Fermi estimation (problem breakdown)

1)Define the Problem: The overarching goal is to segment the electric vehicle (EV) market.

2)Identify Key Metrics:

- Vehicle type (e.g., passenger cars, trucks)
- Geographic regions
- Customer demographics (age, income, occupation)
- Psychographic factors (lifestyle, values, attitudes)
- Behavioral factors (usage patterns, purchase intent)

3)Data Requirements:

- Estimate the need for a comprehensive dataset including EV sales records, vehicle specifications, customer profiles, and market trends.
- High-quality data from diverse sources is essential for accurate segmentation.
- The quantity of data required depends on the granularity of segmentation and the complexity of the analysis.

4)Data Collection:

- Estimate time and effort for data collection, including web scraping, data cleaning, and merging datasets.
- Consider the availability of data sources and the need to ensure data cleanliness and completeness.

5)Feature Engineering:

- Estimate the complexity of feature engineering, including the creation of relevant features from raw data.
- Consider the number of features needed to capture different aspects of the EV market and customer behavior.

6)Algorithm Selection:

- Estimate time and computational resources required to select and implement suitable machine learning algorithms.
- Consider clustering, classification, or regression techniques suitable for market segmentation.

7)Model Training:

- Estimate time and computational resources for training machine learning models on the prepared dataset.
- Model complexity and dataset size will influence the training time.

8)Hyperparameter Tuning:

- Estimate time needed for iterative hyperparameter tuning to optimize model performance.

- Consider the trade-off between model accuracy and computational resources.

9)Model Evaluation:

- Estimate effort required to evaluate model performance using appropriate metrics such as accuracy, precision, and recall.
- Cross-validation and validation set preparation are essential for robust evaluation.

10)Interpretation and Visualization:

- Estimate time and effort needed to interpret and visualize segmentation results.
- Consider how segments will be presented and communicated to stakeholders effectively.