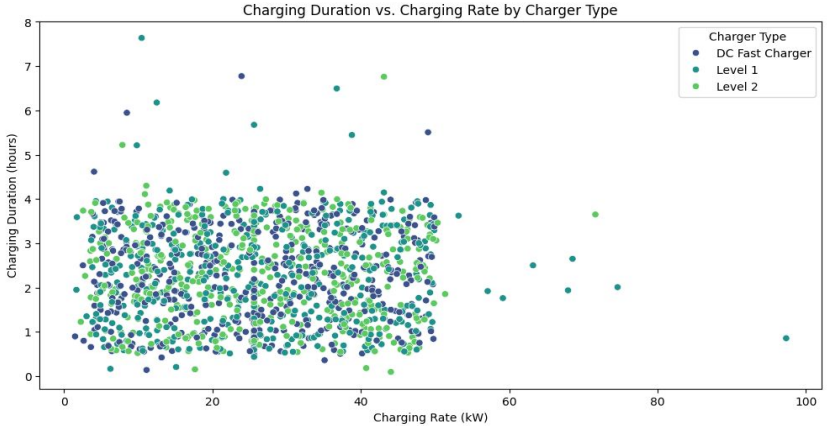
shape	Before	After
No:of Rows	1320	1317
No:of Columns	16	16

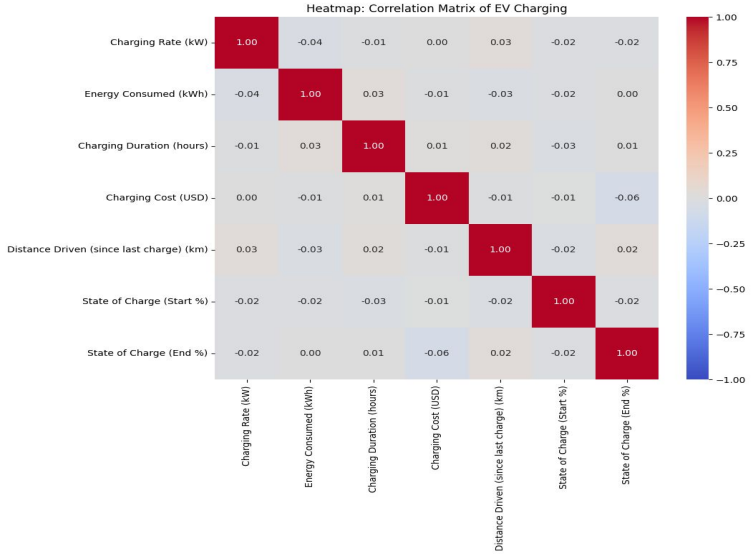
Data Visualization

- Histogram
- Scatterplot
- Heatmap


Distribution of Energy Consumed (kWh)
Frequency
Energy Consumed (kWh)


Energy Consumed vs. Distance Driven by Vehicle Model
Vehicle Model
BMW i3
Hyundai Kona
Chevy Bolt
Nissan Leaf
Tesla Model 3
Energy Consumed (kWh)
Distance Driven (since last charge) (km)


Charging Duration vs. Charging Rate by Charger Type
Charger Type
DC Fast Charger
Level 1
Level 2
Charging Duration (hours)
Charging Rate (kW)


Heatmap: Correlation Matrix of EV Charging
Charging Rate (kW)
Energy Consumed (kWh)
Charging Duration (hours)
Charging Cost (USD)
Distance Driven (since last charge) (km)
State of Charge (Start %)
State of Charge (End %)

Thank You..