ANALYSING FACEBOOK LIVE MERCHANTS

SHARON ZACHARIA

22074407

ABSTRACT

- This analysis utilizes Pandas, Matplotlib, and Scikit-learn to delve into a CSV dataset.
- Following data cleaning and exploration, K-means clustering is applied to unveil patterns.
- A subsequent K-means-based classification model predicts 'status_type.'
- Visualizations such as line plots, bar charts, and scatter plots, including the elbow method, enhance dataset insights and model evaluation.

INTRODUCTION

- This study explores the commitment examples of Facebook Live merchants in Thailand, utilizing a dataset that incorporates pointers like as reactions, remarks, and offers.
- The principal stage is information stacking, in which we use pandas to import the dataset, do a central measurable examination, and check for missing qualities.
- Following that, a bunching examination utilizing K-implies, utilizing factors like reactions, criticism, and offers to identify patterns in dealer movement.
- Moreover, the project utilized curve_fit to fit models, explore time series or trademark associations, and gauge future qualities.

RESULT ANALYSIS

GOAL 1: Data Loading and Exploration

- The dataset summary shows that a few segments, explicitly 'Column1' through 'Column4', have a raised number of values that are absent (7050 contrasted with the all-out lines).
- Each point represents a data entry, and the line shows the trend.
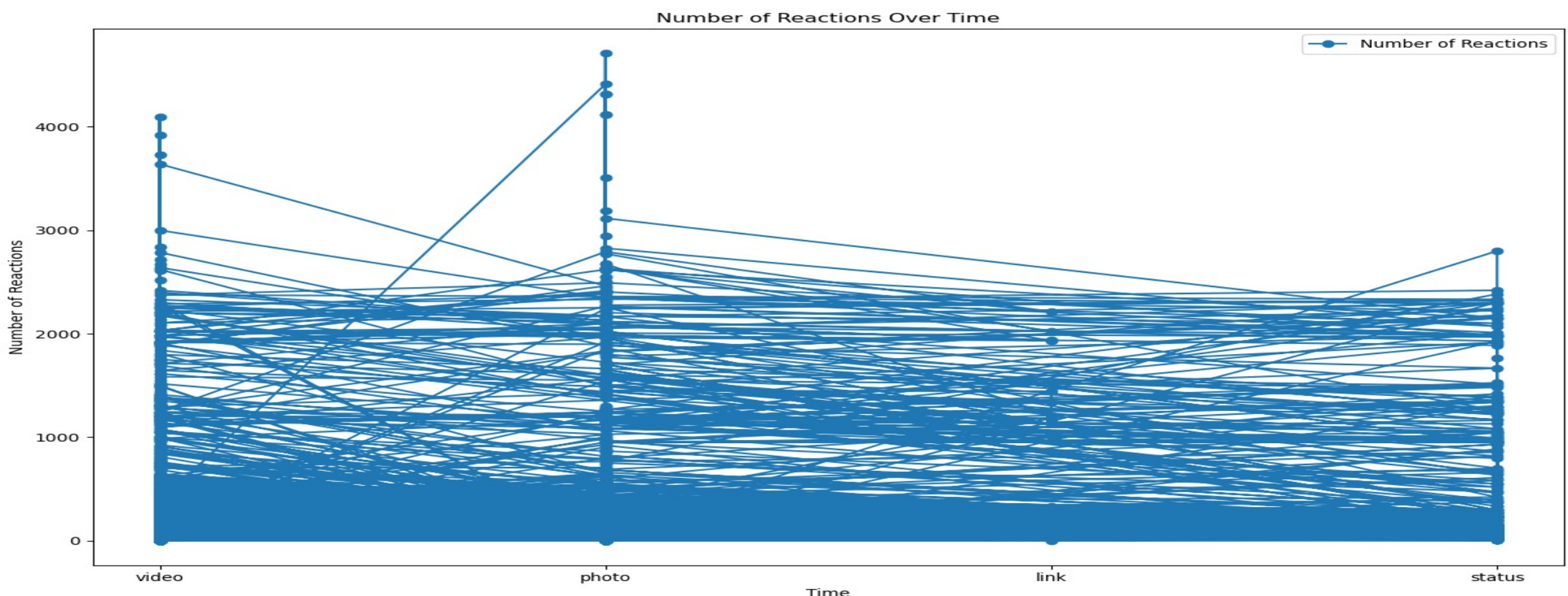
- The plot, which incorporates markers and names, portrays the example in responses, working with time-sensitive examination.
- Each bar represents a different status type.
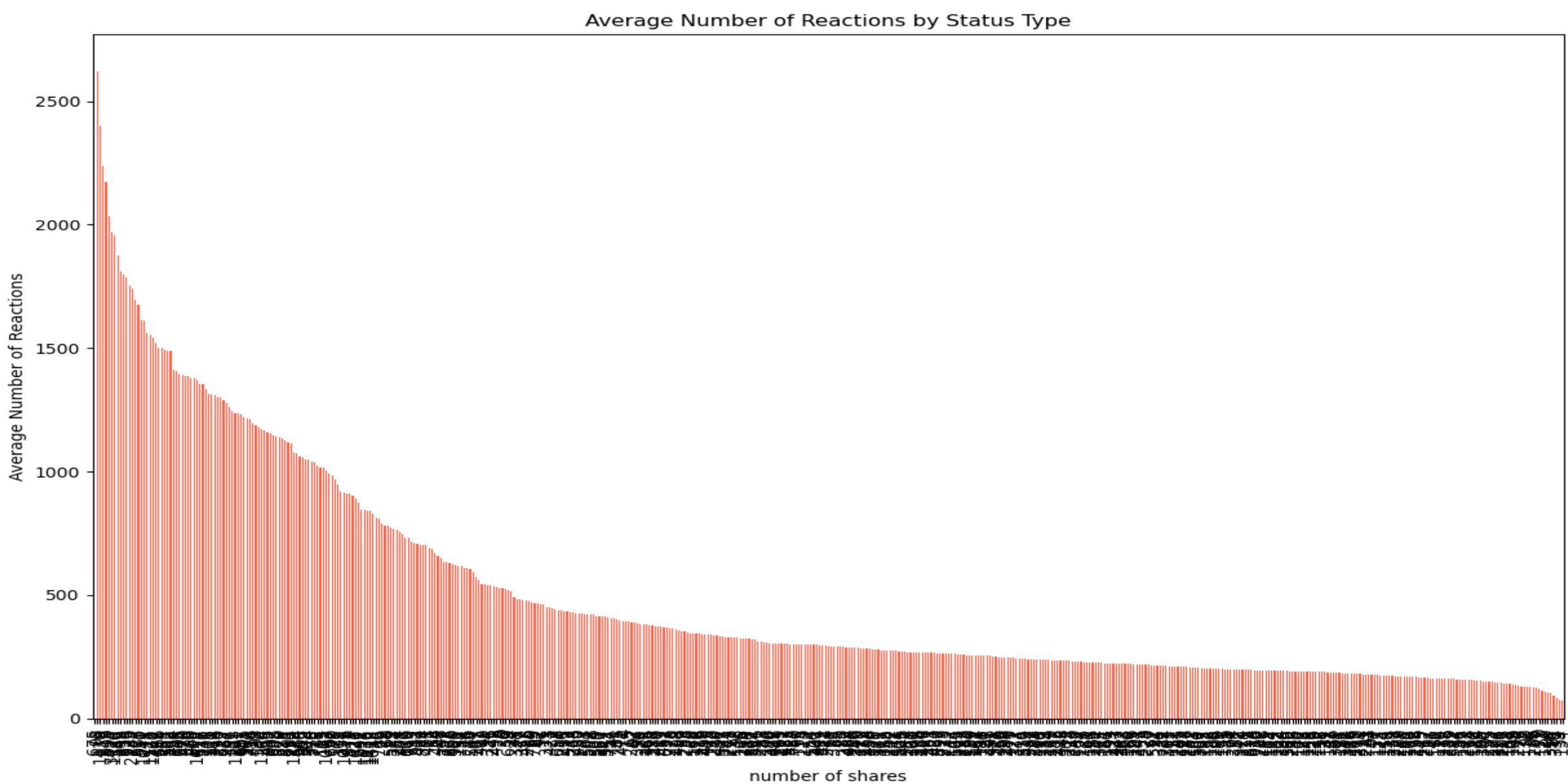

GOAL 2: Clustering

- Clustering is performed utilizing the suitable lines and segments 'num_reactions' and 'num_shares'
- Scatter plot showing the clustering results based on the number of reactions and comments.
- Different colors represent different clusters assigned by K-means.
- X-axis: Number of Reactions, Y-axis: Number of Comments.
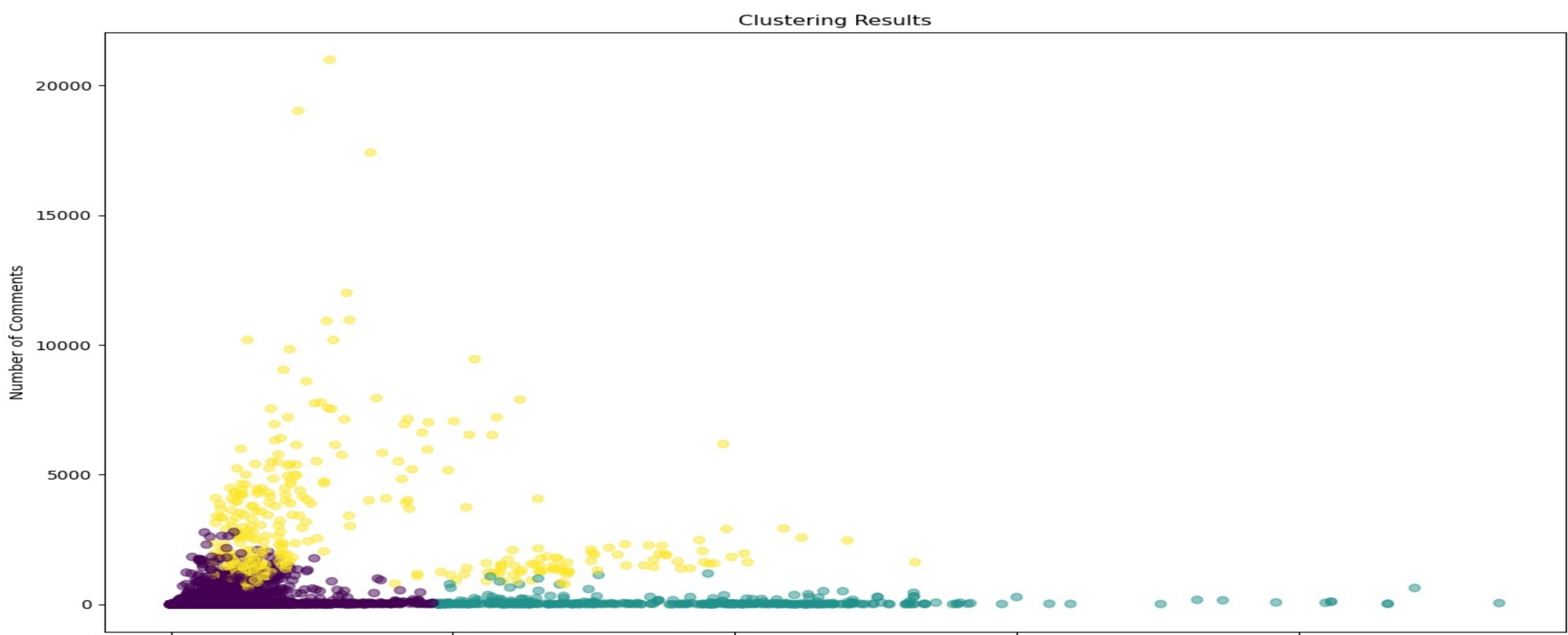
- This outlines the clustering discoveries by charting client responses and remarks.
- The scattered dots demonstrate data of interest, which are variety-coded by cluster.
- Moreover, red 'X' images recognize cluster focuses, which give experiences in client association designs.
- Helps visualize the central tendencies of each cluster.
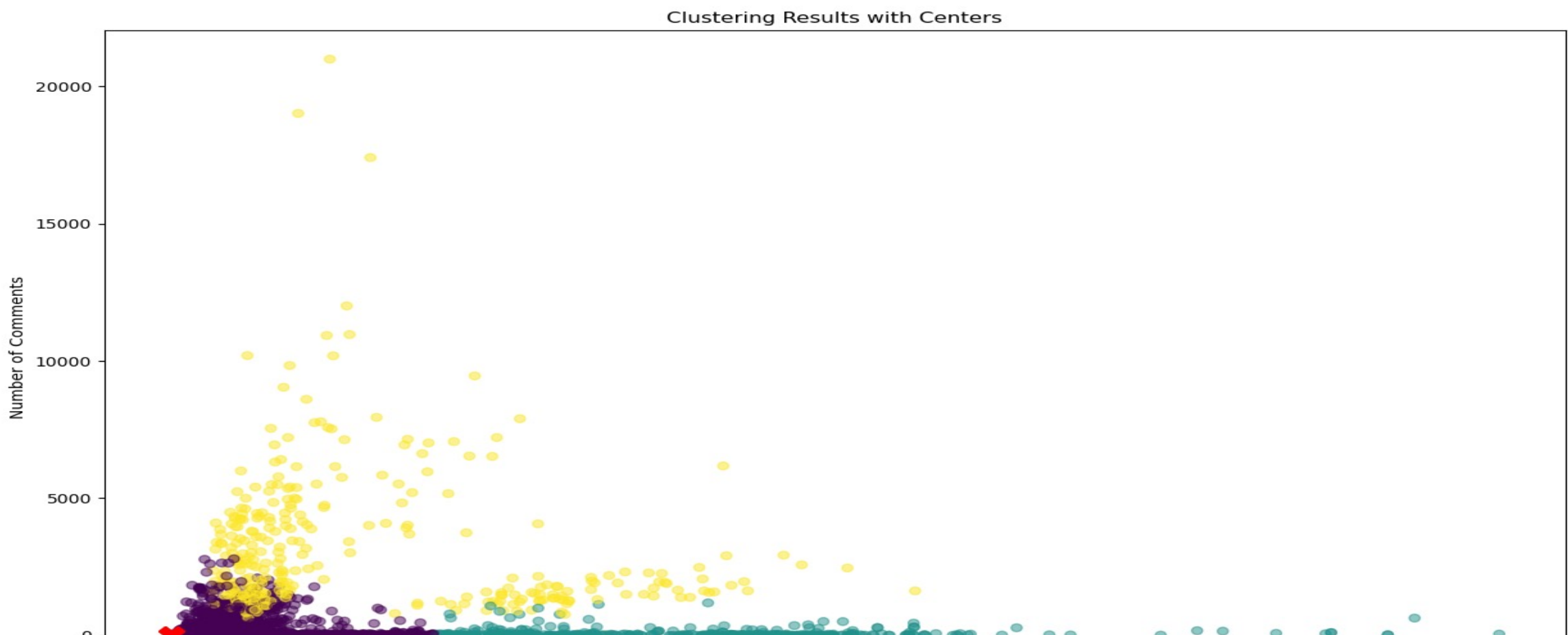
- A plot to determine the optimal number of clusters using the elbow method
- X-axis: Number of clusters, Y-axis: Within-cluster sum of squares (inertia)
- Helps identify the point where adding more clusters provides diminishing returns.
- This utilizes the greatest Elbow Strategy to work out the right measure of clusters utilizing the K-Means division (Liu *et al.* 2023).
- It emphasizes over-cluster sums, computes the all-out number of square squared distances (idleness), as well as shows the outcomes to outwardly distinguish the ideal cluster number.
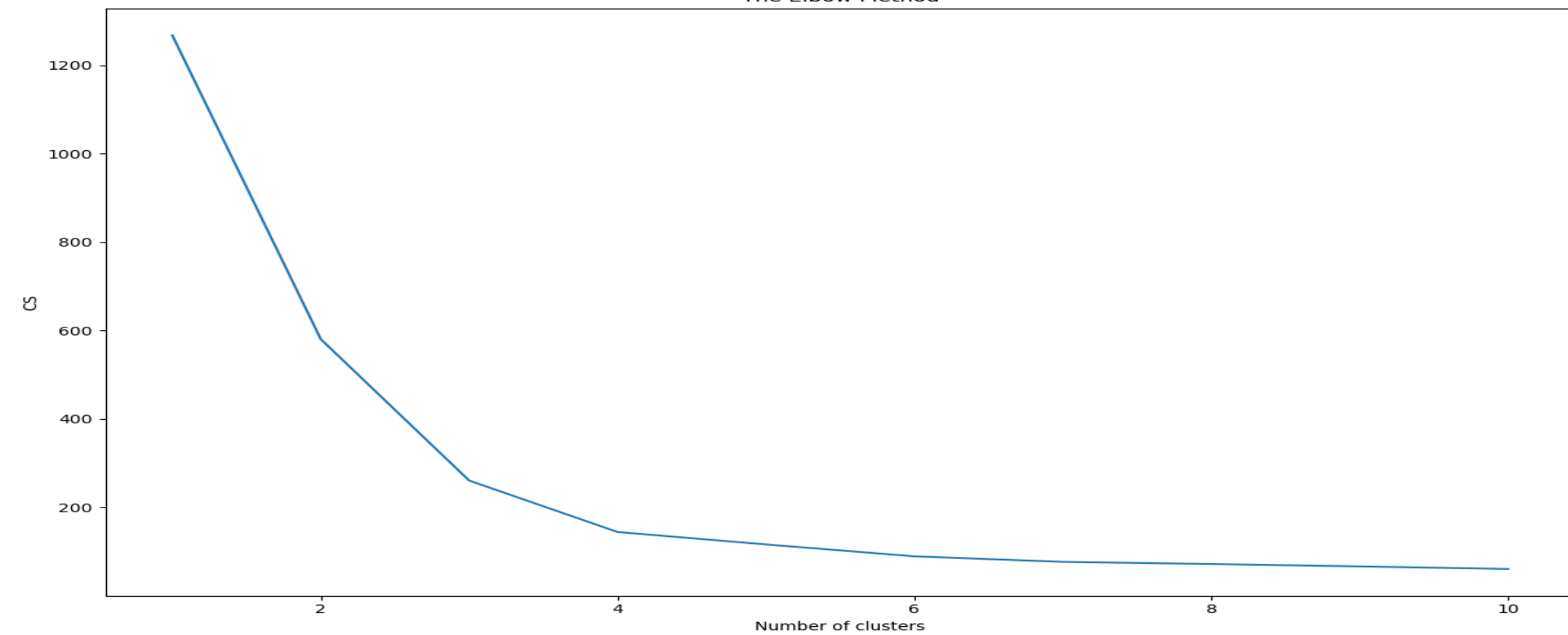
CONCLUSION

- The characterization results showed great precision, showing the way that the model can separate between cooperation levels.

- The discoveries of this examination assist organizations and advertisers with acquiring superior information on internet-based local area elements, permitting them to redo strategies for upgraded crowd investment.

- Addressing the further examination and clustering of new factors might work on the exactness of clustering alongside arrangement, giving more nuanced experiences about the project.