

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

GÉNIE INFORMATIQUE

---

## Rapport du projet de SY09

---

*Authors :*

Crauser JULIEN et Antoine COLLAS

14 avril 2018



# 1 Cuisine

## Question 1

Le jeu de données présent dans le fichier recettes-pays.data contient 51 variables pour 26 individus. Nous pouvons donc noter que nous disposons de plus de variables que d'individus. Parmi ces 51 variables, une seule est qualitative (les origines des recettes) et 50 sont quantitatives. Ces dernières prennent leurs valeurs entre 0 et 1 (0 et 0.82 pour être plus précis). Dans ce jeu de données les recettes ont été agrégées par origine, il y a donc 26 origines (une par ligne de notre tableau individus-variables). De plus, le jeu de données ne présente aucune valeurs manquantes.

## Question 2

Nous réalisons une ACP sur le jeu de données. Nous obtenons 26 axes principaux : autant que d'individus. En effet, comme il y a moins d'individus que de variables, il est suffisant de prendre 26 axes pour représenter tous les individus.

Nous les pourcentages d'inertie expliquée visibles sur la figure 1. Les 3 premiers axes représentent 72% de la variance du nuage de points.

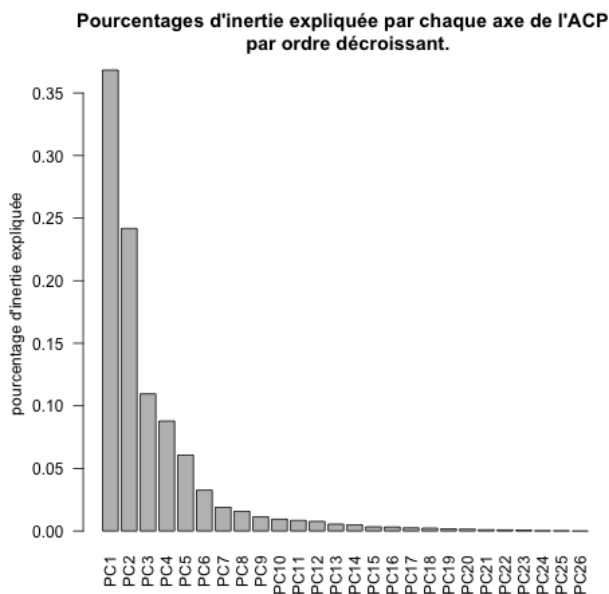


FIGURE 1 – Pourcentages d'inertie expliquée

Comme les 3 premiers axes factoriels représentent 72% de la variance totale, nous analysons le nuage de points seulement avec les 3 premiers plans factoriels. Nous obtenons les plans factoriels suivant :

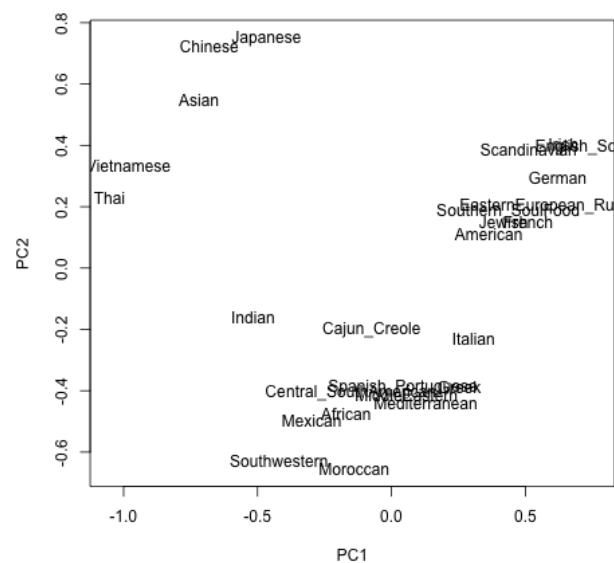


FIGURE 2 – Plan 1,2

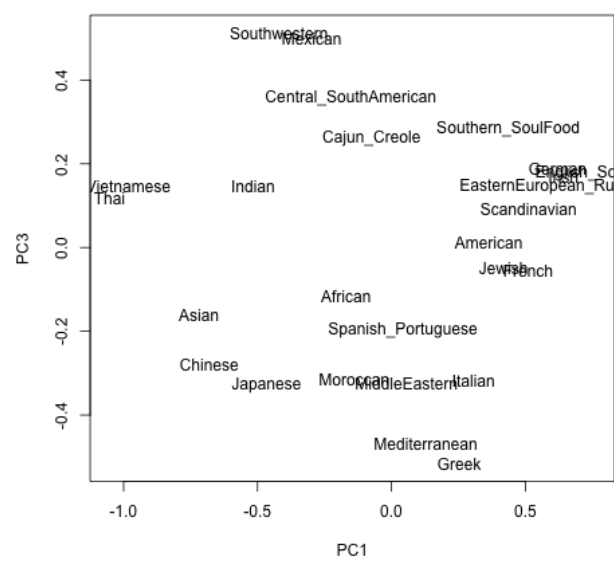


FIGURE 3 – Plan 1,3

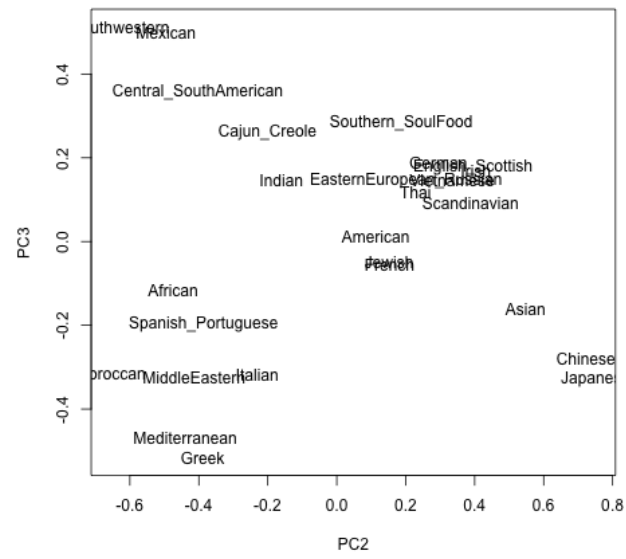


FIGURE 4 – Plan 2,3