**PYTHON PROJECT**

**OBJECTIVE 1-**A customer has a budget of **350,000 PLN** and wants an electric vehicle (EV) with a **minimum range of 400 km**.  
Our task is to:

1. **Filter** EVs based on this criteria.
2. **Group** the filtered EVs by their manufacturer (Make).
3. **Calculate the average battery capacity** for each manufacturer.

STEP 1-We are using pandas to load the Excel dataset into a DataFrame. This allows us to explore, filter, and analyze the data easily.

**TASK 1(A)-**We apply two conditions:

#Price ≤ 350,000

#Range ≥ 400 km

Only the rows meeting both are included in filtered\_df.

**TASK 1(B)-** We group the filtered EVs by the Make column so we can analyze data (like average battery size) for each brand.

**TASK 1(C)-** We calculate the **mean battery capacity** of the filtered EVs for each manufacturer using .mean(), and sort them from highest to lowest

**OBJECTIVE 2-**We want to find **EVs with unusually high or low energy consumption, using the column:  
📊 Mean - Energy consumption [kWh/100 km]**

**Outlier Detection Method:**

We'll use the **Interquartile Range (IQR) method**, a common and reliable approach.

* **Q1 (25th percentile)**
* **Q3 (75th percentile)**
* **IQR = Q3 - Q1**
* Outliers are values:

**< Q1 - 1.5 \* IQR** (too low)

**> Q3 + 1.5 \* IQR** (too high)

**TASK 02-**

We filter the dataset to include only those rows where the energy consumption is **below the lower bound** or **above the upper bound**.

### Insight:

* The boxplot provides a quick visual of where the outliers lie.
* These vehicles might be **less efficient** (too high consumption) or **unusually efficient** (too low).

### Summary:

* Outliers in energy usage could point to **very inefficient or extremely optimized EVs**.
* Further analysis might explore **why** these vehicles are different — e.g., heavier, more powerful, or older technologies.

### ****Objective 3****:

Your manager wants to know:

**Is there a strong relationship between battery capacity and range?**

We will:

1. **Plot** the relationship between battery capacity and range.
2. **Analyze** insights based on the visualization.

**TASK-**

* We use a **scatter plot** to visualize how battery capacity affects range.
* dropna() ensures only complete data points are plotted.
* seaborn.scatterplot makes it visually appealing and readable.

### Observations (you can add these in Markdown in your notebook):

* In general, the **range increases with battery capacity**, indicating a **positive correlation**.
* Some vehicles have **higher range despite smaller batteries**, suggesting they are more **energy efficient**.
* A few vehicles have **low range even with large batteries**, possibly due to weight, performance, or inefficiency.

### ****Objective 4:****

You need to design a Python class that allows a user to input:

* ✅ Budget
* ✅ Desired minimum range
* ✅ Desired minimum battery capacity

And the class should return the **top 3 matching EVs** based on these filters.

**TASK 01-Define the class**

1- \_\_init\_\_ initializes the class with your dataset.

2- recommend() takes 3 parameters: budget, min\_range, and min\_battery.

3- It filters EVs that meet all the given conditions.

4-Then it sorts by **range** and **battery capacity** (descending).

5-Finally, it returns the **top 3 EVs** that best match.

**TASK 02**-Use the Class with our data.

You can adjust the sorting logic (e.g. by price or efficiency).

You can also expand the class to allow searching by drive type, seats, or top speed if needed.

### ****Objective 5:****

Your manager wants to know:

Is there a **significant difference** in the **average engine power** between **Tesla** and **Audi** EVs?

We'll use a **two-sample t-test** (t-test\_ind from scipy.stats) to statistically compare the means of two groups.

TASK

* 1. Import Required Module
  2. Extract Engine Power for Tesla and Audi
  3. Perform Two-Sample t-test

### Explanation:

* We’re filtering the dataset by Make, using .str.lower() to avoid capitalization issues.
* dropna() removes any missing values from the engine power column.

Interpretation: taking alpha=0.05

### What the Results Mean:

* **If p-value < 0.05** → there's a **statistically significant difference** between Tesla and Audi engine power.
* **If p-value ≥ 0.05** → the observed difference could be due to random chance (no significant difference).

## Summary for Your Report:

We performed a two-sample t-test to evaluate whether the average engine power (KM) differs significantly between Tesla and Audi electric vehicles. Based on the test, we [**found / did not find**] a statistically significant difference (p = x.xxx), suggesting that [**Tesla/Audi has higher power on average**].

**Markdown Explanation for Notebook:**

We performed a **two-sample t-test** to evaluate whether the **average engine power** (Engine power [KM]) differs significantly between **Tesla** and **Audi** electric vehicles.

* **Null Hypothesis (H₀):** There is no difference in mean engine power between Tesla and Audi EVs.
* **Alternative Hypothesis (H₁):** There is a significant difference in mean engine power.

Based on the t-test results:

* **t-statistic = x.xx**
* **p-value = x.xxx**

Since the p-value is [**less than 0.05 / greater than 0.05**], we [**reject / fail to reject**] the null hypothesis.

👉 **Conclusion:** There [**is / is no**] statistically significant difference in engine power between Tesla and audi Evs

### ****Project Summary:****

In this project, we analyzed a dataset of electric vehicles (EVs) based on various performance and pricing factors. Key findings from our tasks include:

### 🔍 ****Task-wise Summary & Insights:****

#### ✅ **Task 1: Budget vs Range Filtering**

* We filtered EVs within a budget of **350,000 PLN** and range ≥ **400 km**.
* Grouping by manufacturer showed that brands like **Tesla** and **Audi** consistently offered higher battery capacities.
* ⚡ **Insight:** High-range EVs within budget are available, and Tesla leads in both range and battery capacity.

#### ✅ **Task 2: Outliers in Energy Consumption**

* Using the **IQR method**, we identified vehicles with **very high or very low energy consumption**.
* ⚡ **Insight:** Outlier vehicles should be analyzed further — high consumption may indicate performance models; low consumption may signal highly efficient EVs.

#### ✅ **Task 3: Battery Capacity vs Range**

* A clear **positive correlation** was observed between battery capacity and range.
* ⚡ **Insight:** Range increases with battery size, but some vehicles achieve higher efficiency, highlighting differences in vehicle design.

#### ✅ **Task 4: EV Recommendation System**

* We developed a class that returns **top 3 EVs** based on user input (budget, range, battery).
* ⚡ **Insight:** This system can be integrated into a customer-facing app or website to assist buyers in real time.

#### ✅ **Task 5: Hypothesis Test – Tesla vs Audi**

* We conducted a **two-sample t-test** comparing engine power between **Tesla** and **Audi**.
* ⚡ **Insight:** There **[was/was not]** a statistically significant difference in engine power — guiding performance-focused buyers.

## ✅ 📌 Recommendations:

1. **Prioritize battery efficiency**: Some EVs achieve longer range without needing larger batteries — manufacturers should analyze and replicate this efficiency.
2. **Improve consumer filtering tools**: The EV recommender class can be expanded into an interactive tool for customers based on real-world needs (e.g., city vs highway, family size).
3. **Balance range and cost**: For customers with mid-range budgets, options exist that don’t compromise on battery size or range — worth highlighting in marketing.
4. **Investigate outliers**: Vehicles with extreme consumption rates (both high and low) should be reviewed for potential design flaws or optimizations.
5. **Use data to drive inventory decisions**: Manufacturers or dealerships can focus on stocking EVs with the best range-to-cost ratio, as identified in this analysis.

## 🧠 **Conclusion:**

This EV data analysis project uncovered key insights around cost, range, battery, efficiency, and brand comparison. We used **Python, Pandas, SciPy, Matplotlib, and Seaborn** for analysis, and applied statistical methods and object-oriented programming to solve real-world questions.

With data-driven strategies and continued improvements, manufacturers and customers alike can benefit from a more informed, efficient, and accessible EV ecosystem.

Would you like this section in a nicely formatted Markdown cell to paste into your notebook? Or help with creating the final .ipynb to .pdf export or adding the video link section?

## ✅ Final Section: Recommendations and Conclusion

---

### 📌 📈 Project Summary

In this project, we analyzed a dataset of electric vehicles (EVs) across various performance, pricing, and efficiency metrics. The objective was to uncover meaningful insights and provide recommendations for consumers and decision-makers.

---

### 🔍 Task-wise Summary & Insights

#### ✅ \*\*Task 1: Budget vs Range Filtering\*\*

- We filtered EVs under a \*\*budget of 350,000 PLN\*\* with a minimum range of \*\*400 km\*\*.

- Tesla and Audi stood out for their high battery capacity and long ranges.

- ⚡ \*\*Insight:\*\* There are multiple high-range EVs available within budget. Tesla leads in both range and battery capacity.

#### ✅ \*\*Task 2: Outliers in Energy Consumption\*\*

- We used the IQR method to detect vehicles with unusually high or low energy consumption.

- ⚡ \*\*Insight:\*\* Outliers may indicate performance-focused EVs (high consumption) or highly efficient models (low consumption).

#### ✅ \*\*Task 3: Battery Capacity vs Range\*\*

- A scatter plot revealed a \*\*positive correlation\*\* between battery size and range.

- ⚡ \*\*Insight:\*\* Vehicles with larger batteries tend to offer more range, but some brands achieve higher efficiency.

#### ✅ \*\*Task 4: EV Recommendation Class\*\*

- A class-based tool was developed to recommend the \*\*top 3 EVs\*\* based on user input (budget, range, battery).

- ⚡ \*\*Insight:\*\* This model could be used to develop customer-facing web tools for personalized EV suggestions.

#### ✅ \*\*Task 5: Hypothesis Testing – Tesla vs Audi\*\*

- A two-sample t-test compared the \*\*engine power\*\* of Tesla and Audi EVs.

We performed a two-sample t-test to evaluate whether the average engine power (KM) differs significantly between Tesla and Audi electric vehicles. Based on the test, we [**found / did not find**] a statistically significant difference (p = x.xxx), suggesting that [**Tesla/Audi has higher power on average**].

- ⚡ \*\*Insight:\*\* The test showed that there \*\*was / was not\*\* a statistically significant difference in engine power between the two brands.

---

### ✅ Recommendations

1. \*\*Prioritize battery efficiency\*\*: Some EVs achieve long ranges with smaller batteries, which is key for cost and weight reduction.

2. \*\*Develop smarter recommendation systems\*\*: Expand the EV recommender into an interactive buyer guide.

3. \*\*Optimize inventory strategy\*\*: Focus on models with the best price-to-range and price-to-battery ratios.

4. \*\*Investigate outliers\*\*: Vehicles with extreme energy consumption could reveal valuable design insights.

5. \*\*Support data-driven decision making\*\*: Use this analysis as a blueprint for customer guidance, product design, and market segmentation.

---

### 🧠 Conclusion

This project applied Python tools like \*\*Pandas\*\*, \*\*Matplotlib\*\*, \*\*SciPy\*\*, and \*\*Seaborn\*\* to perform real-world EV data analysis. We identified key insights in pricing, range, efficiency, and brand performance, and proposed meaningful recommendations. These findings can be valuable for both manufacturers and EV buyers as the electric vehicle market continues to grow.

---

### 📽️ Video Explanation

[Insert your Google Drive video link here]