**Tutorial file:** [**Mall Customer Segmentation.ipynb**](https://troy.instructure.com/courses/94127/files/55608845?wrap=1)

**Dataset for above:** [**Mall\_Customers.csv**](https://troy.instructure.com/courses/94127/files/55617112?wrap=1)

**[Download Mall\_Customers.csv](https://troy.instructure.com/courses/94127/files/55617112/download?download_frd=1)**

**Dataset for the assignment:** [**driverdata.csv**](https://troy.instructure.com/courses/94127/files/55608839?wrap=1)

[**Problem Statement**](https://troy.instructure.com/courses/94127/files/55608865?wrap=1)

**[Download Problem Statement](https://troy.instructure.com/courses/94127/files/55608865/download?download_frd=1)**

**The instructions for this assignment are in the documents above.**

**After viewing the Tutorial above, use the dataset and the Problem Statement to perform a cluster analysis.**

**What to submit:**

**Jupyter notebook**

**Videos to watch:**

[**K-means clustering**](https://www.youtube.com/watch?v=4b5d3muPQmA)

**[StatQuest: K-nearest neighbors, Clearly Explained](https://www.youtube.com/watch?v=HVXime0nQeI)**

**[Links to an external site.](https://www.youtube.com/watch?v=HVXime0nQeI)**

**[K-means Clustering Algorithm: Applications, Types, & How Does It Work?](https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm)**

**[Links to an external site.](https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm)**

**[K Nearest Neighbour Easily Explained with Implementation](https://www.youtube.com/watch?v=wTF6vzS9fy4)**

**[Links to an external site.](https://www.youtube.com/watch?v=wTF6vzS9fy4)**

**[K-means clustering: how it works](https://www.youtube.com/watch?v=_aWzGGNrcic)**

**[Links to an external site.](https://www.youtube.com/watch?v=_aWzGGNrcic)**

**Readings:**

**[What Is Clustering?](https://builtin.com/data-science/data-clustering-python)**

**[Links to an external site.](https://builtin.com/data-science/data-clustering-python)**

**[Clustering](https://en.wikipedia.org/wiki/Cluster_analysis)**

**[Links to an external site.](https://en.wikipedia.org/wiki/Cluster_analysis)**

**[Introduction to K-Means Clustering](https://www.pinecone.io/learn/k-means-clustering/)**

**[Links to an external site.](https://www.pinecone.io/learn/k-means-clustering/)**

[**10.pdf**](https://troy.instructure.com/courses/94127/files/55608880?wrap=1)

**[Download 10.pdf](https://troy.instructure.com/courses/94127/files/55608880/download?download_frd=1)**

**(pages 3 - 15)**

**[Customer Segmentation for Wholesalers](https://www.aaarl.ca/post/customer-segmentation-for-wholesalers)**

**[Links to an external site.](https://www.aaarl.ca/post/customer-segmentation-for-wholesalers)**

[**Client\_Segmentation.pdf**](https://troy.instructure.com/courses/94127/files/55058218?wrap=1)

**[Download Client\_Segmentation.pdf](https://troy.instructure.com/courses/94127/files/55058218/download)**

**Business Problem: Optimizing Battery Rental Pricing with Driver Segmentation**

**Background:** Lithionpower, the leading electric vehicle (e-vehicle) battery provider, operates on a rental model for its batteries. The company faces challenges in determining optimal rental pricing for e-vehicle drivers. The battery's lifespan is influenced by various factors such as over speeding and the distance driven per day.

**Objective:** Develop a cluster model that groups drivers based on their driving data, allowing Lithionpower to optimize its battery rental pricing strategy. The goal is to identify patterns in driving behavior to create targeted pricing plans, ensuring fair and competitive rates for drivers while maximizing the company's profitability.

**Data Fields:**

1. **id:** Unique identifier for each driver.
2. **mean\_dist\_day:** Mean distance driven by the driver per day.
3. **mean\_over\_speed\_perc:** Mean percentage of time a driver exceeded 5 mph over the speed limit.
4. **Gender**: Gender of the driver.

**Approach:**

1. **Cluster Analysis:** Utilize machine learning clustering algorithms to group drivers with similar driving behavior.
2. **Segmentation Criteria:**
   * Drivers in the same cluster exhibit similar driving patterns.
   * Clusters can be formed based on mean distance driven and mean over speeding percentage.
3. **Benefits:**
   * **Customized Pricing:** Tailor rental pricing plans based on cluster characteristics.
   * **Resource Optimization:** Efficiently manage battery replacements and charging stations based on cluster demands.
   * **Customer Satisfaction:** Offer fair and competitive pricing, enhancing customer satisfaction and loyalty.
4. **Outcome:**
   * **Optimized Pricing Plans:** Differentiated pricing plans for each cluster to reflect the impact of driving behavior on battery life.
   * **Improved Operational Efficiency:** Streamlined battery replacement and charging logistics based on cluster demands.
   * **Enhanced Profitability:** Maximizing revenue by aligning pricing with driving patterns while ensuring customer satisfaction.

**Key Performance Indicators (KPIs):**

1. **Cluster Homogeneity:** Measure the similarity of driving patterns within clusters.
2. **Customer Satisfaction Index:** Monitor customer feedback and satisfaction with the new pricing plans.
3. **Revenue Growth:** Track the impact of optimized pricing on overall revenue.

By leveraging machine learning clustering techniques, Lithionpower aims to enhance its business model, providing a win-win solution for both the company and its e-vehicle driver customers.