

## **20<sup>TH</sup> MAY** 2.30PM (WEST)

JOIN US <u>HERE</u>

## ROBUST DISTANCES FOR MIXED-TYPE DATA

Data scientists address real-world problems using multivariate and heterogeneous datasets, characterized by multiple variables of different natures. Selecting a suitable distance function between units is crucial, as many statistical techniques and machine learning algorithms depend on this concept. Traditional distances, like classical Gower's or Euclidean, are unsuitable for mixed-type data when underlying correlation or outlying observations are present, and often lead to suboptimal results. In this talk robust distances for mixed-type data will be explored, like robust Generalized Gower's and robust Related Metric Scaling, as well as their performance in clustering and prediction problems.



## SPEAKER AUREA GRANÉ [UNIVERSIDAD CARLOS III DE MADRID, SPAIN]

Aurea Grané is Full Professor of Statistics at the Carlos III University of Madrid. She holds a PhD in Mathematics from the University of Barcelona. Her recent scientific contributions are in the field of data science, where she developed robust statistical learning techniques based on the notion of distance between objects for visualization, prediction, and classification in large heterogeneous multivariate datasets. She currently coordinates the Statistical Data Science research group at Carlos III University of Madrid and the working group on Multivariate Analysis and Classification of the Spanish Society of Statistics and Operations Research.

