

Dataviz - Analyse des Données YouTube 2023

Evan PENNEC - Clément MARTIN



Professeur : Idir BENOURET
EPITA - Ing2

23 Avril 2025

Table des matières

1	Introduction	2
2	Quelles sont les statistiques de la consommation de contenus YouTube au niveau mondial ?	3
2.1	Quels sont les types de contenus les plus regardés ?	3
2.2	Nombre de chaînes par catégories de contenus ?	4
2.3	Abonnés moyens par type de chaîne ?	5
2.4	Le nombre d'abonnés est-il une conséquence directe du nombre de vues ?	6
3	Quelles sont les statistiques des chaînes YouTube par pays ?	7
3.1	Combien y a-t-il de chaînes YouTube par pays dans le dataset ?	7
3.2	Quelles sont les plus grosses chaînes par pays ?	8
3.3	Quelles catégories de contenus sont les plus suivies par pays ?	9
4	Quelle est l'influence de la situation socio-économique sur la consommation de contenus YouTube ?	10
4.1	Quelle est l'influence de la taille de la population d'un pays sur les abonnements ?	10
4.2	Sur les vues ?	11
4.3	Les pays ayant un taux de chômage élevé regardent-ils plus YouTube ?	11
4.4	Et quels types de contenu regardent-ils ?	12
4.5	Quel type de contenu les pays les plus urbanisés regardent-ils ?	13
5	Conclusion	15

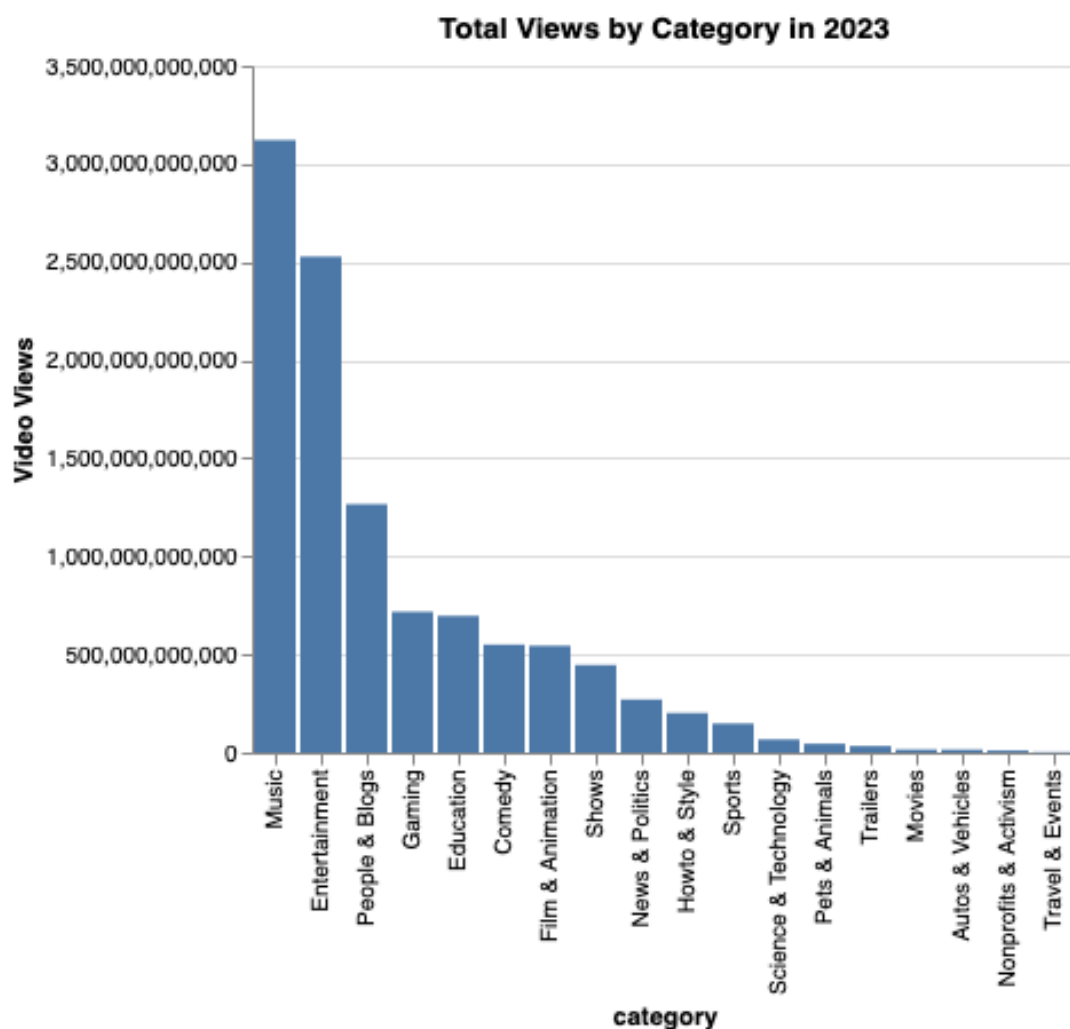
Introduction

Ce rapport analyse les données YouTube de 2023 pour répondre à des questions clés sur la consommation mondiale de contenu, les statistiques des chaînes par pays, et les influences socio-économiques sur les tendances d'utilisation de la plateforme. À travers une série de visualisations réalisées avec Altair, nous explorons les catégories les plus visionnées, le nombre de chaînes par catégorie, les pays dominants en termes de chaînes et de vues, ainsi que l'impact de facteurs comme la population, le chômage et l'urbanisation.

YouTube, en tant que plateforme mondiale de partage de vidéos, offre une richesse de données permettant d'explorer les préférences des utilisateurs à travers le monde. Ce rapport utilise un ensemble de données globales sur les statistiques YouTube de 2023 (`Global_YouTube_Statistics.csv`) pour répondre à trois grandes catégories de questions : les statistiques de consommation, les statistiques par pays, et les influences socio-économiques. Le dataset contient des informations sur les chaînes YouTube les plus influentes en 2023. Il contient 28 colonnes et 995 chaînes YouTube, nous permettant une large analyse. Les visualisations, générées via Python avec les bibliothèques Pandas et Altair, servent de base à notre analyse.

Quelles sont les statistiques de la consommation de contenus YouTube au niveau mondial ?

2.1 Quels sont les types de contenus les plus regardés ?

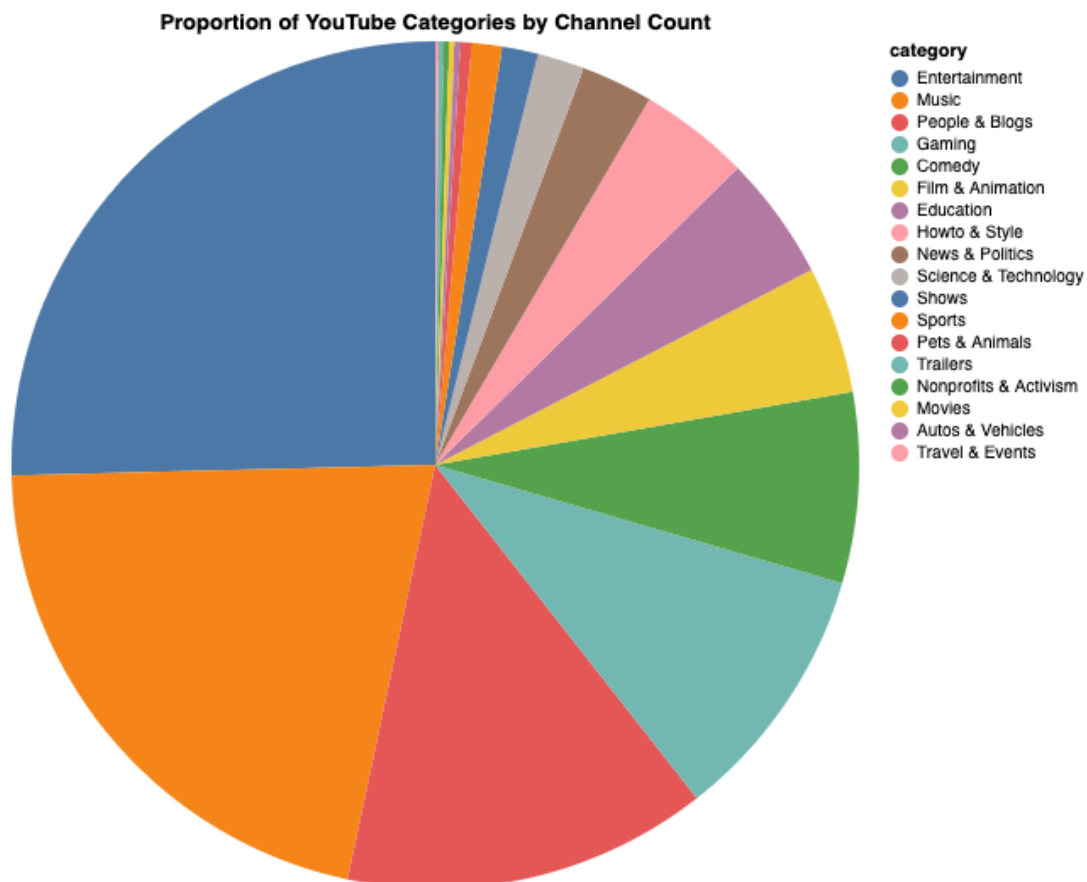


Les catégories de contenus les plus regardées en termes de vues totales sont *Music* (3,12 trillions de vues), *Entertainment* (2,53 trillions de vues), *People & Blogs* (1,27 trillion de vues), *Gaming* (717 milliards de vues) et *Education* (697 milliards de vues).

Nous remarquons que la catégorie *Music* domine largement sur la visualisation ci-dessus, ce qui pourrait signifier que les utilisateurs consomment massivement des clips musicaux en raison de leur accessibilité universelle. YouTube est un moyen d'accès facile pour écouter de la musique gratuitement. *Entertainment* suit, suggérant une forte demande pour des vidéos divertissantes, ce qui paraît plutôt logique vis-à-vis des habitudes de chacun. Après une journée de travail intense et parfois difficile, les gens ne cherchent que la tranquillité en s'asseyant devant une vidéo de divertissement en ne pensant à rien d'autre.

2.2 Nombre de chaînes par catégories de contenus ?

Entertainment (241 chaînes), *Music* (202), *People & Blogs* (132), *Gaming* (94), *Comedy* (69).

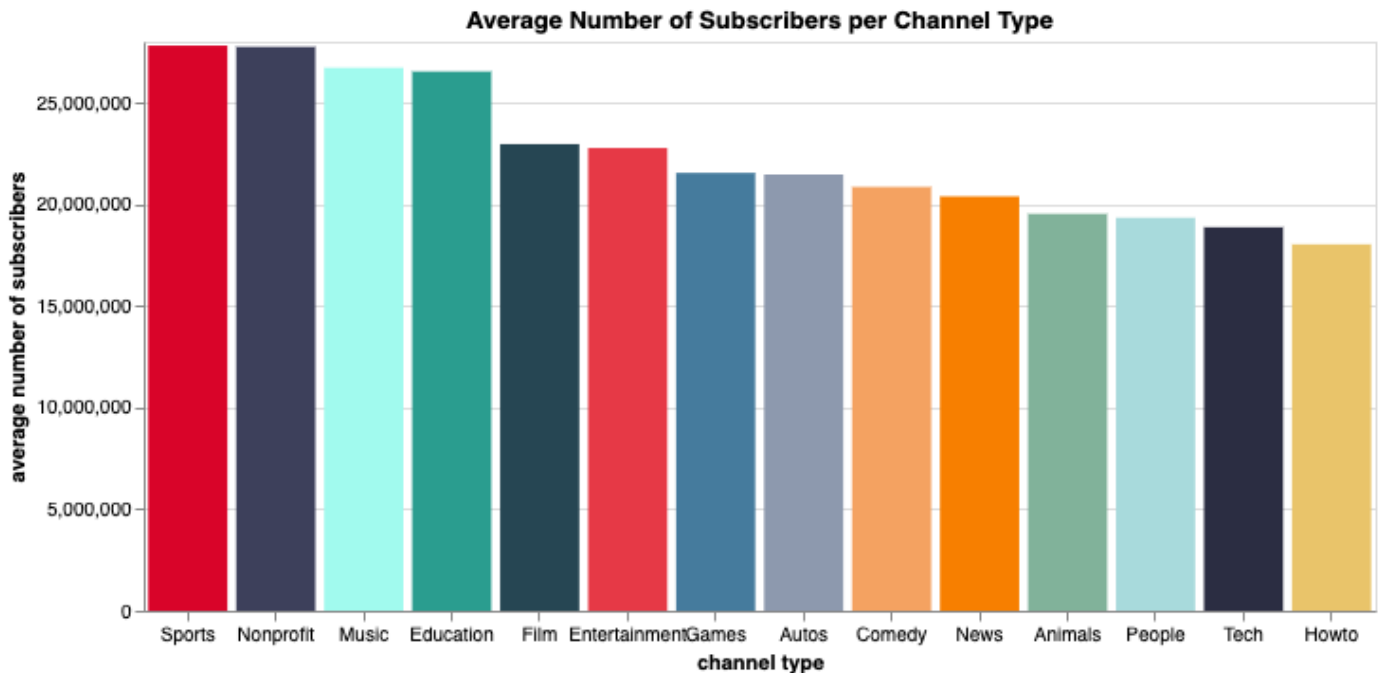


La visualisation montre que *Entertainment* et *Music* attirent le plus de créateurs, ce qui pourrait probablement signifier un fort potentiel de monétisation dû à une forte demande.

- Entertainment englobe une large variété de contenus (vlogs, séries, sketches, etc.), ce qui attire un public diversifié et global. Cette catégorie est souvent perçue comme un "fourre-tout" pour les créateurs qui produisent des vidéos divertissantes sans se limiter à une niche spécifique. Cela pourrait expliquer pourquoi elle attire autant de créateurs : elle offre une flexibilité dans le type de contenu produit et une audience potentielle très large.
- Music reflète l'attrait universel de la musique. Les clips musicaux, les performances live, et les playlists attirent des audiences transculturelles, et les grandes maisons de disques (comme T-Series en Inde) utilisent YouTube comme une plateforme principale pour diffuser leur contenu.

2.3 Abonnés moyens par type de chaîne ?

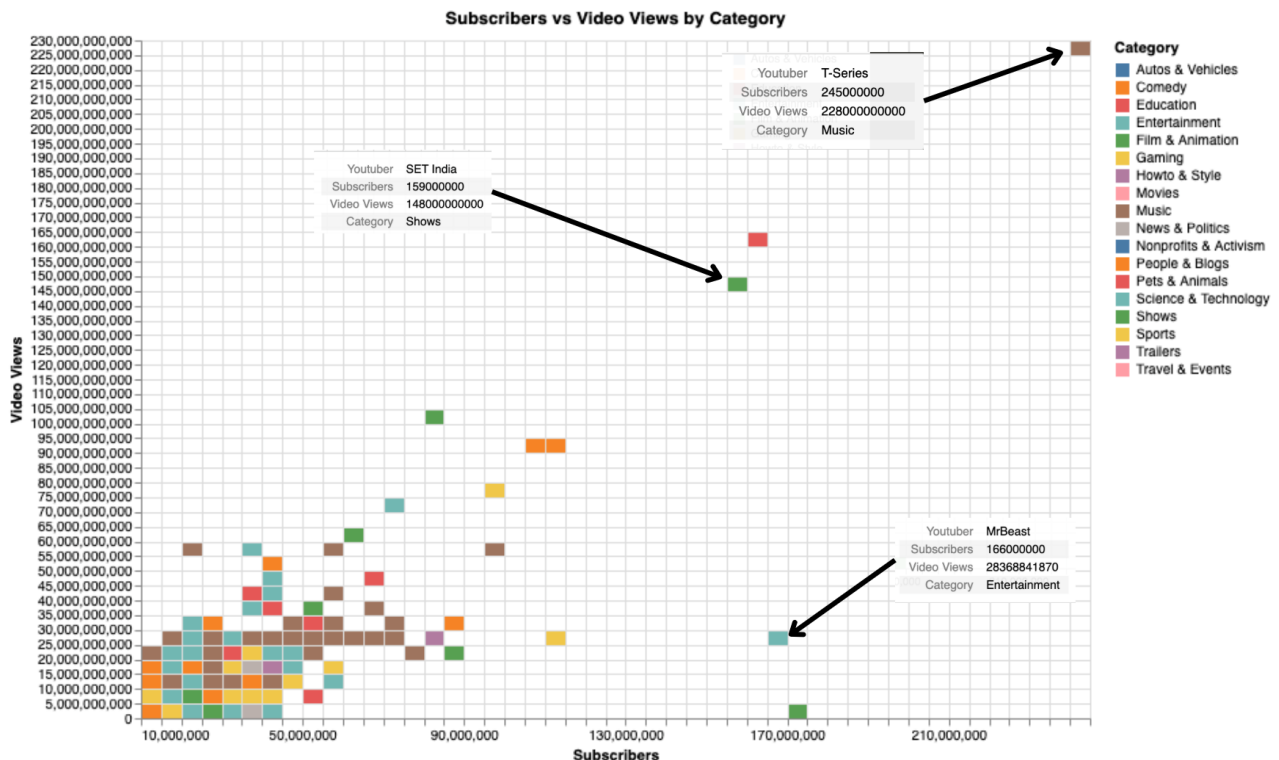
Les abonnés moyens varient en fonction du type de chaîne.



La visualisation montre que *Sports* (27.8 millions) et *Music* (26.7 millions) ont les abonnés moyens les plus élevés.

- *Sports* (27.8 M) et *Music* (26.7 M) dominent, ce qui pourrait signifier que ces types de chaînes fidélisent mieux leurs audiences. Les chaînes sportives, souvent liées à des clubs, des événements ou des personnalités (ex. NBA, FIFA), attirent des fans passionnés qui s'abonnent pour suivre des mises à jour régulières. De même, *Music* bénéficie de la popularité des artistes et des labels (ex. T-Series), qui attirent des abonnés grâce à des sorties fréquentes et un attrait universel.
- *Entertainment* (18.8 M) et *Games* (18.5 M), bien que populaires en nombre de chaînes, ont des abonnés moyens plus faibles. Cela pourrait signifier une forte concurrence dans ces types, diluant l'audience par chaîne. Les créateurs dans *Entertainment* (ex. vlogs, séries) et *Games* (ex. streams) doivent se démarquer dans un marché saturé, ce qui limite leur croissance individuelle.
- *Howto* (13.5 M) et *Tech* (15.2 M) ferment la marche, suggérant que ces types, bien qu'utiles, attirent des audiences plus spécifiques et moins massives, peut-être en raison d'une demande plus fonctionnelle (ex. tutoriels) qu'émotionnelle.

2.4 Le nombre d'abonnés est-il une conséquence directe du nombre de vues ?



Nous remarquons globalement qu'en effet, il existe une corrélation générale entre le nombre d'abonnés et le nombre de vues, mais elle n'est pas toujours directe, comme le montre la visualisation. Certaines chaînes, comme T-Series et SET India, suivent cette tendance, mais des exceptions comme MrBeast montrent que des facteurs autres que les vues influencent les abonnés.

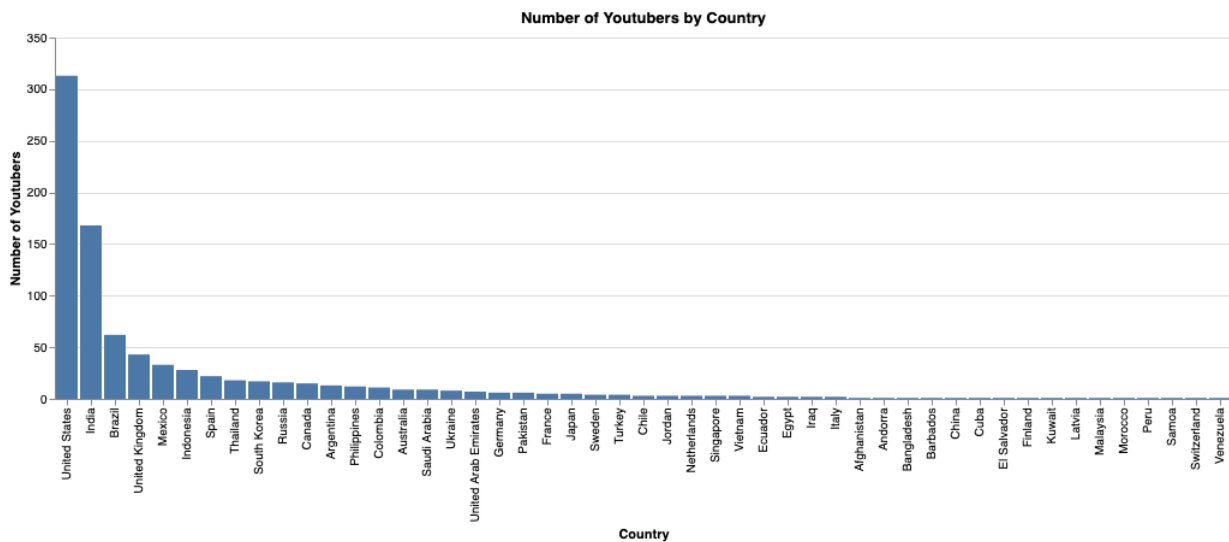
La visualisation illustre la relation entre abonnés et vues par catégorie, avec des points représentant chaque chaîne :

- La majorité des points se concentrent dans la zone inférieure gauche (moins de 50 M abonnés, moins de 50 B vues), mais une tendance positive est visible : les chaînes avec plus d'abonnés ont souvent plus de vues.
- Exceptions : MrBeast se distingue avec un nombre d'abonnés élevé mais des vues bien inférieures à T-Series ou SET India. Cela pourrait s'expliquer par une stratégie de contenu : MrBeast produit des vidéos virales à fort engagement (ex. défis coûteux), attirant des abonnés fidèles, mais publie moins fréquemment (741 uploads) que SET India (116 536 uploads), ce qui limite ses vues totales.

On pourrait donc dire que les abonnés ne découlent pas uniquement des vues. La fréquence de publication, la viralité, et le type de contenu jouent un rôle clé.

Quelles sont les statistiques des chaînes YouTube par pays ?

3.1 Combien y a-t-il de chaînes YouTube par pays dans le dataset ?

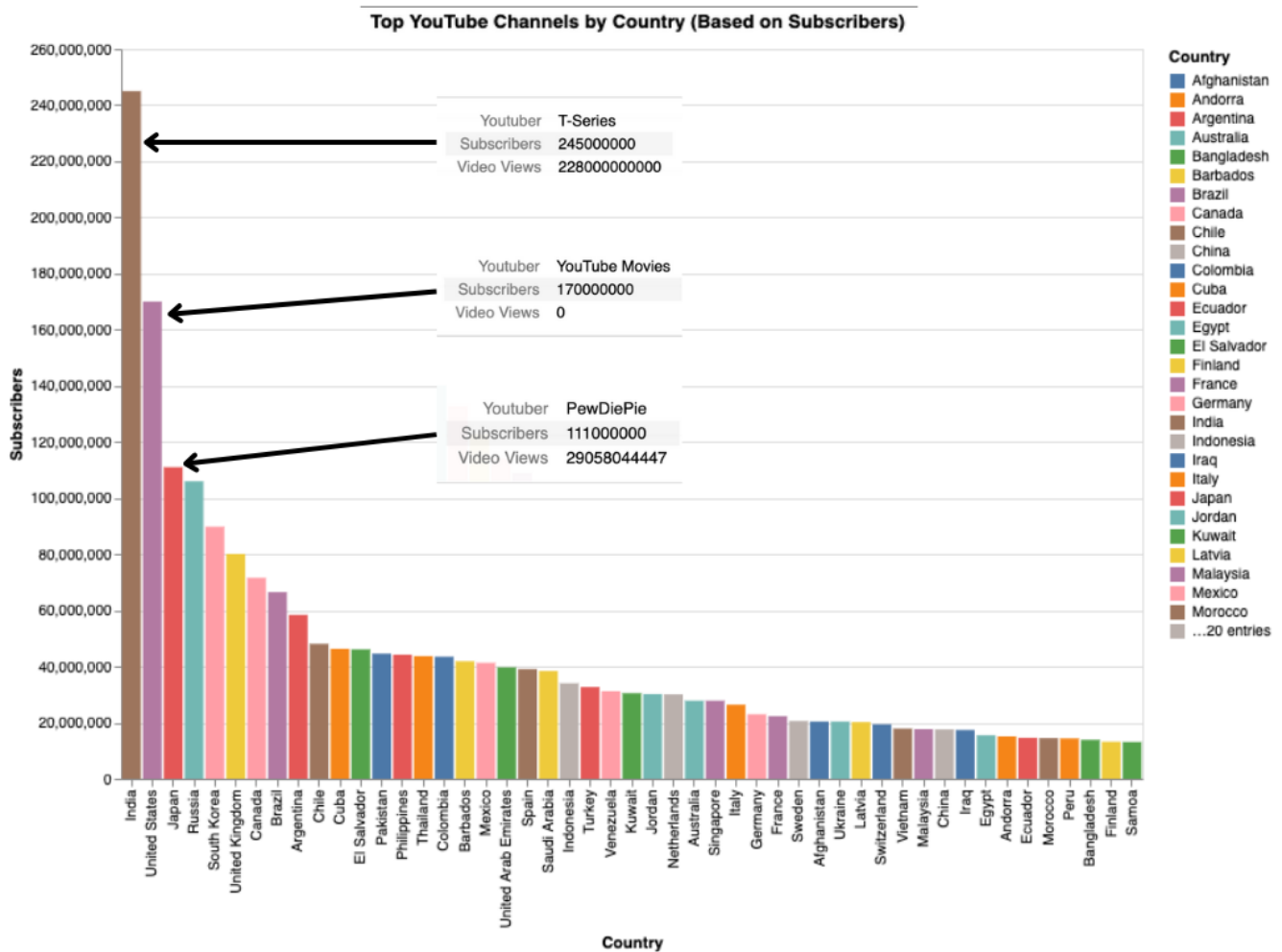


La visualisation montre une forte concentration du nombre de youtubers aux États-Unis (313 chaînes) et en Inde (168 chaînes), qui représentent près de la moitié des chaînes enregistrées dans le dataset. Le Brésil (62) et le Royaume-Uni (43) suivent, mais avec un écart notable. Cette répartition s'explique par :

- Population et accès à Internet : Les États-Unis et l'Inde, avec leurs grandes populations (328 M et 1.37 B) et un accès Internet développé, offrent un vivier de créateurs et d'audience.
- Culture numérique : Les États-Unis, berceau de YouTube, bénéficient d'une culture numérique avancée. L'Inde, avec Bollywood et une industrie musicale florissante, a vu une explosion de contenu local (ex. T-Series).

La domination des États-Unis et de l'Inde suggère que YouTube reste concentré dans les grands marchés numériques, mais la présence de pays comme le Brésil indique une diversification croissante. Également, nous pourrions penser que beaucoup de données n'ont pas été enregistrées dans le dataset.

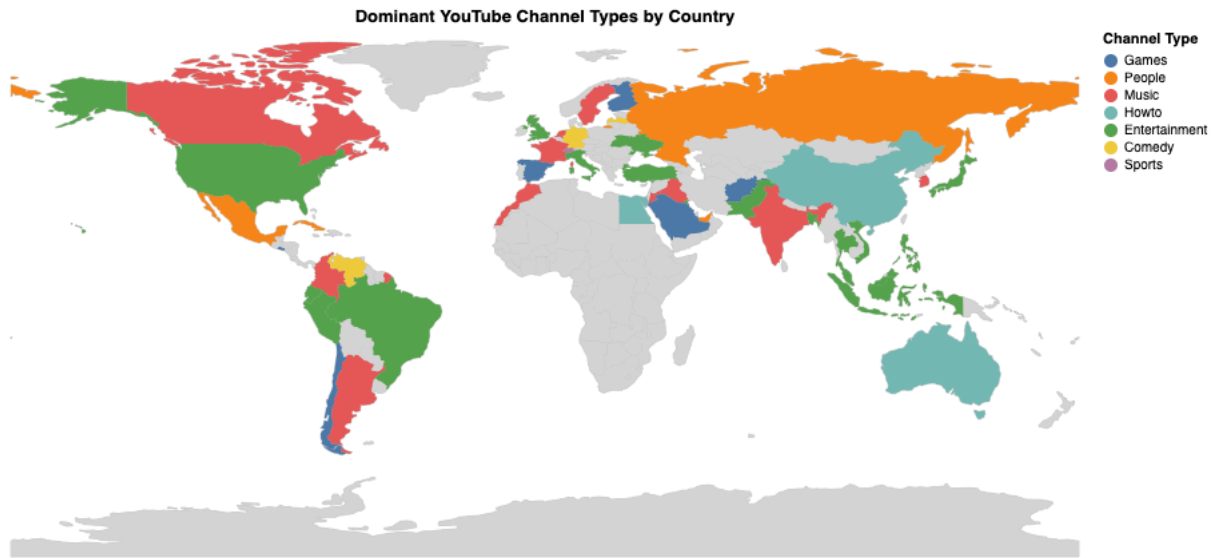
3.2 Quelles sont les plus grosses chaînes par pays ?



Les plus grosses chaînes par pays en 2023 en termes d'abonnés sont T-Series (Inde, 245 M abonnés, 228 B vues), YouTube Movies (États-Unis, 170 M abonnés, 0 vues), et PewDiePie (Japon, 111 M abonnés, 29 B vues).

La visualisation montre une forte disparité entre les pays. Cette visualisation montre à nouveau qu'en 2023, l'Inde et les États-Unis dominaient l'industrie YouTube. Nous retrouvons logiquement les chaînes ayant le plus d'abonnés au niveau des pays les plus peuplés au monde.

3.3 Quelles catégories de contenus sont les plus suivies par pays ?



Les catégories les plus suivies varient par pays : Music domine en Inde, au Canada, et en France ; Entertainment est prédominant aux États-Unis, au Royaume-Uni, et au Brésil ; Games est populaire en Espagne ; Howto domine en Chine et en Australie ; Comedy en Colombie.

La carte montre une répartition géographique des types de chaînes dominants. Globalement, nous remarquons que les différents types de chaînes sont présents un peu partout. Les préférences sont influencées par des facteurs culturels et démographiques. Entertainment et Music attirent des audiences globales, tandis que des types comme Games ou Howto répondent à des besoins spécifiques, offrant des opportunités pour des créateurs ciblés.

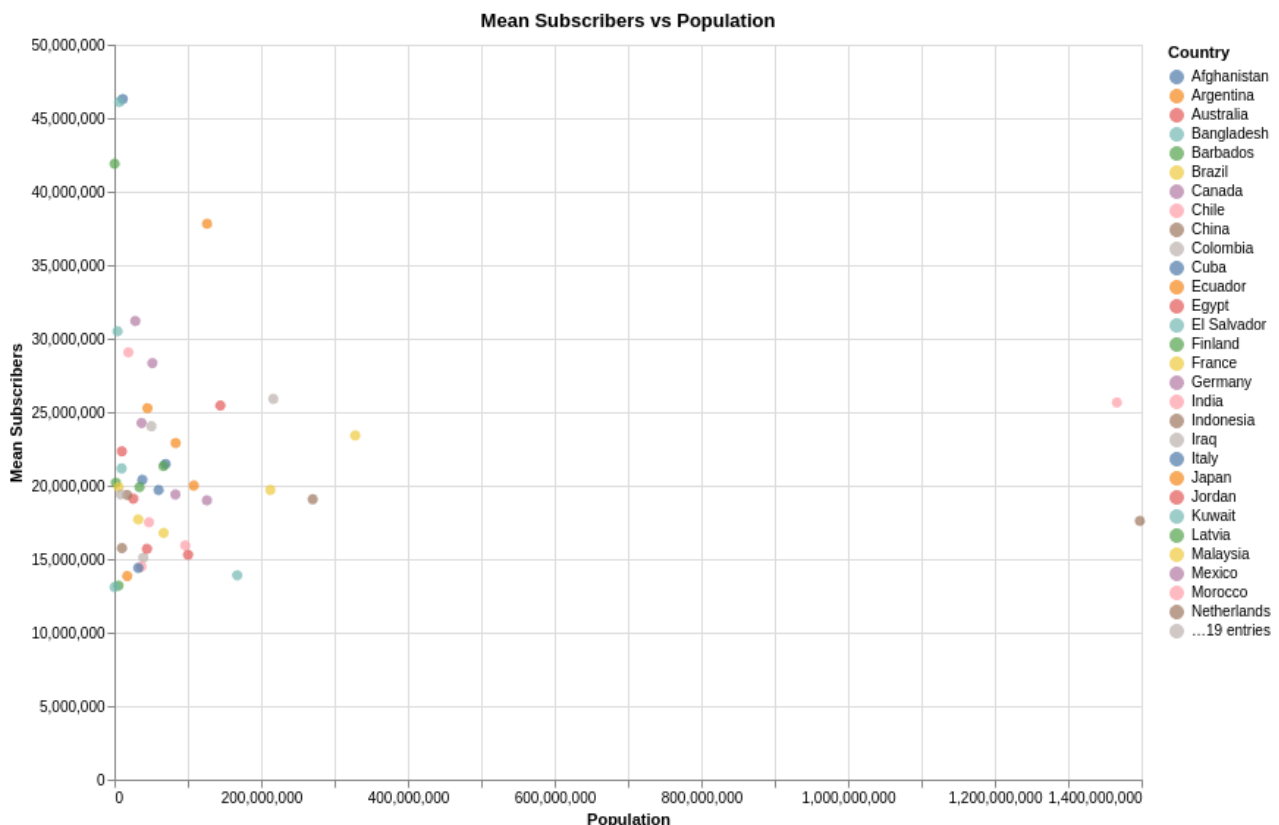
En France, en 2023, la chaîne la plus populaire était une chaîne de musique. Cela peut ne pas forcément refléter totalement la réalité, car les jeunes français aiment regarder des vidéos de divertissement avec une gamme de youtubeurs variés tels que Squeezie ou Amixem. Cependant, cela confirme le fait que les jeunes regardent les vidéos sans forcément s'abonner à la chaîne du créateur de contenu, tandis que les chaînes de musique ont plus de succès en termes d'abonnement.

Quelle est l'influence de la situation socio-économique sur la consommation de contenus YouTube ?

Le dataset comprend des données sur les situations socio-économiques des pays de création des chaînes YouTube. Nous allons maintenant essayer de comprendre quelles sont les influences de ces situations sur les performances et de type de contenus des chaînes présentes dans le dataset.

4.1 Quelle est l'influence de la taille de la population d'un pays sur les abonnements ?

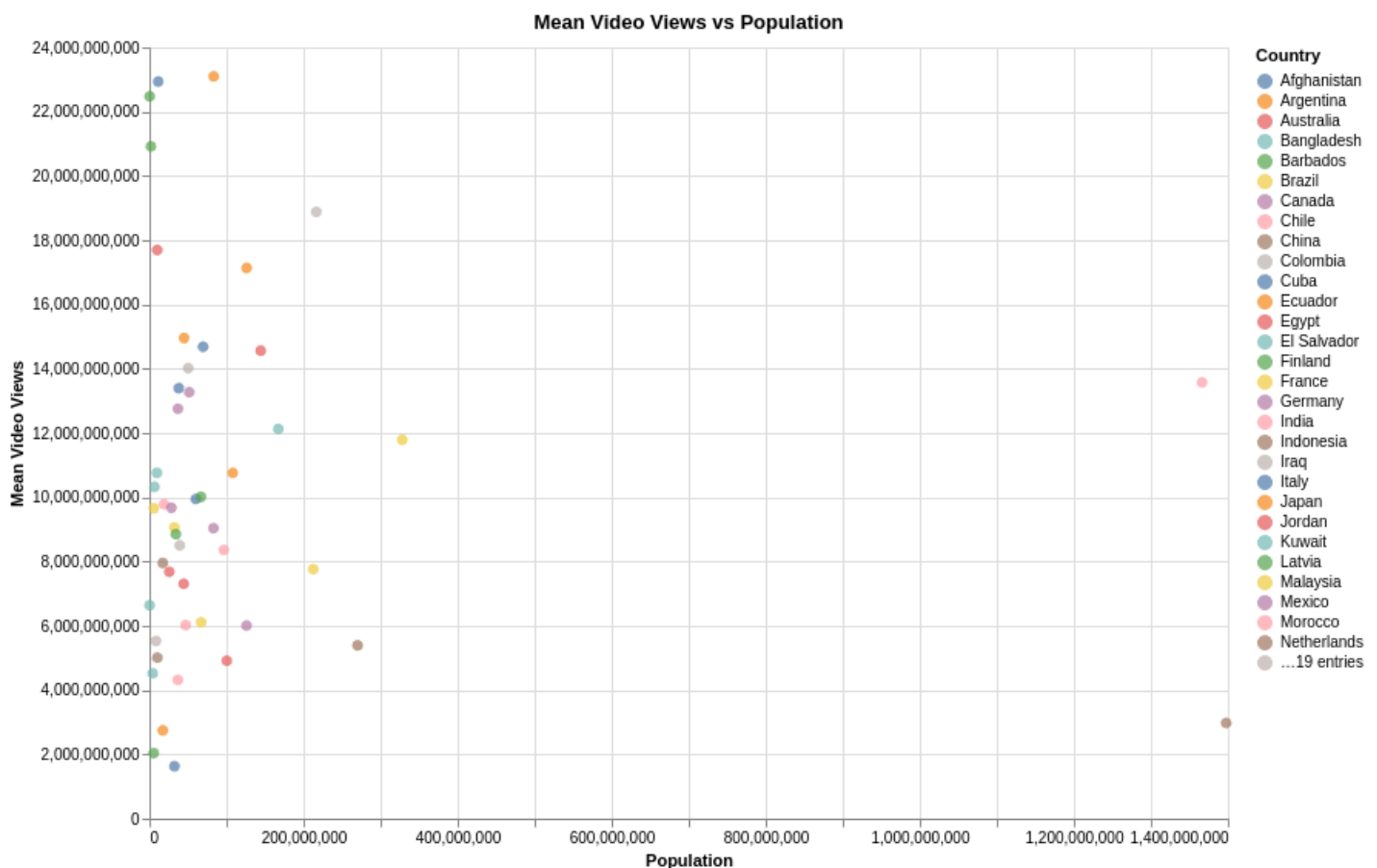
Dans cette partie, nous nous demandons si le nombre d'habitants d'un pays influe sur le nombre d'abonnements des chaînes YouTube ayant été créées dans ces mêmes pays. Nous avons donc calculé la moyenne des abonnements par populations et donc par pays.



Cette visualisation montre qu'il n'y a pas de lien évident entre la population d'un pays et le nombre d'abonnés. En effet, des pays ayant une faible population comme Cuba (11.333.483 d'habitants) ont des nombres d'abonnements moyens plus élevés que des pays ayant des populations plus importantes comme la Chine (1.397.715.000 d'habitants), et inversement.

4.2 Sur les vues ?

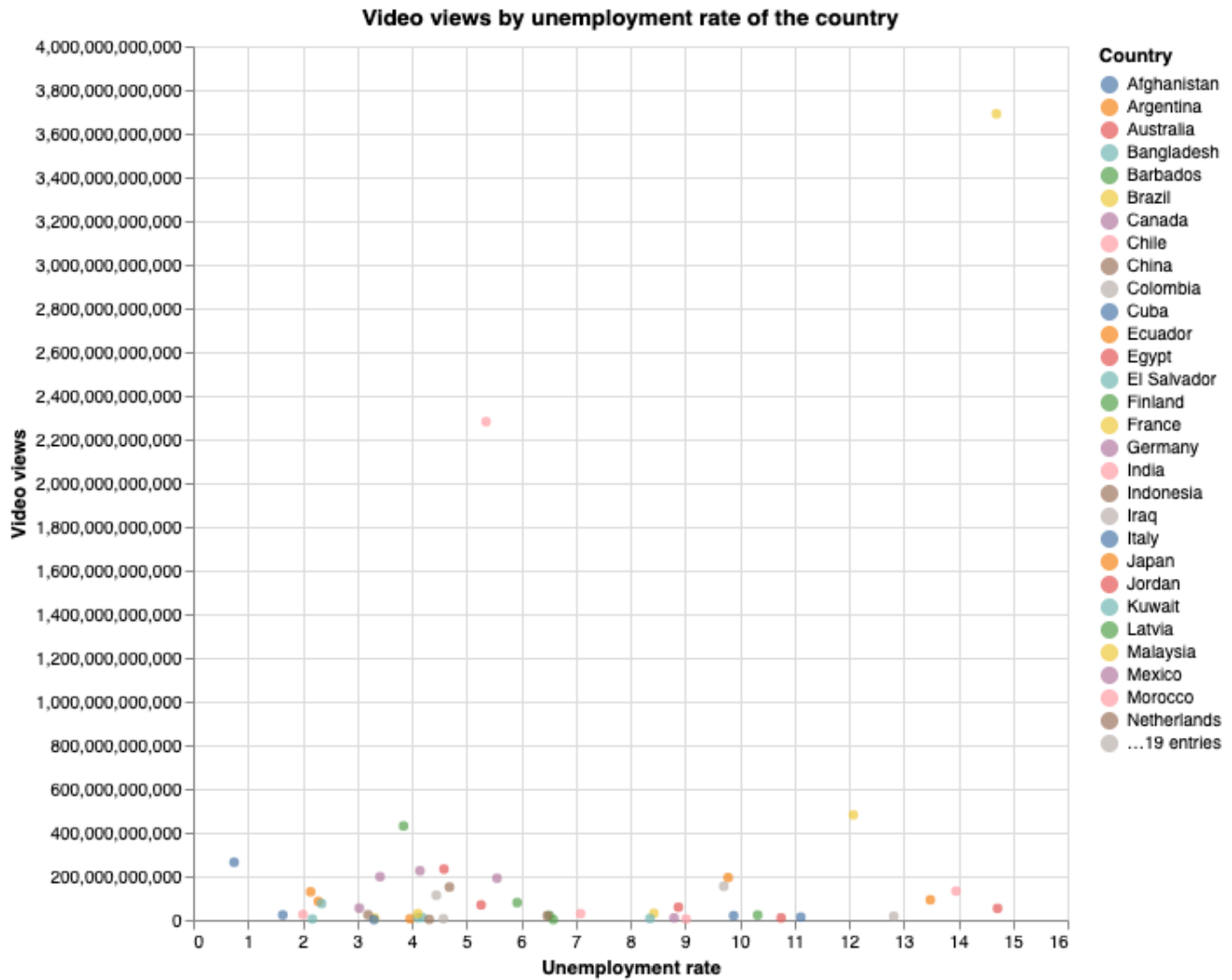
Nous pouvons nous poser la même question que précédemment, mais avec les vues. De même, nous avons calculé la moyenne des vues des chaînes.



De même que précédemment, la visualisation ne permet pas d'identifier un lien entre la population d'un pays et les vues. Le pays ayant la moyenne de vues la plus élevée (23.099.811.190 vues) est la Turquie qui a une population de 'seulement' 83.429.615 habitants.

4.3 Les pays ayant un taux de chômage élevé regardent-ils plus YouTube ?

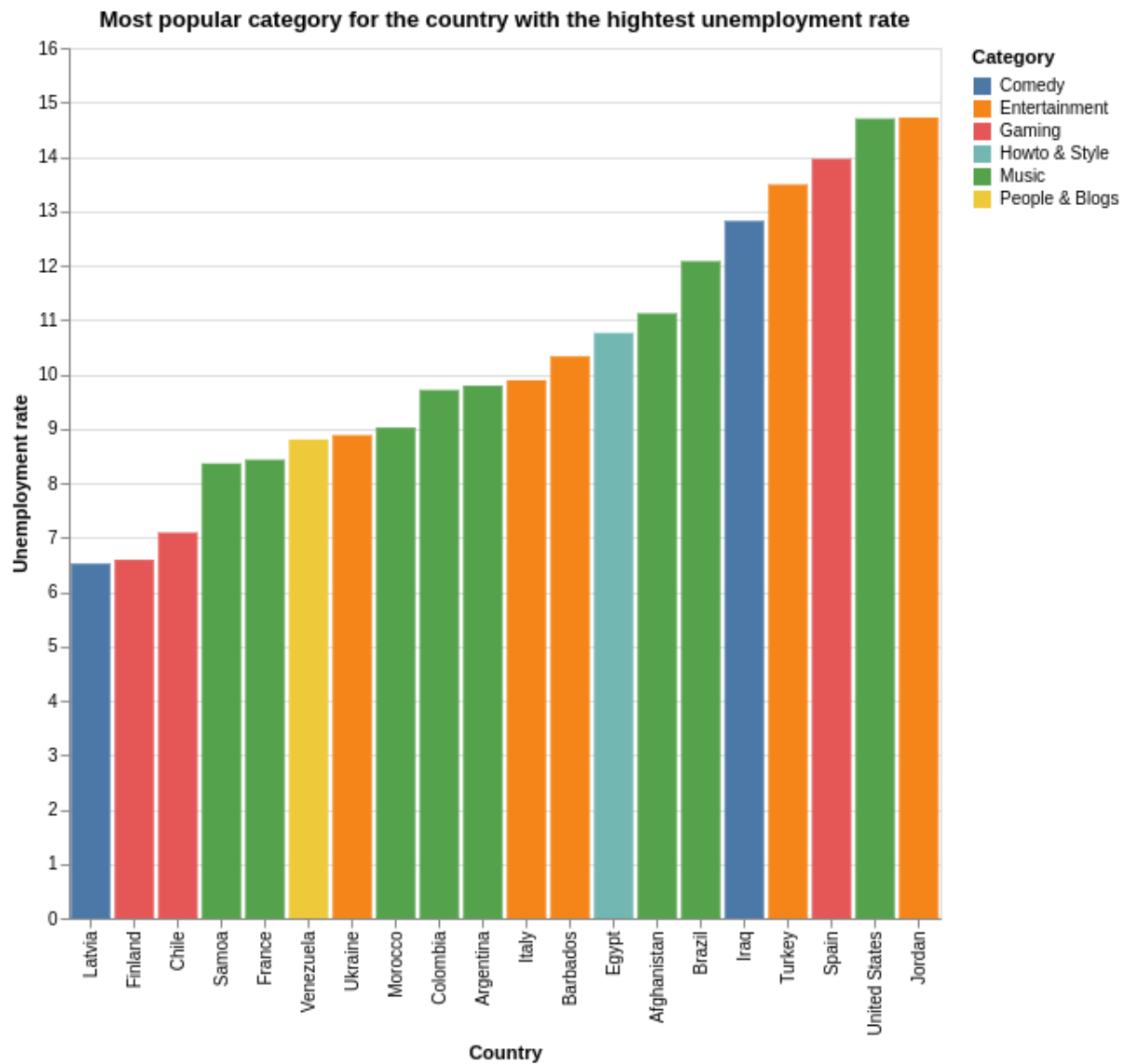
Le dataset comprend des données sur le taux de chômage de la population de chaque pays dans lequel les chaînes YouTube ont été créées, et donc en général sur le public. Nous avons regardé si ce taux de chômage avait une influence sur les statistiques YouTube.



Nous pouvons voir que le taux de chômage n'a pas d'influence particulière sur le nombre de vues. En excluant les États-Unis (3.69×10^{12} views, unemployment rate : 14,7 %) et l'Inde (2.28×10^{12} views, unemployment rate : 5.36 %) qui ne sont pas représentatifs du reste des données, aucun lien visible ne se fait entre le taux de chômage et le nombre de vues.

4.4 Et quels types de contenu regardent-ils ?

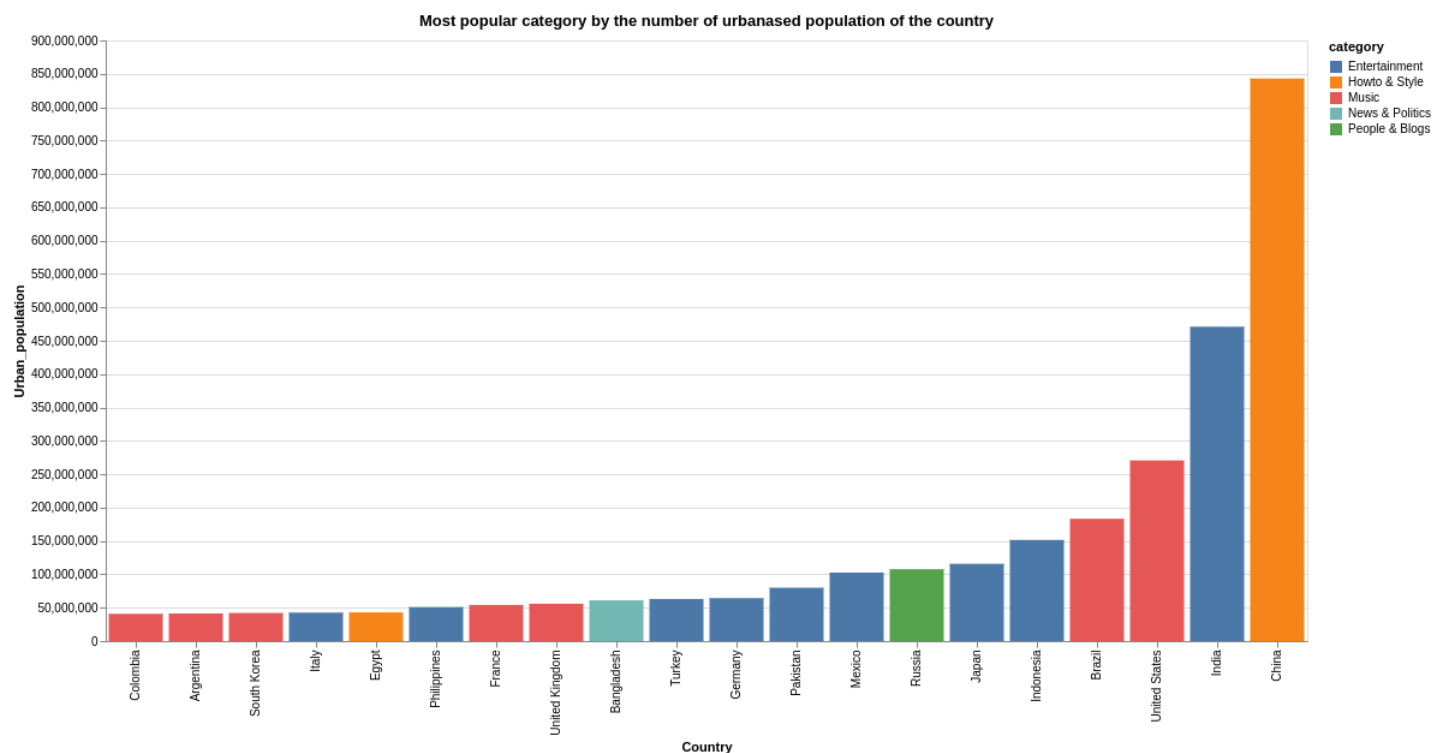
Nous pouvons maintenant nous demander si le type de contenu se dégage pour les pays ayant un taux de chômage élevé. Ici, nous allons arbitrairement nous restreindre aux 20 pays ayant un taux de chômage le plus élevé.



Nous pouvons voir que le type de contenus majoritaires est *Music* (8 pays sur 20) ce qui diffère de nos résultats au niveau global qui, rappelons-le, plaçait *Entertainment* devant.

4.5 Quel type de contenu les pays les plus urbanisés regardent-ils ?

Une autre donnée présente dans le dataset est le nombre de personnes urbaines dans les pays. Nous allons une nouvelle fois arbitrairement nous restreindre aux 20 pays ayant les plus de personnes urbaines.



Nous pouvons voir que ce sont les contenus de type *Entertainment* qui sont les plus représentés (9 sur 20 pays). La proportion de contenu issu de la catégorie *Entertainment* approche ici les 50% alors qu'au niveau global, la catégorie *Entertainment* ne représentait qu'un peu plus de 25% des contenus.

Conclusion

L'analyse des données YouTube 2023 révèle plusieurs leçons clés sur la consommation mondiale de contenus, les dynamiques par pays, et les influences socio-économiques.

Tout d'abord, les catégories Music et Entertainment dominent en termes de vues et de chaînes, portées par leur attrait universel et leur potentiel de monétisation, notamment dans des pays à forte population comme les États-Unis et l'Inde. Ces deux pays concentrent près de la moitié des chaînes du dataset, soulignant l'importance des grands marchés numériques et de la culture locale.

Ensuite, bien que le nombre d'abonnés et de vues soit souvent corrélé, des exceptions comme MrBeast montrent que la viralité, la fréquence de publication, et la fidélisation jouent un rôle crucial. Sur le plan socio-économique, ni la taille de la population ni le taux de chômage ne semblent influencer directement les vues ou abonnements, mais l'urbanisation favorise les contenus Entertainment, probablement en raison d'un accès accru à Internet.

Enfin, les préférences par pays reflètent des facteurs culturels et démographiques, avec Music dominant dans des contextes variés et Games ou Howto attirant des audiences plus spécifiques. Ces résultats mettent en lumière l'importance des dynamiques démographiques, culturelles, et économiques dans la consommation de contenus YouTube, offrant des insights précieux pour les créateurs et les plateformes cherchant à optimiser leur portée.