



ULST Timisoara  
**Multidisciplinary Conference on  
 Sustainable Development**  
 21-22 May 2026



**MULTIVARIATE ASSESSMENT OF THE PHYSICO-CHEMICAL COMPOSITION OF COMMERCIAL CHEESE VARIETIES: A PCA AND CLUSTER ANALYSIS APPROACH**

Virgil Dacian Lalescu<sup>1</sup>, Erisilia Alexa<sup>1</sup>, Cocan Ileana<sup>1</sup>, Maria Rada<sup>2</sup>, Petculescu Iasmina – Mădălina<sup>3</sup>, Ariana Bianca Velciov<sup>1\*</sup>

<sup>1</sup>University of Life Sciences “King Mihai I” from Timișoara  
<sup>2</sup>University of Medicine and Pharmacy “Victor Babeș”,  
<sup>3</sup>Politehnica University of Timișoara



**Abstract:**

This study aimed to characterize and classify 20 diverse cheese varieties based on five key physico-chemical parameters: moisture (U), dry matter (SU), protein (P), fat (F), and energy value (E). Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) were employed to identify the main drivers of variability and to group the samples based on their nutritional similarities. The application of multivariate statistical tools provides a robust framework for differentiating cheese varieties. These results facilitate a deeper understanding of how specific macro-nutrient ratios contribute to the overall nutritional identity of different cheese categories, offering valuable insights for both the dairy industry and clinical nutrition.

**Introduction**

- Cheese represents one of the most complex categories of dairy products, with thousands of varieties globally characterized by distinct sensory and nutritional profiles. The nutritional density of cheese is determined by a multitude of factors, including the animal species, milk composition, and, most significantly, the technological parameters of production such as curd cooking temperature, acidification rate, and ripening duration.
- From a nutritional perspective, the partitioning of macronutrients (proteins and fats) relative to moisture content determines the caloric density and physical texture of the final product. While traditional classification often relies on a moisture-free fat basis (MFFB), this approach frequently fails to capture the nuanced correlations among multiple nutritional variables.

**Materials and methods**

A total of 20 commercial cheese varieties were selected for analysis, representing a broad spectrum of textures and maturation stages. Recommended standard laboratory procedures were used to evaluate the concentration of nutritional parameters in the analyzed cheese samples. The dataset was constructed based on standardized nutritional labels and analytical reports, focusing on five fundamental physico-chemical parameters: moisture (U%), dry matter (SU, %), protein (P, g/100g), fat (F, g/100g), energy value (E, kcal/100g). The optimal number of principal components was determined from the Scree Plot. The interpretation of the dimensions was based on the quality of representation and the percentage contribution of each variable to the respective axes. Subsequently, a Hierarchical Cluster Analysis (HCA) was conducted to identify natural groupings among the cheese varieties. The classification was based on Euclidean distances, using Ward's minimum-variance method to minimize intra-cluster heterogeneity and maximize inter-group distance. The optimal number of clusters was validated against the PCA biplot results to ensure a robust taxonomic framework. All computational procedures were performed in the R environment (version 4.4.1).

**Results and discussions**

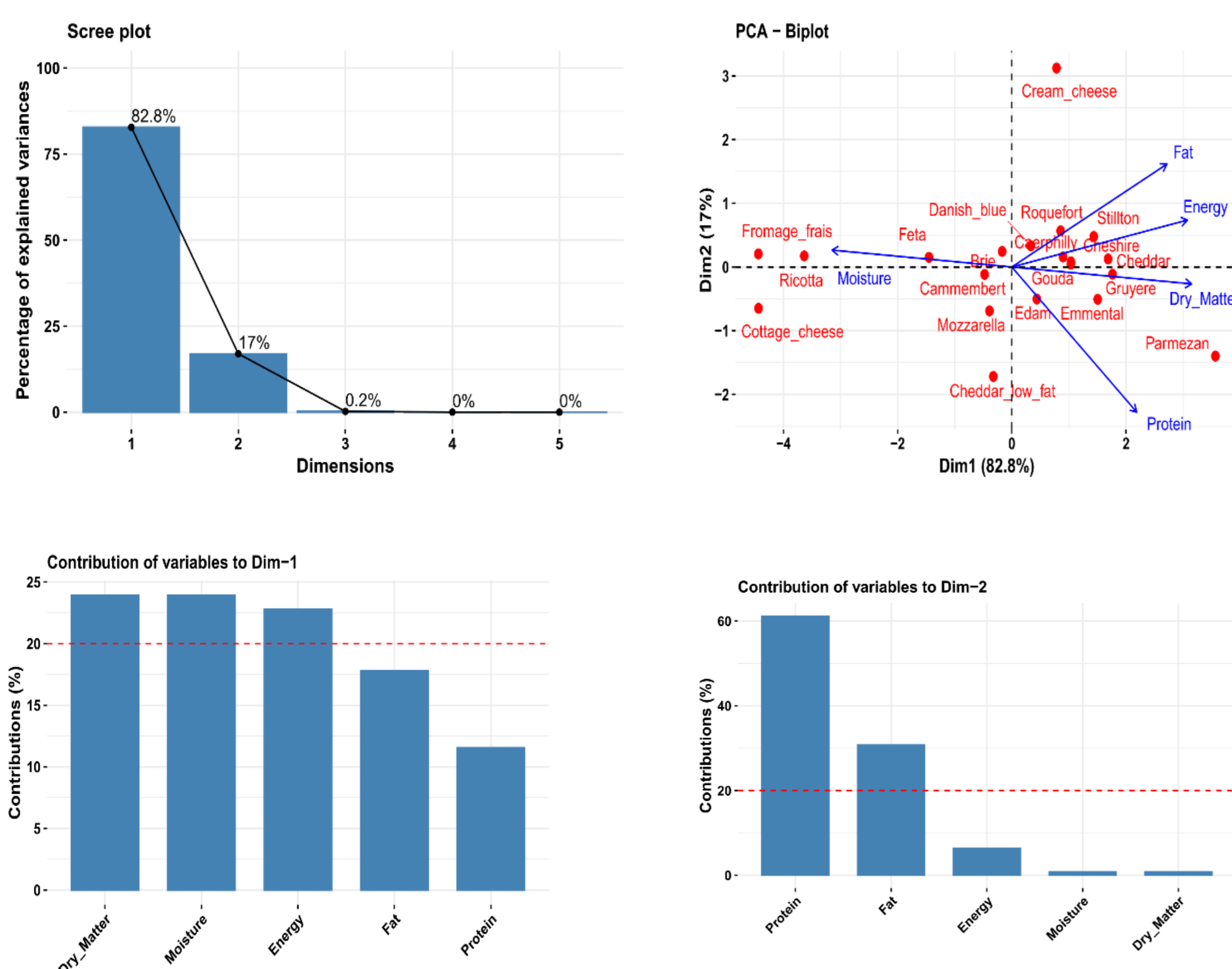


Figure 1. (A) Scree plot of PCA. (B) Biplot of PCA. (C) Contribution of variables to the first dimension of PCA. (D) Contribution of variables to the second dimension of PCA

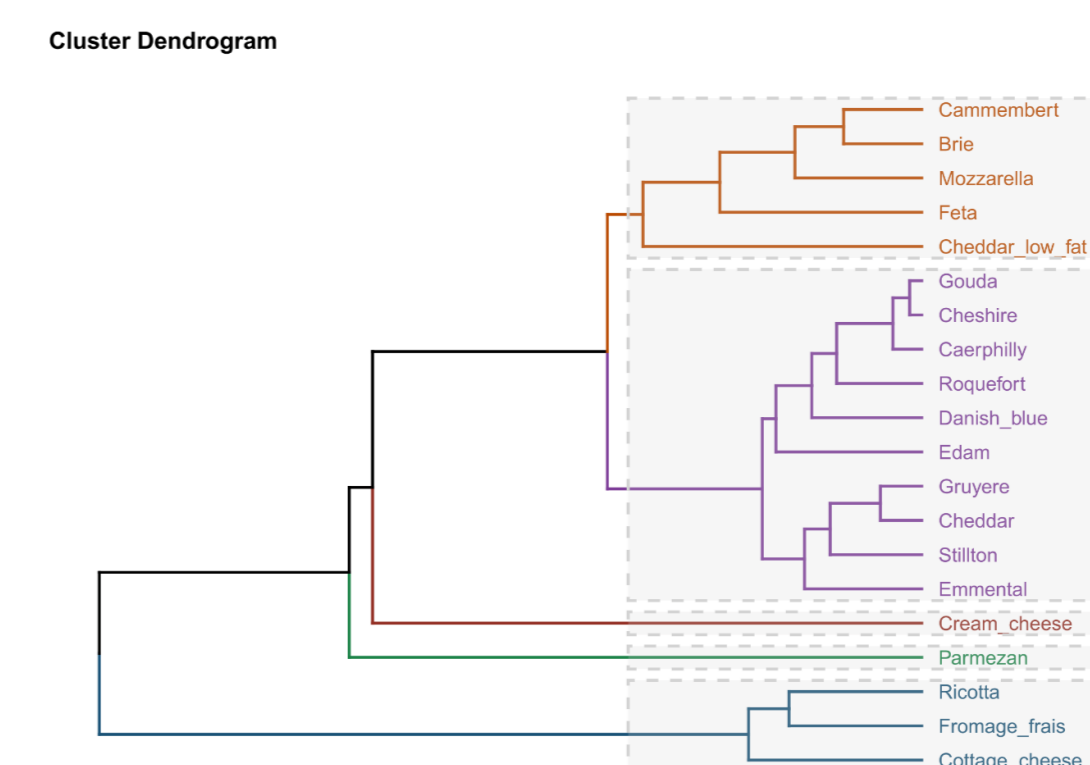


Figure 2 Cluster dendrogram for cheese

**Conclusions**

- The application of PCA and HCA successfully mapped the nutritional space of 20 diverse cheese varieties, explaining 99.8% of the total variance through only two dimensions.
- We identified that moisture content is the primary determinant of cheese classification, dictating the overall energy density (PC1).
- The protein-to-fat ratio provides a secondary layer of differentiation, allowing for the isolation of specific varieties, such as Cream cheese and Parmesan, as nutritional outliers.
- The HCA provided a clear five-cluster solution that transcends traditional artisanal categories, offering a more precise tool for nutritional labeling and the development of specialized dairy products.
- These findings demonstrate that multivariate analysis is an objective and powerful method for food categorization, providing valuable data for producers and nutritionists alike.