ELECTRIC VEHICLE ANALYSIS AND CLASSIFICATION

Electric Vehicles (EVs): run fully or partly on electricity stored in batteries.

It is important because:

- Reduce use of petrol/diesel
- Cut air pollution
- More energy-efficient than normal cars
- Cheaper to run and maintain

However, understanding and classifying EVs based on their features (range, battery, type, etc.) is essential for research, consumer decision-making, and policy-making.

PROJECT OBJECTIVES

- **Analyze** the key features of Electric Vehicles (EVs).
- **Classify** EVs based on important parameters such as range, battery capacity, and type.
- **Provide insights** into EV market trends and categories for better understanding and decision-making.

PROBLEM STATEMENT

- More people are using EVs, so we need to study their performance and market trends.
- Without clear classification, it is hard for consumers to compare EVs (range, battery, charging time, price).
- Makers and government need proper insights to improve EVs and plan policies.
- Without such analysis, it becomes difficult for buyers, researchers, and policymakers to take the right decisions

METHODOLOGY

- **Data Collection** :

  From datasets, manufacturer websites, and government EV portals
- **Data Preprocessing** :

  Cleaning and normalizing data for analysis
- **Feature Selection** :

  Key features: battery size, motor power, range, charging time, price

- **Classification Algorithm** :

  Applied ML models Random Forest.

- **Evaluation Metrics :**

  Measured using accuracy, precision, recall

DATASET DESCRIPTION

- **Dataset:** Cheapestelectriccars-EVDatabase.csv
- **Features (columns):**
  - **Name** – Car name (e.g., Nissan Leaf, Tesla Model 3)
  - **Subtitle** – Vehicle type and battery capacity (e.g., Battery Electric Vehicle | 58 kWh)
  - **Acceleration** – Time taken to reach 0–100 km/h (performance indicator)
  - **TopSpeed** – Maximum speed of the vehicle (km/h)
  - **Range** – Distance the car can travel on a single charge (km)
  - **Efficiency** – Energy consumption (Wh/km), shows how efficient the car is
  - **FastChargeSpeed** – Charging speed when using fast chargers (km/h)
  - **Drive** – Type of drivetrain (Front Wheel Drive, All Wheel Drive)
  - **Number of Seats** – Seating capacity of the vehicle
  - **Price in Germany** – Cost of the vehicle in respective markets
  - **Price in UK-** Cost of the vehicle in respective markets

DATA PREPROCESSING

- Handling Missing Values : Filled empty cells with 0 to avoid errors in analysis.
- Extracting Manufacturer & Model : Separated car manufacturer and model name from the "Name" column.
- Cleaning Price Data : Removed currency symbols (€ / £) and commas. Converted prices into numeric format for analysis.
- Final Dataset : Clean, structured data with features ready for classification.

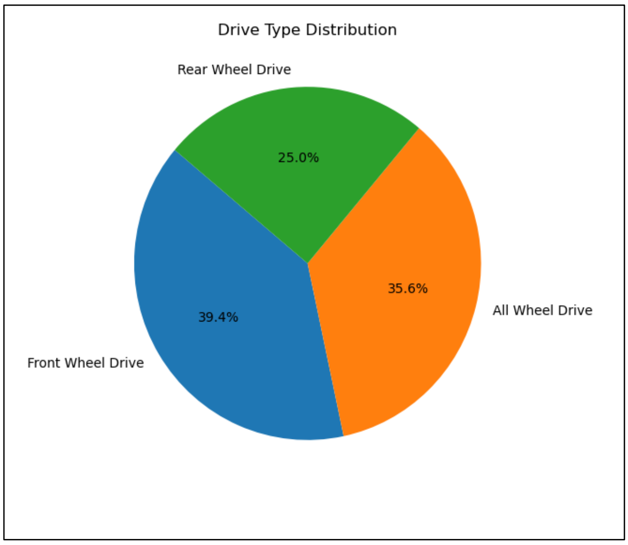
DATASET(BEFORE PREPROCESSING):

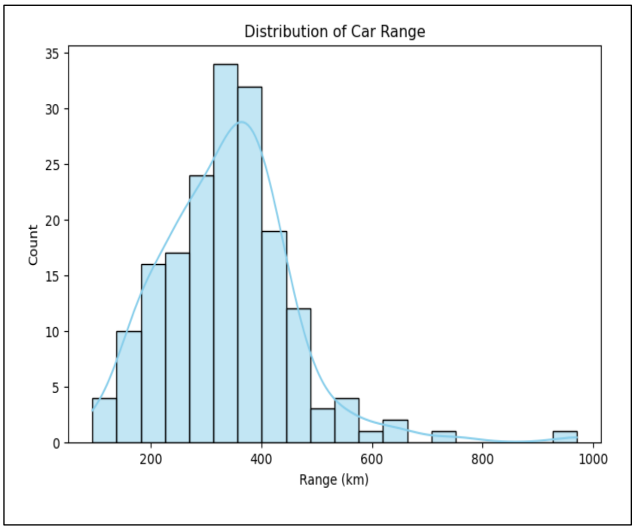
	Name	Subtitle	Acceleration	TopSpeed	Range	Efficiency	FastChargeSpeed	Drive	NumberofSeats	PriceinGermany	PriceinUK
0	Opel Ampera-e	Battery Electric Vehicle | 58 kWh	7.3 sec	150 km/h	335 km	173 Wh/km	210 km/h	Front Wheel Drive	5	€42,990	NaN
1	Renault Kangoo Maxi ZE 33	Battery Electric Vehicle | 31 kWh	22.4 sec	130 km/h	160 km	194 Wh/km	-	Front Wheel Drive	5	NaN	£31,680
2	Nissan Leaf	Battery Electric Vehicle | 36 kWh	7.9 sec	144 km/h	220 km	164 Wh/km	230 km/h	Front Wheel Drive	5	€29,990	£25,995
3	Audi e-tron Sportback 55 quattro	Battery Electric Vehicle | 86.5 kWh	5.7 sec	200 km/h	375 km	231 Wh/km	600 km/h	All Wheel Drive	5	NaN	£79,900
4	Porsche Taycan Turbo S	Battery Electric Vehicle | 83.7 kWh	2.8 sec	260 km/h	390 km	215 Wh/km	860 km/h	All Wheel Drive	4	€186,336	£138,830

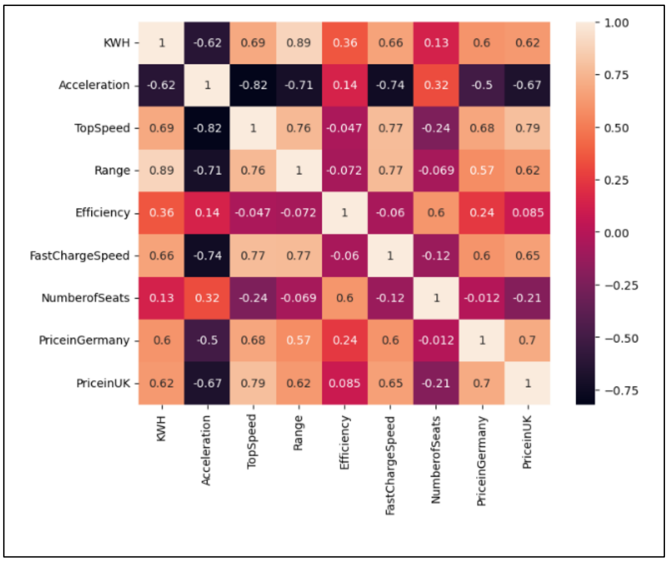
DATASET(AFTER PREPROCESSING):

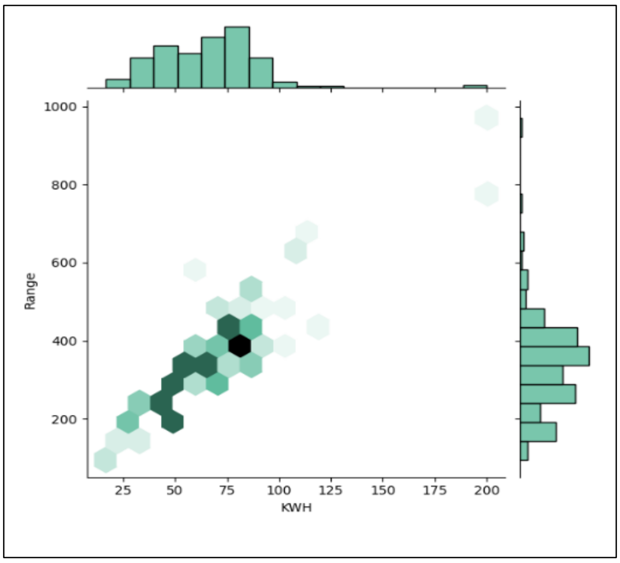
	Name	Subtitle	Acceleration	TopSpeed	Range	Efficiency	FastChargeSpeed	Drive	NumberofSeats	PriceinGermany	PriceinUK	Manufacturer	Model
0	Opel Ampera-e	Battery Electric Vehicle | 58 kWh	7.3 sec	150 km/h	335 km	173 Wh/km	210 km/h	Front Wheel Drive	5	42990	0	Opel	Ampera-e
1	Renault Kangoo Maxi ZE 33	Battery Electric Vehicle | 31 kWh	22.4 sec	130 km/h	160 km	194 Wh/km	-	Front Wheel Drive	5	0	31680	Renault	Kangoo Maxi ZE 33
2	Nissan Leaf	Battery Electric Vehicle | 36 kWh	7.9 sec	144 km/h	220 km	164 Wh/km	230 km/h	Front Wheel Drive	5	29990	25995	Nissan	Leaf
3	Audi e-tron Sportback 55 quattro	Battery Electric Vehicle | 86.5 kWh	5.7 sec	200 km/h	375 km	231 Wh/km	600 km/h	All Wheel Drive	5	0	79900	Audi	e-tron Sportback 55 quattro
4	Porsche Taycan Turbo S	Battery Electric Vehicle | 83.7 kWh	2.8 sec	260 km/h	390 km	215 Wh/km	860 km/h	All Wheel Drive	4	186336	138830	Porsche	Taycan Turbo S

VISUALIZATION



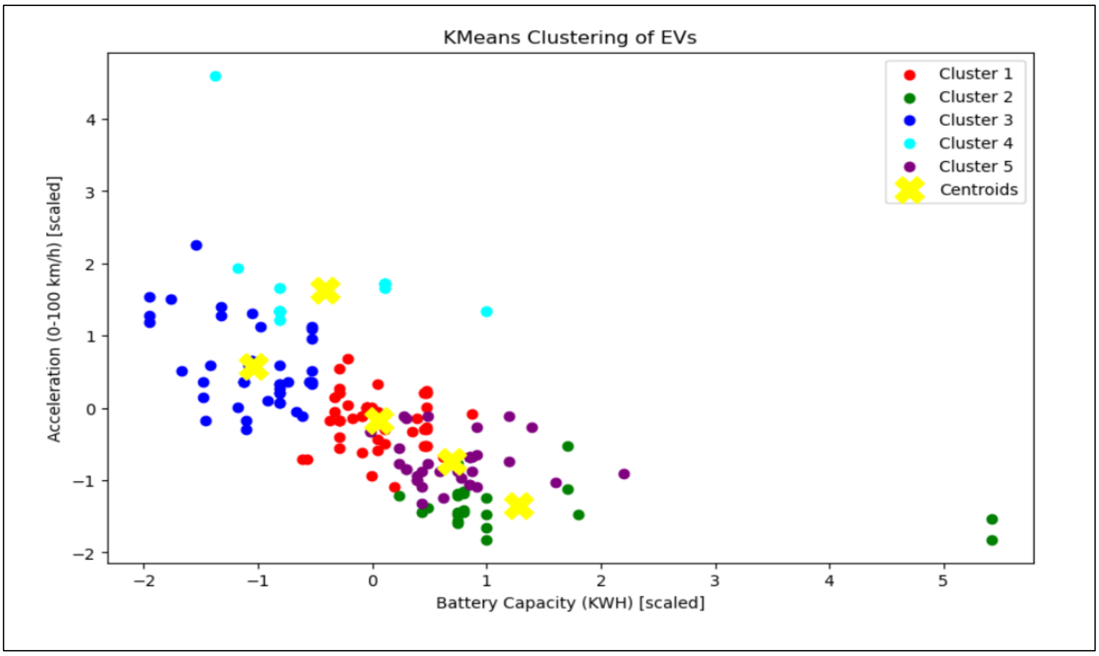






KMEANS CLUSTERING

- EVs grouped into **5 clusters** based on battery capacity & acceleration
- **Cluster 1 (Red) :** Moderate capacity & acceleration
- **Cluster 2 (Green) :** High capacity, slower acceleration
- **Cluster 3 (Blue) :** Small battery, better acceleration
- **Cluster 4 (Cyan) :** Outliers with high acceleration
- **Cluster 5 (Purple) :** Balanced, mid-range Evs
- **Centroids (Yellow X) :** Represent the “average EV” in each cluster



RANDOM FOREST ALGORITHM

- It makes **many decision trees**.
- Each tree looks at **different random data** and **different random features**.
- For prediction:

  Classification → Each tree gives a vote → the majority wins.

  Regression → Each tree gives a number → take the average.
- Many trees together = **better accuracy & less overfitting**.

RANDOM FOREST PREDICTION OF EV PRICES

- Built a Random Forest model to predict UK EV prices.
- Achieved $R^2 = 0.82$ (good accuracy).
- Average prediction error ~ £7k.
- Can be used for price estimation, comparison, and business planning.

```
#create a Random Forest Regressor model.
from sklearn.ensemble import RandomForestRegressor
# create regressor object
model = RandomForestRegressor(n_estimators = 300, random_state = 0)
# fit the regressor with x and y data
model.fit(X_train, y_train)
```

```
RandomForestRegressor
Parameters
```

```
#Mean Absolute Error(MAE)
from sklearn.metrics import mean_absolute_error
print("MAE: ",mean_absolute_error(y_test,y_pred))
```

```
MAE:  7165.767572718257
```

```
#Mean Squared Error(MSE)
from sklearn.metrics import mean_squared_error
print("MSE: ",mean_squared_error(y_test,y_pred))
```

```
MSE:  107966388.72466356
```

```
#Root Mean Squared Error(RMSE)
print("RMSE: ",np.sqrt(mean_squared_error(y_test,y_pred)))
```

```
RMSE:  10390.687596336613
```

```
#R Squared (R2)
from sklearn.metrics import r2_score
r2 = r2_score(y_test,y_pred)
print("R2: ",r2)
```

```
R2:  0.8158774004518805
```

RANDOM FOREST – DRIVE TYPE CLASSIFICATION

- Accuracy: 84%
- FWD predictions are most accurate (F1 = 0.94)
- AWD also strong (F1 = 0.86)
- RWD weaker (F1 = 0.62)
- Useful for classifying EVs by drive type from technical specs
-

Classification Report:

```
# Report
print(classification_report(y_test, y_pred))

                   precision    recall  f1-score   support

   All Wheel Drive       0.82      0.90      0.86        20
 Front Wheel Drive       0.94      0.94      0.94        16
  Rear Wheel Drive       0.71      0.56      0.62         9

          accuracy                           0.84        45
         macro avg       0.82      0.80      0.81        45
      weighted avg       0.84      0.84      0.84        45
```

OVERALL CONCLUSION

- EVs grouped into **5 market segments** using clustering.
- Drive type classification achieved **84% accuracy** (best for FWD).
- Price prediction model explained **82% of variation** with ~£7k–£10k error.
- Models are **reliable tools for segmentation, classification, and pricing decisions** in the EV market.