



Now Playing



Filière GéoDataScience  
ENSG - Géomatique

Mélodie FLEURY  
Chaima HASDI  
Fosio ULUTUIPALELEI

# Analyse géostatistique : Spotify vs FMA

Décembre 2023



01

**Introduction**

02

**Présentation des données**

03

**Analyse statistique**

04

**Analyse spatiale**

05

**Conclusion**



Now playing



# Introduction

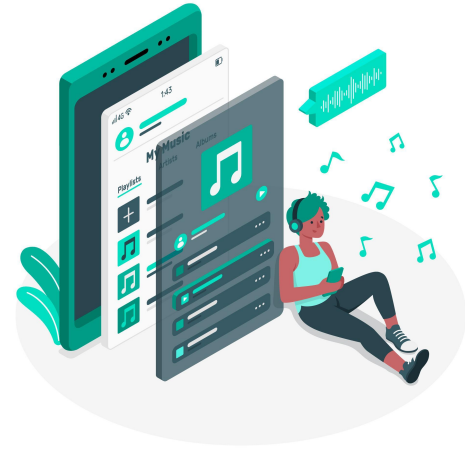
1:23



2:36



- Fin année 2000 : Émergence de plateformes d'écoutes payantes ou gratuites avec la numérisation de la musique
- Possibilité d'accès aux données relatives au comportements des utilisateurs de ces plateformes



**“Quels sont les facteurs contribuant à la popularité d’une musique considérée comme un hit comparés à une chanson amateur selon différentes échelles spatiales ?”**



# Présentation des données

1:23



2:36

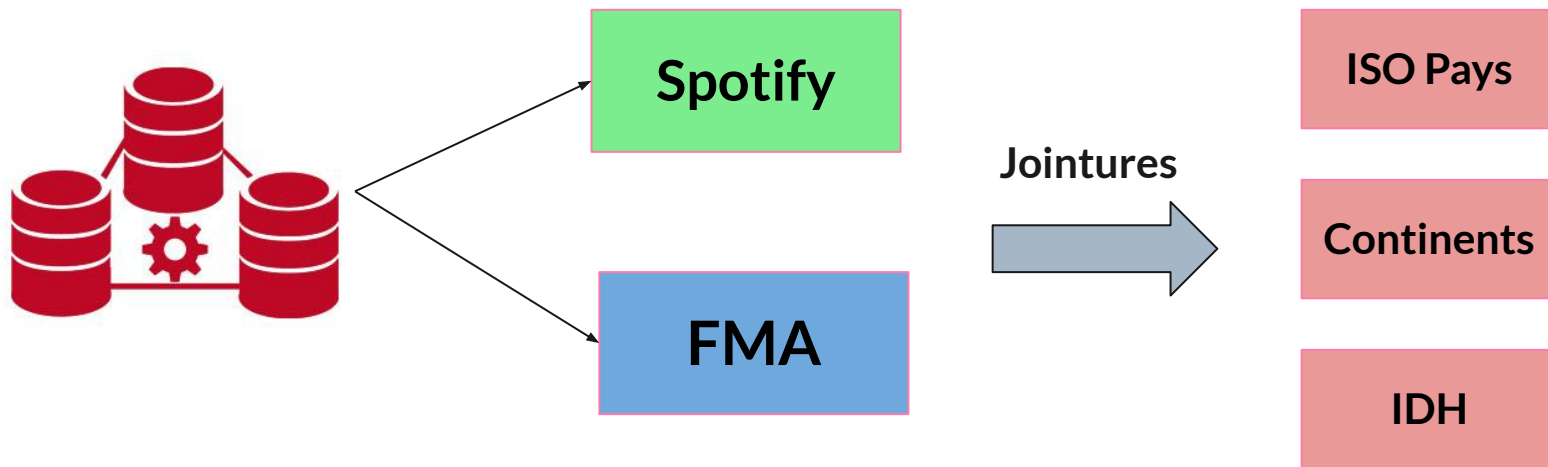


- 18 - 28 Octobre 2023
- Kaggle.com



- 100 000 chansons
- Avril 2017
- Répertoire Github

Sélection d'attributs





# Analyse statistique

1:23



2:36



- Attribut principale : “listen”
- Aplatissement vers la gauche des données => Jeu de donnée hétérogène

Listen	
Min	12
Median	694
Max	543252

=> peu de chansons beaucoup écoutés

/	Speechiness	Liveness
Min	0,022	0,025
Median	0,040	0,12
Max	0,96	0,98

- Musique caractérisée par la présence de musique
- Majorité des chansons produites en studio

- Peu de corrélations entre les variables : **14%** > 0,1 et différent de 1

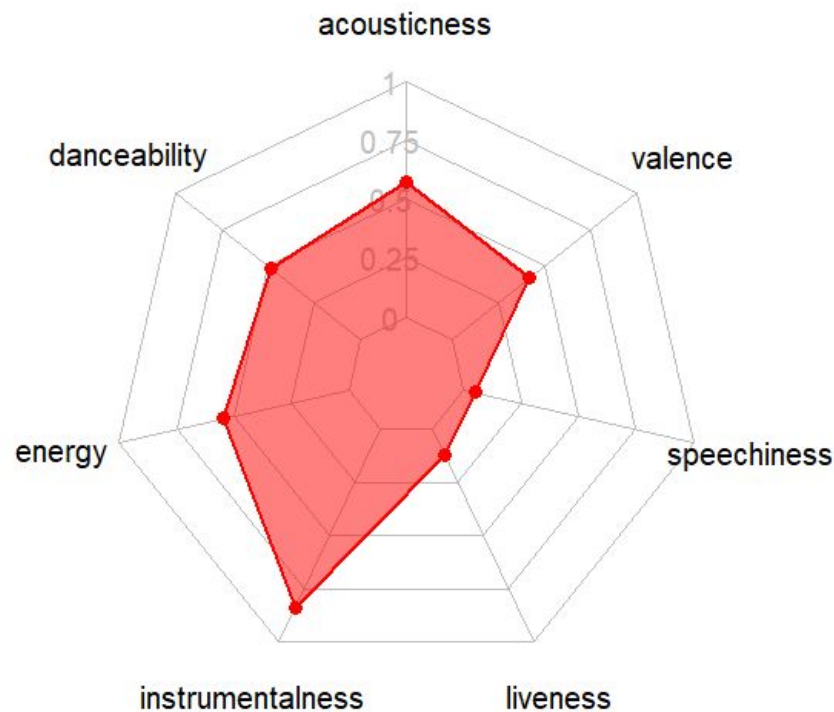
=> Jeu de données très hétérogène et assez indépendant

- Corrélation forte entre “listens” et “favorites” : **0,84**

=> Plus une chanson est écoutée, plus elle est mise en favori

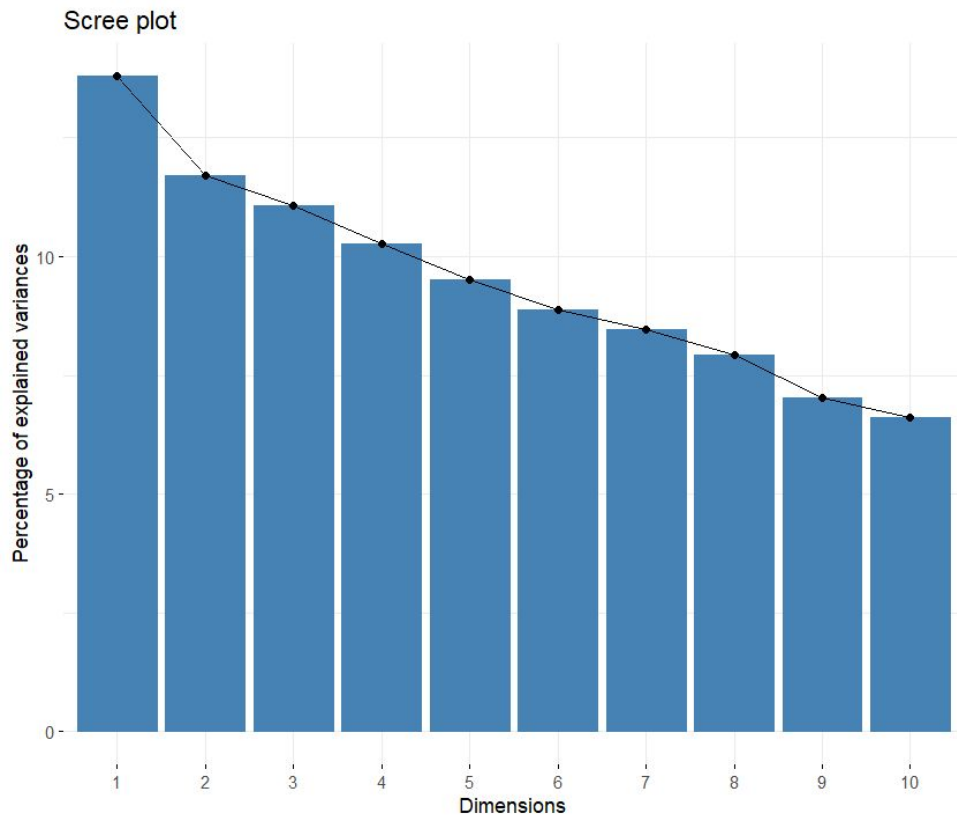
- Graphique radar des caractéristiques musicales (de 0 à 1)
- Prédominance de “instrumentalness”  
=> chansons peu de paroles

=> Lien entre “instrumentalness” et popularité d’une chanson



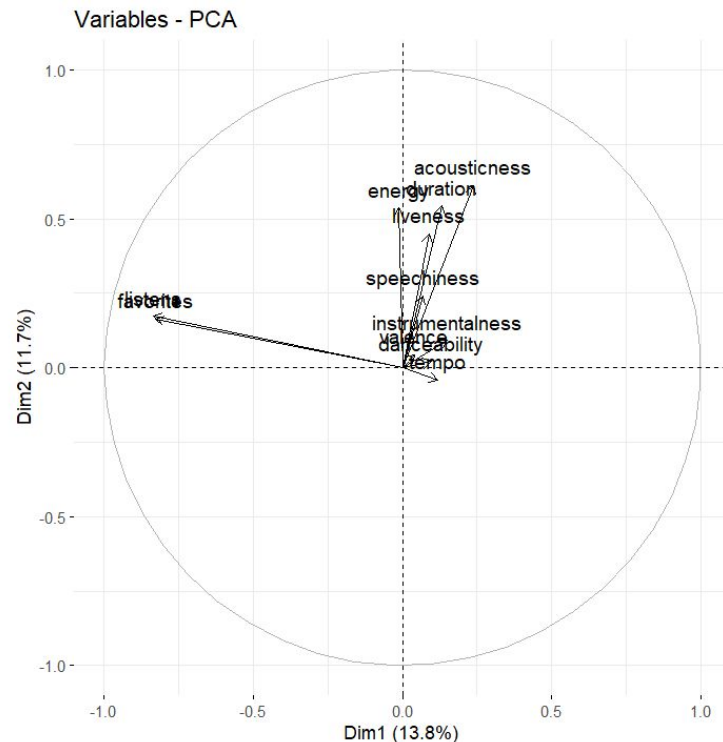
- Objectif : Réduire la dimensionnalité tout en restituant au mieux l'information
  - Scree plot : Capture de l'inertie équivalente
- => Composantes ne résument pas bien l'information des variables

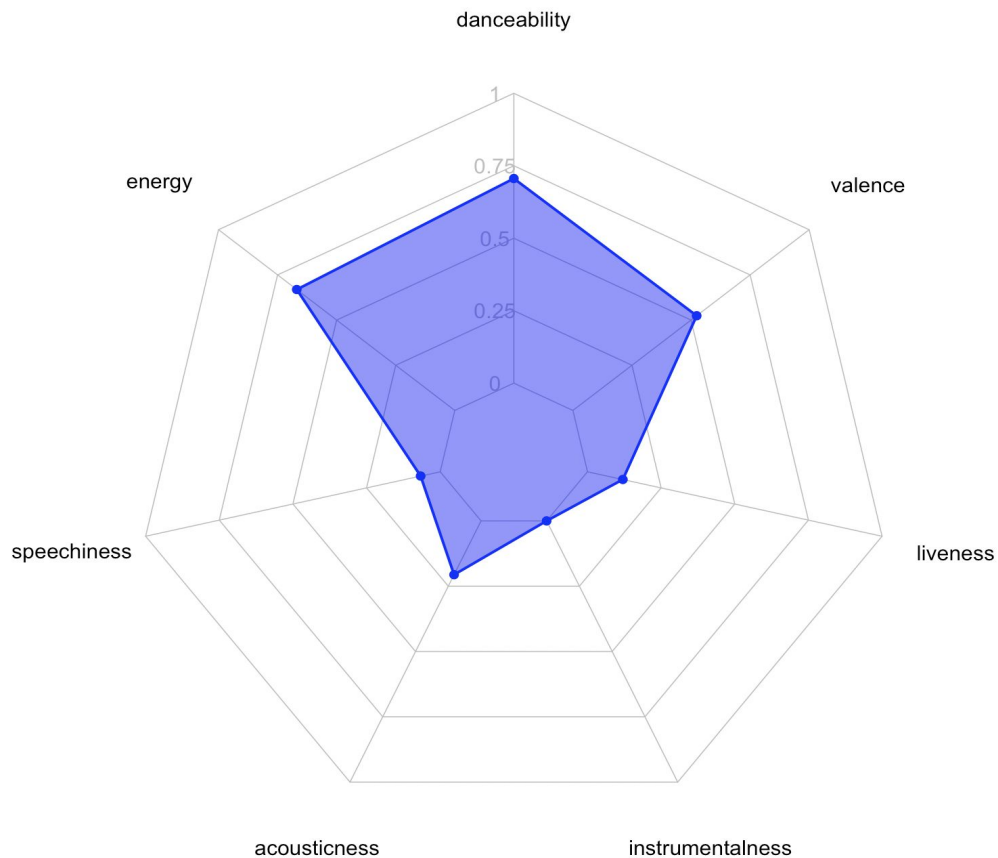
=> Jeu de données hétérogène

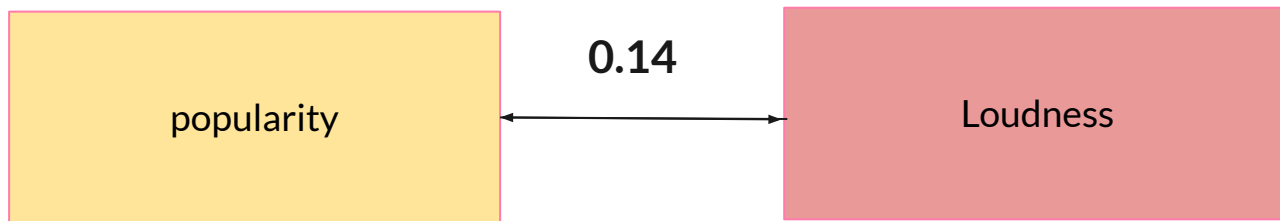


- Étude sur 25,5 % de l'inertie totale => Resté critique des résultats
- Mauvaise représentation globale des variables sauf pour "listen" et "favorites"
- 2 groupes => Écoute aléatoire des auditeurs ou variété grande des chansons proposée sur FMA

=> Pas de liens entre facteurs musicaux et de popularité

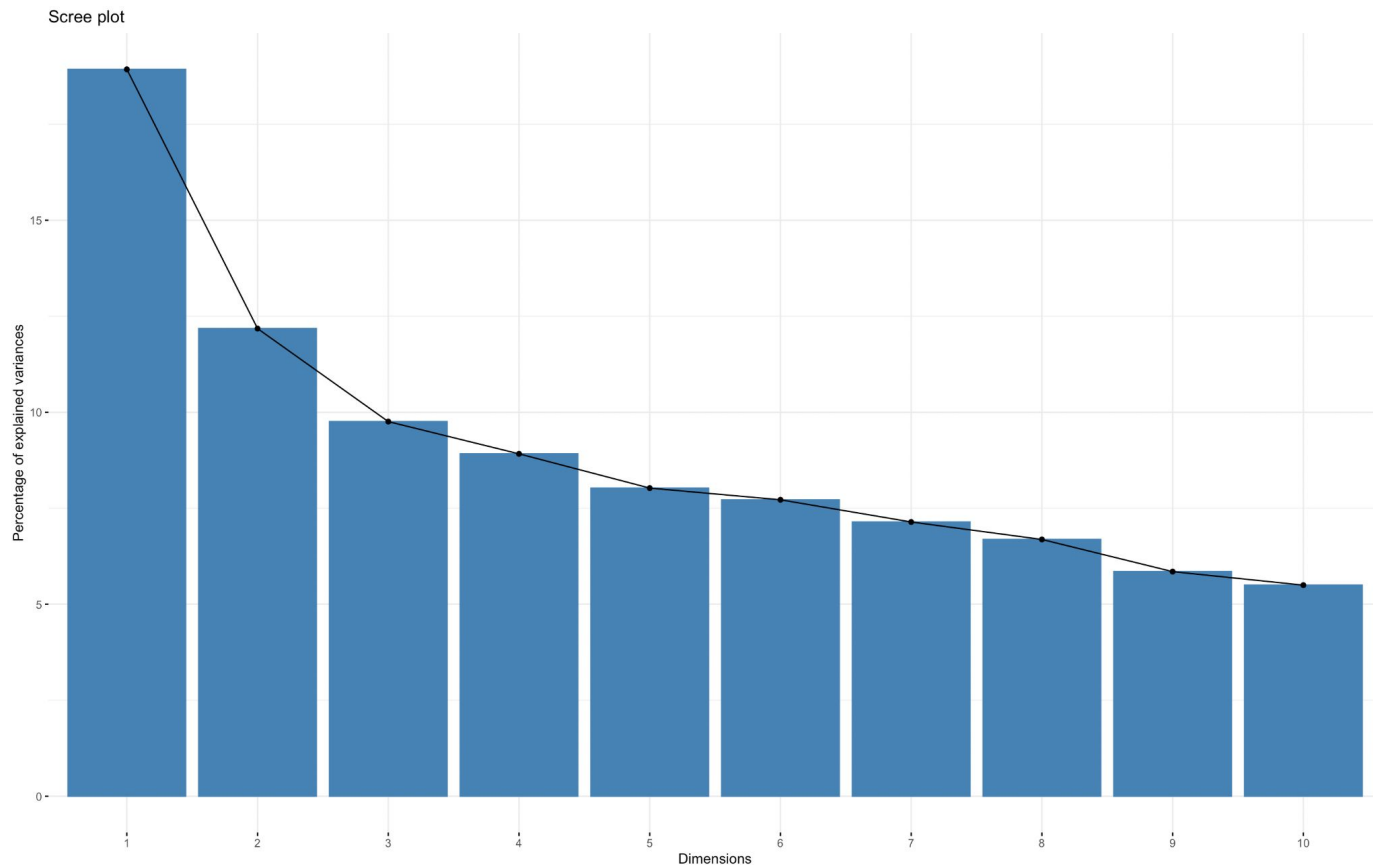




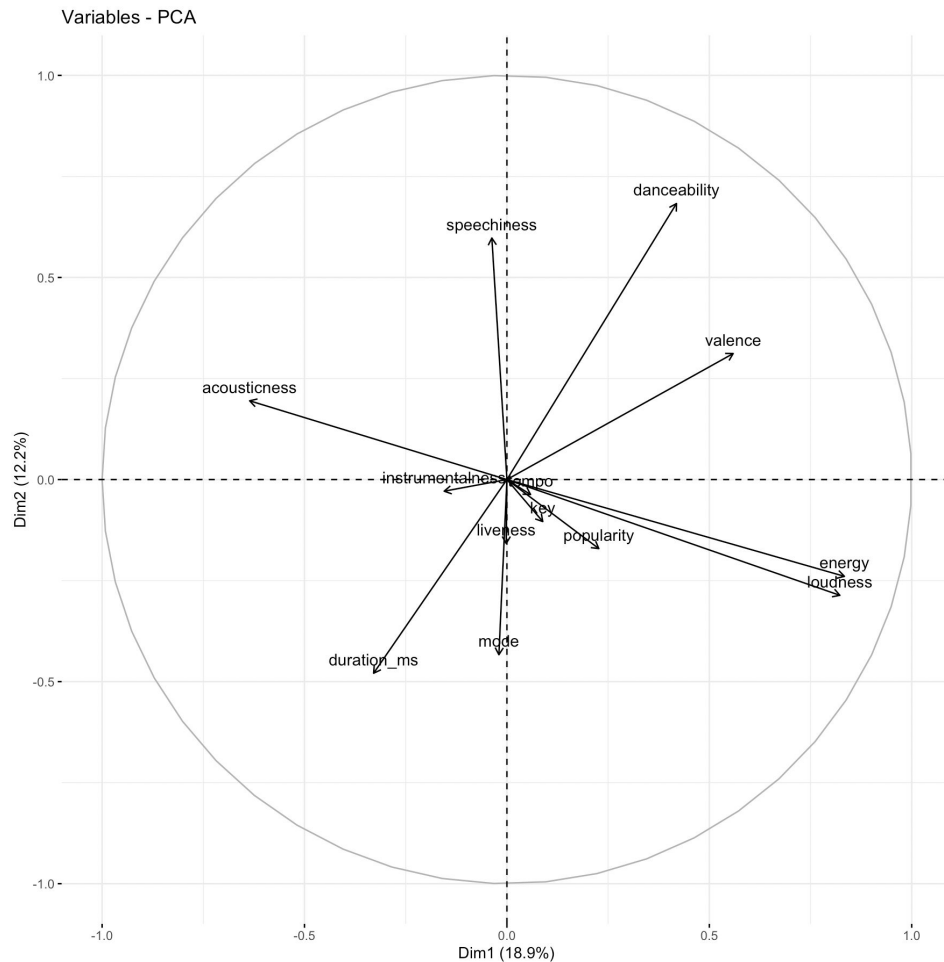


- Très faibles corrélation avec les autres attributs

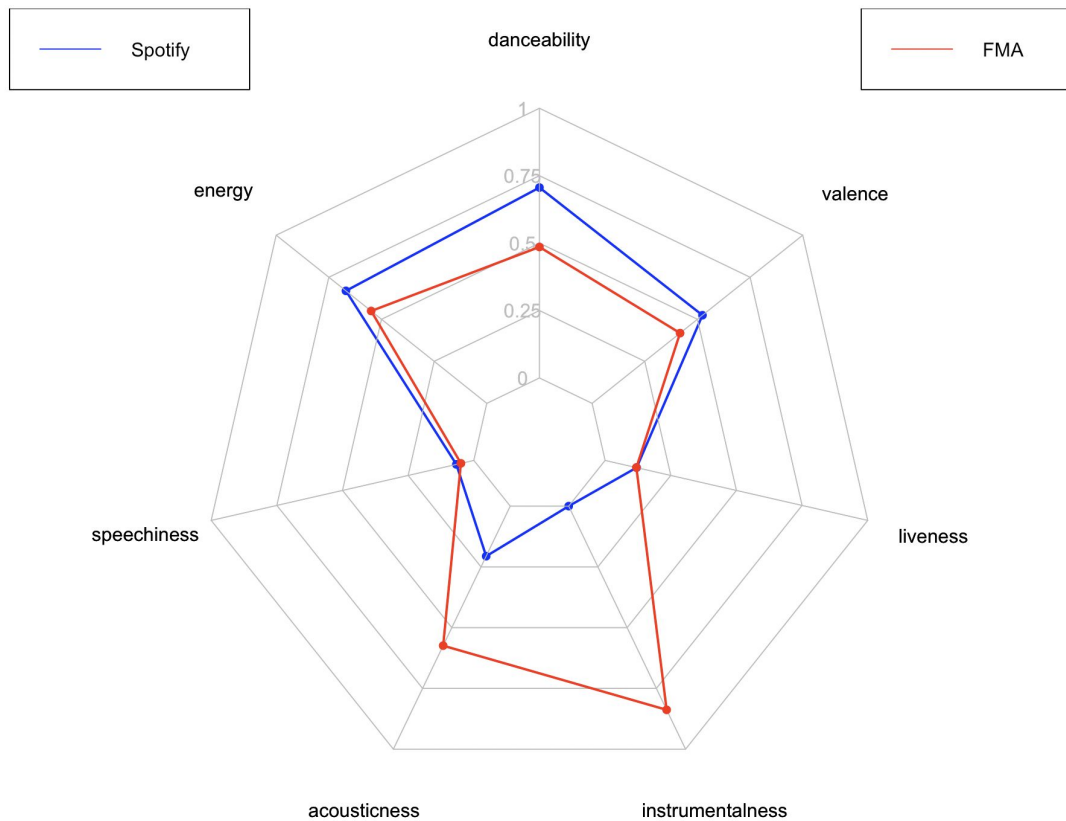
- 2 ou 3 composantes principales







- Inertie captée = 30%





# Analyse spatiale

Étude de l'IDH et des croyances populaires  
Nord/Sud associées

1:23



2:36

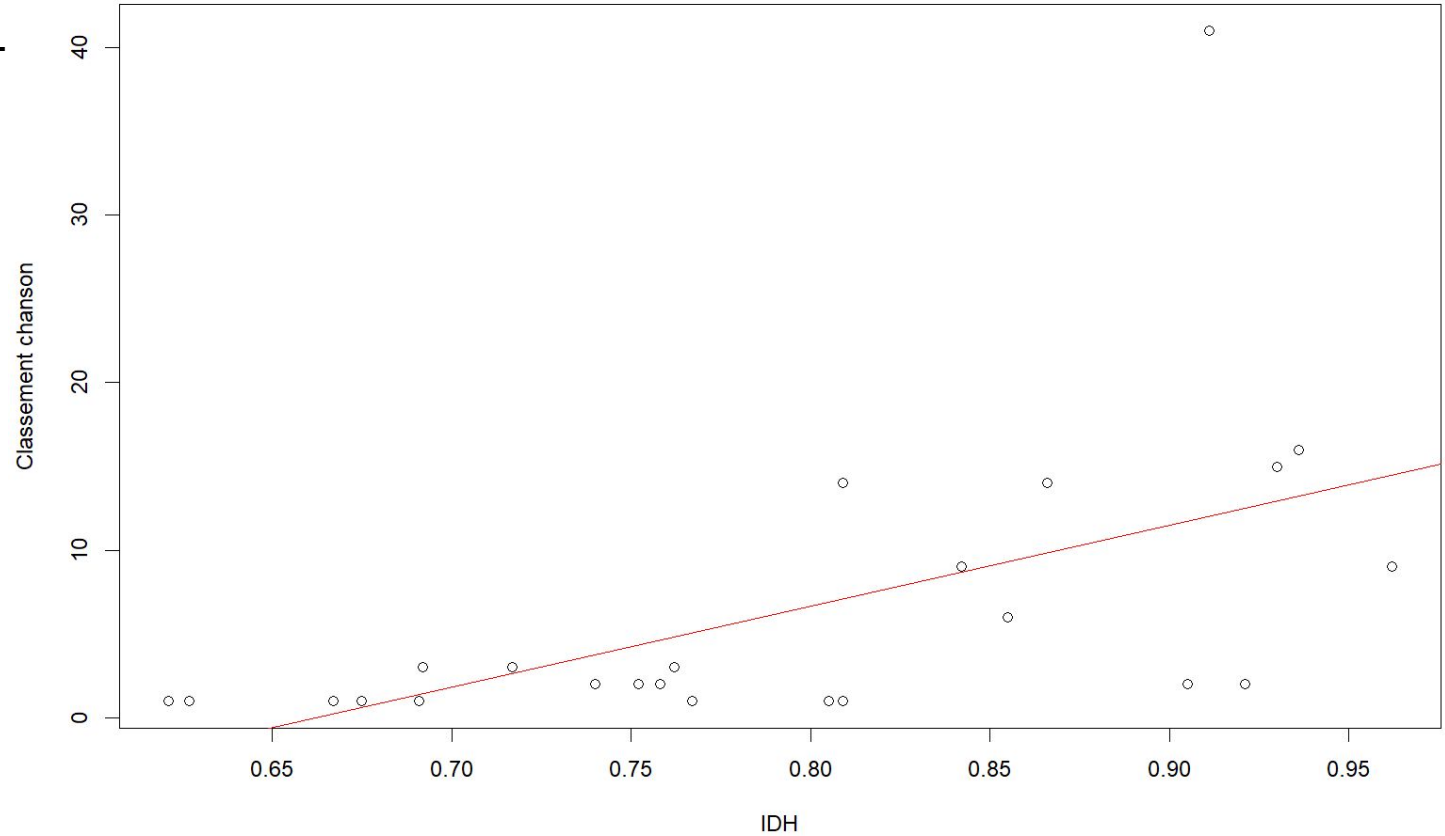


# Relation popularité d'une chanson & IDH du pays



La p-value = **0.0051**

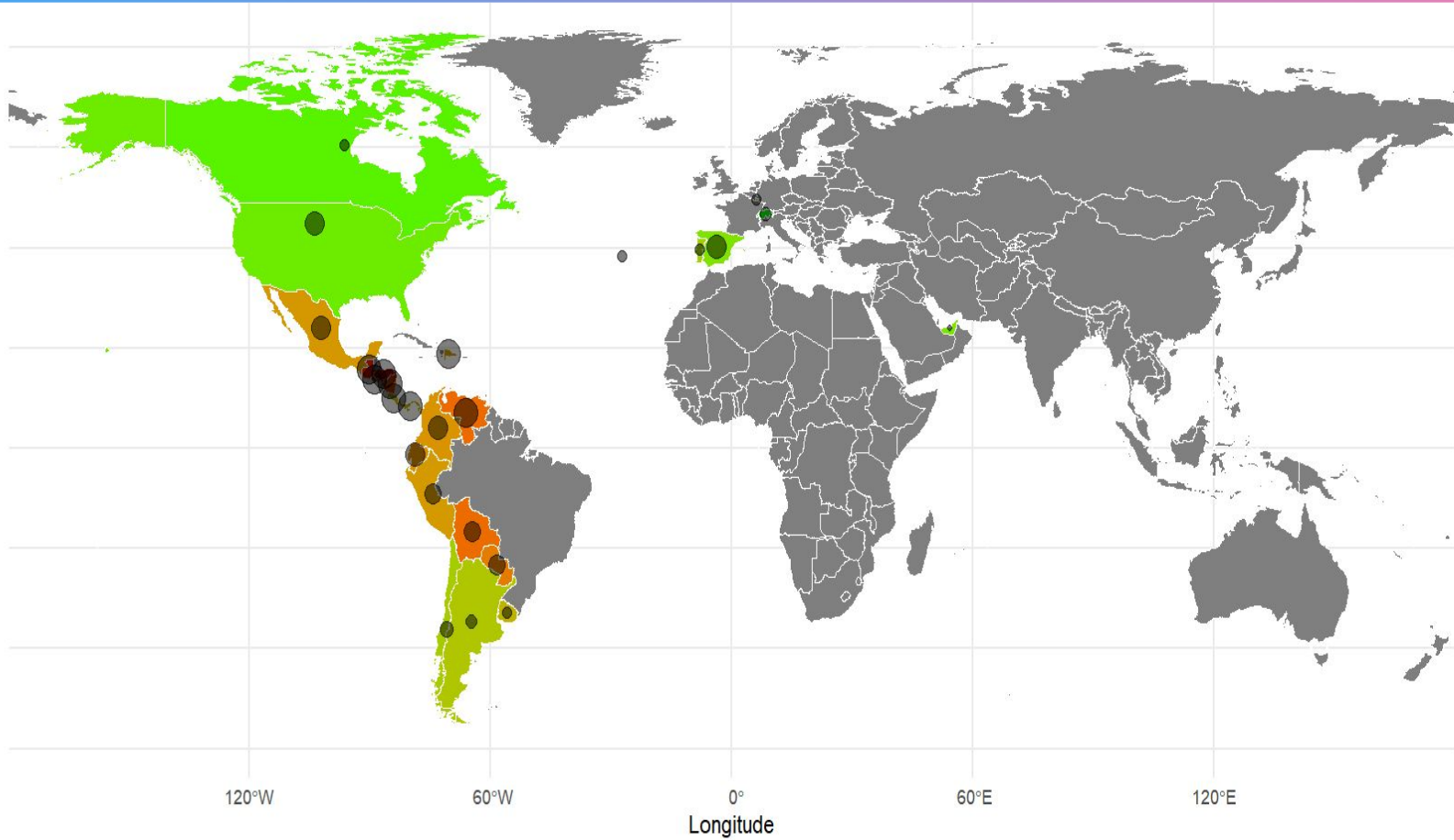
Le  $R^2 = 0.3054$



# < Relation popularité d'une chanson & IDH du pays



Latitude

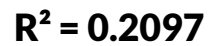


Popularité

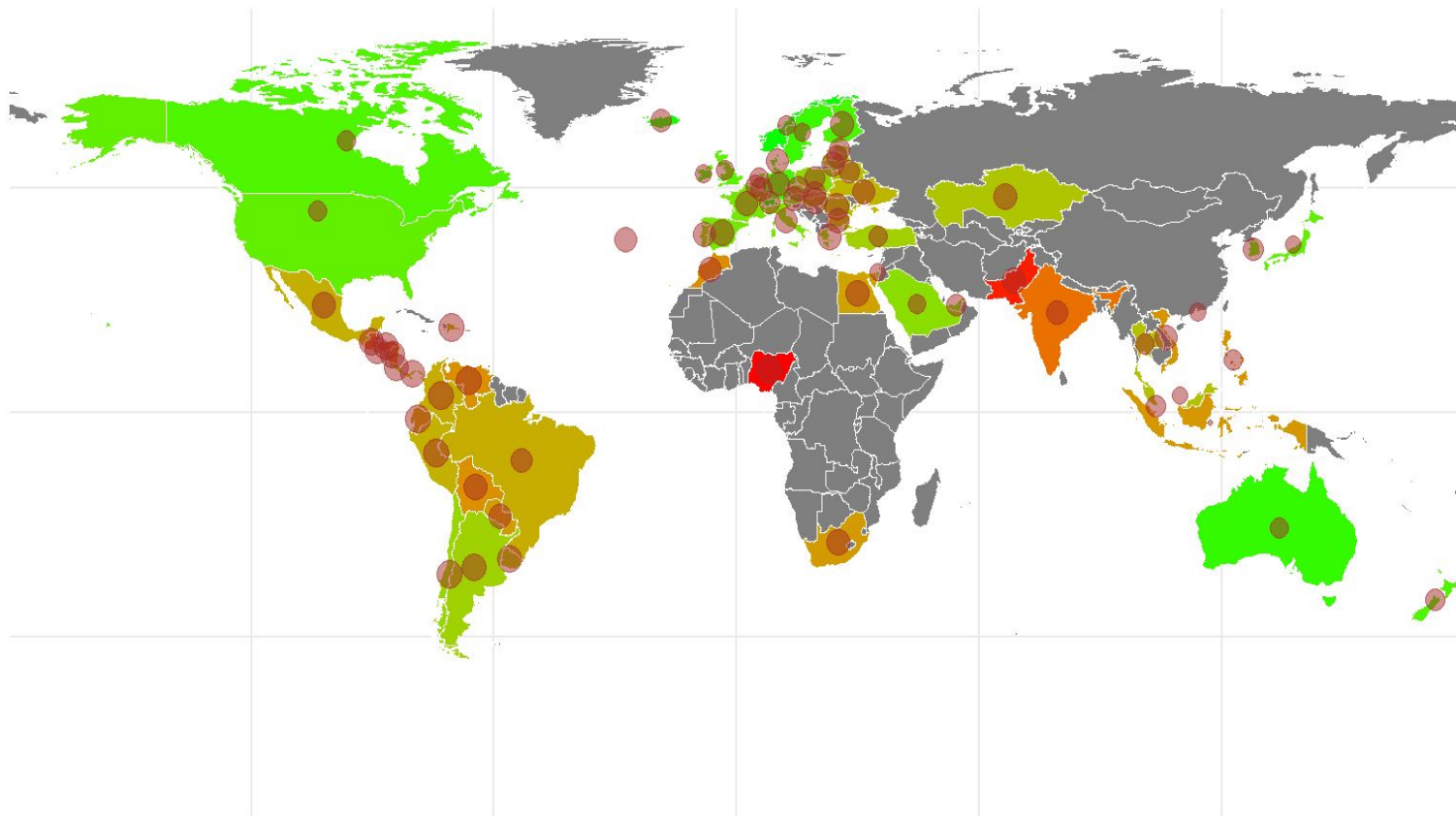


IDH





Latitude





Now playing



# Conclusion

1:23



2:36







Facteurs musicaux



Facteur social



Facteur choix plateforme

## LES LIMITES

- Biais sélection et promotion des chansons sur Spotify → impact des labels
- Moyen d'accès au jeu de données FMA par auteur répertoire Github

## LES PERSPECTIVES

- Analyse d'attributs non-traités :
  - => "is\_explicit", "daily\_movement" pour Spotify
  - => "language\_code" et "date\_recorded" pour FMA
- Améliorer les résultats d'ACP FMA : vérifier si distribution aléatoire pour FMA

# Merci de votre attention

Avez-vous des questions ?



Now Playing



Filière GéoDataScience  
ENSG - Géomatique

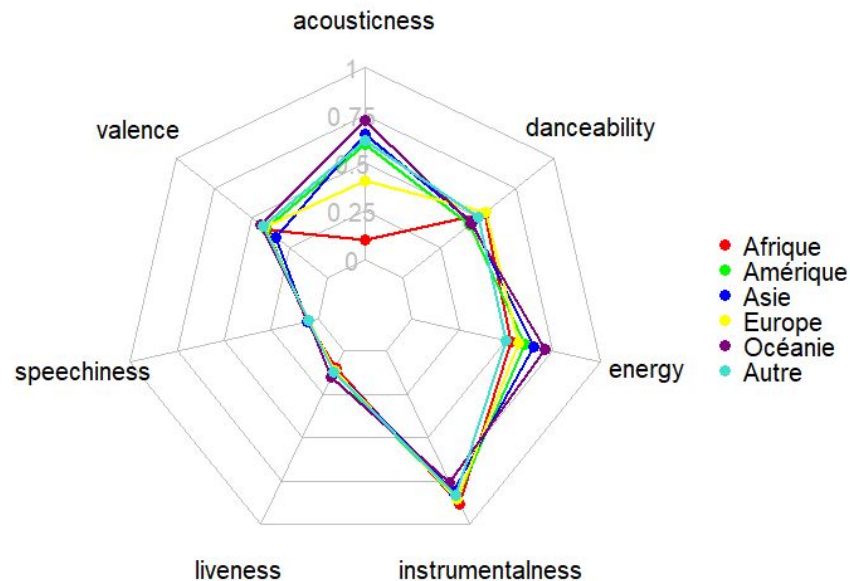
Mélodie FLEURY  
Chaima HASDI  
Fosio ULUTUIPALELEI

# Analyse géostatistique : Spotify vs FMA

Décembre 2023



- Graphique en radar des profils médian des chansons FMA en fonction du continent
- Catégorie "Autre" => Pas de pays associé à la chanson
- Peu de différences de profil entre les continents ("valence", "liveness")
- "Acousticness" pour l'Afrique et élevé pour l'Océanie => valeurs peu représentatives (avec 3 et 50)



=> Pas de lien avéré entre facteur géographique et la popularité d'une chanson

