**Segmentation based on EV Dataset with Updated Registrations**

1. **Analysis and Segmentation of Vehicle and Charging Station Data to Profile Potential EV Markets:**

Understanding the distribution of vehicle types and charging stations across states in India provides valuable insights for targeted EV promotion strategies. By analyzing factors such as vehicle counts, charging station availability, and regional distribution, we can identify high-potential segments for EV adoption and infrastructure development.

Key insights from the dataset suggest:

* States with higher numbers of two-wheelers and charging stations indicate readiness for EV proliferation.
* States with significant counts of goods vehicles and public service vehicles represent commercial markets for EV adoption.
* Regional disparities exist, with Tier 2 states showing a balanced distribution of vehicles and charging stations, highlighting emerging markets.

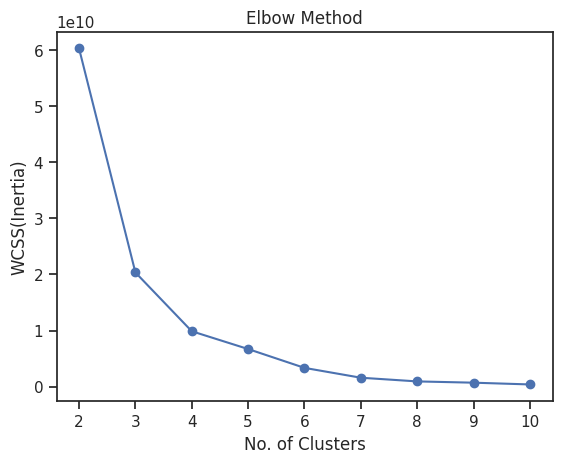
1. **Data Collection and Pre-processing:**

• The dataset contains vehicle counts and charging station data across various states.  
• Initial steps involve checking for missing values and ensuring data consistency.  
• Numerical features such as 'Two Wheeler', 'Three Wheeler', 'Four Wheeler', 'Goods Vehicles', 'Public Service Vehicle', 'Special Category Vehicles', 'Ambulance/Hearses', 'Construction Equipment Vehicle', 'Other', and 'total-charging-stations' are included.  
• Categorical features like 'State Name' are used for labeling but excluded from clustering.

1. **Clustering Process:**

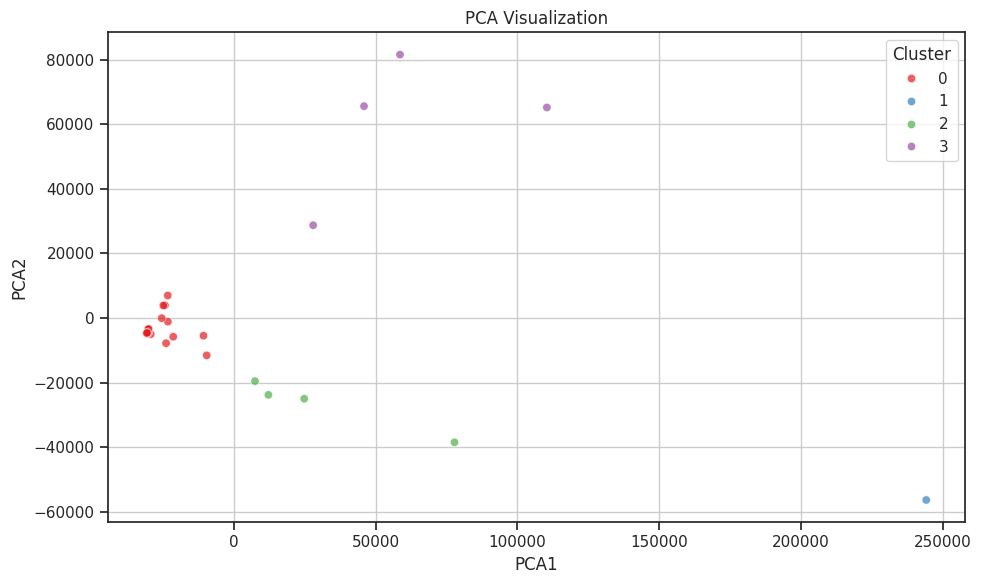
• To determine the optimal number of clusters, methods such as the elbow method and silhouette score are employed.

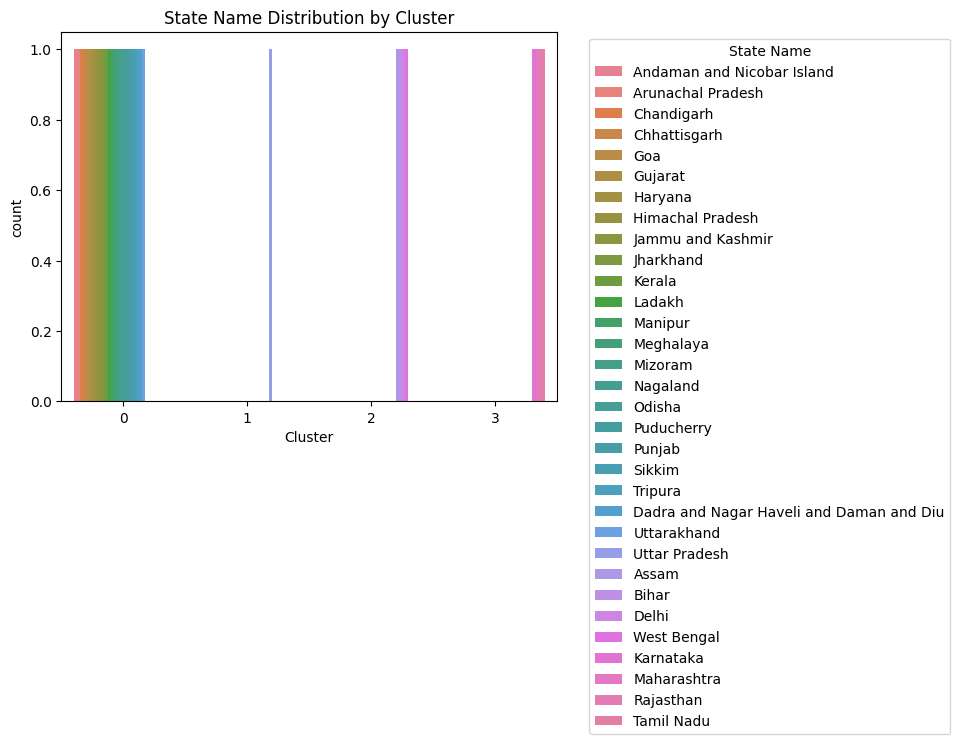
* **Elbow Method**
  + Visualizes the optimal cluster count by plotting the Within-Cluster Sum of Squares (WCSS) against potential cluster numbers
  + The "elbow point" (where the rate of decrease sharply changes) indicates the optimal number. Implemented using matplotlib for visualization
* **Silhouette Score**
  + Measures how similar each point is to its own cluster compared to other clusters
  + Higher scores indicate better clustering configurations.
* Based on this analysis, the ideal number of clusters is identified (for example, 4 clusters).
* KMeans clustering is applied to the scaled features, and a new 'Cluster' label is assigned to each state.

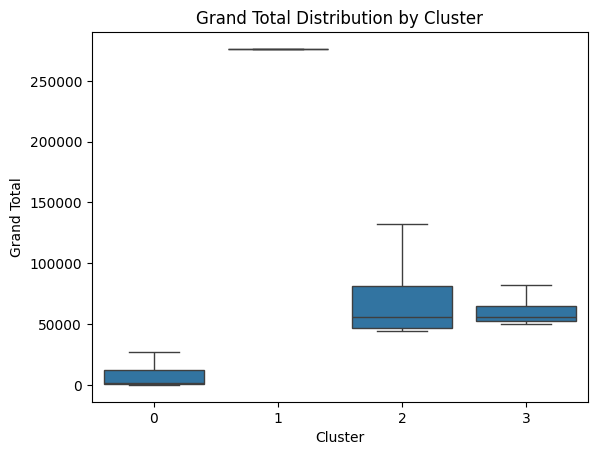


1. **Profiling and Visualization:**

• The resulting clusters are profiled by aggregating features like vehicle counts and charging station numbers.  
• Seaborn and matplotlib libraries are used to visualize the distribution of vehicle types and charging stations across clusters.  
• PCA reduces the high-dimensional data into two principal components, enabling a 2D scatter plot for visual inspection of cluster separations.







1. **Key Findings:**

* Cluster 0: States with moderate vehicle counts and charging stations, typically Tier 2 regions with emerging EV infrastructure.
* Cluster 1: States with high vehicle counts across categories and abundant charging stations, indicating ready markets.
* Cluster 2: States with lower vehicle counts but growing infrastructure, representing potential markets for EV expansion.
* Cluster 3: States with lower vehicle counts but growing infrastructure, representing high scale markets for EV expansion.

1. **Strategic Recommendations for EV Marketing and Infrastructure:**

* Focus on Cluster 1 states for expanding premium and mass-market EV models, emphasizing existing infrastructure.
* Promote cost-effective and utility-based EV solutions in Cluster 2 and 3 states, with incentives and awareness campaigns.
* Invest in infrastructure development in emerging clusters to accelerate adoption.
* Tailor marketing messages highlighting cost savings, environmental benefits, and convenience of EVs based on regional vehicle usage patterns.
* Marketing should focus on pointing out the cost-effectiveness of electric vehicles as compared to fuel vehicles, listing the areas where costs will be cut. This property must be marketed in all segments.
* Also emphasize the reduction in pollution an EV will bring being advantageous both to the customers and their surroundings.