Cluster Analysis and Latent Class Analysis

Instructor:	Professor Dr Robin Samuel, University of Luxembourg
Date & Time:	Monday, 06 May, 2024, 09.00 a.m. to 05.00 p.m. (s.t.) Tuesday, 07 May, 2024, 01.00 p.m. to 05.00 p.m. (s.t.)
Place:	BAGSS, Feldkirchenstraße 21, 96050 Bamberg, Room FG1/00.06
Registration:	To register, please send an email to courses@uni-bamberg.de by April 08, 2024. Registration is mandatory. The number of participants is limited to 16.

Short Outline

This course offers a practical introduction to identifying groups in your data.
When confronted with large datasets containing numerous variables, casual browsing or superficial data examination will be insufficient for discovering similar cases that may form groups. Cluster analysis and latent class analysis will be helpful in these cases.

Cluster analysis is a bottom-up approach, which employs different algorithms to identify similar cases within the data, such as individuals or organizations, resulting in distinct clusters. Cluster analysis is a type of unsupervised machine learning.
Latent class analysis follows a top-down approach by assuming a probabilistic model to explain group membership. It utilizes the data distribution and allows for the inclusion of covariates, providing goodness of fit measures that facilitate the comparison of different solutions.
We will cover hierarchical cluster analysis, non-hierarchical clustering, fuzzy clustering, latent class analysis, as well as latent profile analysis and longitudinal applications.
Upon completion of this course, participants will have a good understanding of the possibilities and challenges of cluster analysis and latent class analysis

The workshop is mainly aimed at doctoral students of BAGSS and LIfBi. In case of any vacant places, it is open to anyone interested.

Prerequisites

Participants should be familiar with multivariate statistics. If you have never been exposed to t-tests, ANOVA, and (OLS) regression, this is not the course for you. We will use the software R and R Studio for demos and short exercises. While some familiarity with R is useful, this is not strictly necessary if you have some knowledge of working with other statistical software packages using syntax (e.g., Mplus, Stata, etc.) and are willing to learn. If you have never used R and R Studio or have not used them for some time, please install/update them and have a little look around before the first day of the course.
You will find some helpful materials here: https://stats.idre.ucla.edu/r/

About the Trainer

Robin Samuel has been an Associate Professor at the Department of Social Sciences at the University of Luxembourg since 2016, where he was appointed Head of the Centre for Childhood and Youth Research in 2020. His research is mainly concerned with the collection and analysis of large data sets. In some projects, he also applies experimental designs and qualitative methods. He is currently investigating the appropriateness of certain statistical models to study social inequalities in a range of health outcomes. His substantive research interests include social inequality, work, health, well-being, and sustainability, often with a focus on young people.