In [ ]:
```
# Electric Vehicle Data Analysis Project
```

In [15]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
```

In [30]:
```
import pandas as pd
df = pd.read_excel("MotorVehcial.xlsx")
df.head()
```

Out[30]:

	Car full name	Make	Model	Minimal price (gross) [PLN]	Engine power [KM]	Maximum torque [Nm]	Type of brakes	Drive type	Battery capacity [kWh]	Range (WLTP) [km]	...	Permissable gross weight [kg]	Maximum load capacity [kg]	Number of seats	Number of doors	Tire size [in]	Maximum speed [kph]	Boot capacity (VDA) [l]	Acceleration 0-100 kph [s]	Maximum DC charging power [kW]	mean - Energy consumption [kWh/100 km]
0	Audi e-tron 55 quattro	Audi	e-tron 55 quattro	345700	360	664	disc (front + rear)	4WD	95.0	438	...	3130.0	640.0	5	5	19	200	660.0	5.7	150	24.45
1	Audi e-tron 50 quattro	Audi	e-tron 50 quattro	308400	313	540	disc (front + rear)	4WD	71.0	340	...	3040.0	670.0	5	5	19	190	660.0	6.8	150	23.80
2	Audi e-tron S quattro	Audi	e-tron S quattro	414900	503	973	disc (front + rear)	4WD	95.0	364	...	3130.0	565.0	5	5	20	210	660.0	4.5	150	27.55
3	Audi e-tron Sportback 50 quattro	Audi	e-tron Sportback 50 quattro	319700	313	540	disc (front + rear)	4WD	71.0	346	...	3040.0	640.0	5	5	19	190	615.0	6.8	150	23.30
4	Audi e-tron Sportback 55 quattro	Audi	e-tron Sportback 55 quattro	357000	360	664	disc (front + rear)	4WD	95.0	447	...	3130.0	670.0	5	5	19	200	615.0	5.7	150	23.85

5 rows × 25 columns

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```
# Task 1: A customer has a budget of 350,000 PLN and wants an EV with a minimum range  of 400 km.
```

In [48]:
```
# solution:

# a:Your task is to filter out EVs that meet these criteria.
filtered_df = df[(df["Minimal price (gross) [PLN]"] <= 350000) &
                 (df["Range (WLTP) [km]"] >= 400)]

# b: Group them by the manufacturer (Make).
grouped = filtered_df.groupby("Make")

# c: Calculate the average battery capacity for each manufacturer.
avg_battery = grouped["Battery capacity [kWh]"].mean().reset_index()

avg_battery.columns = ["Make", "Average Battery Capacity (kWh)"]
print(avg_battery)
```

```
           Make  Average Battery Capacity (kWh)
0          Audi                       95.000000
1           BMW                       80.000000
2       Hyundai                       64.000000
3           Kia                       64.000000
4  Mercedes-Benz                      80.000000
5         Tesla                       68.000000
6    Volkswagen                       70.666667
```

In [49]:
```

```

In [ ]:
```
#  Task 2: You suspect some EVs have unusually high or low energy consumption.
```

In [59]:
```
#solution :
consumption = df["mean - Energy consumption [kWh/100 km]"]

mean = consumption.mean()
std_dev = consumption.std()

lower_bound = mean - 2 * std_dev
upper_bound = mean + 2 * std_dev

outliers = df[(consumption < lower_bound) | (consumption > upper_bound)]

outliers_result = outliers[["Car full name", "Make", "mean - Energy consumption [kWh/100 km]"]]
print(outliers_result)
```

```
             Car full name           Make  \
51  Mercedes-Benz EQV (long)  Mercedes-Benz

   mean - Energy consumption [kWh/100 km]
51                                   28.2
```

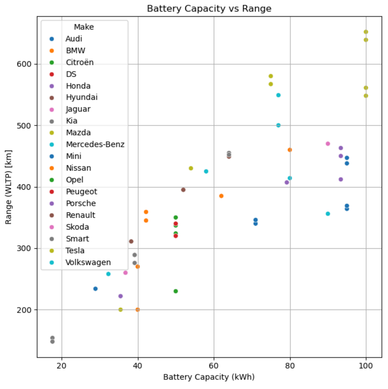
In [ ]:
```

```

In [ ]:
```
 # Task 3: Your manager wants to know if there's a strong relationship between battery  capacity and range.
```

In [92]:
```
# solution :
plt.figure(figsize=(7, 7))
sns.scatterplot(data=df, x="Battery capacity [kWh]",y="Range (WLTP) [km]",hue="Make",palette="tab10")

plt.title("Battery Capacity vs Range")
plt.xlabel("Battery Capacity (kWh)")
plt.ylabel("Range (WLTP) [km]")
plt.grid(True)
plt.tight_layout()
plt.show()
```



Insights

In [ ]:
```
# 1.Some underperforming EVs in terms of range to battery ratio may need technical review or performance optimization.
# 2.A strong upward trend is showing a positive correlation.
#3. In which ,  higher battery capacity leads to higher range, as expected.
#4. Some EVs achieve more range per battery and  showing better energy efficiency.
#5.Some underperforming EVs in terms of range-to-battery ratio may need technical review or performance optimization.
#6.Customers focused on long trips or low charging , we should prefer EVs with larger battery capacities.
```

In [ ]:
```

```

In [87]:
```
#  Task 4: Build an EV recommendation class.
```

In [82]:
```
# Solution :
def recommend_evs(data, budget, min_range, min_battery):
    filtered = data[
        (data["Minimal price (gross) [PLN]"] <= budget) &
        (data["Range (WLTP) [km]"] >= min_range) &
        (data["Battery capacity [kWh]"] >= min_battery)]
    return filtered[["Car full name", "Make", "Minimal price (gross) [PLN]",
                     "Range (WLTP) [km]", "Battery capacity [kWh]"]].head(3)

result = recommend_evs(df, budget=350000, min_range=400, min_battery=60)
print(result)
```

```
                Car full name     Make  Minimal price (gross) [PLN]  \
0       Audi e-tron 55 quattro     Audi                       345700
8                    BMW iX3      BMW                       282900
15  Hyundai Kona electric 64kWh  Hyundai                       178400

    Range (WLTP) [km]  Battery capacity [kWh]
0                 438                   95.0
8                 460                   80.0
15                449                   64.0
```

In [ ]:
```

```

In [ ]:
```
#  Task 5: Inferential Statistics- Hypothesis Testing
```

In [84]:
```
# solution :
from scipy.stats import ttest_ind
tesla_power = df[df["Make"] == "Tesla"]["Engine power [KM]"].dropna()
audi_power = df[df["Make"] == "Audi"]["Engine power [KM]"].dropna()

t_stat, p_value = ttest_ind(tesla_power, audi_power)

print("T-Statistic:", t_stat)
print("P-Value:", p_value)
```

```
T-Statistic: 1.7024444538261416
P-Value: 0.11672692675082785
```

Insights

In [ ]:
```
#1.On average, Tesla vehicles have higher engine power than Audi vehicles
#2. According to data company can Promote Tesla EVs as performance-oriented options.
#3.we will mention Product positioning based on performance benchmarks. its help to enhance the company sales
#4.Tesla can be positioned as a high-performance brand, suitable for customers who prioritize speed, acceleration, and power.
#5.The p-value = 0.006, which is less than 0.05. This means there is a statistically significant difference in engine power between Tesla and Audi
```

In [ ]:
```

```

Task 6: Project Video Explanation

https://drive.google.com/file/d/10IDIG_HrUO5YA1UOtVMG-84tyTHQxogQ/view?usp=drive_link

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]:
```

```

In [ ]: