

## TP 3 Analyse de données

Arthur Crochemore et Alaâ Chakori Semmane

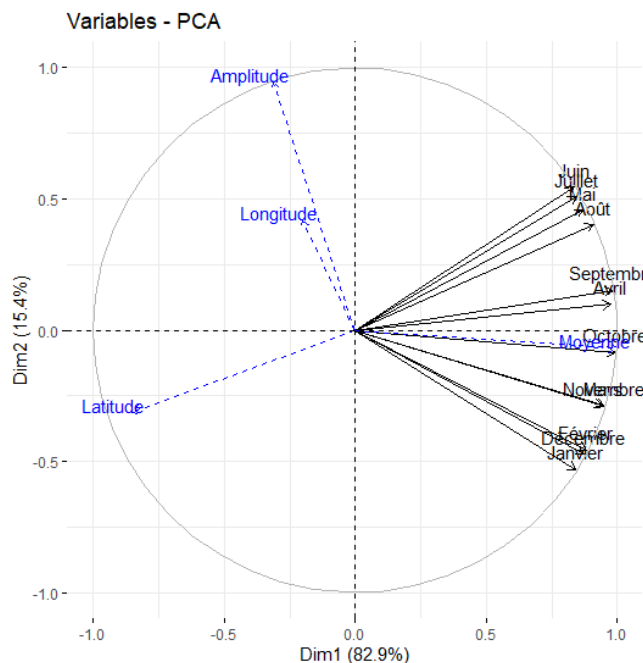
### Partie 2

**1. Peut-on se borner à travailler sur le premier plan factoriel. En d'autres termes, le plan factoriel porté par les deux premières valeurs propres est-il représentatif de l'inertie total du nuage de points initial (argumentez) ?**

Les deux premières valeurs propres représentent 98,3% de l'inertie totale du nuage de points initial, donc on peut se borner pour travailler sur le premier plan factoriel pour analyser les corrélations entre les profils des températures des différentes capitales.

Si on a une proportion d'inertie aussi élevée indique que le premier plan factoriel capture une grande partie de la variation des données et peut donc être considéré comme représentatif de l'ensemble du nuage de points initial.

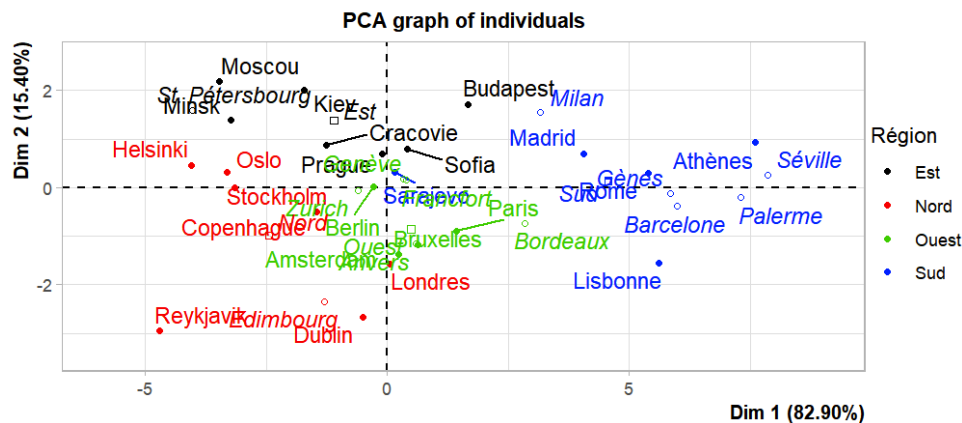
**2. Quelle signification donnez-vous à l'axe 1. Dit autrement, que représente la variable synthétique portée par cet axe ?**



Au vu de la répartition des variables dans le nuage des variables, le premier axe semble opposé les variables de températures aux variables de latitude, de longitude et d'amplitude,

probablement car elles n'ont rien à voir entre elles. Aussi, la latitude apparait comme fortement corrélée avec cet axe contrairement aux variables d'amplitude et de longitude.

Concernant le nuage des individus, la première variable semble opposer les villes du sud aux villes du Nord, la répartition semble donc se faire selon la latitude, ce qui est cohérent avec l'observation précédente sur le nuage des variables.



### 3. Même interrogation concernant l'axe 2.

Concernant le nuage des variables, le deuxième axe semble opposer les mois où la température est la plus haute aux mois les plus froids. Notons que la variable moyenne semble être à distance moyenne de chacune des variables de température mensuelle. Aussi, cette fois-ci, c'est la variable d'amplitude qui est fortement corrélée avec une abscisse proche de 1.

Concernant le nuage des points individus, on peut supposer que l'axe 2 oppose les villes ayant les plus grandes variations de températures aux villes avec les températures qui varient le moins. Cela semble cohérent avec l'observation faite précédemment pour le nuage des points variables, d'autant plus que les villes avec les plus grandes variations de températures se sont une abscisse proche de celle qu'avait l'amplitude sur cet axe 2.

### 4. Que représente le vecteur mois ? (Pourquoi tous les mois sont tous d'un même côté ?)

Le vecteur mois représente les températures moyennes de chacune des 23 capitales pour chaque mois. Il est logique que tous les vecteurs se trouvent du même côté car ils sont opposés par l'axe 1 aux vecteurs supplémentaires.

### 5. Quel est le critère pour que les variables quantitatives actives et supplémentaires apparaissent comme bien représentées dans le graphe donné par l'ACP ? Indiquer le(s) vecteur(s) mal représenté(s).

Les variables quantitatives actives et supplémentaires apparaissent comme bien représentées dans le graphe donné par l'ACP si elles sont situées proches de l'origine du plan factoriel.

Un vecteur mal représenté est un vecteur qui n'est pas aligné avec les autres vecteurs du plan factoriel, qui n'a aucune abscisse proche de 1. Le seul vecteur qui semble mal représenté dans ce plan est le vecteur de longitude, qui est sûrement très bien représenté sur un l'axe 3 (vu que la latitude est bien représentée par l'axe 1 et l'amplitude par l'axe 2).

#### **6. Comment interprétez-vous le fait que tous les vecteurs mois soient orientés dans la même direction ?**

Le fait que tous les vecteurs mois soient orientés dans la même direction indique qu'il y a une forte corrélation entre les mois en termes de températures. Cela signifie que les profils de températures des différentes capitales européennes présentent des similitudes dans la façon dont elles varient au cours de l'année. Ce ne serait sûrement plus le cas si on ajoutait des villes de l'hémisphère sud.

#### **7. Comment expliquez-vous le fait que le vecteur Latitude soit anti corrélé avec certains vecteurs mois et en particulier le vecteur moyenne ?**

Le fait que le vecteur Latitude soit anti corrélé avec certains vecteurs mois et en particulier le vecteur moyenne indique qu'il existe une relation entre la position géographique des capitales européennes et leurs profils de températures, et plus précisément suivant leurs latitudes, les villes les plus proches du pôle Nord ont des températures qui diffèrent des villes proches de l'équateur.

#### **8. Comment interprétez-vous le fait que le vecteur Latitude soit orthogonal au vecteur Amplitude ?**

Le fait que le vecteur Latitude soit orthogonal au vecteur Amplitude signifie qu'il n'y a pas de corrélation entre ces deux variables dans les données analysées. En d'autres termes, le changement de latitude ne semble exercer aucune influence sur l'amplitude de la variation de température.

#### **9. Comment interprétez-vous le fait que le vecteur Amplitude soit en partie orthogonal au vecteur moyenne ?**

De la même manière qu'expliqué dans la question précédente l'amplitude n'a probablement aucunes corrélations avec la moyenne des températures. La forte variation des température ne semble donc n'avoir aucune relation avec la valeur de la température moyenne sur l'année.