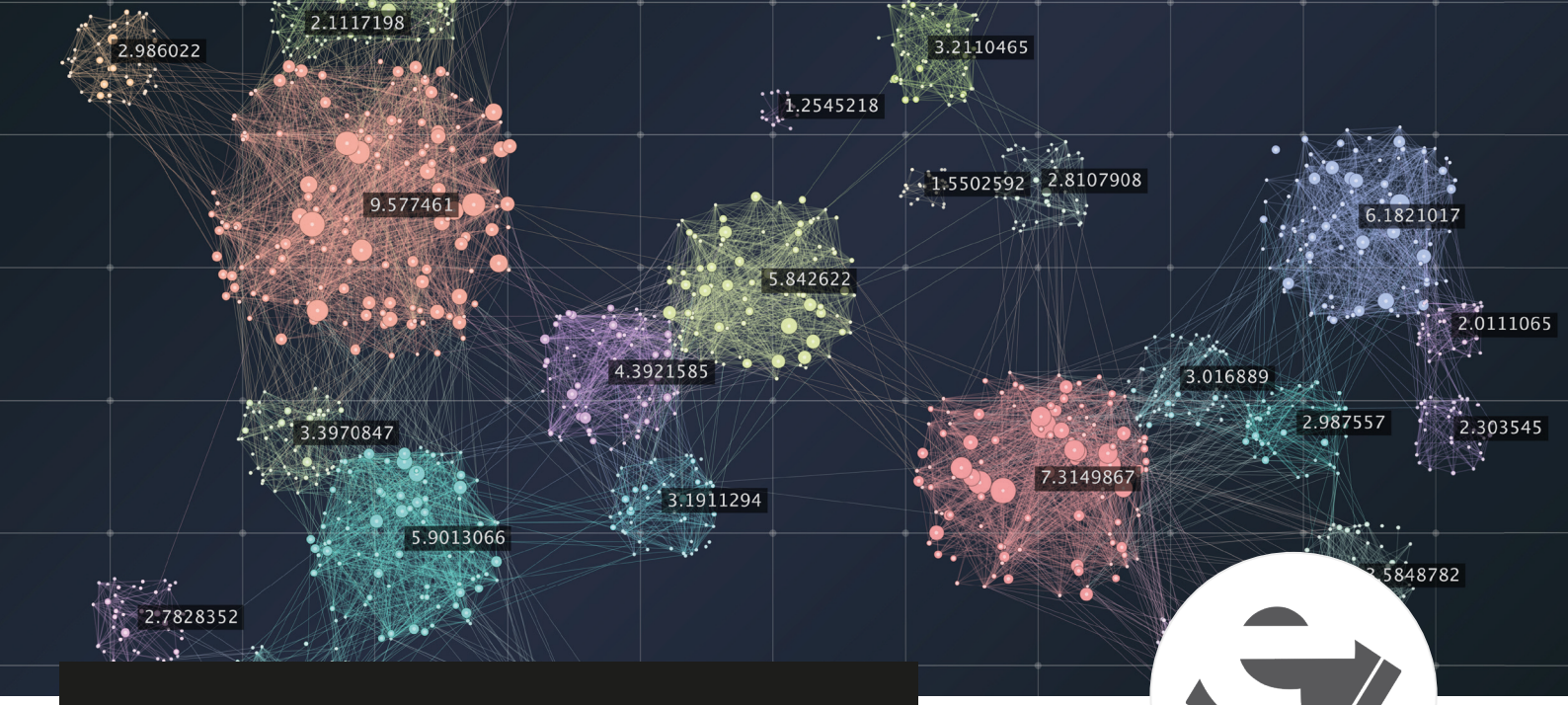


Analyse design space data easily with Multivariate Clustering techniques for efficient decision making

by Mauro Munerato, Alessandro Viola
ESTECO

When dealing with enormous amounts of data, analysing them is a challenging task. Clustering techniques come in handy for such data analyses. Assume you have a lot of books in your house and want to categorize them to discover the book you need quickly and easily. Obviously, randomly placing the books on the bookshelf is not particularly effective. Rather, you could organize your books by subject, by genre, by colour, or by any other criterion. The book collection method is another type of clustering where you group books according to your favourite criterion of similarity: any book in the group is a good representative of the book group.

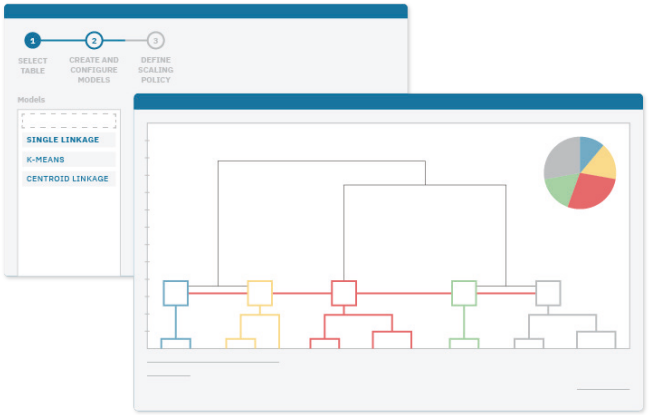
The same concept can be applied to simulation-driven product development where cluster analysis allows you to identify groups of designs with similar features within a complex dataset.

For instance, imagine you are a CAE engineer working on a vehicle development project using parametric simulation and optimization techniques to improve design performance. After an optimization run, you may end up with thousands of designs where each design represents a possible car configuration. Manually analysing each design is impractical; you need an automatic way to detect the fundamental differences between them. A clustering tool allows you to do that: you need only select the variables (discern the similarities) and a few other parameters, and the tool will do the trick. It will group the designs into a smaller number of clusters so that you only need to compare the representatives of each cluster. Once you have chosen the most relevant cluster for your goals, you can further examine the designs it contains.

Clustering: a multivariate analysis method to facilitate design categorization

Multivariate Analysis (MVA) refers to statistical techniques used to analyse datasets with a lot of variables to consequently identify patterns. It allows you to better understand the design space and the relationships between the variables before formulating the optimization study, making the process more efficient.

Among MVA techniques (such as classification, multiple correlations analysis, and dimensionality reduction), clustering is a well-known unsupervised learning task in data mining. This technique can be used for pre-processing and post-processing, which means that you can cluster before the optimization (starting from the input space parameters), or you can group the designs resulting from an exploration. Both pre-processing and post-processing approaches reduce the amount of data



to be managed, simplifying the analysis by identifying those designs that represent the variability of the entire dataset.

modeFRONTIER process automation and design optimization software contains an MVA environment that includes a tool to perform clustering analysis. It allows users to organize designs into groups according to a given rule (hierarchical and partitive clustering) and looks for clusters of data.

How modeFRONTIER's clustering analysis tool works

Among other things, clustering in modeFRONTIER can be used to:

- Group data based on the position of each design in the input design space. These grouping could be used to train Response Surface Models (RSMs) in different areas of the design space.
- Identify the areas of the design space with the best values for the objectives. The ultimate goal is to reduce the variable space and subsequently explore this reduced area with a Design of Experiments (DOE) or an exploration algorithm.

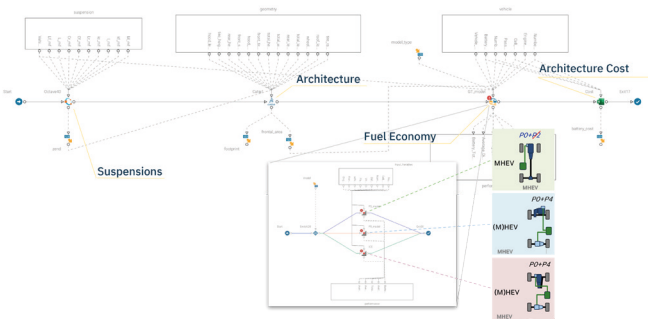
Once you have selected a dataset to cluster to find groups of designs with similar input and output properties, you can run clustering models using several hierarchical or k-means clustering algorithms available in modeFRONTIER. k-means (or partitive) clustering is more suitable for clustering a large dataset, while hierarchical clustering produces results that are easily visualized using a dendrogram chart.

modeFRONTIER provides a single intuitive interface that unites hierarchical and partitive clustering, enabling you to simultaneously train multiple clustering models, each with a different combination of algorithms, variables, and scaling policies. After training the models, the final steps of the wizard prompt you to create a dendrogram chart for the hierarchical approaches and a DB-index chart for the partitive approaches.

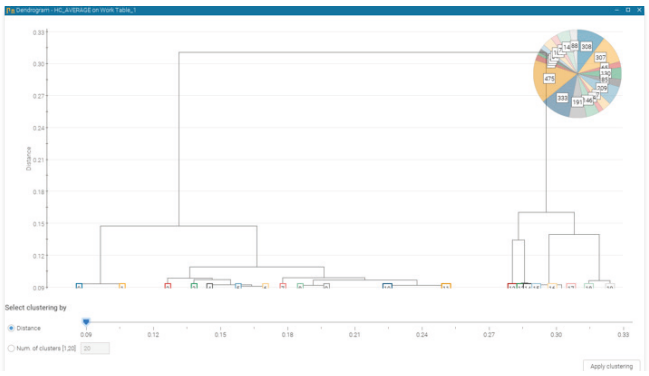
Hyundai: Clustering for vehicle architecture trade space analysis

Research engineers at Hyundai Motors and Kia R&D centre wanted to rapidly investigate and identify the global optimum design region at the early stage of a vehicle design, focusing on mechanical package design, system selection, and attribute modelling. The analysis involved components such as suspension, fuel economy, battery, and architecture costs. They used modeFRONTIER to perform Trade Space Analysis (TSA) to identify a set of system parameters, attributes, and characteristics to satisfy the required vehicle performance during the conceptual product development phase. In practice and starting from an automated multidisciplinary modeFRONTIER workflow, they ran a DOE to evaluate all the possible vehicle configurations. They then applied advanced post-processing techniques such as Clustering and Multi-Criteria Decision Making (MCDM) to group similar designs and rank all reasonable alternatives on the basis of given preferences.

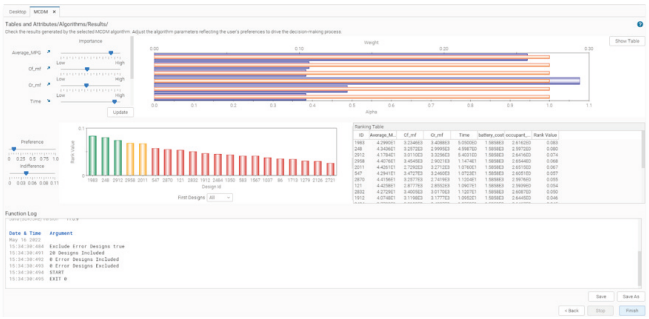
Initially, they generated 3,000 different vehicle configurations based on a low fidelity model to make a decision based on exploration of a vast domain. The first phase of data mining required wise filtering to reduce the dataset while maintaining its variability. They therefore applied the



modeFRONTIER multidisciplinary design exploration workflow used for trade space analysis



The dendrogram chart shows the 20 clusters obtained from the hierarchical clustering analysis.



MCDM rank chart shows the clusters' centroids ranking considering given preferences.

hierarchical clustering technique to the 3,000 designs using a variety of requirements (fuel economy, performance, cost, and comfort) and grouped similar designs into 20 clusters.

Finally, they performed an MCDM analysis on those clusters to support decision making by ranking all reasonable design alternatives on the basis of user-defined preferences. This allowed them to reject more than half of the designs immediately and to concentrate subsequent data analysis efforts on a smaller but representative dataset. In conclusion, clustering methods make the design optimization process more efficient. Multiple design samples with similar characteristics can be grouped into different clusters. This allows you to reduce both the complexity of the design space and the number of computations required during an exploration or optimization run.

For more information:
info@esteco.com