



Université Claude Bernard



Lyon 1

## Analyse des Données Musicales de Spotify

Mecheri Jawad, Mayouf Lotfi, Ishak Bouchlaghem

## Résumé

Dans ce rapport, nous présentons une analyse approfondie des données musicales provenant de l'API Spotify. L'étude couvre la préparation et le nettoyage des données, une analyse exploratoire détaillée des tendances musicales, une segmentation des morceaux basée sur le clustering, ainsi qu'une analyse des collaborations artistiques à l'aide de graphes. Les résultats mettent en évidence les tendances clés de l'industrie musicale contemporaine, l'importance des caractéristiques audio dans la popularité des morceaux, et la structure complexe des réseaux de collaborations entre artistes. Nous avons travaillé ensemble de manière équitable sur l'ensemble du projet, sans qu'aucune personne ne soit affectée seule à une partie spécifique.

# Table des matières

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>3</b>  |
| <b>2</b> | <b>Présentation des Données</b>  | <b>3</b>  |
| 2.1      | Origine des Données . . . . .  | 3         |
| 2.2      | Description des Jeux de Données . . . . .                                    | 3         |
| 2.2.1    | Dimensions Initiales . . . . .   | 3         |
| 2.2.2    | Colonnes Principales . . . . .   | 3         |
| 2.2.3    | Étendue Temporelle . . . . .   | 4         |
| 2.3      | Nettoyage et Préparation des Données . . . . .                               | 4         |
| <b>3</b> | <b>Analyse exploratoire des données musicales</b>                            | <b>5</b>  |
| 3.1      | Évolution de la popularité et des collaborations . . . . .                   | 5         |
| 3.2      | Analyse des caractéristiques musicales par année . . . . .                   | 6         |
| 3.3      | Corrélations entre popularité et caractéristiques musicales . . . . .        | 6         |
| 3.4      | Genres musicaux et popularité . . . . .                                      | 8         |
| 3.5      | Durée et caractéristiques des morceaux populaires . . . . .                  | 9         |
| <b>4</b> | <b>Analyse des clusters : Une segmentation claire des morceaux musicaux</b>  | <b>10</b> |
| 4.1      | Caractéristiques moyennes des clusters . . . . .                             | 10        |
| 4.2      | Popularité des clusters . . . . .  | 11        |
| 4.3      | Répartition des genres par cluster . . . . .                                 | 12        |
| 4.4      | Visualisation 2D des clusters et insights . . . . .                          | 12        |
| 4.5      | Synthèse des résultats . . . . .   | 13        |
| 4.6      | Insights clés et recommandations . . . . .                                   | 13        |
| <b>5</b> | <b>Analyse des communautés et collaborations artistiques via les graphes</b> | <b>14</b> |
| 5.1      | Structure globale du graphe . . . . .  | 14        |
| 5.2      | Analyse des principales communautés . . . . .                                | 16        |
| 5.2.1    | Communauté 10 (Reggaeton/Latin) . . . . .                                    | 16        |
| 5.2.2    | Communauté 0 (Rap/Hip-Hop international) . . . . .                           | 17        |
| 5.2.3    | Communauté 3 (Pop/Dance) . . . . .   | 18        |
| 5.2.4    | Communauté 2 (Pont Latino/Occidental) . . . . .                              | 19        |
| 5.2.5    | Communauté 23 (Mexicain/Latin) . . . . .                                     | 20        |
| 5.3      | Communautés spécifiques et isolées . . . . .                                 | 21        |
| 5.3.1    | Communauté 17 (Dance/Funk brésilien) . . . . .                               | 21        |
| 5.3.2    | Communauté 24 (Rap German/Français) . . . . .                                | 22        |
| 5.4      | Insights généraux . . . . .  | 23        |
| <b>6</b> | <b>Conclusion</b>  | <b>23</b> |

# 1 Introduction

L'industrie musicale a connu des transformations majeures avec l'avènement des plateformes de streaming comme Spotify, qui offrent un accès sans précédent à une vaste bibliothèque musicale et aux métadonnées associées. Comprendre les tendances musicales actuelles, les caractéristiques qui rendent un morceau populaire, et les dynamiques de collaborations entre artistes est essentiel pour les acteurs du secteur. Ce rapport vise à analyser les données musicales fournies par l'API Spotify afin de mettre en lumière ces aspects. Nous aborderons la présentation des données, leur nettoyage et préparation, une analyse exploratoire des tendances musicales, une segmentation des morceaux via le clustering, et une analyse des communautés artistiques à travers les graphes.

## 2 Présentation des Données

### 2.1 Origine des Données

Les données utilisées dans ce projet proviennent d'un dataset issu de l'API Spotify, une plateforme reconnue pour son immense bibliothèque musicale et ses métadonnées associées. L'API Spotify fournit un accès programmatique aux informations des artistes, des morceaux et des caractéristiques audio détaillées, permettant ainsi une analyse approfondie des tendances musicales. La documentation complète de l'API est disponible sur Spotify Developer.

### 2.2 Description des Jeux de Données

Deux fichiers principaux ont été exploités pour cette étude :

1. **artists.csv** : Contient les informations sur les artistes.
2. **tracks.csv** : Contient les informations sur les morceaux musicaux et leurs caractéristiques.

#### 2.2.1 Dimensions Initiales

- **artists.csv** : 1 104 349 lignes et 5 colonnes.
- **tracks.csv** : 586 672 lignes et 20 colonnes.

#### 2.2.2 Colonnes Principales

**artists.csv** :

- **id** : Identifiant unique de l'artiste.
- **followers** : Nombre de followers Spotify de l'artiste.
- **genres** : Genres associés à l'artiste (par exemple, pop, rap).
- **name** : Nom de l'artiste.
- **popularity** : Score de popularité de l'artiste (0 à 100).

**tracks.csv** :

- **id** : Identifiant unique du morceau.
- **name** : Nom du morceau.
- **popularity** : Popularité du morceau (0 à 100).
- **release\_date** : Date de sortie du morceau.
- Caractéristiques audio :
  - **danceability** : Mesure la qualité dansante d'un morceau (0 à 1).
  - **energy** : Mesure de l'intensité et de l'activité (0 à 1).
  - **acousticness** : Probabilité qu'un morceau soit acoustique (0 à 1).
  - **instrumentalness** : Probabilité qu'un morceau soit instrumental (0 à 1).
  - **tempo** : Vitesse globale d'un morceau en BPM.

### 2.2.3 Étendue Temporelle

Les morceaux présents dans le jeu de données couvrent une période allant de **1922** (date de sortie la plus ancienne) à **2021** (année la plus récente incluse dans le dataset). Cela permet une analyse des tendances musicales sur plus d'un siècle.

## 2.3 Nettoyage et Préparation des Données

### 1. Nettoyage des Genres :

- Les genres, initialement stockés sous forme de chaînes de caractères, ont été convertis en listes Python pour permettre une manipulation facile.
- Les artistes sans genre associé ont été supprimés.

### 2. Gestion des Valeurs Manquantes :

- Suppression des lignes contenant des valeurs manquantes après nettoyage des genres.

### 3. Conversion des Dates :

- La colonne `release_date` a été convertie en format datetime pour permettre des analyses temporelles. L'année (`release_year`) et le mois (`release_month`) ont été extraits.

### 3 Analyse exploratoire des données musicales

L'objectif de cette section est de mieux comprendre les tendances musicales, les caractéristiques des morceaux populaires et l'évolution des collaborations entre artistes. Cette analyse est appuyée par des visualisations pour illustrer les résultats clés.

#### 3.1 Évolution de la popularité et des collaborations

L'analyse de la popularité moyenne des morceaux par année (2000-2020) révèle une croissance significative depuis 2016. Cette tendance est accompagnée d'une forte augmentation des collaborations entre artistes, qui connaissent une montée spectaculaire à partir de 2017, culminant en 2020 avec plus de 5000 collaborations. Ces résultats soulignent le rôle croissant des collaborations dans la popularité musicale récente.

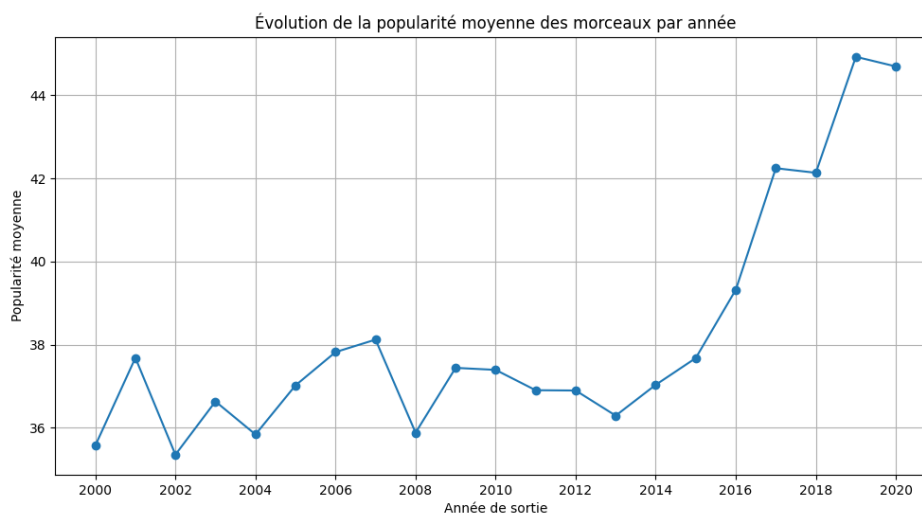


FIGURE 1 – Évolution de la popularité moyenne des morceaux par année

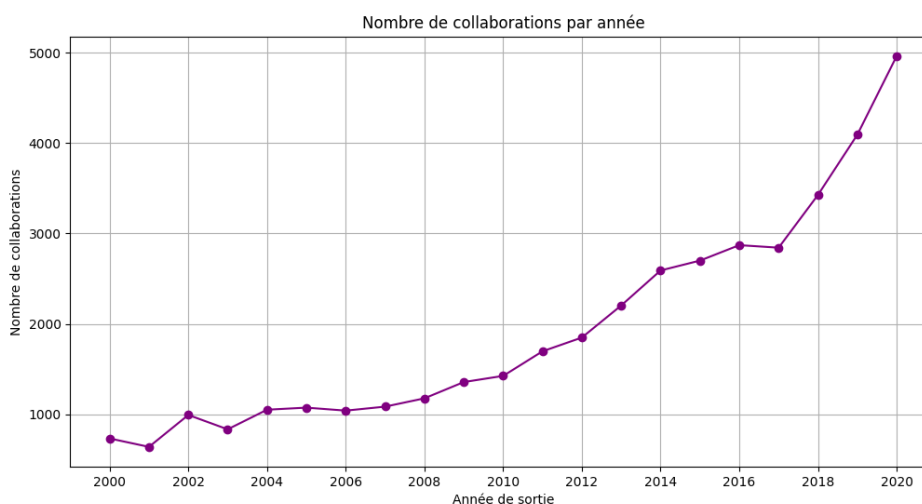


FIGURE 2 – Nombre de collaborations par année

### 3.2 Analyse des caractéristiques musicales par année

Les caractéristiques musicales des morceaux populaires montrent des évolutions marquées entre 2000 et 2020 :

- **Danceability** : Une augmentation constante depuis 2015, reflétant une préférence pour des morceaux accessibles pour la danse.
- **Energy** : Un pic observé entre 2012 et 2014 avant une légère diminution.
- **Valence** : Une baisse progressive, suggérant une tendance vers des morceaux perçus comme moins joyeux.

Ces résultats traduisent des évolutions dans les préférences des auditeurs et l'adaptation des artistes à ces tendances.

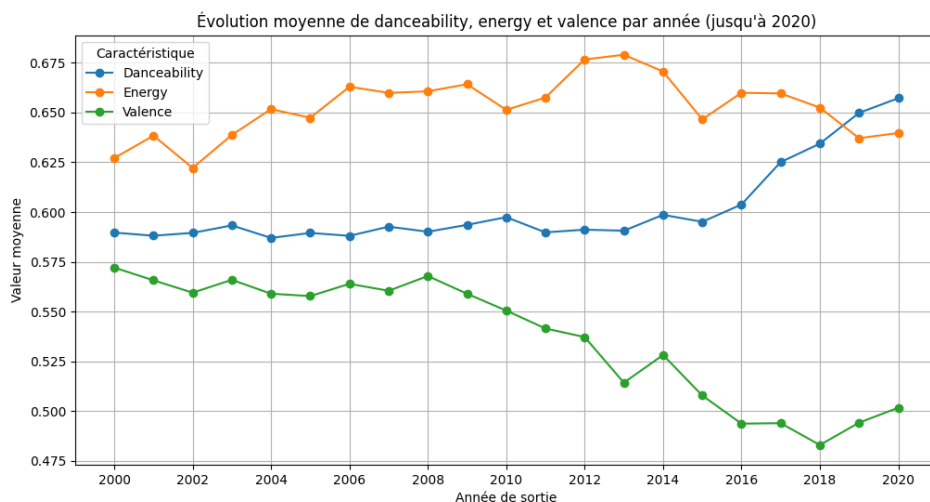


FIGURE 3 – Évolution moyenne de danceability, energy et valence

### 3.3 Corrélations entre popularité et caractéristiques musicales

L'analyse des corrélations de Pearson met en évidence des relations intéressantes :

- Une forte corrélation positive entre la **danceability** et la **popularité** (+0.899), suggérant que les morceaux dansants sont plus susceptibles de devenir populaires.

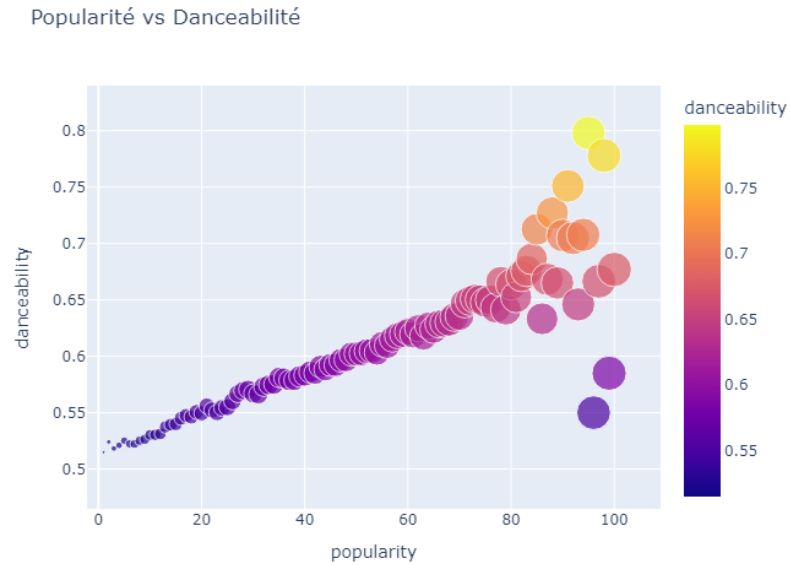


FIGURE 4 – Corrélation entre la danceability et la popularité

- Une corrélation négative entre l'**instrumentalness** et la popularité (-0.869), indiquant que les morceaux avec moins de paroles (plus instrumentaux) sont moins populaires.

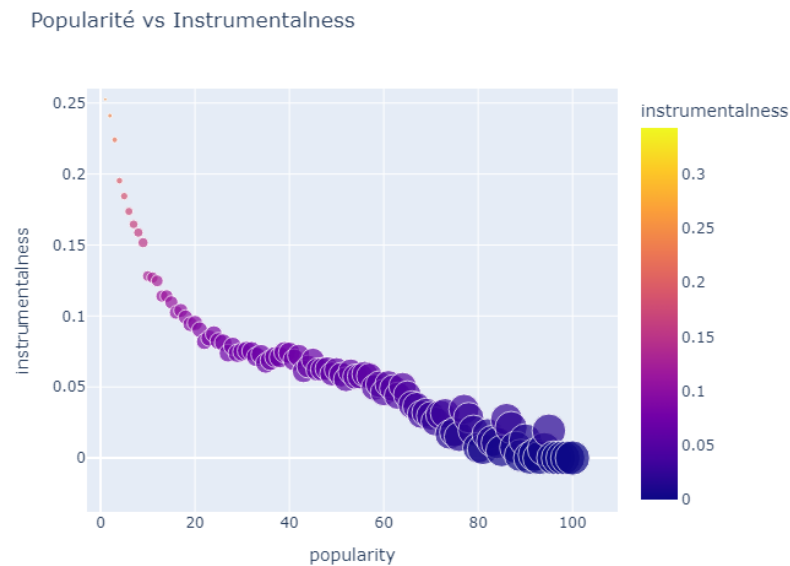


FIGURE 5 – Corrélation entre l'instrumentalness et la popularité



### 3.4 Genres musicaux et popularité

Le **pop** domine les genres des morceaux populaires, suivi par des sous-genres tels que le **dance pop**, le **rock**, et le **latin pop**. Ces genres répondent probablement à des attentes universelles des auditeurs. Par ailleurs, la distribution du nombre de genres par morceau montre que la majorité des morceaux appartiennent à moins de trois genres, soulignant une spécialisation des productions musicales.

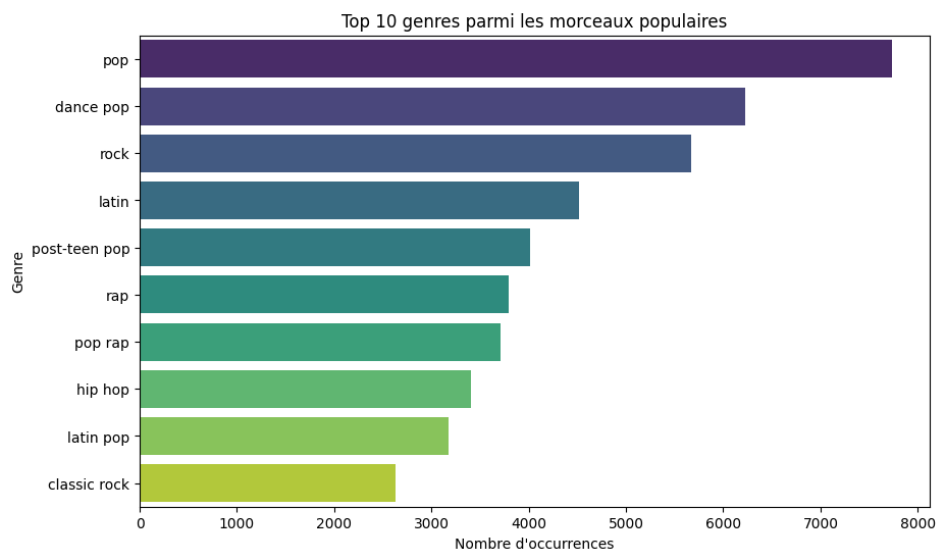


FIGURE 6 – Top 10 des genres parmi les morceaux populaires

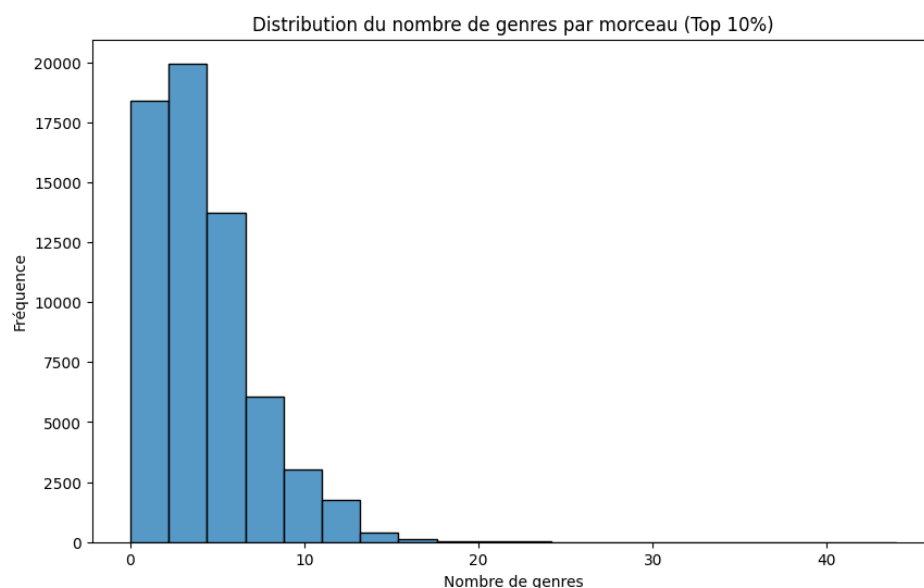


FIGURE 7 – Distribution du nombre de genres par morceau

### 3.5 Durée et caractéristiques des morceaux populaires

Les morceaux populaires sont généralement courts, avec une durée moyenne de **3 à 4 minutes**, correspondant au format standard de l'industrie musicale. Une analyse des caractéristiques des morceaux du top 10% montre une **danceability** et une **valence** modérées, ainsi qu'une énergie suffisante pour capter l'attention des auditeurs.

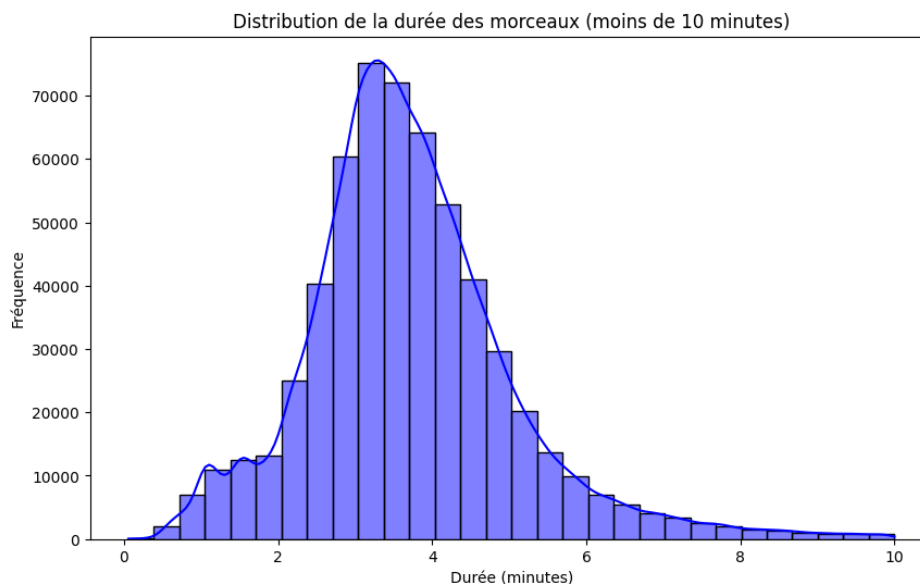


FIGURE 8 – Distribution de la durée des morceaux

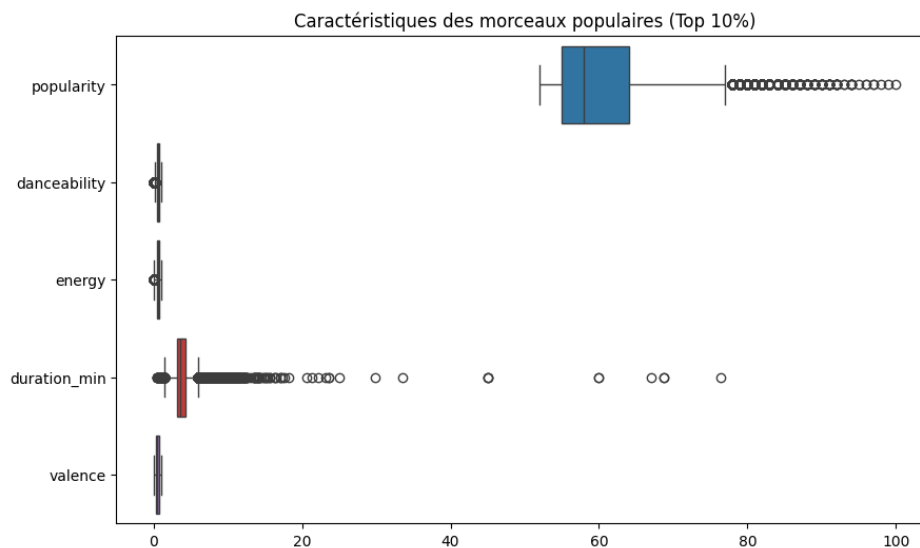


FIGURE 9 – Caractéristiques des morceaux populaires (Top 10%)

L'analyse exploratoire a mis en lumière des tendances clés dans l'industrie musicale contemporaine :

- L'importance croissante des collaborations dans la musique populaire, particulièrement depuis 2016.

- Une évolution des caractéristiques musicales, avec une hausse de la **danceability** et une légère diminution de la **valence**, reflétant les préférences changeantes des auditeurs.
- La prédominance de certains genres comme le **pop**, ainsi qu'une spécialisation des morceaux dans des formats spécifiques, notamment en termes de durée et de caractéristiques audio.

Ces observations fournissent des bases solides pour une exploration plus approfondie dans l'industrie musicale.

## 4 Analyse des clusters : Une segmentation claire des morceaux musicaux

L'analyse des clusters permet de regrouper les morceaux en groupes homogènes basés sur leurs caractéristiques musicales. Cette segmentation a été réalisée à l'aide de l'algorithme **K-Means**, après une étape de **normalisation des données** via la méthode **StandardScaler** pour garantir une échelle uniforme des caractéristiques (comme la danceability, l'énergie ou le tempo). Cette méthodologie a permis de structurer l'ensemble des morceaux en **quatre clusters distincts**, reflétant des profils d'écoute variés. L'interprétation détaillée des clusters est présentée ci-dessous.

### 4.1 Caractéristiques moyennes des clusters

Les morceaux ont été répartis en quatre groupes à partir des caractéristiques musicales suivantes : **danceability**, **energy**, **instrumentalness**, **acousticness**, **loudness**, **speechiness**, et **tempo**. La méthode du **coude** (Elbow Method) a été utilisée pour identifier le nombre optimal de clusters, en observant la diminution de l'inertie au fur et à mesure de l'augmentation du nombre de clusters. Une fois les clusters définis, les caractéristiques moyennes de chaque cluster ont été analysées pour mieux comprendre les différences et similitudes entre les morceaux.

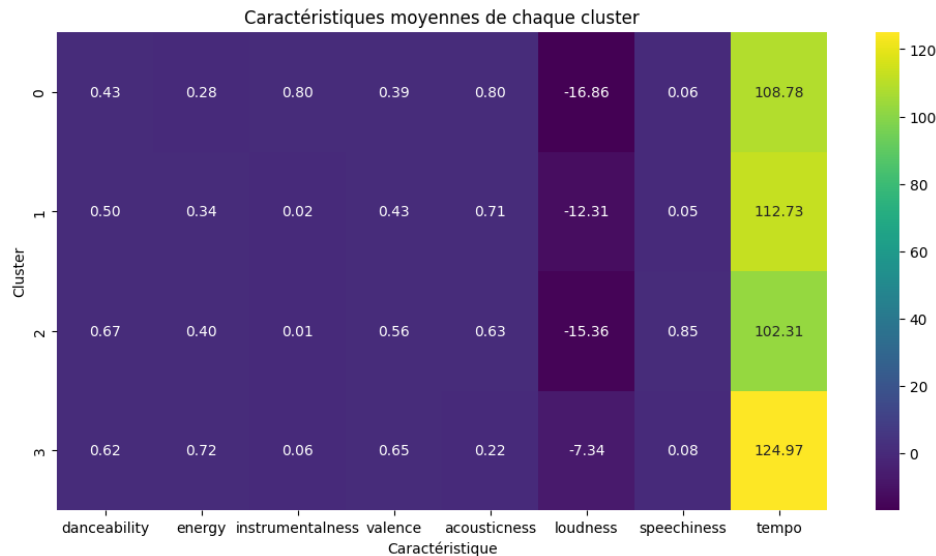


FIGURE 10 – heatmap mettant en évidence les différences principales entre les clusters

#### 1. Cluster 0 : Calme, instrumental et acoustique

- **Danceability** : Faible (0.42)
- **Energy** : Très faible (0.28)
- **Instrumentalness** : Élevée (0.80)

- **Acousticness** : Très élevée (0.80)
  - **Loudness** : Faible (-16.86 dB)
  - **Speechiness** : Très faible (0.057)
  - **Type de morceaux** : Représente des morceaux apaisants, souvent instrumentaux et acoustiques, typiques des genres classiques ou orchestraux.
2. **Cluster 1 : Mélodieux, vocal et acoustique**
- **Danceability** : Moyenne (0.49)
  - **Energy** : Faible (0.34)
  - **Instrumentalness** : Très faible (0.02)
  - **Acousticness** : Moyenne à élevée (0.71)
  - **Loudness** : Moyenne (-12.31 dB)
  - **Speechiness** : Très faible (0.054)
  - **Type de morceaux** : Ce cluster inclut des morceaux vocaux et mélodieux avec un accent acoustique, comme la pop acoustique et le soft rock.
3. **Cluster 2 : Narratif et dansant**
- **Danceability** : Élevée (0.67)
  - **Energy** : Moyenne (0.40)
  - **Instrumentalness** : Très faible (0.005)
  - **Acousticness** : Moyenne (0.63)
  - **Loudness** : Moyenne (-15.36 dB)
  - **Speechiness** : Très élevée (0.85)
  - **Type de morceaux** : Les morceaux dans ce cluster ont une forte composante narrative ou vocale, comme les genres spoken word ou le rap.
4. **Cluster 3 : Énergique et dansant**
- **Danceability** : Élevée (0.62)
  - **Energy** : Très élevée (0.72)
  - **Instrumentalness** : Faible (0.056)
  - **Acousticness** : Très faible (0.22)
  - **Loudness** : Élevée (-7.34 dB)
  - **Speechiness** : Moyenne (0.08)
  - **Type de morceaux** : Ce cluster est composé de morceaux dynamiques et puissants, souvent associés à des genres comme le rock, la pop ou l'électro.

## 4.2 Popularité des clusters

Une analyse complémentaire a été réalisée pour évaluer la **popularité moyenne** des morceaux dans chaque cluster. Les résultats montrent que le **Cluster 3**, qui regroupe des morceaux énergiques et dansants, est le plus populaire avec une popularité moyenne de **33.03**, tandis que le **Cluster 0** (calme et instrumental) est le moins populaire avec une moyenne de **15.38**.

| Cluster | Popularité moyenne |
|---------|--------------------|
| 3       | 33.03              |
| 1       | 23.01              |
| 2       | 20.61              |
| 0       | 15.38              |

TABLE 1 – Popularité moyenne par cluster

### 4.3 Répartition des genres par cluster

Les genres dominants associés à chaque cluster offrent une perspective enrichissante pour l'analyse :

1. **Cluster 0 :**
  - **Genres dominants :** Classical, Orchestral Performance, Jazz.
  - Correspond parfaitement aux caractéristiques calmes, acoustiques et instrumentales du cluster.
2. **Cluster 1 :**
  - **Genres dominants :** Adult Standards, Vocal Jazz, Lounge.
  - Représente des morceaux vocaux doux et mélodieux.
3. **Cluster 2 :**
  - **Genres dominants :** Hörspiel, Spoken Word, Barnmusik.
  - Inclut des morceaux narratifs ou éducatifs avec une forte composante vocale.
4. **Cluster 3 :**
  - **Genres dominants :** Rock, Classic Rock, Album Rock.
  - Associe des morceaux énergiques et dynamiques, cohérents avec ses caractéristiques.

### 4.4 Visualisation 2D des clusters et insights

Une réduction dimensionnelle via l'analyse en composantes principales (PCA) a permis de représenter les clusters dans un espace 2D. Cette visualisation met en lumière leur positionnement et les chevauchements éventuels.

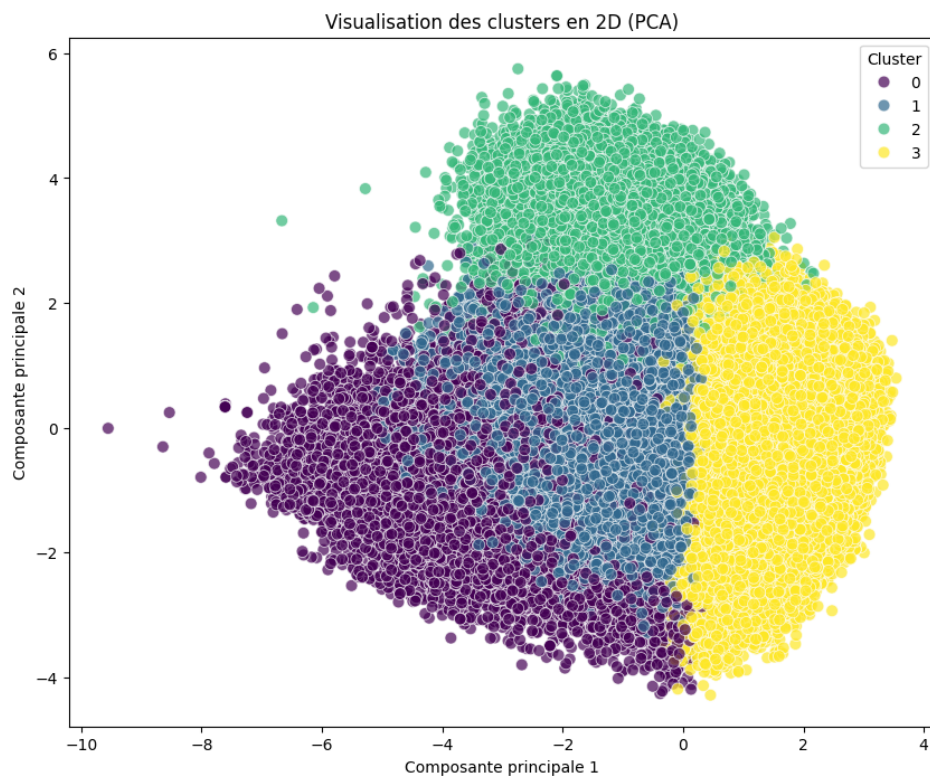


FIGURE 11 – Visualisation 2D des clusters

#### Chevauchements et transitions musicales :

- Les chevauchements entre clusters, particulièrement entre les clusters 1 et 2, reflètent des transitions musicales naturelles, notamment entre morceaux calmes, narratifs et mélodieux.
- Le **Cluster 3**, distinct et dynamique, montre peu de chevauchements, soulignant son caractère unique.

## 4.5 Synthèse des résultats

| Cluster | Type de morceaux                          | Popularité moyenne | Genres dominants                        |
|---------|---|--------------------|---|
| 0       | Calme, instrumental, acoustique           | 15.38              | Classical, Orchestral Performance, Jazz |
| 1       | Mélodieux, vocal, acoustique              | 23.01              | Adult Standards, Vocal Jazz, Lounge     |
| 2       | Narratif, parlé, contenu lyrique dominant | 20.61              | Hörspiel, Spoken Word, Barnmusik        |
| 3       | Énergique, dansant, dynamique             | 33.03              | Rock, Classic Rock, Album Rock          |

TABLE 2 – Synthèse des résultats par cluster

## 4.6 Insights clés et recommandations

### 1. Personnalisation des recommandations :

- Les clusters peuvent être utilisés pour recommander des morceaux adaptés aux préférences des auditeurs. Les morceaux proches des frontières des clusters peuvent être suggérés pour explorer des styles similaires.

### 2. Applications commerciales :

- **Cluster 0** : Idéal pour des playlists de relaxation ou d’ambiance.
- **Cluster 1** : Convient aux amateurs de musique vocale et acoustique.
- **Cluster 2** : Recommandé pour les contenus narratifs ou éducatifs.
- **Cluster 3** : Parfait pour des playlists dynamiques et festives.

Cette analyse met en évidence la robustesse de la segmentation réalisée, avec des clusters bien définis et des transitions musicales reflétant des préférences diversifiées des auditeurs.

## 5 Analyse des communautés et collaborations artistiques via les graphes

Dans cette section, nous examinons la **dynamique des collaborations entre artistes** à travers une analyse des graphes. En utilisant les 1000 artistes les plus populaires identifiés dans les données, en se basant sur les morceaux sortis après 2017 (période marquant l'explosion des collaborations, selon l'analyse précédente). nous avons construit un graphe représentant leurs collaborations. Les nœuds représentent les artistes, tandis que les liens entre eux symbolisent leurs collaborations. Les **communautés ont été identifiées grâce à l'algorithme Louvain**, et les métriques de centralité telles que la **centralité du degré** et la **betweenness centrality** ont été utilisées pour mettre en évidence les rôles clés des artistes dans le réseau.

### 5.1 Structure globale du graphe

Pour analyser les collaborations artistiques, nous avons utilisé le logiciel **Gephi** pour visualiser et explorer les connexions entre artistes. Nous avons construit un graphe non orienté, la spatialisation des nœuds a été réalisée avec les algorithmes **Yifan Hu Proportional**

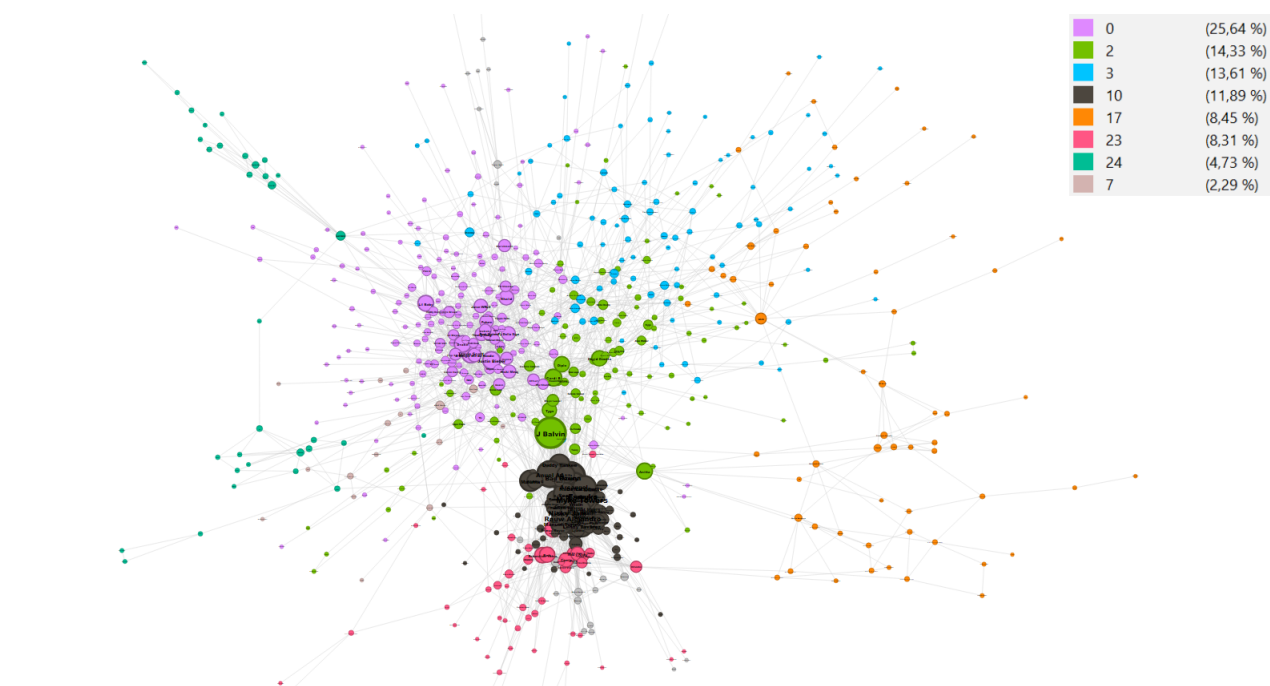


FIGURE 12 – Visualisation des communautés d'artistes dans le graphe

Chaque couleur représente une communauté, la taille des nœuds reflète la **centralité du degré**. On observe une structure dense où les artistes se regroupent majoritairement en fonction de leur genre ou de leur origine géographique.

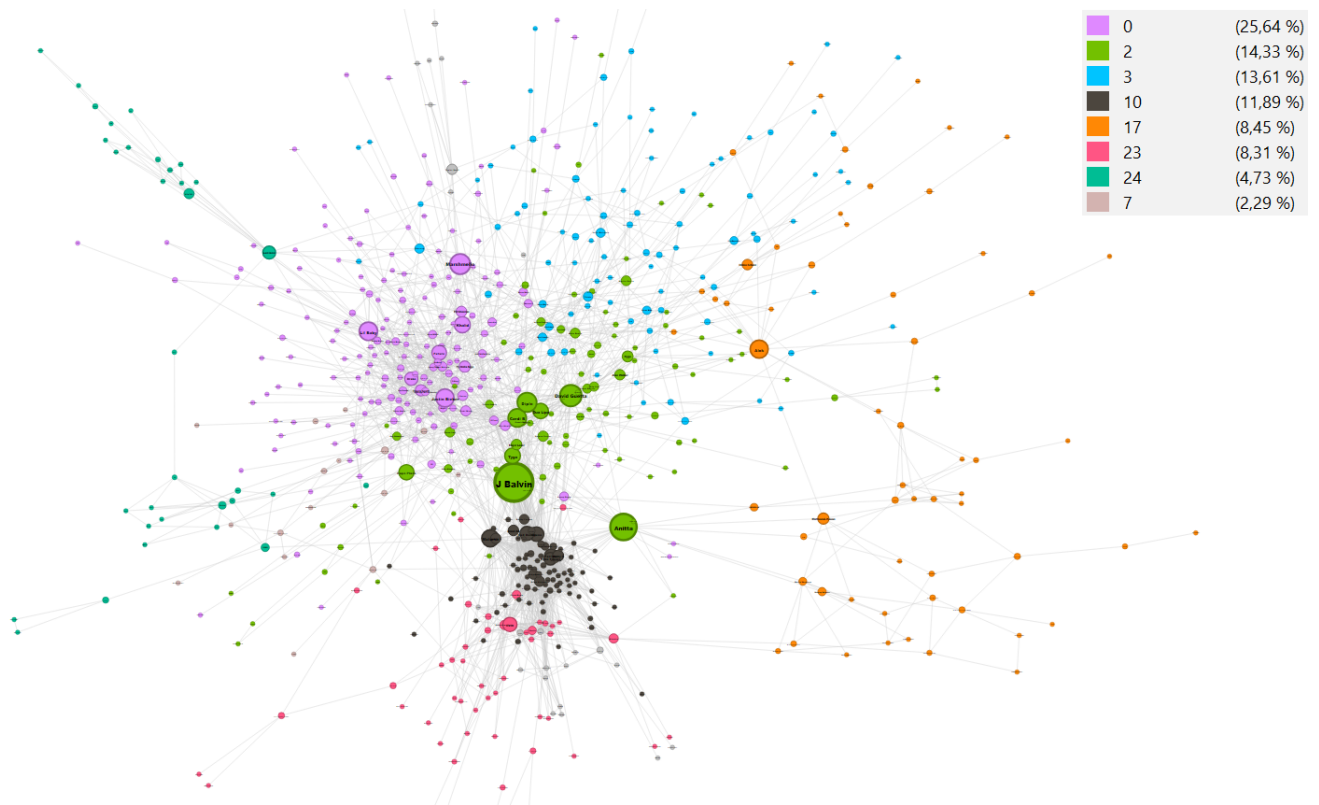


FIGURE 13 – Visualisation des communautés d'artistes (taille de noeud sur betweenness)

Dans ce deuxième graphe la taille des nœuds reflète la **betweenness**. Les nœuds stratégiques avec des **valeurs élevées de betweenness** apparaissent souvent comme des ponts entre différentes communautés.

— **Clusters principaux :**

- Les principales communautés identifiées sont :
  - **Communauté 10** (Reggaeton/Latin)
  - **Communauté 0** (Rap/Hip-Hop international)
  - **Communauté 3** (Pop/Dance)
  - **Communauté 2**, située au centre du graphe, jouant un rôle pivot en connectant les trois grands clusters.

— **Position des communautés :**

- Les communautés proches culturellement ou musicalement, comme la **10 (Latin)** et la **23 (Mexicain/Latin)**, sont logiquement rapprochées. En revanche, les communautés comme la **17 (Dance brésilien)** ou la **24 (Rap German/Français)**, sont plus isolées en raison de leur niche régionale.



### 5.2.1 Communauté 10 (Reggaeton/Latin)

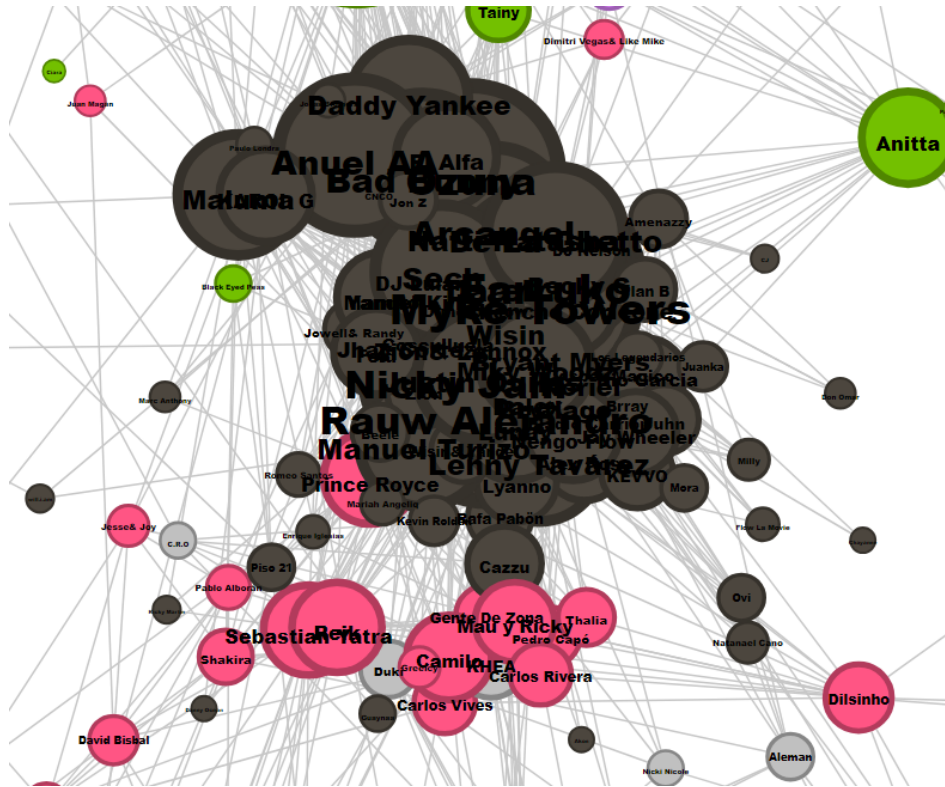


FIGURE 14 – Communauté 10 (Reggaeton/Latin) (taille de noeud sur centralité du degré)

- **Caractéristiques :**
  - Densité élevée avec des artistes très connectés, comme **Bad Bunny**, **Ozuna**, et **Daddy Yankee**.
  - **Centralité du degré élevée** pour ces artistes, qui dominent les collaborations au sein de cette communauté.
- **Position :**
  - Proche de la **communauté 23**, en raison de leur lien avec la scène musicale latino-américaine.
- **Observation clé :**
  - Cette communauté illustre un fort niveau de collaborations internes entre artistes populaires, reflétant une dynamique où les figures majeures de la scène reggaeton et trap, dans la musique latino, travaillent fréquemment ensemble, renforçant leur dominance et leur influence au sein de ce genre.

### 5.2.2 Communauté 0 (Rap/Hip-Hop international)

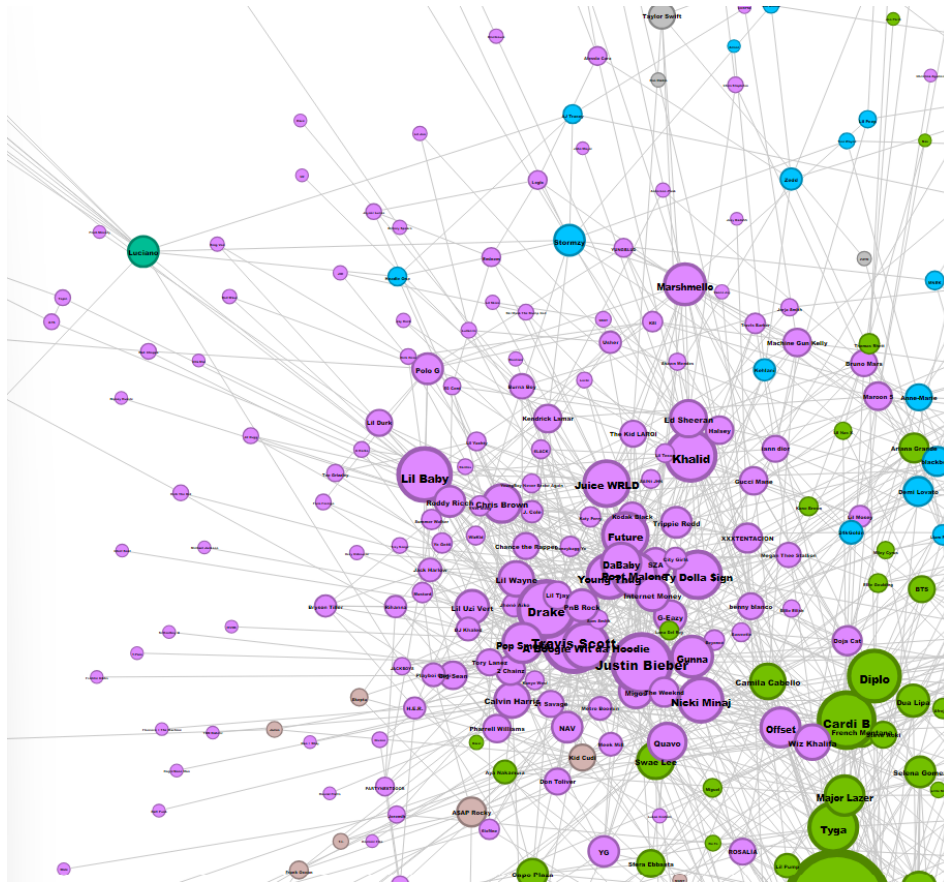


FIGURE 15 – Communauté 0 (Rap/Hip-Hop international) (taille de noeud sur centralité du degré)

- **Caractéristiques :**
  - Une structure plus dispersée mais marquée par une forte influence grâce à des artistes comme **Justin Bieber, Drake, Travis Scott, Nicki Minaj, et Lil Baby**.
  - Ces artistes ont une **centralité du degré élevée**, ce qui montre leur rôle clé dans la densité des collaborations.
- **Position :**
  - Représentant une scène musicale globale, cette communauté est un pilier de l'industrie musicale occidentale.
- **Observation clé :**
  - La diversité des collaborations dans cette communauté témoigne de l'hybridation des genres entre rap, pop et autres musiques populaires.

### 5.2.3 Communauté 3 (Pop/Dance)

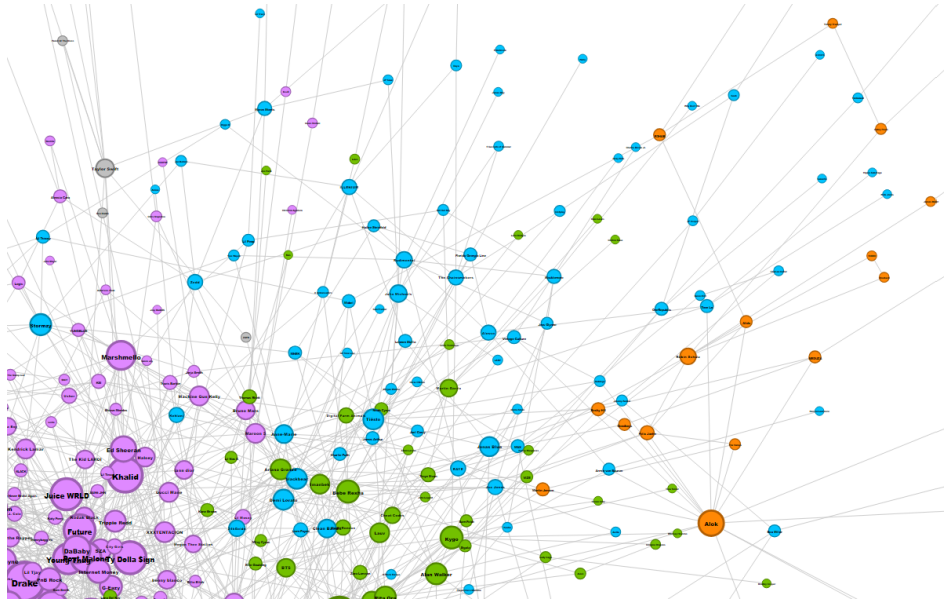


FIGURE 16 – Communauté 3 (Pop/Dance) (taille de noeud sur centralité du degré)

- **Caractéristiques :**
  - Structure **éparpillée**, avec une **centralité du degré plus faible** que les communautés 10 et 0.
  - Artistes clés comme **Tiësto**, **Jonas Blue**, **Chainsmokers**, et **blackbear**, qui dominent les collaborations dans la musique électronique, dance et pop.
- **Position :**
  - Relativement **périphérique**, mais bien connectée à la **communauté 2**, qui agit comme un pont central vers d'autres genres musicaux.
- **Observation clé :**
  - Cette communauté illustre l'intégration de la musique **pop** et **dance** avec d'autres styles, notamment via des collaborations fréquentes avec des artistes de la **musique latine** et du **rap**. Cela reflète une influence grandissante de la pop et de la dance sur les tendances musicales internationales.

### 5.2.4 Communauté 2 (Pont Latino/Occidental)

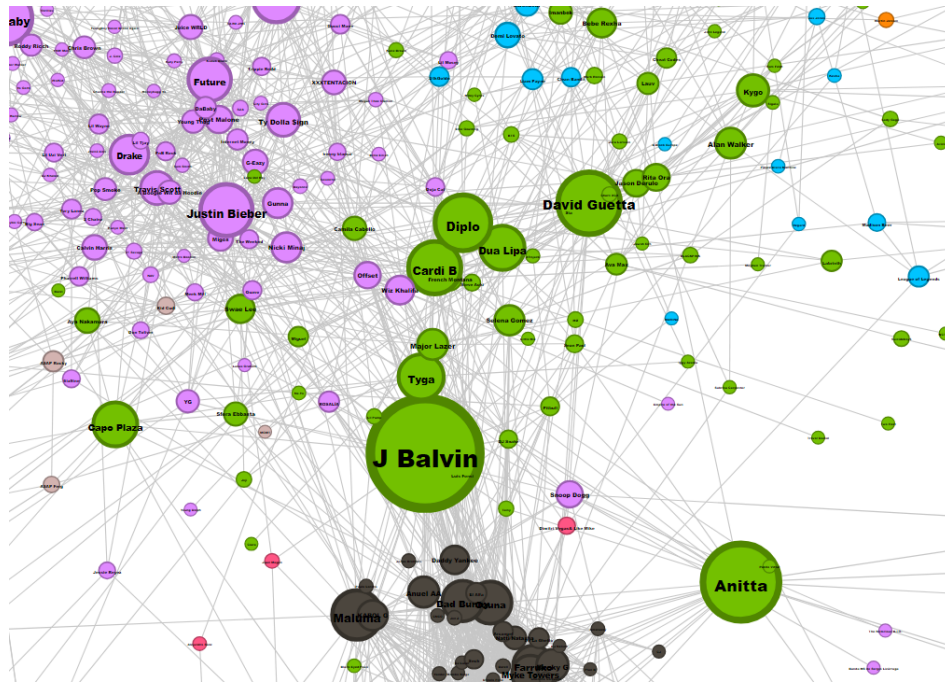


FIGURE 17 – Communauté 2 (Pont Latino/Occidental) (taille de noeud sur betweenness)

— **Caractéristiques :**

- Une des communautés les plus stratégiques, située au centre du graphe, avec une **betweenness centrality élevée** pour des artistes comme **J Balvin** et **Anitta**.
- **J Balvin**, artiste latino international, joue un rôle crucial en reliant la communauté latino (10) à la scène occidentale (0 et 3).
- **Anitta**, autre figure clé, sert de pont non seulement entre les communautés 10 et 2, mais aussi vers la **communauté 17**, représentant la musique brésilienne.
- **David Guetta**, pivot de la scène électro, consolide les connexions entre des genres variés comme la dance et la pop.

— **Position :**

- Rôle pivot dans le graphe global, en connectant des genres latins aux genres internationaux.

— **Observation clé :**

- Cette communauté met en lumière l'essor de la musique latine sur la scène mondiale . Les collaborations fréquentes entre artistes populaires mettent en lumière son rôle central dans la fusion des genres et son impact sur les marchés internationaux.

### 5.2.5 Communauté 23 (Mexicain/Latin)

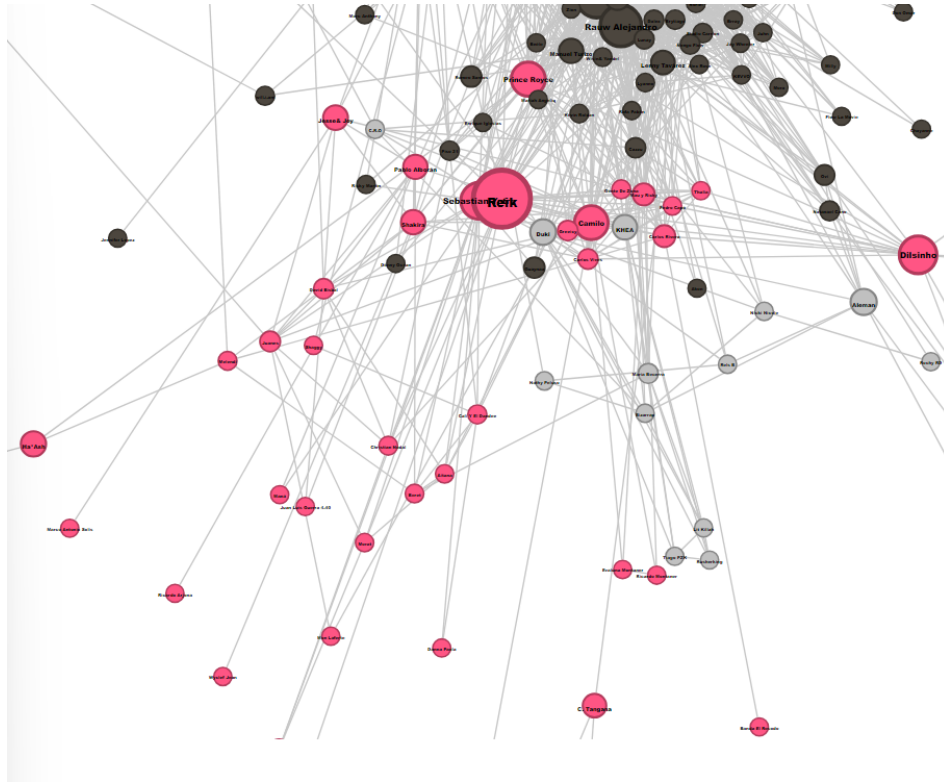


FIGURE 18 – Communauté 23 (Mexicain/Latin) (taille de noeud sur betweenness)

- **Caractéristiques :**
  - Située sous la communauté 10, en raison de leurs similarités culturelles et musicales.
  - **Reik**, groupe de pop mexicain, agit comme un pont entre les deux communautés grâce à une **betweenness centrality** élevée.
- **Position :**
  - Représentative de la scène mexicaine, connectée à la scène reggaeton via des collaborations stratégiques.
- **Observation clé :**
  - Cette communauté illustre le rôle de la musique mexicaine comme un élément complémentaire à la scène latine globale.

## 5.3 Communautés spécifiques et isolées

### 5.3.1 Communauté 17 (Dance/Funk brésilien)

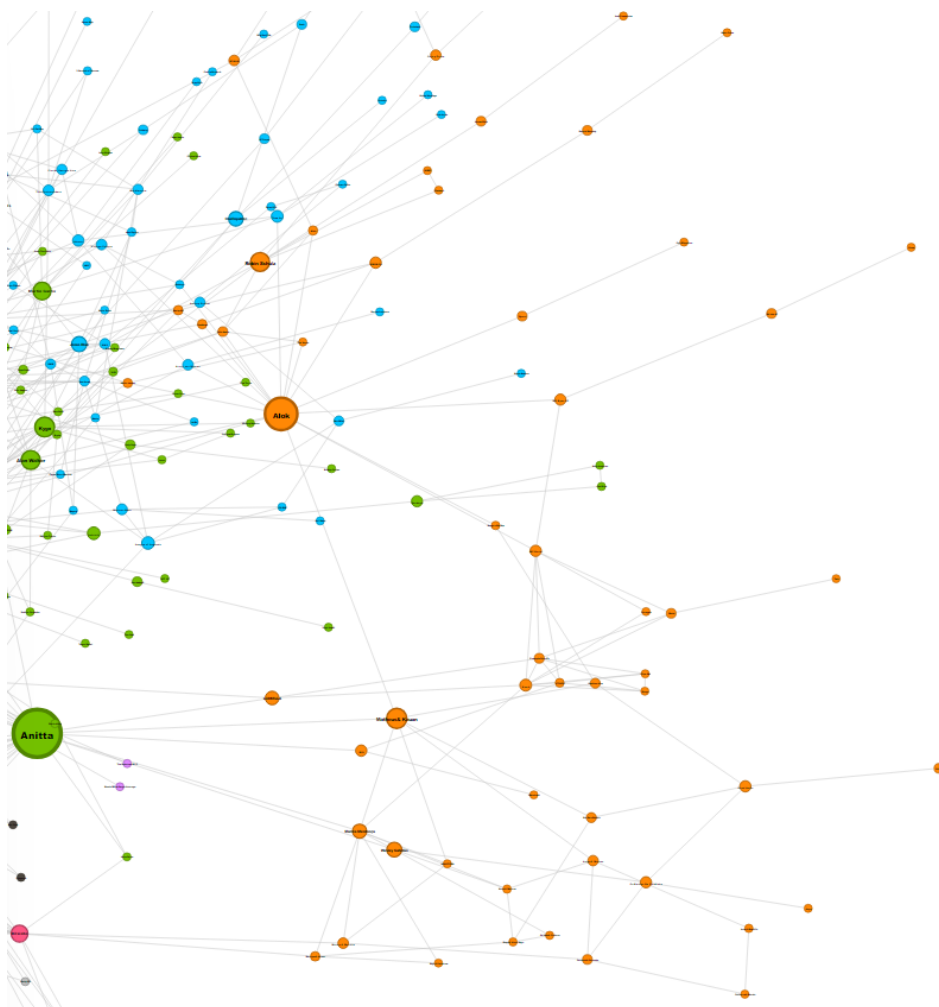


FIGURE 19 – Communauté 17 (Dance/Funk brésilien) (taille de noeud sur betweenness)

- **Caractéristiques :**
  - Relativement écartée et éloignée du centre du graphe, avec des valeurs de **centralité du degré** et de **betweenness** faibles.
  - Acteur principal : **Alok**, qui représente la musique électronique et le funk brésilien.
- **Position :**
  - Peu connectée aux communautés principales, mais reliée à la communauté 2 via **Anitta**.
- **Observation clé :**
  - Cette communauté met en avant la spécificité régionale de la scène musicale brésilienne.



### 5.3.2 Communauté 24 (Rap German/Français)

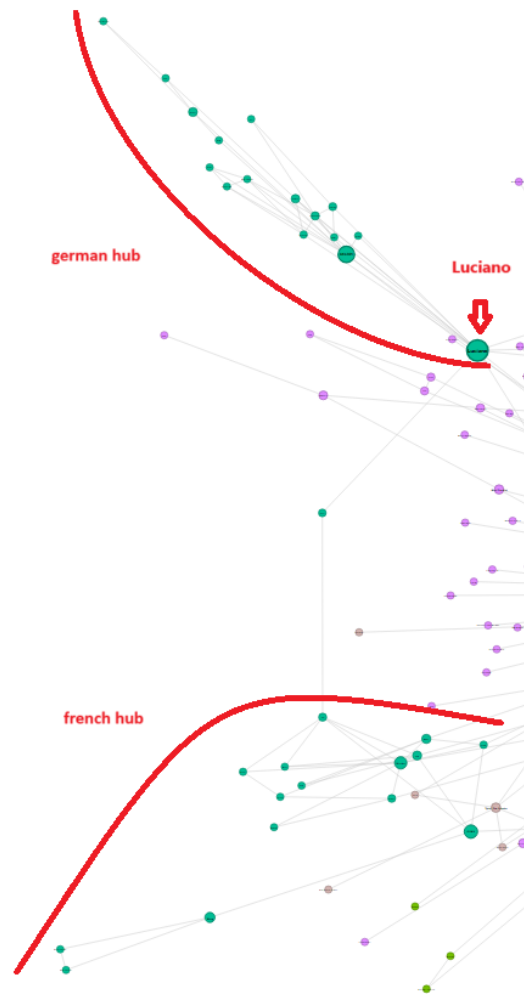


FIGURE 20 – Communauté 24 (Rap German/Français) (taille de nœud sur betweenness)

- **Caractéristiques :**
  - Comprend deux hubs majeurs : le rap allemand (**Luciano**) et le rap français .
  - Artiste clé : **Luciano**, qui agit comme un pont vers d'autres communautés grâce à une **betweenness centrality élevée**.
- **Position :**
  - À l'extrémité du graphe, relativement isolée, mais connectée à d'autres communautés via des artistes internationaux comme **Gazo**.
- **Observation clé :**
  - Cette communauté illustre des niches régionales, avec un mélange de scènes locales et de collaborations internationales limitées.

## 5.4 Insights généraux

1. **Centralité et influence des communautés principales :**
  - Les communautés 10, 0, et 3 dominent par leur densité et leur influence, avec des artistes clés comme **Bad Bunny** et **Drake**.
  - La **communauté 2** agit comme un pivot, reliant les grandes scènes latines et occidentales, avec des artistes clés comme **J Balvin**.
2. **Artistes ponts et collaborations stratégiques :**
  - **J Balvin** et **Anitta** jouent des rôles cruciaux en tant que ponts entre plusieurs communautés, en connectant les genres brésiliens (17) et latins (10) avec la scène internationale.
  - **Reik** (communauté 23) illustre le rôle des groupes mexicains dans le lien entre la musique pop mexicaine et le reggaeton.
3. **Spécificités régionales :**
  - Des communautés comme la 24 (Rap German/Français) ou la 17 (Dance/Funk brésilien) montrent des particularités régionales, mais restent connectées grâce à des artistes stratégiques.
4. **Importance des collaborations inter-communautaires :**
  - Ces collaborations renforcent l'interconnexion entre genres et régions, permettant une hybridation musicale et une globalisation des tendances musicales.

## 6 Conclusion

L'analyse des données musicales de Spotify a révélé une structure complexe et interconnectée de l'industrie musicale moderne. Les tendances observées indiquent une importance croissante des collaborations entre artistes, une évolution des préférences musicales vers des morceaux plus dansants et énergiques, et une segmentation claire des morceaux basée sur leurs caractéristiques audio. L'analyse des clusters a permis de comprendre comment les morceaux peuvent être regroupés pour améliorer les recommandations musicales, tandis que l'étude des graphes de collaborations artistiques a mis en évidence le rôle central de certains artistes dans la diffusion et l'hybridation des genres musicaux. Ces insights sont précieux pour les professionnels de l'industrie musicale, les plateformes de streaming, et les chercheurs intéressés par la dynamique musicale contemporaine.