

TP 1

Introduction

Langage de programmation : Python

But

Dans ce premier TP, Il s'agit d'évoquer les principales structures de données liées à notre domaine telles que :

Les vecteurs, les listes, les matrices de données, les fonctions, le graphisme.

Nous allons nous intéresser aux :

Nuages des points, Graphisme, quelques Fonctions de base

C'est-à-dire :

Plus précisément, nous nous intéressons à la :

- **Représentation graphique de points**
- **Représentation graphique de nuages de points ou bien dispersion de points**
- **Déclaration de matrices de données**
- **Lire les données par lignes et colonnes**

Donc, avant de commencer, il faut tout d'abord installer les bibliothèques suivantes :

- Module **math** : Donne accès aux fonctions sur les réels.
- Module **statistics** : Permet de calculer les valeurs statistiques basiques
- Module **geometry** : variables, fonctions qui calculent l'aire, classes (carré, triangle, ...)
- Bibliothèques **numpy, Scipy, Pandas** : Donnent des packages avec beaucoup de modules pour le calcul numérique scientifique.
- Bibliothèque **matplotlib** : Pour tracer les courbes.

Les utiliser par le biais du suffixe **import** afin de les importer et les exploiter.

Dans ce premier exercice proposé, vous allez essayer de voir de près et d'appliquer les différentes notions vues en cours sur les matrices des données. Aussi, savoir lire et interpréter les résultats obtenus. Vous pouvez faire appel à quelques fonctions prédéfinies en exploitant les bibliothèques de base pour l'analyse factorielle ou bien faire des changements sur les fonctions prédéfinies pour les adapter à nos concepts. Un autre point qui est très important en analyse des données est de visualiser les données et savoir interpréter les graphes obtenus.

Exercice 1

Considérons la matrice des données X suivante, résultat de notes obtenues par 9 étudiants dans 5 modules. Les modules sont Multimedia, Maths, Système, Réseau et Autre.

Modules Individus	Mul	Maths	Système	Réseau	Autre
<i>E 1</i>	<i>6</i>	<i>6</i>	<i>5</i>	<i>5.5</i>	<i>8</i>
<i>E 2</i>	<i>8</i>	<i>8</i>	<i>8</i>	<i>8</i>	<i>9</i>
<i>E 3</i>	<i>6</i>	<i>7</i>	<i>11</i>	<i>9.5</i>	<i>11</i>
<i>E 4</i>	<i>14.5</i>	<i>14.5</i>	<i>15.5</i>	<i>15</i>	<i>8</i>
<i>E 5</i>	<i>14</i>	<i>14</i>	<i>12</i>	<i>12</i>	<i>10</i>
<i>E 6</i>	<i>11</i>	<i>10</i>	<i>5.5</i>	<i>7</i>	<i>13</i>
<i>E 7</i>	<i>5.5</i>	<i>7</i>	<i>14</i>	<i>11.5</i>	<i>10</i>
<i>E 8</i>	<i>13</i>	<i>12.5</i>	<i>8.5</i>	<i>9.5</i>	<i>12</i>
<i>E 9</i>	<i>9</i>	<i>9.5</i>	<i>12.5</i>	<i>12</i>	<i>18</i>

- 1) Déclarer la matrice des données X ainsi que sa transposée X^t . Afficher les deux matrices.
- 2) Donner la liste des individus.
- 3) Extraire les variables dans un vecteur.
- 4) Accéder aux individus : 4, 5 et 7.
- 5) Créer un tableau noté $X(j)$ qui nous permet d'avoir pour chaque variable les informations suivantes : la moyenne arithmétique, la variance, l'écart type. Dans vos résultats (calculs) fixez le nombre de chiffres décimaux à quatre. Afficher ce tableau.
- 6) Calculer l'individu moyen. Afficher le résultat.
- 7) Ecrire une fonction qui calcule la variance des 5 variables de la matrice donnée. Prenez quatre chiffres décimaux.
- 8) Calculer toutes les covariances possibles entre les variables données. Afficher les résultats obtenus dans une matrice notée V .
- 9) Commenter les résultats obtenus de la matrice V .
- 10) Calculer les coefficients de corrélations des variables. Afficher toutes les corrélations dans une matrice notée R .
- 11) Que remarquez-vous ? Commenter les résultats obtenus.
- 12) Ces dernières questions seront consacrées à la représentation graphique d'un nuage de points ou bien points de dispersion ainsi qu'à la représentation simultanée dans plusieurs sous fenêtres.

Représenter graphiquement les individus dans l'espace \mathbb{R}^2 des couples des variables :
 (X^1, X^5) , (X^2, X^4) , (X^3, X^4) .
- 13) Commenter les graphes obtenus. Interpréter la dispersion du nuage de points.