

THESIS / THÈSE

MASTER EN SCIENCES DE GESTION À FINALITÉ SPÉCIALISÉE

L'impact du streaming musicale dans nos habitudes de consommations de la musique

VERMEERSCH, Alexandre

Award date:
2022

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



L'impact du streaming musicale dans nos habitudes de consommations de la musique.

Alexandre Vermeersch

Directeur: Prof. N. Steils

Mémoire présenté
en vue de l'obtention du titre de
Master 120 en sciences de gestion, à finalité spécialisée
en Business Analysis & Integration

ANNEE ACADEMIQUE 2021-2022

Université de Namur, ASBL
Faculté des Sciences économiques, sociales et de gestion – Département des Sciences de gestion

Rempart de la Vierge 8, B-5000 Namur, Belgique, Tel. +32 [0]81 72 49 58/48 41

Résumé :

Ce travail est une étude des consommations actuelles de la musique et de comment le streaming musical a impacté ces habitudes de consommation. Une revue de littérature sera présentée dont le but sera de poser un cadre théorique au travail. Dans cette revue, l'industrie de la musique sera analysée ainsi que le monde du streaming et ses technologies. Elle permettra également au lecteur d'en savoir plus sur les systèmes de recommandations musicales, leurs constructions, leurs impacts et d'éventuelles critiques à son égard.

Enfin le travail présentera une enquête qui a été faite montrant les possibles facteurs pouvant impacter le fait de se faire guider ou non pour écouter de la musique. Des analyses statistique seront faites et celles-ci permettront de savoir si oui ou non il existe des liens de corrélations entre différents facteurs et le fait d'être enclin à se faire guider.

Abstract :

This work is a study of current music consumption and how music streaming has impacted these consumption habits. A literature review will be presented in order to provide a theoretical framework for the work. In this review, the music industry will be analyzed as well as the streaming world and its technologies. It will also allow the reader to learn more about music recommendation systems, their construction, their impacts and possible criticisms towards it.

Finally, the work will present a survey that has been done showing the possible factors that can impact the fact of being guided or not to listen to music. Statistical analyses will be made and these will allow us to know whether or not there are correlations between different factors and the fact of being inclined to be guided.

Remerciement

Je tiens avant toute chose à remercier les différentes personnes qui m'auront aidé lors de la rédaction de ce mémoire.

Je voudrais d'abord remercier ma directrice de mémoire, Madame Nadia Steils, pour sa disponibilité, sa patience et ses conseils avisés qui m'auront permis d'alimenter ma réflexion et de toujours aller plus loin dans mon travail.

J'adresse mes remerciements à l'Université de Namur pour l'ensemble des compétences qu'elles m'ont transmises durant ces années d'étude.

Enfin je tiens à remercier ma famille et mes amis pour leur bienveillance et leur soutien.

Tables des matières

1) Introduction	5
2) Revue de littérature	7
2.1) Focus sur le marché actuel de la musique et sur sa consommation	7
2.2) Les plateformes de streaming musical	10
2.3) Influence des plateformes de streaming sur le monde de la musique	11
2.4) Les technologies des plateformes de streaming musical : datafication, algorithmes et affordance	13
2.5) Technologie appliquée pour les consommateurs de musique	14
2.6) Les profils d'auditeurs	15
2.7) Deux processus de création d'algorithmes : le cas de Pandora Internet Radio et de Spotify	16
2.7.1) Pandora Internet Radio	16
2.7.2) Spotify et la "Discover Weekly"	17
2.8) Critique de l'efficacité des recommandations musicales et de la raison algorithmique	18
2.9) Perspective d'avenir	19
2.10) Enquête sur l'autonomie de l'auditeur	20
2.11) La question de l'hyper-choix	20
3) Questions de recherche et hypothèses	22
4) Méthodologie	24
5) Analyse de données	28
5.1) Préparation des données	28
5.2) Statistiques descriptives	29
5.3) Test sur les hypothèses	32
5.3.1) ANOVA : Hypothèse numéro 1 - La catégorie d'âge et le fait d'être enclin à se faire guider pour écouter de la musique	32
5.3.1.1) Prémisses à l'analyse ANOVA	32
5.3.1.2) Analyse ANOVA	34
5.3.1.3) Test de comparaisons multiples	34
5.3.1.4) Analyse du R ²	35
5.3.1.5) Conclusion ANOVA	36
5.3.2) ANOVA : Hypothèse numéro 2 - Le temps de consommation et le fait d'être enclin à se faire guider pour écouter de la musique et Hypothèse numéro 4 - Le fait de consommer les produits actuels de la plateforme de streaming et le fait d'être enclin à se faire guider pour écouter de la musique	36
5.3.2.1) En conclusion de ces tests d'ANOVA	37
5.3.3) Analyse par la régression simple : Hypothèse numéro 5, lien entre hyper-choix et le fait d'être enclin à se faire guider	37

5.3.3.1) Rappel des hypothèses	37
5.3.3.2) Etape 1 : Evaluation de la pertinence du modèle de régression	38
5.3.3.3) Etape 2 : Évaluation de l'ajustement des données au modèle de régression	38
5.3.3.4) Conclusion	40
5.3.4) Analyse par la régression simple : hypothèse numéro 3, lien entre l'adéquation de l'offre des plateforme de streaming et le fait d'être enclin à se faire guider	40
5.3.4.1) Conclusion	41
6) Conclusion	42
6.1) Retour sur la discussion	42
6.2) Limites	44
6.3) Recommandations managériales	45
7) Bibliographie	46
7.1) Bibliographie de la littérature	46
7.2) Bibliographie des sites internet consultés	49
8) Annexes	52

1) Introduction

En 1954, Leonard Bernstein, un célèbre compositeur américain, disait : “On ne vend pas la musique, on la partage.”

Même si sur le principe on ne peut qu’être d’accord, la musique en tant que bien immatériel ne se vend pas, elle s’écoute et se partage. Dans les faits, ça n’a pas toujours été le cas.

Depuis 1877 et la création du phonographe par Thomas Edison (L.Tournès, 2011), la musique n’a eu de cesse de se renouveler tant en matière de style, qu’en matière de support d’écoute. Le gramophone, les vinyles, les cassettes compactes, le walkman de Sony, les CD ont jusqu’aux années 2000 tous impacté le marché de la musique qui jusque-là se vendait bien, atteignant même un revenu de 24 milliards d’euros en 1999.

Cependant comme s’il s’agissait là d’une prédiction, les mots de monsieur Bernstein ont trouvé un sens quand, en 1999, le site de téléchargement illégal Napster est né. Ce site qui sera une sorte de fusion entre les technologies *peer to peer* et MP3 permettra à tous ses utilisateurs d’avoir accès à un catalogue immense de musique et ce pour 0 euro (L. Tournès, 2011).

La musique va, à ce moment-là, juste se partager à travers le monde entier et ce sans en tirer des bénéfices. Le marché de la musique, lui, va connaître une importante crise économique, probablement la plus grosse et la plus longue de son histoire. En chutant entre 1999 et 2014 de 24 milliards de dollars à 14,2 milliards de dollars (IFPI, 2022), l’industrie sera pendant longtemps plongée dans le noir et il faudra attendre 2008 et la création d’un site par un suédois nommé Daniel Ek pour entrevoir une éclaircie (Englebert, 2018).

Avec la création de Spotify, Daniel Ek ne le sait pas encore mais il vient d’inventer un *business model* qui va régner sur le monde de la musique pendant de longues années tout en sauvant son marché. Parfois considérés comme un problème au niveau des rémunérations des artistes par exemple (Boucher, 2020), Spotify et le streaming musical en général auront, sans contestation possible, permis au marché de la musique de connaître une hausse dans ses revenus globaux et ce depuis 2014. Pourtant, ce *business model* n’est toujours pas basé sur de la vente de musique. En effet, sur ces sites et applications, on ne peut pas acheter de musique. Leonard Bernstein aurait donc encore en partie raison.

En partie car il est désormais possible d’actualiser sa citation : maintenant, on ne vend pas la musique, on vend le partage de celle-ci.

Mais alors comment la vend-on ? Quelles sont les techniques de ventes et comment impactent-elles la vie, les choix et la consommation des auditeurs ?

Avec plus de 520 millions d’abonnés à travers le monde aujourd’hui et une hausse de 109,5 millions par rapport à l’année passée (AFP, 2022), les services de streaming musicaux sont omniprésents dans nos vies de tous les jours. Travailler à répondre à ces différentes questions est particulièrement intéressant.

Ce sont les raisons qui m’ont poussé à faire ce travail qui tentera d’un côté d’expliquer et de résumer les techniques de ventes nouvelles qu’ont apporté la technologie et les services de streaming. Et d’un autre côté, d’étudier comment le consommateur est impacté par ces

technologies ; bref, quel est l'impact des services de streaming musicaux sur la consommation de musique.

Pour étudier tous ces sujets, le mémoire sera divisé en plusieurs parties.

D'une part, une revue de littérature qui permettra de poser un cadre théorique sur le marché de la musique et le marché du streaming mais également de définir les différentes technologies utilisées, comment elles sont utilisées, par qui et pour qui. La revue de littérature permettra également de comprendre quels sont déjà les impacts sur les consommations en revenant sur des enquêtes et des travaux qui ont déjà été effectués. Ce travail reviendra également sur le thème de l'hyper-choix, un phénomène de plus en plus présent dans nos vies de tous les jours et donc également présent dans le marché de la musique.

D'autre part, ce mémoire étudiera les comportements d'écoute des auditeurs et l'influence qu'ont les services de streaming musical sur ces comportements afin de voir s'il existerait des tendances générales et ainsi de déceler quels impacts ont les services de streaming actuellement sur nos consommations d'écoute. Pour étudier cela, une enquête quantitative a été effectuée et les résultats de celle-ci seront donc dévoilés et analysés.

Enfin, une conclusion générale viendra boucler le travail en discutant des recherches déjà faites et des résultats de notre recherche. Cette conclusion permettra également d'énoncer les limites éventuelles de notre travail et de se pencher sur les perspectives d'avenir du marché de la musique.

2) Revue de littérature

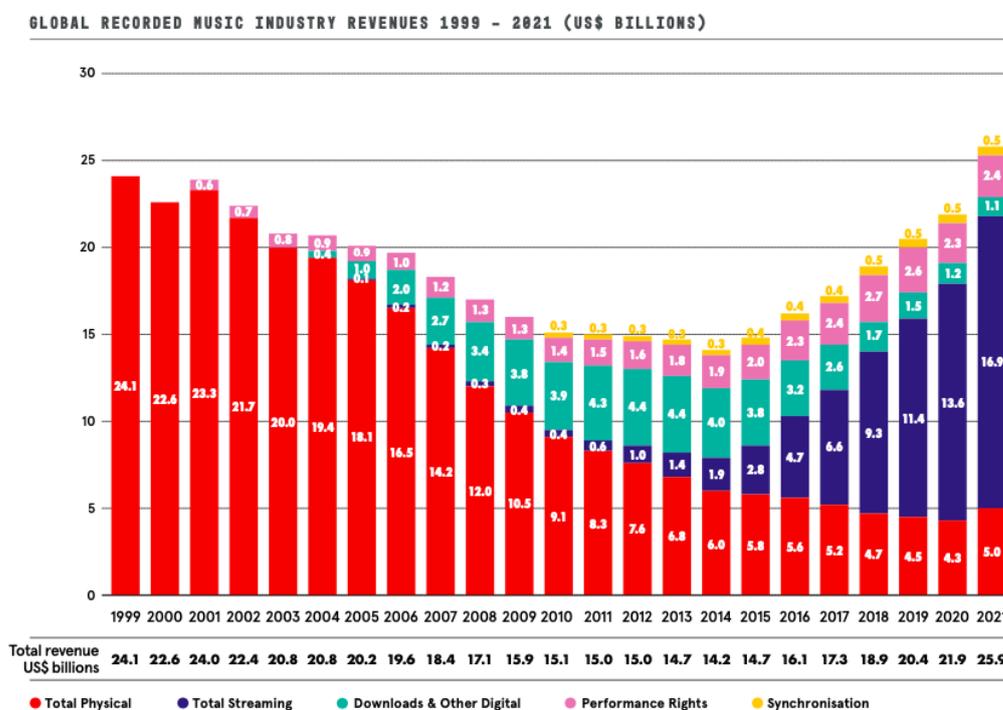
2.1) Focus sur le marché actuel de la musique et sur sa consommation

Avant de commencer cette revue de littérature et donc de se plonger un peu plus dans mon mémoire, il est bon d'appliquer un cadre théorique au marché actuel de la musique.

L'IFPI est la Fédération internationale de l'industrie phonographique, active depuis 1933. Il s'agit donc là du plus gros organisme syndical international pour les artistes.

Le rapport qu'elle sort chaque année est une source pertinente et riche pour connaître la situation du marché de la musique.

Actuellement, et ce après une crise sanitaire mondiale, le marché de la musique connaît une hausse de 18,5% de ses revenus globaux. Son revenu total a ainsi atteint 25,9 milliards de dollars en 2021 alors qu'il s'élevait à 21,9 milliards de dollars en 2020, à 20,4 milliards de dollars en 2019 et 18,9 milliards en 2018. Bref, à l'inverse de nombreux autres secteurs économiques, le secteur de la musique n'aura pas été touché négativement par la crise sanitaire mondiale de la COVID-19. En fait, ce revenu est même le plus haut enregistré depuis les années 2000 (IFPI, 2022).



Figure

e 1.1- Histogramme du revenu global du marché de la musique en milliards de dollars en 2022.

Ainsi, comme on peut le voir sur ce graphique qui rapporte les revenus totaux de l'industrie de la musique, celle-ci connaît une baisse de 1999 à 2014 et, depuis cette même année, elle est en constante hausse (IFPI, 2022).

Cette croissance peut être corrélée à l'augmentation des recettes des plateformes de streaming qui sont maintenant devenues les premières actrices du marché musical avec la part de revenu la plus importante de ce marché (IFPI, 2022).

En effet, depuis 2018, la part générée par le streaming dans les revenus globaux de l'industrie de la musique ont peu à peu augmenté, jusqu'à en représenter plus de la moitié 49,2% en 2018, 55,8% en 2019, 62,1% en 2020 et 65% en 2021 (IFPI, 2022). Une nouvelle preuve que les plateformes de streaming sont désormais omniprésentes dans le marché de la musique.

Ces 65% peuvent être divisés en deux parties : une première qui représente les revenus liés aux abonnements aux services de plateforme musicale (47,3%) et une deuxième qui représente les revenus liés aux publicités sur ces mêmes plateformes de streaming musical (17,7%), revenus donc liés aux services gratuits, sans abonnement, que proposent les plateformes de streaming musical (IFPI, 2022).

Mais avec 65%, le streaming ne constitue pas le seul segment de cette industrie musicale. Les 35% restants sont composés :

- des droits d'exécution (9,4%) qui peuvent être assimilés aux droits d'auteurs et sont donc des revenus que les différents acteurs de la musique touchent quand une de leurs œuvres est diffusée dans des lieux publics (radios, concerts...) (Loïs, 2022);
- des droits de synchronisation (2,1%) qui peuvent se résumer aux revenus que touchent les différents acteurs de la musique pour que leurs œuvres soient utilisées dans des productions audiovisuelles (films, séries...);
- les revenus des ventes physiques (19,2%) (CD, vinyles, cassettes...);
- les revenus liés aux ventes digitales de la musique (le fait d'acheter une version digitale d'un album ou d'un titre) (4,3%).

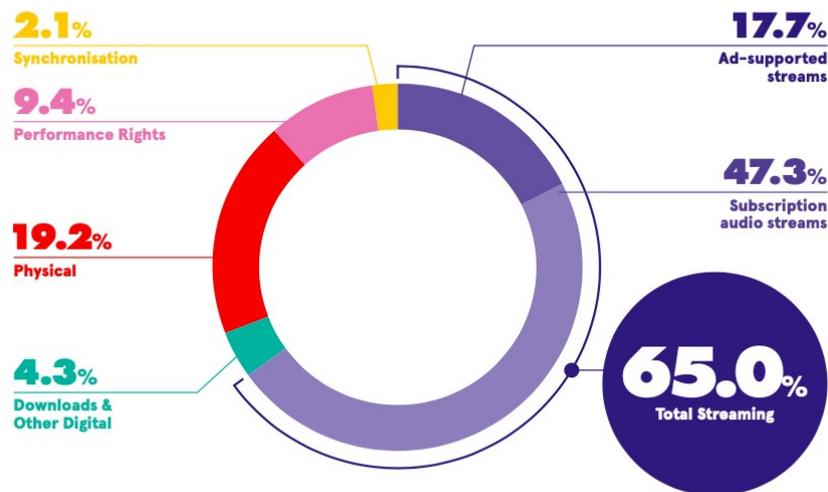


Figure 1.2- Graphique en secteur des parts de revenus des secteurs de la musique dans le marché de la musique en 2022.

Cette hausse des revenus de l'industrie de la musique est présente dans tous les continents du monde avec une hausse de (IFPI, 2022) :

- 16,1% en Asie ;
- 31,2% en Amérique latine ;
- 35% au Moyen-Orient et en Afrique du Nord ;
- 9,6% en Afrique subsaharienne ;
- 15,4% en Europe ;
- 4,1% en Australie ;
- 22% au Canada et aux USA.

Ces chiffres reflètent bien que la musique en tant que marché est présente partout. Cela peut être mis en relation avec le top 10 des marchés de la musique en 2021 (IFPI, 2022) :

- | | |
|-------------------|----------------|
| 1 USA | 6 Chine |
| 2 Japon | 7 Corée du Sud |
| 3 Grande-Bretagne | 8 Canada |
| 4 Allemagne | 9 Australie |
| 5 France | 10 Italie |

Enfin, cette hausse du revenu global du marché de la musique peut être expliquée par le fait que 4 segments sur 5 qui composent l'industrie de la musique sont en augmentation (IFPI, 2022)

- Le streaming connaît une hausse de 24,3%.

- Les ventes physiques connaissent une hausse de 16,1%.
- Les droits d'exécution connaissent une hausse de 4%.
- Les droits de synchronisation connaissent une hausse de 22%.
- Les ventes digitales connaissent une baisse de 10,7% ce qui peut sembler logique vu la hausse des plateformes de streaming et l'offre quasiment identique qu'ils proposent.

Le marché de la musique est donc de nouveau un marché grandissant et cela peut être expliqué par l'avènement du streaming qui offre des nouvelles technologies et un nouveau *business model* aux consommateurs.

Du système de playlist aux systèmes de recommandations musicales et donc aux algorithmes, les plateformes de streaming sont devenues les moyens les plus utilisés de consommer de la musique. C'est donc naturellement que les consommateurs les ont adoptées comme ils ont, à l'époque, adopté les autres nouveautés technologiques.

La suite de la revue de littérature se focalise plus sur les plateformes de streaming et leurs technologies.

2.2) Les plateformes de streaming musical

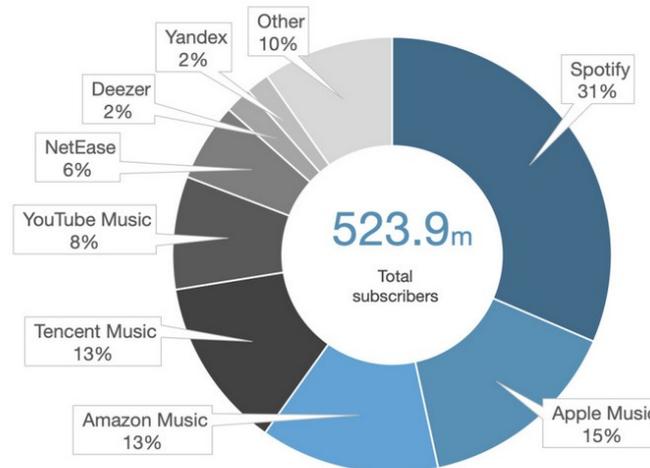
Qu'entendons-nous par plateformes de streaming musical ? Il s'agit d'un site d'hébergement sur internet qui propose des services de streaming c'est-à-dire un service de lecture de contenu audio voire vidéo sans avoir à télécharger ce même contenu. Ces fonctionnalités nécessitent donc une connexion à internet afin de pouvoir profiter du service de streaming. On peut aujourd'hui différencier deux types d'offres :

- les abonnements à ces services sous forme payante avec, à une certaine échéance (le plus souvent mensuelle), un tarif à payer (allant de 4,99 euros à 19,99 euros) (Husson, 2022);
- une deuxième offre qu'on peut qualifier de *freemium* où l'auditeur aura accès à la musique avec certaines contraintes au niveau des choix possibles (il ne pourra pas choisir la musique qu'il désire écouter), au niveau de la qualité (la qualité sonore pour un abonnement payant est plus élevée que pour un abonnement gratuit). Cette offre implique également des coupures de publicités audio pendant l'écoute (Husson, 2022).

Il existe actuellement plusieurs acteurs sur le marché du streaming musical, Spotify en est le leader mondial. Le graphique ci-dessous montre les parts de marché du streaming musical à la fin de l'année 2021. Cela nous permet également de dresser une liste non-exhaustive des acteurs actuels du marché qui sont par ordre de part de marché d'abonnements : Spotify, Apple Music, Amazon Music, Tencent Music, Youtube Music, NetEase, Deezer, Yandex. (Mulligan,2021)

Global streaming music subscription market, Q2 2021

Global streaming music subscription market, Q2 2021, global



Source: MIDiA Research Music Subscriber Market Share Model 11/21

MIDIA.

Figure 1.3- Graphique en secteur des parts d'abonnées en pourcentage du marché du streaming musical.

Il s'agit donc d'un marché concurrentiel où les différents acteurs essaient de se démarquer bien que le produit vendu (streaming de musique) reste en définitive le même. Pour se démarquer, ces différents acteurs essaient de se différencier par leurs stratégies marketing, leurs offres d'abonnements mais aussi via leurs technologies.

2.3) Influence des plateformes de streaming sur le monde de la musique

Avec 65% des parts des revenus globaux de la musique et plus de 520 millions d'abonnés dans le monde, les services de streaming musical sont devenus indispensables pour les acteurs de l'industrie musicale.

Les plateformes de streaming ont un certain "pouvoir", notamment celui de faire ou défaire un artiste. Les travaux de A. Maaso et A.N. Hagen en 2019 dans lesquels une enquête qualitative a été menée en passant des entretiens auprès d'acteurs de l'industrie musicale

peuvent confirmer cette hypothèse. Un manager de musique allant jusqu'à dire que "Si quelqu'un à Spotify décide de ne pas te placer en playlist, le titre peut tomber totalement dans l'oubli" (Maaso et Hagen 2019).

Ce "pouvoir" est en partie dû au système de plateforme digitale. Le terme plateforme qui est emprunté au monde industriel est un terme neutre (Gillepsie, 2010). Or cette nouvelle façon de "vendre" (ici on pourrait même parler de prêter car rien n'est vendu sur les plateformes de streaming musical) n'est pas neutre. La plateforme en tant que telle peut mettre en avant des artistes, des musiques, des albums ou des playlists. Dès lors on parle de "plateformisation de production culturelle" (Nieborg et Poel, 2018).

Les plateformes comme Spotify, Deezer ou Apple Music seraient donc les nouvelles radios grâce à leur influence sur le monde de la musique car elles peuvent changer les habitudes de consommation ou faire découvrir des musiques, albums et artistes rien qu'en les mettant en avant dans le menu d'accueil de la plateforme (Prey, Esteve Del Valle, Zwerwer 2022).

De plus, les services de streaming proposent également un système de playlists qui est actuellement très populaire. Par exemple, la playlist "Today's top hits" de Spotify qui est générée par les algorithmes est plus écoutée que n'importe quelle radio américaine (Music Business Worldwide, 2018). Cela change donc forcément la donne et la vision des acteurs de la musique sur ces services. Le fait est que ces playlists sont devenues un moyen incontournable pour les artistes d'être écoutés. Cependant, même si les utilisateurs peuvent créer eux-mêmes leurs playlists et les partager (et donc on pourrait penser, mettre en avant les artistes de leurs choix), cette représentation des auditeurs n'est, dans les faits, pas présente (Prey et al., 2022).

En effet, la plupart des playlists les plus écoutées sur les plateformes sont créées par les plateformes elles-mêmes, renforçant ainsi ce pouvoir qu'elles ont sur la popularité des artistes. Par exemple, les 35 playlists les plus écoutées de Spotify sont toutes créées par le service. Ce sont toutes des playlists où la plateforme a le "pouvoir" de décider si tel ou tel titre peut en être (et donc si tel ou tel titre peut jouir de la renommée de celle-ci). C'est également le cas pour 99 des 100 premières playlists de la plateforme qui appartiennent également à Spotify (Prey et al., 2022).

Cela devient alors un "jeu" pour les managers et gérants de labels. Par exemple, du fait de leur importance, les playlists du type "Top 50" sur les différentes plateformes offrent une

grande visibilité et le fait d’y entrer confirme le succès d’un titre. Sachant cela, les managers et gérants de labels feront tout pour y faire entrer leurs titres. Dans ce but, s’il manque par exemple 50 000 écoutes à un titre pour être dans le “Top 50”, ils essaieront de placer ce même titre dans une playlist plus “nichée” afin de récolter les écoutes manquantes et d’ensuite pouvoir espérer rentrer dans la playlist plus grande. (Maaso et Hagen 2019).

2.4) Les technologies des plateformes de streaming musical : datafication, algorithmes et affordance

Tous les services de streaming musical sont construits de la même manière : une interface visible qui propose des caractéristiques techniques (comme les boutons, la *scroll bar*, les icônes...), des aspects réglementaires (le fait qu’on soit obligé de créer un compte personnel) et des fonctions interactives qui permettent aux auditeurs de partager, organiser et chercher leurs musiques voire même d’être créatifs (via la création de playlist par exemple) (Van Dijk, 2014).

En plus de cela, il existe également une interface invisible, une sorte de formule secrète qui fait que certains contenus vont être mis en avant. Elle serait à la base des recommandations musicales automatisées et serait composée des différents algorithmes (Van Dijk, 2014).

On pourrait compter actuellement deux types généraux d’algorithmes : les “*trending algorithms*” dont le but serait de proposer à l’auditeur des contenus en fonction des chansons populaires du moment et les “*collaborative filtering*” dont le but est de proposer des musiques aux auditeurs en fonction des écoutes d’autres auditeurs (Maaso et Hagen, 2019). Par exemple, si l’auditeur A écoute et aime une musique XY alors on essaiera de lui proposer une musique XX car l’auditeur B qui a également écouté et aimé la musique XY aime et écoute aussi la musique XX.

Il y a 3 concepts à bien comprendre : la datafication, les algorithmes et les affordances.

La datafication permet d’expliquer comment un certain nombre de phénomènes de notre vie quotidienne peuvent être traqués, quantifiés et analysés dans le but d’ensuite pouvoir faire des conclusions prédictives (Mayer-Schönberger et Cukier, 2013). Ce phénomène est de plus en plus présent dans beaucoup de sociétés digitales telles que les réseaux sociaux ou autres et permet d’élargir le business model des services de streaming musicaux. Grâce aux données qu’ils récoltent à chaque écoute, ils peuvent savoir quelles musiques sont écoutées, par quels profils, quand et où. Toutes ces informations pourront ensuite être vendues à des

“stakeholders” de l’industrie de la musique tels que les artistes, les labels et les autres acteurs de l’industrie de la musique (Maaso et Hagen, 2019). Un exemple de ce phénomène est le service “*Spotify for artist*” qui, depuis 2017, permet aux artistes d’avoir accès à leurs données d’écoute. Des informations qui peuvent être cruciales pour par exemple l’organisation d’une tournée de concerts ou d’une sortie d’album (Maaso et Hagen, 2019).

La datafication devient primordiale dans le fait de gérer son label. Selon des acteurs de l’industrie de la musique interviewés par A. Maaso et A.N.Hagen en 2019, “la datafication est devenue primordiale pour le planning relationnel et stratégique, la communication, la coopération, l’exécution des lancements de musiques, le marketing, les publications, le ciblage d’audience et les approches des nouveaux marchés”. Certains labels utiliseraient même le machine learning pour signer des nouveaux artistes.

Il n’existe à ce jour pas de consensus scientifique sur ce que sont les algorithmes (Bucher, 2018). Cependant, certaines définitions existent tout de même. Selon Striphas en 2015, “certaines des définitions les plus utilisées des algorithmes les décrivent comme des processus formels ou des ensembles de procédures étape par étape qui sont souvent exprimés mathématiquement ou codés pour transformer des données d’entrée en un résultat souhaité basé sur des calculs spécifiques”.

L’affordance pourrait être résumée comme ce qui est offert : l’environnement, le contexte, les relations sociales... (Gibson, 1986). Dans le cadre de la musique, l’affordance serait donc ce qui est donné aux auditeurs comme le menu d’accueil de la plateforme par exemple. Les algorithmes pourraient modifier ce qui est “donné” et donc ainsi changer l’environnement musical des auditeurs et donc l’affordance.

2.5) Technologie appliquée pour les consommateurs de musique

Qu’en est-il alors des consommateurs de musique ? Les plateformes de streaming, en plus de mettre en avant des musiques, artistes ou albums de par la simple présentation du menu d’accueil (Gillepsie, 2010), ont développé des systèmes de recommandations musicales.

Avant de continuer, il serait donc bon de définir ce qu’est un système de recommandation. Parmi les lectures que j’ai pu faire à ce sujet, une idée d’une thèse d’Octavian Rolland Arnautu m’a paru très explicite. Il voit les systèmes de recommandation comme un DJ qui préparerait de la musique pour un évènement. Celui-ci choisit la musique en fonction de son

expérience dans le domaine musical, de l'événement et des personnes présentes à cet événement mais également en fonction de la réaction de la foule (O. R. Arnautu, 2012).

Un système de recommandation au final est donc un service qui grâce aux technologies qu'il a pu développer, aux informations liées aux profils des auditeurs qu'il a pu obtenir et aux réactions de ces mêmes auditeurs vous proposera une musique plutôt qu'une autre.

Plusieurs questions sont donc à traiter : comment sont vus les profils d'auditeurs, quelles sont les techniques et technologies développées par les services de streaming musical et comment sont-elles appliquées ?

2.6) Les profils d'auditeurs

Il est toujours bon de faire un petit retour historique pour expliquer les phénomènes d'aujourd'hui. Dans les années 1950 et avec l'industrialisation, les consommateurs étaient vus comme des masses de consommateurs (Williams, 1958).

Près de 60 ans plus tard, l'industrialisation a laissé place aux médias et à la personnalisation toujours plus présente grâce notamment aux technologies et à la digitalisation des médias et des services qui la rendent possible. Cependant il est bon de spécifier que sur ces plateformes, il n'y a en fait pas d'individus mais seulement des manières de voir les individualités (Prey, 2018).

De plus, avec l'avènement d'internet, ces individus seraient désormais plus vus comme des "utilisateurs" voire des "producteurs". Car au final, ils interagissent désormais avec les médias en ligne et par la même occasion produisent des données qui permettent à ceux-ci de créer des profils cibles et une personnalisation de plus en plus précise (Bruns, 2008).

Enfin, un concept est né de ces personnalisations : l'identité algorithmique. L'identité est créée grâce à nos actions et nos usages en lignes qui nous classifient donc dans un type mesurable comme par exemple : "une femme universitaire de la trentaine", "un papa sportif du dimanche" (Cheney-Lippold, 2017)...Ces profils sont ensuite utilisés dans le but de nous définir et de créer des filtres pour qu'ensuite un média puisse nous vendre quelque chose ou communiquer directement par rapport à notre identité (Prey, 2018). Cette identité algorithmique n'est jamais définitive et peut totalement changer en fonction de nos actions et usages (Cheney-Lippold, 2017).

Les plateformes de streaming faisant partie des “nouveaux médias”, ils ne dérogent donc pas à la règle et créent également des identités algorithmiques. Ceci dans le but d’avoir le système de personnalisation le plus précis possible et ainsi se démarquer de la concurrence des services de streaming musical grâce à un meilleur service de recommandations musicales, ou d’avoir des identités algorithmiques permettant pour les non-abonnés d’avoir des publicités en lien avec qui ils sont. Voire même, comme dit précédemment, afin de vendre des informations vitales aux acteurs de l’industrie de la musique.

Maintenant que nous savons ce qu’est une identité algorithmique, il est intéressant de savoir comment on la crée et comment les plateformes de streaming musical l'utilisent pour leurs services de recommandations musicales. Le travail de Robert Frey en 2018 “*Nothing personal: algorithmic individuation on music streaming platforms*” est une source très utile car elle explique le processus de création des recommandations musicales de deux plateformes de streaming musicales : Pandora Internet Radio et Spotify. Cela nous permettra également de savoir ce qu’est une identité algorithmique au niveau du monde de la musique.

2.7) Deux processus de création d’algorithmes : le cas de Pandora Internet Radio et de Spotify

2.7.1) Pandora Internet Radio

Pandora Internet Radio est une webradio très populaire aux États-Unis et une pionnière dans le streaming mais également la personnalisation de son service. Ils ont créé une database appelée “Music Genome Project” dont le principe n’est pas de classer les musiques par genre ni par *collaborative filtering* ou classement mais par traits ou gènes musicaux (Prey, 2018).

Un gène musical pourrait, par exemple, être le tempo, le fait que la musique contiennent des sons de guitare ou même le sexe de l’artiste. Il existe donc une multitude de gènes qui diffèrent selon les genres musicaux. Le rock en compterait 150, le rap 350 et le jazz 400 par exemple (Glaser et Prior, 2006).

Ces gènes seraient construits par des employés de Pandora et il faut en général une analyse et construction de gène de 20 à 30 minutes pour une musique de 4 minutes. Chaque gène présent dans la musique est classé manuellement de 0 à 5. Une fois que la musique est analysée, elle est placée topographiquement dans un ensemble d’autres musiques déjà analysées en utilisant une fonction distancielle (Prey, 2018).

Quand un auditeur choisit de lancer une musique, Pandora Internet Radio trouve rapidement les musiques comprenant les mêmes gènes que celle lancée et peut ensuite en proposer une similaire si l'auditeur a aimé la musique ou différente si ce n'est pas le cas. Pour savoir si un auditeur a aimé la musique, ils utilisent le système de "like" (pouce en l'air si c'est aimé et pouce vers le bas si ce n'est pas aimé). Grâce à ce système, Pandora Internet Radio arrive à définir quel gène est plus important qu'un autre en définissant un poids pour chaque gène (Prey, 2018).

Bref, la plateforme de streaming utilise un système de feedback continu afin d'améliorer son système de recommandation. Et pour reprendre l'exemple du DJ : Pandora Internet Radio (le DJ), décide de passer des musiques grâce à ses connaissances (les gènes déjà trouvés de son répertoire musical) à des auditeurs différents (dont les musiques déjà écoutées donnent des informations par rapport à leurs goûts de certains gènes) et modifie ses playlists de lecture en fonction de la réaction de la foule (le système de like).

2.7.2) Spotify et la "Discover Weekly"

Spotify est le leader mondial du streaming musical avec plus de 406 millions d'utilisateurs mensuels (Statista, Spotify Nombre d'utilisateurs actifs dans le monde, 2022). Il est donc logique que leur playlist phare personnalisée "Discover Weekly" dont le but est de proposer une playlist de 30 musiques chaque lundi aux utilisateurs en fonction de ce qu'ils aiment, connaisse le même succès avec 2,3 milliards d'heures d'écoute depuis 2015 (Spotify.com, 2020).

Mais comment fonctionne l'algorithme pour cette playlist ? Il faut remonter à 2014 avec l'achat par Spotify d'Echo Nest, une start-up américaine dont l'activité se concentre sur l'analyse de data pour le comprendre. Echo Nest utilise un système informatique acoustique qui analyse et classe les musiques en fonction de certains facteurs auditifs comme le tempo ou la possibilité de danser sur la musique (Prey, 2018). Le système informatique intègre donc tous ces éléments (une musique moyenne en possède 2000) pour ensuite faire des connexions avec les musiques possédant les mêmes éléments (Prey, 2018).

Mais ce n'est pas tout, la société utilise également internet et plus précisément les réseaux sociaux en faisant des analyses sémantiques sur les conversations en ligne qu'il y a quotidiennement partout dans le monde. Ainsi, les employés analysent les tweets, les blogs, les revues musicales et les différentes discussions publiques afin de créer une matrice

quantifiable en compilant les mots-clés utilisés pour ensuite créer des liens avec d'autres musiques où ces mots-clés ont également été utilisés (Prey, 2018).

Une fois qu'ils ont créé une sorte de carte du monde de la musique (en créant des sortes de mini-univers dans lesquels des musiques présentent les mêmes caractéristiques), il ne reste plus qu'à savoir où l'utilisateur se trouve sur cette carte et ainsi lui proposer les musiques qui lui conviennent. Ils font cela en capturant toutes les informations et données que l'utilisateur crée en likant, en recherchant, en passant des musiques ou en écoutant des musiques sur leur plateforme (Prey, 2018).

2.8) Critique de l'efficacité des recommandations musicales et de la raison algorithmique

Voici donc comment Spotify et Pandora Internet Radio utilisent leurs systèmes de recommandations musicales.

En plus de ces deux types de systèmes algorithmiques, d'autres plus connus sont également utilisés. Le *collaborative filtering*, dont le but est "d'établir des proximités statistiques entre les clients selon leurs paniers d'achats" reste le plus utilisé (Beuscart, Coavoux, Maillard, 2019).

On est maintenant en droit de se demander si ces systèmes de recommandations sont réellement efficaces. Mais également de poser d'éventuelles critiques sur ces systèmes de recommandations.

Mesurer l'efficacité des recommandations algorithmiques est compliqué. D'une part, il est difficile d'obtenir des données à ce propos (Beuscart, Coavoux, Maillard, 2019), et d'autre part il est difficile d'obtenir les constructions algorithmiques qui ont été créées par les plateformes, ces dernières refusant de révéler leurs constructions algorithmiques afin de garder leurs avantages compétitifs (Maaso et Hagen, 2019).

De plus, même s'il est à ce jour possible de savoir quel titre a été écouté grâce aux recommandations musicales, il est compliqué de savoir parmi ces titres lesquels n'auraient pas été écoutés s'il n'y avait pas eu de système de recommandation musicale (Sharma, Hofman, Watts, 2015). Par contre, on sait que la recommandation algorithmique d'un point de vue général affecte les ventes et qu'elle est plus influente que la recommandation humaine. Et ce même si cette dernière est perçue comme moins possessive et ayant plus d'expertise que la recommandation algorithmique (Senecal et Nantel, 2004).

Plusieurs critiques peuvent être émises par rapport aux systèmes de recommandations algorithmiques. D'une part, le fait que celles-ci agissent sans qu'on comprenne comment ces algorithmes fonctionnent. Il y aurait un manque de transparence et d'*accountability* de ces systèmes qui pourtant influencent les choix des consommateurs. (Pasquale, 2016)

Une autre critique qui peut se faire porte sur la diversité des contenus proposés par les recommandations algorithmiques et plus précisément le système de recommandation musicale basé sur le filtrage collaboratif. Comme le disent J-S. Beuscart, S. Coavoux et S. Maillard en 2019, "parce qu'il est fondé sur la cooccurrence des consommations, le filtrage collaboratif peut favoriser l'enfermement dans un portfolio de produits très similaires. Alors qu'il est pensé comme une façon de favoriser la découverte d'artistes peu connus, il peut aussi paradoxalement renforcer la place des stars" (Celma et Cano, 2008).

Enfin, il y a autant de systèmes de recommandations qu'il y a de plateformes de streaming musical (comme vu précédemment avec l'exemple de Pandora Internet Radio et de Spotify) mais il y a également un caractère pluriel et difficile à comprendre de ces algorithmes qui sont interconnectés, imbriqués dans d'autres technologies (l'exemple d'Echo Nest qui utilise les réseaux sociaux est pertinent) changeant aussi de par leurs aptitudes à proposer des musiques différentes en fonction du contexte d'écoute, du lieu où se trouve l'auditeur, des usages du consommateur. Bref, les algorithmes agissent comme des boîtes noires (Gillespie, 2016) et même si des travaux empiriques ont été faits comme celui de Prey en 2018 (travaux sur les systèmes algorithmiques de Pandora Internet Radio et Spotify), il est aujourd'hui toujours impossible de savoir comment fonctionnent réellement ces systèmes (Beuscart et al., 2019).

2.9) Perspective d'avenir

Les systèmes de recommandations algorithmiques devraient être de plus en plus présents sur les plateformes de streaming musical. Quand on voit l'évolution de ces dernières, on peut confirmer cette tendance : alors que les pages d'accueil de Spotify et Deezer présentaient par exemple, avant les albums écoutés, les titres téléchargés ou les artistes suivis, ces mêmes pages présentent désormais les playlists recommandées, les artistes et albums suggérés (Beuscart, et al., 2019). Des sites comme Youtube proposent depuis un moment une écoute totalement guidée à la fin de chaque titre s'il n'y a pas d'action de la part de l'utilisateur (Eriksson, Fleischer, Johansson, Snickars, Vonderau, 2019). Spotify espère arriver au projet

“Zero Button Music Player” et donc être un service qui en l’absence d’action de la part d’un utilisateur pourrait continuer de lire des musiques basées sur l’algorithme (Melançon, 2016). Bref, on se dirigerait vers des écoutes de plus en plus guidées où l’autonomie de l’auditeur serait réduite.

2.10) Enquête sur l’autonomie de l’auditeur

Il est pertinent de faire le point concernant cette autonomie de l’auditeur et les consommations réelles par rapport aux services de streaming musicaux.

En 2019, J-S. Beuscart, S. Coavoux, S. Maillard ont effectué une enquête sur un panel de 4000 utilisateurs en France suivis pendant 5 mois à partir d’avril 2014. Voici un des résultats les plus intéressants de cette enquête.

Trois quarts des 20 millions d’écoutes enregistrées pendant le laps de temps seraient en fait autonomes, c’est-à-dire écoutées sans la présence de système de recommandation. Même si les services de streaming musicaux veulent de plus en plus une écoute guidée, dans les faits ce n’est toujours pas le cas (Beuscart, et al., 2019).

Les résultats de cette enquête ont, comme toutes enquêtes, leurs limites. D’une part, le fait que les auditeurs enquêtés soient uniquement des utilisateurs de Deezer fait qu’il est impossible de tirer des conclusions sur les services de streaming musicaux en règle générale. D’autre part, la date de l’enquête (2014) fait qu’il n’est pas possible de savoir si les résultats de cette enquête sont toujours d’actualité, les technologies avancent rapidement et les systèmes algorithmiques ne dérogent pas à la règle. Il est donc important de réactualiser ces données (Beuscart, et al., 2019).

2.11) La question de l’hyper-choix

J’ai jusqu’ici abordé plusieurs sujets importants pour le développement de mon mémoire : le marché de la musique, le marché du streaming, les influences actuelles du streaming sur la musique, les technologies du streaming appliquées au monde de la musique et aux consommateurs. Une dernière question me semble encore importante : celle de l’hyper-choix.

Avant de définir l’hyper-choix il est bon de définir le choix. “Le choix serait les options dont disposent les individus ainsi que la capacité que ces individus ont à les comparer pour ensuite prendre une décision. Le choix est considéré par la société occidentale comme fondamental pour la liberté, l’autonomie et le bien-être des individus” (Leotti, Iyengar, Ochsner, 2010).

Le choix n'est pas réparti équitablement dans la société (Graham, 2018). Cependant, c'est quelque chose de complexe car même si nous pensons que plus de choix est toujours mieux, c'est en réalité problématique (Graham, 2018). En effet, trop de choix a tendance à submerger les consommateurs et également à sous-estimer les avantages réels du choix (Graham, 2018). Ainsi, la notion d'hyper-choix est née et pourrait être définie comme "un phénomène situé dans lequel les individus sont confrontés à des options trop nombreuses ou trop variées, qui à leur tour dépassent leur capacité à faire des choix. Le choix est donc positionné comme "hyper" en termes d'échelle et de vitesse à laquelle il s'étend et se différencie" (Larceneux, 2007).

Cette notion d'hyper-choix est désormais aussi d'actualité dans le monde de la musique, l'auditeur a accès à plusieurs millions de titres sur chaque plateforme (Annexe 1) et donc est confronté tous les jours au choix de la musique qu'il veut écouter parmi cette bibliothèque. Il est donc intéressant de savoir à quel degré cet hyper-choix est présent et comment les auditeurs consomment de la musique dans cet environnement. Le travail de Sophie Maisonneuve "L'économie de la découverte musicale à l'ère numérique : Une révolution des pratiques amateurs ?" est pour cela un bon point de départ.

Pour ce travail, elle a effectué, entre 2013 et 2015, des entretiens semi-directifs approfondis et « équipés » auprès de jeunes de 17 à 34 ans et, de sa confrontation avec les résultats d'enquêtes précédentes, plusieurs constats ont été faits. Le premier constat est que même si découvrir la musique est facilité par beaucoup de technologies (les recommandations et les algorithmes), le fait d'être aidé par leurs proches, leur savoir-faire est toujours existant (Maisonneuve, 2019). Le deuxième constat est que "dans la plupart des cas étudiés, il n'y a pas de spécificité des pratiques à l'ère numérique, mais des réinvestissements nouveaux, construction de ressources nouvelles par appropriation de nouvelles technologies, constituées en dispositifs inédits, mais adaptés à des dispositions et procédés antérieurs" (Maisonneuve, 2019). Le dernier constat est que le numérique ne serait pas perçu comme une technologie déterminante, ni un dispositif sociotechnique unique et homogène mais plutôt comme une ressource et une aide supplémentaire à la consommation de musique. Ces technologies seraient utilisées de manière différente par les auditeurs en fonction de différents facteurs comme les dispositions à découvrir des nouvelles musiques ou à consommer celles-ci et d'un facteur de confiance qu'ils ont par rapport à ces technologies (Maisonneuve, 2019). Enfin, l'auteure souligne que "si la circulation des goûts, des choix et des dispositifs de

recommandation a changé d'échelle, cette enquête montre la tension persistante, à l'échelle individuelle de l'économie domestique de la découverte, entre offre *mainstream* et recherche de singularisation, entre hyperchoix et repérage – tension résolue par les « arts de faire » de chaque consommateur, à partir de ressources communes, voire partagées, et configurées en fonction des dispositions (compétences, degrés d'engagement, expériences passées et ressources) de chacun” (Maisonneuve, 2019). Même si les plateformes de streaming musical offrent désormais des choix presque infinis, même si l'accès à la musique est plus facile et moins cher, le fait de consommer sa musique et de choisir/découvrir des artistes ou des titres demande toujours un risque (qui a certes évolué) tant l'hyper-choix que proposent ces plateformes de streaming musical nous demande plus d'énergie (Maisonneuve, 2019).

3) Questions de recherche et hypothèses

Après lecture de la littérature scientifique, plusieurs points importants sont à souligner. Il y a évidemment des conclusions déjà simples que l'on peut faire. Oui, le marché de la musique a connu d'énormes changements technologiques. Oui, le streaming musical est devenu le moyen le plus utilisé pour consommer de la musique. Oui, les technologies sont de plus en plus présentes, de plus en plus avancées dans ce marché et oui les plateformes de streaming musical veulent que l'écoute guidée devienne un standard de nos consommations d'écoutes. Cependant certaines questions restent à explorer... Même si une enquête a déjà été réalisée en 2014 par rapport au nombre d'écoutes réellement guidées sur un panel de 4000 auditeurs pendant 4 mois et que ses résultats sont que $\frac{3}{4}$ des écoutes restent autonomes (Beuscart, et al., 2019), il peut être intéressant de réactualiser cette enquête mais également de savoir comment les auditeurs sont enclins à se faire guider. Est-ce qu'il y aurait des critères socio-démographiques par rapport à l'utilisation ou non des systèmes de recommandations (l'âge, le sexe...) ou des critères liés à la consommation (des liens entre le temps de consommation, le fait de créer ses playlists, est ce que si les recommandations sont déjà en adéquation avec ce que j'écoute, je voudrai ensuite plus me faire guider) ? Enfin, une dernière question de recherche liée cette fois-ci plus au phénomène d'hyper-choix serait de savoir si les personnes qui sont sujets à l'hyper-choix sur les plateformes de streaming sont plus enclines au fait de se faire guider. Tous ces éléments pourraient être résumés dans ma question de recherche principale qui est : Est-ce que des variables peuvent impacter le fait de vouloir se faire guider par les systèmes de recommandations musicales et si oui, quelles sont-elles ?

Pour répondre à cette question, en plus de la revue de littérature qui pose un cadre théorique sur les études et recherches qui ont déjà été faites, j'ai fait une étude quantitative qui est pour moi le meilleur moyen d'englober un plus large échantillon d'auditeurs.

Cela me permet de collecter des données sur les comportements actuels de consommation. Cela me permet également de pouvoir quantifier d'une manière générale les perceptions face aux systèmes de recommandations musicales.

Enfin, cela me permet de pouvoir vérifier ou infirmer certaines hypothèses qui ont donc été créées pour cette question.

Une première hypothèse liée à des questions socio-démographiques : Est-ce que l'âge influence le fait d'être enclin à se faire guider par un système de recommandation musical. On peut donc la transformer en :

H1 : L'âge n'influence pas le fait d'être enclin à se faire guider pour écouter de la musique.

Cette hypothèse est basée sur les observations de la vraie vie et à une idée reçue qui serait que les jeunes de manière générale sont plus à l'aise avec la technologie et donc l'utilisent plus. Ainsi j'aimerais savoir si l'âge peut être un facteur déterminant sur le fait d'être enclin ou non à se faire guider pour écouter de la musique.

Les prochaines hypothèses sont liées plus à des caractères de consommation et de satisfaction liées aux plateformes de streaming musical.

La deuxième hypothèse testée est :

H2 : Le temps de consommation n'influence pas le fait d'être enclin à se faire guider pour écouter de la musique.

Par déduction logique, on pourrait penser que si des utilisateurs ont un temps de consommation élevé sur les plateformes de streaming musical, ils devraient utiliser plus de ressources proposées par ces plateformes et donc utiliser les systèmes de recommandations.

La troisième hypothèse est liée aux résultats que produisent déjà les systèmes de recommandations musicales et est un lien avec la satisfaction actuelle face à ceux-ci.

H3 : Plus les recommandations musicales actuelles sont en adéquation avec ce que l'auditeur écoute, plus il sera enclin à se faire guider par des systèmes de recommandations musicales.

Cette hypothèse me permettrait de savoir s'il existe un lien avec la performance des systèmes de recommandations musicales et leur utilisation future.

La quatrième hypothèse teste un type de consommation qui est l'écoute des playlists. Il est possible sur les plateformes de streaming d'écouter des playlists provenant de 3 sources différentes : celles créées par soi-même, celles créées par d'autres utilisateurs et celles créées par la plateforme. Etant donné que cette dernière est le produit de la plateforme, est-ce qu'il y a un lien avec le fait de consommer déjà des musiques proposées par la plateforme et le fait de vouloir consommer plus de produits de la plateforme (ici les systèmes de recommandations) ?

H4 : Consommer des playlists créées par la plateforme de streaming n'influence pas sur le fait d'être enclin à se faire guider pour écouter de la musique.

Enfin, ma dernière hypothèse est plus liée à la question d'hyper-choix. Larceneux en 2007 considère "l'hyper-choix comme un phénomène situé dans lequel les individus sont confrontés à des options trop nombreuses ou trop variées, qui à leur tour dépassent leur capacité à faire des choix." Autrement dit, les personnes confrontées à l'hyper-choix pourraient ne plus savoir quoi choisir à cause du choix très varié de musiques proposées. Dès lors, est-ce que ces auditeurs sont plus enclins à aimer les systèmes de recommandations musicales qui peuvent être perçus comme une potentielle solution à leurs problèmes d'hyper-choix ? La dernière hypothèse peut, je l'espère, amener certains éléments de réponse à cette question.

H5 : Plus les personnes sont sujettes à l'hyper-choix, plus elles aimeront être guidées par des systèmes de recommandations musicales.

Cette hypothèse permettrait de donner d'éventuels liens de corrélation entre l'utilisation des systèmes de recommandations musicales et le sentiment d'hyper-choix.

4) Méthodologie

Pour répondre à ces différentes hypothèses, j'ai donc décidé de faire une étude quantitative. Comme dit précédemment, il existait déjà un cadre théorique bien défini et reconnu, dès lors, l'étude quantitative est la meilleure méthode de recherche pour moi car elle me permet également de pouvoir combiner plusieurs cadres de réponses ce qui, au vu de mes nombreuses hypothèses, est quelque chose dont j'ai besoin. Je pourrai ensuite combiner et analyser ces différents cadres (Giordano et Jolibert, 2016).

Cette étude quantitative se porte sur un échantillon de 121 auditeurs. Etant donné ma volonté d'avoir un échantillon représentatif de la population actuelle, il n'y a pas eu de critères socio-démographiques (âge, sexe, localisation) pour faire ce choix d'échantillon. Cependant, un

critère de consommation sera utilisé, celui de savoir si oui ou non les personnes utilisent les plateformes de streaming musical. Il n'est pour cette étude pas pertinent de connaître l'avis des personnes n'utilisant pas les plateformes de streaming musical sur les systèmes de recommandations musicales.

Pour des raisons pratiques, le public cible sera recruté dans mon entourage via les réseaux sociaux (Facebook, Instagram, LinkedIn) mais également via l'entourage de mes proches via ces mêmes réseaux sociaux.

Étant donné la volonté de questionner les individus sur plusieurs aspects de la problématique du streaming musical, le questionnaire est la méthode quantitative la plus adaptée pour mon travail.

Ce questionnaire (Annexe 2) qui comprendra plusieurs questions comme :

- 1 des questions binaires fermée (oui/non) ;
- 2 des QCM fermés (sans possibilité de répondre "autres") ;
- 3 des QCM ouverts (avec possibilité de répondre "autres") ;
- 4 des QRM ouverts (avec possibilité de répondre "autres") ;
- 5 des échelles de Likert.

La variété du type de questions me permet d'englober une quantité plus importante d'informations.

Les questions binaires sont utilisées pour pouvoir ensuite effectuer des filtres dans ma base de données et pour effectuer des statistiques descriptives.

Les QCM et QRM sont utilisés pour connaître les habitudes de consommations actuelles de l'échantillon (combien de temps en moyenne d'écoute, sur quelle plateforme...) et cela permet d'une part de continuer à décrire l'échantillon lors des analyses descriptives mais également de commencer à pouvoir combiner ces variables avec d'autres informations afin d'établir ensuite des analyses statistiques sur des possibles liens de corrélations entre des facteurs.

Les échelles de Likert permettent de connaître les avis par rapport aux systèmes de recommandations musicales et sur d'autres questions liées à l'hyper-choix. L'échelle de Likert est plus intéressante à utiliser pour l'analyse car elle permet d'avoir des avis nuancés sur des questions et donc des points de vue plus précis.

Après la récolte de données du 27 juillet 2022 au jeudi 4 août 2022 permettant de récolter un total de 121 réponses, il fallait nettoyer la base de données. En effet, lors du lancement de cette enquête beaucoup de questions supplémentaires avaient été posées. Cela dans le but d'une part de définir le plus précisément l'échantillon et les consommations actuelles des répondants mais aussi de pouvoir ouvrir, s'il le fallait, la recherche et la discussion vers d'autres potentielles hypothèses. Pour nettoyer cette base de données une question filtre a été utilisée, permettant de réduire l'échantillon à 114 répondants.

Ensuite, des variables ont été supprimées afin de ne pas sortir du cadre de mes hypothèses.

La création d'une variable dépendante, "Le fait d'être enclin à se faire guider", est un élément important pour le reste de l'analyse car elle deviendra la variable la plus utilisée par la suite.

La procédure pour la création de cette variable est une analyse par composante principale qui est une méthode de synthèse de données quantitative sur trois variables (Guerrien, 2003): "Vous aimez vous faire guider pour écouter de la musique" , "Vous aimez quand une personne de votre entourage vous propose une musique" et "Vous aimez quand votre plateforme de streaming vous propose une musique". Cette méthode me permet de savoir si oui ou non on peut synthétiser les trois variables décrites au-dessus en une nouvelle variable. D'autre part, l'analyse de l'alpha de Cronbach pour mesurer la cohérence interne (ou la fiabilité) des questions posées dans mon questionnaire (Piette, 2019). Après validation des deux tests (qui sont mis dans l'analyse de résultats), la nouvelle variable dépendante "Le fait d'être enclin à se faire guider" a été créée.

La procédure pour la création de la variable en rapport avec l'hyper-choix est la même. Pour ce faire, j'ai également fait une analyse en composante principale et une étude de l'alpha de Cronbach sur 3 variables qui sont "Vous ne savez pas quoi écouter quand vous lancez votre plateforme de streaming musical", "Il y a trop de choix sur les plateformes de streaming musical" et "Vous vous sentez perdu face à la quantité de musiques proposées par les plateformes de streaming musical" qui, après des analyses positives sur les liens de dépendances deviendront la nouvelle variable indépendante " Vous êtes sujet à l'hyper-choix" (les résultats seront également mis dans l'analyse de résultats).

Enfin, la variable "Les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement" est la dernière variable gardée de cette échelle de Likert car elle est directement en lien avec la cinquième hypothèse. Cette variable sera également une variable indépendante.

Dans les questions liées aux statistiques descriptives, certaines variables ont également été supprimées afin de ne garder que celles importantes à l'étude. Un tableau récapitulatif des variables gardées se trouve en annexe (Annexe 3)

Après avoir nettoyé la base de données, plusieurs analyses statistiques ont été effectuées :

- des statistiques descriptives qui permettent d'avoir des informations sur l'échantillon et donc le calcul de la moyenne, du mode, de la fréquence par modalité et de l'effectif par modalité de plusieurs variables qui permettent de décrire l'échantillon comme l'âge, le genre, quelle plateforme est utilisée, via quel support d'écoute et à quel moment, le temps de consommation et l'utilisation ou non des systèmes de recommandations musicales ;
- des analyses plus poussées qui consistent en des tests statistiques sur les différentes variables afin de déterminer s'il existe, ou non des corrélations et donc ensuite de pouvoir rejeter ou non les hypothèses décrites précédemment.

Les tests d'analyses utilisés sont une ANOVA pour l'hypothèse 1, 2 et 3 qui me permet d'analyser les différences de variances entre certaines moyennes de l'échantillon.

C'est une analyse importante à faire car il est parfois possible qu'une différence au niveau des moyennes soit expliquée par une erreur d'échantillonnage. Dès lors, cette différence n'aurait pas de sens. L'ANOVA me permet donc de savoir si les différences entre les moyennes sont statistiquement significatives (Tibco, 2022).

Avant de procéder à l'analyse d'ANOVA pour les différentes hypothèses, il faut d'abord vérifier si elle respecte les conditions de validités qui sont au nombre de 3 (Yergeau, 2021) :

- 1 un test d'indépendance donc le fait qu'il n'y ait pas d'auto-corrélation ;
- 2 un test de normalité, il faut que les résidus suivent une loi normale (attention : l'ANOVA n'est pas très sensible aux écarts de normalité) ;
- 3 l'homogénéité des variances des résidus.

Ce n'est qu'après avoir accepté ces conditions de validité que l'ANOVA peut être effectuée. Des tests dit "ad-hoc" qui consistent en des tests de comparaison multiple par paires et qui nous permettent de savoir où se trouvent ces différences dans l'échantillon. Enfin, une analyse du R carré qui est le coefficient de détermination permettent de connaître la proportion de la variance qui est expliquée par notre variable dépendante.

Ensuite, des régressions simples ont été effectuées afin de savoir s'il existe des liens de dépendances entre les différentes variables indépendantes avec une variable principale dépendante (“Le fait d’être enclin à se faire guider”). Il s’agit cependant de liens de dépendance d’une seule variable (dépendante) vers une seule variable (indépendante) (Yergeau, 2021). Enfin, ces analyses sont interprétées et une conclusion vient terminer l’analyse de résultats. Voici un schéma permettant de voir plus clair dans mes procédures d’analyses :

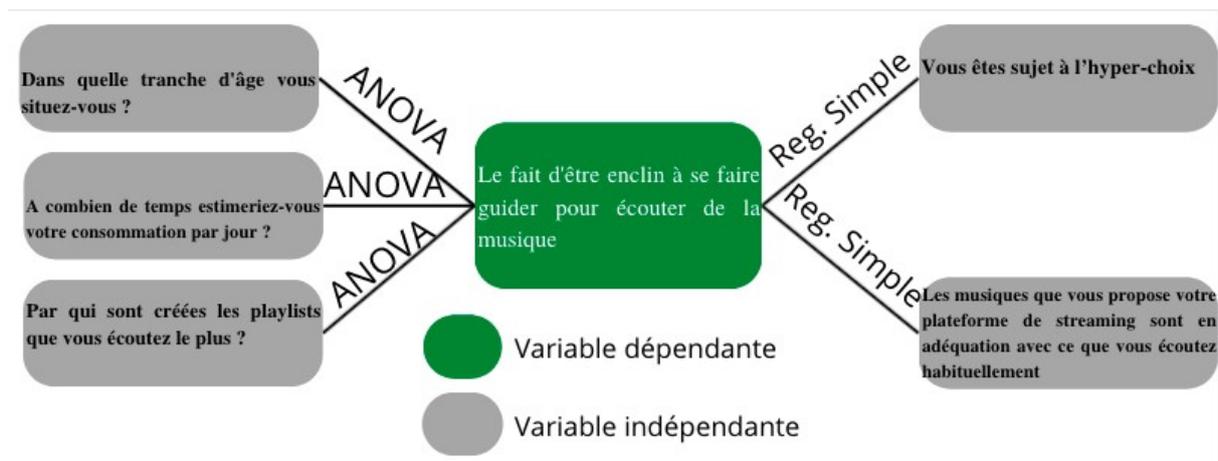


Figure 4.1- Graphe montrant les analyses faites durant ce mémoire.

5) Analyse de données

5.1) Préparation des données

Création d’une variable dépendante : “Être enclin à se faire guider”.

Pour créer cette variable, une analyse par composante principale a été effectuée. Ce qui nous a conduit à faire 3 tests : Une mesure de précision de l’échantillonnage de Kaiser-Meyer-Olkin, un test de sphéricité de Bartlett et un alpha de Cronbach.

Voici un tableau reprenant les valeurs principales :

Nom du test	Résultat	Interprétation
KMO	0,614	Les items peuvent être factorisables dès que la valeur du KMO dépasse 0,5 (Kaiser, 1974)
Test de sphéricité de Bartlett	<0,0001	Nous pouvons rejeter l’hypothèse nulle voulant

		qu'il s'agisse d'une matrice identité.
Alpha de Cronbach	0,713	L'alpha de Cronbach dépasse le seuil minimum requis de 0,70 (Nunnaly, 1978)

Tableau 5.1- Tableau reprenant les résultats d'analyses pour la création d'une variable.

Grâce à ces données, on sait maintenant qu'on peut créer une nouvelle variable dépendante. La deuxième variable créée est une variable indépendante liée à l'hyper-choix. Les procédures et analyses étant les mêmes, vous pourrez les retrouver en annexe (Annexe 4) sachez cependant que la création de la variable était également possible.

5.2) Statistiques descriptives

Pour cette section liée aux statistiques descriptives, seuls les graphiques qui sont en lien avec les hypothèses seront présentés. Ces derniers sont sans chiffres, vous pouvez retrouver les données en annexe (Annexe 5) pour plus d'informations.

Effectif (A quel genre appartenez vous ?)

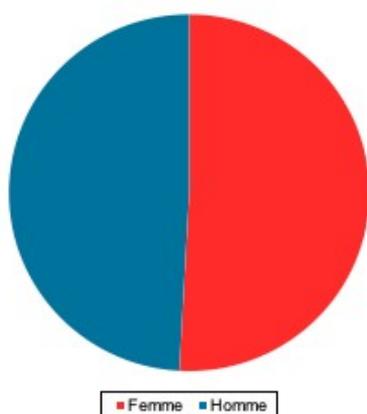


Figure 5.2- Graphique en secteur montrant la proportion d'hommes et catégories femmes des répondants.

Effectif (Dans quelle tranche d'âge vous situez-vous ?)

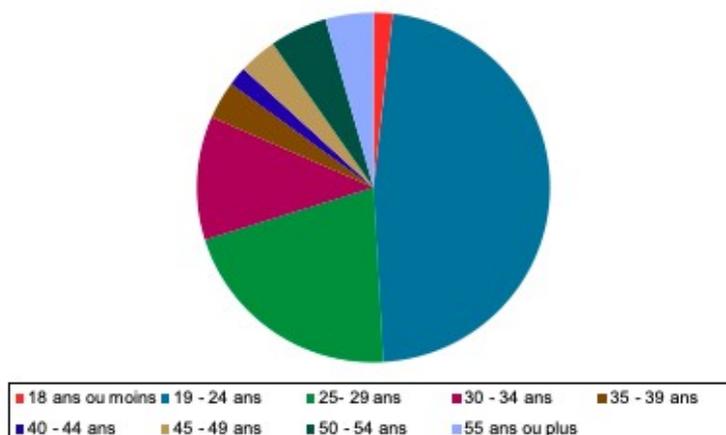


Figure 5.3 - Graphique en secteur montrant les proportions des d'âges des répondants.

On peut voir que parmi les 114 répondants (121 répondants moins 7 répondants ayant répondu non au fait d'utiliser les services de streaming musicaux), il y a une bonne proportionnalité d'hommes et de femmes. Au total il y a 58 femmes et 56 hommes qui ont

répondu au questionnaire ce qui représente presque une proportion de 50% pour chaque genre (50,88% et 49,12%). Parmi ces répondants, la plus grande tranche d'âge présente est les 19-24 ans suivis ensuite de 25-29 ans et de 30-34 ans. Ensemble, ces trois tranches d'âge représentent presque 80% de la totalité des réponses.

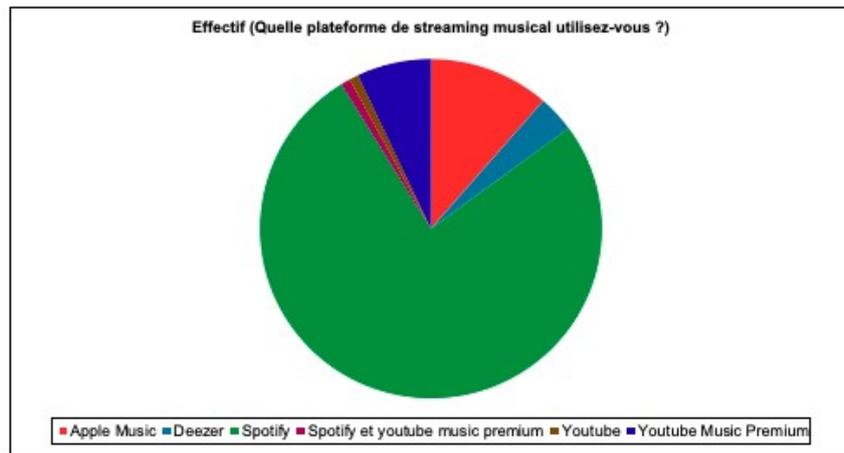


Figure 5.4 - Graphique en secteur des différentes plateformes de streaming musical utilisées par les répondants à l'enquête.

Spotify est déjà le leader du marché du streaming musical avec 31% de parts de marchés (MiDiA, 2021). Une tendance qui est également présente dans les résultats de mon enquête. En effet, Spotify est la plateforme la plus utilisée de mes répondants avec 87 répondants et donc une fréquence de 76%.

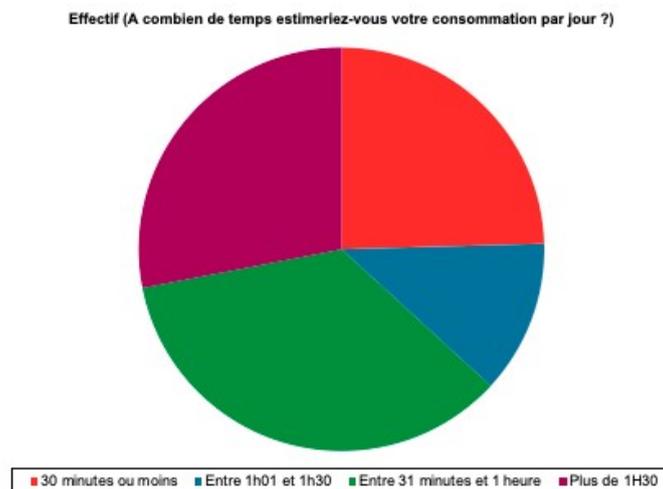


Figure 5.5 - Graphique en secteur du temps de consommation moyen journalier des répondants à l'enquête.

Les proportions au niveau du temps de consommation par jour sont les plus égalitaires de mon enquête. Le temps de consommation le plus présent est "Entre 31 minutes et 1 heure"

avec 35% du taux de réponses, suivi de “Plus de 1H30” avec 28% et “30 minutes ou moins” avec 24%.

Les box-plots suivants décrivent les trois variables quantitatives présentes dans les hypothèses.

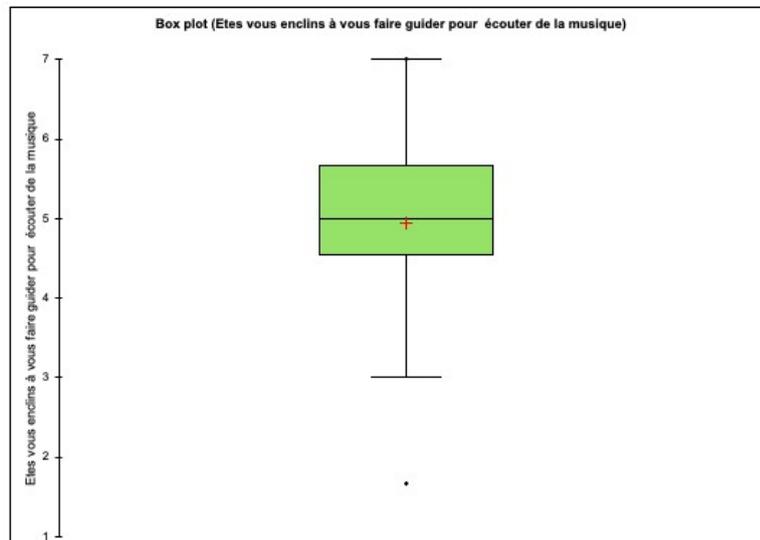


Figure 5.6 - Box Plot des réponses de l'enquête à la variable “êtes vous enclins à vous faire guider pour écouter de la musique”.

Ce box-plot permet de visualiser les tendances de la nouvelle variable créée (Êtes-vous enclin à vous faire guider pour écouter de la musique ?). La moyenne des réponses est de 4,949/7 et elle est proche de la médiane qui est de 5/7. La “boîte verte” représente 50% des réponses et est comprise entre 4,542 et 5,667 ce qui veut dire que 50% des réponses à cette nouvelle variable se trouve entre ces deux notes sur sept.

Le box-plot en annexe (Annexe 6) quant à lui représente l'adéquation des propositions de musiques des plateformes de streaming aux écoutes habituelles des répondants. La moyenne des réponses à cette variable est de 5,07/7. La médiane de 5/7 et 50% des réponses sont comprises entre 5 et 6 sur 7. Ce qui représente donc que la moitié des répondants sont soit “Plutôt d'accord” ou “D'accord” avec le fait que les plateformes de streaming musicale proposent des musiques en adéquation avec ce que les gens écoutent.

Enfin, le dernier box-plot (Annexe 7) représente la deuxième variable créée (Êtes-vous sujet à l'hyperchoix ?). On peut voir qu'en général la tendance a l'air basse. Avec une moyenne de 3,035/7, une médiane de 3/7 et 50% des réponses qui se situent entre 1,75/7 qui est donc entre

“Pas du tout d’accord” et “Pas d’accord” et 4/7 qui représente la neutralité dans mon échelle de Likert, on peut dire que la plupart des répondants ne sont pas sujets à l’hyperchoix.

5.3) Test sur les hypothèses

5.3.1) ANOVA : Hypothèse numéro 1 - La catégorie d’âge et le fait d’être enclin à se faire guider pour écouter de la musique

5.3.1.1) Prémisses à l’analyse ANOVA

Avant de procéder à l’analyse d’ANOVA pour cette hypothèse, il faut d’abord vérifier si elle respecte les conditions de validité qui sont au nombre de 3.

- 1 Pour le test d’indépendance, ayant fait moi même créé et collecté les données de l’enquête, je peux affirmer qu’il y a bien une personne différente par répondant.
- 2 Pour le test de normalité, j’ai procédé à un test de Shapiro-Wilk pour voir si la variable dont provient l’échantillon suit une loi normale. Voici les résultats :

Test de Shapiro-Wilk (Résidu) :	
W	0,971
p-value (bilaté)	0,015
alpha	0,05
Interprétation du test :	
H0 : La variable dont provient l'échantillon suit une loi Normale.	
Ha : La variable dont provient l'échantillon ne suit pas une loi Normale.	
Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.	

Table 5.7 - Tables des valeurs au test de Shapiro-Wilk.

Avec une p-value de 0,015 qui est inférieure au niveau de signification alpha = 0,05 (Yergeau, 2021), on peut rejeter l’hypothèse nulle selon laquelle notre variable dont provient l’échantillon suit une loi normale. Cependant, l’ANOVA n’étant pas très sensible aux écarts de normalité, je décide de regarder également à un graphique de P-Plot (Yergeau, 2021), me permettant de savoir graphiquement si la variable dont provient l’échantillon suit une loi normale.

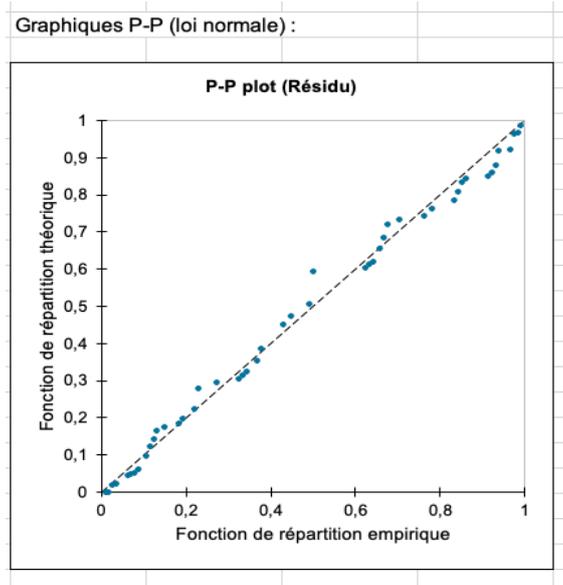


Figure 5.8 - Graphique P-Plot.

On peut voir sur ce graphique que la plupart des échantillons suivent la loi normale. Dès lors, je décide de valider ce test.

3 Enfin, pour tester l'homogénéité des variances de résidus, je procède à un test de Levene.

Test de Levene (Médiane) / Test bilatéral (Etes vous enclins à vous faire guider pour écouter de la musique) :	
F (Valeur obs)	0,811
F (Valeur criti)	2,028
DDL1	8
DDL2	105
p-value (bilaté)	0,594
alpha	0,05
Interprétation du test :	
H0 : Les variances sont identiques.	
Ha : Au moins l'une des variances est différente d'une autre.	
Etant donné que la p-value calculée est supérieure au niveau de signification seuil alpha=0,05, on ne peut pas rejeter l'hypothèse nulle H0.	

Table 5.9 - Table reprenant les valeurs au test de Levene.

Je peux voir grâce à ce test que la p-value est de 0,594 ce qui est supérieur au niveau de signification alpha de 0,05 et donc je ne peux pas rejeter l'hypothèse nulle selon laquelle les variances sont identiques. Mon test est donc valide (Yergeau, 2021). L'ANOVA respectant les 3 tests de fiabilité, je la considère comme valide et ses résultats peuvent donc être vérifiés.

5.3.1.2) Analyse ANOVA

Avant de commencer l'analyse ANOVA, il est bon de répéter quelles sont mes hypothèses.

H1 : L'âge n'influence pas le fait d'être enclin à se faire guider pour écouter de la musique.

H1 : Les moyennes sont donc semblables entre les différentes catégories d'âge.

H1a: L'âge influence le fait d'être enclin à se faire guider pour écouter de la musique.

H1a: Les moyennes sont donc différentes.

Pour tester cette hypothèse, j'ai donc procédé à une analyse ANOVA dont les résultats sont les suivants :

Analyse de la variance (Etes vous enclins à vous faire guider pour écouter de la musique) :					
Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	8	19,705	2,463	3,080	0,004
Erreur	105	83,980	0,800		
Total corrigé	113	103,684			
<i>Calculé contre le modèle $Y = \text{Moyenne}(Y)$</i>					

Table 5.10 - Table reprenant les valeurs à l'analyse de la variance.

On peut voir sur ce tableau plusieurs informations :

- D'une part, la p-value est égale à 0,004 et donc plus petite que le seuil de significativité alpha qui est de 0,05. On peut rejeter l'hypothèse nulle en prenant un risque de 0,4% de nous tromper (Yergeau, 2021).
- D'autre part on peut également analyser deux chiffres, la variabilité intra-groupe qui est ici égale 0,800 et la variabilité entre les moyennes des groupes qui est ici égale à 2,463. Ces deux chiffres nous permettent de calculer le facteur F qui est ici égal à 3,080. Plus cette valeur est haute, plus les moyennes des groupes sont éloignées les unes des autres (Yergeau, 2021).

5.3.1.3) Test de comparaisons multiples

L'ANOVA ne nous permettant que de savoir s'il existe des différences de moyenne statistiquement significative, il faut procéder à des tests "ad-hoc" pour savoir où sont situées ces différences (Yergeau, 2021).

Contraste	Différence	Différence standardisée	Valeur critique	Pr > Diff
18 ans ou moins vs 55 ans ou plus	2,433	2,842	3,284	0,005
18 ans ou moins vs 40 - 44 ans	1,667	2,425	3,284	0,017
18 ans ou moins vs 45 - 49 ans	1,583	2,361	3,284	0,020
18 ans ou moins vs 35 - 39 ans	1,417	1,765	3,284	0,080
18 ans ou moins vs 25- 29 ans	0,840	1,362	3,284	0,176
18 ans ou moins vs 50 - 54 ans	0,778	1,251	3,284	0,214
18 ans ou moins vs 19 - 24 ans	0,728	1,213	3,284	0,228
18 ans ou moins vs 30 - 34 ans	0,705	1,145	3,284	0,255
30 - 34 ans vs 55 ans ou plus	1,728	2,673	3,284	0,009
30 - 34 ans vs 40 - 44 ans	0,962	2,425	3,284	0,017
30 - 34 ans vs 45 - 49 ans	0,878	2,393	3,284	0,018
30 - 34 ans vs 35 - 39 ans	0,712	1,240	3,284	0,218
30 - 34 ans vs 25- 29 ans	0,135	0,527	3,284	0,599
30 - 34 ans vs 50 - 54 ans	0,073	0,271	3,284	0,787
30 - 34 ans vs 19 - 24 ans	0,023	0,109	3,284	0,914
19 - 24 ans vs 55 ans ou plus	1,705	2,698	3,284	0,008
19 - 24 ans vs 40 - 44 ans	0,938	2,520	3,284	0,013
19 - 24 ans vs 45 - 49 ans	0,855	2,510	3,284	0,014
19 - 24 ans vs 35 - 39 ans	0,688	1,235	3,284	0,220
19 - 24 ans vs 25- 29 ans	0,112	0,516	3,284	0,607
19 - 24 ans vs 50 - 54 ans	0,049	0,214	3,284	0,831
50 - 54 ans vs 55 ans ou plus	1,656	2,538	3,284	0,013
50 - 54 ans vs 40 - 44 ans	0,889	2,191	3,284	0,031
50 - 54 ans vs 45 - 49 ans	0,806	2,137	3,284	0,035
50 - 54 ans vs 35 - 39 ans	0,639	1,101	3,284	0,273
50 - 54 ans vs 25- 29 ans	0,062	0,231	3,284	0,818
25- 29 ans vs 55 ans ou plus	1,593	2,460	3,284	0,016
25- 29 ans vs 40 - 44 ans	0,826	2,076	3,284	0,040
25- 29 ans vs 45 - 49 ans	0,743	2,016	3,284	0,046
25- 29 ans vs 35 - 39 ans	0,576	1,003	3,284	0,318
35 - 39 ans vs 55 ans ou plus	1,017	1,230	3,284	0,221
35 - 39 ans vs 40 - 44 ans	0,250	0,385	3,284	0,701
35 - 39 ans vs 45 - 49 ans	0,167	0,264	3,284	0,792
45 - 49 ans vs 55 ans ou plus	0,850	1,217	3,284	0,227
45 - 49 ans vs 40 - 44 ans	0,083	0,175	3,284	0,862
40 - 44 ans vs 55 ans ou plus	0,767	1,073	3,284	0,286
Niveau de signification corrigé :			0,001	

Table 5.11 - Table montrant les résultats au test de Bonferroni.

Ce tableau analyse les différences entre chaque catégorie d'âge selon un test de Bonferroni. On peut voir que les plus grandes différences sont entre la catégorie d'âge "18 ou moins" et toutes les catégories d'âges à partir de 35 ans. De plus, la catégorie d'âge "55 ans ou plus" a les plus grandes différences entre toutes les catégories d'âge.

5.3.1.4) Analyse du R^2

Enfin, il est possible d'analyser à quel degré les réponses sont expliquées par le modèle. Pour cela il faut analyser le R carré qui est le coefficient de détermination. Il permet de connaître la proportion de la variance qui est expliquée par notre variable dépendante (Yergeau, 2021).

Observations	114
Somme des poids	114
DDL	105
R ²	0,190
R ² ajusté	0,128
MCE	0,800
RMCE	0,894
MAPE	16,093
DW	2,002
Cp	9,000
AIC	-16,841
SBC	7,785
PC	0,949

Table 5.12 - Table montrant les résultats de l'ANOVA dont le R².

Ici, le R carré est de 0,190. Le R qui est donc la racine de 0,190 est égale à $\sqrt{0,190} = 0,436$.

La statistique r indique donc que la différence entre les groupes se situe entre l'effet de moyenne et de grande taille (Cohen, 1988).

5.3.1.5) Conclusion ANOVA

Il existe donc des différences significatives par rapport au fait d'être enclin à se faire guider en fonction de la catégorie d'âge dans laquelle le répondant appartient. Cependant ces différences sont légères et surtout présentes pour les catégories d'âges "18 ans ou moins" et "55 ans ou plus".

5.3.2) ANOVA : Hypothèse numéro 2 - Le temps de consommation et le fait d'être enclin à se faire guider pour écouter de la musique et Hypothèse numéro 4 - Le fait de consommer les produits actuels de la plateforme de streaming et le fait d'être enclin à se faire guider pour écouter de la musique

Pour ces deux analyses ANOVA, les procédures restent les mêmes que pour l'analyse sur l'hypothèse numéro 1, elles seront placées en annexe (Annexe 8)

Voici cependant un tableau reprenant les informations importantes et leurs conclusions :

Nom du test	Résultat	Validité/ Interprétation
HYPOTHÈSE 2		
VALIDITÉ GÉNÉRALE DE L'ANALYSE	Positive	Valide

Analyse ANOVA	p-value = 0,712	Non-rejet de l'hypothèse nulle
---------------	-----------------	--------------------------------

Table 5.13 - Table montrant les résultats principaux de l'ANOVA à l'hypothèse 2.

HYPOTHÈSE 4		
VALIDITÉ GÉNÉRALE DE L'ANALYSE	Positive	Valide
Analyse ANOVA	p-value = 0,386	Non-rejet de l'hypothèse nulle

Table 5.14 - Table montrant les résultats principaux de l'ANOVA à l'hypothèse 4.

5.3.2.1) *En conclusion de ces tests d'ANOVA*

Pour conclure par rapport à ces différents tests d'ANOVA on pourrait dire qu'une seule variable influence le fait d'être enclin à se faire guider ou non, celle de la catégorie d'âge. Pour les variables liées au temps de consommation et au fait de consommer des playlists créées par la plateforme, les tests d'ANOVA ne nous permettent pas de rejeter les hypothèses nulles. Nous ne pouvons pas dire si statistiquement ces variables influencent le fait d'être enclin à se faire guider pour écouter de la musique ou non.

5.3.3) *Analyse par la régression simple : Hypothèse numéro 5, lien entre hyper-choix et le fait d'être enclin à se faire guider*

5.3.3.1) *Rappel des hypothèses*

H5 : L'hyper-choix n'influence pas le fait d'être enclin à se faire guider.

H5a : L'hyper-choix influence le fait d'être enclin à se faire guider.

5.3.3.2) *Etape 1 : Evaluation de la pertinence du modèle de régression*

Analyse de la variance (Êtes vous enclins à vous faire guider pour écouter de la musique) :					
Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	1	11,019	11,019	13,318	0,000
Erreur	112	92,665	0,827		
Total corrigé	113	103,684			
Calculé contre le modèle $Y=Moyenne(Y)$					

Table 5.15 - Table reprenant les résultats de l'analyse de la variance de la variable "Êtes vous enclins à vous faire guider pour écouter de la musique".

Ce tableau s'interprète de la même manière que celui pour l'analyse ANOVA. La valeur de F est de 13,318 et est significative car p est plus petit que l'alpha de 0,05 (Yergeau 2021).

Nous pouvons donc rejeter l'hypothèse nulle formulée plus haut et il existe donc une relation statistiquement significative entre le fait d'être sujet à l'hyper-choix et le fait d'être enclin à se faire guider pour écouter de la musique.

5.3.3.3) Étape 2 : Évaluation de l'ajustement des données au modèle de régression

"Le R^2 (coefficient de détermination) donne une idée du % de variabilité de la variable à modéliser, expliqué par la variable explicative. Plus ce coefficient est proche de 1, meilleur est le modèle" (XLStat, 2022).

Coefficients d'ajustement (Êtes vous enclins à vous faire guider pour écouter de la musique) :					
Observations	114				
Somme des p	114				
DDL	112				
R^2	0,106				
R^2 ajusté	0,098				
MCE	0,827				
RMCE	0,910				
MAPE	17,341				
DW	1,986				
Cp	2,000				
AIC	-19,621				
SBC	-14,149				
PC	0,926				

Table 5.16 - Table reprenant les résultats des coefficients d'ajustement pour la variable "Êtes vous enclins à vous faire guider pour écouter de la musique".

Ici le R^2 est de 0,106. Nous pouvons donc dire que le fait d'être sujet à l'hyper-choix peut expliquer près de 10,6% de la variation du fait d'être enclin à se faire guider pour écouter de la musique. Le reste de la variabilité est dû à des effets (autres variables explicatives) qui ne sont pas pris en compte dans cet exemple (Maalouf, 2017).

Paramètres du modèle (Êtes vous enclins à vous faire guider pour écouter de la musique) :						
Source	Valeur	Erreur standard	t	Pr > t	Borne inférieure	Borne supérieure
Constante	4,301	0,196	21,894	<0,0001	3,912	4,691
êtes vous sujet à l'hyperchoix	0,213	0,058	3,649	0,000	0,097	0,328

Table 5.17 - Table reprenant les paramètres du modèle

Ce tableau permet de faire des prévisions ou des simulations. Les valeurs importantes à regarder sont le 4,301 et le 0,213 représentant l'ordonnée à l'origine et la pente de la droite, le bêta (Yergeau, 2021). Le 0,213 indique que si le répondant est sujet à une unité en plus de la variable hyper-choix, alors il augmentera le fait d'être enclin à se faire guider de 0,213 (sur une échelle de 7 points).

De plus, voici l'équation finale de prévision :

$$\text{Intention se de se faire guider} = 4,301 + 0,213 * \text{Hyper-choix} + e$$

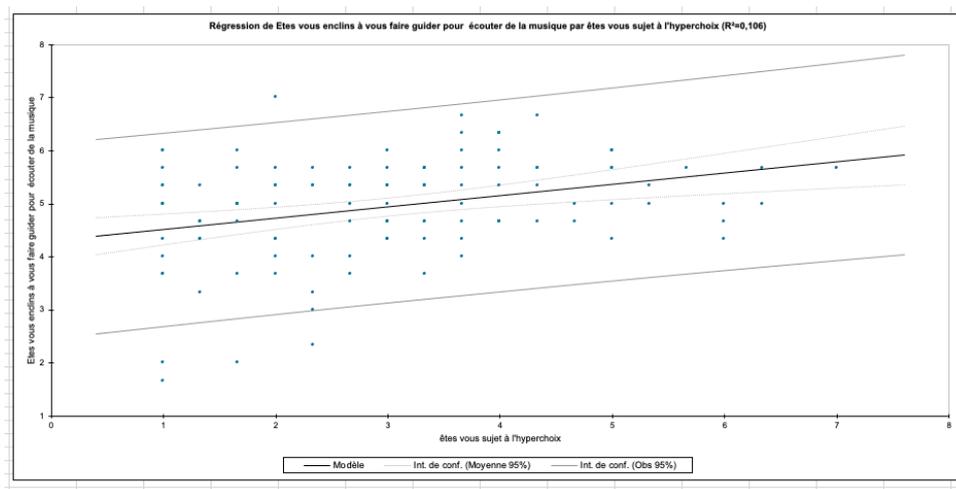


Figure 5.17 - Graphique nous montrant les tendances des réponses de la variable “Êtes vous enclins à vous faire guider pour écouter de la musique” et la variable “Êtes vous sujet à l’hyper-choix”.

Enfin ce dernier graphique permet de mieux visualiser les données au niveau des résidus. On peut voir ici qu’il n’y a pas réellement de tendance linéaire et que, même si les observations sont d’une manière générale groupées, cinq observations sont en dehors des limites de l’intervalle de confiance de 95%.

5.3.3.4) Conclusion

En conclusion, la variable de l'hyper-choix impacte faiblement et positivement l'intention de se faire guider. De plus, le fait d'être sujet à l'hyper-choix permet d'expliquer 10% de la variabilité du fait d'être enclin à se faire guider pour écouter de la musique. Pour expliquer la variabilité restante, d'autres sources de variabilité doivent donc être prises en compte dans le modèle (XLStat, 2021).

5.3.4) Analyse par la régression simple : hypothèse numéro 3, lien entre l'adéquation de l'offre des plateformes de streaming et le fait d'être enclin à se faire guider

Étant donné les procédures qui sont les mêmes que pour l'analyse par la régression simple de l'hypothèse numéro 5, seul un tableau reflétant les données les plus importantes sera présenté. L'analyse complète se trouvera en annexe (Annexe 9).

Nom du test	Résultat	Interprétation
Analyse de la variance	$p=0,000$	La p-valeur étant plus petite que l'alpha de 0,05, nous pouvons rejeter l'hypothèse nulle.
Coefficient de détermination	$R^2= 0,117$	La variable indépendante explique 11,7% de la variation du fait d'être enclin à se faire guider pour écouter de la musique
Équation de prévision	Intention se de se faire guider = $3,853 + 0,321 * \text{l'adéquation des propositions} + e$	

Table 5.18- Table reprenant les résultats de la régression linéaire pour la variable "L'adéquation de l'offre des plateformes".

5.3.4.1) Conclusion

En conclusion, la variable de l'adéquation des propositions de la plateforme de streaming impacte faiblement et positivement l'intention de se faire guider.

De plus, “le fait que les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement” permet d'expliquer 11,7% de la variabilité du fait d’être enclin à se faire guider pour écouter de la musique. Pour expliquer la variabilité restante, d'autres sources de variabilité doivent donc être prises en compte dans le modèle.

6) Conclusion

6.1) Retour sur la discussion

Il est maintenant temps de faire un point sur les choses déjà connues et celles que nous avons découvertes, apprises lors de notre enquête.

Le marché musical est en pleine expansion. Les revenus globaux augmentent et on peut mettre cela en relation avec le monde du streaming musical qui lui aussi est en expansion. Le nombre d'abonnés augmente ainsi que la part de revenus générés dans ce revenu global. Les plateformes de streaming sont désormais le moyen le plus utilisé pour écouter de la musique et les technologies liées aux plateformes de streaming comme les playlists ou les systèmes de recommandations musicales sont de plus en plus employées.

Justement, par rapport à ces systèmes de recommandations, qu'avons nous appris ?

D'une part, on comprend vaguement comment un système de recommandation pourrait fonctionner. Même si rien ne nous permet de quantifier quelles données sont les plus utilisées, comment ils sont créés et quel impact ils ont réellement dans nos habitudes de consommation (Beuscart et al., 2019). Les deux exemples de Pandora Web Radio et de la "Discover Weekly" de Spotify nous permettent de voir, premièrement, à quel point un système de recommandations peut être différent en fonction des plateformes de streaming et deuxièmement, que les recommandations algorithmiques ne sont pas simples comme on pourrait le penser avec le *collaborative filtering* ou les *trending algorithms*.

D'autre part, nous comprenons bien que l'usage des systèmes algorithmiques va être de plus en plus présent dans nos consommations de la musique mais également nos consommations de tous les jours et que c'est la volonté claire des plateformes de streaming musical (on peut prendre l'exemple du projet "Zero Button Music Player" de Spotify).

Le consommateur, lui, n'a toujours pas suivi la tendance et ne semble pas utiliser de manière proactive les systèmes de recommandations musicales. L'enquête faite par S. Beuscart en 2019 nous permettait de savoir que les consommateurs n'utilisaient pas forcément ces systèmes et que seul ¼ des auditeurs avaient recours aux recommandations musicales (Beuscart et al., 2019). Notre travail a voulu aller plus loin en essayant de connaître l'intention à se faire guider et ce qui pourrait l'influencer. Nos résultats nous permettent donc de dire plusieurs choses.

Premièrement il existe des liens entre trois variables indépendantes et la variable dépendante “être enclin à se faire guider” qui représente l’intention de se faire guider.

Quels sont ces liens :

- un lien faible entre la catégorie d’âge et le fait d’être enclin à se faire guider.
- une corrélation faible et positive entre la variable dépendante et le fait d’être sujet à l’hyper-choix.
- une corrélation faible et positive entre l’adéquation des propositions musicales des plateformes dans leurs playlists et le fait d’être enclin à être guidé.

Pour les deux autres variables indépendantes : “le temps de consommation” et “Le fait de consommer les produits actuels de la plateforme de streaming “ rien ne nous permet d’affirmer qu’il existe des liens statistiquement parlant.

Bref, ce travail n’a pas permis de trouver un facteur pertinent qui expliquerait le fait de vouloir être guidé pour écouter de la musique.

Par contre, il a permis de confirmer ou de donner des nuances à certaines théories:

Larceneux en 2007 disaient de l’hyperchoix que c’était ”un phénomène situé dans lequel les individus sont confrontés à des options trop nombreuses ou trop variées, qui à leur tour dépassent leur capacité à faire des choix. “ On peut voir avec les résultats de mon enquête qu’il est vrai que la capacité à faire des choix diminue avec la hausse du sentiment d’hyper choix. Cela viendrait donc confirmer le fait qu’une fois sujet à l’hyperchoix, une personne est moins capable de faire de choix et donc plus encline à se faire guider.

Maisonneuve en 2019 disait que “Ces technologies (les recommandations musicales) seraient utilisées de manière différente par les auditeurs en fonction de différents facteurs comme les dispositions à découvrir des nouvelles musiques ou à consommer celles-ci et d’un facteur de confiance qu’ils ont par rapport à ces technologies”. Notre enquête a, en effet, réussi à trouver des liens (même s’ils sont faibles) entre différents facteurs et le fait d’être enclin à se faire guider. Les facteurs de confiance ou de disposition à découvrir ne faisant pas partie des facteurs étudiés, nous pouvons cependant dire qu’il est vrai que ces technologies sont utilisées de manière différentes par les auditeurs en fonction d’autres facteurs comme l’adéquation de l’offre, le fait d’être sensible à l’hyperchoix ou encore la catégorie d’âge. Cela ouvre donc la discussion mais cela montre également que plusieurs facteurs sont à prendre en compte et que les auditeurs utilisent la plupart du temps les recommandations musicales pour des raisons différentes en fonction de leurs sensibilités à ces facteurs.

6.2) Limites

Il existe des limites à notre enquête. Le format de l'enquête par exemple ne permet pas de s'assurer de la bonne compréhension des questions par les répondants ni d'ailleurs du fait que les répondants aient répondu avec sérieux. Certaines questions sont difficiles à analyser, c'est le cas pour les questions ouvertes.

En outre, l'échantillon est limité et ne contient que 121 répondants qui ont une bonne proportion pour le genre des répondants mais dont 80% ont moins de 35 ans. Cela pourrait constituer un problème au niveau de la représentativité de la population.

Enfin, les méthodes d'analyses statistiques ont toutes leurs défauts et limites.

Aussi bien l'ANOVA que la régression simple ne permettent l'étude que d'une variable indépendante vers une autre variable dépendante. Cela ne nous permet pas d'analyser l'influence de plusieurs variables indépendantes vers une variable dépendante (Debois, 2022). Or, dans nos recherches nous avons vu que trois variables ont un lien statistique avec notre variable dépendante. Quand serait-il si les trois variables avaient été prises ensemble ?

L'impact des valeurs aberrantes sur les régressions linéaires peuvent également poser des problèmes (Sparrow, 2022) car elles peuvent avoir des effets énormes sur la régression.

La méthode de construction des deux nouvelles variables "être enclin à se faire guider" et "être sujet à l'hyper-choix" s'est faite grâce à une analyse par composante principale. D'une part, cette analyse a également ses limites. Par exemple, l'ACP ne prend en compte que les corrélations linéaires (Rangeon, 2022). D'autre part, nous pouvons discuter des valeurs des KMO qui sont de 0,614 et de 0,688. Certes ces valeurs nous permettent selon Kaiser en 1974 de créer sans se tromper statistiquement ces deux nouvelles variables. Cependant, ces valeurs restent faibles et ne prennent donc pas en compte toutes les nuances des variables qui composent les nouvelles créées.

Si des recherches complémentaires devaient être faites, le questionnaire pourrait être revu. Permettant un focus plus important sur les questions liées à l'hyper-choix et au niveau des questions utilisées pour les analyses. D'autres analyses pourraient d'ailleurs être effectuées comme une analyse MANOVA permettant de pouvoir exploiter plusieurs variables indépendantes à la fois.

6.3) Recommandations managériales

Au vu des résultats de mon enquête et des informations retenues avec ma revue de littérature, des recommandations managériales vont être dirigées, premièrement, vers les acteurs de la musique comme les gérants de labels ou les managers et deuxièmement vers les plateformes de streaming musical.

Par rapport aux acteurs de la musique, étant donné la part croissante et déjà importante du streaming musical dans le monde de la musique actuel, je dirais qu'il ne serait pas étonnant qu'à l'avenir les sorties musicales ne se fassent que via les plateformes. Le fait de vouloir entrer dans les playlists éditoriales est aussi quelque chose qui doit être au centre des stratégies de développements des artistes. En effet, avec un nombre d'abonnés croissants et un nombre d'auditeurs dépassant le nombre d'auditeurs des grandes radios. Il est clair qu'un artiste arrivera désormais à toucher une audience plus grande en se développant sur ces plateformes de streaming musicales.

Par rapport aux plateformes de streaming, continuer à développer les systèmes de recommandations afin que ceux-ci deviennent plus importants dans les habitudes de consommation des auditeurs.

Selon mon enquête, le public jeune a une moyenne à la variable "êtes vous enclin à vous faire guider" plus grande. Dès lors, se focaliser sur un public plus jeune pourrait être opportun. Le fait d'effectuer des stratégies différentes pour proposer des recommandations musicales en fonction de la catégorie d'âge dans laquelle se trouve l'auditeur est également pertinent.

Continuer à perfectionner les services de recommandations musicales car on peut voir que l'adéquation de l'offre de la plateforme de streaming est corrélée positivement avec le fait d'être enclin à se faire guider. Une offre de recommandation plus précise pourrait donc augmenter le taux d'utilisation des services de recommandations musicales.

Enfin, la recommandation humaine étant toujours aussi intéressante voir perçue comme ayant plus d'expertise (Senecal et Nantel, 2004) pourquoi ne pas essayer de combiner les deux en personnifiant les systèmes de recommandations. Par exemple, si vous avez écouté la chanson XX de l'artiste XY, sachez que l'artiste XY adore la musique ZZ. Cela pourrait ainsi permettre de rendre ces systèmes de recommandations plus "accessible"

7) Bibliographie

7.1) Bibliographie de la littérature

- Arnautu, O. R. (2012). Mures : Un système de recommandation de musique. *Université de Montréal*.
- Beuscart, J. S., Coavoux, S., & Maillard, S. (2019). LES ALGORITHMES DE RECOMMANDATION MUSICALE ET L'AUTONOMIE DE L'AUDITEUR Analyse des écoutes d'un panel d'utilisateurs de streaming. *La Découverte*, 213, 17-47.
- Boucher, Mike ; Bakole Baende-Mbele, Yvan. *Comment vivre de sa musique à l'ère du streaming digital ?*. Louvain School of Management, Université catholique de Louvain, 2020. Prom. : Belleflamme, Paul. <http://hdl.handle.net/2078.1/thesis:24708>
- Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and Beyond : From Production to Producers (Digital Formations)* (New éd.). Peter Lang Inc., International Academic Publishers.
- Celma, S., & Cano, P. (2008). From hits to niches ? *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition - NETFLIX '08*. <https://doi.org/10.1145/1722149.1722154>
- CHENEY-LIPPOLD, J. (2017). *We Are Data*. NYU. Press. <https://doi.org/10.2307/j.ctt1gk0941>
- Clarke, M. (2010). *Challenging Choices : Ideology, Consumerism and Policy* (First éd.). Policy Press.
- Culture is Ordinary (1958). (2014). *Raymond Williams on Culture & ; Society : Essential Writings*, 1-18. <https://doi.org/10.4135/9781473914766.n1>
- Englebert, Simon. *Le streaming musical : à qui profite-t-il ? Analyse du secteur, du business model et des enjeux vis à vis des parties prenantes*. Louvain School of Management, Université catholique de Louvain, 2018. Prom. : Percy, Thierry. <http://hdl.handle.net/2078.1/thesis:14318>

- *The Ecological Approach To Visual Perception* by Gibson, James J. (1986) Paperback (New Ed). (1986). Psychology Press.
- Gillespie, T. (2010). The politics of 'platforms'. *New Media & ; Society*, 12(3), 347-364. <https://doi.org/10.1177/1461444809342738>
- Giordano, Y., & Jolibert, A. (2016). Pourquoi je préfère la recherche quantitative/Pourquoi je préfère la recherche qualitative. *Revue internationale P.M.E. : Économie et gestion de la petite et moyenne entreprise*, 29(2), 7. <https://doi.org/10.7202/1037919ar>
- Graham, T. (2018). Technologies of choice : The shaping of choice on the World Wide Web. *Doctoral Dissertation, The University of Queensland*. <https://doi.org/10.14264/uql.2016.865>
- *IFPI Global Music Report*. (2022).
- Jacobsen, B. (2019). Bucher, T. (2018). If... then : Algorithmic power and politics. New York : Oxford University Press. 216 pp. *Communications*, 44(2), 251-253. <https://doi.org/10.1515/commun-2019-2027>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/bf02291575>
- Kaun, A. (2014). Jose van Dijck : Culture of Connectivity : A Critical History of Social Media. Oxford : Oxford University Press. 2013. *MedieKultur : Journal of media and communication research*, 30(56). <https://doi.org/10.7146/mediekultur.v30i56.16314>
- Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose : the origins and value of the need for control. *Trends in Cognitive Sciences*, 14(10), 457-463. <https://doi.org/10.1016/j.tics.2010.08.001>

- Ludovic, T. (2011). *Musique ! : DU PHONOGRAPHE AU MP3 (SCIENCES HUMAINES (A))* (AUTREMENT éd.). AUTREMENT.
- Maasø, A., & Hagen, A. N. (2019). Metrics and decision-making in music streaming. *Popular Communication*, 18(1), 18-31. <https://doi.org/10.1080/15405702.2019.1701675>
- Maisonneuve, S. (2019). L'économie de la découverte musicale à l'ère numérique. *Réseaux*, n° 213(1), 49-81. <https://doi.org/10.3917/res.213.0049>
- Mayer-Schönberger, V., & Cukier, K. (2013). Big Data : A Revolution that Will Transform how We Live, Work, and Think. *Books*, N° 52(3), 20-23. <https://doi.org/10.3917/books.052.0020>
- Melançon, L. (2016, mars). *Melançon, L. (2016, mars). La nubémorphose des pratiques de consommation musicale : le cas des plateformes de diffusion en continu.*
- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production : Theorizing the contingent cultural commodity. *New Media & ; Society*, 20(11), 4275-4292. <https://doi.org/10.1177/1461444818769694>
- Pasquale, F. (2016). *The Black Box Society : The Secret Algorithms that Control Money and Information*, by Frank Pasquale. Cambridge : Harvard University Press, 2015. 320 pp. ISBN 978-0674368279. *Business Ethics Quarterly*, 26(4), 568-571. <https://doi.org/10.1017/beq.2016.50>
- Prey, R. (2017). Nothing personal : algorithmic individuation on music streaming platforms. *Media, Culture & ; Society*, 40(7), 1086-1100. <https://doi.org/10.1177/0163443717745147>
- Prey, R., Esteve Del Valle, M., & Zwerwer, L. (2020). Platform pop : disentangling Spotify's intermediary role in the music industry. *Information, Communication & ; Society*, 25(1), 74-92. <https://doi.org/10.1080/1369118x.2020.1761859>

- Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, 80(2), 159-169. <https://doi.org/10.1016/j.jretai.2004.04.001>
- Sharma, A., Hofman, J. M., & Watts, D. J. (2015). Estimating the Causal Impact of Recommendation Systems from Observational Data. *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. <https://doi.org/10.1145/2764468.2764488>
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4-5), 395-412. <https://doi.org/10.1177/1367549415577392>
- Werner, A. (2020). Organizing music, organizing gender : algorithmic culture and Spotify recommendations. *Popular Communication*, 18(1), 78-90. <https://doi.org/10.1080/15405702.2020.1715980>

7.2) Bibliographie des sites internet consultés

- A. (2022, 21 janvier). Le streaming musical fait place à de plus en plus de concurrents. RTBF. Consulté le 25 août 2022, à l'adresse <https://www.rtf.be/article/le-streaming-musical-fait-place-a-de-plus-en-plus-de-concurrents-10919379>
- Addinsoft. (s. d.). Régression linéaire simple dans Excel | XLSTAT Centre d'aide. Consulté le 8 août 2022, à l'adresse <https://help.xlstat.com/fr/6705-regression-lineaire-simple-dans-excel>
- Debois, S. (2022, 14 juillet). Les 9 avantages et inconvénients des questionnaires. Pointerpro. Consulté le 13 août 2022, à l'adresse <https://pointerpro.com/fr/blog/les-9-avantages-et-inconvenients-des-questionnaires/>

- Guerrien, M. (2003, 31 juillet). L'intérêt de l'analyse en composantes principales (ACP) pour la rec. . . Journal Open Editions. Consulté le 6 août 2022, à l'adresse <https://journals.openedition.org/cal/7364>
- Husson, G. (2022, août 10). Spotify vs Deezer vs Apple Music. . . : quel service de streaming de musique choisir ? Frandroid. Consulté le 22 août 2022, à l'adresse https://www.frandroid.com/android/288776_streaming-musical-offre-choisir
- Loïs, H. (2022, 9 mai). Les DROITS en MUSIQUE : comment ça marche ? Droits éditoriaux, droits phonographiques et droits de synchronisation. Groover Blog. <https://blog.groover.co/conseil/les-droits-en-musique/#3.%20Le%20droit%20de%20synchronisation>
- Maalouf, J. P. (2018a, mai 18). Introduction aux statistiques descriptives illustrée avec XLSTAT - Webinars XLSTAT. YouTube. Consulté le 7 août 2022, à l'adresse <https://www.youtube.com/watch?v=o-arbQtFUCA>
- Maalouf, J. P. (2018b, juin 15). Introduction à l'analyse de Variance ou ANOVA illustrée avec XLSTAT - Webinars XLSTAT. YouTube. Consulté le 8 août 2022, à l'adresse <https://www.youtube.com/watch?v=4Aa3ek3mmz4&t=1649s>
- Mulligan, M. (2022, 18 janvier). Music subscriber market shares Q2 2021. MiDIA. Consulté le 20 août 2022, à l'adresse <https://www.midiaresearch.com/blog/music-subscriber-market-shares-q2-2021>
- Piette, P. (2020, 17 février). Un coefficient de Cronbach ça sert à quoi ? Actukiné. Consulté le 6 août 2022, à l'adresse <https://www.actukine.com/un-coefficient-de-cronbach-ca-sert-a-quoi/>
- Qu'est-ce que l'analyse de la variance (ANOVA) ? (s. d.). TIBCO Software. Consulté le 7 août 2022, à l'adresse <https://www.tibco.com/fr/reference-center/what-is-analysis-of-variance-anova>

- Rangeon, N. (2022, 9 mai). Soyez attentif aux spécificités de l'ACP. OpenClassrooms. Consulté le 13 août 2022, à l'adresse <https://openclassrooms.com/fr/courses/4525281-realisez-une-analyse-exploratoire-de-donnees/5280463-soyez-attentif-aux-specificites-de-lacp>
- Sparrow, J. (2022, 6 juillet). ML – Avantages et inconvénients de la régression linéaire – Acervo Lima. Acervo Lima. Consulté le 13 août 2022, à l'adresse <https://fr.acervolima.com/ml-avantages-et-inconvenients-de-la-regression-lineaire/>
- Spotify Users Have Spent Over 2.3 Billion Hours Streaming Discover Weekly Playlists Since 2015. (2022, 20 juillet). Spotify. Consulté le 20 juillet 2022, à l'adresse <https://newsroom.spotify.com/2020-07-09/spotify-users-have-spent-over-2-3-billion-hours-streaming-discover-weekly-playlists-since-2015/>
- Statista. (2022, 14 février). Spotify : nombre d' ; utilisateurs actifs dans le monde 2015–2021. Consulté le 20 juillet 2022, à l'adresse <https://fr.statista.com/statistiques/574665/spotify-nombre-d-utilisateurs-actifs-dans-le-monde/>
- Yergeau, E. (s. d.). SPSS Espace Web. SPSS Espace Web. Consulté le 7 août 2022, à l'adresse <https://spss.espaceweb.usherbrooke.ca/analyse-de-variance/>

8) Annexes

Annexe 1 - Nombre de titres par plateformes (Frandroid, 2022)

Plateforme	Taille du catalogue
Amazon Music	Plus de 90 millions de titres
Apple Music	Plus de 90 millions de titres
Deezer	Plus de 90 millions de titres
Qobuz	Plus de 80 millions de titres
Spotify	Plus de 80 millions de titres
Tidal	Plus de 90 millions de titres
YouTube Music	Aucun chiffre donné

Annexe 2 - Table reprenant les questions de l'enquête, leurs types et les réponses possibles.

Question	Type de questions	Réponses possibles
A quel genre appartenez-vous ?	QCM ouvert	Homme / Femme / Autres (à spécifier)
Dans quelle tranche d'âge vous situez-vous ?	QCM Fermé	18 ans ou moins / 19 - 24 ans / 25-29 ans / 30 - 34 ans / 35 - 39 ans / 40 - 44 ans / 45 - 49 ans / 50 - 54 ans / 55 ans ou plus
Utilisez vous des plateformes de streaming musical ?	QCM Fermé	Oui / Non

Quelle plateforme de streaming musical utilisez-vous ?	QCM ouvert	Spotify / Apple Music / Deezer / Youtube Music Premium / Amazone Music Unlimited / Qobuz / Tidal / Autres (à spécifier)
Quelle offre avez-vous sur votre plateforme de streaming musical ?	QCM fermé	Payante / Gratuite
Sur quel support écoutez-vous de la musique via votre plateforme de streaming musical ?	QRM ouvert	Via une tablette / Via un smartphone / Via un PC / Via une application dans une voiture / Via une smart TV / Via une console de jeu / Autres (à spécifier)
A quelle fréquence écoutez-vous de la musique sur une plateforme de streaming ?	QCM fermé	Une fois par jour / Plusieurs fois par jour / Une fois par semaine / Plusieurs fois par semaine / Une fois par mois /Plusieurs fois par mois
A combien de temps estimeriez-vous votre consommation par jour ?	QCM ouvert	30 minutes ou moins / Entre 31 minutes et 1 heure / Entre 1h01 et 1h30/ Plus de 1H30 / Autres (à spécifier)
A quel moment écoutez-vous le plus souvent de la musique via votre plateforme de streaming musical ?	QCM ouvert	Au matin / Pas de moment précis / En soirée / L'après-midi / Pendant la nuit / Pendant le temps de midi / Autres (à spécifier)
Dans quel cadre écoutez-vous de la musique	QRM ouvert	Lors de déplacement (voiture, bus, avion, train, marche,...), / En faisant du sport / En travaillant / Pas de cadre spécial / Lors d'une activité sociale (fête, dîner entre ami(e)s, boire un verre,...) / Pour dormir / En travaillant / Autres (à spécifier)
Quel style de musique écoutez-vous le plus sur ces plateformes de streaming ?	QCM ouvert	Rap / Rock / Pop / Hip-Hop / Jazz / Classique/ Folk / Metal / RnB / Blues / Country / Reggae / Electro / Funk / Disco / Musique de films séries/ Raï / K-Pop / Latine / Autre (à spécifier)
Par quel procédé consommez-vous de la musique sur les plateformes de streaming musical ?	QRM ouvert	Par recherche de titres / Par rechercher d'artistes / Par recherche d'albums / Via des playlists / Écoute aléatoire / Via vos téléchargements récents/ Autre (à spécifier)
Écoutez vous des playlists	QCM fermé	Oui / Non

Vous arrive t'il de créer vos playlists ?	QCM fermé	Oui / Non
Si oui, quel critère est le plus important pour vous quand vous créez vos playlists ?	QCM Ouvert	Mes connaissances musicales / Les recommandations de la plateforme de streaming / Par style de musique / Par période de sortie de musiques / Sur base de recommandations de votre entourage / Autres (à spécifier)
Par qui sont créées les playlists que vous écoutez le plus	QCM fermé	Par vous-même / Par d'autres utilisateurs / Par la plateforme
Pour quelle raison écoutez-vous des musiques via des playlists créées par un tiers (plateforme ou autre utilisateur) ?	QCM Ouvert	Par facilité / Parce que ma plateforme de streaming m'en fait la publicité / Les musiques proposées sont celles qui me plaisent/ Pour découvrir des nouvelles musiques / Autre (à spécifier)
Vous utilisez les systèmes de recommandations musicales	QCM Fermé	Oui / Non
Selon vous, quel est le meilleur cadre pour avoir une écoute "guidée"	QCM ouvert	Lors de déplacement (voiture, bus, avion, train, marche,...), / En faisant du sport / En travaillant / Pas de cadre spécial / Lors d'une activité sociale (fête, dîner entre ami(e)s, boire un verre,...) / Pour dormir / En travaillant / Autres (à spécifier)
Vous utilisez les systèmes de recommandations musicales :	QCM ouvert	Pour découvrir des nouveaux artistes ou de nouvelles musiques / Par facilité / Par manque d'idée / Les musiques proposées sont celles qui me plaisent
Vous savez comment les systèmes de recommandations fonctionnent	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalement d'accord
Vous trouvez que les systèmes de recommandations sont intrusifs	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalement d'accord
Vous aimez avoir le choix des musiques que vous	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4

écoutez		= Neutre (ni d'accord, ni pas d'accord) / 5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Cela ne vous dérange pas d'avoir une écoute complètement guidée	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous aimez vous faire guider pour écouter de la musique	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous aimez quand un membre de votre entourage vous propose une musique	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous aimez quand votre plateforme de streaming vous propose une musique	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Les systèmes de recommandations vous ont aidé à découvrir des musiques/ artistes	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous ne savez pas quoi écouter quand vous lancez votre plateforme de streaming musical	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Il y a trop de choix sur les plateformes de streaming musical	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous vous sentez perdu face à la quantité de	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4

musiques proposées par les plateformes de streaming musical		= Neutre (ni d'accord, ni pas d'accord) / 5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Les plateformes de streaming ont remplacé les autres supports d'écoute (CD, radio, vinyle, cassette,...)	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Je ne sais plus me passer de mes plateformes de streaming musical	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
J'ai du plaisir à écouter des albums en entier	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Le format playlist a remplacé le format album	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Vous réécoutez souvent les mêmes musiques	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord
Les plateformes de streaming musical ont changé vos habitudes d'écoute	Likert	1 = Pas du tout d'accord / 2= Pas d'accord / 3 = Plutôt pas d'accord/ 4 = Neutre (ni d'accord, ni pas d'accord) /5 = Plutôt d'accord/ 6 = D'accord / 7= Totalemt d'accord

Annexe 3 : Noms des variables gardées, type et l'utilisation qui en sera faite.

<u>Nom de la variable</u>	<u>Type de questions</u>	<u>Utilisation pour :</u>
A quel genre appartenez-vous ?	Binaire	Statistiques

		descriptives
Dans quelle tranche d'âge vous situez-vous ?	QCM fermé	Statistiques descriptives Analyse statistique
Utilisez-vous des services de streaming musicale ?	Binaire	Filtre
Quelle plateforme de streaming musical utilisez-vous ?	QCM ouvert	Statistiques descriptives
Quelle offre avez-vous sur votre plateforme de streaming musical ?	Binaire	Statistiques descriptives
Sur quel support écoutez-vous de la musique via votre plateforme de streaming musical ?	QRM ouvert	Statistiques descriptives
A combien de temps estimez-vous votre consommation par jour ?	QCM fermé	Statistiques descriptives Analyse statistique
Écoutez-vous des playlists ?	Binaire	Filtre
Par qui sont créées les playlists que vous écoutez le plus ?	QCM fermé	Analyse statistique
Vous aimez vous faire guider pour écouter de la musique	Échelle de Likert	Création variable dépendante
Vous aimez quand un membre de votre entourage vous propose une musique	Échelle de Likert	Création variable dépendante
Vous aimez quand votre plateforme de streaming vous propose une musique	Échelle de Likert	Création variable dépendante

Les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement	Échelle de Likert	Analyse statistique
Vous ne savez pas quoi écouter quand vous lancez votre plateforme de streaming musical	Échelle de Likert	Création variable indépendante
Il y a trop de choix sur les plateformes de streaming musical	Échelle de Likert	Création variable indépendante
Vous vous sentez perdu face à la quantité de musiques proposées par les plateformes de streaming musical	Échelle de Likert	Création variable indépendante

Nom du test	Résultat	Interprétation
KMO	0,688	Les items peuvent être factorisables dès que la valeur du KMO dépasse 0,5 (Kaiser, 1974)
Test de sphéricité de Bartlett	<0,0001	Nous pouvons rejeter l'hypothèse nulle voulant qu'il s'agisse d'une matrice identité.
Alpha de Cronbach	0,846	L'alpha de Cronbach dépasse le seuil minimum requis de 0,70 (Nunnaly, 1978)

Annexe 4 - Procédure excel pour la création de deux nouvelles variables "Être enclin à se faire guider pour écouter de la musique" et "Êtes vous sujet à l'hyper-choix"

Filtrer les facteurs Nombre maximum = 5
 Normalisation : (n)
 Type de biplot : Biplot de distance / Coefficient = Automatique



Matrice de corrélation (Pearson (n))

Matrice de corrélation (Pearson (n)) :

Variables	Vous ne savez pas	Il y a trop de choix sur les	Vous vous sentez perdu
Vous ne savez pas quoi écouter quand vous lancez votre plateforme de streaming musical	1	0,566	0,591
Il y a trop de choix sur les plateformes de streaming musical	0,566	1	0,783
Vous vous sentez perdu face à la quantité de musiques proposées par les plateformes de streaming musical	0,591	0,783	1

Test de sphéricité de Bartlett :

Kh ² (Valeur observée)	158,250
Kh ² (Valeur critique)	7,815
DDL	3
p-value (bilatérale)	<0,0001
alpha	0,05

Interprétation du test :

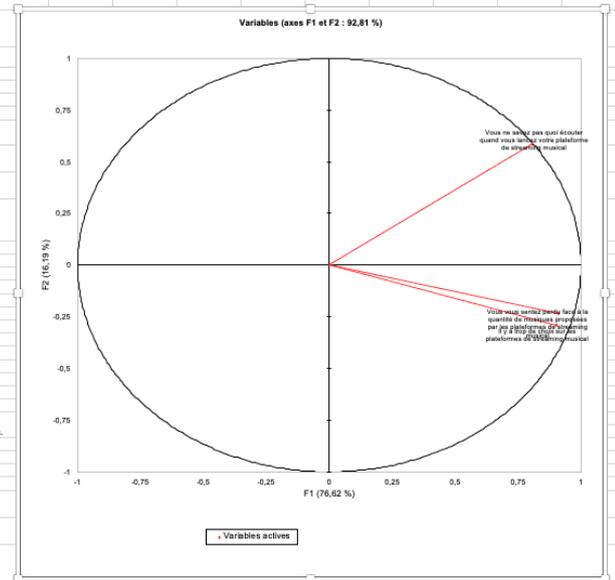
H0 : Il n'y a pas de corrélation significativement différente de 0 entre les variables.

Ha : Au moins l'une des corrélations entre les variables est significativement différente de 0.

Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.

Mesure de précision de l'échantillonnage de Kaiser-Meyer-Olkin :

Vous ne savez pas quoi écouter quand vous lancez votre plateforme de streaming musical	0,843
Il y a trop de choix sur les plateformes de streaming musical	0,652
Vous vous sentez perdu face à la quantité de musiques proposées par les plateformes de streaming musical	0,641
KMO	0,688



Matrice de corrélation (Pearson (n))

Matrice de corrélation (Pearson (n)) :

Variables	Vous aimez vous faire	Vous aimez quand un	Vous aimez quand votre
Vous aimez vous faire guider pour écouter de la musique	1	0,307	0,532
Vous aimez quand un membre de votre entourage vous propose une musique	0,307	1	0,522
Vous aimez quand votre plateforme de streaming vous propose une musique	0,532	0,522	1

Test de sphéricité de Bartlett :

Kh ² (Valeur observée)	72,491
Kh ² (Valeur critique)	7,815
DDL	3
p-value (bilatérale)	<0,0001
alpha	0,05

Interprétation du test :

H0 : Il n'y a pas de corrélation significativement différente de 0 entre les variables.

Ha : Au moins l'une des corrélations entre les variables est significativement différente de 0.

Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.

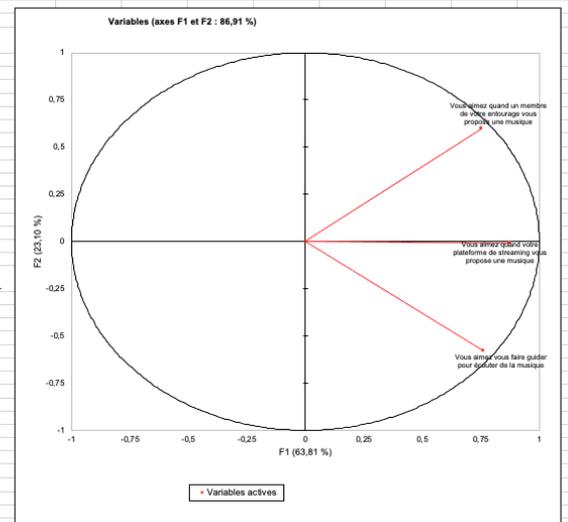
Mesure de précision de l'échantillonnage de Kaiser-Meyer-Olkin :

Vous aimez vous faire guider pour écouter de la musique	0,641
Vous aimez quand un membre de votre entourage vous propose une musique	0,647
Vous aimez quand votre plateforme de streaming vous propose une musique	0,577
KMO	0,614

Analyse en Composantes Principales :

Valeurs propres :

	F1	F2	F3
Valeur propre	1,914	0,693	0,393
Variabilité (%)	63,810	23,101	13,089
% cumulé	63,810	86,911	100,000



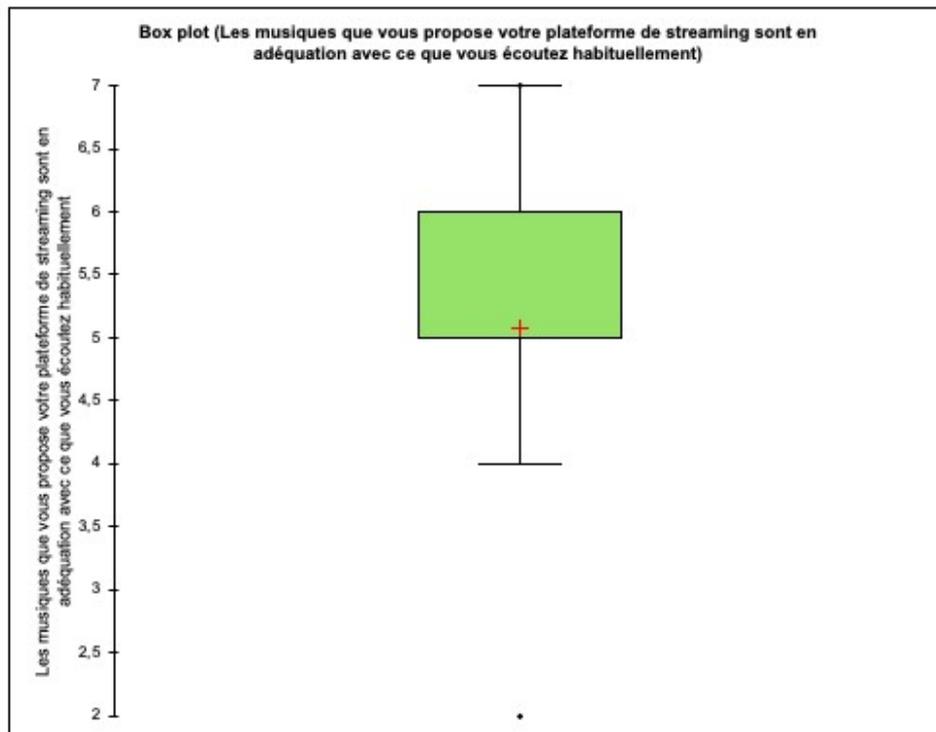
Annexe 5 - Tables reprenant les valeurs des graphiques des statistiques descriptives

Statistiques descriptives (Données qualitatives) :									
Variable\Statistique	Nb. d'observations	Somme des poids	Nb. de modalités	Mode	Mode (effectif)	Modalités	Effectif par modalité	Fréquence par modalité (%)	Borne sup. des proportions (95%)
A quel genre appartenez-vous ?	114	114	2	Femme	58	Femme	58,000	50,877	0,601
						Homme	56,000	49,123	0,583
Statistiques descriptives (Données qualitatives) :									
Variable\Statistique	Nb. d'observations	Somme des poids	Nb. de modalités	Mode	Mode (effectif)	Modalités	Effectif par modalité	Fréquence par modalité (%)	Borne sup. des proportions (95%)
A combien de temps estimeriez-vous votre consommation par jour ?	114	114	4	es et 1 heure	40	Jtes ou moins	28,000	24,561	0,325
						Entre 1h01 et 1h30	14,000	12,281	0,183
						Entre 31 minutes et 1 heure	40,000	35,088	0,438
						Plus de 1H30	32,000	28,070	0,363

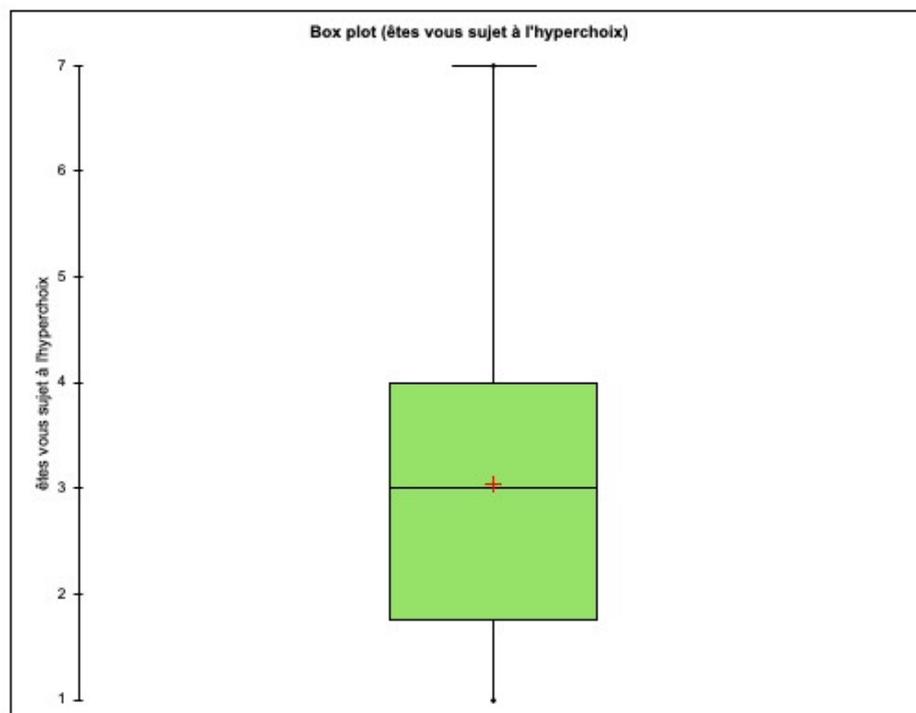
Statistiques descriptives (Données qualitatives) :									
Variable\Statistique	Nb. d'observations	Somme des poids	Nb. de modalités	Mode	Mode (effectif)	Modalités	Effectif par modalité	Fréquence par modalité (%)	Borne sup. des proportions (95%)
Dans quelle tranche d'âge vous situez-vous ?	114	114	9	19 - 24 ans	54 ans ou moins		2,000	1,754	0,042
						19 - 24 ans	54,000	47,368	0,565
						25 - 29 ans	24,000	21,053	0,285
						30 - 34 ans	13,000	11,404	0,172
						35 - 39 ans	4,000	3,509	0,069
						40 - 44 ans	2,000	1,754	0,042
						45 - 49 ans	4,000	3,509	0,069
						50 - 54 ans	6,000	5,263	0,094
						55 ans ou plus	5,000	4,386	0,081

Statistiques descriptives (Données qualitatives) :									
Variable\Statistique	Nb. d'observations	Somme des poids	Nb. de modalités	Mode	Mode (effectif)	Modalités	Effectif par modalité	Fréquence par modalité (%)	Borne sup. des proportions (95%)
Quelle plateforme de streaming musical utilisez-vous ?	114	114	6	Spotify	87	Apple Music	13,000	11,404	0,172
						Deezer	4,000	3,509	0,069
						Spotify	87,000	76,316	0,841
						Spotify et youtube music premium	1,000	0,877	0,026
						Youtube	1,000	0,877	0,026
						Youtube Music Premium	8,000	7,018	0,117

Annexe 6 - Box Plot de la variable “Les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement.”



Annexe 7 - Box Plot de la variable “Êtes vous sujet à l’hyperchoix”



Annexe 8 - Analyse ANOVA de l’hypothèse 2 et 3.

ANOVA : Hypothèse numéro 2 - Le temps de consommation et le fait d'être enclin à se faire guider pour écouter de la musique

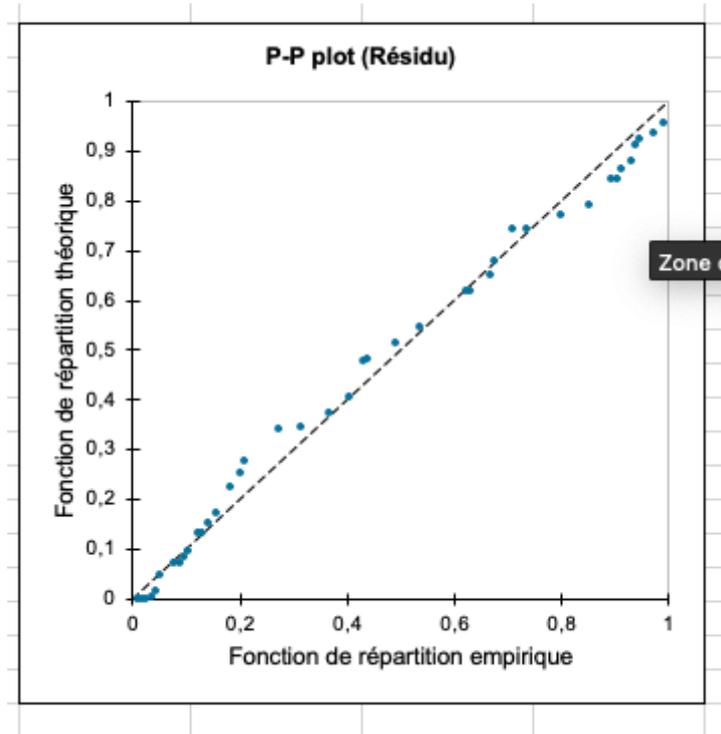
Prémises à l'analyse ANOVA

Avant de procéder à l'analyse d'ANOVA pour cette hypothèse, il faut d'abord vérifier si elle respecterait les conditions de validités qui sont au nombre de 3 :

- 1 Pour le test d'indépendance, ayant fait moi même créer et collecter les données de l'enquête, je peux affirmer qu'il y a bien une personne différente par répondant.
- 2 Pour le test de normalité, j'ai procédé à un test de Shapiro-Wilk pour voir si la variable dont provient l'échantillon suit une loi normale. Voici les résultats :

Test de Shapiro-Wilk (Résidu) :	
W	0,944
p-value (bilaté)	0,000
alpha	0,05
Interprétation du test :	
H0 : La variable dont provient l'échantillon suit une loi Normale.	
Ha : La variable dont provient l'échantillon ne suit pas une loi Normale.	
Etant donné que la p-value calculée est inférieure au niveau de signification $\alpha=0,05$, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.	

Avec une p-value de 0,000 qui est inférieure au niveau de signification $\alpha = 0,05$ (Yergeau, 2021), on peut rejeter l'hypothèse nulle selon laquelle notre variable dont provient l'échantillon suit une loi normale. Cependant, l'ANOVA n'étant pas très sensible aux écarts de normalité (Yergeau, 2021), je décide de regarder également à un graphique de P-Plot, me permettant de savoir graphiquement si la variable dont provient l'échantillon suit une loi normale.



On peut voir sur ce graphique que la plupart des échantillons suivent la loi normale. Dès lors, je décide de valider ce test.

3 Enfin, pour tester l'homogénéité des variances de résidus, je procède à un test de Levene.

Test de Levene (Médiane) / Test bilatéral (Etes vous enclins à vous faire guider pour écouter de la musique) :	
F (Valeur obs)	0,590
F (Valeur critik)	2,687
DDL1	3
DDL2	110
p-value (bilaté)	0,623
alpha	0,05
Interprétation du test :	
H0 : Les variances sont identiques.	
Ha : Au moins l'une des variances est différente d'une autre.	
Etant donné que la p-value calculée est supérieure au niveau de signification seuil alpha=0,05, on ne peut pas rejeter l'hypothèse nulle H0.	

Je peux voir grâce à ce test que la p-value est de 0,623 ce qui est supérieur au niveau de signification alpha de 0,05 et donc je ne peux pas rejeter l'hypothèse nulle selon laquelle les variances sont identiques. Mon test est donc valide (Yergeau, 2021).

L'ANOVA respectant les 3 tests de fiabilité, je la considère comme valide et ses résultats peuvent donc être vérifiés.

Analyse ANOVA

Avant de commencer l'analyse ANOVA, il est bon de répéter quelles sont mes hypothèses.

Avant de commencer l'analyse ANOVA, il est bon de répéter quelles sont mes hypothèses.

H0 : Le temps de consommation n'influence pas sur le fait d'être enclin à se faire guider pour écouter de la musique.

H0 : Les moyennes sont donc semblables entre les différentes catégories de temps de consommation.

Ha: Le temps de consommation influence le fait d'être enclin à se faire guider pour écouter de la musique

Ha: Les moyennes sont donc différentes entre les différentes catégories de temps de consommation.

Analyse de la variance (Etes vous enclins à vous faire guider pour écouter de la musique) :					
Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	3	1,280	0,427	0,458	0,712
Erreur	110	102,404	0,931		
Total corrigé	113	103,684			
Calculé contre le modèle $Y = \text{Moyenne}(Y)$					

On peut voir sur ce tableau plusieurs informations.

- D'une part, la p-value est égale à 0,458 et donc plus grande que le seuil de significativité alpha qui est de 0,05. Donc on ne peut pas rejeter l'hypothèse nulle (Yergeau, 2021).
- D'autre part on peut également analyser trois chiffres, la variabilité intra-groupe qui est ici égale 0,931 et la variabilité entre les moyennes des groupes qui est ici égale à 0,427. Ces deux chiffres nous permettent de calculer le facteur F qui est ici égal à 0,458. Plus cette valeur est haute, plus les moyennes des groupes sont éloignées les unes des autres (Yergeau, 2021). Ici, la valeur est plus basse que 1, les moyennes des groupes ne sont donc pas éloignées les unes des autres.

Analyse du R^2

Enfin, il est possible d'analyser à quelle degré nos réponses sont expliquées par notre modèle. Pour cela il faut analyser le R carré qui est le coefficient de détermination. Il permet de connaître la proportion de la variance qui est expliquée par notre variable dépendante (Yergeau, 2021).

Coefficients d'ajustement (Etes vous enclins à vous faire guider pour écouter de la musique) :	
Observations	114
Somme des p	114
DDL	110
R ²	0,012
R ² ajusté	-0,015
MCE	0,931
RMCE	0,965
MAPE	17,877
DW	2,045
Cp	4,000
AIC	-4,229
SBC	6,716
PC	1,059

Ici, le R carré est de 0,012. Le R qui est donc la racine de 0,012 est égale à $\sqrt{0,012} = 0,109$. La statistique r indique donc que la différence entre les groupes se situe entre l'effet de petite et de moyenne taille (Cohen, 1988).

ANOVA : Hypothèse numéro 3 - Le fait de consommer les produits actuels de la plateforme de streaming et le fait d'être enclin à se faire guider pour écouter de la musique

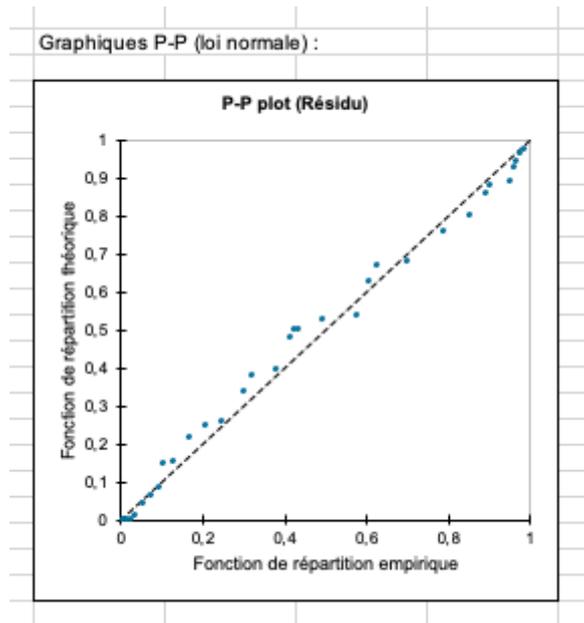
Prémises à l'analyse ANOVA

Avant de procéder à l'analyse d'ANOVA pour cette hypothèse, il faut d'abord vérifier si elle respecterait les conditions de validités qui sont au nombre de 3 :

- 1 Pour le test d'indépendance, ayant fait moi même créer et collecter les données de l'enquête, je peux affirmer qu'il y a bien une personne différente par répondant.
- 2 Pour le test de normalité, j'ai procédé à un test de Shapiro-Wilk pour voir si la variable dont provient l'échantillon suit une loi normale. Voici les résultats :

Test de Shapiro-Wilk (Résidu) :	
W	0,943
p-value (bilaté)	0,000
alpha	0,05
Interprétation du test :	
H0 : La variable dont provient l'échantillon suit une loi Normale.	
Ha : La variable dont provient l'échantillon ne suit pas une loi Normale.	
Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.	

Avec une p-value de 0,000 qui est inférieure au niveau de signification alpha = 0,05 (Yergeau, 2021), on devrait rejeter l'hypothèse nulle selon laquelle notre variable dont provient l'échantillon suit une loi normale. Cependant, l'ANOVA n'étant pas très sensible aux écarts de normalité, je décide de regarder également à un graphique de P-Plot (Yergeau, 2021), me permettant de savoir graphiquement si la variable dont provient l'échantillon suit une loi normale.



On peut voir sur ce graphique que la plupart des échantillons suivent la loi normale. Dès lors, je décide de valider ce test.

3 Enfin, pour tester l'homogénéité des variances de résidus, je procède à un test de Levene.

Test de Levene (Médiane) / Test bilatéral (Etes vous enclins à vous faire guider pour écouter de la musique) :	
F (Valeur obs)	2,490
F (Valeur critik)	3,085
DDL1	2
DDL2	102
p-value (bilaté)	0,088
alpha	0,05
Interprétation du test :	
H0 : Les variances sont identiques.	
Ha : Au moins l'une des variances est différente d'une autre.	
Etant donné que la p-value calculée est supérieure au niveau de signification seuil alpha=0,05, on ne peut pas rejeter l'hypothèse nulle H0.	

Je peux voir grâce à ce test que la p-value est de 0,088 ce qui est supérieur au niveau de signification alpha de 0,05 et donc je ne peux pas rejeter l'hypothèse nulle selon laquelle les variances sont identiques. Mon test est donc valide (Yergeau, 2021).

L'ANOVA respectant les 3 tests de fiabilité, je la considère comme valide et ses résultats peuvent donc être vérifiés.

Analyse ANOVA

Avant de commencer l'analyse ANOVA, il est bon de répéter quelles sont mes hypothèses.

H0 : Consommer des playlists créées par la plateforme de streaming n'influence pas sur le fait d'être enclin à se faire guider pour écouter de la musique.

H0 : Les moyennes sont donc semblables entre les différents répondants peu importe qu'ils écoutent des playlists créées par la plateforme ou non.

Ha: Consommer des playlists créées par la plateforme de streaming influence le fait d'être enclin à se faire guider pour écouter de la musique.

Ha: Les moyennes sont donc différentes entre les différents répondants si ils écoutent des playlists créées par la plateforme.

Analyse de la variance (Etes vous enclins à vous faire guider pour écouter de la musique) :					
Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	4	3,537	0,884	1,044	0,389
Erreur	100	84,711	0,847		
Total corrigé	104	88,248			
Calculé contre le modèle $Y=Moyenne(Y)$					

On peut voir sur ce tableau plusieurs informations.

- D'une part, la p-value est égale à 0,389 et donc plus grande que le seuil de significativité alpha qui est de 0,05. Donc on ne peut pas rejeter l'hypothèse nulle (Yergeau, 2021).
- D'autre part on peut également analyser trois chiffres, la variabilité intra-groupe qui est ici égale 0,884 et la variabilité entre les moyennes des groupes qui est ici égale à 0,847. Ces deux chiffres nous permettent de calculer le facteur F qui est ici égal à 1,044. Plus cette valeur est haute, plus les moyennes des groupes sont éloignées les unes des autres (Yergeau, 2021). Ici, la valeur est légèrement plus haute que 1, les moyennes des groupes ne sont donc fort éloignées les unes des autres.

Analyse du R^2

Enfin, il est possible d'analyser à quelle degré nos réponses sont expliquées par notre modèle. Pour cela il faut analyser le R carré qui est le coefficient de détermination. Il permet de connaître la proportion de la variance qui est expliquée par notre variable dépendante.

Régression de la variable Etes vous enclins à vous faire guider pour écouter de la musique :	
Coefficients d'ajustement (Etes vous enclins à vous faire guider pour écouter de la musique) :	
Observations	105
Somme des p	105
DDL	100
R^2	0,040
R^2 ajusté	0,002
MCE	0,847
RMCE	0,920
MAPE	16,347
DW	2,066
Cp	5,000
AIC	-12,545
SBC	0,724
PC	1,056

Ici, le R carré est de 0,040. Le R qui est donc la racine de 0,040 est égale à $\sqrt{0,040} = 0,632$.

La statistique r indique donc que la différence entre les groupes est de grande taille (Cohen, 1988).

Annexe 9 - Analyse par la régression simple de l'hypothèse numéro 3

Analyse par la régression simple : hypothèse numéro 3, lien entre l'adéquation de l'offre des plateforme de streaming et le fait d'être enclin à se faire guider

Étape 1 : Evaluation de la pertinence du modèle de régression

Analyse de la variance

Analyse de la variance (Etes vous enclins à vous faire guider pour écouter de la musique) :					
Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	1	12,137	12,137	14,849	0,000
Erreur	112	91,547	0,817		
Total corrigé	113	103,684			
Calculé contre le modèle Y=Moyenne(Y)					

Ce tableau s'interprète de la même manière que celui pour l'analyse ANOVA. La valeur de F est de 14,849 et est significative car p est plus petit que l'alpha de 0,05 (Yergeau, 2021).

Nous pouvons donc rejeter l'hypothèse nulle formulée plus haut et il existe donc une relation statistiquement significative entre le fait d'être sujet à l'hyper-choix et le fait d'être enclin à se faire guider pour écouter de la musique.

Étape 2 : Évaluation de l'ajustement des données au modèle de régression

Coefficients d'ajustement (Etes vous enclins à vous faire guider pour écouter de la musique) :	
Observations	114
Somme des p	114
DDL	112
R ²	0,117
R ² ajusté	0,109
MCE	0,817
RMCE	0,904
MAPE	16,316
DW	2,036
Cp	2,000
AIC	-21,006
SBC	-15,533
PC	0,914

Ici le R^2 est de 0,117. Nous pouvons donc dire que la variable indépendante "les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement" peut expliquer près de 11,7% de la variation du fait d'être enclin à se faire

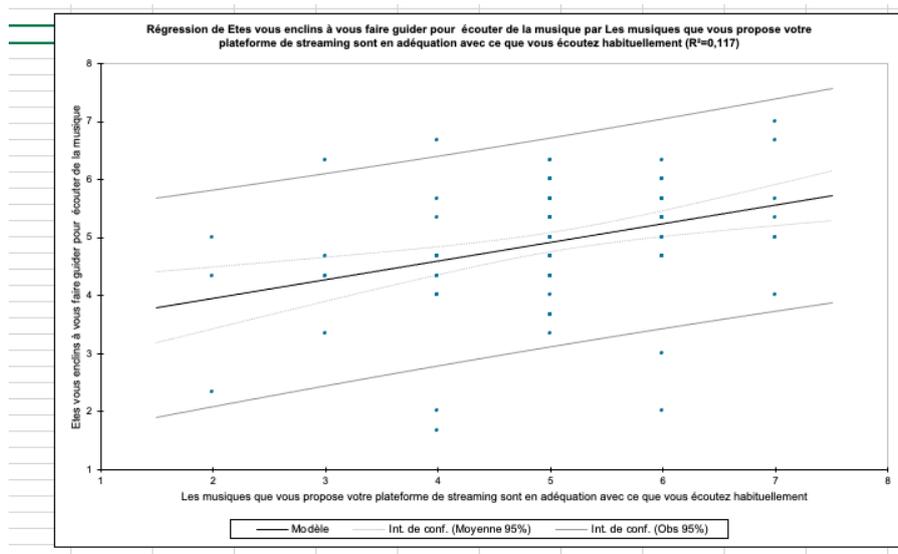
guider pour écouter de la musique. Le reste de la variabilité est dû à des effets (autres variables explicatives) qui ne sont pas pris en compte dans cet exemple (Maalouf, 2017).

Paramètres du modèle (Etes vous enclins à vous faire guider pour écouter de la musique) :						
Source	Valeur	Erreur standard	t	Pr > t	Borne inférieure	Borne supérieure
Constante	3,317	0,431	7,690	<0,0001	2,463	4,172
Les musiques	0,321	0,083	3,853	0,000	0,156	0,487

Ce tableau permet de faire des prévisions ou des simulations. Les valeurs importantes à regarder sont le 3,317 et le 0,321 représentant l'ordonnée à l'origine et la pente de la droite, le beta (Yergeau, 2021). Le 0,321 indique donc que si le répondant est sujet à une unité en plus de la variable "le fait que les musiques que vous propose votre plateforme de streaming sont en adéquation avec ce que vous écoutez habituellement", alors il augmentera le fait d'être enclin à se faire guider de 0,321 (sur une échelle de 7 points)

De plus, nous pouvons écrire une équation finale de prévision :

$$\text{Intention de se faire guider} = 3,853 + 0,321 * \text{l'adéquation des propositions} + e$$



Enfin, ce dernier graphique permet de mieux visualiser les données au niveau des résidus. On peut voir ici qu'il n'y a pas réellement de tendance linéaire et que même si les observations sont d'une manière générale groupées, 6 observations sont en dehors des limites de l'intervalle de confiance de 95%.