

Rapport

Information - Visualisation

Léo MARACHE - Timothée POULAIN

Utilisateurs cibles (What ? Why ?)

Notre système de visualisation est destiné aux **professionnels de la musique** souhaitant produire des morceaux à partir de titres existants. Afin de les aider dans leur processus de création :

- composition,
- mixage,
- et mastering;

le producteur doit connaître un certain nombre de paramètres du(des) titre(s) sur le(s)quel(s) il se base :

- Le **tempo** (BPM) : indispensable pour l'étape de composition;
- La **clé** du morceau (la note racine autour de laquelle toute la composition s'articule) : aussi importante pour la phase de composition;
- Le **gain** (dB) : qui est une valeur à fixer au moment du mastering du morceau;

L'idée est donc de permettre à l'utilisateur de visualiser toutes ces données, qu'il pourra filtrer en fonction du **genre** et éventuellement de la période temporelle.

Traitement des données

Nous sommes tout d'abord partis de l'extrait de la base de données Wasabi au format *.rds*. Le traitement des données a ensuite suivi le cheminement suivant :

- Formatage des différentes données brutes (R):

La principale difficulté ici a été le traitement de la liste des genres ainsi que de la clé à partir de la colonne *chords*, qui a dû être traitée à l'aide d'expressions régulières.

- Sélection des colonnes nécessaires (R):

Les colonnes suivantes ont été retenues : title, genres[0 à 7], bpm, length, gain, rank, publicationDate, key

- Export en JSON (R)
- Conversion en CSV (Outil en ligne)
- Suppression des doublons (Excel)

Certains morceaux apparaissent plusieurs fois dans la base de données, mais avec des paramètres différents. Nous avons décidé de les garder, car les versions alternatives sont souvent des versions remasterisées pour coller aux standards d'une nouvelle époque, et c'est ce que l'utilisateur recherche. En revanche, les doublons strictement identiques ont été supprimés.

- Extraction de l'année (Excel):

A partir du champ *publicationDate*, extraction de l'année dans une nouvelle colonne *year*.

- Correction des données erronées (Excel)

Il s'agissait uniquement d'enlever les virgules dans le titre de certains morceaux, ce qui créait des erreurs puisque les données étaient stockées au format CSV, avec une virgule pour séparateur.

How ?

Visualisation - Léo - Scatter plot

Quelles codifications visuelles ?

Le graphe lui-même représente le **tempo** en fonction du **gain**. Ce choix des axes est en effet le plus réaliste pour un professionnel de la musique, car ce sont les deux paramètres les plus importants pour son travail.¹

Chaque morceau est représenté par un cercle sur le graphe.

Le **genre** d'un morceau est représenté par la couleur du cercle, et permet donc de distinguer visuellement la répartition des genres selon le gain et le tempo.

La **longueur (durée)** d'un titre est quant à elle représentée par la taille du cercle correspondant.

Une légende à la droite du graphe permet de consulter :

- la liste des genres (avec code couleur) des morceaux représentés
- et une indication visuelle sur la correspondance taille-durée du morceau.

Quelles tâches utilisateur et techniques d'interaction ?

Overview : Initialement, le scatter plot permet de visualiser les morceaux dans leur ensemble.

Zoom : littéralement une interaction de **zoom et navigation** sur le graphe

Filter : Non implémenté => idée : **cliquer** sur un cercle pour n'afficher que les morceaux du même genre

Details on demand : une **tooltip** s'affiche lors d'un **hover** sur un cercle, celle-ci contient toutes les informations qui nous intéressent

¹ Notons qu'il aurait aussi été pertinent de représenter sur un axe les clés des morceaux, cependant de par le faible nombre de titres de la base de données fournies possédant une clé; elle n'aurait pas représenté un grand intérêt.

Visualisation - Timothée - Parallel coordinates + Box-And-Whisker Plot

Quelles tâches utilisateur et techniques d'interaction ?

Overview : Le diagramme Parallel Coordinates, affichant une ligne par morceau. La couleur des lignes correspond au genre “principal”, mais n’a pas d’utilité réelle sur ce diagramme (la variable visuelle “couleur” n’est pas utilisée).

Zoom + Filter : Lors de la sélection d’un genre dans le menu déroulant, ou lors de l’application de filtres sur les colonnes du diagramme Parallel Coordinates.

Details on demand : A tout moment, la moitié basse de la visualisation donne de nombreuses informations sur la sélection en cours, et ce quels que soient les filtres appliqués :

- Tableau des 5 morceaux les plus populaires;
- Médiane/Quartiles du tempo;
- Médiane/Quartiles de la longueur;
- Médiane/Quartiles du gain.

Pistes d’amélioration :

- Ajout de champs de recherche

La recherche de genres dans une liste déroulante n’est pas très propre, d’autant plus que la liste n’est pas triée. L’ajout d’un champ de recherche permettrait à l’utilisateur d’être beaucoup plus efficace dans ses recherches. Il pourrait ainsi entrer “rock” et tomber sur les styles qui contiennent cette chaîne de caractère : “Pop rock”, “Hard rock”, etc.

- Exploration de sous-genres

Comme chaque morceau se voit attribuer plusieurs genres, il serait cohérent de proposer, à chaque genre sélectionné, une liste des genres avec lesquels il cohabite le plus fréquemment. Cela permettrait à l’utilisateur d’affiner petit à petit sa recherche.

- Utilisation concrète de la variable visuelle de couleur

Chaque morceau, représenté par une ligne sur le diagramme Parallel Coordinates, est d’une couleur spécifique en fonction de son genre. Cependant, l’information n’est pas utile en l’état. Une utilisation intelligente serait, d’afficher les morceaux correspondant aux sous-genres d’un genre sélectionné avec une couleur différente et une opacité plus faible, afin que l’utilisateur puisse se rendre compte des tendances en fonction des sous-genres.

Annexes

Liens :

Code source

- [Repo GitHub du projet](#)

Sources des graphs D3.js

- [Parallel coordinates](#)
- [Box-And-Whisker plot](#)