

Prédiction de popularité de musiques

Proposition de problématique
Projet DASI

FOREST Philippe
FRANCOIS-MARTIN Camille
GOMEZ Guillian
ROUESNE Martin



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Sommaire

1. Contexte
2. Problématique
3. Résultats attendus
4. Données du projet
5. Enjeux et impacts
6. Conclusion
7. Bibliographie



1 - Contexte



Digitalisation de l'industrie

523 M d'abonnés aux plateformes de streaming en 2021

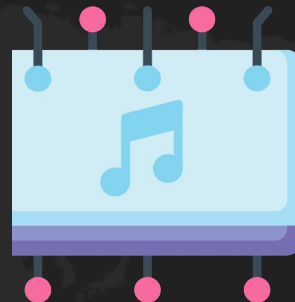
Marché de **34,53 MM \$** en 2022,
103 MM \$ en 2030



Influence des réseaux sociaux

Nouvelles opportunités marketing

Plus de données sur les consommateurs



Essor de l'IA et du Big Data

Prédictions sur la popularité d'un single

Analyser les tendances du marché

2 - Problématique

Approche principale

A partir de **nouvelles productions musicales** d'un artiste, comment orienter **géographiquement** les décisions marketing à faire pour **maximiser la popularité** de sa sortie?

Problématiques supplémentaires

Quelles communautés/réseaux à viser pour une campagne marketing en ligne?

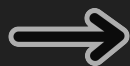
A quelle période faudrait-il sortir une nouvelle production?

3 - Résultats attendus

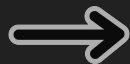
Objectifs



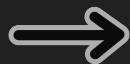
Segmenter les musiques par **zones géographiques**



Cibler la **meilleure zone** pour une musique donnée

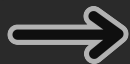


Prédire la **meilleure période**

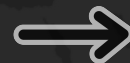


Prédire **réseaux/communautés** à viser sur internet

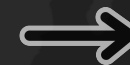
Indicateurs



Augmentation du **nombre d'écoutes** par musique



Répartition géographique (%) de l'audience en accord avec la stratégie conseillée



Augmentation de **l'engagement de l'audience (%)** sur les réseaux ciblés (comparé aux autres réseaux)

4 - Données du projet



Spotify API Database

Métriques Spotify décrivant la musique (tempo, valence, énergie...) et infos sur l'artiste, le genre musical, la date de sortie, etc.



Billboard 100 scraping

Plusieurs sites fournissent des classements par pays (Spotify, Youtube Music), mais plus difficile d'avoir l'évolution sur le temps



Bonus : social media scraping

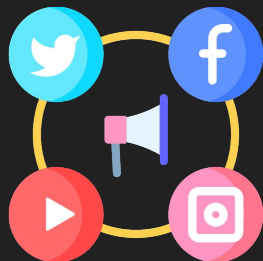
option 1 : Interactions avec l'artiste, notoriété

option 2 : Analyse de contenu sur les réseaux



5 - Enjeux et impacts

Opportunités



Faciliter la découverte pour des nouveaux artistes, **améliorer la visibilité** d'artistes déjà connus



Soutient le travail des équipes marketing, découverte d'un **nouveau potentiel d'audience**

Risques potentiels



Renforcement des genres populaires par pays, **pousse moins à la variété** pour favoriser le succès commercial



Classification **potentiellement non pertinente** (ex: rap français en Espagne?)

Biais des données utilisateurs Spotify (40% EU, 40% NA/SA, 20% others)

6 - Conclusion

Permettre aux **artistes et aux maisons de disques** d'orienter
leur campagne marketing vers la bonne audience
géographique adaptée au style musical de leurs prochaines
productions.

7 - Bibliographie

<https://datareportal.com/social-media-users>

<https://www.socialtracker.io/>

<https://towardsdatascience.com/country-wise-visual-analysis-of-music-taste-using-spotify-api-seaborn-in-python-77f5b749b421>

<https://www.businessofapps.com/data/music-streaming-market/>

<https://charts.spotify.com/home> (can't scrap legally)

<https://stackoverflow.com/questions/66880585/how-to-scrape-music-charts-insights-page-from-charts-youtube-com>

8 - Evaluation des slides

Vous avez 10 mins de présentation à faire.
Nous vous conseillons de faire des slides pour :

Soutenance 1 « Business Understanding »		
Phases CRISP-DM		Critères d'évaluations
« Business understanding » : Formulation d'un problème qui pourrait valoriser les données pour l'entreprise	Le contexte de l'étude est bien défini. La piste de valorisation s'inscrit dans la stratégie d'une organisation. Les hypothèses importantes sont identifiées, en particulier la disponibilité et l'accès des données. Les objectifs et les critères de succès sont explicités, le résultat attendu est bien défini : est-ce une analyse des usages, une compréhension des profils d'utilisateurs, le génération d'un score de propension à acheter, de la prédiction, etc..	CG1 : La problématique est-elle bien formulée ? Est-ce que le contexte de l'étude est présenté ?
		CG5 : Des pistes de valorisation sont-elles proposées ? Le résultat attendu est-il bien défini : la cible est-elle clairement décrite ? quel bénéfice quantifiable peut-elle en tirer ?
		CG11 : Les enjeux et impacts organisationnels sont-ils présentés ?

- introduire le contexte, l'écosystème, avec chiffres et références à l'appui ;
- présenter le problème ;
- le résultat attendu ;
- et montrer que votre projet est faisable : vous avez des données et vous avez identifié les variables a priori pertinentes ;
- identifier les opportunités et risques potentiels : enjeux et impacts sociétaux, organisationnels
- et pour finir une slide de conclusion où vous formulez explicitement votre valeur ajoutée : "avec notre analyse, nous allons offrir/permètre/améliorer (proposition de valeur) à/de (votre cible)".