

# PROJET DATAVIZ - AVANCEMENT & CADRAGE

## TODO :

- Remplir les différentes sections demandées
- Récupérer nos données d'écoute
- 

## CADRAGE

### Quel est le problème abordé / à quel besoin répondez vous ?

Notre projet aborde le défi de la collecte et de l'analyse des données comportementales liées à la consommation musicale sur des plateformes de streaming telles que Spotify, YouTube Music et Deezer. Ces plateformes génèrent de grandes quantités de données sur les habitudes d'écoute des utilisateurs, mais ces informations sont souvent difficiles à exploiter. Notre objectif est de simplifier l'accès à ces données, de les organiser de façon utile et de les présenter de manière visuelle et interactive, permettant une analyse approfondie. Nous souhaitons offrir à chaque membre de l'équipe une vue détaillée de ses propres habitudes d'écoute tout en facilitant la comparaison entre membres pour mettre en lumière les préférences individuelles et les tendances collectives. Cela révèle des similitudes et des divergences sur les genres, artistes, morceaux et moments de la journée propices à l'écoute. Ce projet répond à plusieurs besoins : une visualisation claire et accessible des données, ce qui est essentiel pour les utilisateurs non experts en données, et la possibilité de comparer les comportements d'écoute entre membres afin de découvrir des points communs et de renforcer la dynamique d'équipe à travers la musique.

### À qui s'adresse la visualisation, quelles tâches seront effectuées au travers de votre projet ?

En termes de cible, cette visualisation s'adresse principalement aux membres de l'équipe de projet, qui pourront mieux comprendre les préférences musicales des uns et des autres, et aux analystes de données ou à des personnes intéressées par les habitudes de consommation. Par exemple, un membre de l'équipe pourrait utiliser ces visualisations pour mieux comprendre ses habitudes personnelles, tout en ayant un aperçu plus global de l'équipe. Les visualisations permettront également de guider des décisions plus informées dans des contextes collaboratifs, comme la sélection de musiques ou de playlists communes pour des événements de groupe ou des moments d'échange.

Les tâches qui seront effectuées à travers ce projet comprennent plusieurs étapes clés :

- Collecte des données : Nous devons d'abord rassembler les données provenant des différentes plateformes (Spotify, Deezer, YouTube Music), souvent via des API spécifiques à chaque service, afin d'accéder aux informations sur les morceaux écoutés, les genres, les artistes, ainsi que les moments de la journée où l'écoute est la plus active.
- Traitement et nettoyage des données : Après la collecte des données, il est essentiel de les structurer, de les nettoyer et de les préparer pour la visualisation. Cela inclut l'élimination des doublons, la gestion des valeurs manquantes et la normalisation des informations pour les rendre comparables.
- Création des visualisations dynamiques : L'étape suivante consiste à concevoir des visualisations qui permettent une exploration facile des données, telles que des graphiques à barres pour les genres et artistes préférés, des cartes thermiques pour les moments d'écoute, ou encore des diagrammes circulaires pour les répartitions par genre. L'objectif est de rendre ces visualisations interactives, de manière à ce que chaque utilisateur puisse naviguer à travers les différentes vues pour explorer ses propres données et celles des autres membres.
- Conception de dashboards interactifs : Un aspect central de ce projet est la création de dashboards interactifs qui permettent à l'utilisateur de filtrer, explorer et comparer les données selon différents critères. Ces dashboards offriront des outils de personnalisation, permettant de zoomer sur des périodes spécifiques, de comparer les écoutes entre

membres de l'équipe, et de visualiser les données sous différents angles (par genre, par artiste, etc.).

- Analyse des résultats : Après la création des visualisations, une analyse des résultats peut être effectuée afin d'interpréter les tendances observées. Par exemple, un membre de l'équipe pourrait découvrir qu'il écoute principalement de la musique le soir, tandis qu'un autre écoute plus tôt dans la journée. Ces découvertes peuvent ouvrir des discussions sur les moments propices pour organiser des activités en équipe ou des moments de détente autour de la musique.

### **Sources de données choisies**

**Lister les sources de données identifiées, pour chacune identifier leur intérêt principal et leurs limites potentielles. Clarifier si vous souhaitez en fusionner plusieurs ou pensez privilégier une en particulier. En cas de surprise/problème imprévu lié aux données, quel est votre plan de secours ?**

Les sources de données identifiées incluent principalement les demandes RGPD, qui permettent de récupérer nos historiques d'écoute, et les API des plateformes musicales (Spotify, Deezer, YouTube Music) pour obtenir des informations complémentaires comme les statistiques et les genres musicaux.

L'intérêt principal des demandes RGPD est qu'elles offrent une vue exhaustive et indépendante des restrictions liées à l'utilisation des API. Cependant, des opérations de prétraitement sont à prévoir, notamment en raison des différences de format entre les plateformes.

### **Travaux important liés au projet**

**Lister 3 projets liés au votre, par exemple : des projets avec des jeux de données similaires, des projets proposant des techniques de visualisation que vous trouvez intéressantes, des articles scientifiques présentant une technique ou une description du problème qui vous sera utile... Pour chaque travail, expliquer en quoi il vous semble intéressant par rapport à votre projet, et en quoi il pourrait être amélioré.**

#### **Spotistats (<https://spotistats.app/>)**

Spotistats est une application qui permet d'obtenir des statistiques détaillées sur ses habitudes d'écoute sur Spotify. Elle propose des visualisations claires et compréhensibles, notamment des graphiques sur les artistes ou morceaux

les plus écoutés et les périodes d'écoute. Cependant, certaines fonctionnalités comme la visualisation des albums sont réservées à une version payante.

Ce projet est pertinent pour notre travail car il démontre comment simplifier des données complexes en graphiques facilement compréhensibles. L'approche analytique de Spotistats, avec des représentations claires des tendances musicales, peut inspirer nos visualisations interactives en D3.js. Par exemple, l'analyse temporelle des habitudes d'écoute pourrait être traduite en graphiques animés, rendant les données plus dynamiques.

Cependant, on peut dépasser ses limites en ajoutant la possibilité de comparer les statistiques de plusieurs utilisateurs. Une telle fonctionnalité permettrait de croiser les données et d'identifier des similitudes ou des divergences dans les goûts musicaux. Par ailleurs, en intégrant des options interactives plus sophistiquées, comme des filtres ou des animations, on pourrait offrir une expérience utilisateur encore plus enrichie.

### Galaxy of Music (<https://galaxy.spotifytrack.net/>)

Galaxy of Music propose une visualisation immersive où les morceaux sont présentés sous forme de nœuds connectés dans un graphe galactique, avec un arrière-plan étoilé. Les relations entre les morceaux sont basées sur leurs genres ou d'autres attributs musicaux, créant une expérience visuelle attrayante et exploratoire.

Ce projet est intéressant pour nous car il illustre comment représenter des relations complexes dans un format graphique et immersif. En utilisant D3.js, on pourrait adopter une structure similaire pour visualiser les connexions entre genres, artistes ou albums dans notre base de données. L'aspect ludique et esthétique de la galaxie peut également rendre notre projet plus engageant et mémorable.

Toutefois, cette visualisation peut être améliorée en enrichissant les données présentées. Par exemple, on pourrait inclure des informations supplémentaires comme des exemples sonores, des visuels d'albums ou des statistiques sur la popularité des morceaux. De plus, intégrer des filtres interactifs ou des animations permettant d'explorer différentes catégories de musique (genres, périodes) rendrait l'exploration encore plus intuitive et captivante.

### Last.fm (<https://www.last.fm/>)

Last.fm est une plateforme qui analyse les habitudes d'écoute des utilisateurs grâce au scrobbing. Elle propose diverses visualisations, comme des graphiques sur les artistes les plus écoutés, les morceaux tendance et les cycles temporels des écoutes. En plus de ces outils analytiques, elle offre un aspect social où les utilisateurs peuvent comparer leurs goûts musicaux.

Ce projet est particulièrement pertinent pour nous car il combine des visualisations de données avec une dimension communautaire. On pourrait s'inspirer des graphiques de tendances temporelles pour représenter l'évolution des préférences musicales ou des habitudes d'écoute sur des périodes spécifiques. L'aspect social, permettant de comparer ses données avec d'autres, pourrait également enrichir notre projet en ajoutant une dimension collaborative.

Cependant, les graphiques de Last.fm, bien que fonctionnels, manquent souvent d'esthétisme et de dynamisme. En utilisant D3.js, on pourrait rendre ces visualisations plus modernes et immersives, avec des transitions fluides et des représentations visuelles engageantes. En outre, ajouter des visuels associés aux morceaux ou aux artistes, comme des pochettes d'album ou des extraits audio, améliorerait considérablement l'expérience utilisateur.

## Organisation

- Quels moyens de communications avez vous mis en place (email, tracking github, slack...) ?

Nous communiquons entre nous sur discord, ou à vive voix durant les pauses.

- Quelles sessions de travail avez vous prévu hors du cours d'ici à la soutenance de janvier ?

Pour les séances de travail en groupe, voici les prévisions pour la semaine du 9/12: mardi matin, aprem de 15h30 à 17h15, jeudi matin avant 9h45, jeudi aprem et vendredi aprem.

On souhaiterait arriver à un rythme de deux séances par semaine pour réussir à prendre un peu d'avance avant les vacances de noël.

- Quels rôles avez vous identifiés au sein du groupe (design, développement D3, pré-traitement des données, suivi, etc.) ? Ces

rôles ne sont pas exclusifs, et il est attendu que tout le monde contribue à la conception et au code des visualisations. Nous utilisons les fonctions de suivi de projet de github pour évaluer en partie la contribution de chacun au projet.

création d'un github pour partage du code, communication via discord

Thomas : github master (dev D3)

Céline : team leader, designer

Matthieu : prétraitement des données (mise en commun entre les différentes sources de données)

Clément : dev D3

Chacun de son côté a proposé des idées de visualisations différentes, et on a discuté sur la position des différentes visualisations et leurs possibles interactions afin d'apporter un dynamisme agréable.

### Scan des esquisses finales

Cf ci-dessus

**Quelles sont nos habitudes musicales ?**

**Quels genres musicaux sont les plus représentés dans nos écoutes ? Comment se comparent nos données avec celles d'autres utilisateurs ? = Quels genres musicaux sont les plus représentés dans nos écoutes et comment se comparent-ils avec les données des autres utilisateurs ?**

**Y a-t-il des tendances saisonnières (par exemple, certains genres plus écoutés en été/hiver) ? Quels sont les artistes et chansons les plus écoutés sur une période donnée (mois, année) ? Comment évoluent nos préférences musicales dans le temps ? Comment se comparent nos données avec celles d'autres utilisateurs ? = Quelles sont les tendances saisonnières, les artistes et chansons les plus écoutés, et comment nos préférences musicales évoluent-elles dans le temps par rapport à celles des autres utilisateurs ?**

**Quels nouveaux artistes ou genres avons-nous découverts récemment ?**

**À quoi ressemble notre diversité musicale ?**

**Quelle est la répartition des genres musicaux écoutés ?**

**Avons-nous une préférence pour certains artistes ou styles en particulier, ou est-ce que nos écoutes sont diversifiées ?**

**En comparant avec des données globales issues des services de streaming, nos préférences sont-elles uniques ou alignées avec les tendances générales ?**

**Quel est le volume de notre consommation musicale ? = Combien d'heures de musique avons-nous écouté en moyenne par jour, semaine ou mois ? Quels jours de la semaine ou moments de la journée sont les plus actifs en termes d'écoute ? Comment se comparent nos données avec celles d'autres utilisateurs ? = Quel est le volume de notre consommation musicale, en termes d'heures écoutées en moyenne par jour, semaine ou mois, quels jours de la semaine ou moments de la journée sont les plus actifs, et comment ces données se comparent-elles à celles d'autres utilisateurs ?**

## **AVANCEMENT**

**3/12**

### **Etape 1**

Comment pourrait-on représenter nos habitudes d'écoute?

- Les top genres
- Les top artistes
- Les top musiques

Comment pourrait-on rendre compte des genres musicaux que nous écoutons le plus?

Comment pourrait-on remarquer une tendance saisonnière dans nos tendances d'écoute?

Comment pourrait-on étudier l'évolution temporelle de nos tendances d'écoute ?

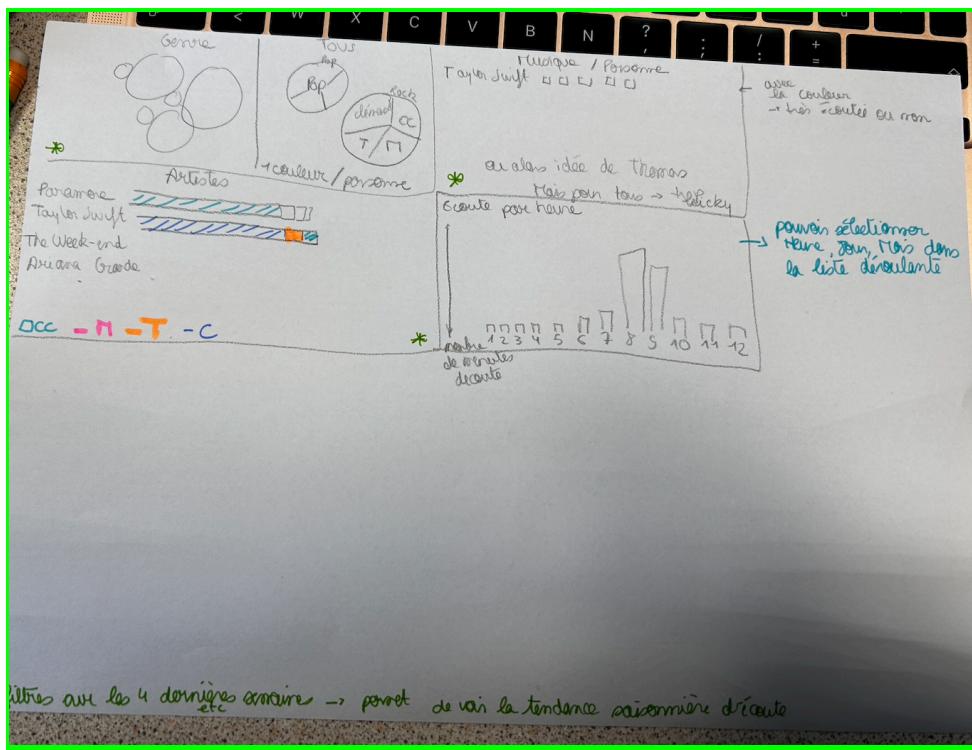
Comment pourrait-on comparer nos tendances d'écoute avec les tendances générales?

Comment pourrait-on comparer nos tendances d'écoute entre nous ?

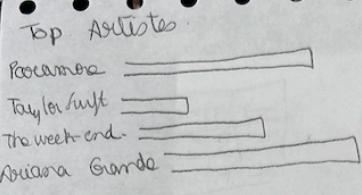
Comment pourrait-on démontrer les heures/jours à laquelle nous écoutons le plus de musique?

Comment pourrait-on afficher une moyenne par jour/mois/semaine de notre niveau d'écoute?

### **Etape 2**

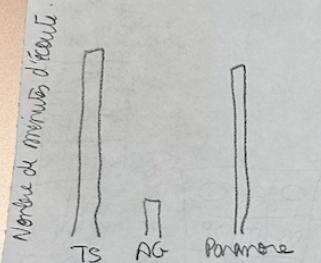


### Etape 3

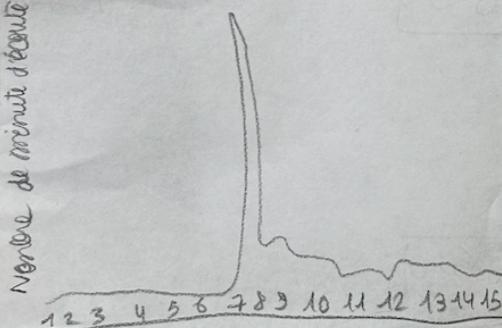


Avantage : on peut mettre plusieurs personnes dessous.

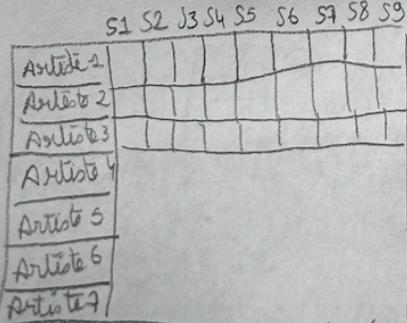
## Top Artists



*Gastric peristole.*



Top song



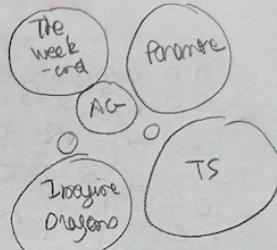
des plus au moins écrits



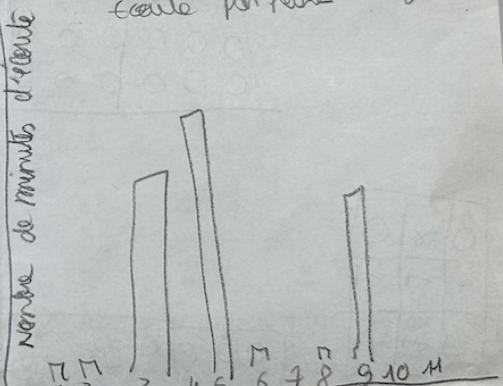
que / Personne  
□ E □ C

idée de th  
is pour tou  
-nne

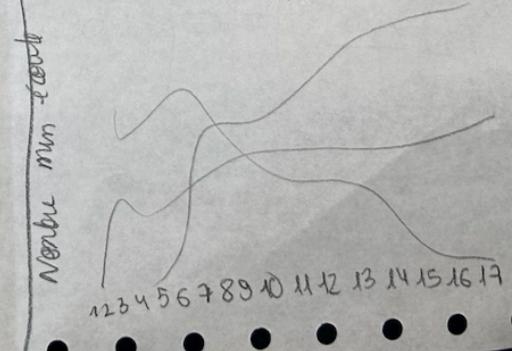
## Top Artists



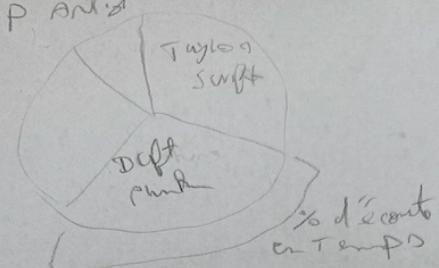
Favorait permettre de corriger par rapport aux deux  
étoile par ferme généraux.



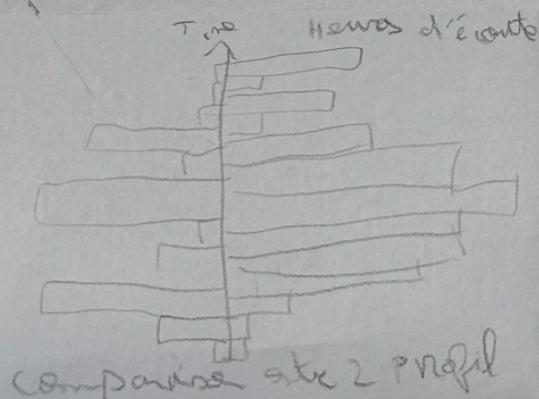
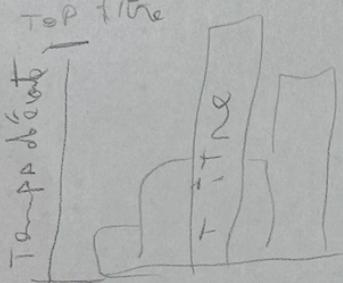
## Comparaison d'écoute par paire entre 2 personnes



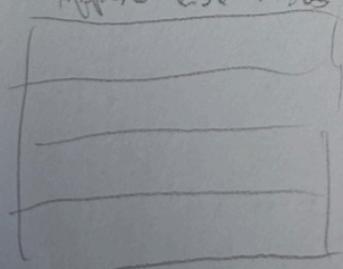
TOP Artist



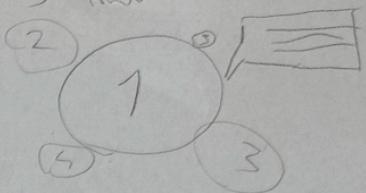
TOP Titre



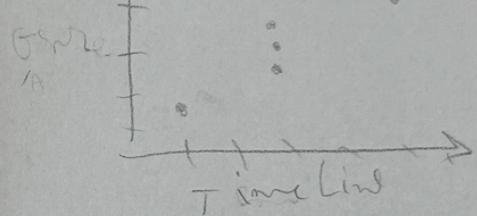
"Artistes comparaison  
effets liste artistes de connivance"



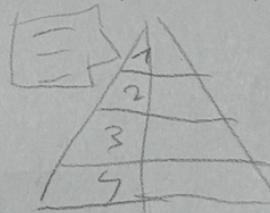
TOP 5 musiques



TOP Artistes  
Artistes +  
Genre



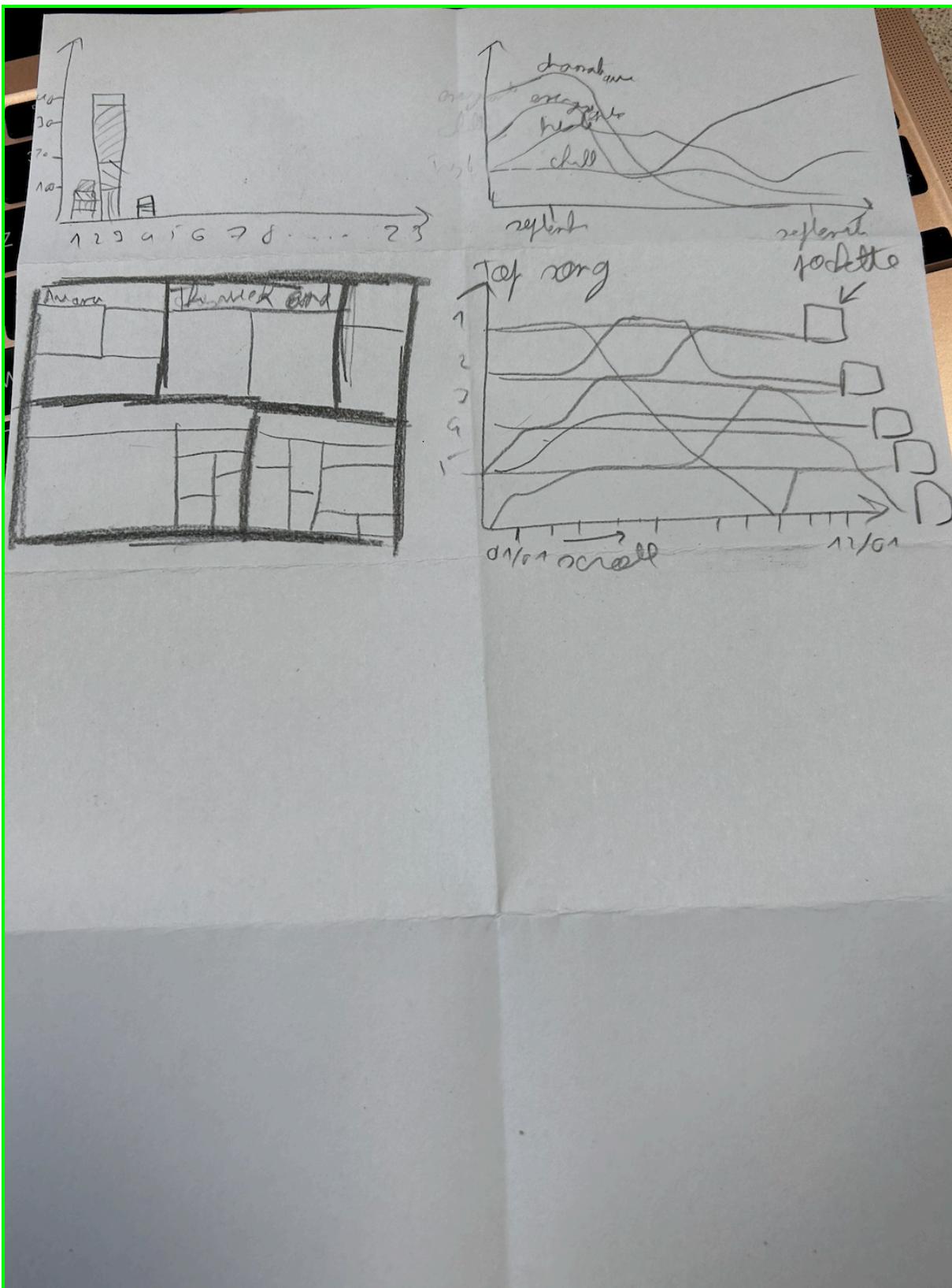
TOP comparaison



classe mat

TOP Artistes





control

>

<

w

v

x

c

z

y

g

h

b

n

i

o

u

e

ø

æ

å

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

ø

æ

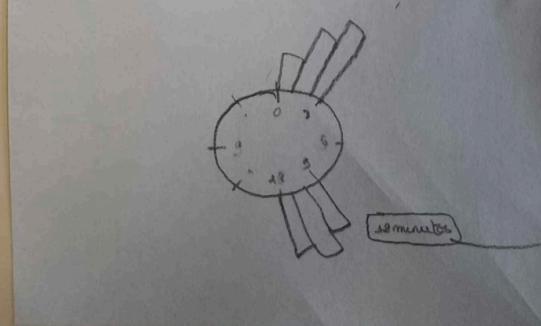
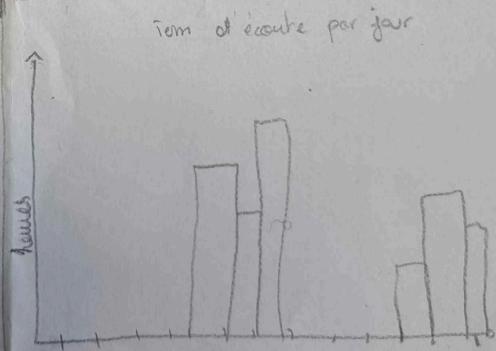
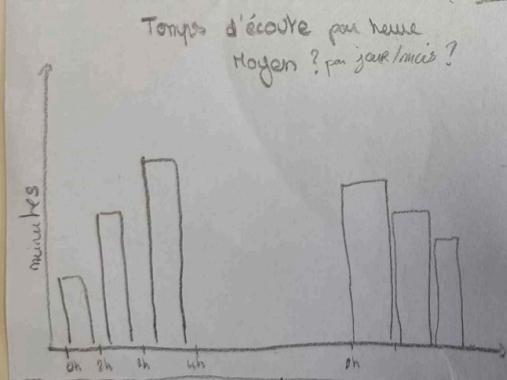
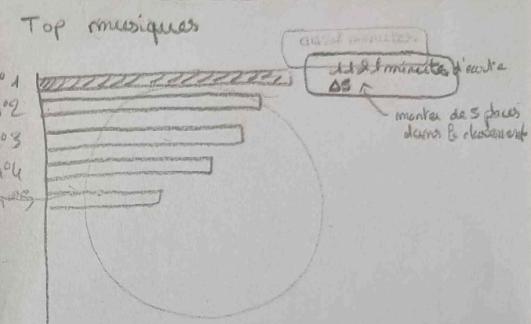
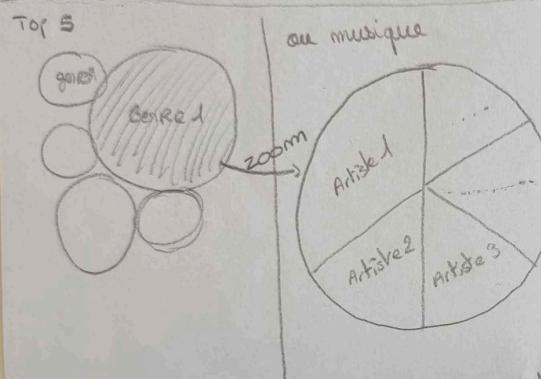
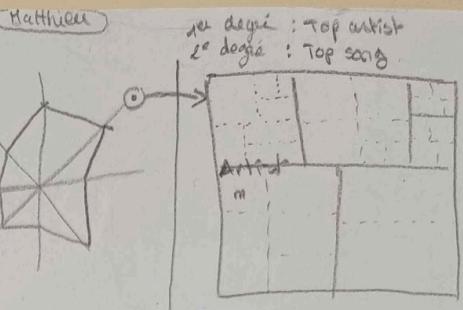
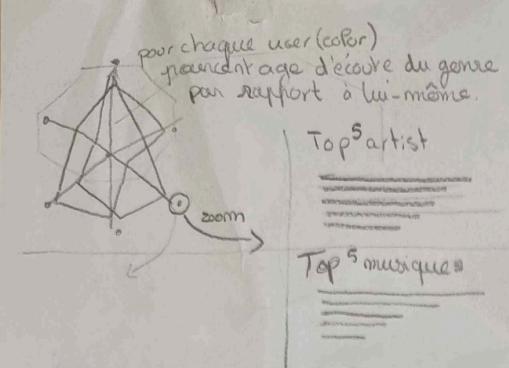
ø

æ

ø

æ

ø



- musique 1
- musique 2
- musique 3
- 
- 
-

## Exemple de visualisation

- Représentation des top songs

