BADJI MOKHTAR UNIVERSITY
FACULTY OF ECONOMICS AND MANAGEMENT SCIENCES
Department: Financial Sciences

Exercises Serie N03.
Data Analysis, Statistics, and Applications

Level: L3
Instructor: Dr. Soumaya Allaoui

Exercise 1:

1. What is the main objective of Principal Component Analysis (PCA)?
2. How does PCA help in dimensionality reduction?
3. What are principal components, and how are they ordered?
4. What is the indicator of association between quantitative variables?
5. What is the role of eigenvalues and eigenvectors in PCA?
6. When should we apply the operations of centering and scaling to the data before applying PCA?

Exercise 2:

As a manager, you aim to understand how the climatic characteristics of cities influence consumer behavior and local needs. The following data represents precipitation (P, in cm), maximum temperatures ($t_{\max}$), and minimum temperatures ($t_{\min}$), in ($^{\circ}\mathrm{C}$) recorded in various cities in 2024:

	p	tmax	tmin
Ajaccio	12.04	23.7	5.9
Brest	17.18	15.5	−1.8
Dunkerque	11.83	13.1	2.8
Nancy	6.23	13.5	−2.4
Nice	16.99	21.1	7.2
Toulouse	3.87	20.3	−0.9

Questions

1. Calculate the statistics (mean, variance, and standard deviation) of P, tmax, and tmin.
2. Calculate the dispersion matrix corresonding to this problem statement.
3. Calculate the correlation matrix and , and then comment on the result.

We aim to study Principale Composent Analysis of the given problem statement.

1. Compute the variances (eigenvalues) of the principal components for this problem statement.
2. Deduce the total variance (inertia) in this Principal Component Analysis (PCA).
3. Determine the percentage of variance explained by each principal component.
4. Based on this, decide how many dimensions should be retained.
5. Find the normalized eigenvector corresponding to the first principal component.