

Stakeholder Report: Financial Clustering and Predictive Modeling

This report summarizes insights from applying clustering and predictive modeling to financial data. The goal is to help decision makers understand performance patterns and identify firms with strong profit potential.

Executive Summary

We identified four financial groups using K Means clustering. These clusters revealed differences in revenue and profitability that helped improve predictions for identifying high margin firms. A Random Forest model performed best, showing that cluster information added measurable predictive value.

Feature Discovery

The elbow method and silhouette score supported selecting four clusters as the best structure for the data.

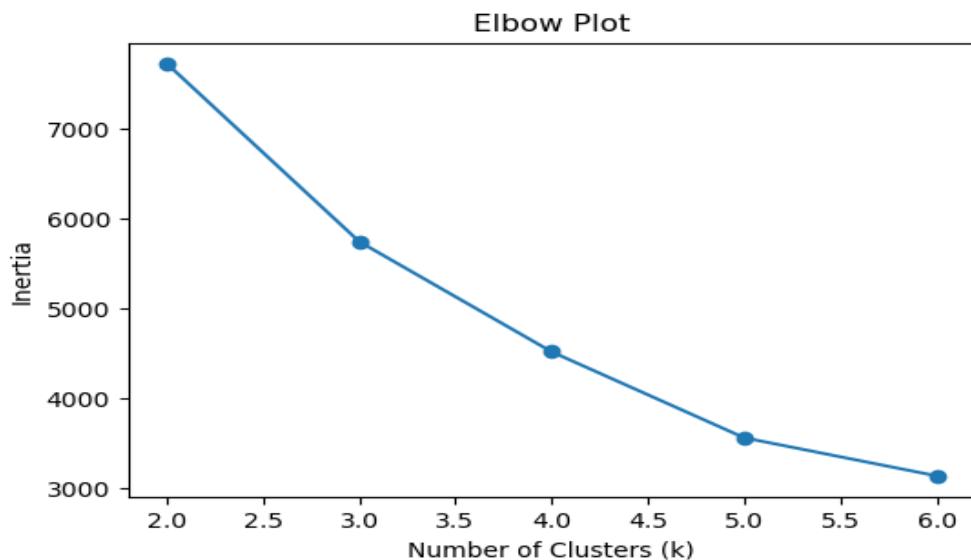


Figure 1. Elbow Plot suggesting $k = 4$.



Figure 2. Silhouette Scores confirming $k = 4$.

A PCA scatterplot shows that the four clusters separate clearly, meaning they represent real financial differences.

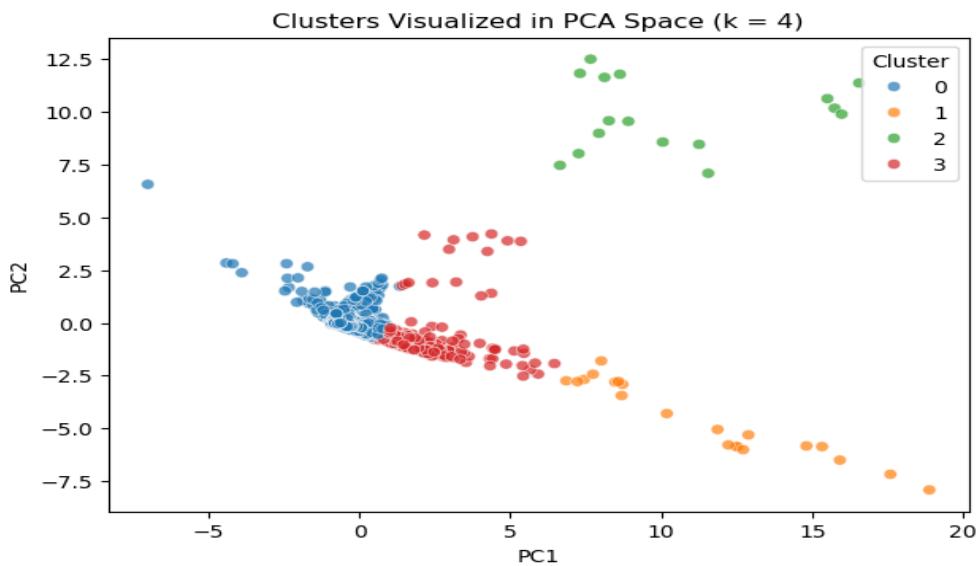


Figure 3. PCA visualization of the four clusters.

We examined revenue and net income across clusters to better understand each group.

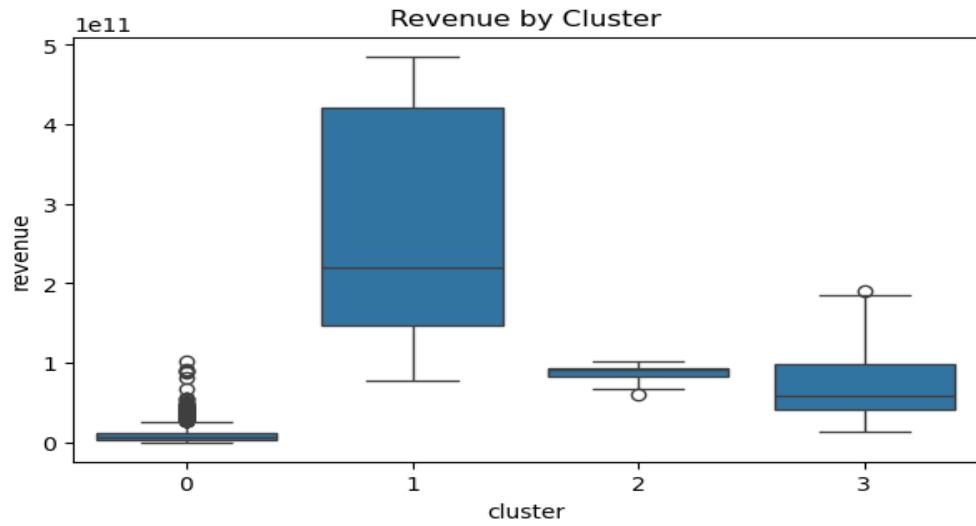


Figure 4. Revenue distribution by cluster.

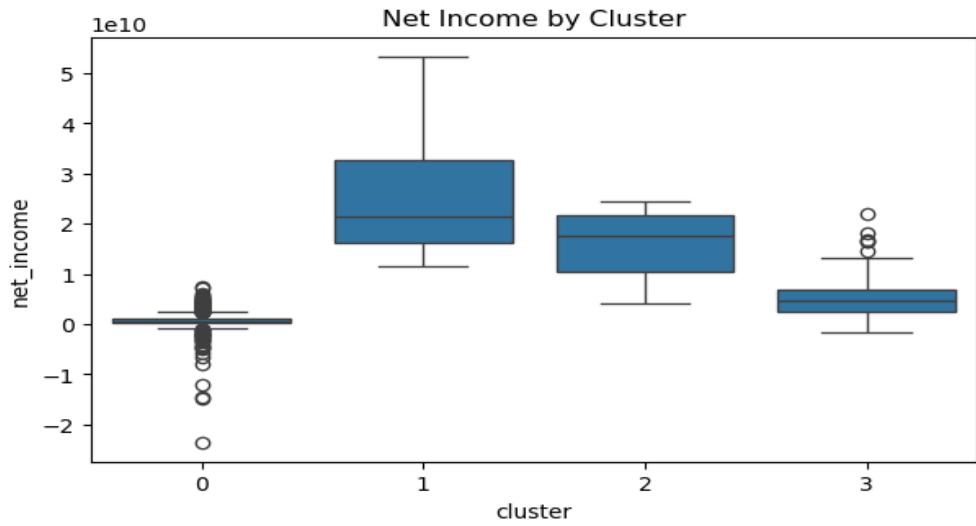


Figure 5. Net income distribution by cluster.

Predictive Modeling

We predicted whether firms were high margin. The Random Forest model performed best and showed that the cluster feature improved accuracy. This confirms the value of the unsupervised learning step.

Business Recommendations

- Prioritize firms in top performing clusters for investment or partnership.
- Monitor low performing clusters for financial risks.
- Use cluster membership in future forecasting models.
- Develop pricing or cost strategies based on differences between clusters.
- Track movement between clusters to identify financial improvement or decline.

Limitations and Next Steps

The dataset lacks industry labels and may include outliers. Adding sector information, more financial variables, and time based data would strengthen future analysis.