Data Science Engineering Club Project

Financial Customer Segmentations – New Dimensions

Vasily Ogievsky, Pietro Mascolo, Severino Tessarin, Fitri Kenny

11th August 2018

Problem Definition

Companies are facing increased uncertainty over the cash flows of their business due to market disruptions: price fluctuations, early redemptions, etc.

As a result traditional system defined categories are becoming more irrelevant. The same applies to the financial portfolios.

Problem: safeguard value of portfolios against market movements in the future.

Project Proposal

Explore number of clustering approaches to identify segmentations of financial portfolio without prior knowledge about the data. And try some tests for the quality of the clustering results: persistence, etc.

Data

Daily financial asset prices were used.

Methodology

Different clustering approaches were considered by project members on the raw data and derived variables, such as: K-means with different number of clusters, PCA, DTW, etc.

Results

More complicated approaches are harder to explain. How to pick the best one? Persistence tests based on percentage of repeat winners could provide an insight.

Clustering Results: Vasily Ogievsky

Clustering analysis done in SAS. Tested K-means with set number of clusters (5 and 10). Also tried a version of DTW algorithm based on similarity of time-series.
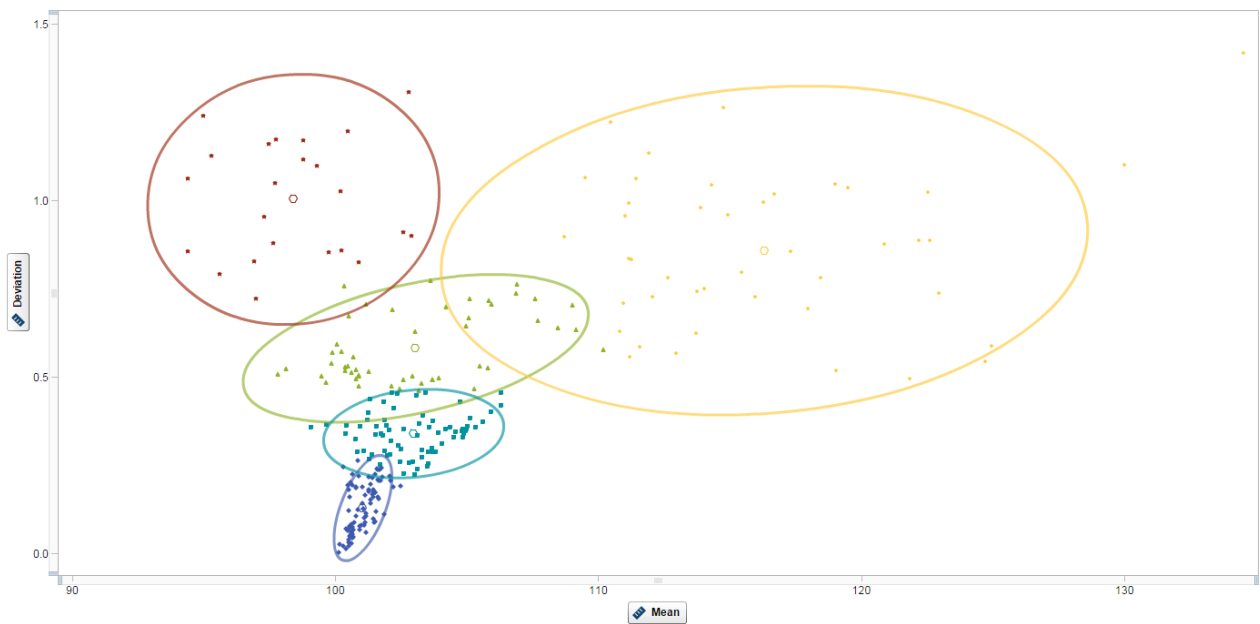
	Percentage Repeat W	Chi-Square	Yates
kmeans - 5 clusters	0.84	2766.47	2777.13
kmeans - 10 clusters	0.74	646.19	647.92
dtw - 10 clusters	0.45	538.50	539.87

Repeat winners analysis was used to determine persistence of bonds within clusters and the quality of the clustering algorithm chosen. K-means with 5 set of clusters shows highest persistence.

Conclusion and Business Impact

- Different approaches with aim are to correspond the method with business focus as close as possible.

- Persistent that gives clear defined segments to indicate the diversification that allows being better managing the fund.

- The results could be used for performance comparisons which benefit the fund managers.

- Data quality could skew the results,e.g. missing data, clean against unclean prices, etc.

- To discover the customers' needs and wants, through the ultimate in predictive customer modeling.

- To understand the customer's behavior that creates emotionally-intelligent communications across multiple channels to engage in real time.

- Monetize to improve every customer metric: increase conversions and reduce churn. (Ref: Optimove.com)

QUESTIONS