MASTER IN BUSINESS ENGINEERING PROFESSIONAL FOCUS IN DATA SCIENCE

A study of brands communication strategies
the case of streaming services on Twitter

LUCCA, Hugo

Award date:
2022

Awarding institution:
University of Namur

A study of brands communication strategies : the case of streaming services on Twitter

Hugo LUCCA

Directeur: Prof. I. LINDEN

Mémoire présenté
en vue de l'obtention du titre de
Master 120 en ingénieur de gestion, à finalité spécialisée
en data science

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

Acknowledgements

First of all I would like to thank Mrs Linden, my supervisor. She was extremely present and available throughout the preparation and writing of this thesis. Her help has been invaluable and I am very grateful to her.

I would also like to thank my fellow promotion students who helped me, informed me and gave me their opinion when I needed it.

Summary

This thesis consists of an analysis of communication on social networks in the context of the streaming services sector. The 3 biggest players are studied: Netflix, PrimeVideo and Disney+. Each of the brands is studied via two accounts, one English-speaking and one French-speaking. The analyses will consist of different techniques such as text-mining, data-mining and a more qualitative analytical approach. We obtain tools such as metrics, the most used words of each account or a categorization of the tweets of an account. These are used to conclude different things about the communication of brands in this sector. Indeed, the aim of this thesis is to establish a communication profile for each brand, to classify these brands according to strategies and also to identify socio-cultural differences between 2 accounts of the same brand.

Contents

1 Introduction

Social networks have been growing steadily over the last 10 years. Starting as simple platforms for virtual contact and sharing content with friends, the large-scale communication they allow has quickly attracted public entities such as politicians, celebrities but also companies. Indeed, the latter can now communicate on a large scale without spending astronomical amounts of money on advertising (on television for example). Nowadays, creating and maintaining a community is essential in many businesses, especially those providing services to individuals. The reason why brands rely so much on influencers, streamers, youtubers and other forms of creators with a built up community is because the effectiveness of this marketing technique has been proven. The key is to have a loyal community that feels close to the brand and even feels affection towards it. The most powerful and even indispensable tool for this is social networks.

They are also used by brands to implement a global business strategy, particularly in terms of the link with the consumer: is it possible to be as close as possible to the consumer and therefore respond to all the private messages received on the company's official account? Or on the contrary, does it only serve to communicate major announcements and does not interact with other users? Given the number of strategic issues represented by these platforms, communication on them had to be carefully managed. This is how the job of community manager was born.

At the same time as the development of social networks another type of service has appeared and has grown at least as fast, streaming video on demand so called SVOD. Today everyone knows Netflix but 10 years ago the service as we know it today was barely born. In this new sector of streaming services, competition is fierce, especially between the 3 biggest actors of the sector Netflix, Amazon Prime Video, Disney+. This market continues to grow while the television market has been declining for several years. These effects can certainly be explained by each other, as for many users streaming services are replacing traditional television. Netflix currently has 221.8 million users, Amazon Prime Video 175 million and Disney+ 129.8 million. From now on, even the cinema will be included in these services. Indeed, some films that were initially supposed to be released in cinemas have finally been released directly on these services (Mulan, Black Widow, Soul, ...) as well as blockbusters that were not planned to be released in the cinema.

These platforms have completely changed consumers' habits and attitudes towards the way they watch programs. One example is the phenomenon of binge watching, which has become increasingly prevalent, particularly since the COVID 19 crisis, as the study by Rubenking and Bracken (2021) [1] shows. In addition, series are becoming increasingly popular with the population and there are also more and more of them, given the quantity of production that these platforms produce. This creates a fierce competition because these platforms produce a lot of content but the time available to users is far from sufficient to watch all the content they are interested in. Moreover, the platforms are not only competing with each other, there is also television, despite its decline in popularity as mentioned above, but also all cultural entertainment in general, such as video games, reading, etc.

It is therefore absolutely necessary for these platforms to stand out from their competitors and acquire a significant market share in a market that is already on the verge of saturation despite its youth. This is where social networks come into play and their role of community acquisition indirectly brings new users.

If Netflix is the leader today, it is of course because this brand was the first to insert itself into the market, but also because of other factors, notably its communication. Without a doubt, one of the key points of Netflix's success in recent years is its communication. Netflix has one of the most atypical digital communications, using a lot of imagination to make its followers react. That's why I focused my research on the sector of streaming services. This thesis has for theme «A study of brands communication strategies : the case of streaming services on Twitter»

During this analysis, several questions are answered. Can we profile each brand in terms of communication ? Do the big 3 services have an established strategy and if so, how do they compare? Are the strategies common to the different national accounts of these brands? Which brands stand out for their communications?

This work is organized as follows : we first study the different techniques used to analyse social network posts in previous studies of this kind. Then we develop the data collection methodology used throughout the research. We then attempt to answer the questions posed above through the analysis of this data. Finally, we discuss the results obtained and attempt to establish first a profile for each brand and then the important communication themes in this sector as well as a the strategy of each service.

2 State of the art

In this section, we draw a survey of the state of the art on a double perspective. This includes articles concerning brand communication strategies as well as articles with a similar purpose to ours and with significant similarities as the sector studied (e.g. SVOD sector). The latter serve both as inspiration and as a comparison, i.e. certain elements deemed beneficial have been partially reproduced in the study. On the contrary, some ways of doing things have been ignored or improved because they were considered as not optimal or even bad.

2.1 Strategy Classification

As cited in the introduction, one of the aim of this thesis is to classify each brand in a communication strategy. In order to perform that, we must have a typology of classification. To find one, we can be inspired by the previous works on this subject.

In a similar approach to ours, Wu and al. (2012) [2] studies the communication of 3 major fast food pizza brands using text mining techniques. One of them is the classification of text based on the themes evoked by the posts and finally classify them in 5 main themes. This makes it possible to identify the specificities of the market in terms of communication and to evaluate the priorities of each theme for each actor within their messages. Are they more inclined towards deliveries (often responses to users) or towards promotional messages about their products? The aim is to establish a profile.

In a less automated way, Oliveira and Figueira (2015) [3] also used text classification on high school posts to obtain different categories. Indeed, this classification is based on the use of different keywords that defined a post as belonging to a class if their presence is verified. However, this classification is used to cluster schools according to their communication strategies. Indeed, these clusters are created according to the recurrences at which each theme is cited by each school. In short, two similar strategies are considered if they evoke each theme in a similar proportion. This second step seems to be less relevant in our case since the number of objects studied is quite limited.

In their paper, Floreddu and Cabiddu (2016) [6] identify the best performing communication strategies according to the reputation of the companies. In order to perform this, they establish certain types of strategies to classify the different brands' communication in. These strategies are as follows: Egocentric, Conversational, Selective, Openness, Secretive, and Supportive. Egocentric is a strategy that consists only of sharing information about their own products, releases, events, ... without interacting with the users. In a Conversational strategy, the firms creates a true relationship with its clients by responding at every messages. A Selective communication only answer to positive comments from customers. Openness is used by companies having the desire to be as transparent as possible and so they do not delete posts and answer publicly to all posts either positive or negative. The Secretive strategy consists in managing negative posts and conflicts by another channel especially private channels. Finally, the Supportive strategy helps the customer during his purchase and answers to its questions.

2.2 Similar Works

During the exploration of the state of the art, it turned out that several researchers had already approached the analysis of the streaming services sector in more or less similar ways in terms of communications. Some of them studied a part of the market players, while others focused only on one of them, Netflix.

One of the theses studies is very close to our subject of study. In fact, in the text of Martinez-Sanchez and al. (2021) [4], communications on Facebook and Instagram of different streaming services (Netlfix, Amazon Prime Video, Disney+, HBO) are studied. However, the ambitions of the analysis are not the same as the ones we have, they only use a few metrics (number of posts per day, evolution of followers,...) as well as the most used word and hashtag. The study period is both very specific but very interesting because it takes place during the global containment of 2020 and this is also when Disney+ launched in Europe. They notice that this period had a big impact on the sector in terms of audio visual content consumed but also on the social media accounts. A lot of new followers came during this period and the statistics of the tweets were higher than before. They also conclude that despite Disney's remarkable entry into the market Netflix is still the undisputed leader. PrimeVideo remains far behind Netflix but ahead of DisneyPlus. However, the growth during Disney's launch phase suggests that Prime will soon be overtaken in terms of popularity.

In the thesis of Scerbinina (2019) [7], they study the communication of Netflix on Twitter. They also compare the communication between two regional accounts of Netflix : Netflix US and Netflix India. It was done by collecting +- 1000 posts per account with a non-automated survey. No text-mining or data-mining techniques are used (except for basic metrics such as number of each type of post, their respective percentages, number of pictures, ...). The analysis is therefore done in a human way by consulting the tweets. Concerning the regionals communication, it was found that there are notable differences between these 2 accounts. Firstly, in terms of content within the tweets with references to each culture. Secondly, in terms of post management with a higher frequency of posts for the American account. Also, an important use of original tweets (= made by the brand) for the Indian account while the US part uses much more retweets from other users. The global analysis of Netflix showed that they use a lot of pictures and videos within their tweets. Furthermore, it was identified that these accounts use a lot of pop culture references in the content of their tweets, especially from their own series but not only. Moreover, it seems that they communicate much more on their original productions than on their other content. During their analysis, 3 types of tweets were identified from Netflix: 'All about Netflix', 'Community relations' and 'The world around us'. The first being tweets about release announcements and other brand related items. The second is about interactions between the brand and its community. The last one represents tweets mentioning elements outside the audio-visual context or at least not the main topic such as tweets about socio-cultural elements like elections etc.

Another study specifically focused on Netflix was conducted by E. Fernández-Gómez and J. Martín-Quevedo (2018) [9]. It studies the strategy used by the brand to create engagement among its followers. Through different techniques such as data mining or text mining it was concluded that in order to promote interaction there are different means such as: creativity, direct questioning, humour, content promotion and mastering the characteristics of social networks (hasthags, memes, ...). These are the ways that Netflix uses to interact according to their study and those that work best.

The specificity of our study is that we use some of the elements mentioned in these different texts

in a complementary way. Indeed, most of these texts only explore one specific facet, which makes the study less complete than what we are doing. The study will use a combination of data-mining, text-mining and qualitative research to profile each account and then classify them according to their profiles and strategies.

3 Methodology

As a reminder my research question is «A study of brands communication strategies : the case of streaming services on Twitter». In order to address this, it is essential to collect data in order to implement text mining. Throughout this section, we first define the text-mining which is the main analysis technique used in our study. In a second time, we explain our thinking process about our data selection in terms of scope, temporality and so on. We then describe our cleaning process of the datasets extracted and the challenges related to it. After that, we present the different data added in our datasets by features extraction. Finally, we resume our different datasets and their size.

3.1 Definition : Text Mining

In this thesis, we use text-mining to perform our analyses as it has been the case in several pervious studies (Wu and al. (2012)[2], Oliveira and Figueira (2015) [3]). Text mining is a process of extracting information from textual data. This includes various techniques such as Natural Language Processing (sentiment analysis, tokenization, stemming ...) aiming then to use data analysis processes such as machine learning or data mining in order to extract information from these data. The final goal of text mining is to "identify patterns, implicit rules, recurrences, causal links that are hidden behind the use of text", Hotho (2005) [8]. In short, it is simply the application of techniques for collecting, cleaning and analysing numerical data in the case of textual data. Text-mining is really relevant to use on social media data as these platforms are a source of unlimited data to analyse. In our study, we use text mining to identify the most words used by each account but also to perform a Latent Semantic Analysis. We also identify the most popular tweets and study their content.

3.2 Data Selection

However, in order to use text-mining techniques, it is of course necessary to have data that allows it. This means text data from the social network communication of different accounts. The first step is to choose the social network from which the data are extracted. In this study, we decide to study communication on Twitter for several reasons. The first of these is the accessibility and richness of data extraction offered by the Twitter API. The second is the study tools that Twitter's features such as retweets and replies provide, which are quite easy to identify and therefore analyse. The last one is the nature of these social networks, Facebook users are mainly present on the platform to maintain contacts with their acquaintances. Twitter, on the other hand, is used extensively to follow users interests, the media people they like, who are more often accessible on Twitter.

The next step is to define the scope of the study, whether in terms of time, quantity or actors examined. As far as the accounts studied are concerned, it is in the first instance that the Netflix France account is taken into account as it is this account that has pushed for this sector to be chosen as part of this thesis. In Europe, 3 services stand out from the others in terms of popularity: Netflix, PrimeVideo, Disney+. It is these three services that we study. in our case An expansion to more services has been considered as other services are doing well in Europe (HBO max, AppleTV+, OCS in French-speaking countries, ...). However, in view of the various analyses planned, the addition of other accounts would

certainly detract from the quality and depth of the analysis. The more elements to be examined, the more the analysis will remain on the surface. Therefore, the accounts of Netflix, PrimeVideo and Disney+ are the ones we study. However, we do not study only one account per brand but one English and one French account. This, in order to identify potential cultural and strategic differences within the same brand as already done by Scerbinina (2019) [7]. This choice is also a factor in the reason for keeping only the 3 largest players in the sector.

The study period is a fairly important element to determine. Depending on its duration, it can lead to certain objectives, such as firstly defining whether the study is carried out in the short term (defining a communication strategy based on one month's data is not necessarily representative of the overall strategy) or in the long term (a study over several years is very likely to be representative of the overall strategy). In the context of this work, a long duration is more interesting to study, particularly because of text-mining. Since text-mining is based on collected data, the larger the number, the more accurate and relevant the results will be. Therefore, to store as much data as possible, we need to extend the research period as much as possible according to our study desires. As studying the first years of communication in 2014 - 2015 - 2016 is not really relevant. It is certainly not the same way of communicates as today. The ideal would have been to study the last 3 to 4 years. However, Disney+ is a rather young service that landed in France in April 2020 and therefore only has a bit more of 2 years of usable data. In fact, in order to equalise each account as much as possible, the study period starts on 1 January 2020 and ends on 30 June 2022.

Concerning the different type of information extracted, it is important to specify that as our study only concerns the communication of brands in itself, we only extract tweets coming from these brands. In other words, we do not take into account tweets from users interacting with these brands (replies to tweets from these brands for example). Another precision is that all types of post are taken into account, i.e. we extract all the tweets, retweets and replies of each account. The first and most essential element extracted is of course the text of the tweet itself, which allows us to apply various text-mining techniques. Next, the date of creation of the tweet, which is an important element for studies of cycles, evolution over time and pattern identification. Finally, to complete this, metric data such as the number of retweets and likes that the tweet in question has obtained. Nevertheless, it is essential to add a few more in order to make the analysis as accurate as possible. However, we must first clean our datasets of data that could be harmful to our analyses.

3.3 Data Cleaning

Indeed, when extracting a large quantity of data such as this one, it is almost systematic that phenomena present within our datasets can disturb the analysis, such as outliers for example (abnormally low or high numerical data impacting in particular an average). In our case, During the exploration of the data we identified dates on which certain accounts made a lot of tweets compared to other days. Several of the accounts in the study are using automatic replies via a bot. This involves replying a pre-generated sentence to anyone who likes or replies to the initial tweet. The fact that an account uses this kind of process is interesting to note, especially concerning the definition of their strategy, but keeping all of these tweets would partially distort our analyses. Therefore, we decide to keep one copy of each pre-generated tweet and only one within our datasets. This allows us to take into account the content of these

sentences without them weighing too heavily on the whole of the tweets and therefore on the accuracy of our analysis. We favour "manual" communication within our study, that which is done directly by the community managers and therefore reflects the brand's communication philosophy. Two brand use this process, PrimeVideo and Disney+. Concerning Prime, this represents 71,42% and 46,94% (respectively ENG and FR) of all tweets. Disney US is even more dependant to it with 81,28% where DisneyPlusFR is only at 32,28%. It is therefore essential to remove them, otherwise our analyses would only give us results based on these pre-written sentences.

3.4 Features Extraction

In order for the analysis to be complete, we still need to add some tools to it. Indeed, the analysis will be heavily based on the type of post we are going to study, i.e. whether they are tweets, retweets or replies. However, we have no direct way in our dataset to indicate the type of post. We remedy this by creating binary variables for each post as to whether it is a retweet or not and a reply or not. This allows us to divide the datasets into 3 types of posts. Another binary variable created is the one indicating the presence of a video or a photo within a post allowing analysis on the impact of a visual element on the popularity of a tweet. The last data created are the variables containing the day, month and year of creation of the post via the division of the date of creation of the date initially extracted. This separation make it easier for evolution and pattern analysis.

3.5 Final Data

After all these steps, we count a number of posts of :

- **@NetflixFR** : 42406
- **@Netflix** : 14188
- **@PrimeVideoFR** : 8959
- **@PrimeVideo** : 13431
- **@DisneyPlusFR** : 2887
- **@DisneyPlus** : 5597

In comparison, the Scerbinina (2019) [7] study had just over 1000 for both accounts. Martinez-Sanchez and al. (2021)'s [4] study counted about 100 per account and per network. We therefore have much more data and our conclusions could be more reliable and more representative of a long-term communication plan.

Now that the data has been collected according to our wishes, that it has been cleaned and that we have added various elements that help us to develop text and data mining, we can move on to the analysis phase.

4 Analysis

Our analyses are divided in several parts. The first one consists of extracting some basic metrics such as the number of weekly posts as well as the percentage distribution of these numbers by type of posts (tweets, retweets, replies). This allows us to establish an initial profile for each account studied, both in terms of activity and presence (do they make a lot of tweets during the period studied?) and also in terms of communication intentions via the types of posts used. Does an account just tweet without responding to its users or does it focus on that?

In the same vein, we are studying the evolution of the number of posts over the months. This allows us to identify a change in policy in terms of intensity and activity on the social network. Was a brand passive 2 years ago but has completely changed and become very active in recent months? Secondly, to identify cycles over the years, whether specific to each brand or the sector as a whole. Ideally, if a cycle is confirmed, try to find potential reasons for it (a drop in activity as the summer holidays approach, an increase at the beginning of the school year, etc.)

Then, we study the content of tweets by identifying the most used words. Until now we have used simple data mining to extract purely numerical data. In this part, we start text-mining by looking at the content of the tweets, their substance. By determining the most used words we are able to determine the themes on which the brand's communication is most focused on. Is the most used word "available" and therefore the communication is more focused on the announcement of new content for example ?

As a continuation, in order to study the semantics of the tweets, we then want to establish categories within the tweets of each brand. These classes are created via a Latent Semantic Analysis (LSA) model that groups tweets according to the co-occurences of words within a document. We are then able to identify the 10 most used words in each category to better identify the topic for each category.

Once these content analyses are done, we look at the tweets deemed to be the most popular. We establish whether these are more of an announcement, humour, reactions to a societal subject, ... Above all, we try to establish a link with the most used words (are they present in this type of tweet) and the categories identified via the LSA (which categories are more likely to be successful). This not necessarily define the focus of brand communication but rather which types of tweets work best on which accounts.

As soon as that these analyses are complete, we can achieve the goals fixed by this thesis. Firstly, the profiling of each brand in detail, identify its specifities, its personal way of communicate on social medias. Then, the segmentation of communication strategies in the video streaming service sector. This is done via an adapted typology of communication strategies from Floreddu and Cabiddu (2016) [6]. We take into account the interactivity part of their typology (the fact that an account only tweets to give information about itself or on the contrary responds to users) by adding a classification axis of the order of passivity/aggressiveness. That is, the frequency and number of posts that a brand makes over a given period. It seems essential to us to take this aspect into account when developing a strategy model because it shows the extent to which the brand invests and relies on this marketing aspect within its overall strategy. We also add a distinction between brand-oriented and product oriented communication. In a third step, we study the differences in strategies between the English and French accounts of each brand. We identify whether a brand seems to have an international communication strategy or whether it is specific to each regional account. We also check these cultural differences within the tweets themselves.

In other words, does each account use socio-cultural phenomena specific to each region within its tweets (e.g., tweets on presidential elections)?

4.1 Metrics

4.1.1 Frequency and type of posts

As mentioned in the introduction to the analysis, we start with data-mining and more precisely with the extraction of several basic metrics. We have in fact already started in the methodology with the total number of posts collected for each brand. However, throughout this thesis we distinguish three types of posts: tweets, retweets and replies to tweets. The first indicator is therefore the number of posts of each type made on a weekly basis by each account studied.

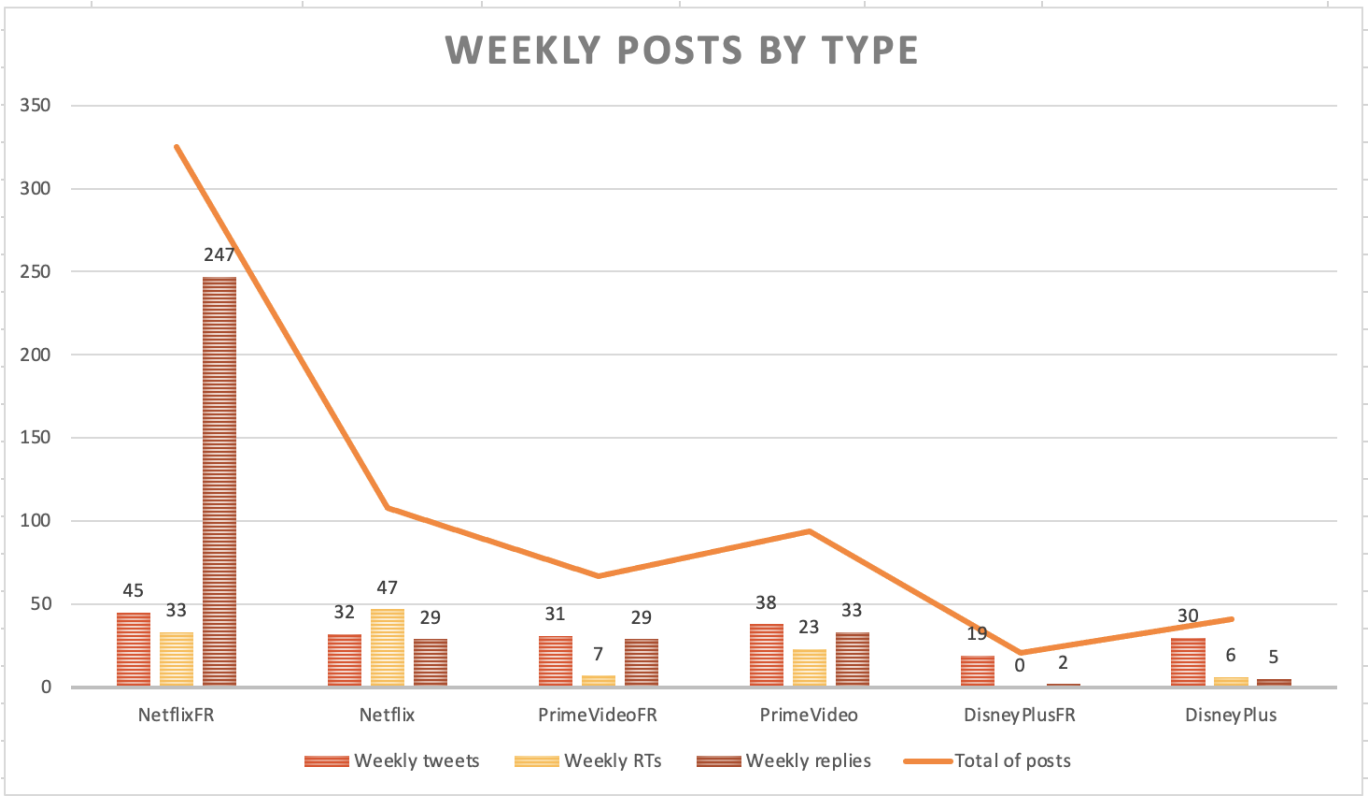

Figure 1: Weekly posts per type

We can already identify differences in terms of pure activity with Netflix France being 3 times more active than the second account which is its English counterpart. It seems that the 3 brands follow a ranking of activity both on the English and French side: Netflix in first, PrimeVideo in second and finally Disney each being quite far from the one ahead (except for PrimeVideo English). It can be seen that apart from Netflix France, the French accounts seem less active than their English counterparts. We can also identify the first strategic elements specific to each brand, such as the fact that the Disney brand as a whole focuses mainly on tweets, with even a non-use of retweets for the French account. For Prime and the American Netflix account there seems to be a certain balance between tweets and other

types of posts (note the low number of retweets from @PrimeVideoFR). As for NetflixFR, we note the enormous importance given to replies with a bit less than 250 replies per week. In order to confirm these observations and to have a better vision of each post use policy, here is a table representing the percentage of each type of post within the global dataset.

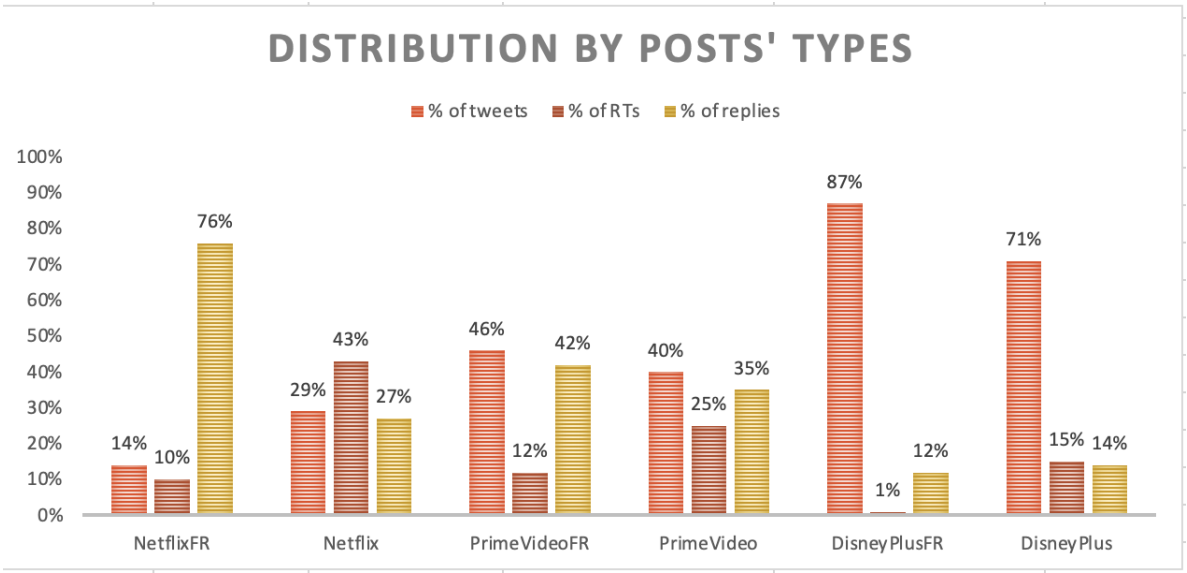

Figure 2: Distribution of posts types

This confirms our previous observations with a particular attraction on the side of Netflix France, a balance for the 3 following and finally a property towards the original tweets for Disney. It is important to note that despite the fact that the Netflix France account has "only" 14% of tweets and 10% of retweets it remains respectively 1st and 2nd in weekly frequency for these posts types. Also, it is important to note that Netflix US uses more retweets than original tweets. This means that their communication is based more on other users' tweets than their own. In summary, these figures should be put into perspective with the total number of posts for each account.

4.1.2 Followers count

The indication of the number of followers is certainly one of the essential metrics for brands. Their primary goal in their communication is to acquire more and more followers and to reach a majority of the population via their tweets and thus potentially acquire new customers.

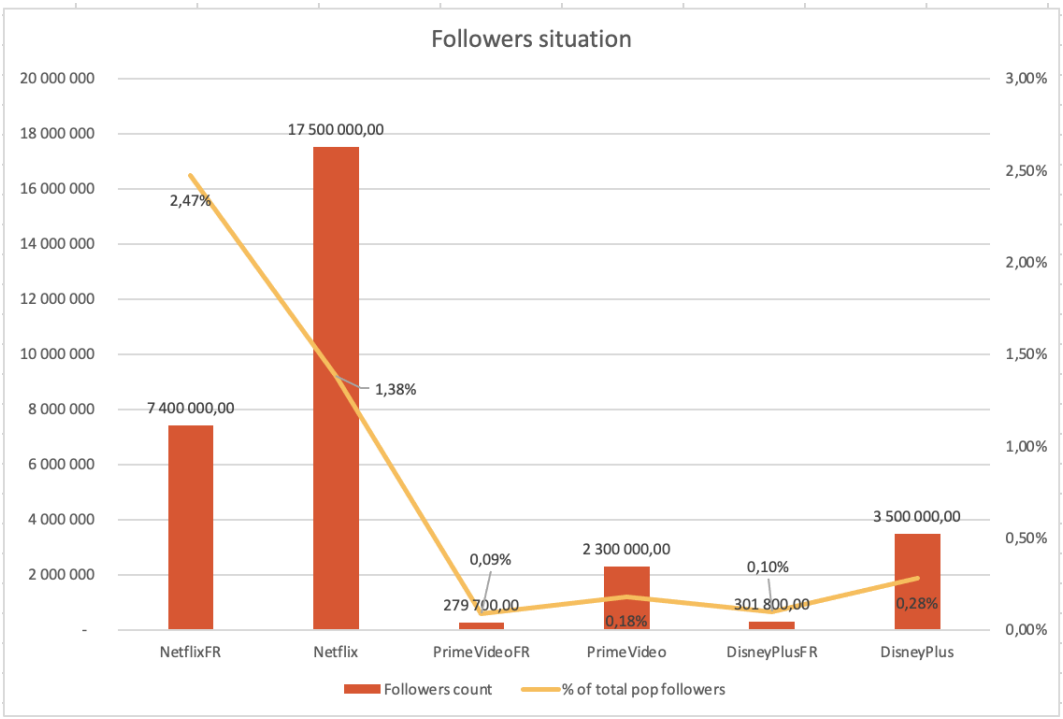

Figure 3: Followers count and reach on linguistic population

We first look at the number of followers in itself. We can notice that in this field Netflix, whether English or French, largely surpasses its competitors with 6 and 5 times more followers for Netflix US than its English-speaking competitors. NetflixFR even reaches more than 23 times the number of followers of its competitors, which is a gigantic gap. In fact, the number of followers of Netflix France is quite huge in view of its target which is the French-speaking population, we can see in particular that it exceeds the two English accounts of Dinsey and PrimeVideo (combined). It seems interesting to us therefore to put into perspective the number of followers in relation to the global population of French and English speakers The number of French and English speakers is 321 000 000 and 1 268 000 000 respectively . Once we divide the number of followers of each account by the total population for each language, we get quite interesting results. Indeed, we can see that the two brands that are lagging behind are PrimeVideo and Disney with a much lower French-speaking side than their English-speaking counterparts. As for Netflix, the result is totally opposite with a French account having a fairly important reach on the whole French-speaking population and a US Netflix quite far behind despite a fairly important score.

4.1.3 Multimedia content usage

Within the communication, we can distinguish 2 types of pure content : textual and multimedia elements. We focus our next metric on this second type. The table below indicates the proportion of tweets containing at least 1 picture or video. A picture or video inside a tweet is often a synonym to an announcement of either a release of a new series/season saying it is available or a teasing of an upcoming new series/season.

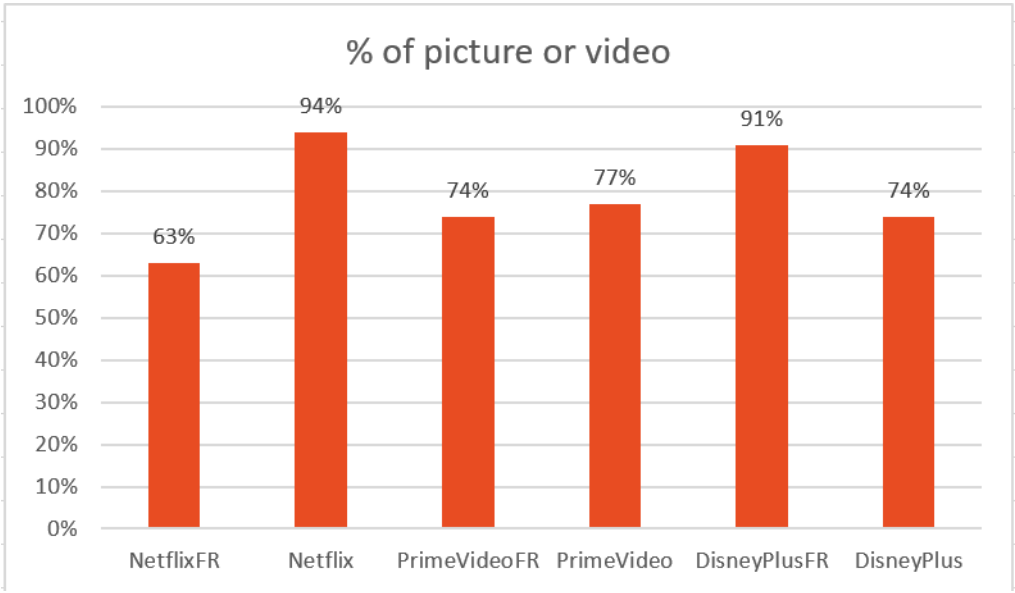

Figure 4: Percentage and number of tweets using a multimedia content

At a glance, we can see that overall the use of multimedia content is very popular within the sector with an average of 79% presence of photo or video. Some brands even exceed 90% such as Netflix US and DisneyPlusFR. One account seems to be particularly less committed to the omnipresence of this type of media within their communication in the presence of Netflix.

4.1.4 Popular tweets

One of the main goals of communication on social networks is to create a larger community every day. To do this, you need to gain more and more visibility and one of the ways to do this is to create tweets that go viral, that create a buzz. It is this kind of tweets that we are interested in with the following metric: the number of tweets with more than 10,000 likes. In addition, we are interested in the potential presence of multimedia within these popular tweets and their possible link to their success.

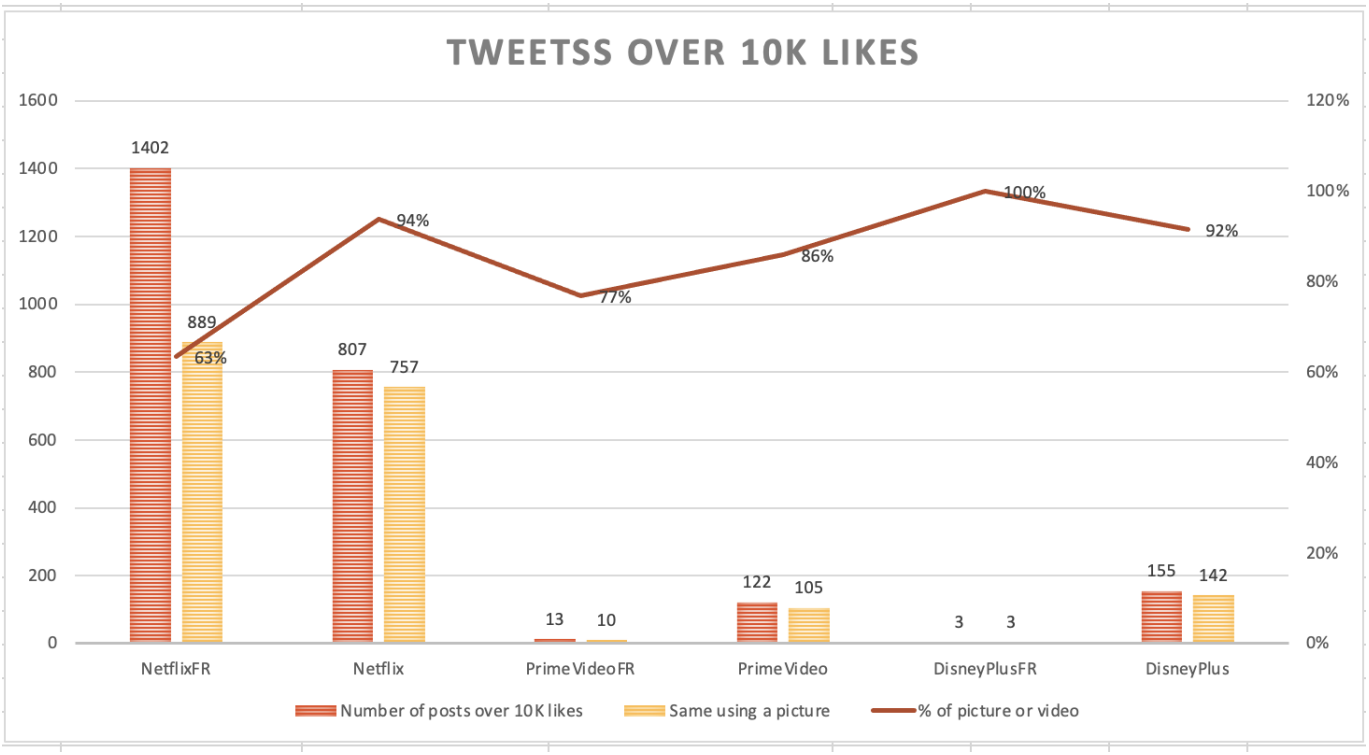

Figure 5: Number of tweets of having at least 10K likes and the percentage of it using multimedia content

We can already notice a rather marked difference between the statistics of the two Netflix accounts and their respective counterparts. Indeed, the latter, as for the other previous metrics, outperforms its opponents. Once again, the French account of the brand with the red N is largely ahead of its counterpart despite the fact that for the two other services it is the English accounts that are largely ahead. We can also see that these buzz tweets are highly dependent on multimedia content. Indeed, we can see in the table that the percentage of popular tweets using a photo or a video is quite high. In fact, on average, 85% of popular tweets use multimedia content. We can however note that once again NetflixFR seems less dependent on this kind of media.

4.1.5 Metrics - Summary

Overall, we were able to identify several interesting elements through our different metrics. Firstly, it is quite clear that Netflix has better popularity and presence statistics. Indeed, both accounts are well ahead of their respective opponents in terms of popularity (followers, reach on the language population, tweets with more than 10K likes). They are also ahead of their opponents in terms of frequency of tweets, especially the French account which has a huge number of weekly tweets. Talking about NetflixFR, we can notice that it gives a lot of importance to the interaction with its users via replies. Concerning PrimeVideo, the conclusions are quite similar for both accounts with a certain balance for the use of the types of posts. PrimeVideo seems to be the less popular account in French and English speaking countries (although PrimeVideoFR has more popular tweets than DisneyFR which is more recent). Concerning Disney, we note that both accounts bring a lot of importance to their own original tweets, more so for

DisneyFR. The latter also brings very little interest to retweets with only 1% of posts being retweets. Disney is still more popular than Prime despite being much more recent.

4.2 Evolution plots

In this section, we first look at the general trends (troughs and peaks of activity). We then quickly study the individual evolutions for each account in order to identify potential adaptations (increase or decrease of post frequency).

4.2.1 General trend

We can observe an overall decrease in the number of posts within the sector from early 2020 to early 2022. This is largely explained by the influence of NetflixFR which appears to have reduced its number of posts since June 2020. Given that, as mentioned earlier, this account is by far the most active, a drop in its numbers greatly impacts those of the entire sector. Other accounts were also quite active during this period such as Netflix (US) and the 2 Disney accounts. This is explained in particular by the global containment in progress during this period. Indeed, as everyone was stuck at home, the consumption time of content on these platforms exploded as explained C. Labrovitz (2020). So the platforms communicated a lot via social networks in order to attract even more people, who had more time to spend than before. A second element specific to DisneyPlus also plays a part in their fairly large number of tweets, namely the launch of the service in each region. The service launched on 19 November 2019 in the US and 7 April 2020 in France. It is therefore logical that communication is quite extensive at the beginning of the platform's life.

Concerning peak and off-peak periods, it would seem that summer, and more specifically the month of June, is a less favourable period for aggressive communication. On the contrary, the back-to-school periods (September, October) as well as the end of year festivities (despite January 2021 being at its lowest) seem to be the times chosen by the brands to increase the pace. This is certainly correlated to the different releases that the platforms can make. Indeed, although we have not found any figures on this, we can imagine that platforms release more content during the end-of-year holidays than during the summer. These cold periods are more favourable to the consumption of this type of content than the summer period.

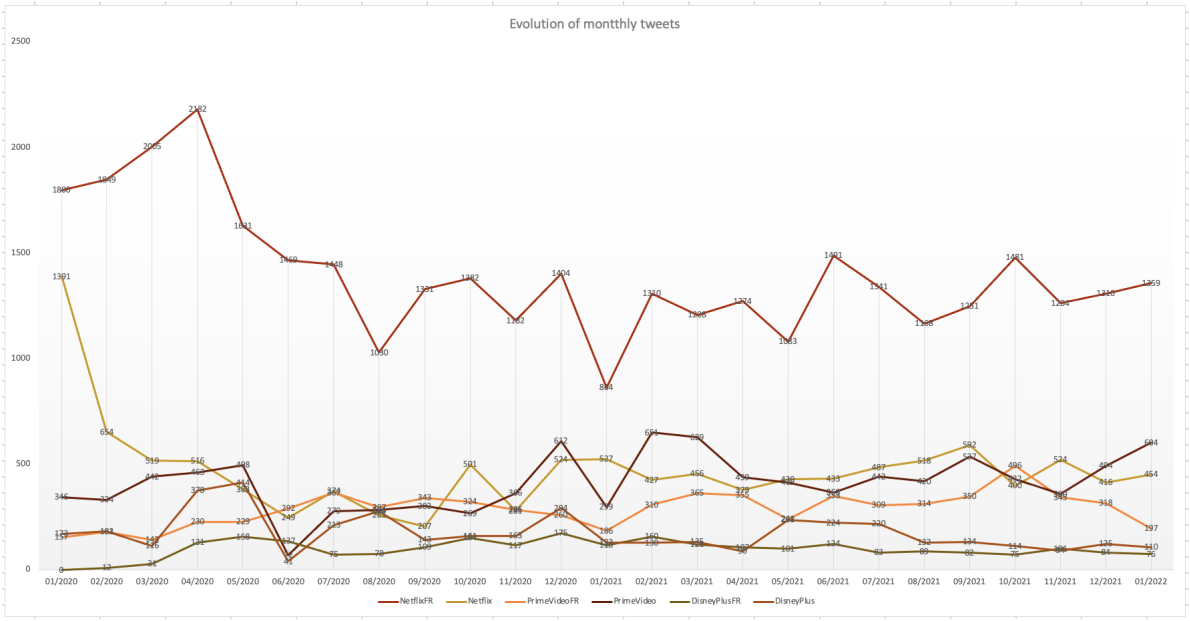

Figure 6: Evolution of posts over last 2 and half years

4.2.2 NetflixFR - Figure 21

As mentioned previously, NetflixFR was particularly active during the global lockdown that lasted from March to June 2020. After this period, we can note a certain stability with between 1000 and 1400 posts per month (except January 2021 with 800 posts which seems to be an anomaly). If we put the containment period aside as a rather exceptional situation, there has been no real growth or decline over time. At NetflixFR, August and November seem to be the regular low periods. While June, October and December seem to be the regular peaks.

4.2.3 Netflix - Figure 22

NetflixUS also appears to be fairly stable despite a slight drop in the summer of 2020 and a record number of tweets in January of that year. Over the period studied, it is quite complicated to establish peak and trough periods. There does not appear to be any real rise or fall over time, apart from the anomalous January 2020.

4.2.4 PrimeVideoFR - Figure 23

In the case of PrimeVideoFR, it seems that over the period studied we can identify an overall growth in the number of posts emitted. The month of January seems to be a fairly slow period for this account.

4.2.5 PrimeVideo - Figure 24

Compared to previous years, the curve of posts is a little more unstable depending on the month. Like its French counterpart, PrimeVideo seems to have increased its posting frequency over the last two years. June could be identified as a low point, with a huge drop in 2020 and a slight drop in 2021.

4.2.6 DisneyPlusFR - Figure 25

We find an even more unstable curve than the previous one with strong variations from month to month. We can also identify a general decrease which could be explained by the first anniversary of the platform (installed therefore less need to be known via social networks). It seems that the summer is a low period for this account but it is quite complex to clearly identify a repeated peak over the period studied.

4.2.7 DisneyPlus - Figure 26

Like its French counterpart, DisneyPlus has a curve that varies enormously from one month to the next and seems to experience a fairly significant decrease over the period studied.The only identifiable peak in the period is in May.

4.2.8 Evolution - Summary

The three brands seems to have a common evolution over the time studied for their both accounts. Netflix does not variate a lot between the after lockdown period and the middle of 2022. Prime have a slight increase over the period studied and finally Disney is really irregular while having a slight decrease.

4.3 Top words used

In this section, we look at the top 20 words used by each account to identify recurring content within the tweets. We will identify whether these words are more focused on promoting content or interacting with the user or otherwise.

4.3.1 NetflixFR

The first thing that can be noticed is that the vast majority of these words are more appropriate for content advertising. It seems that in the vast majority of tweets from NetflixFR (to be distinguished from posts which are all possible publications on Twitter : tweets, retweets, reply) contain promotion for their platform with the arrival of a new program in the catalogue. This is represented by most of the words present in this table such as: "dispo", "demain", "arrive", "disponible", ... Also noteworthy is the presence of "9h01" which is the release time for original content produced by Netflix. A rather important element is present in this graph but difficult to identify without having seen the other graphs. It is the word most used by Netflix France which is "saison". But you will see in the next graphs that for every other account the most used word is the name or part of the name of the service. This should be seen in the context of NetflixFR's overall use of tweets, which represents "only" 12% while the replies represent 79%.

Looking at each word within this graph, we notice the presence of the word "dispo" which seems to stand out. Indeed, this word is not as formal as all the others present. In fact, you will see that it is the only one in the set of 6 graphs representing each account that can be designated as informal. Moreover, this one is far from being last in the ranking since it is the second most used word by Netflix France community managers. This could possibly suggest a desire to make their communication less formal,

more transparent and that the user does not have the impression of reading tweets from a multinational but from a Twitter user like any other.

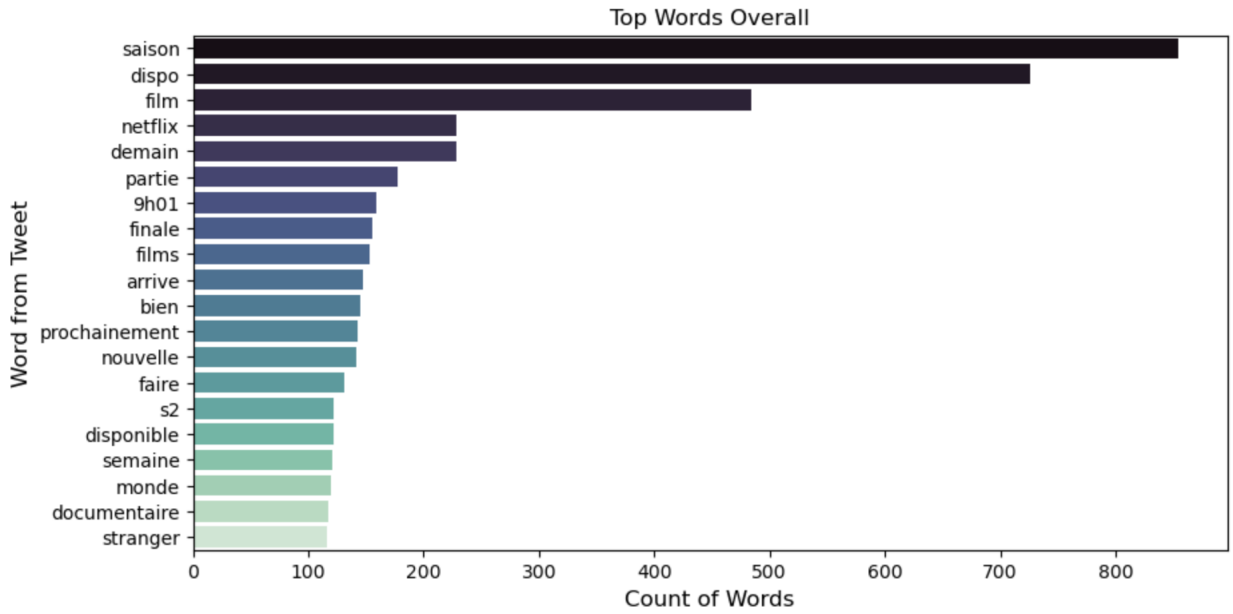

Figure 7: Top words of NetflixFR

4.3.2 Netflix

The English account also seems to be very focused on promoting content with the presence of words such as: "new", "premieres", "watch", "today", ... As mentioned earlier, by far the most used word is the trademark itself.This should also be seen in the context of the percentage that tweets represent within the communication of Netflix US. Indeed, tweets represent 28% for 42% of retweets. This means that the communication is based more on the tweets of others than on original tweets.

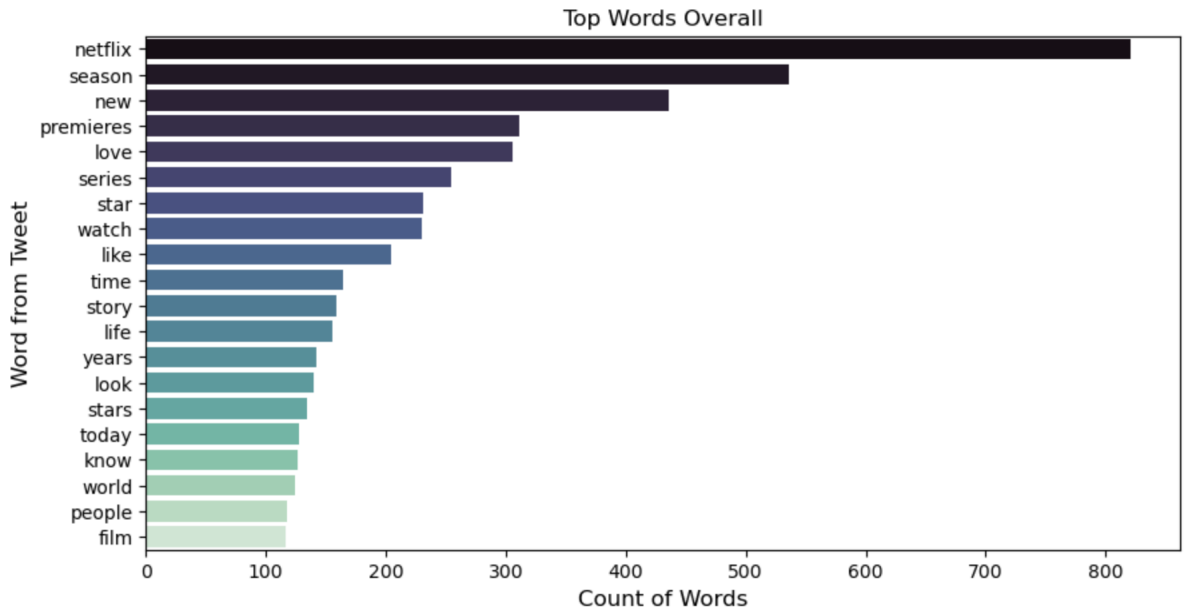

Figure 8: Top words of Netflix

4.3.3 PrimeVideoFR

The same is true for the French account of Amazon PrimeVideo, with tweets focusing on the release of new products. This can be seen through the presence of the words : "disponible", "nouvelle", "découvrir", ... A rather interesting element to note compared to previous and most subsequent accounts is the position of the word "film". Indeed, it is in 3rd position before a word like "saison" which rather refer to series. We can therefore make the hypothesis that PrimeVideo communicates more about its films than its series, which potentially reflects their content production and acquisition strategy as well. This is to be verified with the analysis of the brand's American account.

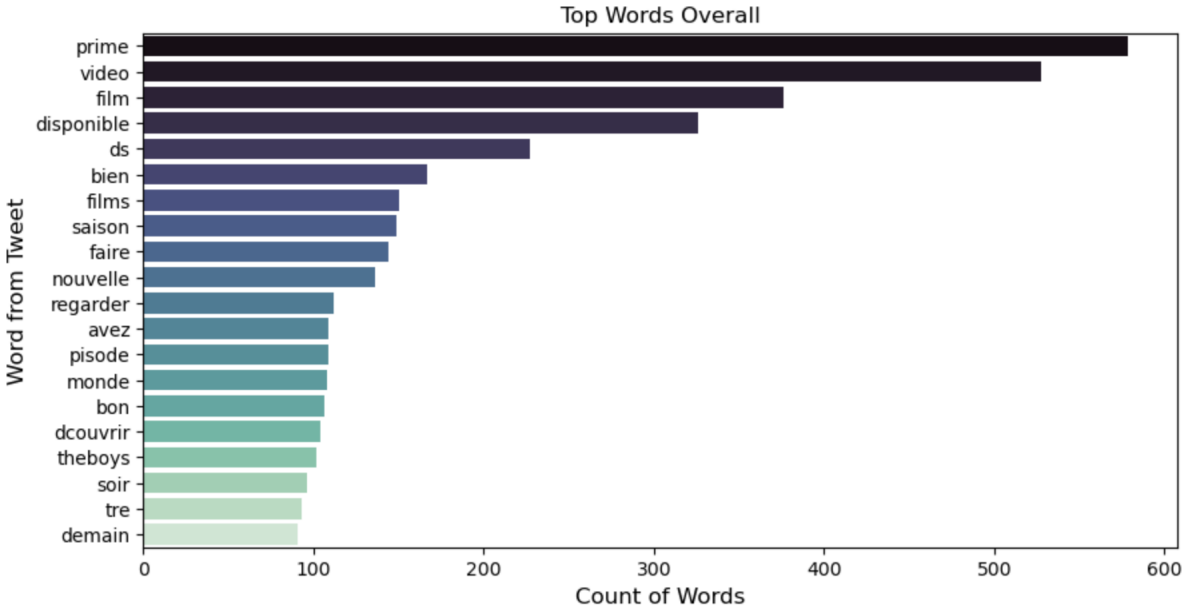

Figure 9: Top words of PrimeVideoFR

4.3.4 PrimeVideo

Once again, we find words with promotional purposes such as: "watch", "new", "coming", ... However, for the first time since this part of the analysis, we find words with an interactive connotation. More precisely, these are the words "youre" and "favourite". The first is a word that suggests that PrimeVideo is directly challenging its users, while the other suggests questions about its users' favourite content (e.g. "What's your favourite film?"). Speaking of films, the phenomenon we identified earlier in the French account seems to be confirmed. The word "movie" is in 4th position (and "film" in 12th position), ahead of "season" and no other word referring to a series such as "episode" or "series" is present.

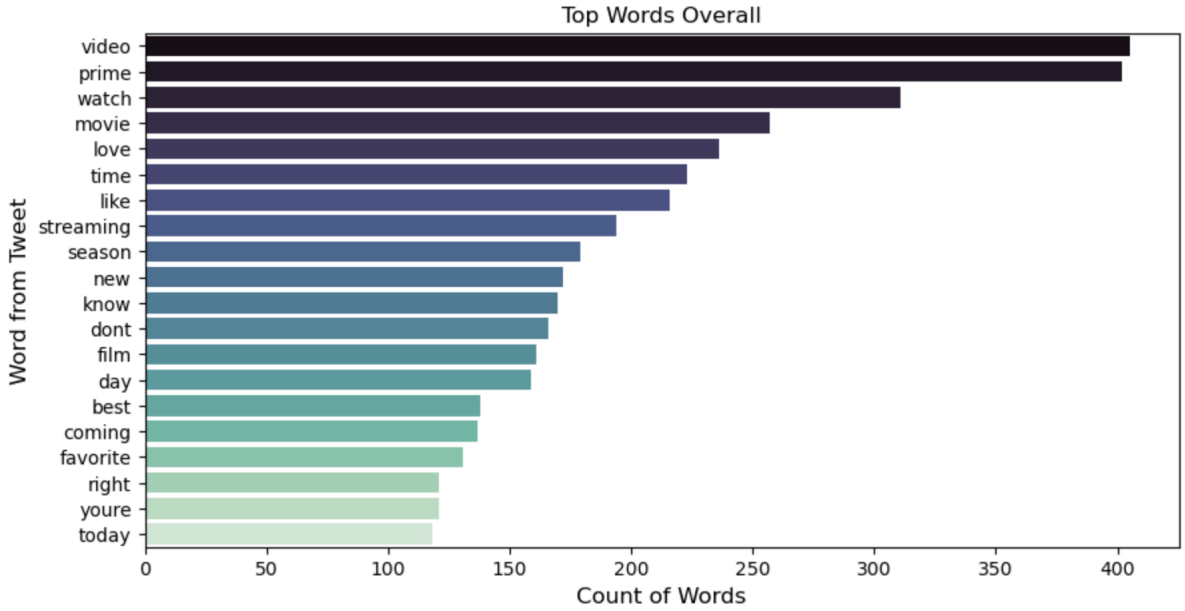

Figure 10: Top words of PrimeVideo

4.3.5 DisneyPlusFR

As before, the main theme remains the highlighting of new content on the platform via words like : "disponible", "retrouvez", découvrez", ... There are several major differences. The first is the frequency of top words. Until now, the most recurrent word appeared between 400 and 850 times. Here, this word appears more than 2000 times. The same can be said for the other words in the ranking (2nd, 3rd, ...). This is quite strange when you know that DisneyPlusFR is the account with the fewest posts and proportionally the fewest tweets. This means that within a tweet, certain words are used extremely frequently. Some of these words are actually hashtags, such as the word "disneyplus" which appears in almost every tweet. Another difference is the important presence of brand references. Indeed, we have already identified that a general trend is that the most used word is the name of the service. However, in other cases this stopped there, here in addition to the name "disneyplus" are the words "disney", "star", "disneyplusstar" which are different types of product of the brand.

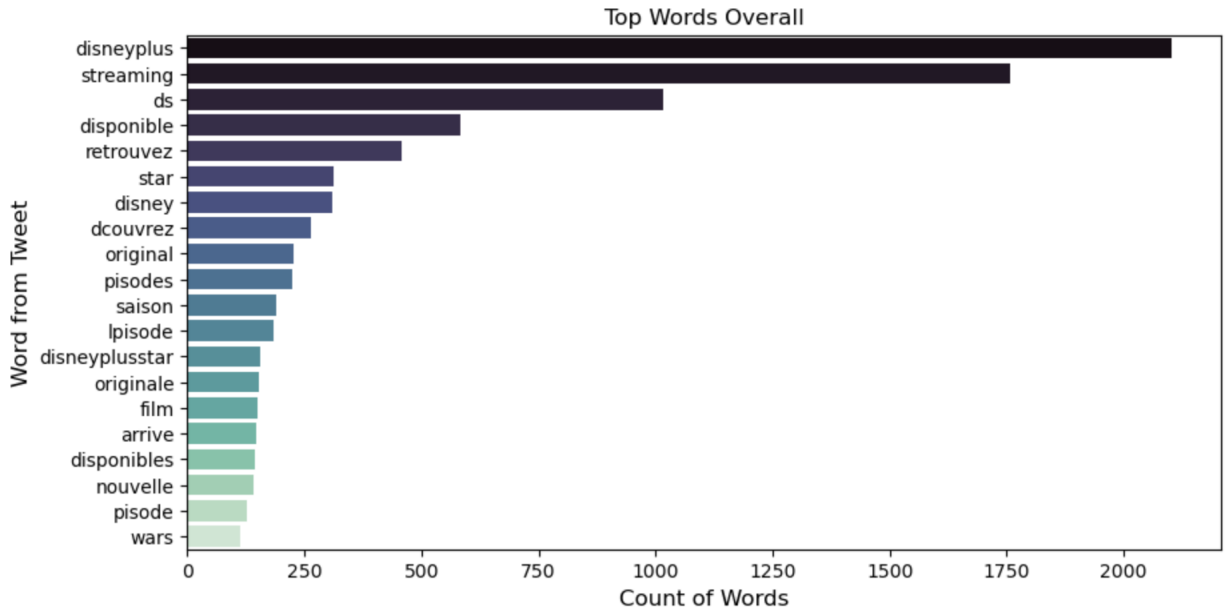

Figure 11: Top words of DisneyPlusFR

4.3.6 DisneyPlus

Overall, the same trends can be observed for the US account. As usual, the emphasis of the communication is on announcements. The frequency of top words is also higher than for the other brands with the first one at more than 2500 words again. Given that this account is second to last in terms of posts made, we can conclude the same thing before, i.e. words that come up almost every tweet like hashtags for example. Also like its French counterpart, DisneyPlus US integrates a lot of references to its brand and products via words like: "marvelstudios", "marvel", "studios", "disney" in addition to the word "disneyplus" which is first in the ranking and name of the service.

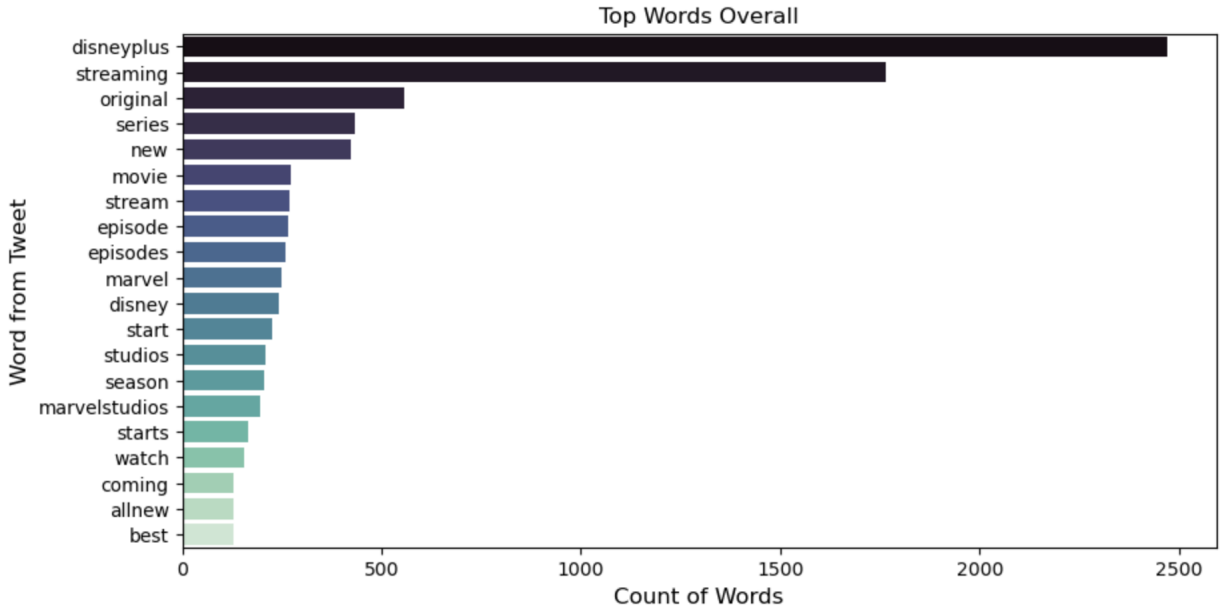

Figure 12: Top words of DisneyPlus

4.3.7 Top Words - Summary

It is clear that all accounts communicate extensively about the release of content and the promotion of it. That's the big thing they all have in common, most of their most used words are about this type of tweet. However, we can notice some specificities in some and the lack of originality in others.

NetflixFR stands out in particular by the presence of the word "dispo" showing its desire to get closer to its users and to use a familiar language for this purpose.

Netflix US doesn't stand out, using only common words to all the other accounts.

The two Prime accounts are also fairly traditional, although it should be noted that they focus on films rather than series. The English account also seems to ask questions of its users.

As for the Disney accounts, they seem to be very stereotypical. Indeed, the number of occurrences of their top words are huge compared to the others despite being the least active. They should therefore have a certain pre-made typology that they reuse a lot. They also put a lot of emphasis on the brand itself rather than on its products by using words like "Disney", "DisneyPlus", "Marvel", ...

4.4 LSA categorisation

In this section, we distinguish different categories of tweets within each brand's communication. To do this, we use a Latent Semantic Analysis (LSA) model which classify these tweets into a category according to the number of classes previously defined. It is important to know that this model is not initially intended for this type of use. It is usually used to identify broad themes within documents. For example, the LSA is perfect for identifying which scientific articles deal with economics, science,

psychology, etc. It is therefore potentially less efficient with the use we make of it, but the results are still satisfactory.

Our goal in this section is first to classify the tweets as mentioned above but also to link these categories to our previous analyses. We therefore try to identify within the different categories potentially top words identified in the previous step or to note a category of tweet showing interactions already studied in the Metrics section.

We also identify the most recurrent words within each category. We use these to define the nature of each class and thus find its meaning, purpose, ... Is it a category of promotional tweets or rather tweets that promote interaction? In our graphs we only display the top 3 words for aesthetic reasons, but you should know that we use the top 10 words to identify the nature of the tweets in the category.

4.4.1 LSA definition

As said previously, LSA is designed to classify a series of documents within different topics. It works via two main steps: the creation of a Document-Term Matrix and the Single Value Decomposition.

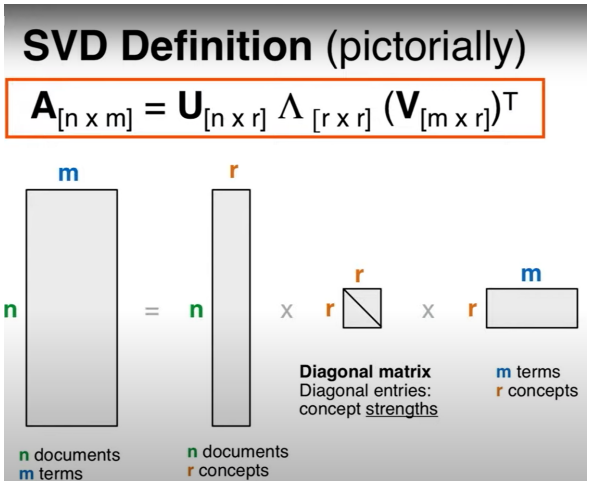

Figure 13: Mathematical representation of LSA by https://www.youtube.com/watch?v=M1duqgg8-IM

The Document-Term Matrix is built via a simple principle. The lines represent the documents studied. Here, a tweet, with for example "Have a nice day". The columns represent all the terms present in the different documents. In our example, column 1 represents the term "Have", column 2 "a", column 3 "nice" and column 4 "day". Then we find in the matrix the number of occurrences of each term in the document.

Via this matrix we make an SVD which allows us to obtain 3 new matrices. We keep the first two for the following. The first one is the matrix showing the link between a document and a topic which will help us to define in which topic we should classify such documents. But before that, we need to multiply it with the second matrix which is the strength of the topics. It represents the most distinct topics identified in the first place. This means that if two topics are completely different and do not use any words in common, they will be given priority in this matrix. The more topics are added, the less we are expected to find topics that are totally different from each other. Once the number of topics has been defined, the two matrices must be multiplied to identify whether a document belongs to a topic.

4.4.2 NetflixFR

NetflixFR	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	dispo	cest	toujours	saison	film	tout	série	bien	déjà	documentaire	1273
Topic 2	saison	demain	partie	finale	prochainement	academy	umbrella	juin	lucifer	education	961
Topic 3	demain	aujourd'hui	épisodes	dance	act	booth	kissing	trailer	dernier	cyril	81
Topic 4	quoi	regarde	soir	weekend	dimanche	minuit	matin	ca	dès	hier	173
Topic 5	série	film	ca	plus	si	tout	cette	fait	si	bien	2634
Topic 6	c'est	ici	rappel	dispooooo	mois	dispooooooo	disponible	juste	qu'on	daredevil	128
Topic 7	bon	tous	bonjour	weekend	diimanche	merci	surtout	courage	dispooooo	souhaite	242
Topic 8	stranger	thing	volume	love	death	robot	mai	vol	juillet	épisode	196
Topic 9	netflix	note	seul	séries	prochainement	tv	illimité	tarif	film	blinder	158

Figure 14: LSA of NetflixFR

NetflixFR has the most categories in the accounts studied as they have different themes evoked quite interesting for our thesis. Indeed, it has 9 categories with 3 dominating others, the topics 1, 2 and 5 which are related to promotion of content. In these categories we can observe some words as "saison", "dispo", "finale", "film" that are present in the top words ranking which is quite normal as these categories are really above the other ones. Some other minor topics are also related to this as topic 6 or topic 8. It confirms that most of the tweets made by NetflixFR concerns their series, releases and so on.

However, we can identify some interesting categories not linked at all with new content. It is the case for topic 4 and 7. Indeed, these classes seem to represent 2 recurrent tweet that Netflix France posts one which would certainly be something like "Que regardez vous ce soir ?" or one of its variants and the second which certainly looks like this: "Bon weekend à tous !". These typical tweets demonstrate the interaction the account wants to establish with its users via 2 slightly different means. The first challenges and questions the reader directly and therefore encourages a response from the user. The latter then feels listened to, heard by the brand, all the more so if it responds in turn (which is the case of NetflixFR as we have seen in the Metrics section). The second invites less response but more attachment. Indeed, wishing something as banal as a good weekend may seem at first sight useless for a brand for marketing purposes. But by accumulating this kind of tweet bringing attention to the followers it strengthens the link between the brand and the customer and therefore the loyalty. These two categories show that interaction is indeed a crucial element of NetflixFR's strategy by creating tweets that incite a reaction on which the account itself can react in turn.

Another somewhat surprising topic is the number 9s. This one seems to be about platform pricing and uniqueness. Looking at the top 10 most used words we can see words like "illimité", "tv", "séries", "film". So this seems to be a more branded type of promotional tweet like "Seul sur Netflix vous pourrez retouver toutes vos séries et films au tarif de 13,99€ par mois". We aslo find the word "note" which maybe suggests that they ask feedback for the quality of their contents.

4.4.3 Netflix

Netflix	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	season	netflix	new	star	premiere	&	look	series	time	know	3039
Topic 2	aisha	cole	floor	flow	flourish	florida	florence	flora	flor	flop	1
Topic 3	love	blind	nelflix	story	kill	boy	heartstopper	account	premiere	true	263
Topic 4	today	year	like	thing	ago	day	stranger	happy	watching	perfect	388
Topic 5	premiere	heist	la	casa	money	papel	end	ready	april	february	175
Topic 6	hour	right	watch	episode	academy	question	umbrella	ozark	drop	got	191
Topic 7	thing	stranger	ozark	epsiode	final	streaming	howozarkends	season	vol	ok	105

Figure 15: LSA of Netflix

The English account has a category that largely dominates the others and this also concerns announcements, releases. The only difference with the French account seems to be the presence of the brand name within the 3 most used words. This could be related to the fact that the account communicates a lot about Netflix original content, the programs they produce and broadcast exclusively. This also seems to be consistent with what we identified in the previous analysis section. Indeed, this category, which is much more represented, includes the 3 most present words in this category, which are the 3 top words for all tweets combined that we observed previously.

The other categories also seem to be related to the content of the platform, such as Topic 3, 5 and 7 referring respectively to "Love is blind", "La Casa de Papel" and "Stranger Things", all of which are popular Netflix original programs, which supports our previous observation. This is also the case for topic 6 talking about "Umbrella Academy" and "Ozark".

Class number 4 seems to be a type of recurrent tweet (2nd most present category) celebrating the anniversaries of series by reminding that a content was released such a number of years ago. Again we can link this to promotion even if it is less obvious and certainly less effective than direct announcements of current series. It also contains words identified as top words like "years" and "today". So it seems that this is indeed a habit that the platform's community managers have taken on, bringing older programs back into the light.

Topic 2 seems to be an anomaly as it contains only 1 tweet containing weird words with a lot beginning by "flo...". We can then assume that thi was a special twweet made for a challenge or something like that and that this tweet uses a lot of unique words not used in other tweets. This could explain it is a topic by itself.

4.4.4 PrimeVideoFR

PrimeVideoFR	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	prime	video	plus	série	ca	film	tout	si	cette	saison	3360
Topic 2	lesanneauxdepouvoi	leseigneurdesanneau	anneaux	lotronprime	bluffant	seigneur	apercu	casquette	lenvers	deuxième	41
Topic 3	freedbritney	toternstopsecretmp	voxographies	fermées	festivalcannes	festival	fervante	ferragni	ferragnez	ferontelles	3
Topic 4	cest	theboys	demain	maintenaint	ca	quon	therealfakelive	theterminallist	épisode	lheure	263
Topic 5	disponible	maintenant	dès	nouveau	lintégrale	original	film	série	famille	nouvelle	236
Topic 6	film	quel	acteur	titre	préféré	regarder	règlement	ditesnous	noël	complet	235

Figure 16: LSA of PrimeVideoFR

Once again, one category seems to be well ahead of the others in terms of presence. However, this time the category is slightly different from the content announcement. Indeed, this is more in topic 5. In fact, topic 1 promotes the brand as a whole and not a specific type of content. It seems that the priority of Amazon's service is the strength of the brand and not its products. This is in line with what we had previously identified in the top words as the top 2 of the ranking "prime" and "video" (i.e. the brand name) are found in this category alongside the word "plus". We don't find words as "disponible", "maintenant" synonyms of release of a content. This typology of tweet is then more axed on promote the brand via a tweet like "Retrouvez toutes vos séries, films et encore plus sur Prime Video".

It is however important to note the presence of topics 2 and 4 referring to two franchises that are hugely popular at Prime, The Boys with topic 4 and The Lord Of The Rings via topic 2. This shows that although the brand seems to be at the centre of their communication, the products are not put forward,

simply less so. Especially since topic 2 concerns The Lord Of The Rings but more precisely a series that is not yet available, which will be available on September 2nd. We can therefore imagine that the communication on this topic will increase.

Topic 6, on the other hand, seems to represent tweets specific to the interaction with the followers of the account, asking in particular which films they prefer, but also the actors or the series if we observe the top 10 words. In fact, this is not really surprising since Prime's French account is the second in terms of percentage of replies according to our metrics. This would mean that the account is encouraging users to reply to the ones they are going to reply to themselves.

We do not consider topic 3 as relevant because it is not sufficiently represented.

4.4.5 PrimeVideo

PrimeVideo	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	video	prime	watch	streaming	coming	new	look	arrives	episode	pm	3597
Topic 2	movie	time	favorite	best	like	know	scene	right	film	tv	417
Topic 3	rocketman	themarvelousmrsmaizombieland		matrix	thewildthornberrysm	thewildthornberrys	savingface	oi	moneyball	onemississippi	19
Topic 4	love	selahwatchparty	sylvie	modern	sylvies	lizzo	black	beingthericardos	mess	fall	289
Topic 5	day	today	year	happy	birthday	great	photo	turn	mood	shrek	479
Topic 6	season	happy	theboystv	boy	Friday	primevideo	episode	outerrange	ready	Monday	529

Figure 17: LSA of PrimeVideo

The trend noticed for the French account seems to be confirmed in part for the American account. Indeed, within the topic being largely dominant we still find the brand via the 2 words "prime" and "video". This still refers to a communication focused on the brand. Whereas product-oriented communication, represented by topic 6, is much less present than the other. Once again, the observations made for the top words are confirmed as in this category the 3 most used words are also the most used words for all tweets combined.

We also find a category almost identical to one of the French-speaking account, topic 2, which reflects the interaction by asking for their preference in terms of films, programs, etc. This time, however, we can find the words of the category within the top words of the account. Indeed, the words "favorite" and "movie" are both in the ranking, which was only the case for the word "film" for the French account. This was not enough to affirm that it was linked.

Then we also find elements identified in top words in the presence of topic 4 referring to watch parties on Twitch. Indeed, Prime uses influencers a lot in their communication via watch parties, i.e. broadcasting audiovisual content via a community platform such as Twitch (also owned by Amazon) in order to watch a content with hundred of thousands of other persons simultaneously. We therefore find this in the Twitter communication via hashtags such as "#selahwatchparty". It should be noted that the French account also does this via French influencers but this looks not sufficiently reflected in their tweets.

As for topic 2, despite the fact that it is not very informative, it shows us the use of hashtags in communication, which has not always been the case until now.

4.4.6 DisneyPlusFR

DisneyFR	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	disneyplus	streaming	dès	maintenant	série	disponible	retrouvez	disney	découvrez	original	1867
Topic 2	star	bacth	bad	war	l'épisode	disneyplusstar	disponible	disneyplus	streaming	saisons	74
Topic 3	tous	épisodes	retrouvez	disponibles	streaming	disneyplus	programme	premier	deux	encore	169
Topic 4	war	star	bad	batch	disney	arrive	disneyplus	mai	original	série	79
Topic 5	wandavision	épisode	marvel	soldat	falcon	disneeyplus	nouvel	streaming	moon	studio	87
Topic 6	disponible	streaming	disneyplus	soul	règlement	exclusivement	paraèles	parallèles	décembre	sort	246

Figure 18: LSA of DisneyPlusFR

As far as DisneyPlusFR is concerned, globally all the categories concern the promotion of content but also of the brand itself. Here, unlike Prime, the two seem to be mixed. Some categories are similar with topics 1 and 2 concerning the whole service while the others concern specific content like WandaVision, Star Wars The Bad Batch, The Mandalorian, ... This account is also no exception to the rule regarding the link between topic 1 and top words. Indeed, the 3 words of the category are the 3 top words of the service. It is all the more logical given the number of appearances of the top words more important than for the other brands. This is in line with our hypothesis that certain words are used extremely often and that there is a certain format defined at Disney.

This lack of diversity in the categories seems to be a reflection of the same lack of diversity in the communication of the account. In other words, the only way the model manages to find new categories is to divide the ads according to the series concerned by its ads.

The presence of Disney's brand elements is noteworthy. We find the words "disneyplus" referring directly to the service and also "disney" referring to both the brand and some of the available content produced directly by the platform.

Moreover, the word "disneyplus" is in the top 10 words of all of the categories. It prooves that whatever the theme of the tweet the service include its name and promote its brand at any occasion.

4.4.7 DisneyPlus

Disney	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10	Number of posts
Topic 1	disneyplus	streaming	original	start	series	episode	new	movie	season	disney	2808
Topic 2	marvel	studio	moonknight	disneyplus	streaming	episode	wandavision	multiverse	making	assembled	136
Topic 3	season	episode	therightstuffseries	themandalorian	war	star	new	summerofdisneyplus	final	hsmtmts	157
Topic 4	disney	access	premier	mulan	stream	required	fee	additional	exclusively	info	329
Topic 5	wandavision	marvelstudios	lokiwednesdays	loki	episode	avenger	disneyplus	streaming	monstersatwork	time	195
Topic 6	disneybundle	hulu	espnplus	stream	hodwarddoc	best	favorite	theprincessbride	disneyplus	congratulation	367

Figure 19: LSA of DisneyPlus

The observation is broadly the same for the American account. One part of the categories refers to the communication of the whole service and the other part to specific series. It may be interesting to note that the series mentioned in the classifications of the two accounts are not totally identical. This also applies to the link with the top words, the situation is the same (i.e. the 3 words of topic 1 are the 3 top words).

The emphasis on brand-specific elements is as pronounced as for the French account. In addition to "disneyplus" (that is a bit less used than in France) and "disney" that we had already seen with the

previous account, we find "marvel", "marvelstudios".

4.4.8 LSA - Summary

Overall, it is clear that all accounts have different categories of tweets used for promotion and are often the most represented. However, these categories may differ slightly from account to account. Indeed, some accounts such as those of Netflix issue more communication about their product, either in a general way by announcing an availability or by directly quoting the name of a series.

We also find tweets favourable to the exchange between user accounts, notably at NetflixFR and Prime accounts.

We also identified an almost omnipresence of the DisneyPlus brand within their tweets, which seems to come first.

Finally, some brands have very specific categories, such as watch parties at PrimeVideo or the mention of pricing and quality at NetflixFR.

4.5 Most popular tweets

This is our final stage of analysis. So far we have used quantitative analysis techniques, including data and text mining. This stage is different as we rely more on qualitative analysis with human perception which is therefore more subjective. We first determine the 10 most popular tweets of each account via the number of likes. We only take into account likes because if retweets are also counted the results are partially distorted via contests with the condition of retweeting. Once these tweets are extracted, we analyse qualitatively the recurrences on the typology of these tweets: announcements, humour, social topics, ... We also make links with previous analyses, notably with top words and categorisation: do these tweets use top words and/or can they be attached to a category?

4.5.1 NetflixFR - Appendix 7.2.1

The first thing noticed when looking at all of the most popular tweets from NetflixFR is how few photos or videos are present. Actually, only 4 of the tweets contain audiovisual material. In comparison, except for one account, the others have between 7 and 9 tweets with this type of media. We had already observed this phenomenon in the Metrics section on the number of tweets with more than 10000 likes. This seems to confirm that the French account of Netflix seems to need less visual elements in order to make important statistics.

Concerning the background of the tweets in question, we can find a mix of different types of tweets.

The first are references to Netflix original programs with tweets n° 1 and 3 which are respectively references to Stranger Things and Kogustaki Mucize a Turkish film. It turns out that the tweet with the most likes is about Stranger Things, the most popular series of the moment, which broke all records in streaming. In fact, this tweet rose to the top extremely quickly as it was dated 20 June 2022 and the study period ends on 30 June of the same year. So it took only 10 days to be the most popular tweet.

The second type of tweet is of course the announcement of new content, releases, ... We find 4 of them in the 10 in the presence of the top 2, 5, 7 and 9. Note that the top 2 and 5 are linked. We have seen that most of the tweets issued were about announcements, so it is logical that this category is the majority of the most popular posts. We therefore find a category identified in the LSA stage but also some top words like "9h01" typical to this account because this is the time of release in France of programs produced by Netflix. We also find words like "demain" or "partie".

The third and more subjective theme is humour in various forms. The numbers of the tweets in this category are tweets 4, 6, 8, 10. In order to better understand the humour in these tweets, we quickly clarify some elements. Tweet number 4 is actually a tweet quoting another tweet that has since been deleted. This tweet came from an account posing as NetflixFR and announcing a documentary on the singer Wejdene. Netflix then denied the claim, citing a popular clip of the singer's music. Tweet number 6 is quite special as it concerns the arrival of an equally special Japanese program featuring guinea pigs being cars. I decided to put it in this section because the French account itself seems to be taking it in second degree, given the text accompanying the tweet (i.e. "??????????..."). Tweet number 8 refers to an explosion noise heard in Ile de France that day (which turned out to be a plane breaking the sound barrier). Netflix took the opportunity to make a link with its universe and in particular its iconic noise introducing each of its production.

Finally, the last tweet of the top concerns a recurrent request that Netflix users make to the platform: the addition of the anime "One Piece". The brand responds to all these people at once by using what is called ASCII (a character encoding) forming a character carrying a sign. NetflixFR thus uses a lot of second degree in its communication (4 tweets of this type on a par with the ads). The account does this in different ways by making references to phenomena, whether on the internet (Wejdene) or in real life (plane breaking the sound barrier) while linking these events to an element of its brand. She also does this via videos taking her own productions at face value, which might seem marketingly questionable. Indeed, we are used to a brand highlighting one of its products in a clean and traditional way. Here it's the opposite, the brand puts itself in the place of the user and does not do it in a classic way. And obviously, according to the statistics, it works. The last tweet, although humorous above all, also reflects the desire to interact with its followers. Of course the account could simply ignore all requests for content but prefers to respond with a direct tweet and a touch of humour.

4.5.2 Netflix - Appendix 7.2.2

At Netflix US we find the 3 categories we identified at its French counterpart but also a new one.

Indeed, the American account seems to be involved in some social issues as proven by tweets number 1 and 3. The first one in support of the Black Lives Matter movement that emerged during the year 2020 after the death of George Floyd in the United States. This tweet exceeded one million likes and is the only one to have done so within our study. The second was a tweet that was quoted as saying that there were too many gay people in Netflix series. The original tweet has since been deleted. Netflix thus seems committed to the protection and representation of minorities. This can be seen both in their communication but also in their series, since they are represented in many of their productions.

The second category concerns the use of humour. This part is represented by tweets 2 and 6. One refers to a real life event using a scene from the Squid Game series. The other uses off-set footage of

the Stranger Things season 4 villain's costume to make the situation unusual. In both cases, the aim is to make the user laugh while making a reference to them and thus promoting one of its productions.

Another category that we also find and that we should have for each account in view of our previous analyses are the ads. The top 5, 8 and 10 illustrate this with fairly classic ads. We can find the presence of top words like "season", "love" or "life". We can also link tweet number 8 with a topic identified in the LSA, number 4 having as 3 main words "la", "heist" and "casa" which are all present in this tweet.

Finally the last category, the simple reference. Tweets 4, 7, and 9 are references to elements of series produced by Netflix that have not been classified as humorous (subjectively).

4.5.3 PrimeVideoFR - Appendix 7.2.3

The top 10 tweets from PrimeVideoFR are quite surprising and interesting compared to previous analyses. It turns out that 9 of the tweets are humorous and only one tweet is about an ad. Moreover, the latter is the 10th tweet. It seems that PrimeVideoFR's followers are particularly sensitive to the account's humour. Moreover, most of the tweets are related to memes that are quite popular in the internet culture. This is the case for tweets 2 to 9.

We dot not explain the different meanings of these memes as it would take too long. Indeed, there would be a lot to explain as the references are often related to events outside the streaming industry itself. The important thing to remember is that the account's community managers know the internet culture and how to use it wisely. Thanks to this, the account is once again detached from the cold image of a business brand present only for promotion and marketing.

However, is this not too much? Of the top 10 tweets, only one was promotional. Moreover, this announcement does not even concern an Amazon production but the acquisition of broadcasting rights for some Ligue 1 matches. Compared to the other accounts, there are very few tweets about announcements, and even too few if we assume that the main purpose of these accounts is to promote their content. It is therefore interesting to look at the reasons why these announcements are not in the top tweets. From a positive point of view, we could simply imagine that the humour is effective enough to make users react a lot. From a negative point of view, we can make the hypothesis that the series and films announced on PrimeVideoFR are not expected and appreciated enough to make users react. In reality, we think this is probably due to a combination of both, an excellent sense of humorous opportunity coupled with a lack of compelling content on the part of PrimeVideoFR.

4.5.4 PrimeVideo - Appendix 7.2.4

The type of PrimeVideo's top 10 differs slightly from that of its French counterpart. We certainly find humour as we have seen with the French account but in a lesser presence. Indeed, half of the tweets can be considered as having the aim of making people laugh.

The notable difference is in the ads, both in terms of content and form. It turns out that the other half of the tweets from the American account are about the release of an Amazon PrimeVideo production. Not only are there many more announcements, but they are all about an original production and not about third party content as was the case with PrimeVideoFR.

In fact, on closer inspection, all of these Prime ads are actually about the same program. A movie called "My Policeman" which will be released on November 4th 2022. It is quite difficult to explain the popularity of this as yet unreleased film and its failure to make the French top ten. The most plausible explanation would be the presence of the singer "Harry Styles" in the lead role and the fact that he is very popular in the United States and potentially less so in France.

4.5.5 DisneyPlusFR - Appendix 7.2.5

At DisneyPlusFR, we find a lot of ads. Of the 10 tweets, 7 have an advertising purpose. What is, however, totally different from the other accounts is the timing of these tweets. Indeed, all these tweets are in a range of more or less 40 days. This period actually corresponds both to the launch of the platform, which is often mentioned within these tweets, but also to the confinement that most of the planet (and thus the French-speaking world by extension) has experienced. The launch effect coupled with the fact that people were at home a lot (and therefore had more time to spend on the networks) certainly boosted the statistics. This is why, after more than 2 years of existence, these tweets remain the most popular. We could also have made the same hypothesis as for PrimeVideoFR (i.e. the lack of attractive content) but this seems unthinkable coming from Disney which owns and creates extremely strong and popular licences (e.g. Star Wars, Marvel, Pixar, ...).

We can also find, in the same period, a tweet encouraging interaction. Tweet No. 8 asks, on the eve of the platform's release, what the users' priority program is. Once again, this reinforces the effect of the announcement and the enthusiasm of the time, but this time by asking followers.

Finally, the last two tweets are quite special as they don't really belong to any of the traditional categories we have encountered so far. One of them, which is none other than the top 2, is a bot system, an automatic response, which notifies people who liked the initial tweet of the release of an episode. This is quite interesting as it is a phenomenon that we encountered in our Methodology section as it posed analysis problems. The presence of this tweet in the top proves that this technique is effective and wanted by the users themselves. The second type of original tweet could look like an advert, but this one is not about the content directly. This one is rather negative as it concerned the postponement of the release of the service by the French government. This request was made in view of the bandwidth usage generated by the start of the containment.

4.5.6 DisneyPlus - Appendix 7.2.6

Looking at the most popular tweets from DisneyPlus, we can see straight away that, compared to other accounts, nothing is very surprising. Almost all of them concern announcements of series or films.

Only 2 tweets do not fit into this category. The third tweet is a tribute to Chadwick Boseman who died prematurely. He played the role of Black Panther, a hero of Marvel, a studio owned by Disney. The eighth is a quote from a program in which singer Taylor Swift is taking part.

4.5.7 Most popular tweets - Summary

We were able to identify different types of popular tweets: advertisements, humorous tweets, tweets about societal phenomena, and references to each brand's series.

Some accounts are strongly focused on one type of tweet, such as PrimeVideoFR with humour or DisneyPlus US with ads.

Others use a mix of several of these types, such as the two Netflix and PrimeVideo US accounts.

The French-speaking Disney account also sends out a lot of promotional tweets, but these are fairly focused on the launch of the platform and not the arrival of new ones after the platform's life has begun.

5 Results

In this section, we discuss the different results that we have been able to obtain in our previous analyses. In the course of these analyses, we have already suggested points of interest, links between the analyses and hints about their meaning. Here, we summarise the results obtained in all the analyses carried out. This allows us to take an overview of the analyses and to define the overall situation of communication in the streaming services sector. But also to answer the different questions we asked ourselves at the beginning of the thesis.

Firstly, we start with an overview of the situation on the social networks of streaming services and profling each account. Profiling means that we precisely determines the caracteristics of each account and all the elements of their communication strategy. Secondly, with the help of the previous profiling, we define a global classification of communication strategies that we use to classify the different accounts studied in this thesis. Finally, we identify the differences between the accounts of the same services and their socio-cultural identities.

5.1 Sector summary

Throughout this thesis we have been talking about the situation of the streaming service sector in terms of digital communication. It is now time to take stock of the latter in a rather global way by focusing on specific points, which we could not do next as we will define more global strategies. For this purpose, we take into account all the analyses made in the previous chapter and establish first a report on the whole sector and then a profile for each account studied.

As we have mentioned many times in our analysis, it is clear that the communication of all the accounts remains mainly focused on promotion. Whether this is primarily about the products offered (series and films) or about the brand, all the accounts send out a majority of tweets for advertising purposes, as our Latent Semantic Analysis has shown us.

We have observed that the month of June seems to be a slow period for the accounts. In fact, only one account breaks the rule, NetflixFR, for which this account is not a low but a slight peak. It seems that this month is even quieter for the others since we recall that NetflixFR is the most active account by far (the posts of NetflixFR represents 48.22% of all posts collected). If this month is low for the sector, it was therefore necessary that the other accounts are in a big slump to compensate for the slight peak of the French account of Netflix.

Speaking of which, if there is one brand that dominates the others in terms of statistics it is Netflix. The American account is quite ahead of its competitors: 1st in number of weekly posts, 1st in followers with more than 5x than the 2nd American, account using the most multimedia, American account with the most tweets with at least 10K likes. What is even more impressive is the French side of Netflix. The domination is also present but it is even more accentuated than on the American side: 5x more posts than PrimeVideoFR and 15x more than DisneyPlusFR, 25x more followers than the second, more than 1400 posts with more than 10K likes. We will talk more about the statistics of these two accounts in the definition of their profile but it is clear that on the field of social networks Netflix dominates largely. As for the hierarchy between the other two, Disney seems to be slightly ahead of the Amazon-owned service in terms of popularity (followers) despite being long established. Indeed, DisneyPlus is 2.5 years old in

America and is a bit more than 2 years old in Europe but both accounts already have more followers than PrimeVideo. This is certainly explained by the fact that Disney was already a strongly established and appreciated brand in the audiovisual sector and that they had already acquired a community strongly loyal to their product.

5.1.1 NetflixFR

We have just touched on it, but it is clear that the Netflix brand seems to be outperforming its competitors. This confirms what Martinez-Sanchez and al. (2021) [4] have identified in their study of the streaming service market in 2020. The French-language account performs particularly well in the French-speaking region, but also in comparison to the American accounts. Indeed, NetflixFR is certainly the most active account within our study but the second account, which is none other than its American counterpart, has actually 3 times less posts. This account is therefore extremely active compared to all other members of the sector studied. In fact, when you look at the numbers more closely, in terms of tweets and retweets NetflixFR is not necessarily far ahead of its competitors. The difference in the number of posts is actually in the number of replies the account makes to users. Indeed, NetflixFR has more than 240 weekly replies to its credit. This is 7.5x more than the next best performer in this area (PrimeVideo US).

The profile of Netflix France is therefore partly different. The account does everything that every other account does equally or better (e.g. number of weekly tweets) but adds a huge amount of interaction with its users. Their goal is not just to post a tweet and let people react to it and hope it gets good stats. No, for them, the release of a new tweet is just the beginning of the work as they respond to the comments that their followers have made to that tweet. We were also able to identify in our Latent Semantic Analysis that the account was posting tweets that prompted responses from users, which usually resulted in a new response from NetflixFR. According to Fernández-Gómez and Martín-Quevedo (2018)[9], it is one of the most efficient way to encourage interaction between brands and users.

It is clear that the French-speaking account of Netflix is by far the most active and interactive. However, is it for all that the most successful, is their quantity does not harm their quality? Well it is clear that no, as we have already mentioned several times, NetflixFR is the most followed French account with 7.5 million followers. Moreover, it is the account that reaches the largest part of its reachable population since 2.47% of the French-speaking world follows NetflixFR (assuming that a person has only one account which is not necessarily the case). It is also clear that the account producing the most tweet buzz (with more than 10000 likes) is also this one with more than 1400.

An interesting phenomenon concerning this account is the non-dependence on audio-visual elements. Indeed, compared to other accounts it uses quite a few (63% of tweets containing it) and even in buzz tweets (same percentage). This proves that NetflixFR finds other ways than just putting a picture of an ad, a popular character or something else to make popular tweets. They're not taking the easy way out as much as their competitor. This result is different from what might have been expected from the findings of Scerbinina (2019) [7]. Indeed, according to his study on American and Indian accounts, Netflix uses a lot of visuals in its communication. But this does not seem to be the case for the French account.

In the same study, 3 types of tweets are identified within Netflix's communication. Those concerning the brand and its products (films, series, ...), those pronouncing the interaction and finally those

talking about elements more distant from the streaming sector such as socio-cultural elements. During our study we could only identify 2 of these 3 categories (the first two ones) within the communication of NetflixFR which partially validates these results.

Finally the last strategic element of Netflix France is definitely its image. The brand wants to be close to its users, as we have seen with its number of interactions, but this is also reflected in the content of its tweets. Indeed, during the "Top Words" stage we mentioned the fact that the only more or less informal word we could identify came from this account. Moreover, this word, "dispo", was no less than the second most used word. This shows that NetflixFR doesn't necessarily want to use a very clean and formal language that one would expect from a brand as strong as theirs. On the contrary, it seeks to make users feel close to the brand through this process, breaking down the wall between the strong and a priori inaccessible brand and the average user. It also achieves this through humour, which we have also discussed extensively. Fernández-Gómez and Martín-Quevedo (2018) [2] have also noted a rather conspicuous use of humour at Netflix which achieve to create engagement from the followers.

In short, Netflix France is the account that stands out the most within our study, often going against the grain of the sector as for the use of photos and video. Everything being largely dominant in all areas: activities, popularity, community, ... They use strategies such as humour and informality to create closeness in addition to their numerous responses to their followers. Their focus therefore seems to be on the community and the interaction they want to have with it.

5.1.2 Netflix

Netflix is of course dominant in the US market as is its French-speaking counterpart. However, there are some profile differences between the 2 profiles of these accounts.

Netflix has a more classic profile with a certain balance in the use of different types of posts. However, there is an important use of retweets compared to the others. This account therefore makes extensive use of the content created by its followers in its own communication by means of retweets. This had already been identified in the work of Scerbinina (2019) [7]. In fact it uses more retweets than original tweets which means that their entire communication relies more on other people's tweets than their own.

This account interacts, despite everything, moderately with its followers according to the metrics (27% of replies). However, this is not represented in the categories identified in the LSA, which shows that it is not regular enough. However, we can count their high number of retweets as a form of interaction since it directly involves the community in their communication. This is still less powerful in terms of exchange than a direct response.

We can once again make the link with the study of Scerbinina (2019) [7] which had identified 3 types of tweets issued by the American account of Netflix: the tweets concerning everything that revolves around the Netflix brand (programs, releases, ...), the tweets inciting interaction and finally those not really concerning the streaming but social phenomena. Regarding the first type, it is quite easy to confirm the existence of this category since it is the one that is extremely represented in the "Top Words" and "LSA" parts of our analysis. Indeed, most of the top words and almost all the categories of our LSA concern this type of tweets. The second type, interaction, is not directly identified in our analysis but is

assumed in view of the percentage of responses made out of the total number of posts (27%), which is certainly not huge but sufficient for it to exist. Finally, the last type, which we were unable to identify for the French account, is well present in our observations for the United States. Moreover, these types of tweets are among those that have collected the most likes as we can see in the "Most popular tweets" part of our analysis. We can therefore say that our results are in line with those of Scerbinina (2019) [7].

In the same study it is mentioned that Netflix is very dependent on the use of photos and videos in its communication. During our analysis we were able to study this element via our "Metrics" section. We can observe that 94% of the posts contain them, which largely confirms the conclusions of Scerbinina (2019) [7].

If we were to summarize Netflix's profile in a few words, it is an account that uses all types of posts with a certain balance while using more than the average tweets from other users in its communication. Their original tweets are heavily focused on promoting content via images and video with an occasional but notable focus on societal phenomena.

5.1.3 PrimeVideoFR

PrimeVideoFR's profile is actually quite varied. Rather active without being extremely so like NetflixFR, a certain balance between tweets and replies with less interest in retweets.

PrimeVideoFR's communication is, like the rest of the sector, primarily focused on promoting its content, as shown by the analyses made via the "Top Words" and the "LSA". It should be noted that there are two types of promotion in the communication of Amazon's service: those focused on the brand (PrimeVideo) and those focused on the products (films, series, etc.). With the first one being the most dominant according to our figures.

In these two stages we also identified a type of tweet that called on the opinions of their followers on themes such as their favourite work or actor.

This confirms the results obtained in the "Metrics" section since we identified 46% of tweets and 42% of replies. Almost all of our analyses show that PrimeVideo's appeal lies both in promotion and also quite strongly in interaction.

However, when we look at the top 10 most popular tweets we find almost no such tweets. Only one advert (the 10th) and no tweets challenging the user, almost only humorous tweets. The presence of this kind of tweets is not reflected in the previous analyses, which is actually understandable since humour has different forms and the lexical field used can be extremely varied.

It is interesting to ask why only humorous tweets are in this ranking because, overall, a lot of accounts have humorous tweets in their top 10 tweets but it's only a portion of the top 10, not nearly all of them. So what is different about the French account of Amazon's service? We have already mentioned some hypotheses in the "Most popular tweets" analysis, but in our opinion it is a mixture of two elements. Firstly, it seems that PrimeVideoFR's community manager(s) have a good knowledge of internet culture, proving it by using popular social media memes effectively. Their humour is therefore particularly appreciated by Twitter users where the use of memes is king. Secondly, it is clear that in our opinion PrimeVideo lacks iconic content with high popularity and visibility for which a release announcement could generate many likes. There is however one huge exception to this: The Lord of the

Rings series. Indeed, a Lord of the Rings series has been announced and despite the huge popularity of the saga the tweet announcing it does not seem to have made it into the top 10. We'll see on the day of its release on September 2, 2022 if the tweet notifying the availability of the series on the platform can do better.

One element not to be overlooked that we have at the beginning of this thesis is the use of bot. Indeed, PrimeVideoFR is one of the accounts concerned, with 46% of the posts being bot related. It is quite difficult to conclude anything about this. The use of this can be explained by a lot of different things: a lack of means, a lack of content, a strategy that is deemed to be effective, ... It is however certain that these are not the most federative tweets because they are not personal.

In summary, PrimeVideoFR is an account that is necessarily focused on promoting content (like all the actors studied) but also on promoting the brand itself. It also invites its followers to react via direct tweets to them and direct replies to their tweets. The latter seem to particularly appreciate the humour generated by the account as it is these tweets that they like the most.

5.1.4 PrimeVideo

The profile of PrimeVideo US is, according to our analysis, broadly similar to the French account but we can see some differences on certain points. We also find a fairly balanced profile that does not particularly emphasise a certain type of post in terms of frequency or neglect any.

Speaking of frequency, PrimeVideo is a fairly active account, ranking 3rd among the accounts studied and second among English-language accounts to very few Netflix. This is in line with what has been observed by Martinez-Sanchez and al. (2021) [4], i.e. significant activity but nowhere near that of Netflix.

Like its French-speaking counterpart, Prime US uses promotion in two distinct forms, the first focused on the brand and the second on the content offered by the brand. Here too the first is more present than the second.

Following on from these commonalities, it turns out that PrimeVideo also has a category dedicated to user interaction. In fact, the words in this category are literal translations of the words used by the French account. We can therefore imagine that the typology of these tweets has been strategically shared by one of the two accounts, the other having simply translated it.

However, several notable differences can be identified within the communication of the 2 accounts. The first is in the most popular tweets of each account. The French account has almost only humorous tweets, whereas the top 10 tweets of the English account are a mix of content promotion and humour that is still expected.

PrimeVideo Us also uses bots and even more that its french account with 71,42% of posts being pre-generated.

An element that may seem insignificant but which demonstrates a new strategic element not seen in other brands is the presence of watch parties. Indeed, we identified in the LSA section that the English account was communicating on the watch parties it was setting up on the Twitch platform. This is a communication and acquisition strategy that we did not see in their competitor. Beyond the concept of

watch parties, we can also identify the partnership with influencers, a phenomenon that is exploding and that has shown its marketing efficiency (O. L. Vyatkina (2020) [11]). However, only PrimeVideo uses it enough, in its communication at least, for us to be able to identify it (of course the other services have already dealt with influencers: Disney, Netflix)

In summary, PrimeVideo has a fairly classic communication style being quite active, community-based and humorous without being strongly focused on any of these points. Their strength is therefore to master different points in an equivalent way. However, they are not as expert in one area as Netflix France is in interaction or PrimeVideo France in humour. They seem to be quite forward thinking with influencer partnerships and watch parties. Creating both a new marketing strategy and a new way of consuming audiovisual content.

5.1.5 DisneyPlusFR

The French account of DisneyPlus France is by far the least active. With only 21 posts per week, Disney does not seem to give a huge importance to communication on social networks. Moreover, when we look more precisely at the types of posts made by the account, it is clear that the goal of their communication is not community. On average, no retweets and only 2 replies per week. This means that the main and extremely dominant purpose of DisneyPlusFR is to promote content. The results obtained in the "Top Words", "Latent Semantic Analysis" and "Most popular tweets" sections confirm these findings.

The momentum of the account seems to be fading with a decrease in the number of monthly tweets since the beginning of 2021. Furthermore, it seems that the only time Disney has managed to unite a community is at its launch, as the 10 most popular tweets come from this period. This is in line with the study made by Martinez-Sanchez and al. (2021) [4] which identified a more than satisfactory launch period for DisneyPlus in terms of statistics on its networks (followers, likes, etc.). In view of their spectacular arrival, this same study assumed that DisneyPlus would quickly overtake PrimeVideo in terms of popularity, especially in terms of followers. This is the case in the French-speaking market at least.

This account is also concerned by the use of posts automation. It is the less dependent user but it still represents 32,28% of its posts.

There is really not much more to say about this account. So, if one were to summarise the profile of DisneyPlusFR, it is clear that this account serves as a simple showcase for the content offered on the platform. The only role of the account is to promote the products of the service and to create excitement around them. And even in this unique role the account does not seem to be very active.

5.1.6 DisneyPlus

Despite being twice as active as its fancophone counterpart, DisneyPlus US is not the most active. 5th in our study and last of the American accounts.

As far as the types of posts used are concerned, the situation is similar overall, with a lot of original tweets. However, the use is still present, even if it is slight, which is also the case for replies.

The Top Words and the LSA are globally very similar to the French account. A lot of words and categories referring to the promotion of content and especially the brand. The most popular tweets are different as they are not in the same period which is logical as the service was launched at the end of 2019 in America, a period not studied in the case of this thesis. However, the purpose of these tweets remains the same: promotion.

Despite low activity and a more recent presence than its competitor AmazonPrime, DisneyPlus has more followers and has issued more tweets with over 10,000 likes. This can be explained by the strength of the Disney brand, which is certainly one of the best known in the world. Martinez-Sanchez and al. (2021) [4] predicted this surpassing in view of the early life of DisneyPlus.

It is by far the account using the most bots in its communication as 81,28% of its entire communication is generated by it.

In summary, DisneyPlus is an account that is used almost exclusively for announcing series, films and other releases, as well as for publicising the brand. Community interaction seems very limited and put in the background. Communication does not seem to be at the heart of Disney's marketing since their activity on the networks seems to be quite low.

5.2 Socio-cultural differences

In this section, we define, for each of the brands, the differences or similarities between their accounts. The aim is to identify whether the accounts of the same brand seem to follow a common strategy, i.e. whether there is an international strategy or whether each account in each region has its own strategy and has more freedom at this level. At a more micro level, we observe whether the different accounts adapt to the culture of each region within their tweets.

5.2.1 Netflix

Strategically, the platform's two accounts seem to be quite different. The French account is initially much more active than its American counterpart, although it is not passive at all. The two accounts seem to want to interact with their community in different ways. On the one hand, we have NetflixFR doing it via direct replies to users with an average of 247 replies per week. On the other hand, Netflix uses mostly retweets highlighting users' opinions on such or such series. One responds directly to its community while the other uses content created by the community directly for its own communication. A notable difference is also in the brand image that the two display. The French account seems to want to detach itself from its status as a huge company, taking things very much in the second degree. Netflix US uses some of these processes also to a lesser extent like humour but also seems to use its status to convey societal messages. The english account seems to have less of a desire to be extremely familiar with its users. Scerbinina (2019) [7] also identified differences in communication by region by looking at the US and Indian accounts.

As for the content of the tweets produced by the accounts, it is clear that each one adapts to the culture of the region it reaches. As far as NetflixFR is concerned, we have 2 quite adequate examples within the most popular tweets. Indeed, the top 4 and 8 tweets refer to phenomena specific to the French-speaking world in the presence of Wejdene, a French singer, and the passage of a plane breaking the

sound barrier in Paris. As for the American account, its most liked tweet with more than 1 million likes concerns one of the most publicized American social phenomena in recent years: Black Lives Matter.

5.2.2 PrimeVideo

Both accounts seem to be rather brand promotion oriented and have quite similar metrics without being totally identical. Their activity and level of interaction are quite close and their way to do it look similar (a small disinterestedness of retweets for PrimeVideoFR but not a significant difference). We can therefore say that both accounts seem to follow a common strategy consciously or not.

The content of the tweets seems to be quite different between the accounts, thus adapting to their region. PrimeVideoFR uses a lot of humour as we have already mentioned and this humour is mostly from the French internet culture like tweet number 6 referring to Woippy a rather colourful French town which Twitter has seized upon to make memes. Of the 10 most popular tweets on PrimeVideoFrance 7 are related to French internet culture or elements of French culture. In the case of PrimeVideo US this demarcation is less marked because it is less diverse (as far as we can tell with the scope of this study). Indeed, the most popular tweets of the American account are largely related to a specific program "My policeman" which seems to be very appreciated by their community. Amazon's service seems to have understood this and communicates a lot around it. However, this is not at all the case with the French account and we can explain this by a cultural difference. This would be because the popularity of the lead actor, Harry Styles, a singer, in the US is not of the same magnitude as it is in France.

5.2.3 DisneyPlus

It looks like that the 2 accounts have a similar strategy. Indeed, both are brand-oriented in their promotion. Furthermore, both accounts are not very active on their Twitter accounts and do not interact much. It would therefore seem that the entire DisneyPlus service does not place much importance on the use of social networks in their communication plan.

DisneyPlus is the brand that uses the least amount of cultural elements from each region in its tweets, according to our analysis. The two accounts seem to be used only to announce the release of programs and not to create any really engaged community.

5.3 Classification of strategies

We have just reviewed the profiles of the different brands in our study an their cultural differences. With the help of these, but also of a typology of classifications of communication strategies, we divide the different accounts into different categories.

5.3.1 Typology of classifications

Our classification model is based on three main points: activity, interaction and finally promotion orientation. The first two variables are intensity scales (low, medium, high) while the last one is binary (brand-oriented or product-oriented). These levels are estimated according to the averages of all the

players studied (taking into account the fact that NetflixFR has huge figures otherwise all the others would be at low levels). The first two points are inspired by typologies of classifications established in previous work of this kind. Wu and al. (2012) [2] took into account the importance of the quantity of tweets sent, while Floreddu and Cabiddu (2016) [6] took into account the level of interaction of the brand with users (Wu and al. (2012) [2] also did this in a less direct way but via the themes of the tweets). The third point was added to our classification after analysis and in view of the differences in the promotional focus of the various brands studied.

5.3.2 Classification of accounts

Here is the classification we have established:

Brand vs Product	Activity Level	Interaction Level		
		High	Medium	Low
Brand oriented	High			
	Medium		PrimeVideo PrimeVideoFR	
	Low			DisneyPlus DisneyPlusFR
Product oriented	High	NetflixFR		
	Medium		Netflix	
	Low			

Figure 20: Classification of each accounts

As we have already gone into detail on each of the profiles of these accounts we will not repeat everything that has been said. However, it is important to explain our reasoning in terms of classification.

NetflixFR : Product oriented as no real brand promotion was identified as really important in the "Top Words" or in the "LSA". Regarding activity and interactivity, it is clear that the account largely dominates on these points as our metrics have shown.

Netflix : Product oriented as no real brand promotion was identified as really important in the "Top Words" or in the "LSA". There is a lot of activity and interactivity on this account, but it is far from its French counterpart and not that far from PrimeVideo accounts, so it was put in medium (would potentially be high if NetflixFR wasn't there).

PrimeVideoFR : Brand oriented as the most represented category of the LSA is about the brand

and not about the products or the availability of a programme adding the fact that their two most used words are "premium" and "video". The level of activity is medium as it is equal to the sector average (removing NetflixFR which greatly increases it). The level of interaction is medium with 54% of posts not being tweets but replies and retweets.

PrimeVideo : Overall the same thing for the same reasons since as we mentioned in the previous section the two PrimeVideo accounts seem to follow a common strategy. The differences are on things not taken into account by the table such as humour, topics, ...

DisneyPlusFR : The account is slightly more brand-oriented than product-oriented as its top words and LSA show. As far as activity and interaction are concerned, it is clear that this account is the worst performer in the study by a long way. In fact, its number of posts is 3 times lower than the average (without NetflixFR) and the interaction is almost zero.

DisneyPlus : This account is certainly the most brand-oriented, given the number of words in their top (marvel, marvelstudios, ...) referring directly to the brand and not the products. The activity and interaction, although better than the DisneyPlus France account, are still far from the average and this is why both are ranked low.

These results are in line with what we have just defined in the differences between regional accounts. Indeed, we found that Netflix did not have an international strategy while the other two services, PrimeVideo and DisneyPlus, did. In our table, this result is graphically reflected by the sharing of the same box for the PrimeVideo and DisneyPlus accounts while the two Netflix accounts are in different boxes.

6 Conclusion

Throughout our study we have studied the communication of the biggest players in the streaming service sector. We used different approaches via data-mining, text-mining and finally qualitative analysis. All this with the aim of establishing a precise profile for each account studied. Secondly, to classify these accounts according to global strategies as well as to identify potential international or individualistic strategies towards the brand. All this while linking to as much previous relative work as possible.

Our analyses have allowed us to establish different conclusions, both on the sector itself and on the different accounts that represent it. The sector experienced a great peak of communication during the first half of 2020, corresponding globally to the containment of COVID-19. Community managers certainly aligned the pace of their communication according to the growth of content watched during this period [10]. We have also noticed that whatever the account, what dominates above all is the promotion of content, which is logical since, after all, communication is only a tool for acquiring and keeping market share.

Netflix is totally dominant in their communication over their competitors on the English-speaking account but especially on the French-speaking territory where they dominate by several dozen lengths. The brand is also by far the closest to its followers interacting and incorporating them into their communication according to the account. This is to create brand loyalty and therefore maintain market share. This success is also commercially speaking since Netflix remains the streaming service with the most users despite the fact that the trend is starting to reverse. It remains to be seen whether this commercial success is partly explained by their communication or whether, on the contrary, it is the fact that the platform is already very well established that brings success to its social networks.

PrimeVideo, the service of the giant Amazon and currently the second largest streaming service in the world, is quite present on the networks with a shared strategy between the French and English account. It dominates Disney's accounts in terms of activity, but is still last in terms of followers. It seems that despite its efforts the brand is paying for its lack of original content and brand power that Disney and to a lesser extent Netflix can have. However, in addition to a global account strategy in terms of interactions and activity, humorous tweets, especially on the French side, seem to be particularly appreciated. This lack of content is nevertheless felt even commercially since the platform is certainly second but its growth is much less than that of DisneyPlus since the latter is a much younger service.

DisneyPlus does not seem to place much emphasis on communication and even less on interacting and creating a bond with its community. And yet despite this neglect, the growth in terms of followers that the accounts have experienced since their recent creation is quite impressive. All the more so when you put them in perspective with the little effort they seem to put into them. This is most likely due to the power that the Disney brand represents and the quality of the old content that it brings. Indeed, we can assume that a large part of its success does not come from the fact that it creates new programs but from the emotional link that a part of the population has with its old films and other cartoons that have marked several generations.

In two such young and changing sectors as social networking and streaming services, it will be extremely interesting to see how the future develops on their side. It is actually very difficult to predict what communication on social networks will look like in 5 years time. Indeed, 5 years ago, communi-

cation did not look like at all the one we have today, memes hardly existed, brands were still very much keeping their cold and imposing image of big company. It is even more difficult to combine this with the streaming sector, which is also not completely stable and anchored, as we can see with the arrival of new business models at Netflix and Disney . It would therefore be very interesting to carry out a similar study again in 5 years' time and compare the results. For the time being, it is clear that these two sectors will continue to grow side by side.

7 Appendix

7.1 Individual Evolutions

7.1.1 NetflixFR

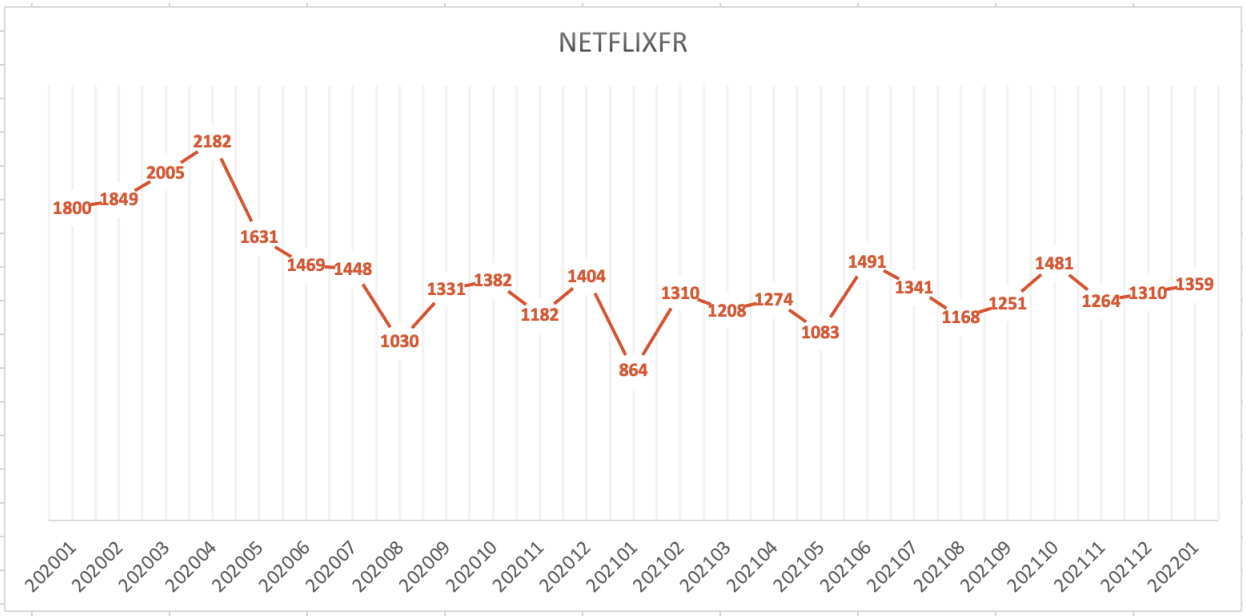

Figure 21: Evolution of @NetflixFR's posts over last 2 and half years

7.1.2 Netflix

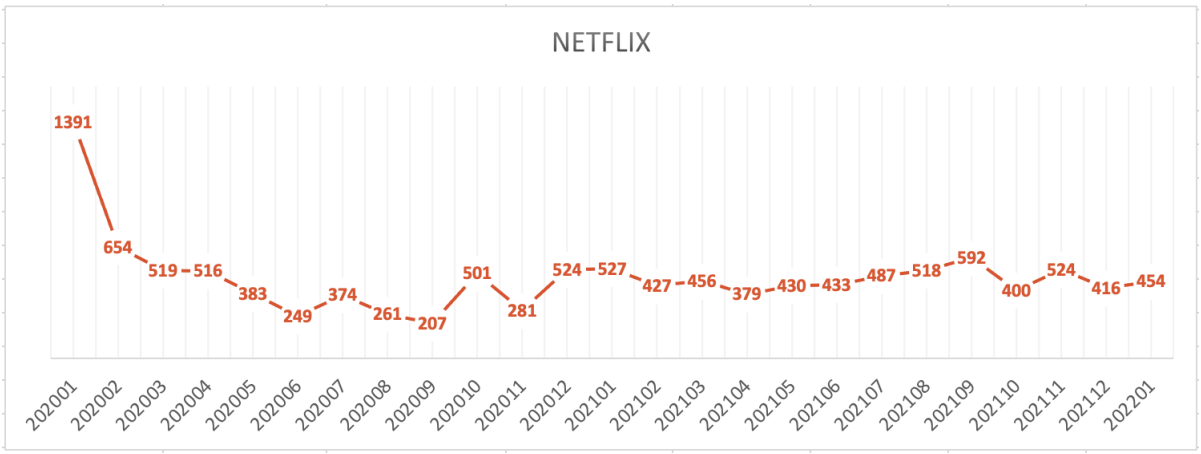

Figure 22: Evolution of @Netflix's posts over last 2 and half years

7.1.3 PrimeVideoFR

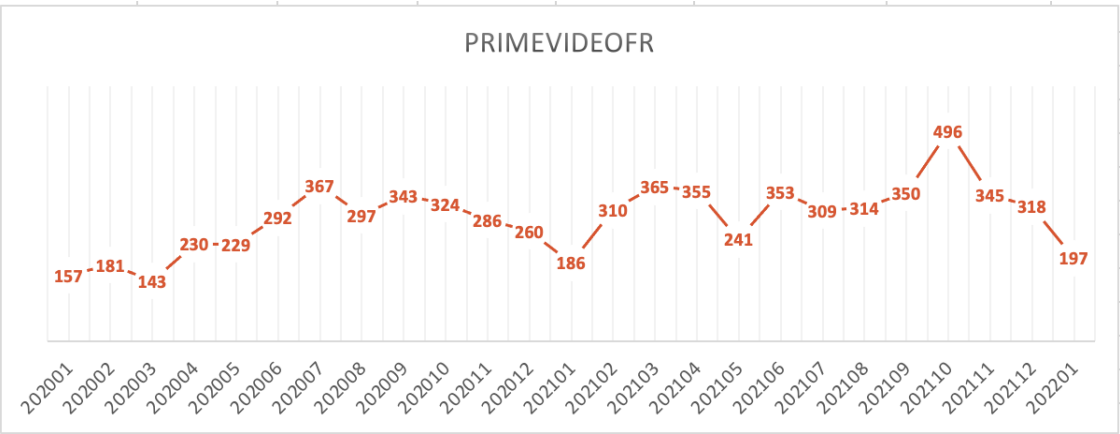

Figure 23: Evolution of @PrimeVideoFR's posts over last 2 and half years

7.1.4 PrimeVideo

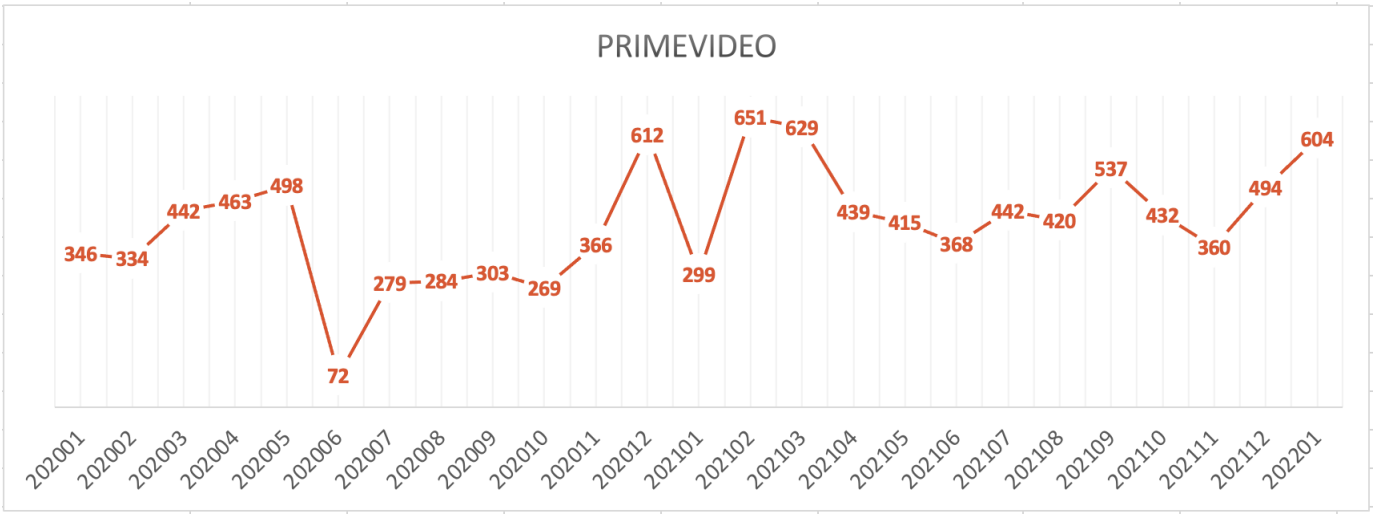

Figure 24: Evolution of @PrimeVideo's posts over last 2 and half years

7.1.5 DisneyPlusFR

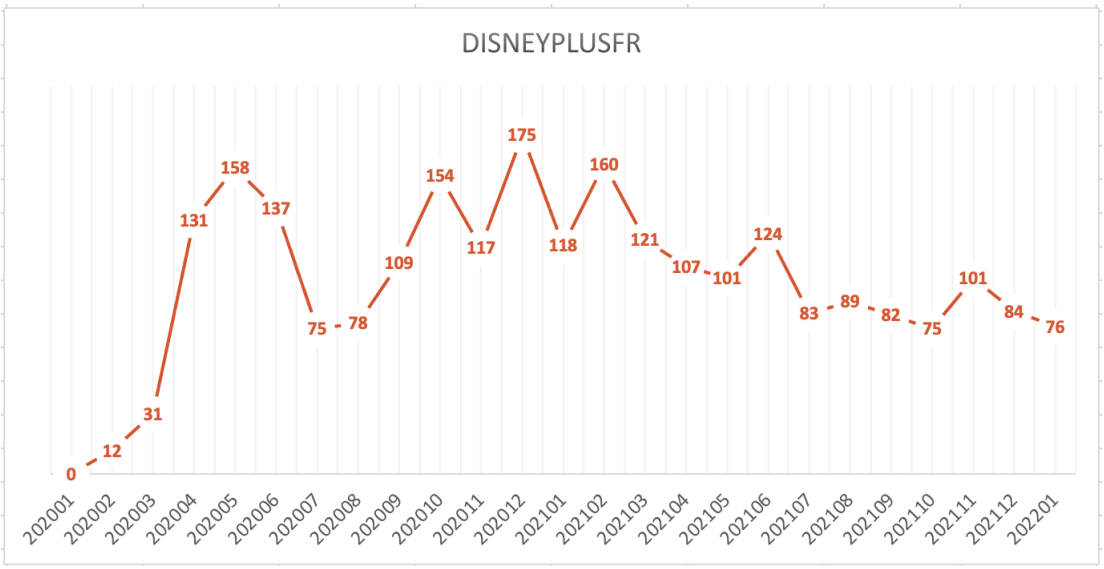

Figure 25: Evolution of @DisneyPlusFR's posts over last 2 and half years

7.1.6 DisneyPlus

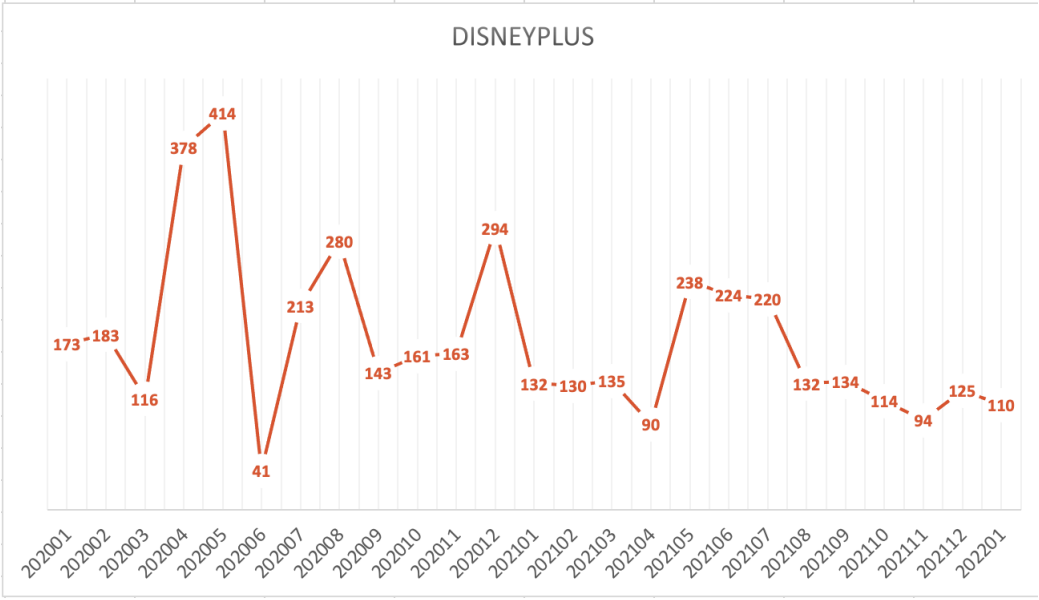

Figure 26: Evolution of @DisneyPlus's posts over last 2 and half years

7.2 Most popular

.

7.2.1 NetflixFR

Top 1

Figure 27: https://twitter.com/NetflixFR/status/1538940995986784257

Top 2

Figure 28: https://twitter.com/NetflixFR/status/1242496403802411009

Top 3

Figure 29: https://twitter.com/NetflixFR/status/1244663314032009216

Top 4

Figure 30: https://twitter.com/netflixfr/status/1277911602587631616

Top 5

Figure 31: https://twitter.com/NetflixFR/status/1241024200678289412

Top 6

Figure 32: https://twitter.com/NetflixFR/status/1512725755570106368

Top 7

Figure 33: https://twitter.com/NetflixFR/status/1219167917054529536

Top 8

Figure 34: https://twitter.com/NetflixFR/status/1311250641243836417

Top 9

Figure 35: https://twitter.com/NetflixFR/status/1245757891866476544

Top 10

Figure 36: https://twitter.com/NetflixFR/status/1375412216569139203

7.2.2 Netflix

Top 1

Figure 37: https://twitter.com/netflix/status/1266829242353893376

Top 2

Figure 38: https://twitter.com/netflix/status/1445061159158042625

Top 3

Figure 39: https://twitter.com/netflix/status/1258156383792279552

Top 4

Figure 40: https://twitter.com/netflix/status/1537581997987483648

Top 5

Figure 41: https://twitter.com/netflix/status/1536015471429840896

Top 6

Figure 42: https://twitter.com/netflix/status/1537209786067914752

Top 7

Figure 43: https://twitter.com/netflix/status/1466498273351331843

Top 8

Figure 44: https://twitter.com/netflix/status/1289214267640565761

Top 9

Figure 45: https://twitter.com/netflix/status/1542244315845431296

Top 10

Figure 46: https://twitter.com/netflix/status/1527635263958097920

7.2.3 PrimeVideoFR

Top 1

Figure 47: https://twitter.com/PrimeVideoFR/status/1462723030896025606

Top 2

Figure 48: https://twitter.com/PrimeVideoFR/status/1484177854724145158

Top 3

Figure 49: https://twitter.com/PrimeVideoFR/status/1512107862817914890

Top 4

Figure 50: https://twitter.com/PrimeVideoFR/status/1425103993798287368

Top 5

Figure 51: https://twitter.com/PrimeVideoFR/status/1531652589418315777

Top 6

Figure 52: https://twitter.com/PrimeVideoFR/status/1516422682333564931

Top 7

Figure 53: https://twitter.com/PrimeVideoFR/status/1462790771845775363

Top 8

Figure 54: https://twitter.com/PrimeVideoFR/status/1346136558756192256

Top 9

Figure 55: https://twitter.com/PrimeVideoFR/status/1429836327986495489

Top 10

Figure 56: https://twitter.com/PrimeVideoFR/status/1403428703351283719

7.2.4 PrimeVideo

Top 1

Figure 57: https://twitter.com/PrimeVideo/status/1323411528318480384

Top 2

Figure 58: https://twitter.com/PrimeVideo/status/1498875384833708032

Top 3

Figure 59: https://twitter.com/PrimeVideo/status/1536785617559986176

Top 4

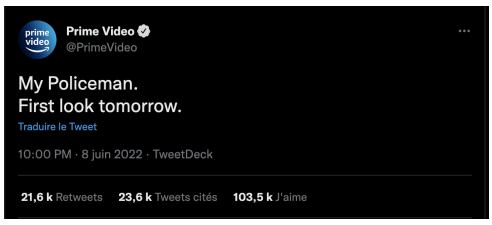

Figure 60: https://twitter.com/PrimeVideo/status/1534626340611051520

Top 5

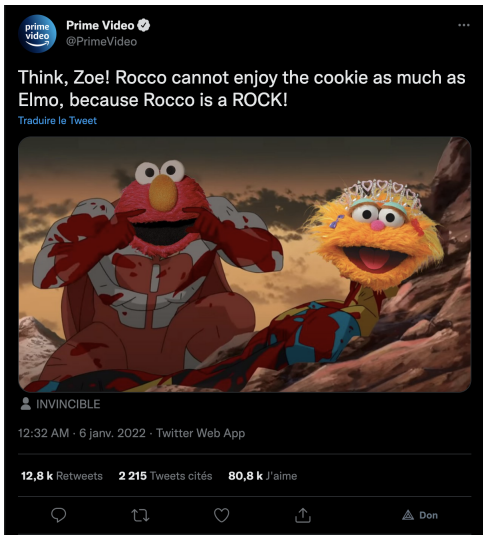

Figure 61: https://twitter.com/PrimeVideo/status/1478871982531776512

Top 6

Figure 62: https://twitter.com/PrimeVideo/status/1318597849014243329

Top 7

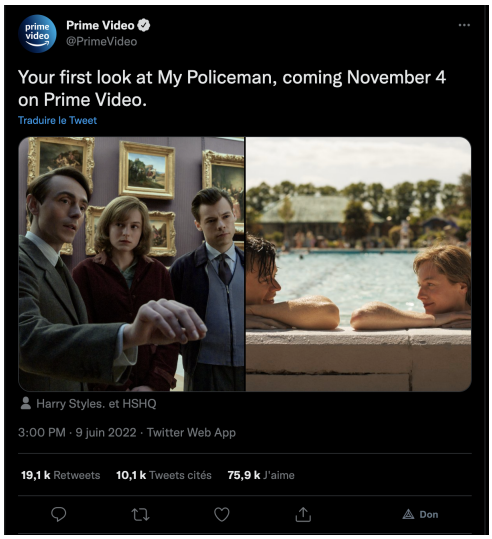

Figure 63: https://twitter.com/PrimeVideo/status/1534883060046409729

Top 8

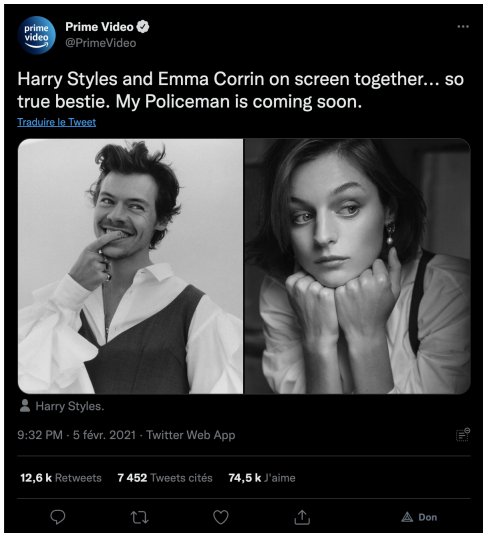

Figure 64: https://twitter.com/PrimeVideo/status/1357789204147703808

Top 9

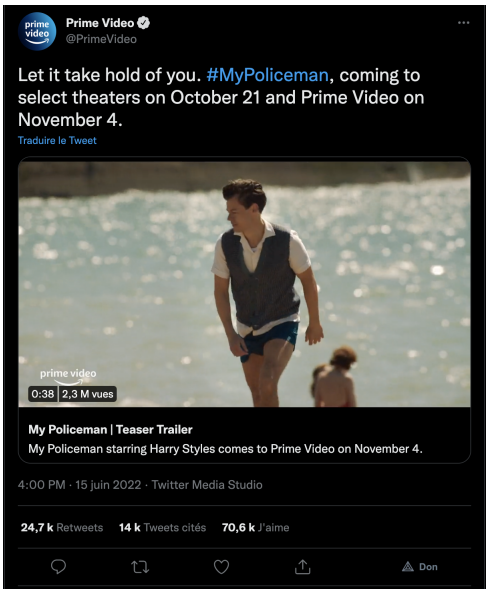

Figure 65: https://twitter.com/PrimeVideo/status/1537072548340547585

Top 10

Figure 66: https://twitter.com/PrimeVideo/status/1541138934851375104

7.2.5 DisneyPlusFR

Top 1

Figure 67: https://twitter.com/DisneyPlusFR/status/1247410735526948864

Top 2

Figure 68: https://twitter.com/DisneyPlusFR/status/1318507263259279362

Top 3

Figure 69: https://twitter.com/DisneyPlusFR/status/1235501542758891523

Top 4

Figure 70: https://twitter.com/DisneyPlusFR/status/1245018015336239105

Top 5

Figure 71: https://twitter.com/DisneyPlusFR/status/1247101921023930374

Top 6

Figure 72: https://twitter.com/DisneyPlusFR/status/1248536841093070849

Top 7

Figure 73: https://twitter.com/DisneyPlusFR/status/1239808687092498433

Top 8

Figure 74: https://twitter.com/DisneyPlusFR/status/1247207470247936000

Top 9

Figure 75: https://twitter.com/DisneyPlusFR/status/1241418712512696320

Top 10

Figure 76: https://twitter.com/DisneyPlusFR/status/1240322342775070721

7.2.6 DisneyPlus

Top 1

Figure 77: https://twitter.com/disneyplus/status/1289092933598081025

Top 2

Figure 78: https://twitter.com/disneyplus/status/1491528787346923521

Top 3

Figure 79: https://twitter.com/disneyplus/status/1299557758983831559

Top 4

Figure 80: https://twitter.com/disneyplus/status/1212432983157694464

Top 5

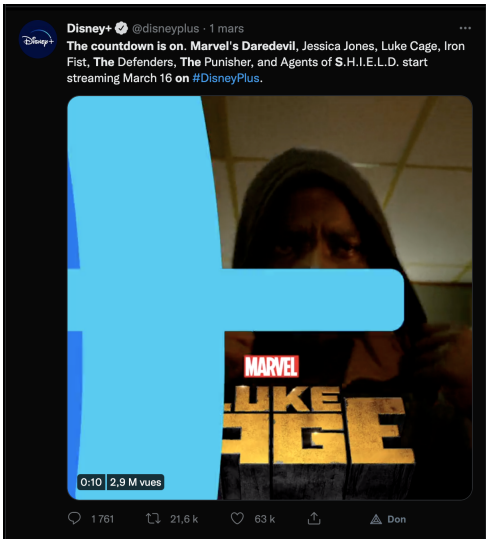

Figure 81: https://twitter.com/disneyplus/status/1498689548536057860

Top 6

Figure 82: https://twitter.com/disneyplus/status/1486033396937547780

Top 7

Figure 83: https://twitter.com/disneyplus/status/1522260015398064129

Top 8

Figure 84: https://twitter.com/disneyplus/status/1331736259660709888

Top 9

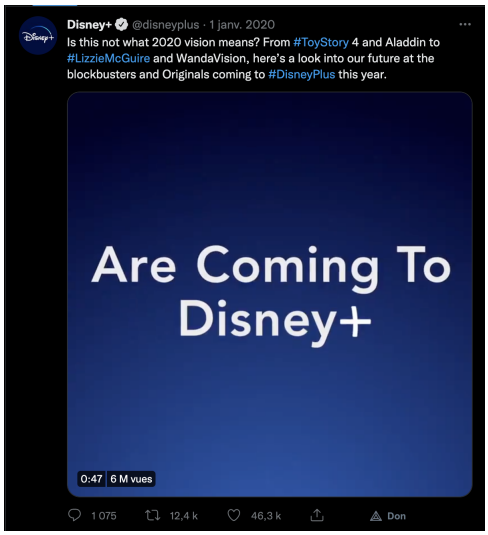

Figure 85: https://twitter.com/disneyplus/status/1212372907399118848

Top 10

Figure 86: https://twitter.com/disneyplus/status/1357377125133651968

7.3 Code

7.3.1 Data extraction

Example for one account :

```
df_net_06_21=pd.read_csv('/Users/hugolucca/Documents/Memoire/PythonCode-master/T

df_net_10_21=pd.read_csv('/Users/hugolucca/Documents/Memoire/PythonCode-master/T

df_net_02_22=pd.read_csv('/Users/hugolucca/Documents/Memoire/PythonCode-master/T

list_net.append(df_net_06_21)

list_net.append(df_net_10_21)

list_net.append(df_net_02_22)

df_net = pd.concat(list_net, ignore_index=True)

df_net.drop_duplicates(subset ="id", keep = 'first', inplace=True)
df_net
```

7.3.2 Data cleaning

```
def clean_bots(list_df):
    for dataframe in list_df:
        all_text=[]
        count=0
        df_test=dataframe
        for row in df_test.itertuples():
            text_to_rep=''
            clean_split=[]
            splitted=row.text.split()
            for el in splitted:
                if '@' not in el:
                    clean_split.append(el)
            for el in clean_split[:7]:
                text_to_rep+=el+' '


            if text_to_rep not in all_text:

                all_text.append(text_to_rep)

            else:

                count+=1
                df_test.at[row.Index, 'text'] = ''
                df_test.at[row.Index, 'created_at']=''
                df_test.at[row.Index, 'retweet_count']=0
                df_test.at[row.Index, 'favorite_count']=0
                df_test.at[row.Index, 'id']=0
        print(count)
        index_list = df_test.index[ df_test.loc[:,'text'] == '']
        df_test.drop(index_list,0, inplace=True)
```

7.3.3 Features extraction

```
def create_columns(dataframe):
    dataframe['is_pic_vid']=0
    for i in dataframe.index:

        if 't.co' in dataframe['text'][i]:
            dataframe['is_pic_vid'][i]=1
```

```
    dataframe['is_reply']=0
    for i in dataframe.index:

        if dataframe['text'][i][0]=='@':
            dataframe['is_reply'][i]=1

    dataframe['is_retweet']=0
    for i in dataframe.index:
        if dataframe['text'][i][:2]=='RT':
            dataframe['is_retweet'][i]=1

#netfr
groups = df_netfr.groupby(df_netfr.is_reply)
noreply_netfr_df = groups.get_group(0)
replies_netfr_df=groups.get_group(1)

groups = df_netfr.groupby(df_netfr.is_retweet)
noretweet_netfr_df = groups.get_group(0)
retweets_netfr_df=groups.get_group(1)

groups = noreply_netfr_df.groupby(noreply_netfr_df.is_retweet)
only_tweets_netfr_df = groups.get_group(0)
only_retweets_netfr_df=groups.get_group(1)

groups = noretweet_netfr_df.groupby(noretweet_netfr_df.is_reply)
not_useful_netfr = groups.get_group(0)
only_replies_netfr_df=groups.get_group(1)

#net


groups = df_net.groupby(df_net.is_reply)
noreply_net_df = groups.get_group(0)
replies_net_df=groups.get_group(1)

groups = df_net.groupby(df_net.is_retweet)
noretweet_net_df = groups.get_group(0)
retweets_net_df=groups.get_group(1)

groups = noreply_net_df.groupby(noreply_net_df.is_retweet)
only_tweets_net_df = groups.get_group(0)
only_retweets_net_df=groups.get_group(1)
```

```
groups = noretweet_net_df.groupby(noretweet_net_df.is_reply)
not_useful_net = groups.get_group(0)
only_replies_net_df=groups.get_group(1)

#primefr
groups = df_primefr.groupby(df_primefr.is_reply)
noreply_primefr_df = groups.get_group(0)
replies_primefr_df=groups.get_group(1)

groups = df_primefr.groupby(df_primefr.is_retweet)
noretweet_primefr_df = groups.get_group(0)
retweets_primefr_df=groups.get_group(1)

groups = noreply_primefr_df.groupby(noreply_primefr_df.is_retweet)
only_tweets_primefr_df = groups.get_group(0)
only_retweets_primefr_df=groups.get_group(1)

groups = noretweet_primefr_df.groupby(noretweet_primefr_df.is_reply)
not_useful_primefr = groups.get_group(0)
only_replies_primefr_df=groups.get_group(1)


#prime
groups = df_prime.groupby(df_prime.is_reply)
noreply_prime_df = groups.get_group(0)
replies_prime_df=groups.get_group(1)

groups = df_prime.groupby(df_prime.is_retweet)
noretweet_prime_df = groups.get_group(0)
retweets_prime_df=groups.get_group(1)

groups = noreply_prime_df.groupby(noreply_prime_df.is_retweet)
only_tweets_prime_df = groups.get_group(0)
only_retweets_prime_df=groups.get_group(1)

groups = noretweet_prime_df.groupby(noretweet_prime_df.is_reply)
not_useful_prime = groups.get_group(0)
only_replies_prime_df=groups.get_group(1)

#disneyfr
groups = df_disneyfr.groupby(df_disneyfr.is_reply)
```

```
noreply_disneyfr_df = groups.get_group(0)
replies_disneyfr_df=groups.get_group(1)

groups = df_disneyfr.groupby(df_disneyfr.is_retweet)
noretweet_disneyfr_df = groups.get_group(0)
retweets_disneyfr_df=groups.get_group(1)

groups = noreply_disneyfr_df.groupby(noreply_disneyfr_df.is_retweet)
only_tweets_disneyfr_df = groups.get_group(0)
only_retweets_disneyfr_df=groups.get_group(1)

groups = noretweet_disneyfr_df.groupby(noretweet_disneyfr_df.is_reply)
not_useful_disneyfr = groups.get_group(0)
only_replies_disneyfr_df=groups.get_group(1)

#disney
groups = df_disney.groupby(df_disney.is_reply)
noreply_disney_df = groups.get_group(0)
replies_disney_df=groups.get_group(1)

groups = df_disney.groupby(df_disney.is_retweet)
noretweet_disney_df = groups.get_group(0)
retweets_disney_df=groups.get_group(1)

groups = noreply_disney_df.groupby(noreply_disney_df.is_retweet)
only_tweets_disney_df = groups.get_group(0)
only_retweets_disney_df=groups.get_group(1)

groups = noretweet_disney_df.groupby(noretweet_disney_df.is_reply)
not_useful_disney = groups.get_group(0)
only_replies_disney_df=groups.get_group(1)




all_df=[df_netfr,df_net,df_primefr,df_prime,df_disneyfr,df_disney]
all_df_fr=[df_netfr,df_primefr,df_disneyfr]
all_df_en=[df_net,df_prime,df_disney]

only_tweets_list_fr=[only_tweets_disneyfr_df,only_tweets_primefr_df,only_tweets_
only_tweets_list_en=[only_tweets_disney_df,only_tweets_prime_df,only_tweets_net_
```

```
only_replies_list_fr=[only_replies_disneyfr_df,only_replies_primefr_df,only_repl
only_replies_list_en=[only_replies_disney_df,only_replies_prime_df,only_replies_

only_retweets_list_fr=[only_retweets_disneyfr_df,only_retweets_primefr_df,only_r
only_retweets_list_en=[only_retweets_disney_df,only_retweets_prime_df,only_retwe
```

7.3.4 Metrics

Type of posts

```
#calculer le nombre de posts de chaque type (tweet, retweet, r ponse)
posts_types={}

def count_tweets_replies_rts(dataframe):

    posts_types[dataframe.name]={'tweets':0,'rts':0,'rep':0}

    for row in dataframe.itertuples():
        if row.is_reply==0 and row.is_retweet==0:
            posts_types[dataframe.name]['tweets']+=1

    for row in dataframe.itertuples():
        if row.is_reply==1:
            posts_types[dataframe.name]['rep']+=1

    for row in dataframe.itertuples():
        if  row.is_retweet==1:
            posts_types[dataframe.name]['rts']+=1

#calculer le nombre moyen de posts hebdo de chaque type (tweet, retweet, r ponse
avg_weekly={}
def weekly_stats(dataframe):
    avg_weekly[dataframe.name]={'weekly_tweets':0,'weekly_rts':0,'weekly_rep':0}
    avg_weekly[dataframe.name]['weekly_tweets']=posts_types[dataframe.name]['twe
    avg_weekly[dataframe.name]['weekly_rts']=posts_types[dataframe.name]['rts']//
    avg_weekly[dataframe.name]['weekly_rep']=posts_types[dataframe.name]['rep']//


#calculer le pourcentage de posts de chaque type (tweet, retweet, r ponse)
perc_posts_types={}

def percentage_of_posts_types(dataframe):
    perc_posts_types[dataframe.name]={'perc_tweets':0,'perc_rts':0,'perc_rep':0}
```

```
    total_posts=posts_types[dataframe.name]['tweets']+posts_types[dataframe.name
    perc_posts_types[dataframe.name]['perc_tweets']=round(posts_types[dataframe.
    perc_posts_types[dataframe.name]['perc_rts']=round(posts_types[dataframe.nam
    perc_posts_types[dataframe.name]['perc_rep']=round(posts_types[dataframe.nam
```

Pictures and videos

```
def count_perc_pics(list_df):
    for dataframe in list_df:
        nb=0
        nb_pic=0
        for row in dataframe.itertuples():

            nb+=1
            if row.is_pic_vid:
                nb_pic+=1
        print(dataframe.name,nb_pic/nb)
```

Tweets over 10K likes

```
def count_nb_tweets_over_10K(list_df):
    for dataframe in list_df:
        nb=0
        nb_pic=0
        for row in dataframe.itertuples():
            if row.favorite_count>10000:
                nb+=1
                if row.is_pic_vid:
                    nb_pic+=1
        print(dataframe.name,nb,nb_pic)
```

Population reached

```
#data direclty extracted from web on 30th June
nb_franco=300000000
nb_anglo=1268000000

nb_follows_netfr=7400000
nb_follows_primefr=279700
nb_follows_disneyfr=301800
nb_follows_net=17500000
nb_follows_prime=2300000
nb_follows_disney=3500000
```

```
ratio_netfr=nb_follows_netfr/nb_franco
ratio_primefr=nb_follows_primefr/nb_franco
ratio_disneyfr=nb_follows_disneyfr/nb_franco
ratio_net=nb_follows_net/nb_anglo
ratio_prime=nb_follows_prime/nb_anglo
ratio_disney=nb_follows_disney/nb_anglo

print('NetflixFR : '+str(ratio_netfr*100)+'\t'+'PrimeVideoFR : '+str(ratio_prime
```

7.3.5 Evolutions

```
def nb_posts_per_m(list_df):

    for dataframe in list_df:
        test={}



        for row in dataframe.itertuples():
            var_row=str(row.year)+str(row.month)

            if var_row not in test:
                test[var_row]=1
            elif var_row in test:
                test[var_row]+=1
        print(dataframe.name,test)
```

7.3.6 Top words

```
import spacy
import fr_core_news_md
import en_core_web_sm



def get_most_used_words(list_df,lang):
    if lang=='fr':

        nlp = fr_core_news_md.load()

    elif lang=='en':
        nlp = en_core_web_sm.load()
```

```
for df in list_df:
    all_sentences = []

    for word in df['text']:
        all_sentences.append(word)


    lines = list()
    for line in all_sentences:
        words = line.split()
        for w in words:
            lines.append(w)



    #Removing Punctuation

    lines = [re.sub(r'[^A-Za-z0-9]+', '', x) for x in lines]

    lines

    lines2 = []

    for word in lines:
        if word != '':
            lines2.append(word.lower())
    lines2

    stem2 = []
    list_banned_words=['cest','the','srie','1','2','3','4','5','6','7','8','
    for word in lines2:
        if word not in nlp.Defaults.stop_words and word not in list_banned_w
            stem2.append(word)

    stem2

    df = pd.DataFrame(stem2)

    df = df[0].value_counts()
```

```
#df
#df['freq'] = df.groupby(0)[0].transform('count')
#df['freq'] = df.groupby(0)[0].transform('count')
#df.sort_values(by = ('freq'), ascending=False)



from nltk.probability import FreqDist

freqdoctor = FreqDist()

for words in df:
    freqdoctor[words] += 1

import matplotlib.pyplot as plt; plt.rcdefaults()
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns


df = df[:20,]
plt.figure(figsize=(10,5))
sns.barplot(df.values, df.index, palette="mako")
plt.title('Top Words Overall')
plt.ylabel('Word from Tweet', fontsize=12)
plt.xlabel('Count of Words', fontsize=12)
plt.show()
```

7.3.7 LSA

Imports

```
import pandas as pd
import glob
import re
import os
import seaborn as sns
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
import numpy as np
import string
from mpl_toolkits.mplot3d import Axes3D
```

```
from sklearn.manifold import TSNE
from nltk.tokenize import word_tokenize

import pandas as pd
import matplotlib.pyplot as plt
import matplotlib
import seaborn as sns
from wordcloud import WordCloud, STOPWORDS
from textblob import TextBlob
import spacy
import scattertext as st
from IPython.display import display, Markdown
from nltk.sentiment.util import *
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
from textblob import TextBlob
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
import numpy as np
matplotlib.style.use('ggplot')
%matplotlib inline

from sklearn.decomposition import TruncatedSVD
from sklearn.decomposition import LatentDirichletAllocation
from sklearn.manifold import TSNE
from collections import Counter
from sklearn.ensemble import RandomForestClassifier
from sklearn.pipeline import Pipeline
from sklearn.metrics import accuracy_score, classification_report
from sklearn.feature_selection import chi2
```

Data preparation

```
def remove_picture_url(text):

    text = re.sub(r"pic.twitter.com\S+", '', text)
    return text




def remove_url(text):

    text = re.sub(r"http\S+","", text)
```

```
    return text



# removing punctuation
def remove_punc(text):

    text = "".join([char for char in text if char not in string.punctuation])
    text = re.sub('[0-9]+', ' ', text)
    return text



# Tokenization
def tokenization(text):

    text = re.split('\W+', text)
    return text



# Defining stopwords

def remove_stopwords(text,lang):

    if lang=='fr':
        stopword = nltk.corpus.stopwords.words('french')
    elif lang=='en':
        stopword = nltk.corpus.stopwords.words('english')

    text = [word for word in text if word not in stopword]
    return text



# Stemming
ps = nltk.PorterStemmer()

def stemming(text):

    text = [ps.stem(word) for word in text]
    return text
```

```
# Lemmatization
wn = nltk.WordNetLemmatizer()

def lemmatizer(text):

    text = [wn.lemmatize(word) for word in text]
    return text


def treat_text_en(list_df):
    for dataframe in list_df:
        dataframe['text'] = dataframe['text'].apply(lambda x: remove_picture_url

        dataframe['text'] = dataframe['text'].apply(lambda x: remove_url(x))

        dataframe['tweet_punct'] = dataframe['text'].apply(lambda x: remove_punc

        dataframe['tweet_tokenized'] = dataframe['tweet_punct'].apply(lambda x:

        dataframe['tweet_nonstop'] = dataframe['tweet_tokenized'].apply(lambda x

        dataframe['tweet_stemmed'] = dataframe['tweet_nonstop'].apply(lambda x:

        dataframe['tweet_lemmatized'] = dataframe['tweet_nonstop'].apply(lambda
        dataframe.head()

def treat_text_fr(list_df):
    for dataframe in list_df:
        dataframe['text'] = dataframe['text'].apply(lambda x: remove_picture_url

        dataframe['text'] = dataframe['text'].apply(lambda x: remove_url(x))

        dataframe['tweet_punct'] = dataframe['text'].apply(lambda x: remove_punc

        dataframe['tweet_tokenized'] = dataframe['tweet_punct'].apply(lambda x:

        dataframe['tweet_nonstop'] = dataframe['tweet_tokenized'].apply(lambda x

        dataframe['tweet_stemmed'] = dataframe['tweet_nonstop'].apply(lambda x:
```

```
        dataframe['tweet_lemmatized'] = dataframe['tweet_nonstop'].apply(lambda
        dataframe.head()

def convert (list_df):
    for dataframe in list_df :
        dataframe['created_at'] = pd.to_datetime(dataframe['created_at'], infer_
        dataframe['favorite_count'] = pd.to_numeric(dataframe['favorite_count'],
        dataframe = dataframe.dropna()

def clean(list_df):
    for dataframe in list_df:

        dataframe['test']=''
        for i in dataframe.index:
            test2=''
            for el in dataframe['tweet_lemmatized'][i]:
                test2+=el+' '
            dataframe['test'][i]=test2
        dataframe
```

Useful functions

```
def get_keys(topic_matrix):

    keys = topic_matrix.argmax(axis=1).tolist()
    return keys

def keys_to_counts(keys):

    count_pairs = Counter(keys).items()
    categories = [pair[0] for pair in count_pairs]
    counts = [pair[1] for pair in count_pairs]
    return (categories, counts)

def get_top_n_words(n, keys, document_term_matrix, tfidf_vectorizer,n_topics):

    top_word_indices = []
    for topic in range(n_topics):
        temp_vector_sum = 0
        for i in range(len(keys)):
            if keys[i] == topic:
                temp_vector_sum += document_term_matrix[i]
        temp_vector_sum = temp_vector_sum.toarray()
        top_n_word_indices = np.flip(np.argsort(temp_vector_sum)[0][-n:],0)
```

```
        top_word_indices.append(top_n_word_indices)
    top_words = []
    for topic in top_word_indices:
        topic_words = []
        for index in topic:
            temp_word_vector = np.zeros((1,document_term_matrix.shape[1]))
            temp_word_vector[:,index] = 1
            the_word = tfidf_vectorizer.inverse_transform(temp_word_vector)[0][0
            topic_words.append(the_word.encode('utf-8').decode('utf-8'))
        top_words.append(" ".join(topic_words))
    return top_words
```

Display

```
def plot_cat(dataframe):

    print(dataframe.name)
    tfidf_vectorizer = TfidfVectorizer(stop_words='english', use_idf=True, s


    reindexed_data = dataframe['test']
    reindexed_data = reindexed_data.values

    document_term_matrix = tfidf_vectorizer.fit_transform(dataframe['test'])
    n_topics = 6

    lsa_model = TruncatedSVD(n_components = n_topics)
    lsa_topic_matrix = lsa_model.fit_transform(document_term_matrix)




    lsa_keys = get_keys(lsa_topic_matrix)
    lsa_categories, lsa_counts = keys_to_counts(lsa_keys)

    top_n_words_lsa = get_top_n_words(10, lsa_keys, document_term_matrix, tfi
    for i in range(len(top_n_words_lsa)):
        print("Topic {}: ".format(i+1), top_n_words_lsa[i])

    top_3_words = get_top_n_words(3, lsa_keys, document_term_matrix, tfidf_v
    labels = ['Topic {}: \n'.format(i) + top_3_words[i] for i in lsa_categor
    print(lsa_counts)

    fig, ax = plt.subplots(figsize=(16,8))
```

```
        sns.barplot(lsa_categories, lsa_counts)
        ax.set_xticks(lsa_categories)
        ax.set_xticklabels(labels)
        ax.set_ylabel('Number of tweet text')
        ax.set_title('LSA topic counts')
        plt.show()
```

7.3.8 Most popular tweets

```
def get_most_popular_tweets(list_df):
    for dataframe in list_df:
        print(dataframe.name)
        list_most_pop=[]
        while len(list_most_pop)<20:
            most_pop=0
            for i in dataframe.index:
                if dataframe['favorite_count'][i]>most_pop and (i not in list_mo
                    most_pop=dataframe['favorite_count'][i]
                    max_id=i

            list_most_pop.append(max_id)


        print(dataframe.loc[list_most_pop])
```

References

[1] Rubenking, B. and C. C. Bracken (2021). Binge watching and serial viewing: Comparing new media viewing habits in 2015 and 2020. Addictive Behavior Reports Volume 14.

[2] Wu, H., Z. Shenghua , L. Ling (2012). Social media competitive analysis and text mining: A case study in the pizza industry. International Journal of Information Management.

[3] Oliveira, L. and A. Figueira (2015). Benchmarking analysis of social media strategies in the Higher Education Sector. Procedia Computer Science 64 (2015) 779 – 786.

[4] Martinez-Sanchez, M. E., R. Nicolas-Sans, J. Bustos Diaz (2021). Analysis of the social media strategy of audio-visual OTTs in Spain: The case study of Netflix, HBO and Amazon Prime during the implementation of Disney +. Technological Forecasting & Social Change.

[5] Lin, Y., C. Weitong, L. Xue , Z. Wanli, Y. Winghao (2018). A survey of sentiment analysis in social media.

[6] P B. Floreddu and F. Cabiddu , (2016),"Social media communication strategies", Journal of Services Marketing, Vol. 30 Iss 5 pp.

[7] Scerbinina (2019). Netflix relationship marketing on Twitter.

[8] A. Hotho, A. Nurnberger, G. Paass (2005). A Brief Survey of Text Mining.

[9] E. Fernández-Gómez and J. Martín-Quevedo , (2018),"La estrategia de engagement de Netflix España en Twitter", El Profesional de la Informacion.

[10] C. Labovitz , (2020), Early effects of COVID-19 lockdowns on service provider networks: the networks soldier on!, Nokia Blog.

[11] O. L. Vyatkina , (2020), The Impact Of Influencer Marketing On The Global Economy, Global Challenges and Prospects of the Modern Economic Development.

[12] Netflix France [@NetflixFR]. (2020-2022). [Twitter account : https://twitter.com/NetflixFR]. Twitter.

[13] Netflix US [@Netflix]. (2020-2022). [Twitter account; https://twitter.com/Netflix]. Twitter.

[14] PrimeVideo France [@PrimeVideoFR]. (2020-2022). [Twitter account : https://twitter.com/PrimeVideoFR]. Twitter.

[15] PrimeVideo US [@PrimeVideo]. (2020-2022). [Twitter account : https://twitter.com/PrimeVideo]. Twitter.

[16] DisneyPlus France [@DisneyPlusFR]. (2020-2022). [Twitter account : https://twitter.com/DisneyPlusFR]. Twitter.

[17] DisneyPlus US [@DisneyPlus]. (2020-2022). [Twitter account : https://twitter.com/DisneyPlus]. Twitter.