



Track 1 : Analyse de données

MTG miku, Child of Song, illustrated by Aya Kato

Analyse de données (en anglais !)

Évaluation :

Présentation orale : Analyse statistique des données en anglais

Objectifs visés :


Être capable d'appliquer les outils statistiques simples à un ensemble de données
Savoir présenter des résultats d'analyse de données en anglais

Compétences et AC validés :

Compétence 4 AC2 et AC3
Compétence 6 AC1 et AC2

Avis du public

Quel est l'intérêt de l'analyse de données avant la conception d'un système de recommandation ?



Quel est l'intérêt de l'analyse de données avant la conception d'un système de recommandation ?

- Trouver des corrélations à exploiter
- Déceler les failles éventuelles
- Creuser de nouvelles idées

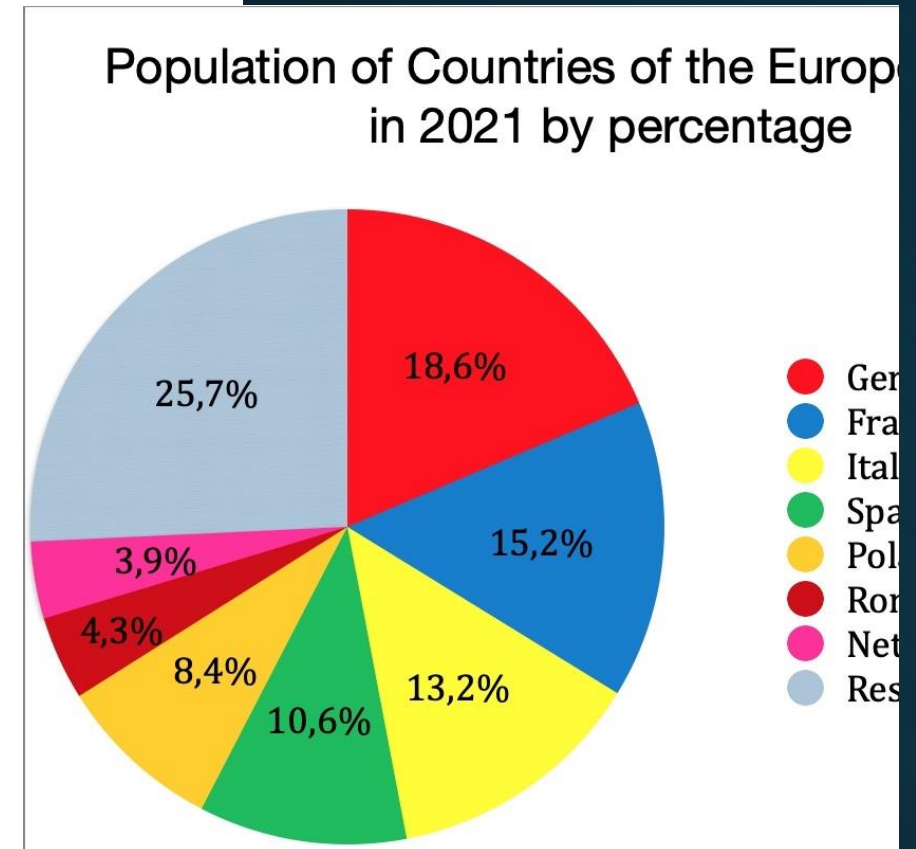
Attentes minimales d'un graphique

- Titre du graphique
- Axes avec graduation et titres
- Légende
- Lisibilité

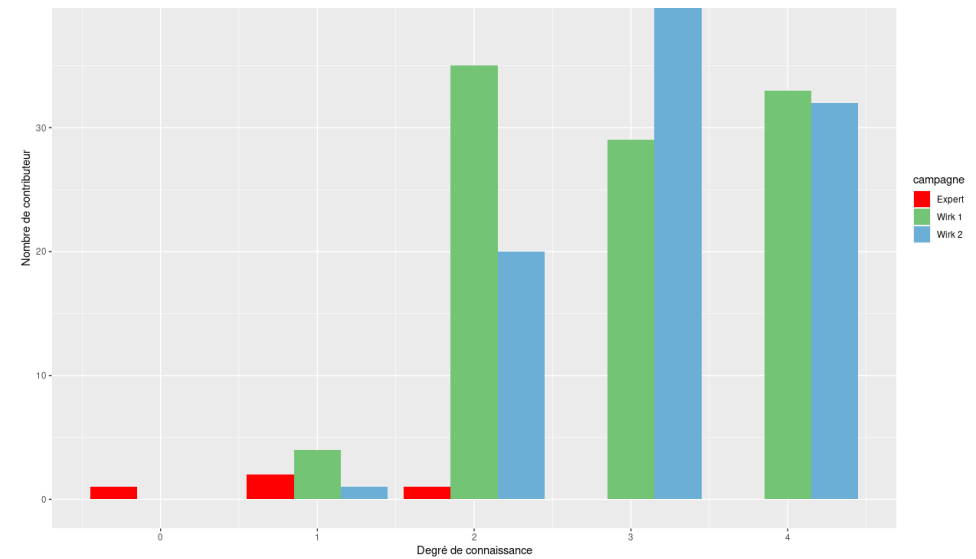
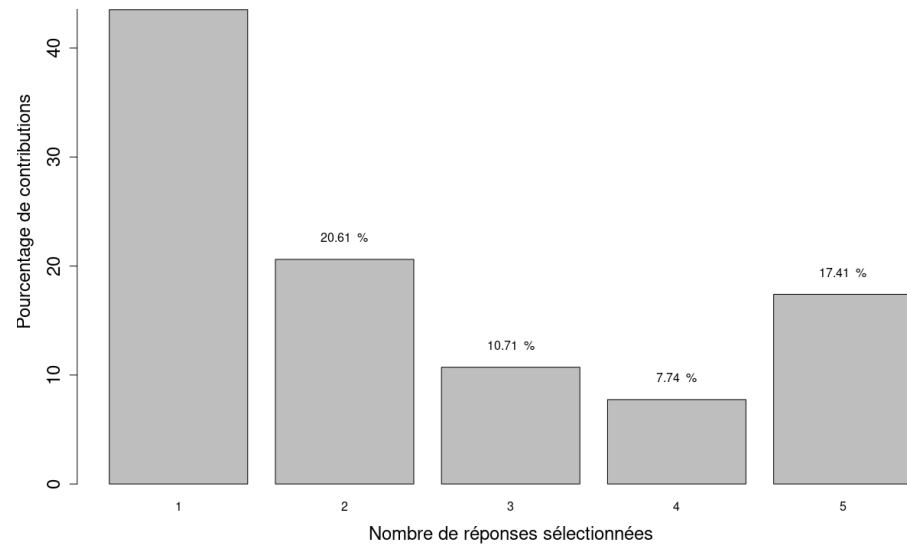


Pie chart

- Représentation visuelle de l'importance d'une catégorie
- En python
- `Pandas.DataFrame.plot.pie`
- `Matplotlib.pyplot.pie`



(Grouped) Bar chart



- En python
- `Pandas.DataFrame.plot.bar`
- `Matplotlib.pyplot.bar`

Box plot chart

Visualisation résumée des statistiques :

Min, Max, Q1, Q3 et médiane

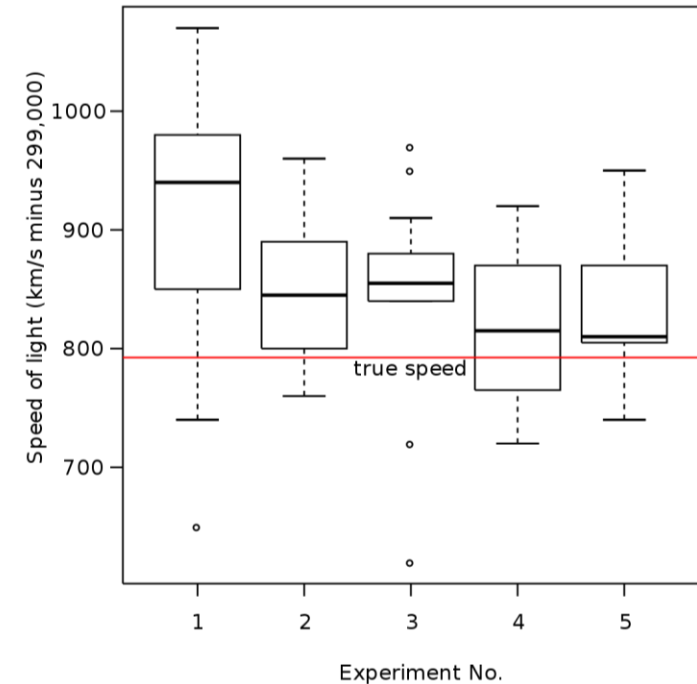
Attention :

$\text{Min} = Q1 - 1,5 * IQ$

$\text{Max} = Q3 + 1,5 * IQ$

En python

- `Pandas.DataFrame.box plot`
- `Matplotlib.pyplot.boxplot`



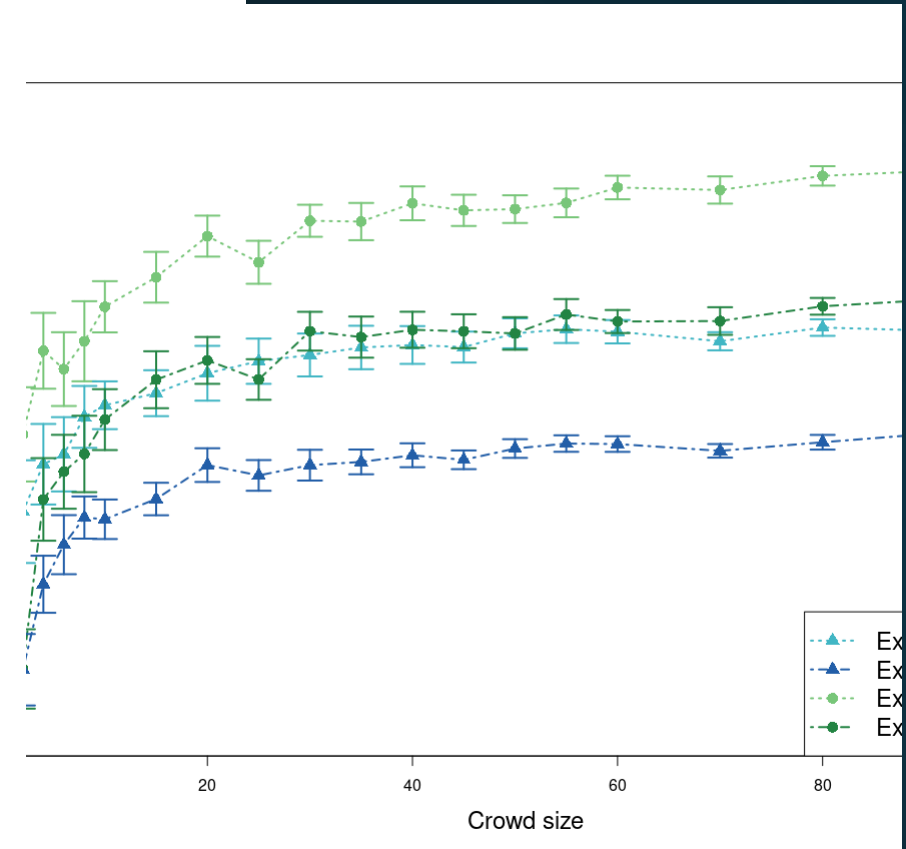
Plot chart

Représentation graphique avec un intervalle de confiance à 95%

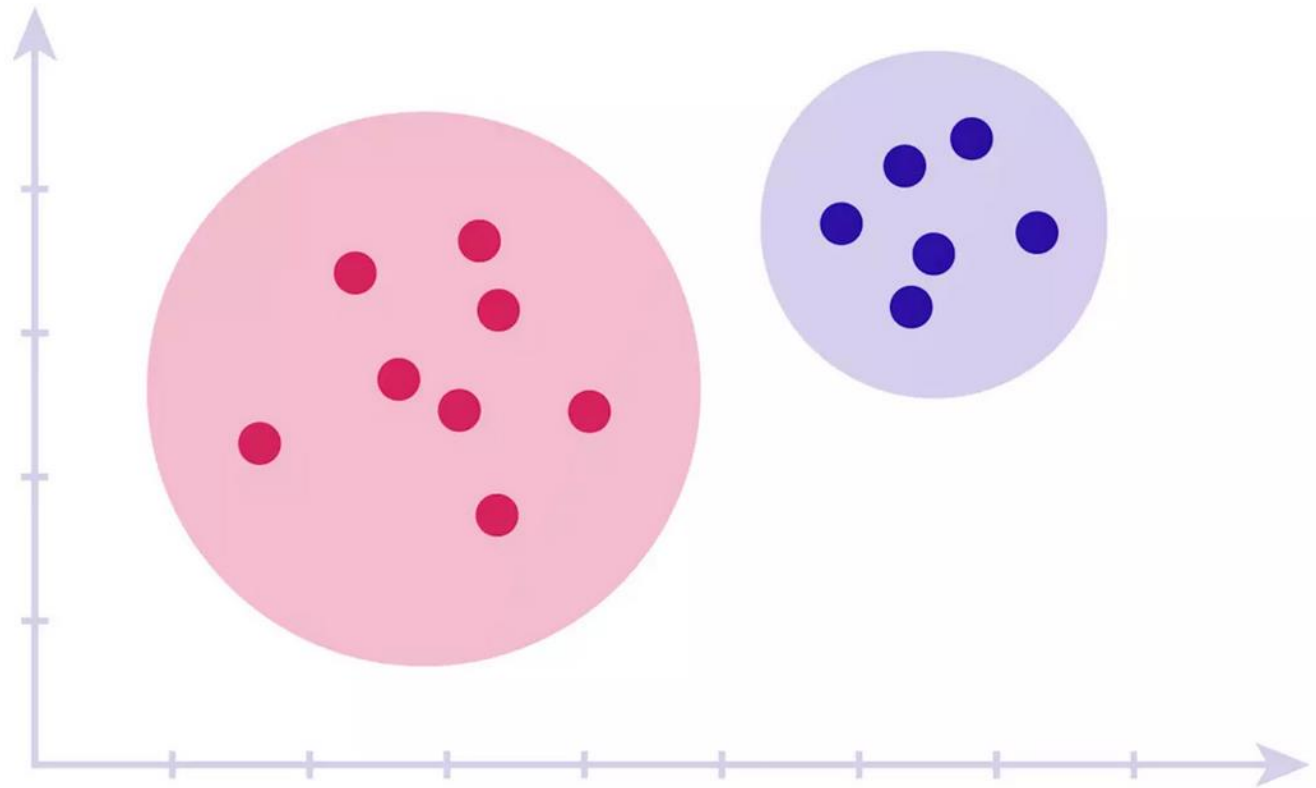
$$(\pm) 1.96 * (\text{std}(X)/\text{sqrt}(\text{length}(x)))$$

En python

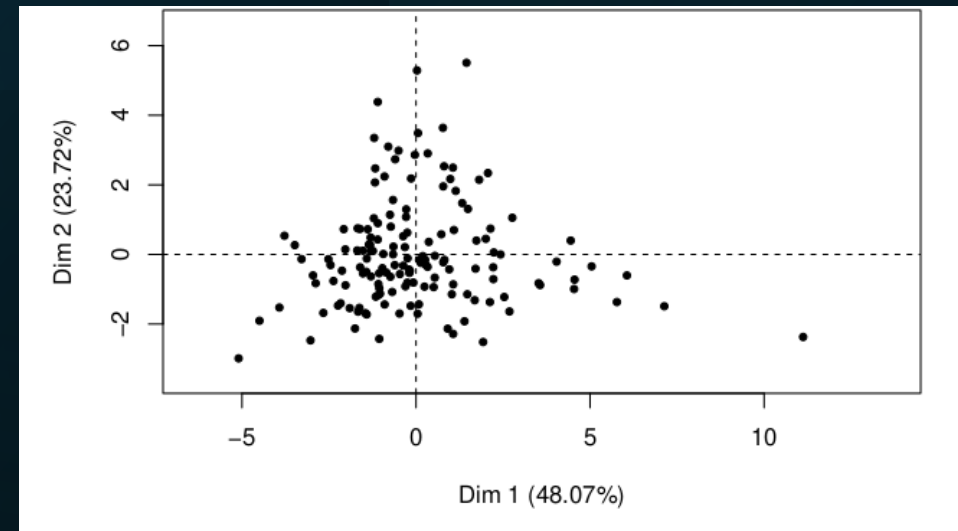
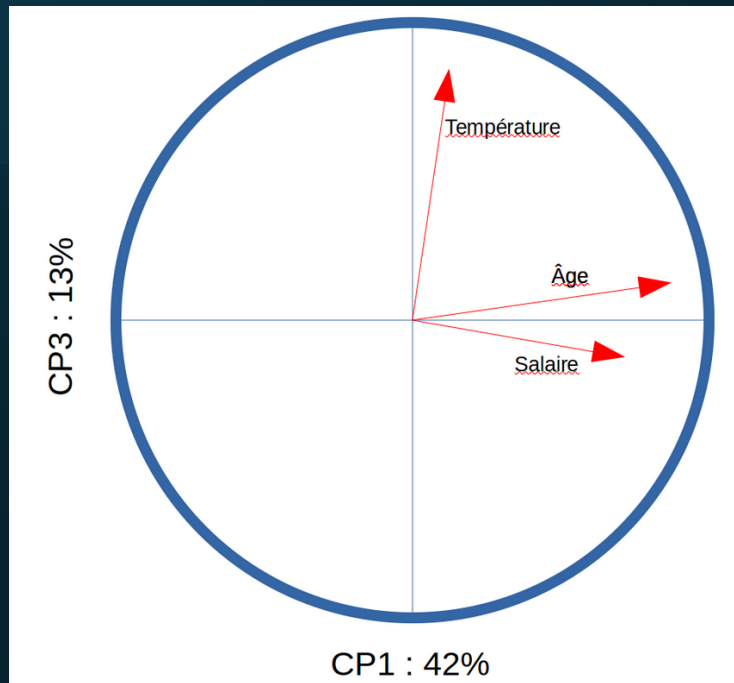
- `Pandas.DataFrame.plot`
- `Matplotlib.pyplot.plot`



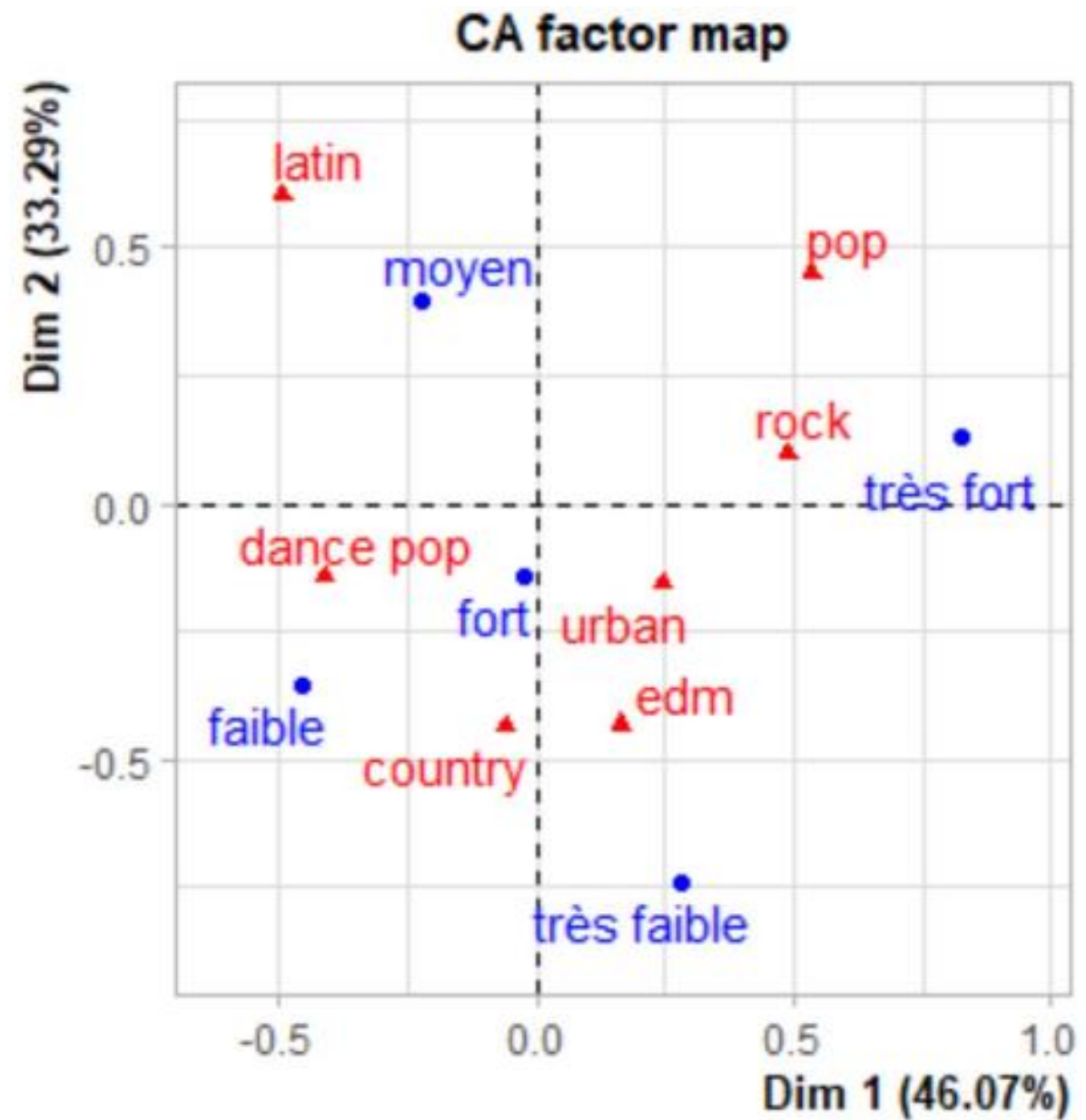
Clustering



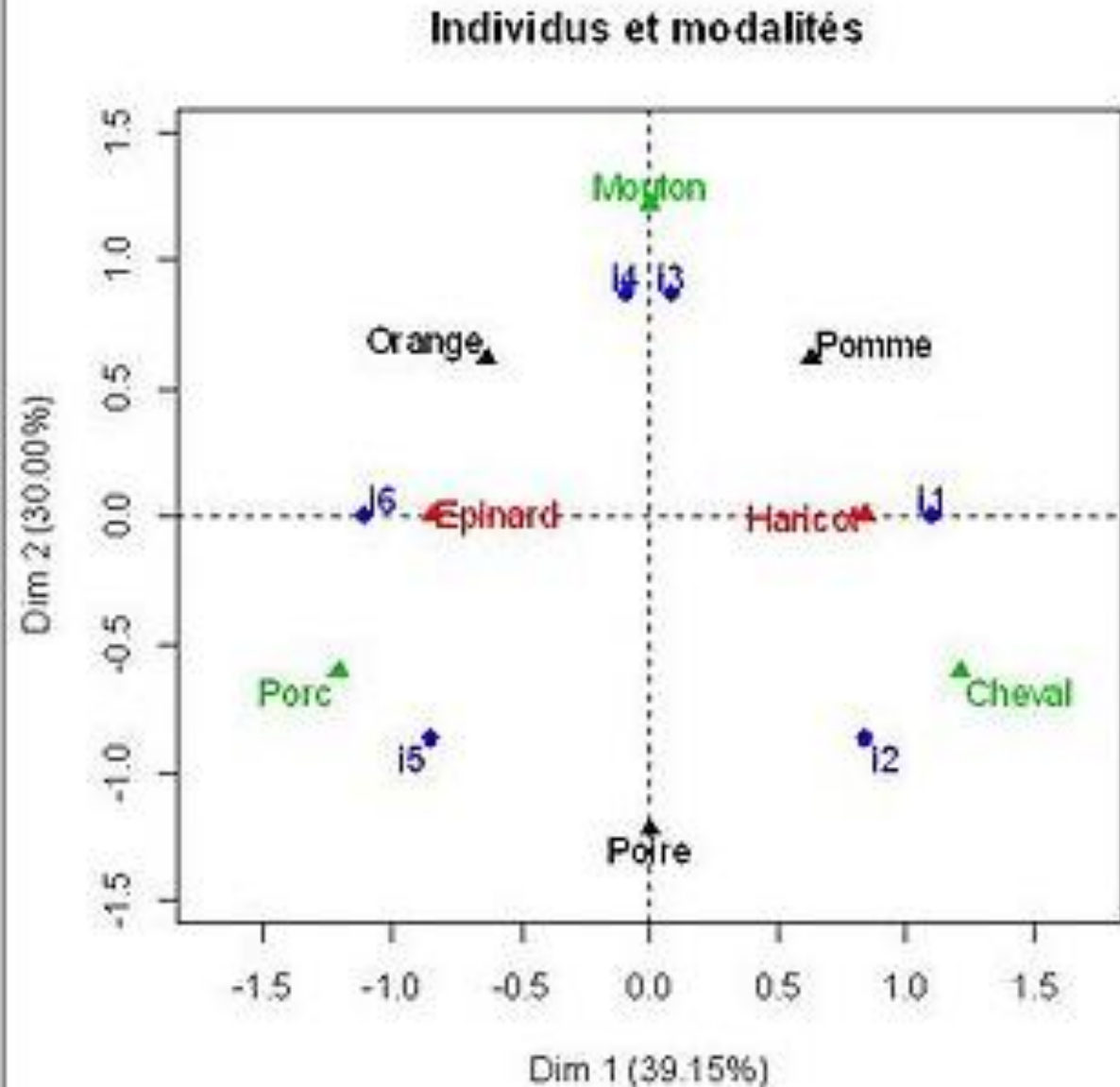
Analyse en composante principale



Analyse Factorielle des correspondances



Analyse en composantes Multiples



Contraintes

FI : MuSE

- 6 graphes différents au minimum
- Au minimum 2 graphes de technique d'IA :
 - 1 pour des données quantitatives
 - 1 pour des données qualitatives

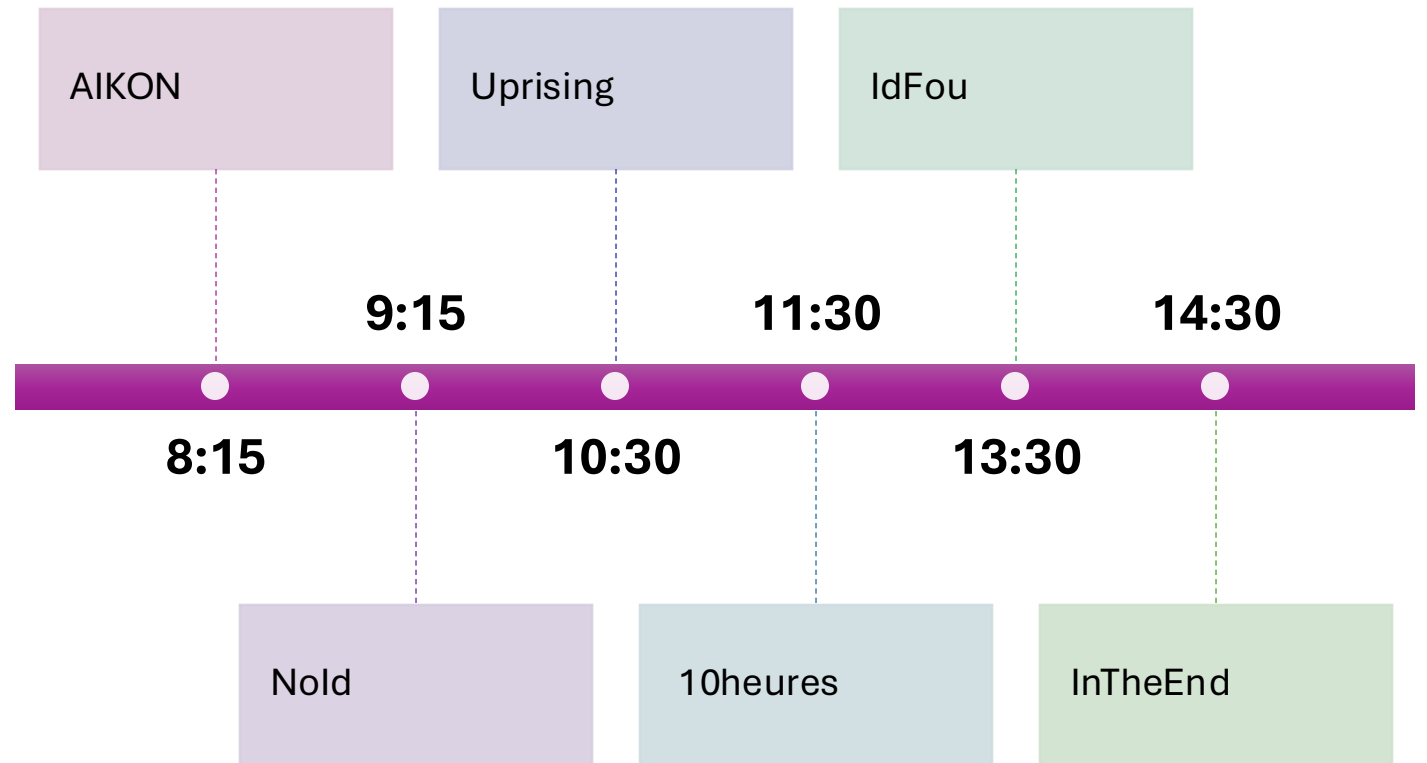
Alt: MYLi

- 2 graphes min dont un par IA

Critères d'évaluation

- Nettoyage des données avant analyse
- Sélection d'ensemble de données appropriées
- Compréhension des données en anglais
- Génération des graphiques
- Explication claire des résultats en anglais
- Titres et légendes en anglais
- Pertinence des méthodes employées
- Recul avec la réalité
- Analyse pour l'exploitation en vue du système de recommandation
- Code
- Gestion du temps de parole (**FI : 25 min de présentation + 15 min questions**) (**ALT : 10 min + 5min**)
- Réponses aux questions
- Qualité du support de présentation

Ordre de passage 17/10



Ordre de passage 21/11

