

# *Bad Gyal*

Anàlisi de les plataformes i serveis de streaming utilitzats per l'artista:

Estudi de la conversa i l'impacte a Spotify, YouTube i Twitter

**AUTOR: JAUME SACRISTAN VIA - TUTOR: ADRIÀ PADILLA MOLINA**  
**TFG COMUNICACIÓ INTERACTIVA - UNIVERSITAT AUTÒNOMA DE BARCELONA**

# Treball de Fi de Grau

## Títol

Anàlisi de les plataformes i serveis de *streaming* utilitzats per l'artista Bad Gyal: estudi de la conversa i l'impacte a Spotify, YouTube i Twitter.

## Autoria

Jaume Sacristan Via

## Professorat tutor

**Adrià Padilla Molina**

## Grau

|                                   |   |
|-----------------------------------|---|
| Comunicació Audiovisual           |   |
| Periodisme                        |   |
| Publicitat i Relacions Pùbliques  |   |
| Comunicació Interactiva           | X |
| Comunicació de les Organitzacions |   |

## Tipus de TFG

|          |   |
|----------|---|
| Projecte |   |
| Recerca  | X |

## Data

|                      |   |
|----------------------|---|
| 1 de juny de 2023    | X |
| 28 de juliol de 2023 |   |

# Full resum del TFG

## Títol del Treball Fi de Grau:

|                    |  |                      |                                   |   |
|--------------------|--|----------------------|-----------------------------------|---|
| Català:            | Anàlisi de les plataformes i serveis de <i>streaming</i> utilitzats per l'artista Bad Gyal: estudi de la conversa i l'impacte a Spotify, YouTube i Twitter.            |                      |                                   |   |
| Castellà:          | Análisis de las plataformas y servicios de <i>streaming</i> utilizados por la artista Bad Gyal: estudio de la conversación y el impacto en Spotify, YouTube y Twitter. |                      |                                   |   |
| Anglès:            | Analysis of platforms and streaming services used by the artist Bad Gyal: study of conversation and impact on Spotify, YouTube and Twitter.                            |                      |                                   |   |
| Autoria:           |  | Jaume Sacristan Via  |                                   |   |
| Professorat tutor: |  | Adrià Padilla Molina |                                   |   |
| Curs:              | 2022/23  | Grau:                | Comunicació Audiovisual           |   |
|                    |  |                      | Periodisme                        |   |
|                    |  |                      | Publicitat i Relacions Pùbliques  |   |
|                    |  |                      | Comunicació Interactiva           | X |
|                    |  |                      | Comunicació de les Organitzacions |   |

## Paraules clau (mínim 3)

|           |   |
|-----------|---|
| Català:   | <b>Bad Gyal, xarxes socials, interacció, Big Data</b>   |
| Castellà: | <b>Bad Gyal, redes sociales, interacción, Big Data</b>  |
| Anglès:   | <b>Bad Gyal, social networks, interaction, Big Data</b> |

## Resum del Treball Fi de Grau (extensió màxima 100 paraules)

|         |  |
|---------|--|
| Català: | L'objectiu principal del treball és analitzar la conversa a les plataformes i serveis de <i>streaming</i> al voltant de l'artista Bad Gyal. Això es fa per poder entendre com s'organitzen els usuaris entre sí i de quina manera interactuen amb el contingut relacionat amb l'artista. Les dades del treball s'han extret directament de les API de les plataformes i xarxes socials analitzades: Spotify, YouTube i Twitter. S'ha creat un codi de programació en Python per extreure les dades i per poder comprendre-les s'han creat gràfics amb Gephi i Tableau. Finalment, s'estreuen les conclusions de l'estat de Bad Gyal en aquestes plataformes. |
|---------|--|

|           |   |
|-----------|---|
| Castellà: | <p>El objetivo principal del trabajo es analizar la conversación en las plataformas y servicios de <i>streaming</i> de la artista Bad Gyal. Esto se hace para poder entender cómo se organizan los usuarios entre sí y cómo interactúan con el contenido relacionado. Los datos del trabajo se han extraído directamente de las API de las plataformas y redes sociales analizadas: Spotify, YouTube y Twitter. Se ha creado un código de programación en Python para extraer los datos y para poder comprenderlos se han creado gráficos con Gephi y Tableau. Finalmente, se extraen las conclusiones del estado de Bad Gyal en plataformas.</p> |
| Anglès:   | <p>The main objective is to analyze the conversation on platforms and streaming services around the artist Bad Gyal. This is done to be able to understand how users organize themselves and how they interact with content related to the artist. The data has been extracted directly from the APIs of the analyzed platforms and streaming services: Spotify, YouTube and Twitter. A programming code has been created in Python to extract the data and to understand them, graphics have been created with Gephi and Tableau. Finally, an analysis of the state of these platforms around the artist is extracted.</p>                       |

# ANÀLISI DE LES PLATAFORMES I SERVEIS DE STREAMING UTILITZATS PER L'ARTISTA BAD GYAL:

ESTUDI DE LA CONVERSA I L'IMPACTE A SPOTIFY,  
YOUTUBE I TWITTER

Jaume Sacristan Via  
UNIVERSITAT AUTÒNOMA DE BARCELONA  
Grau en Comunicació Interactiva

# Índex de continguts

|   |    |
|---|----|
| Índex d'il·lustracions .....                                  | 3  |
| 1. Introducció .....  | 7  |
| 1.1. Introducció del tema .....                               | 8  |
| 1.2. Objectius .....  | 8  |
| 1.3. Estructura del treball.....                              | 9  |
| 2. Marc teòric .....  | 12 |
| 2.1. Consum musical i Spotify .....                           | 12 |
| 2.2. Les plataformes i els serveis d'àudio sota demanda ..... | 13 |
| 2.3. Metodologia d'anàlisi de plataformes .....               | 15 |
| 2.4. Bad Gyal.....  | 16 |
| 2.5. Plataformes i indústria musical.....                     | 17 |
| 3. Objectius i preguntes de recerca .....                     | 20 |
| 3.1. Objectius quantitatius .....                             | 20 |
| 3.2. Objectius qualitatius .....                              | 25 |
| 4. Metodologia .....  | 27 |
| 4.1. Obtenció de les dades.....                               | 27 |
| 4.2. Visualització de les dades .....                         | 27 |
| 4.3. Mètriques i KPI .....                                    | 28 |
| Objectius quantitatius.....                                   | 28 |
| Objectius qualitatius .....                                   | 29 |
| 5. Anàlisi .....  | 30 |
| 5.1. Spotify .....  | 30 |
| 5.1.1. Llistat de cançons a Spotify .....                     | 31 |

|  |     |
|--|-----|
| 5.1.2. Anàlisi de la relació amb altres artistes a Spotify ..... | 41  |
| 5.1.3. Anàlisi de la relació amb altres gèneres a Spotify.....   | 46  |
| 5.2. YouTube.....  | 50  |
| 5.2.1. Font de l'extracció de dades .....                        | 51  |
| 5.2.2. Anàlisi de la informació del canal .....                  | 52  |
| 5.2.3. Anàlisi de la informació de cada vídeo.....               | 54  |
| 5.3.5. Anàlisi de les interaccions de Bad Gyal a YouTube.....    | 70  |
| 5.3. Twitter.....  | 76  |
| 5.3.1. Anàlisi de l'evolució temporal a Twitter .....            | 77  |
| 5.3.2. Anàlisi del tipus de contingut a Twitter.....             | 81  |
| 5.3.3. Anàlisi del tipus d'interaccions a Twitter.....           | 84  |
| 5.3.4. Anàlisi de l'engagement rate a Twitter .....              | 87  |
| 5.3.5. Anàlisi de les relacions entre usuaris a Twitter .....    | 90  |
| 5.3.6. Anàlisi del contingut dels tweets .....                   | 105 |
| 6. Conclusions .....   | 112 |
| 6.1. Anàlisi dels objectius .....                                | 112 |
| 6.2. Conclusions generals .....                                  | 119 |
| Bibliografia.....  | 122 |
| Annexos.....   | 126 |

# Índex d'il·lustracions

|  |    |
|--|----|
| II·Il·lustració 1: Evolució temporal per dia. Font: Elaboració pròpia.....   | 34 |
| II·Il·lustració 2: Evolució temporal per cançó i àlbum per trimestre. Font: Elaboració pròpia.....   | 35 |
| II·Il·lustració 3: Evolució temporal per cançó i àlbum per trimestres 2022-2023. Font: Elaboració pròpia.....                              | 36 |
| II·Il·lustració 4: Evolució temporal per cançó i àlbum per trimestres 2022-2023 (sense àlbums recopilatoris). Font: Elaboració pròpia..... | 37 |
| II·Il·lustració 5: Evolució temporal per cançó i àlbum per trimestres sense recopilatoris. Font: Elaboració pròpia.....                    | 38 |
| II·Il·lustració 6: Publicacions per dia de la setmana. Font: Elaboració pròpia .....   | 38 |
| II·Il·lustració 7: Contingut explícit a Spotify. Font: Elaboració pròpia .....   | 39 |
| II·Il·lustració 8: Evolució del contingut explícit per trimestre. Font: Elaboració pròpia. 40  | 40 |
| II·Il·lustració 9: Evolució del nombre de cançons respecte la seva durada. Font: Elaboració pròpia.....                                    | 40 |
| II·Il·lustració 10: Graf d'artistes relacionats amb Bad Gyal per grau de sortida. Font: Elaboració pròpia.....                             | 43 |
| II·Il·lustració 11: Graf d'artistes relacionats amb Bad Gyal per grau d'entrada .....  | 45 |
| II·Il·lustració 12: Graf de gèneres relacionats amb Bad Gyal per grau d'entrada. Font: Elaboració pròpia.....                              | 48 |
| II·Il·lustració 13: Evolució temporal del nombre de publicacions de Bad Gyal per mes a YouTube. Font: Elaboració pròpia. ....              | 60 |
| II·Il·lustració 14: Evolució temporal del nombre de publicacions de Bad Gyal per dia a YouTube. Font: Elaboració pròpia. ....              | 61 |
| II·Il·lustració 15: Evolució temporal del nombre de publicacions de Bad Gyal per any a YouTube. Font: Elaboració pròpia. ....              | 62 |
| II·Il·lustració 16: Nombre de publicacions de Bad Gyal per mes a YouTube. Font: Elaboració pròpia.....                                     | 63 |

|   |    |
|---|----|
| II·Il·lustració 17: Nombre de publicacions de Bad Gyal per dia de la setmana i any a YouTube. Font: Elaboració pròpia.                      | 64 |
| II·Il·lustració 18: Nombre de publicacions de Bad Gyal per mes i canal a YouTube. Font: Elaboració pròpia.                                  | 65 |
| II·Il·lustració 19: Evolució temporal de publicacions per canal a YouTube. Font: Elaboració pròpia.   | 66 |
| II·Il·lustració 20: Nombre de publicacions de Bad Gyal per hora del dia a YouTube. Font: Elaboració pròpia.                                 | 67 |
| II·Il·lustració 21: Vídeos sense llicència. Font: Elaboració pròpia.  | 68 |
| II·Il·lustració 22: Llicència per canal. Font: Elaboració pròpia.   | 68 |
| II·Il·lustració 23: Categoria per canal. Font: Elaboració pròpia.   | 68 |
| II·Il·lustració 24: Comparativa del nombre de views, like i comment dels vídeos de Bad Gyal a YouTube. Font: Elaboració pròpia.             | 72 |
| II·Il·lustració 25: Nombre d'interaccions respecte de les visualitzacions per vídeo de Bad Gyal a YouTube. Font: Elaboració pròpia.         | 72 |
| II·Il·lustració 26: Nombre d'interaccions respecte de les visualitzacions per vídeo i canal de Bad Gyal a YouTube. Font: Elaboració pròpia. | 73 |
| II·Il·lustració 27: Ràtio d'interacció respecte de les visualitzacions per vídeo de Bad Gyal a YouTube. Font: Elaboració pròpia.            | 74 |
| II·Il·lustració 28: Engagement rate per vídeo i canal de Bad Gyal a YouTube. Font: Elaboració pròpia.                                       | 75 |
| II·Il·lustració 29: Evolució temporal de la interacció per any, vídeo i canal a YouTube. Font: Elaboració pròpia.                           | 75 |
| II·Il·lustració 30: Evolució temporal de nombre de tweets per mes i tendència (2016-2023). Font: Elaboració pròpia.                         | 77 |
| II·Il·lustració 31: Nombre de tweets per mes de l'any. Font: Elaboració pròpia.   | 79 |
| II·Il·lustració 32: Nombre de tweets per dia de la setmana. Font: Elaboració pròpia.  | 80 |
| II·Il·lustració 33: Nombre de tweets per hora del dia. Font: Elaboració pròpia.   | 81 |
| II·Il·lustració 34: Tweets amb contingut multimèdia. Font: Elaboració pròpia.   | 82 |
| II·Il·lustració 35: Tipus de contingut multimèdia. Font: Elaboració pròpia.   | 82 |
| II·Il·lustració 36: Nombre de tweets amb contingut sensible (taronja) i sense contingut sensible (blau). Font: Elaboració pròpia.           | 82 |

|  |     |
|--|-----|
| II·Il·lustració 37: Nombre de tweets per direccionalitat de la publicació. Font: Elaboració pròpria.                                   | 83  |
| II·Il·lustració 38: Comparativa de mètriques per a cada tweet (ordenat per impressions). Font: Elaboració pròpria.....                 | 84  |
| II·Il·lustració 39: Comparativa de mètriques per a cada tweet (ordenat per retweets). Font: Elaboració pròpria.....                    | 86  |
| II·Il·lustració 40: Interaccions respecte del nombre d'impressions i el nombre de seguidors de l'usuari. Font: Elaboració pròpria..... | 87  |
| II·Il·lustració 41: Evolució de l'engagement rate (2022-2023). Font: Elaboració pròpria. ....  | 89  |
| II·Il·lustració 42: Nombre d'interaccions per tipus de contingut multimèdia. Font: Elaboració pròpria.....                             | 90  |
| II·Il·lustració 43: Graf amb les mencions de Bad Gyal a Twitter segons el grau d'entrada. Font: Elaboració pròpria.....                | 93  |
| II·Il·lustració 44: Ampliació de les comunitats al grup central del graf. Font: Elaboració pròpria. ....                               | 94  |
| II·Il·lustració 45: Ampliació de la comunitat lila. Font: Elaboració pròpria. ....   | 95  |
| II·Il·lustració 46: Mencions destacades comunitat lila. Font: Elaboració pròpria. ....   | 96  |
| II·Il·lustració 47: Mencions de l'usuari VNegro2. Font: Elaboració pròpria.....  | 97  |
| II·Il·lustració 48: Representació del compte YouTube al graf. Font: Elaboració pròpria. ....   | 98  |
| II·Il·lustració 49: Visió general de la comunitat verda. Font: Elaboració pròpria. ....  | 99  |
| II·Il·lustració 50: Vinculacions de la comunitat verda i lila. Font: Elaboració pròpria... ..  | 100 |
| II·Il·lustració 51: Ampliació de les vinculacions de BadGyalUpdates. Font: Elaboració pròpria. ....                                    | 101 |
| II·Il·lustració 52: Ampliació de bizarrap i gmrtzzz. Font: Elaboració pròpria.....   | 102 |
| II·Il·lustració 53: Comunitat negra. Font: Elaboració pròpria.....   | 103 |
| II·Il·lustració 54: Comunitat taronja. Font: Elaboració pròpria.....   | 104 |
| II·Il·lustració 55: Resposta del model Hate Speech. Font: Elaboració pròpria. ....   | 108 |
| II·Il·lustració 56: Resposta del model Distilbert. Font: Elaboració pròpria. ....  | 108 |
| II·Il·lustració 57: Resposta del model Roberta. Font: Elaboració pròpria. ....   | 109 |

|  |     |
|--|-----|
| II·lustració 58: Comparació de la mitjana de fiabilitat per a cada model. Font:<br>Elaboració pròpia.....      | 110 |
| II·lustració 59: Comparació de la fiabilitat dels models Roberta (eix y i colors) i Distilbert<br>(eix x)..... | 110 |

# 1. Introducció

Aquesta anàlisi de la conversa a les plataformes i serveis musicals sota demanda al voltant de l'artista Bad Gyal, està motivat per dos elements principals: l'interès personal en l'artista i la importància que té en la indústria musical a Espanya i especialment a Barcelona.

Bad Gyal és una artista que segueix des dels seus inicis. Personalment gaudeixo de la seva música i tot el món que l'envolta, ja que és una artista diferent amb la qual sento proximitat perquè és jove i de l'entorn de Barcelona. Bad Gyal té cançons a les principals plataformes i serveis de distribució de música, però també fa altres tipus de continguts com concerts, entrevistes, col·leccions de moda o marxandatge. Així doncs, la proposta artística de Bad Gyal no es limita exclusivament a la música, sinó que també ha estat capaç de crear tot un univers personal al voltant de la seva figura. I ho ha aconseguit fent col·laboracions amb la revista Vogue o fent una col·lecció de moda amb Bershka.

L'elecció d'aquesta artista és deguda al fet que coneix el seu entorn i que hi ha molta activitat al seu voltant, tant a plataformes com als serveis d'àudio sota demanda. Bad Gyal va néixer i créixer al voltant de Barcelona, on la seva carrera és coneguda des de fa molts anys i ha fet diversos concerts. Tot i que ja havia fet petites actuacions fora de Catalunya i arreu del món, a partir de l'any 2019 es va posicionar en la part més alta de les llistes de música urbana amb artistes com Omar Montes i Rauw Alejandro. Els gèneres *reggaeton* i *urban* són els més populars de la música a Espanya i també tenen un gran poder a escala mundial. És per això que l'artista té una gran importància en l'àmbit nacional, i a poc a poc, també de manera internacional.

Finalment, també m'agradaria destacar la meva motivació personal per l'anàlisi i la visualització de les dades de xarxes socials. Especialment amb les eines avançades que s'utilitzaran com APIs, Python, Tableau o Gephi.

## 1.1. Introducció del tema

Aquest estudi es tracta bàsicament d'una anàlisi de plataformes i serveis d'àudio sota demanda d'una marca. Concretament, s'analitza l'artista Bad Gyal, que és una marca en ella mateixa.

Alba Farelo, sota el nom artístic Bad Gyal, és una cantant, productora i compositora catalana. Va néixer el 12 de març del 1997 a Vilassar de Mar, Barcelona, i es va donar a conèixer el 2016 amb el llançament del seu primer tema "Pai", que tenia la base de la cançó "Work" de Rihanna, i que es va fer viral a les xarxes socials.

Les seves cançons s'adscriuen a un gran ventall de gèneres com *el reggaeton, trap, hip hop o R&B*. Però d'estaca especialment per la influència del *dancehall*, un estil jamaicà variant del *reggae* que convida a ballar. Les lletres són diverses i especialment tracten temes del seu dia a dia: treballar, créixer, l'amor i el desamor, la festa, les drogues o els diners. Sovint és emmarcada dins de la categoria *urban* que engloba a Espanya els diferents estils i temàtiques mencionades anteriorment.

L'artista, ja era coneguda en el món *underground* de la música *urban*, però es va donar a conèixer especialment després de signar amb importants discogràfiques i professionalitzar-se. Fet que li ha permès arribar en diferents ocasions a la primera posició de les llistes espanyoles més reconegudes. També ha col·laborat amb artistes com: Omar Montes, Rauw Alejandro, Soto Asa, Quevedo o Karol G.

Actualment Bad Gyal té un gran ventall de col·laboracions musicals i està molt ben posicionada, especialment en l'àmbit nacional. La seva influència també està arribant a nous col·lectius i és present en grans esdeveniments com concerts i festivals a l'alçada dels artistes musicals més importants a escala nacional i també en música en espanyol mundialment.

## 1.2. Objectius

L'objectiu principal és analitzar la conversa a les plataformes i serveis sota demanda al voltant de l'artista Bad Gyal.

Per aconseguir-ho, aquesta anàlisi es divideix en objectius secundaris més petits per tal que siguin assumibles. A continuació s'exposen i en el punt 3 es detalla amb més profunditat com es durà a terme cadascun.

Actualment és molt important en qualsevol anàlisi de continguts i estratègia digital mostrar dades quantitatives, perquè hi ha un gran nombre de paràmetres que es mesuren en aquest entorn. Concretament, els cinc primers objectius secundaris són quantitatius. D'altra banda, és important mesurar la qualitat d'aquestes dades. Per fer-ho, s'han creat els dos últims objectius secundaris (6-7) amb una visió qualitativa.

1. El primer és observar l'evolució temporal de l'artista a les plataformes i serveis d'àudio sota demanda.
2. El segon és analitzar quin tipus de contingut comparteix l'artista a les plataformes YouTube i Twitter i al servei d'àudio sota demanda Spotify.
3. El tercer és detectar el contingut amb més repercussió a YouTube i Twitter.
4. El quart és quantificar la resposta del públic segons l'usuari que publica el contingut.
5. El cinquè és realitzar una anàlisi comparativa entre YouTube, Twitter i Spotify.
6. El sisè es tracta de saber si la conversa és positiva, negativa o neutra per a Bad Gyal a Twitter.
7. El setè es tracta d'analitzar les relacions del públic de Bad Gyal amb artistes musicals i altres temes.

### 1.3. Estructura del treball

Per complir els objectius anteriors, el document d'entrega estarà dividit en sis fases. Després d'aquesta introducció, s'inicia el segon punt, on s'aprofundeix en el tema amb l'establiment del marc teòric. Cal exposar la situació en la qual ens trobem actualment, però també fer una anàlisi de la història més recent per tal de poder contextualitzar les dades que obtinguem i veure quina és la tendència general del sector. Aquest estudi cal fer-lo dels dos temes principals: el món de les plataformes i la indústria musical en

la qual s'emmarca Bad Gyal. Fent especial menció en la relació entre aquests dos mons.

Una vegada establert el marc teòric, cal concretar què volem investigar a través dels objectius secundaris. Per aconseguir destriar el que és important del que no ho és, s'han de formular les preguntes de recerca.

En quart lloc, cal establir la metodologia que s'utilitzarà. Dins d'aquest apartat s'especificarà com obtenim les dades, com les analitzem i per a què utilitzem cada dada dels indicadors triats.

El centre de l'estudi és l'anàlisi. Destaca la importància d'exposar les eines utilitzades per l'extracció, com són la programació de petits codis per comunicar-se amb les API; i també l'exposició de les dades extretes de manera comprensible, el filtratge i la visualització de les dades.

En l'últim apartat s'exposaran les conclusions. Primer, recuperant els objectius secundaris i analitzant-los, i després es farà una valoració final del treball. És important que en aquest apartat se sintetitzi la informació per tal de mostrar unes conclusions clares i comprensibles.

Cal destacar que aquest treball és experimental i que si bé s'ha tractat part de la teoria de manera transversal en el grau de comunicació interactiva, l'evolució i els canvis en l'àmbit del codi i de les plataformes és molt més ràpid. Per aquest motiu, hi ha poca documentació a l'abast de manera tradicional, i generalment s'ha utilitzat pàgines web de les pròpies plataformes i serveis per obtenir dades actualitzades. També s'ha utilitzat de manera general Stack Overflow, la principal comunitat de desenvolupadors per solucionar dubtes i compartir coneixements (Stack Overflow, 2023).

Durant la redacció d'aquest treball s'ha intentat plasmar al màxim tot el procés de documentació, desenvolupament i execució. Però cal tenir en compte que gran part d'aquest temps invertit no sempre es veu reflectit en resultats, especialment en el tractament del codi amb Python, ja que apareixen moltes dificultats i limitacions tècniques. En el treball a continuació es fa una exploració d'un entorn que era gairebé

desconegut per a mi. És per aquest motiu que el codi i els mètodes utilitzats podrien ser millorats per professionals.

Finalment, també cal destacar que no s'ha fet cap despesa en plataformes professionals, ja que no era viable. I també mostro que l'ordinador des d'on s'ha processat i executat el codi no és dels més punters del mercat, i per això certs processos han trigat força temps a executar-se completament, com s'indicarà en l'anàlisi del codi.

## 2. Marc teòric

En el marc teòric es crea una estructura conceptual que es fa servir per a comprendre i explicar la situació del tema d'estudi. S'exposa informació sobre l'artista Bad Gyal, el context històric i la situació en que es troba el seu entorn musical. Aquestes dades han de servir per a guiar la recopilació de dades i per a interpretar els resultats de la investigació. Així, permet establir una base sólida per a la investigació i ajuda a comprendre el que s'exposa, evitant la recerca innecessària en fases posteriors.

### 2.1. Consum musical i Spotify

En l'actualitat, la indústria musical factura més de 851 milions d'euros a Espanya, i es preveu que aquesta xifra segueixi augmentant els pròxims anys. (Statista, 2023b) Però també hi ha altres indicadors de la importància d'aquest sector, i és que gairebé la meitat dels espanyols han escoltat música recentment, i especialment destaquen els joves, que ho fan en *streaming* (fins a un 75%) i diàriament (un 40%) (IP Mark, 2021). D'aquesta manera queda clar que el sector musical és molt important, especialment entre els joves, i ho seguirà sent en els anys vinents.

De tots els serveis dedicats a l'*streaming* de música, destaca Spotify, una empresa d'origen Suec. "Spotify és un servei de música, pòdcasts i vídeos digitals que et dona accés a milions de cançons i altres continguts de creadors de tot el món." (Spotify, 2023a) Aquest servei d'àudio, té un model de monetització *freemium*, el que significa que les opcions principals són d'ús gratuït amb publicitat i les opcions més avançades i sense publicitat requereixen una subscripció Spotify Premium. Segons el dispositiu, aquestes limitacions varien, però en general tot el contingut és en streaming i queda desat en el compte personal, de manera que es pot accedir fàcilment des d'ordinadors, *smartphones*, *tablets*, *smartTV* o cotxes entre d'altres amb Spotify Connect. És un model semblant al lloguer perquè "La nostra llicència implica que no hi ha cap manera d'exportar el nostre contingut fora de l'aplicació". (Spotify, 2023a)

Spotify ha destacat per la seva facilitat d'ús, la personalització de la música a partir d'algoritmes que creen llistes personalitzades, suggeriments de cançons, artistes... A més d'un extens catàleg de cançons, i pòdcasts mundials.

Actualment és el lloc de *streaming* mundial amb més subscriptors (30,5%), doblant a Apple Music (13,7%) i lluny d'altres serveis. (Statista, 2023c) Spotify manté la seva tendència d'augment i el lideratge mundial des de fa anys. (Statista, 2023a) El consum musical a Espanya segueix la mateixa tendència i el mercat digital (75%) està molt per davant del consum en físic (12%) (PROMUSICAE, 2023).

El lideratge d'aquest servei ha fet que se signin acords amb discogràfiques i altres empreses musicals que afecten a les recomanacions dels usuaris i els posicionaments en *streaming*. Aquests acords són opacs i l'accés a l'algoritme i dades dels artistes està molt limitat (Mata Valldaura, 2022). Més endavant, durant l'apartat d'anàlisi, es treballarà amb l'algoritme des de l'API d'Spotify.

## 2.2. Les plataformes i els serveis d'àudio sota demanda

Primer de tot, cal diferenciar certs conceptes, ja que l'ús del terme plataformes pot ser molt ambigu. Així, es prendrà com a referència la definició del terme plataforma sobre la qual es reflexiona al llibre Custodians of the Internet: "Les plataformes allotgen, organitzen i permeten la circulació de continguts compartits o creats per usuaris, i defineixen les interaccions que aquests poden fer. No apliquen comissions a aquests continguts. Finalment processen tota aquesta informació per obtenir dades útils per a l'obtenció de beneficis". (Gillespie 2018 p.18)

D'aquesta manera, caldrà diferenciar els dos tipus d'espais que s'analitzaran en aquest treball. Per una banda hi ha les plataformes Twitter i YouTube que allotgen, organitzen i fan circular el contingut dels usuaris sense comissions. I d'altra banda hi ha Spotify, que no pot ser considerada plataforma perquè només permet la circulació de contingut professional i aplica comissions. Per això cal tractar a Spotify com a un servei de contingut d'àudio professional sota demanda.

L'ús de les xarxes socials està molt estès i cada vegada se li dona més importància. Des del 2017, hi ha hagut un estancament del percentatge de població espanyola a xarxes socials del 85% aproximadament (Statista, 2022). Com que és una oportunitat d'arribar a moltes persones que abans no es tenia tant en compte, les empreses cada vegada es professionalitzen més i la majoria decideixen mantenir o invertir-hi més diners per tal de seguir creixent (IAB Spain, 2022).

En general, les plataformes permeten crear perfils personals, seguir-ne d'altres, veure el contingut generat per altres usuaris i compartir-lo públicament o per missatgeria privada, com és el cas de YouTube i Twitter. Cal destacar que aquests usos són variables en el temps i en cada plataforma. Però també cal tenir en compte que cada una funciona diferent i crea una cultura i un sociolecte concret.

Les plataformes són una font d'entreteniment i recerca del dia a dia de la majoria de persones a Espanya. Però no són importants només per la quantitat de persones que hi participen, sinó també per l'ús que se'n fa. Un 68% de les persones interactuen amb altres persones i més de la meitat utilitzen les xarxes socials com a font d'informació. Però també cal destacar que inclouen activitat publicitària i empresarial de gran part de les empreses (IAB Spain, 2022). Si la majoria de la població es comunica i s'informa a través de les plataformes vol dir que aquestes tenen un gran poder per modelar la realitat social en la qual vivim.

Tenint en compte aquest fet, cal destacar que la majoria de xarxes socials i plataformes s'han creat i es controlen des dels Estats Units d'Amèrica, excepte TikTok que és xinesa i Spotify que és sueca. Això és un aspecte important a tenir en compte respecte al control de dades, ja que la legislació europea és diferent, en general, més protecciónistà de l'usuari. Com que aquest estudi es fa des de la Unió Europea, cal tenir en compte la legislació de protecció de dades. I també cal considerar que les mateixes empreses també restringeixen molt l'accés a aquesta informació.

Un altre aspecte molt important de les empreses darrere de les plataformes és que no són organitzacions sense ànim de lucre, volen obtenir diners, i per tant intenten que l'usuari estigui el màxim de temps a la plataforma, amb l'objectiu d'impactar-lo amb publicitat. Això és degut al fet que el model de negoci de la majoria de xarxes socials

i plataformes consisteix en la inserció publicitària segmentada i personalitzada. Això ha provocat problemes d'addicció, així com una restricció important dels continguts que poden publicar els usuaris per tal que siguin més afins a la publicitat que es mostra.

D'altra banda, la capacitat de les plataformes per inserir publicitat segmentada i personalitzada, també ha provocat la necessitat de fer un seguiment als perfils de marca. I les plataformes també han aprofitat per posar un mur de pagament per accedir-hi. D'aquesta manera, les plataformes s'han anat professionalitzant, i la majoria de marques hi apareixen, sigui amb perfils oficials o a través d'altres marques com els *influencers*.

### 2.3. Metodologia d'anàlisi de plataformes

Aquest tipus d'anàlisi sovint s'anomena *social listening* o *social monitoring*. És una pràctica que es basa en la revisió i anàlisi de les converses i interaccions que es produeixen en plataformes i xarxes socials. Aquesta tècnica és cada vegada més important per a les empreses i organitzacions d'investigació que volen comprendre millor les necessitats i opinions del seu públic, així com analitzar fenòmens socials i d'opinió pública (Brent Barnhart, 2023).

El *social listening* es basa en la idea que les xarxes socials són una font d'informació molt valuosa que sovint no es té en compte i acostuma a ser molt més orgànica i completa que d'altres (estimacions de vendes, enquestes online o presencials). Pot ser utilitzada per a la investigació de mercat (segmentar el públic, fer col·laboracions més adequades), així com gestionar la reputació de l'artista i tenir dades reals com a marca.

Per això, implica la utilització de diferents eines i tècniques per a la recopilació i anàlisi de dades de les plataformes, com ara la cerca i seguiment de paraules clau, la segmentació del públic, l'anàlisi de comunitats o l'anàlisi del sentiment, entre d'altres. Actualment, el *social listening* és una eina valuosa per a la comprensió del públic que permet prendre decisions més informades i efectives.

## 2.4. Bad Gyal

Bad Gyal es va donar a conèixer a YouTube l'any 2016 amb el llançament del seu primer tema "Pai" (4.146.710 visualitzacions), que tenia la base de la cançó "Work" de Rihanna i es va popularitzar entre els joves de Catalunya. Posteriorment, es va donar a conèixer en l'entorn espanyol amb la seva cançó més reconeguda: "Fiebre" (67.154.808 visualitzacions). (Bad Gyal, 2023b)

L'artista, ja era coneguda en el món *underground* de la música *urban*, però es va donar a conèixer especialment després de signar amb la discogràfica Interscope/Aftercluv el 2019 i posteriorment amb Universal Music Latino (Billboard, s.d.; Universal Musica, 2023). Això li ha permès posicionar-se en el capdavant de les llistes a serveis sota demanda com Spotify o Apple Music, però també a YouTube i les principals cadenes de ràdio espanyoles. El lligam amb la discogràfica també ha desencadenat en col·laboracions amb grans artistes com: Omar Montes (61M de visualitzacions), Rauw Alejandro (18M), Soto Asa (2,4M), Quevedo (4,8M) o Karol G (13M) (YouTube, 2023).

Un exemple d'èxit mundial és l'EP "Warm Up", que va debutar en el top 10 de la *playlist* "Global Album Chart" d'Spotify i que actualment acumula més de 150 milions de reproduccions en plataformes. (Bad Gyal, 2023a)

Actualment Bad Gyal té una posició privilegiada, especialment en l'entorn espanyol. La seva influència també està arribant a nous col·lectius i està en grans esdeveniments com concerts i festivals a l'alçada dels grans artistes urbans en l'àmbit mundial. També té futurs projectes com el seu primer disc oficial "La Joia", que llançarà al 2023, i col·laboracions amb artistes mundials com BZRP (LOS 40, 2023).

Bad Gyal és una artista força activa a les plataformes, a més de ser bastant natural en el seu contingut. És força irregular amb les publicacions i els seus perfils inclouen tant publicacions oficials o corporatives com pot ser una portada de cançó; com contingut totalment banal del seu dia a dia com pot ser ella fent un directe a Instagram des d'un bar fumant. Ambdós continguts generen *engagement* i amplien la comunitat de l'artista. En la propera taula es detallen algunes mètriques de l'artista a les seves plataformes oficials.

|                        | <b>IG</b> | <b>FB</b> | <b>TW</b> | <b>TK</b> | <b>YT</b>     | <b>SP</b>      |
|------------------------|-----------|-----------|-----------|-----------|---------------|----------------|
| <b>Followers</b>       | 1,3M      | 38,8k     | 62,2k     | 840,4k    | 826k          | 927'9k         |
| <b>Likes</b>           |           | 31,0k     |           | 9,4M      | S/D           | S/D            |
| <b>Publicacions</b>    | 355       | S/D       | S/D       | S/D       | 78            | 47             |
| <b>Cançó destacada</b> | N/A       | N/A       | N/A       | N/A       | 67M<br>Fiebre | 126M<br>Alocao |
| <b>Creació</b>         | 2016      | 2016      | 2016      | 2020      | 2016          | 2016           |

Taula 1: Comparativa de mètriques de Bad Gyal a les xarxes socials. Font: Elaboració pròpia.

## 2.5. Plataformes i indústria musical

En el sector de la música, actualment, és molt important tenir en compte les plataformes, ja que formen l'opinió dels usuaris i també decideixen què sona al país. Es pot veure clarament que les cançons del *mainstream* espanyol moltes vegades són llançades o filtrades expressament a les xarxes socials perquè es viralitzin i siguin tendència. D'aquesta manera, s'impulsen les cançons abans del llançament i posteriorment s'estenen més fàcilment. També cal tenir en compte que els artistes han de tenir un comportament concret com a personatges públics que són, ja que la comunitat opina a les plataformes, enaltint-los o cancel·lant-los (Gen Playz - RTVE, 2023) (UAB Barcelona, 2023).

El 80% dels usuaris declara que TikTok és la seva principal font per descobrir música, per davant d'altres plataformes digitals, serveis de *streaming* i recomanacions d'amics. A més, un cop descoberta la música nova a l'aplicació, també s'estén l'experiència fora d'aquesta. Els usuarisafegeixen la cançó a preferits (47%), van al perfil de l'artista (46%) i també segueixen a l'artista (43%). D'altra banda, l'estudi també destaca que més d'un 80% dels usuaris se senten més propensos a compartir cançons si els produeixen nostàlgia (Sala de prensa de TikTok, 2021) .

Actualment la indústria musical pensa molt en les plataformes en tot el procés, des de la creació fins a la venda. Passa el mateix que ja va passar anteriorment, les cançons amb més èxit acostumen a estar pensades per al canal de difusió. Als anys 80 les cançons havien d'estar a la ràdio, als 90 i als 2000 havien de fer-se videoclips

espectaculars i aparicions en televisió. A la dècada dels 2010 es va començar a produir el canvi cap a les plataformes d'*streaming* i les xarxes socials; que va ser impulsat exponencialment amb la COVID-19 i l'explosió de TikTok (La Vanguardia, 2022).

El descobriment i expansió dels artistes per mitjans digitals ja fa temps que existeix. Artistes amb llargues carreres s'han fet famosos gràcies a aquestes eines. És el cas de Justin Bieber, que va ser descobert a YouTube per una discogràfica amb la qual va fer grans col·laboracions a principis de la dècada del 2010. Més recentment, un dels primers artistes que van sorgir amb els vídeos curts (TikTok, Instagram), és l'artista Lil Nas X. Es va fer tendència amb un vídeo en que sonava la seva cançó i ell es convertia en *cowboy* a la tornada com a *meme*. Poc després va treure "Old Town Road (Remix)" amb el famós cantant de country Billy Ray Cyrus. En els últims anys, s'ha consolidat, superant el top 15 de les llistes mundials i guanyant nombrosos premis. (*Lil Nas X Talks «Old Town Road» and the Billboard Controversy | Time*, s.d.)

Una cançó que triomfa a les plataformes, també ho fa fora d'aquest entorn, tot i que també n'hi ha algunes que tenen públics diferents i triomfen sense xarxes socials, però cada vegada en són menys. El contingut de les plataformes funciona molt bé per crear tendències (*trends*) però també per mostrar altres facetes més humanes dels artistes que fan empatitzar molt més als usuaris amb l'artista i les seves lletres (Gen Playz - RTVE, 2023). Bad Gyal sapaprofitar molt bé aquestes oportunitats a les xarxes socials. Contínuament crea vídeos curts de diferents estils, però també fa força directes en els quals acostuma a anunciar llançaments o novetats en la seva carrera, tot i que sovint només els fa per interactuar amb els seus fans.

La importància dels *trends* i els balls de les cançons a xarxes socials és molt important. Una bona promoció abans o durant el llançament pot fer que cançons aparentment simples siguin grans èxits, és el cas de la cançó Chiken Teriyaki de Rosalia, una de les més escoltades del disc Motomami amb més de 60 milions de reproduccions únicament a YouTube (Rosalía, 2022).

Però no és un fenomen exclusiu de la generació Z o dels joves, perquè el 67% dels usuaris de TikTok són majors de 25 anys, fet que ha provocat que artistes

internacionals més grans s'hiafegeixin: ABBA, Alejandro Sanz, Laura Pausini, Camila Cabello, Jason Derulo, Aitana, U2, Led Zeppelin o Raphael (Kantar, 2023).

Aquestes noves maneres de promocionar i crear altres productes dels artistes funciona, i això és indiscretible, perquè podem veure noves veus que han sorgit totalment de les plataformes. Això ha fet que productores i discogràfiques ja estiguin utilitzant-les. Però la manera com ho fa cadascuna és diferent, i ja han aparegut casos en els quals artistes consolidats han mostrat el seu desacord amb algunes d'aquestes tècniques. Florence and the Machine o Doja Cat ja van denunciar aquesta situació l'any 2022, però el cas més sonat va ser el de l'artista Halsey. La cantant va declarar que tot i tenir la cançó i el videoclip preparats per al llançament, sense generar un moment de viralitat al seu TikTok, la seva discogràfica (propietària dels drets) no li donava una data de llançament. Finalment, va explicar aquesta situació als seus perfils socials, també a TikTok, i llavors es va fer viral i li van donar una data de llançament (20 minutos, 2022).

Altres discogràfiques, en canvi, no són tan impositives i creuen que el contingut a les plataformes ha de sortir el més natural possible perquè funcioni, no cal que el contingut sigui el més viral de la plataforma, sinó que cal crear contingut complementari que dirigeixi al consumidor cap als serveis de *streaming*: Spotify, Apple Music... Perquè la música a les plataformes també es monetitza actualment, però no és ni de bon tros la quantitat d'ingressos que es genera als serveis de *streaming* (Gen Playz - RTVE, 2023).

# 3. Objectius i preguntes de recerca

L'objectiu principal és analitzar la conversa a les plataformes i serveis sota demanda al voltant de l'artista Bad Gyal. Per aconseguir-ho, aquesta anàlisi es divideix en objectius secundaris més petits per tal que siguin assumibles. Perquè sigui més fàcil aplicar-los, s'han establert unes preguntes de recerca per a cada objectiu secundari. S'ha fet tant als objectius quantitatius com els que són qualitatius per tal d'obtenir una visió més realista.

## 3.1. Objectius quantitatius

Recollint els objectius quantitatius secundaris mencionats anteriorment, obtenim les preguntes de recerca que s'exposen a continuació.

**Objectiu 1: observar l'evolució temporal de l'artista a les plataformes i serveis d'àudio sota demanda.**

El primer que cal fer per entendre la conversa al voltant de Bad Gyal és establir l'espai temporal en el qual s'emmarca. Ja s'han introduït algunes dades al marc teòric, però cal que a part de la carrera de l'artista, també s'analitzin diferents aspectes propis de cada plataforma per veure si hi ha correspondència temporal entre els fets. Així, s'exposen les següents preguntes de recerca.

*En quin període apareix l'artista?*

Aquesta pregunta pretén saber en quin període es parla sobre l'artista Bad Gyal, és a dir, quan apareix el primer contingut i l'últim.

*Hi ha una tendència?*

Cal investigar sobre si el nombre d'aparicions de l'artista ha anat en augment o si, per contra, ha disminuït. Aquesta pregunta fa referència a l'evolució general de tot el període.

*Hi ha pics o valls? A què corresponen?*

Concretament, es vol investigar si hi ha hagut períodes que destaquen per tenir moltes aparicions (pics) o molt poques (valls). També es vol investigar a quin fet corresponen aquests comportaments. Per exemple, podrien correspondre al llançament d'una cançó, un àlbum o una informació destacada.

*Hi ha algun comportament cílic?*

Moltes vegades acostumen a aparèixer comportaments cíclics en les plataformes directament vinculats a la vida humana, com l'augment d'aparicions durant el dia i la disminució durant la matinada. Aquests comportaments també ens permeten estudiar l'impacte segons els països, ja que poden estar en fusos horaris molt diferenciats. Però també poden aparèixer altres tipus de comportaments cíclics com la publicació d'àlbums cada cert temps o la publicació de cançons per aprofitar períodes concrets, com publicacions abans de l'estiu o de l'època de festivals.

*Quant varia la conversa amb els llançaments i amb les col·laboracions?*

Aquesta pregunta fa referència a l'estudi de la resposta dels usuaris en el moment de les publicacions musicals, siguin pròpies o col·laboracions amb altres cantants.

**Objectiu 2: analitzar quin tipus de contingut comparteix l'artista a les plataformes YouTube i Twitter i al servei d'àudio sota demanda Spotify.**

Per comprendre el comportament de l'artista, cal analitzar quin tipus de publicacions fa, ja que hi ha gran varietat de formats com contingut visual, d'àudio, de vídeo o de text. Cada xarxa social i plataforma té el seu tipus de contingut i les seves limitacions. Per això es plantegen les següents preguntes de recerca.

*Quin tipus de contingut hi ha?*

Aquesta pregunta fa referència al tipus de contingut que apareix en cada plataforma. Cal analitzar quin tipus de contingut s'està portant a cada plataforma, perquè malgrat que hi hagi plataformes que permeten la difusió de diferents tipus, pot ser que no s'utilitzin tots o que exclusivament s'utilitzi un.

*Quin és el contingut més nombrós?*

Més concretament, cal establir si tots els tipus de continguts es publiquen amb la mateixa freqüència o si n'hi ha que destaquen en el nombre d'aparicions.

### Objectiu 3: detectar el contingut amb més repercussió a YouTube i Twitter.

Després d'analitzar cadascun dels tipus de contingut que hi ha al voltant de l'artista, cal veure quan i com reaccionen els usuaris. Per això es plantegen les següents preguntes de recerca.

*Quin és el contingut amb més interaccions?*

Per detectar el contingut amb més repercussió cal veure com interactuen directament els usuaris. Per això cal analitzar cada publicació i veure quines tenen un nombre d'interaccions especialment alt, baix o si segueixen dins de la mitjana.

*Quins tipus d'interaccions hi ha?*

Cal analitzar quins tipus de possibles interaccions hi ha en cada plataforma i quines apareixen realment. Alguns exemples poden ser: *likes*, comentaris, *retweets*, *reply* o *quote*.

*Quantes interaccions de cada tipus hi ha?*

També cal mesurar quines interaccions destaquen en nombre, sigui per superar el nombre habitual o per ser inferiors a la quantitat esperable.

*Quina repercussió final hi ha?*

Per entendre finalment la repercussió caldrà calcular els valors d'*engagement* estandarditzats amb una fórmula. A partir d'aquest nombre es podrà comparar d'una manera més clara la repercussió de les publicacions.

#### Objectiu 4: Quantificar la resposta del públic segons l'usuari que publica el contingut.

En el contingut analitzat a Twitter es troben totes aquelles piulades en les quals aparegui el text “Bad Gyal”. Però és molt important valorar i diferenciar aquell contingut que ve directament de l'artista, és a dir, del perfil oficial, d'aquell que no, com poden ser els *tweets* dels fans.

##### *Quins elements estan dins del contingut oficial?*

Cal veure quina quantitat de publicacions es troba dins del contingut publicat per perfils oficials. És de gran importància analitzar aquest tipus de contingut de manera diferenciada de la resta perquè és el contingut que està controlat per la marca que hi ha al darrere de l'artista. Una manera de visualitzar-ho és comparar el percentatge de publicacions oficials respecte del total de publicacions.

##### *Quins elements estan fora del contingut oficial?*

Igualment cal analitzar el contingut que no és oficial, perquè poden haver-hi opinions i elements molt diferenciats als de la imatge que es vol donar de manera oficial. Cal veure quines diferències hi ha entre el tipus de contingut anteriorment mencionat.

##### *A quin tipus de contingut interactuen més els usuaris?*

Quant als elements que cal comparar entre el contingut oficial i la resta de continguts, destaca la manera com interactuen els usuaris.

##### *Quina evolució temporal hi ha en cada tipus de contingut?*

Quan Bad Gyal fa una publicació oficial, cal analitzar la resposta dels seguidors, sigui en interaccions a aquesta publicació o en altres de creació dels mateixos usuaris. També cal obtenir una visió general de l'evolució per als dos tipus de contingut.

## Objectiu 5: realitzar una anàlisi comparativa entre YouTube, Twitter i Spotify.

Dins de les xarxes socials i plataformes analitzades, cal comparar els fets més destacats que s'han pogut observar en les respostes a les preguntes de recerca anteriors.

*Quina evolució temporal hi ha a cada plataforma?*

Cal comparar l'evolució temporal, ja que cada plataforma pot aparèixer, augmentar o disminuir les seves publicacions de diferents maneres. S'ha de comparar per veure si hi ha alguna que acostuma a ser capdavantera, endarrerida o si no tenen una vinculació apparent.

*Quina plataforma té una tendència més creixent?*

Per poder decidir el rumb futur de la carrera de l'artista, és important que es vegi quines xarxes poden donar millors dades. Per això cal veure si la tendència és creixent i a quin nivell.

*Quin tipus de continguts hi ha a cada plataforma?*

També cal comparar el tipus de contingut de cada plataforma per tal d'aconseguir una visió sobre quins continguts funcionen millor.

*Quin és el contingut amb més i menys repercussió a cada plataforma?*

Per tenir una visió del màxim de repercussió al que arriba cada plataforma, cal comparar la publicació amb més *engagement*.

*Quant d'engagement hi ha a cada plataforma?*

Per comparar el nivell d'interacció cal comparar el nombre màxim i la mitjana d'interacció per post de cada plataforma.

## 3.2. Objectius qualitatius

Recollint els objectius qualitatius secundaris mencionats anteriorment, obtenim les preguntes de recerca que s'exposen a continuació.

**Objectiu 6: saber si la conversa és positiva, negativa o neutra per a Bad Gyal a Twitter.**

És molt important que els usuaris interactuïn en el contingut relacionat amb l'artista, però perquè aquest sigui beneficis per a la seva carrera cal que majoritàriament sigui positiu. Cal esbrinar com és aquest tipus de contingut. Per aconseguir-ho es plantegen les següents preguntes de recerca.

*Hi ha més contingut positiu, negatiu o neutral?*

Per poder saber quant de contingut hi ha de cada tipus, cal establir primer, per a cada contingut, si aquest és positiu, negatiu o neutre.

*En quines plataformes opinen més els usuaris?*

També és important comparar les opinions per plataforma, ja que n'hi ha que fomenten aquest tipus de contingut i n'hi ha d'altres que són molt més restrictives.

*Com opinen a cada xarxa social?*

Però a part de la quantitat d'opinions també és molt important la posició d'aquestes, perquè hi ha plataformes on es concentren els usuaris que acostumen a criticar i pot ser que d'altres usuaris com els seus fans es trobin al voltant d'altres xarxes socials.

**Objectiu 7: analitzar les relacions del públic de Bad Gyal amb artistes musicals i altres temes.**

Per poder entendre com actua l'artista amb el seu entorn, així com les seves col·laboracions, cal visualitzar els cercles dels quals s'envolta. Per aquest motiu s'exposen les següents preguntes de recerca.

*Quants vincles i amb quins artistes es generen?*

Per analitzar les relacions entre els artistes cal veure quins estan vinculats entre si i quins no.

*Quina direccionalitat tenen les connexions de Bad Gyal amb altres artistes?*

Principalment cal veure i analitzar les connexions que es produeixen des de l'artista Bad Gyal i cap a Bad Gyal.

*Quina direccionalitat tenen les connexions dels fans de Bad Gyal amb altres usuaris?*

Però també pot ser molt interessant de veure totes aquelles relacions que no es produeixen directament des del compte oficial de Bad Gyal. És a dir des d'aquells que opinen sobre ella.

*Concorden les dades de xarxes socials i col·laboracions que ha fet l'artista?*

Pot ser que es produixin relacions entre l'entorn de Bad Gyal i els cercles d'artistes amb els quals ha fet col·laboracions.

*Com apareixen els artistes amb qui ha fet col·laboracions? Estan molt distants, en grups?*

La proximitat i interrelació dels cercles dels artistes amb els quals ha fet col·laboracions pot ser molt interessant per visualitzar en quins casos ha arriscat més i en quins ha buscado una col·laboració més propera a ella. De la mateixa manera, podria aprofitar aquesta anàlisi per decidir col·laboracions futures.

*Quins gèneres i temàtiques tenen en comú?*

Per trobar les relacions en comú entre cançons, és interessant veure per quins gèneres es vincula més a l'artista Bad Gyal. D'aquesta manera es pot trobar les vinculacions principals per gènere musical.

# 4. Metodologia

La metodologia que se seguirà durant l'anàlisi es dividirà en dues parts: obtenció de les dades i visualització de les dades.

## 4.1. Obtenció de les dades

L'obtenció de les dades es durà a terme a través de les APIs de les plataformes que utilitza l'artista Twitter i YouTube i també el servei de Spotify. Per interactuar amb aquestes dades i automatitzar els processos s'utilitzarà Python. Aquest llenguatge de programació permet obtenir la informació des de les API de les xarxes socials a través de llibreries que optimitzen els processos i els automatitzen. De la mateixa manera, també ens permet convertir la informació a fitxers estandarditzats que puguin ser utilitzats per les aplicacions de visualització de dades.

Per a l'extracció de les dades s'utilitzarà l'accés a les API de cada plataforma a través del codi de programació en Python que s'executarà amb l'aplicació Pycharm. Tot el codi utilitzat ha estat creat i adaptat per a cada tasca del treball de fi de grau. S'ha realitzat una investigació prèvia per saber com crear el codi, i una vegada s'ha creat, s'han fet nombroses proves i correccions abans d'aconseguir el resultat final.

## 4.2. Visualització de les dades

Les dades obtingudes a través d'API i codi, seran massa grans perquè es puguin entendre i extreure'n informació útil. Fins i tot, es podrien considerar *big data*, ja que són molt grans i complexes i requereixen un nivell d'automatització elevat en tots els processos. Per això, serà necessari que siguin representades de manera visual amb programes especials.

És important que es facin unes bones visualitzacions amb les mètriques triades, ja que això ens permetrà veure els valors normals i aquella informació anòmala que destaca i que cal investigar amb més profunditat. També cal tenir en compte que s'haurà de fer una comparació entre les dades de les diferents fonts, perquè no només cal veure

les dades individualitzades per a cada plataforma, sinó que també cal agrupar-les per tenir una visió més holística com indica l'objectiu general del treball.

Per aconseguir fer aquestes visualitzacions s'utilitzaran els programes Tableau i Gephi, que permeten tractar grans volums de dades i personalitzar les visualitzacions de dades (Tableau, 2023) (Gephi, 2023).

### 4.3. Mètriques i KPI

Per afinar més concretament quines dades volem trobar, cal establir les mètriques que utilitzarem (també anomenades KPI). Aquests indicadors ens permetran concretar com i què volem obtenir per complir els objectius i contestar les preguntes de recerca. A continuació, s'exposaran les mètriques relacionades amb cadascun dels objectius i de les preguntes de recerca.

#### Objectius quantitatius

*Objectiu 1: observar l'evolució temporal de l'artista a les plataformes i serveis d'àudio sota demanda.*

- Nombre d'aparicions en publicacions per data i hora.
- Llista cronològica de llançaments musicals, tant originals com col·laboracions.

*Objectiu 2: analitzar quin tipus de contingut comparteix l'artista a les plataformes YouTube i Twitter i al servei d'àudio sota demanda Spotify.*

- Nombre d'elements de cada categoria per plataforma: Foto, Àudio, Vídeo o Text.
- Tipus de contingut a Twitter: *Tweet, Retweet, Quote, Comment*).

*Objectiu 3: detectar el contingut amb més repercussió a YouTube i Twitter.*

- Llista de recompte de cada tipus d'interacció: *Like, Comment, Retweet, Reply o Quote*.
- Valors mitjans, màxims i mínims de cada publicació i per a cada plataforma.
- *Engagement i engagement rate* de cada publicació.
- *Engagement i engagement rate* per tipus de contingut.

- Evolució temporal de l'*engagement* i *engagement rate*.

*Objectiu 4: quantificar la resposta del públic segons l'usuari que publica el contingut.*

- *Engagement* i nombre de publicacions de cada usuari.
- *Engagement* i nombre de publicacions del contingut oficial i no oficial.
- Evolució temporal de les publicacions dels usuaris.
- Tipus de contingut per usuaris.

*Objectiu 5: realitzar una anàlisi comparativa entre YouTube, Twitter i Spotify.*

- Nombre de publicacions per plataforma i data.
- Recompte de tipus de contingut a cada plataforma: foto, àudio, vídeo o text.
- Recompte d'*engagement* de cada plataforma.
- Recompte de tipus d'interacció per plataforma.

## Objectius qualitatius

*Objectiu 6: saber si la conversa és positiva, negativa o neutra per a Bad Gyal a Twitter.*

- Tipus de contingut: positiu, negatiu o neutre.
- Discurs d'odi a la plataforma.
- Quantitat i tipus d'interacció: *retweet*, *like*, *reply* o *quote*.

*Objectiu 7: analitzar les relacions del públic de Bad Gyal amb artistes musicals i altres temes.*

- Llistes de recomanacions d'artistes de Bad Gyal.
- Llistes de recomanacions d'artistes de segon nivell.
- Llista d'artistes amb els quals Bad Gyal ha col-laborat.
- Direccionalitat de les mencions “Bad Gyal” a Twitter.
- Llista d'entitats detectades en el text.
- Gèneres relacionats de les cançons de Bad Gyal.

# 5. Anàlisi

En aquest apartat s'estudiarà l'extracció i els resultats de les dades obtingudes a través de la metodologia esmentada en el punt anterior. L'objectiu és exposar i resumir la informació que s'ha obtingut a partir del gran volum de dades extret. S'analitzarà el procés realitzat per cada plataforma, ordenant la informació extreta de més general a més concreta a Spotify, YouTube i Twitter.

## 5.1. Spotify

Spotify és el principal servei de *streaming* musical sota demanda. És necessari investigar-la per tal d'obtenir els continguts més rellevants de l'artista dins d'aquesta plataforma, que també és el seu canal principal de llançament de les seves cançons. Spotify se centra en el contingut d'àudio i no té activitat social, fora de la que va dirigida a escoltar música. Es poden seguir perfils d'artistes i d'usuaris i veure el contingut que segueixen, les seves llistes i què estan escoltant en directe. No existeixen sistemes de xat ni la capacitat de publicar contingut que no siguin cançons i pòdcasts (per a professionals) i llistes (per a usuaris consumidors i també per a professionals).

L'API de Spotify és força amigable, ja que hi ha una pàgina de desenvolupadors amb explicacions i demostracions interactives que faciliten la feina. (Spotify, 2023b) A més, per a l'ús d'extracció de dades per estadístiques, com es fa en aquest treball, les dades que es poden obtenir són força limitades comparades amb altres API. Cal afegir que hi ha molta documentació externa sobre com utilitzar-la i que hi ha codi preparat per a simplificar certes tasques.

Per aconseguir extreure dades de l'API, cal fer un programa que ens permeti automatitzar els processos per extreure correctament tot el volum de dades que es necessiten. Per facilitar aquesta tasca, s'ha utilitzat la llibreria Spotipy, que permet simplificar processos amb un codi més simple. Així, s'evita implementar complexes sol·licituds HTTP per fer l'autenticació de l'usuari i per a les peticions de dades

(Spotipy, 2023). D'altra banda, per exportar el contingut en un format de taula, també s'ha utilitzat la llibreria Pandas (Pandas, 2023).

A partir de les eines anteriors, s'espera extreure de Spotify les següents dades. En primer lloc, l'evolució temporal, amb la llista de cançons publicades i les seves característiques. En segon lloc, els artistes recomanats a partir del seu perfil. I finalment, els gèneres relacionats amb Bad Gyal.

### 5.1.1. Llistat de cançons a Spotify

Per entendre l'evolució temporal de Bad Gyal a plataformes i serveis de *streaming*, primer de tot cal exposar els llançaments que ha fet en la seva carrera musical. Per aconseguir-ho, s'han extret totes les cançons i la seva informació des de l'API de Spotify. Amb aquestes eines, s'ha elaborat el programa que s'exposa a “Codi 1” per extreure un fitxer CSV que permeti exportar la informació per fer visualitzacions de dades fàcilment.

El codi desenvolupat es pot dividir en tres parts. La primera part prepara el codi per a l'obtenció de les dades. Exporta les llibreries que seran necessàries al llarg del codi i després utilitza Spotipy per accedir a l'API amb les claus que s'han obtingut prèviament. Les claus s'obtenen des de la web de desenvolupadors de Spotify i requereixen un inici de sessió al servei i posteriorment la creació d'una aplicació que generarà les claus “id” i “secret”. Una vegada s'han inserit aquestes claus, ja es pot utilitzar l'API i llavors es preparen els valors que es necessitaran posteriorment com l'*id* de l'artista Bad Gyal o el tipus de contingut que es vol obtenir d'aquest.

En segon lloc, es produeix l'obtenció de les dades a partir de les iteracions amb els bucles *for*. Gràcies al codi de la llibreria Spotipy s'obtenen les cançons amb *artist\_albums* i *album\_tracks*. En ordre d'identació, es recorren els tipus d'àlbums que volem explorar, després l'àlbum i finalment cada cançó. Els àlbums que s'han decidit tractar són tots els tipus que ofereix Spotify (*album*, *single*, *appears\_on* i *compilation*) per tal d'obtenir la visió més completa possible. Això ha provocat que en explorar un àlbum en el qual ha col·laborat Bad Gyal, però que no és seu, carregui totes les

cançons, incloses en les que no participa. Per solucionar aquest problema s'ha fet una iteració més per a cada cançó per tal que només s'afegeixin aquelles en les quals apareix Bad Gyal com a artista.

Durant aquesta segona part, també cal destacar que s'ha utilitzat el recurs dels *Dataframe* de Pandas. Amb les dades que s'han volgut obtenir per a cada cançó de l'artista, s'ha creat un *Dataframe* per a cadascuna. I posteriorment s'afegeix a una llista per no perdre'ls amb la finalització de cada iteració. Cal destacar que per a poder crear el *Dataframe*, s'han hagut d'ajuntar llistes com a un sol *string* amb la funció *join*.

Finalment, la tercera part d'aquest codi realitza la preparació de la llista de *Dataframe* per a una exportació en format de taula. Així, primer es concatenen totes les llistes dels *Dataframe* en un de sol amb la funció *concat*. Posteriorment, s'exporta el *Dataframe* final a un document estandarditzat CSV amb la funció *.to\_csv*.

Cal destacar que el codi té comentaris explicatius per facilitar la seva comprensió. I també, que hi ha algunes accions comentades per tal de simplificar el codi, però que s'han mantingut per poder veure certes accions de control que s'han dut a terme mentre es desenvolupava el codi. Un exemple és el comptador *count* per poder saber que el programa realment està funcionant. O també és un exemple els diferents *print* que trobem al llarg del codi per mostrar l'estat del *Dataframe*.

```

# Importació de les llibreries
import spotipy
from spotipy.oauth2 import SpotifyClientCredentials
import pandas as pd

# Codi d'autenticació
client_id = 'clau_client'
client_secret = 'clau_secreta'
client_credentials_manager = SpotifyClientCredentials(client_id=client_id, client_secret=client_secret)
sp = spotipy.Spotify(client_credentials_manager=client_credentials_manager)

# Establim les variables necessàries
artist_id = '4F4pp8NUW08JuXwnoxglpN'
count = 0
llista_dfs = []

# Agafem els tipus de recopilatoris que volem analitzar
tipus_albums = ['album', 'single', 'appears_on', 'compilation']

# Iterem per obtenir una llista amb les cançons de tots els tipus de recopilatori
for element in tipus_albums: # Iterem per recopilatori
    llista_albums = sp.artist_albums(artist_id, album_type=element)

    for album in llista_albums['items']: # Iterem per album
        llista_pistes = sp.album_tracks(album['id'])

        for pista in llista_pistes['items']: # Iterem per cançó i afegim a la llista
            llista_artistes = []
            llista_markets = []

            for artista in pista['artists']:
                if artista['name'] == 'Bad Gyal':

                    # iterem pels elements que són llistes
                    for artist in pista['artists']:
                        llista_artistes.append(artist['name'])
                    for market in pista['available_markets']:
                        llista_markets.append(market)

                    print(llista_artistes)
                    df = pd.DataFrame({
                        "track": pista['name'],
                        'duration': pista['duration_ms'],
                        'album': album['name'],
                        'position': pista['track_number'],
                        'published': album['release_date'],
                        'artistes': ', '.join(llista_artistes),
                        'explicit': pista['explicit'],
                        'id': pista['id'],
                        'markets': ', '.join(llista_markets)
                    }, index=[0])
                    llista_dfs.append(df)

            #count += 1 # Utilitzem count com a codi de control
            #print(count)

        # Creem el df a partir de la llista
        df_final = pd.concat(llista_dfs)
        print(df_final)

    # Exportem el df a un arxiu .csv
    df_final.to_csv('bad_gyal_pistes_info.csv', index=False)

```

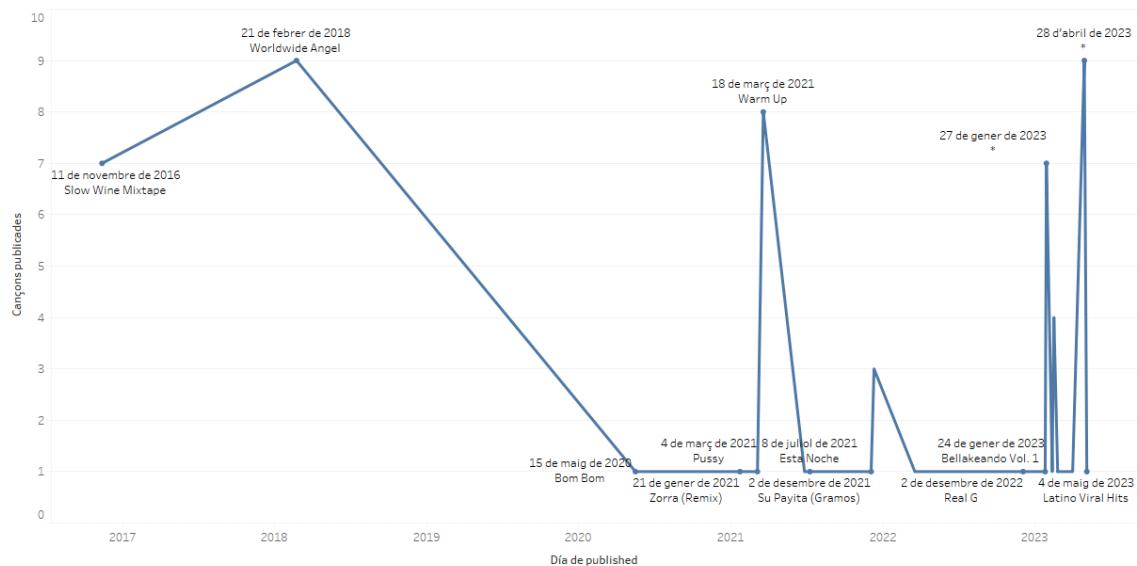
Codi 1: Extracció de cançons i informació a Spotify

### Evolució temporal

Una vegada obtingudes totes les dades de les cançons, es poden explotar per tal d'obtenir diverses informacions. Primer de tot, el que cal analitzar és l'evolució temporal de l'artista Bad Gyal. Per fer-ho, s'ha utilitzat la plataforma de visualització de dades Tableau.

El primer que es pot observar és quan comença a fer publicacions a la plataforma Spotify i quan ha sigut la seva última. En el gràfic següent podem veure el nombre de publicacions de cançons per dia durant tota la seva trajectòria fins a maig de 2023. El primer llançament que es troba és el primer dels seus tres àlbums, Slow Wine Mixtape, publicat l'onze de novembre de 2016.

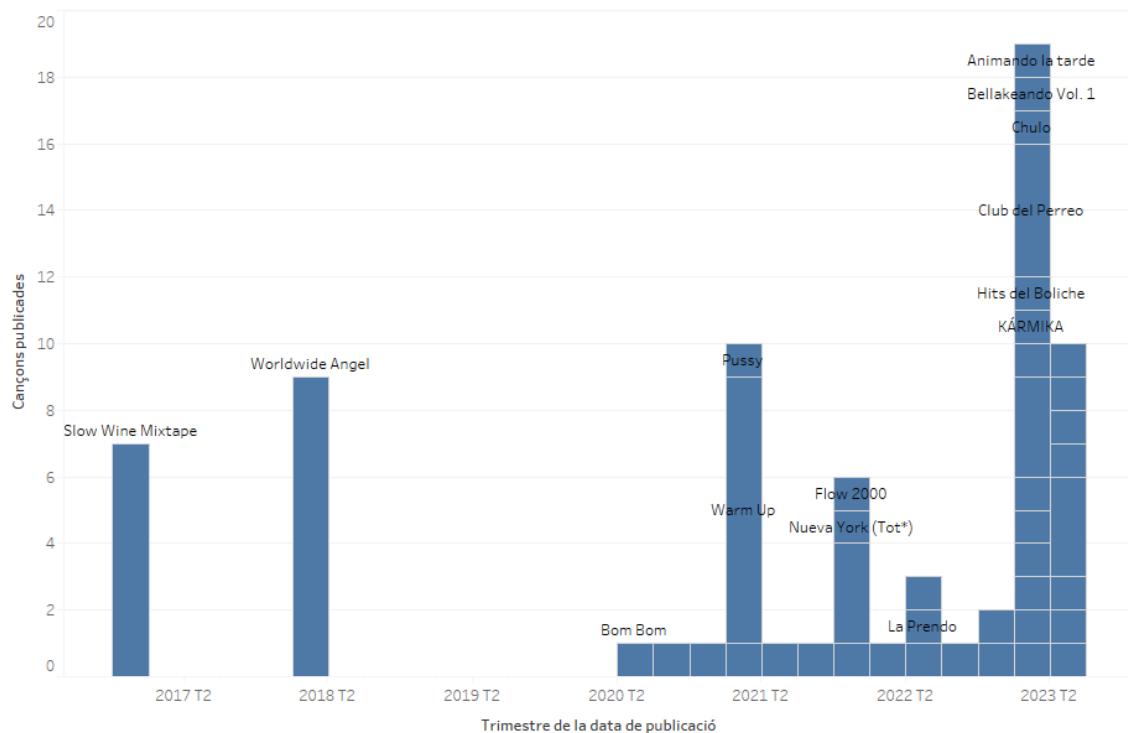
Els primers anys de la seva carrera, les publicacions van estar molt separades temporalment, i principalment van ser àlbums, primer amb Slow Wine Mixtape i després amb Worldwide Angel. No va publicar singles fins al 15 de maig de 2020, quan va publicar Bom Bom. Posteriorment, ha anat publicant singles fins a l'actualitat, a excepció del 18 de març de 2021, quan va publicar el seu tercer àlbum, Warm Up.



II. Il·lustració 1: Evolució temporal per dia. Font: Elaboració pròpia.

En el període final, durant l'any 2023, hi ha molts pics en poc temps. Hi ha dies en els quals s'han publicat moltes cançons de Bad Gyal a la vegada. Per visualitzar millor si es tracta de publicacions de singles o àlbums, s'ha creat la il·lustració 2 on es veuen

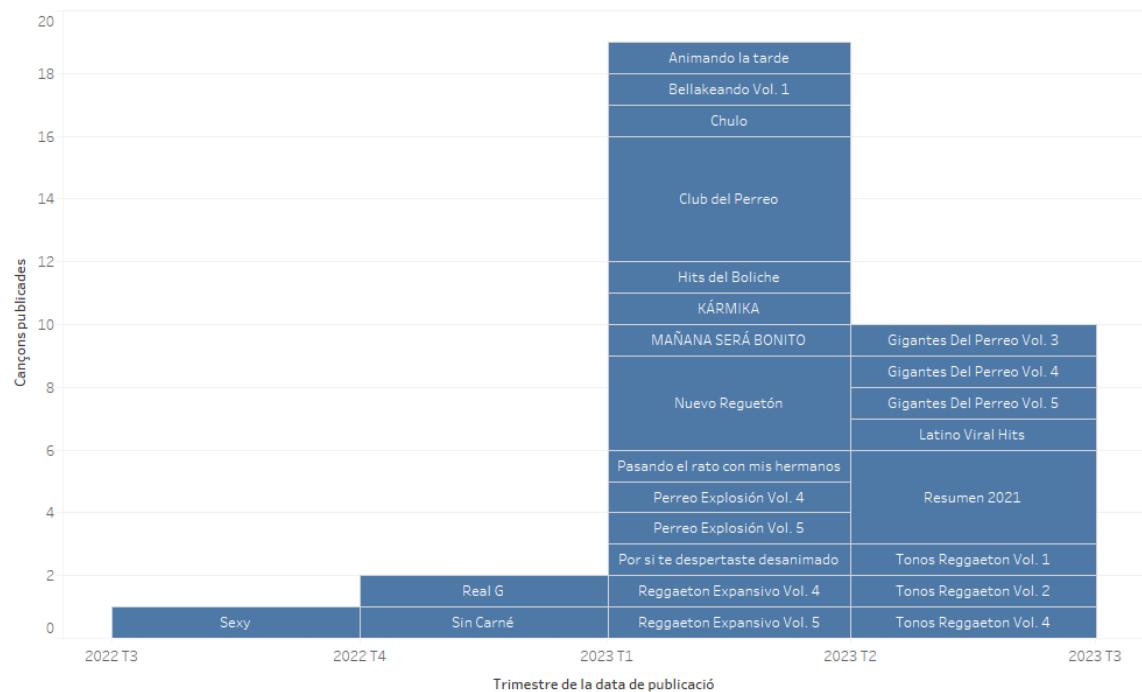
apilades les cançons, separades per els àlbums als quals pertanyen. Per millorar la visualització, s'han agrupat les publicacions per trimestres, de tal manera que es veu més clarament la diferència entre *singles* i àlbums i la freqüència de publicació.



Il·lustració 2: Evolució temporal per cançó i àlbum per trimestre. Font: Elaboració pròpia.

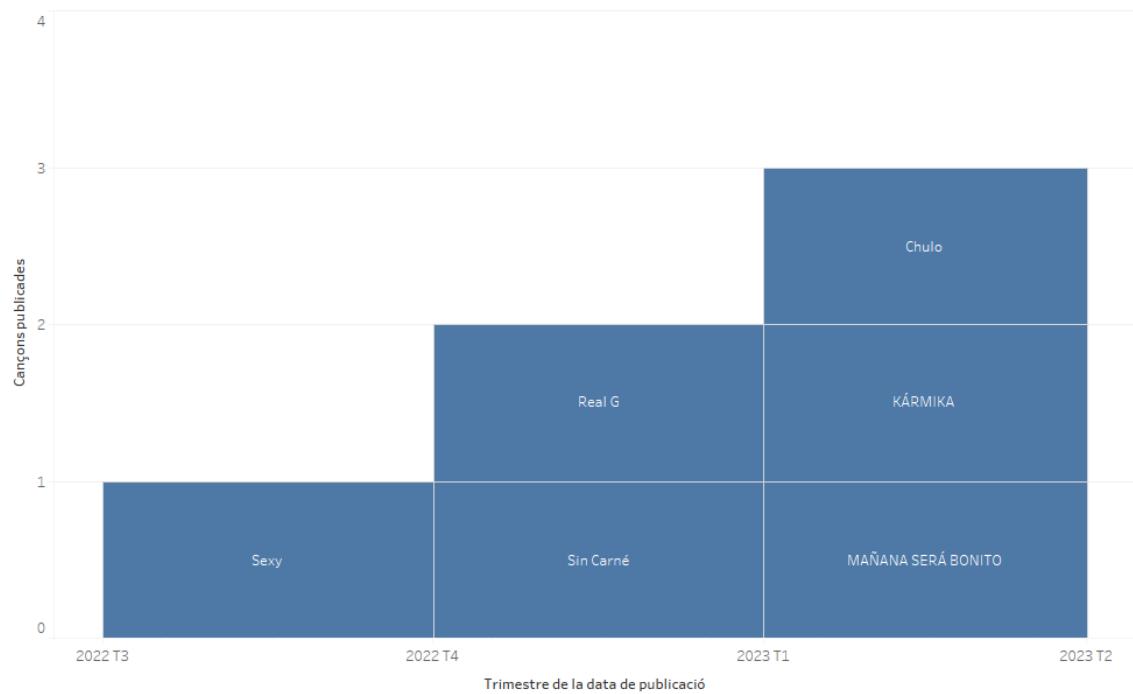
En la part final del gràfic, es pot veure una acumulació molt gran de singles i àlbums que és incomparable a les anteriors dades. Si bé és veritat que Bad Gyal ha fet publicacions cada vegada més freqüents, aquestes dades no són realistes.

Estudiant més a fons l'últim any de les dades obtingudes, en la il·lustració 3 es pot observar que hi ha molts títols d'àlbums que no es corresponen a llançaments propis i que són aparicions en àlbums externs que funcionen com a recopilatoris. Aquest contingut apareix perquè Spotify classifica de la mateixa manera les col·laboracions de cançons oficials amb altres artistes i les aparicions en àlbums de recopilació.



II·lustració 3: Evolució temporal per cançó i àlbum per trimestres 2022-2023. Font: Elaboració pròpia.

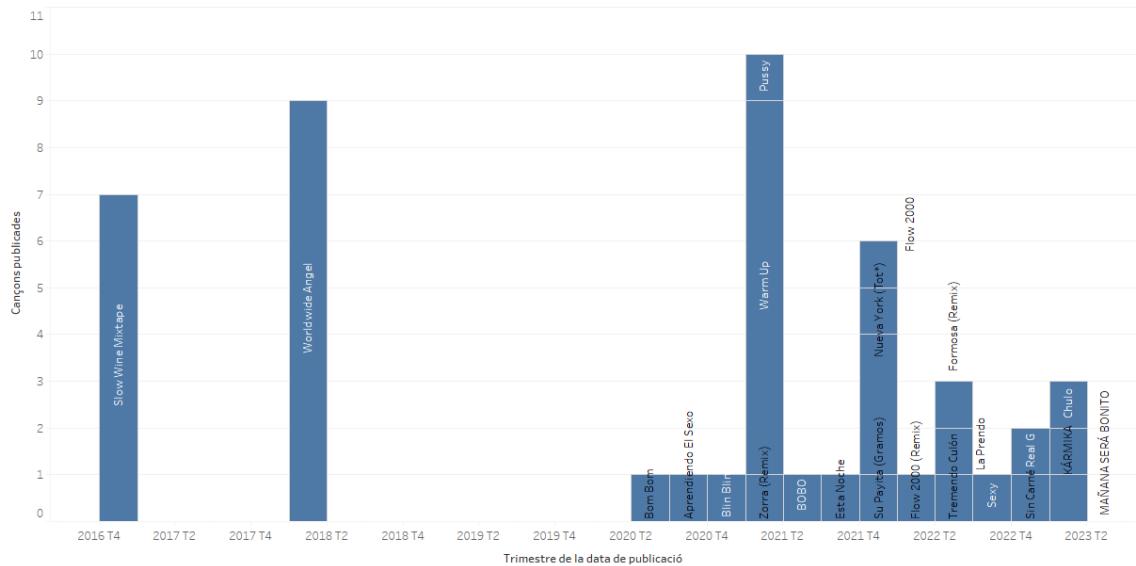
D'aquesta manera, si es netegen les dades per tenir exclusivament les cançons originals sense recopilatori, s'obté la il·lustració 4. I així s'observa que en els últims dos trimestres, exclusivament hi ha sis singles de l'artista, a diferència de les 33 publicacions que apareixen en les dades anteriors.



Il·lustració 4: Evolució temporal per cançó i àlbum per trimestres 2022-2023 (sense àlbums recopilatoris). Font: Elaboració pròpia.

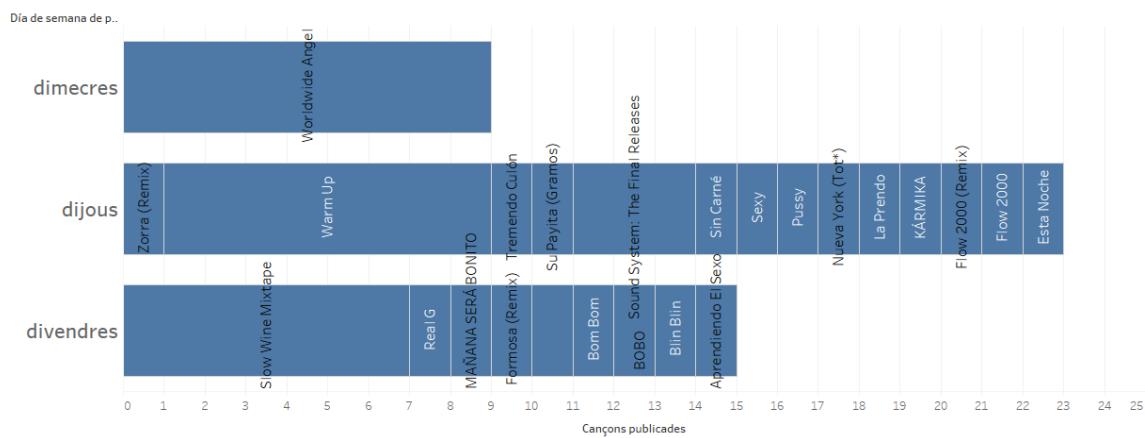
És important diferenciar el contingut dels recopilatoris, però també cal valorar la importància d'aquest fet en la seva evolució i repercussió. Si Bad Gyal comença a aparèixer en altres llistes i àlbums que recopilen els èxits més destacats, vol dir que té una certa importància, que la fa passar per davant d'altres artistes coetanis. Així, es pot afirmar que si bé no és un contingut oficial, aparèixer en les recopilacions vol dir que les seves cançons tenen una repercussió perquè es troben en el capdamunt del seu gènere musical.

Finalment, el gràfic de publicacions de les cançons i els àlbums de Bad Gyal, una vegada esborrades les dades de recopilatoris extraoficials, es mostra a il·lustració 5.



II·lustració 5: Evolució temporal per cançó i àlbum per trimestres sense recopilatori. Font: Elaboració pròpia.

A part de l'evolució temporal, també es pot observar diferents fets destacables. Un exemple és el dia de publicació de les seves cançons. Exclusivament ha fet llançaments els dimecres, els dijous i els divendres. La publicació abans del cap de setmana és una estratègia molt seguida pels artistes actuals, a la que Bad Gyal també s'hi afegeix. Això es fa perquè durant el cap de setmana els usuaris escolten més música nova i també surten més de festa, moment en el qual acostumen a sonar els últims llançaments musicals.

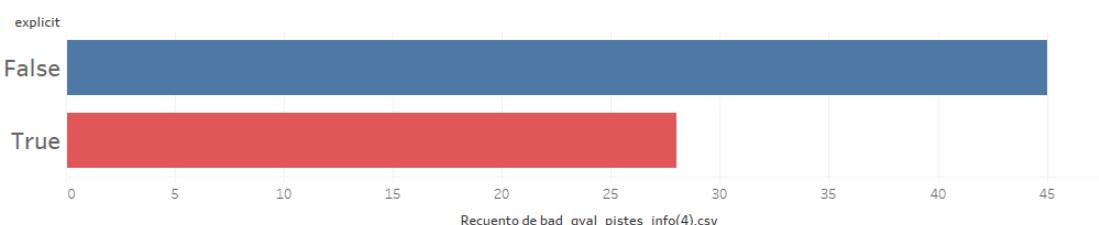


II·lustració 6: Publicacions per dia de la setmana. Font: Elaboració pròpia.

Però Bad Gyal destaca perquè ha fet un gran nombre de llançaments en dijous (23 publicacions); davant dels dimecres, en que únicament va publicar les 9 cançons del seu primer àlbum (2016); i de les que ha fet en divendres (15 publicacions), que és el dia més habitual de llançament entre els artistes. Això podria ser degut a diferents motius com posicionar la seva cançó abans que surtin les competidores del cap de setmana. O també per assegurar l'aparició a les festes que es fan la nit de dijous i divendres, ja que el seu públic és força jove i també surt de festa en dies laborables. Aquest fet és conegut per l'artista, ja que en els seus últims llançaments, ha fet contingut especial en certes discoteques per promocionar-lo les nits anteriors a la seva publicació (LOS 40, 2021).

En relació amb el mes de llançament, també es pot observar que durant l'abril, l'agost i el setembre no ha publicat mai cap cançó. En la resta de mesos sempre ha fet alguna publicació, i destaquen els mesos de febrer i març, així com novembre i desembre. Probablement aquests fets es deuen a que evita publicar cançons en èpoques de vacances en les quals el públic no acostuma a investigar tant les novetats. Així publica molt contingut abans del període habitual de vacances d'estiu i hivern, i evita publicar durant aquest període i el mes següent (agost i setembre a l'estiu, i desembre i gener a l'hivern).

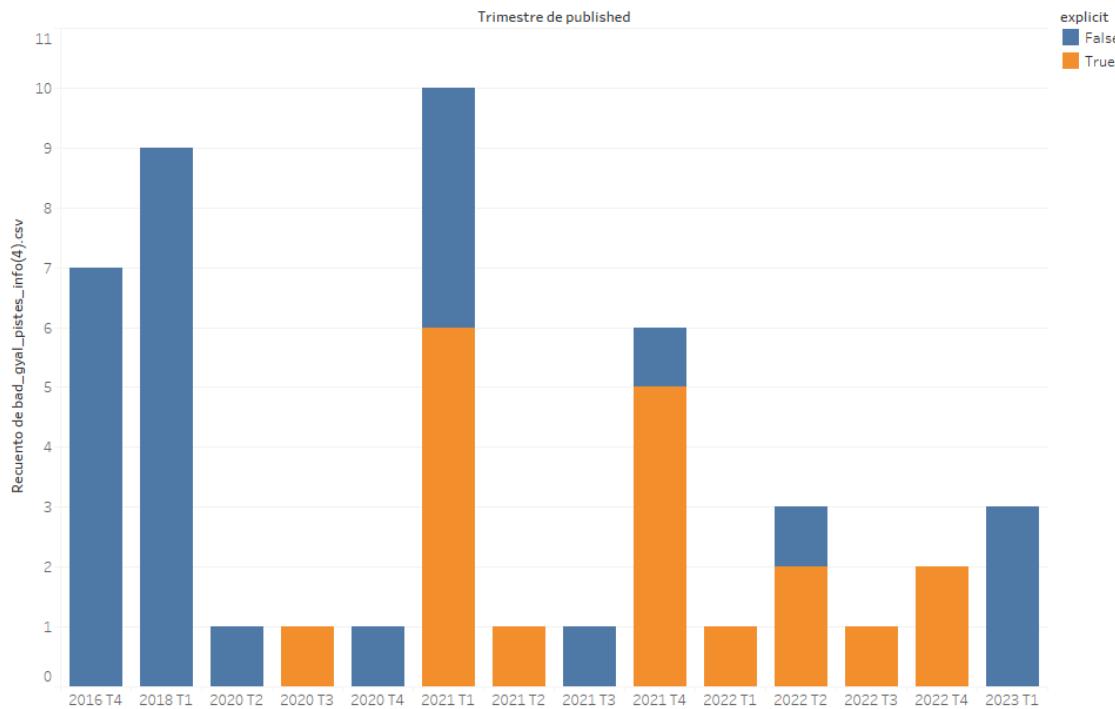
#### *Tipus de publicacions*



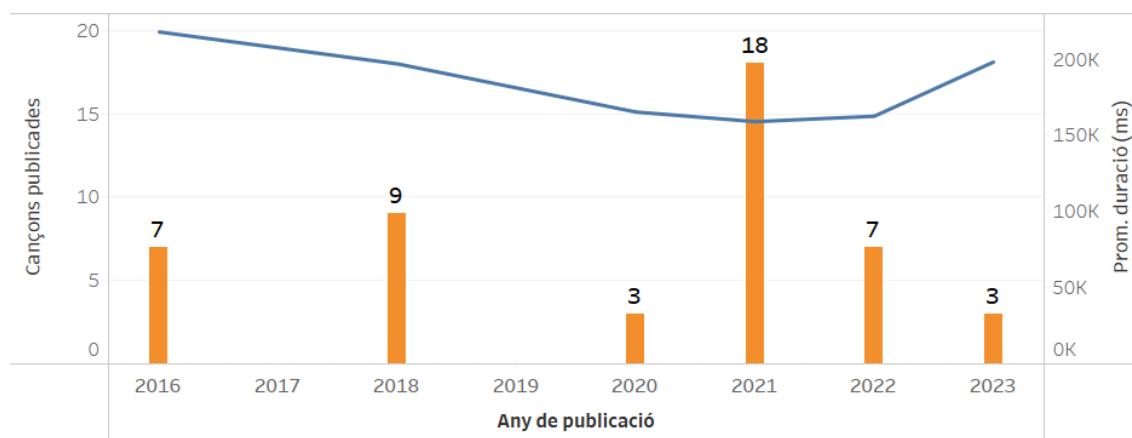
Il·lustració 7: Contingut explícit a Spotify. Font: Elaboració pròpia.

Tot i que el seu públic és jove, és prou madur per a poder escoltar contingut explícit, i entendre'l, ja que part de les seves cançons n'inclouen. En total, un 40,43% de les seves cançons són explícites, un percentatge força alt, ja que les pistes classificades d'aquesta manera poden tenir problemes de publicació en els serveis i també de monetització. Però això no ha aturat a l'artista de crear-ne més, perquè en les cançons

dels últims anys, cada vegada hi ha més publicacions amb contingut explícit, tal com es pot veure en la il·lustració 8.



Il·lustració 8: Evolució del contingut explícit per trimestre. Font: Elaboració pròpia.



Il·lustració 9: Evolució del nombre de cançons respecte la seva durada. Font: Elaboració pròpia.

Per acabar amb les dades de Spotify, cal destacar que la mitjana de la durada de les cançons era més elevada a l'inici, amb el llançament dels seus dos primers àlbums (217 segons i 196 segons respectivament). Així, amb el llançament majoritari de *singles* el 2020, 2021 i 2022, la durada de les cançons ha estat força més curta (al

voltant dels 160 segons). Destaca l'any 2021 perquè va publicar fins a 18 cançons, però molt més curtes (158 segons).

De moment, el 2023 està publicant cançons força més llargues que les anteriors (198 segons). S'haurà d'anar seguint en els mesos vinents si es consolida aquesta tendència, o si, en canvi, torna al que ha fet els últims anys amb moltes més cançons però especialment curtes.

### 5.1.2. Anàlisi de la relació amb altres artistes a Spotify

En aquest apartat, s'analitzarà la relació que té Bad Gyal amb altres artistes. En aquest cas, s'ha volgut crear un graf per poder representar la proximitat més fàcilment. Per aconseguir-ho s'ha utilitzat l'API de Spotify, a la qual s'ha accedit mitjançant la programació d'un codi des de Pycharm.

Primer de tot, ha fet falta extreure totes les dades a partir de l'API i desar-les en un format que funcioni amb Gephi, l'eina de creació de grafs triada. En aquest cas, el fitxer final ha sigut un CSV, amb dues columnes: *source* i *target*.

El codi es divideix en tres parts, la primera part és una preparació dels elements necessaris, de manera semblant al codi anterior. La segona part es tracta d'automatitzar el procés d'extracció dels artistes relacionats amb la funció de Spotify *.artist\_related\_artists* i dos bucles for, el primer per extreure els artistes relacionats amb Bad Gyal directament; i després aquells que Spotify recomana a partir de cada artista recomanat en primer lloc. Finalment, la tercera part crea un *Dataframe* i després extreu un fitxer CSV per poder exportar-ho correctament a Gephi.

El programa que s'ha creat, que està comentat perquè s'entengui millor, és el que s'exposa a Codi 2, a continuació.

```

#Importació de les llibreries
import spotipy
from spotipy.oauth2 import SpotifyClientCredentials
import pandas as pd

#Codi d'autenticació
client_id = 'clau_client'
client_secret = 'clau_secreta'
client_credentials_manager = SpotifyClientCredentials(client_id=client_id, client_secret=client_secret)
sp = spotipy.Spotify(client_credentials_manager=client_credentials_manager)

#Establím les variables necessàries
source_artist_id = '4F4pp8NUW08JuXwnoxglpN'
source_artist_name = sp.artist(source_artist_id)['name']
count = 0
relacions = []
relacionats2n = []

#executem el codi per extreure els artistes relacionats amb Bad Gyal
related_artists = sp.artist_related_artists(source_artist_id) #obtenim els artistes
relacionats = related_artists["artists"] #agafem només la llista d'artistes

for l in relacionats: #iterem per a cada artista
    relacionats2n.append(l["id"]) #afegeim a la llista per explotar-los en segon nivell
    related_artist_name = l["name"] #agafem només el nom
    tupla = (source_artist_name,related_artist_name) #Afegim en tuples totes les relacions
    relacions.append(tupla) #ho afegeim a la llista final

#print(relacionats2n) #printem per comprobar que s'estan afegint bé els de 2n nivell

#per a cada artista relacionat amb BadGyal busquem els seus artistes relacionats també (2n nivell)
for e in relacionats2n:
    related_artists = sp.artist_related_artists(e) #obtenim els relacionats de cada artista
    relacionats = related_artists["artists"] #filtrarem les dades
    source_artist_name = sp.artist(e)['name'] #posem el nom d'origen
    for l in relacionats:
        related_artist_name = l["name"] #afegeim el nom del relacionat
        tupla = (source_artist_name, related_artist_name) #creem la tupla
        relacions.append(tupla) #ho afegeim a la llista final

df = pd.DataFrame.from_records(relacions, columns=['source','target']) #creem el df amb "relacions" i
#(source - target)

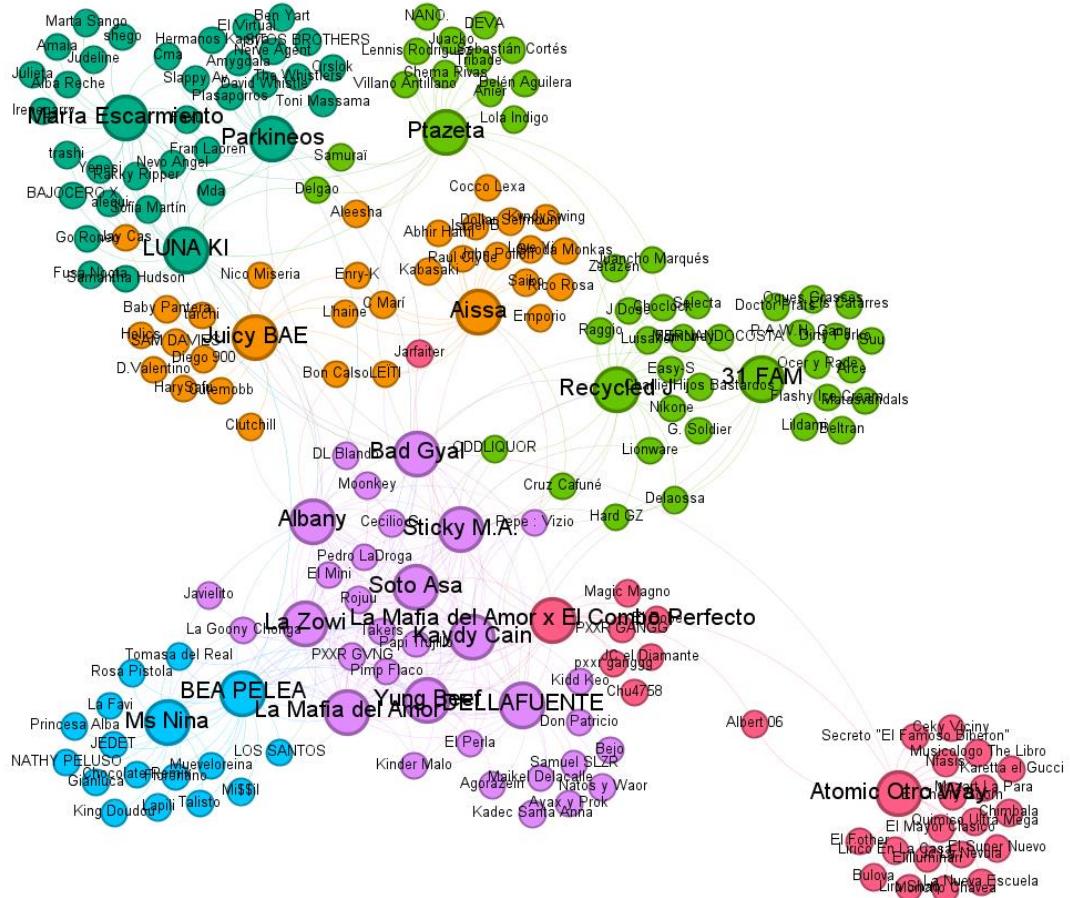
#Imprimim el df
print(df)

#Exportem el df a un arxiu .csv
df.to_csv('bad_gyal_related2nivell.csv', index=False)

```

Codi 2: Extracció d'artistes relacionats. Font: Elaboració pròpria.

Posteriorment, s'ha exportat el fitxer obtingut a Gephi i li he donat format per tal d'obtenir un graf que sigui fàcilment comprensible. S'han creat dues representacions, per poder veure el grau de vinculació amb Bad Gyal, i també la direccionalitat d'aquesta.



II·lustració 10: Graf d'artistes relacionats amb Bad Gyal per grau de sortida. Font: Elaboració pròpia.

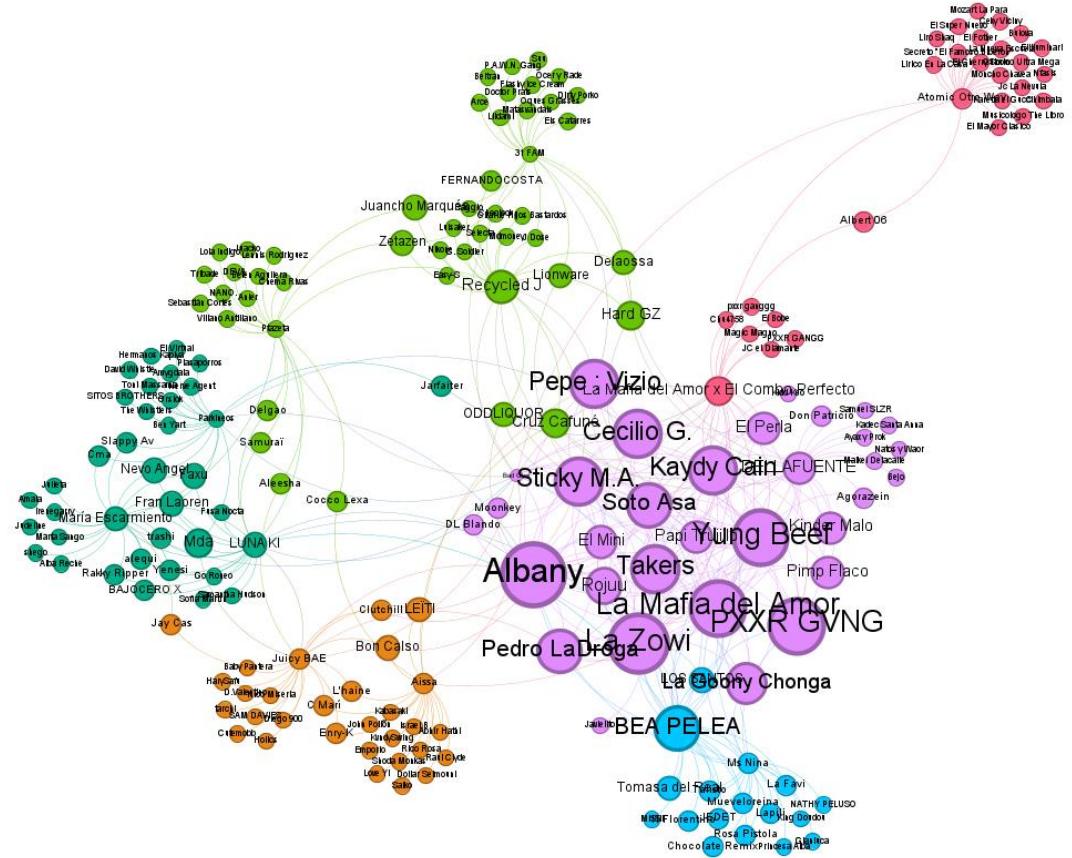
En primer lloc, trobem aquest graf que ens mostra el grau de vinculació amb Bad Gyal segons el grau de sortida. Per entendre el gràfic, és important tenir en compte que per cada artista es proporcionen 20 recomanacions. D'aquesta manera, al centre trobem Bad Gyal, ja que està vinculada amb tots els artistes de primer grau de recomanació, representats amb un node i un títol més grans. A partir de cada artista recomanat es creen 20 suggeriments més, que són representats amb un node més petit. Aquestes recomanacions de segon nivell, poden retornar a altres artistes que apareixen al graf, situació que s'analitzarà més endavant en el segon graf extret de Gephi.

És interessant observar que hi ha recomanacions de primer nivell molt vinculades entre si i d'altres que actuen com a *hub*, arribant a crear illes allunyades. És el cas dels nodes de color vermell, relacionat amb el *reggaeton* més antic (punts centrals) i el *dembow* actual (punts allunyats). Aquest fet, també passa amb els artistes de música en català. Com Bad Gyal és de Barcelona, també està vinculada als artistes catalans, concretament des de la recomanació del grup 31FAM, que actua com a *hub*.

Al centre, trobem l'escena *urban* més propera a Bad Gyal, representada de color morat. Aquestes recomanacions són molt més pròximes a l'artista. I en color turquesa també trobem una comunitat *urban* força vinculada.

A la part superior, de color verd fosc, hi ha la comunitat de música electrònica i *hyperpop*, a la qual Bad Gyal està vinculada a través de Luna Ki, María Escarmiento i Parkineos.

Les comunitats en verd clar, són artistes més vinculats al pop. I la comunitat en taronja es tracta d'artistes de R&B, tot i que també d'altres gèneres com el rap i el trap.



Il·lustració 11: Graf d'artistes relacionats amb Bad Gyal per grau d'entrada

Per poder veure els vincles amb els artistes relacionats de Bad Gyal, és molt important veure quin nombre de recomanacions hi ha entre els artistes del seu entorn. És a dir, en aquest segon gràfic es mostra la mida del node segons el nombre de recomanacions que rebi, o sigui, el grau d'entrada. Per començar es pot observar que les comunitats són aproximadament les mateixes i estan representades del mateix color, tot i que han canviat de lloc. Quant a la posició de Bad Gyal, trobem que el seu node és molt central, com és lògic, però és molt petit. Això significa que els artistes recomanats en primer grau, no tornen a recomanar-la a ella, sinó que recomanen altres artistes.

Destaca el grup central, similar al de l'anterior gràfic i representat de color morat. Aquests artistes tenen uns nodes molt grans, el que vol dir que han estat molt recomanats i també tenen gairebé tots el mateix color, el que significa que també estan

molt interrelacionats. Així, destaquen artistes com Albany, La Zowi, PXXR GVNG, Yung Beef o La Mafia del Amor.

Respecte a la resta de comunitats, no s'observen gaires variacions entre el nombre de recomanacions dels artistes, a excepció de Bea Pelea (color blau) i Recycled J (color verd). Aquests artistes sí que tenen una mida més gran que la de les seves comunitats respectives.

### 5.1.3. Anàlisi de la relació amb altres gèneres a Spotify

Però per entendre l'entorn de Bad Gyal i amb què està relacionat també cal analitzar els seus gèneres pròxims. Per aconseguir-ho, s'ha creat un graf a partir dels gèneres que recomana Spotify. La intenció inicial era obtenir els gèneres que Spotify assignava a cada cançó o els gèneres que recomanava a partir de les cançons de Bad Gyal, però l'API no ho permet. Per obtenir gèneres, únicament es poden extreure a partir del perfil de l'artista. D'aquesta manera, el que s'ha fet és extreure els artistes de cada cançó en una llista. I posteriorment extreure els gèneres amb els quals es relaciona cada artista.

Això s'ha aconseguit amb el mateix mètode que en l'anterior graf. Primer s'ha extret la informació de l'API, s'ha transformat i s'ha exportat a un fitxer CSV amb un programa que es mostra a Codi 3.

```
#Importació de les llibreries
import spotipy
from spotipy.oauth2 import SpotifyClientCredentials
import pandas as pd

#Codi d'autenticació
client_id = 'clau_client'
client_secret = 'clau_secreta'
client_credentials_manager = SpotifyClientCredentials(client_id=client_id, client_secret=client_secret)
sp = spotipy.Spotify(client_credentials_manager=client_credentials_manager)

df_inicial = pd.read_csv("bad_gyal_pistes_info(5).csv", header=0)
llista_artist=[]
llista_ids=[]
llista_genres=[]
count=0

for track in df_inicial['id']:
    #print(track_id)
    track_info = sp.track(track)
    for info in track_info['artists']:
        llista_ids.append(info['id'])

    genres=[]
    for artist_id in llista_ids:
        artist = sp.artist(artist_id)
        genres.append(artist['genres'])
        for g in artist['genres']:
            tupla = (track_info['name'], g) #creem la tupla
            llista_genres.append(tupla)
    count += 1
    print(count)

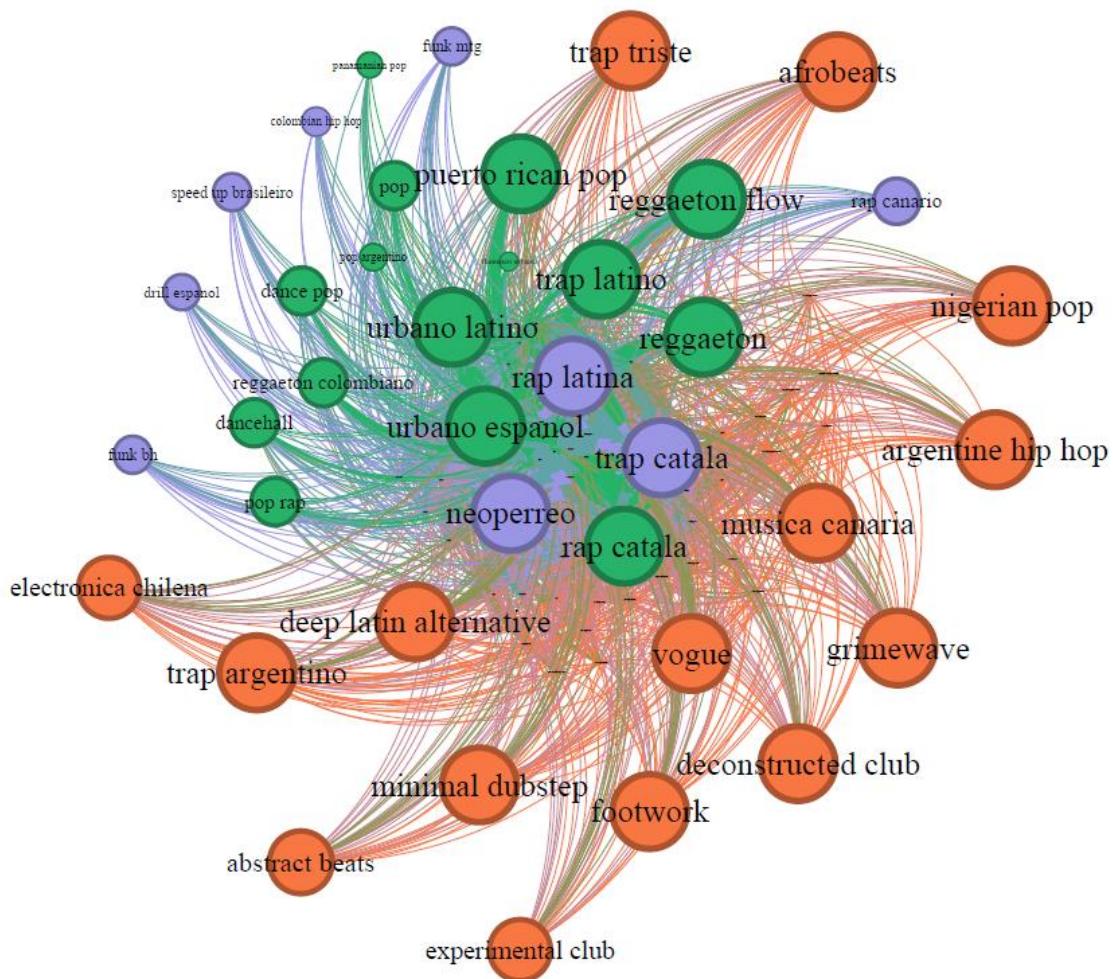
#print(llista_genres)
df = pd.DataFrame.from_records(llista_genres, columns=['source','target']) #creem el df amb "relacions"
#(source - target)

#Imprimim el df
print(df)

#Exportem el df a un arxiu .csv
df.to_csv('bad_gyal_related_genres.csv', index=False)
```

Codi 3: Extracció dels gèneres relacionats. Font: Elaboració pròpia.

Una vegada obtingudes les dades en un fitxer CSV preparat per a Gephi, s'ha importat a aquesta eina. Ajustant els valors i assignant les propietats adequades, així com la creació de comunitats a partir d'estadístiques, s'ha obtingut el següent graf, exposat a II·lustració 12.



II·lustració 12: Graf de gèneres relacionats amb Bad Gyal per grau d'entrada. Font: Elaboració pròpia.

El graf que s'ha obtingut té una forma molt concreta, ja que hi ha uns elements ubicats al centre que estan molt pròxims i després la majoria de nodes es troben allunyats als afores. Al centre hi trobem gèneres molt vinculats amb l'artista, com són: *urbano espanol*, *urbano latino*, *rap latina*, *trap latino*, *reggaeton*, *trap catala*, *rap catala*, *neoperreo*, *puerto rican pop* o *reggaeton flow*. A continuació trobem els nodes de color taronja, que si bé apareixen repetidament en els artistes, no estan gaire recomanats de manera transversal.

En relació amb les comunitats, en trobem tres. Al centre trobem la comunitat de color verd, que està vinculada als gèneres més *mainstream* i aquests nodes es podrien agrupar dins els gèneres *urban* i *pop*, que són els més coneguts i més influents

actualment. Els gèneres més recomanats d'aquesta comunitat estan molt interconnectats i són molt propers a Bad Gyal.

En relació amb la comunitat representada de color lila, únicament té tres gèneres molt recomanats, però aquests estan situats molt al centre, de manera que estan fortament vinculats. Però la resta dels sis gèneres de la comunitat apareixen poc i molt allunyats. En general, es podria conèixer a aquesta comunitat com a una recopilació de gèneres més underground, amb certa vinculació a la música electrònica i al rap, i que també apareixen en l'entorn de Bad Gyal.

Finalment, hi ha una tercera comunitat de color taronja que és més llunyana, però que apareix moltes vegades en les recomanacions. Aquesta comunitat incorpora una gran diversitat de gèneres, entre els quals no es troba cap relació destacable a excepció que es tracta de categories molt específiques.

## 5.2. YouTube

YouTube és la plataforma de vídeo líder a escala mundial. Per això els artistes hi publiquen les seves cançons, i especialment destaca la importància del contingut musical amb imatge, com els videoclips o els *lyric video*. A més, des dels seus inicis ha estat una plataforma molt més oberta i col·laborativa que Spotify, per això és interessant analitzar YouTube, perquè és el lloc on Bad Gyal va començar publicant les seves cançons sense cap suport professional. Les cançons i el nombre de publicacions en una plataforma i una altra poden ser diferents, no tenen per què ser les mateixes. A més també cal destacar que YouTube ofereix moltes més opcions d'interaccions, és fàcil crear canals, fer llistes, pujar vídeos, seguir canals o interaccionar amb m'grades i comentaris. De fet, els comentaris són una eina bàsica per a l'anàlisi de la conversa de Bad Gyal.

L'extracció de dades ha estat força més complicada que la de Spotify, ja que inicialment s'intentava evitar utilitzar l'API per no malgastar recursos, però finalment sí que s'ha utilitzat. Aquesta situació es detalla a l'apartat 5.2.1.

A partir de YouTube es volen extreure diferents dades com la informació dels canals oficials de l'artista, les característiques de cada vídeo i els comentaris que hi ha. Aquesta és una font molt interessant perquè inclou un canal amb més antiguitat que la resta, ja que s'hi va unir el 8 de març de 2016, aproximadament quan publica la seva primera cançó, PAI. YouTube permet extreure dades sobre la seva carrera musical com a Spotify, però també altres tipus de contingut de vídeo com *DJ sets*, directes o videoclips. La plataforma també permet l'obtenció de dades sobre el comportament dels usuaris. Es poden extreure diferents tipus d'interacció com *follows*, *likes* o *comments*. Aquest últim tipus d'interacció permet extreure molta més informació per saber què es pretén amb aquesta interacció.

Més endavant es tractarà tema per tema el contingut que es vol extreure de YouTube. Primer amb l'extracció d'informació del canal i després dels vídeos.

### 5.2.1. Font de l'extracció de dades

Per extreure les dades de YouTube, existeixen moltes eines d'automatització del procés sense necessitat de crear codi. Però degut a l'opacitat, procedència de les dades i disponibilitat del servei no s'han pogut utilitzar. En un principi, es volia fer l'extracció amb YouTube Data Tools, una plataforma prou transparent amb el tractament de les dades (YouTube Data Tools, 2023). Però les dificultats per fer funcionar les consultes van provocar que es descartés l'opció. Entre els motius, hi ha el desfasament del servei, la impossibilitat per repetir processos de manera automatitzada, el temps d'espera o l'aparició d'errors en la resposta.

Pels motius exposats anteriorment, es va decidir d'extreure les dades directament des de l'API de YouTube. Primer va ser necessari configurar tots els serveis, permisos, aplicacions i claus des del web de Google Developers. Google té la informació i els passos sobre com extreure les dades d'una manera molt didàctica, de qualitat i senzilla a la web sobre la seva API (Google, 2023). Resumint, els passos previs de preparació es tracta de crear un perfil de desenvolupador amb un compte de Google, crear unes credencials i autoritzar l'ús de “YouTube Data API v3” (Google Cloud console, 2023).

En segon lloc, i una vegada configurades les credencials, es comença a treballar en el codi. La documentació que aporta el mateix web de YouTube Data Api és molt útil, ja que a través d'un entorn web, permet fer la petició i després traduir-ho al llenguatge de programació adient. En qualsevol cas, per aquest treball no s'ha utilitzat l'entorn de programació des de web, sinó el programa d'edició de codi PyCharm. D'altra banda, el codi d'autenticació sí que s'ha extret des de l'apartat de desenvolupador, així com s'ha utilitzat per cercar les funcions necessàries per extreure la informació.

Finalment, el codi utilitzat per extreure totes les dades del canal i de cada vídeo que apareix al canal de YouTube de Bad Gyal, es mostra a continuació en diferents quadres de codi. S'ha treballat amb documents diferents de programació, així com s'han desat les respostes que dona l'API per no haver de fer peticions a cada prova, perquè hi ha una quota màxima de 10.000 consultes per dia. Durant les proves, s'ha arribat a aquest límit abans de treballar amb documents separats. Arran d'aquest problema, s'ha duplicat l'aplicació des del web de Google Developers i s'han obtingut

unes noves credencials, així com, també s'ha separat les parts de codi i les respostes en fitxers perquè no es consumeixi tan ràpidament la quota.

### 5.2.2. Anàlisi de la informació del canal

Per obtenir una visió general, especialment de certes mètriques pròpies del canal, es va utilitzar el Codi 4 exposat a continuació. Primer de tot s'importen totes les llibreries de Google necessàries per establir comunicació amb l'API de YouTube, després es fa la petició d'autorització, que requereix proporcionar les dades d'accés a l'aplicació creada anteriorment des de la web de Google Developers, és a dir el fitxer JSON amb les claus. Després cal iniciar sessió amb un compte de Google cada vegada que s'executa el codi. Aquest pas ha augmentat molt el temps dedicat a la creació i execució del codi de YouTube.

Finalment, s'executa el codi que s'ha creat concretament per extreure les dades. Per aconseguir la informació del canal s'ha utilitzat la funció “channels.list”. A partir de la resposta, s'ha agafat la part que té informació útil (`response['items']`), i s'ha creat un *Dataframe*. Així, s'ha pogut exportar aquesta informació a un fitxer CSV.

```
# -*- coding: utf-8 -*-

import os

import google_auth_oauthlib.flow
import googleapiclient.discovery
import googleapiclient.errors
import pandas as pd

scopes = ["https://www.googleapis.com/auth/youtube.readonly"]

def main():
    # Disable OAuthlib's HTTPS verification when running locally.
    # *DO NOT* leave this option enabled in production.
    os.environ["OAUTHLIB_INSECURE_TRANSPORT"] = "1"

    api_service_name = "youtube"
    api_version = "v3"
    client_secrets_file = "client_secret.json"

    # Get credentials and create an API client
    flow = google_auth_oauthlib.flow.InstalledAppFlow.from_client_secrets_file(
        client_secrets_file, scopes)

    credentials = flow.run_local_server(port=0)
    youtube = googleapiclient.discovery.build(
        api_service_name, api_version, credentials=credentials)

    request = youtube.channels().list(
        part="snippet,contentDetails,statistics",
        id="UC2ypBaYnDvnLbzyAH8w2jsw"
    )
    response = request.execute()

    #print(response)
    df = pd.json_normalize(response['items'])
    print(df)
    # Exportem el df a un arxiu .csv
    df.to_csv('yt_channel_info.csv', index=False)

if __name__ == "__main__":
    main()
```

Codi 4: Extracció de dades del canal de YouTube. Font: Elaboració pròpia.

A partir d'aquesta informació s'han extret diverses conclusions i s'han preparat altres dades per a la seva posterior utilització. Entre aquests destaquen certes estadístiques com el recompte de visualitzacions, el nombre de subscriptors, i especialment destaca el nombre de vídeos publicats, ja que no es correspon amb el nombre de vídeos que apareixen al seu apartat de vídeos publicats.

Taula 2: Informació del canal de YouTube "Bad Gyal". Font: Elaboració pròpia.

| snipp<br>et.title | snippet.pu<br>blishedAt | snippet.localiz<br>ed.description | statistics.<br>viewCount | statistics.sub<br>scriberCount | statistics.v<br>ideoCount |
|-------------------|-------------------------|-----------------------------------|--------------------------|--------------------------------|---------------------------|
| Bad<br>Gyal       | 8/3/2016<br>14:06:26    | Tightess punani<br>pussy k mana   | 516985244                | 834000                         | 42                        |

Aquest fet va ser un problema important, ja que únicament s'estava carregant una part dels vídeos de l'artista i va ser necessari analitzar bé el fitxer de resposta en JSON que va proporcionar l'API, i també observar els vídeos que hi havia a la pàgina web de YouTube de Bad Gyal. En el desenvolupament del proper codi s'explica la solució a aquest problema.

### 5.2.3. Anàlisi de la informació de cada vídeo

Per extreure les dades de tots els vídeos al canal de Bad Gyal sense acabar amb la quota de consultes a l'API, va ser necessari fer dos fitxers de codi. Primer un per extreure la llista de vídeos que té el canal i exportar-la a un fitxer local, i el segon per extreure la informació que es necessita de cada vídeo. S'exposen els codis continuació, i es fa una valoració respectivament.

```
# -*- coding: utf-8 -*-

import os

import google_auth_oauthlib.flow
import googleapiclient.discovery
import googleapiclient.errors
import pandas as pd
from pyasn1.compat.octets import null

scopes = ["https://www.googleapis.com/auth/youtube.readonly"]
llista_dfs = []
channels = ['UC2ypBaYnDvnlbzyAH8w2jsw', 'UC0mtuk6Mp66DRFcgy81kr3eQ']

def get_video_list_in_dataframe() -> (pd.DataFrame, google_auth_oauthlib.flow.InstalledAppFlow):
    # Disable OAuthlib's HTTPS verification when running locally.
    # *DO NOT* leave this option enabled in production.
    os.environ["OAUTHLIB_INSECURE_TRANSPORT"] = "1"

    api_service_name = "youtube"
    api_version = "v3"
    client_secrets_file = "client_secret_tfg(2).json"

    # Get credentials and create an API client
    flow = google_auth_oauthlib.flow.InstalledAppFlow.from_client_secrets_file(
        client_secrets_file, scopes)

    credentials = flow.run_local_server(port=0)
    youtube = googleapiclient.discovery.build(
        api_service_name, api_version, credentials=credentials)

    for channel in channels:
        request = youtube.search().list(
            part="snippet",
            channelId=channel,
            type='video',
            maxResults=50
        )
        response = request.execute()
        llista_dfs.append(pd.json_normalize(response['items']))
        print(response)

        try:
            hasNextPage = response['nextPageToken']
        except:
            hasNextPage = null

        while hasNextPage:
            request = youtube.search().list(
                part="snippet",
                channelId=channel,
                pageToken=response['nextPageToken'],
                maxResults = 50,
                type="video"
            )
            response = request.execute()
            llista_dfs.append(pd.json_normalize(response['items']))
            print(response)

            try:
                hasNextPage = response['nextPageToken']
            except:
                hasNextPage = null

    # Creem el df a partir de la llista
    df_final = pd.concat(llista_dfs)
    print(df_final)
    df_final.to_csv('yt_videos_list(2).csv', index=False)
    return (df_final, youtube)

# print(response)
# df = pd.json_normalize(response)
# print(df)
# Exportem el df a un arxiu .csv

if __name__ == "__main__":
    get_video_list_in_dataframe()
```

Codi 5: Extracció de la llista de vídeos de Bad Gyal a YouTube. Font: Elaboració pròpia.

El programa de codi 5i té les tres parts que hi ha en qualsevol anterior: preparació de l'accés, obtenció i ordenació de les dades, i exportació. En aquest cas, igual que en l'anterior es crea una funció i després s'executa.

En la part central del codi, destaca la primera iteració: “for channel in channels:”. Ha sigut necessari crear aquest bucle perquè sinó no apareixen tots els vídeos. En el primer codi de YouTube s'ha pogut veure que el nombre de vídeos era molt inferior al que realment apareix si ho consultes manualment. Això és degut al fet que tots els vídeos són de l'artista Bad Gyal i apareixen al canal Bad Gyal, però no tots es gestionen des d'aquest canal. De fet, hi ha molts vídeos de l'artista que pertanyen al seu compte de Vevo, com es pot comprovar simplement veient l'URL. Per això, quan es fa la consulta cal repetir el procés pels dos canals oficials que té l'artista.

Després, cal obtenir els vídeos en si. Però com el límit de respostes són 40, cal utilitzar PageToken per repetir les consultes de l'API fins a arribar a obtenir tots els vídeos. Perquè una vegada acabi no aparegui un error, cal incloure també *try, except* i *while*.

Així s'obtenen tots els identificadors del canal, i posteriorment s'exporten com a CSV.

```
# -*- coding: utf-8 -*-

import os

import google_auth_oauthlib.flow
import googleapiclient.discovery
import googleapiclient.errors
import pandas as pd
import numpy as np
from pyasn1.compat.octets import null

scopes = ["https://www.googleapis.com/auth/youtube.readonly"]
llista_dfs = []
channels = ['UC2ypBaYnDvnLbzyAH8w2jsw', 'UC0mtuk6Mp66DRFcq81kr3eQ']

def main():
    # Disable OAuthlib's HTTPS verification when running locally.
    # *DO NOT* leave this option enabled in production. os.environ["OAUTHLIB_INSECURE_TRANSPORT"] =
    "1"

    api_service_name = "youtube"
    api_version = "v3"
    client_secrets_file = "client_secret_tfg(2).json"

    # Get credentials and create an API client
    flow = google_auth_oauthlib.flow.InstalledAppFlow.from_client_secrets_file(
        client_secrets_file, scopes)

    credentials = flow.run_local_server(port=0)
    youtube = googleapiclient.discovery.build(
        api_service_name, api_version, credentials=credentials)

    count = 0
    df_inicial = pd.read_csv("yt_videos_list(2).csv", header=0)
    print(df_inicial)
    # carrega bé el fitxer

    for song in df_inicial['id.videoId']:
        # print(song)
        request = youtube.videos().list(
            part="snippet,contentDetails,statistics",
            id=song
        )
        response = request.execute()
        # print(response)
        responseo = response['items'][0]
        response = response['items'][0]['snippet']
        #print(responseo)


```

```

if response.get('tags') == None:
    df = pd.DataFrame({
        'title': response['title'],
        'id': response['id'],
        'publishedAt': response['publishedAt'],
        'channelId': response['channelId'],
        'description': response['description'],
        'channelTitle': response['channelTitle'],
        'tags': "None",
        'categoryId': response['categoryId'],
        'liveBroadcastContent': response['liveBroadcastContent'],
        'duration': response['contentDetails']['duration'],
        'caption': response['contentDetails']['caption'],
        'licensedContent': response['contentDetails']['licensedContent'],
        'contentRating': response['contentDetails']['contentRating'],
        'projection': response['contentDetails']['projection'],
        'viewCount': response['statistics']['viewCount'],
        'likeCount': response['statistics']['likeCount'],
        '# favoriteCount': response['statistics']['favoriteCount'],
        'commentCount': response['statistics']['commentCount']
    }, index=[0])
else:
    df = pd.DataFrame({
        'title': response['title'],
        'id': response['id'],
        'publishedAt': response['publishedAt'],
        'channelId': response['channelId'],
        'description': response['description'],
        'channelTitle': response['channelTitle'],
        'tags': ",".join(response['tags']),
        'categoryId': response['categoryId'],
        'liveBroadcastContent': response['liveBroadcastContent'],
        'duration': response['contentDetails']['duration'],
        'caption': response['contentDetails']['caption'],
        'licensedContent': response['contentDetails']['licensedContent'],
        'contentRating': response['contentDetails']['contentRating'],
        'projection': response['contentDetails']['projection'],
        'viewCount': response['statistics']['viewCount'],
        'likeCount': response['statistics']['likeCount'],
        '# favoriteCount': response['statistics']['favoriteCount'],
        'commentCount': response['statistics']['commentCount']
    }, index=[0])
    # print(df)
    llista_dfs.append(df)
    count += 1
    print(count)

    # print(llista_dfs)
final_df = pd.concat(llista_dfs)
#final_df.to_csv('export(4).csv', index=False)
#no funciona: #np.savetxt('export(4).csv', final_df, delimiter='::')
#probem amb json
with open('export(4).json', 'w') as f:
    f.write(final_df.to_json(orient='records', lines=True))

if __name__ == "__main__":
    main()

```

Codi 6: Extracció de la informació dels vídeos de Bad Gyal a YouTube. Font: Elaboració pròpria.

Una vegada estan tots els identificadors de vídeo agrupats en el fitxer, es poden extreure les dades per a cada vídeo. En aquest nou codi, el que es fa és preparar l'api i els fitxers, després iterar per crear el *Dataframe* i finalment exportar-ho en un fitxer JSON.

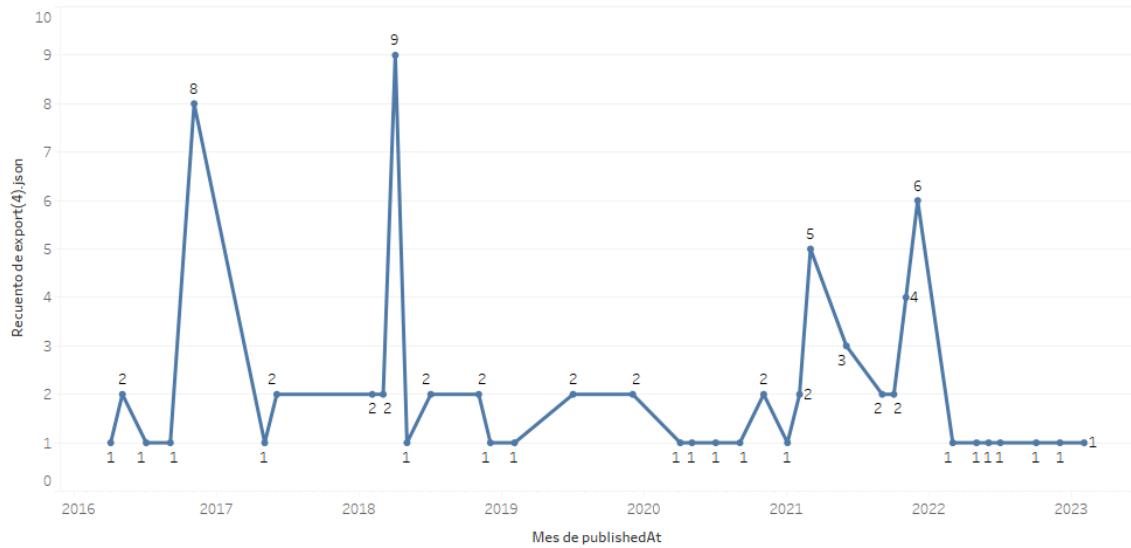
La primera part de preparació és igual que en el codi anterior, però afegint les línies necessàries per carregar la llista d'identificadors extreta abans en un fitxer.

Després s'itera per a cada cançó que hi ha en aquest fitxer de manera que es puguin extreure les dades de cada vídeo. Per aconseguir-ho cal utilitzar la funció “youtube.videos.list”, que permet demanar a l'API que torni totes les dades d'aquests. A partir de la resposta en brut que es proporciona, s'estudia quins valors són d'utilitat i llavors es crea un *Dataframe* per a cada cançó amb aquestes dades. Així, el document final que s'obté ocupa menys espai, és més fàcil d'interpretar i està net i filtrat de dades inútils. Aquest *Dataframe* s'afegeix al final d'una llista, per tal de poder-lo recuperar a l'exportació.

Finalment, es concatenen tots els *Dataframe* de la llista i s'exporta el *Dataframe* final en un fitxer JSON. El canvi al fitxer JSON és degut al fet que hi ha *strings* amb molts caràcters amb contingut divers que interfereixen amb els separadors del text quan es desa en un fitxer CSV com s'ha fet anteriorment.

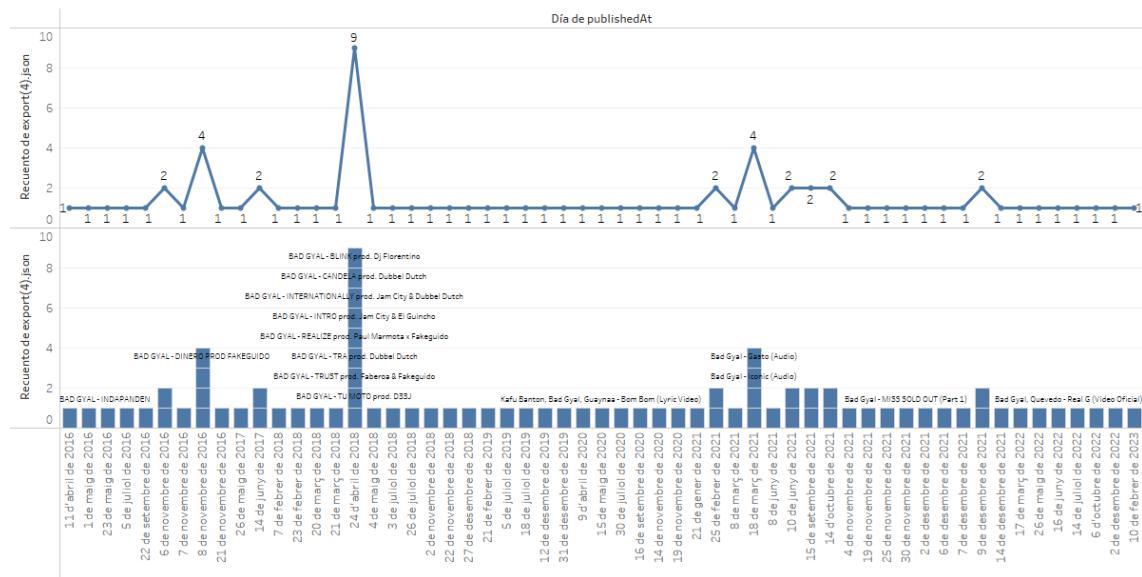
A partir del fitxer final amb totes les dades de cada vídeo, s'ha pogut extreure i visualitzar les dades per a cada vídeo del canal Bad Gyal amb Tableau. Les dades obtingudes a partir de les visualitzacions s'analitzen a continuació, incloent-hi l'anàlisi temporal, de continguts, de canal i d'interaccions.

### Anàlisi de l'evolució temporal



II·lustració 13: Evolució temporal del nombre de publicacions de Bad Gyal per mes a YouTube. Font: Elaboració pròpia.

En aquesta evolució temporal es poden veure tots els vídeos que ha publicat l'artista en ordre cronològic per cada mes. L'inici de les publicacions és l'onze d'abril de 2016, amb el vídeo de PAI, i l'última publicació és Chulo el dia 10 de febrer. L'acumulació acostuma a ser d'entre 0 i 2 vídeos al mes, però hi ha pics que destaquen superant molt aquest nombre. Per veure-ho més concretament s'ha analitzat les publicacions dia a dia, de manera que es mostra si la publicació ha estat simultània o si s'ha publicat cada vídeo en diferents dies.



Il·lustració 14: Evolució temporal del nombre de publicacions de Bad Gyal per dia a YouTube. Font: Elaboració pròpia.

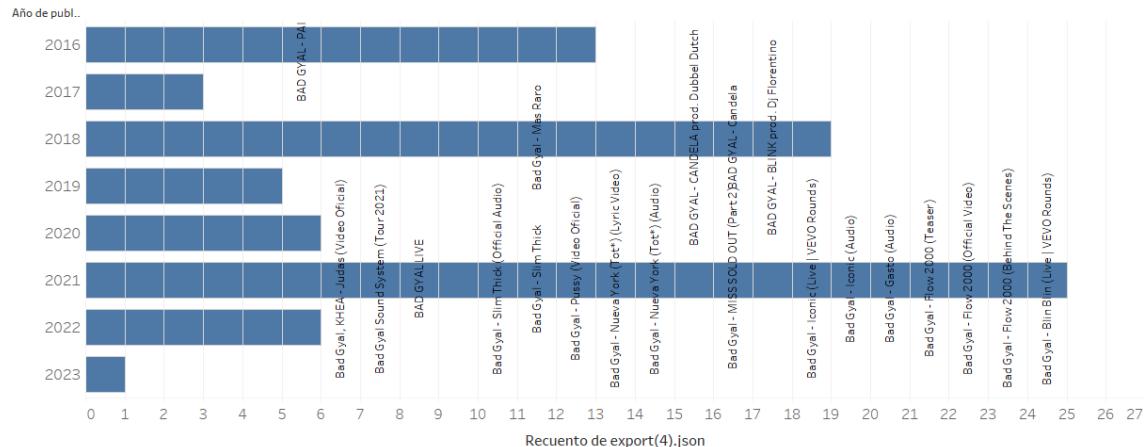
Veient els títols de cada vídeo, es pot observar que quan hi ha més d'una publicació al dia es deu a uns motius concrets i que no és casual. El primer pic és al mes de novembre de 2016. Concretament, les publicacions corresponen als dies 6, 7 i 8 de novembre de 2016, moment en el qual es publiquen els vídeos del primer àlbum, Slow Wine Mixtape.

Al segon pic, durant el mes d'abril de 2018, hi ha la publicació del segon àlbum de l'artista, anomenat Worlwide Angel. En aquest cas, l'artista va decidir publicar els vídeos tots a la vegada el dia 24 d'abril.

El tercer pic, al març del 2021, es correspon a la publicació de les cançons restants del seu tercer àlbum. En els mesos anteriors s'havia començat a publicar els vídeos de l'àlbum, i finalment el 18 de març es van publicar les quatre cançons que faltaven per tenir al complet Warm Up.

Finalment, trobem altres dies en els quals hi ha hagut tres tipus de publicacions: diferents formats, esdeveniments i *teaser*. Per exemple, el 9 de desembre de 2021 es publiquen dos vídeos de la cançó Slim Thick, un és el videoclip i l'altre és el *lyric video*. Com a exemple dels esdeveniments es pot observar que el dia 15 de setembre es va fer un esdeveniment anomenat VEVO Rounds, del qual es van publicar les actuacions en directe de Bad Gyal en dos vídeos. Com a exemple de publicació d'un *teaser* o

*trailer i el llançament de la cançó en el mateix dia, trobem Flow 2000, llançada el 14 d'octubre de 2021.*

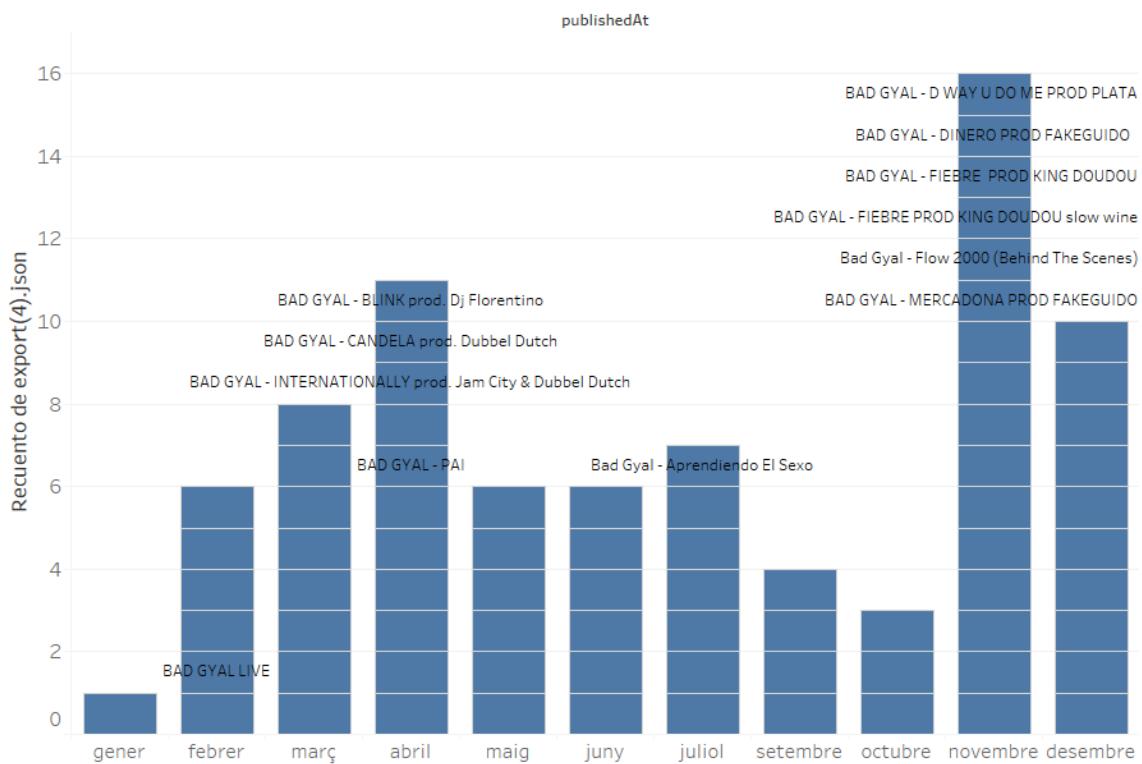


Il·lustració 15: Evolució temporal del nombre de publicacions de Bad Gyal per any a YouTube. Font: Elaboració pròpia.

Vinculat amb les dades obtingudes de Spotify a il·lustració 5, els anys en els quals es publiquen els tres àlbums, són els anys en els quals es publiquen també més cançons. En ordre descendent de suma de cançons els anys són 2021, 2018 i 2016. Tot i que 2021 destaca molt per sobre dels anys anteriors i posteriors, no tot el contingut de vídeos es deuen a l'últim àlbum publicat, ja que només té 8 cançons, i durant el 2021 es van publicar 25 vídeos.

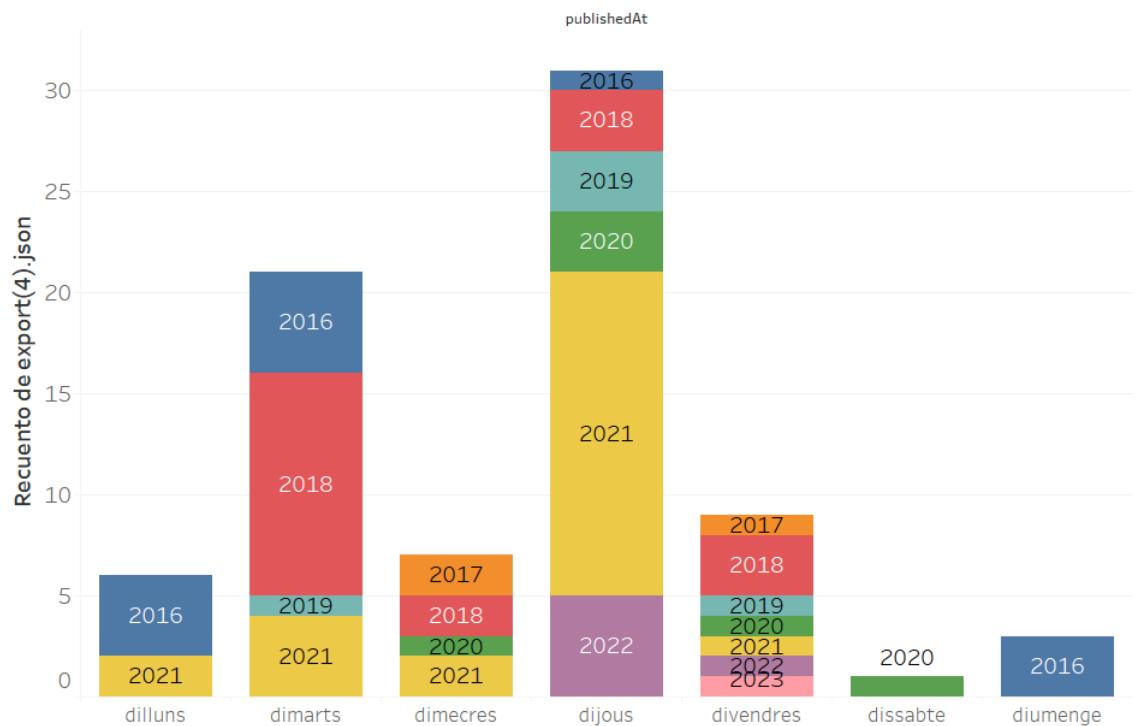
Al 2016 hi ha 13 vídeos publicats, mentre que l'àlbum té 7 cançons. I al 2018 hi ha 19 vídeos publicats mentre que l'àlbum té 9 cançons.

### Anàlisi del moment de publicació



II·Il·lustració 16: Nombre de publicacions de Bad Gyal per mes a YouTube. Font: Elaboració pròpia.

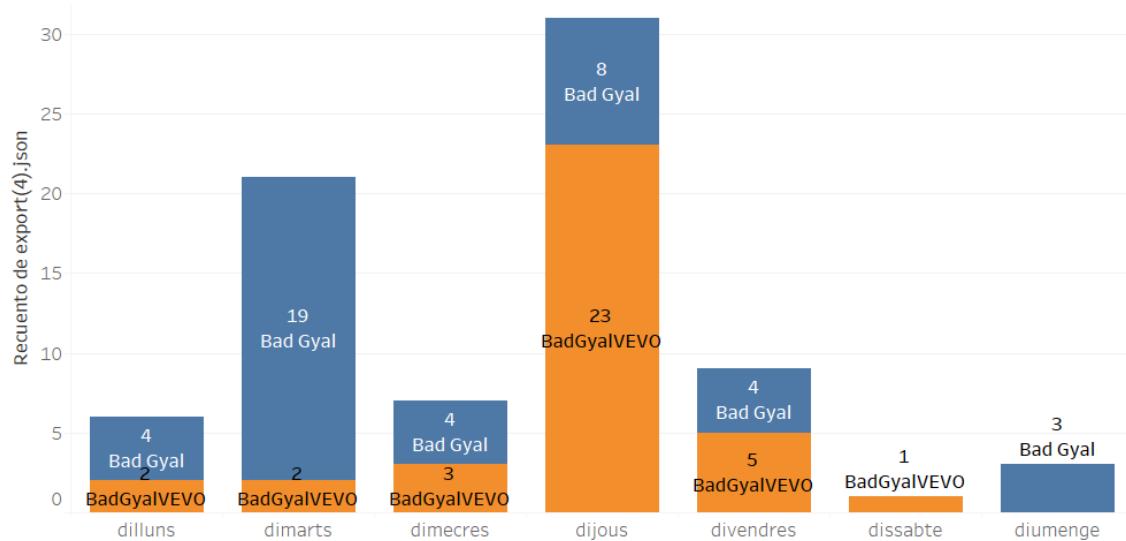
Segons el mes de l'any, hi ha períodes d'especial augment en la quantitat de publicacions. Abans de la Setmana Santa (abril) i abans de Nadal (desembre) hi ha pics de publicacions. El primer pic està més repartit entre els mesos de febrer, març i abril (25 publicacions en total), mentre que el segon pic es concentra molt més entre els mesos de novembre i desembre (26 publicacions). De totes maneres, individualment, la publicació de vídeos a cada mes està bastant igualada, excepte al gener (1 únic vídeo publicat), octubre (3 vídeos) i novembre (16 vídeos).



Il·lustració 17: Nombre de publicacions de Bad Gyal per dia de la setmana i any a YouTube. Font: Elaboració pròpia.

Analitzant els dies de publicació, es pot observar que el dia en el qual es publica més vídeos és el dijous (31), seguit del dimarts (21), i que el dia en el qual es publica menys és dissabte (1) seguit del diumenge (3). La concentració del dimarts es deu a la publicació de l'àlbum Worldwide Angel de manera simultània. Sense comptar aquestes cançons, hi ha una quantitat de publicacions habituals per ser un dia d'inici de setmana.

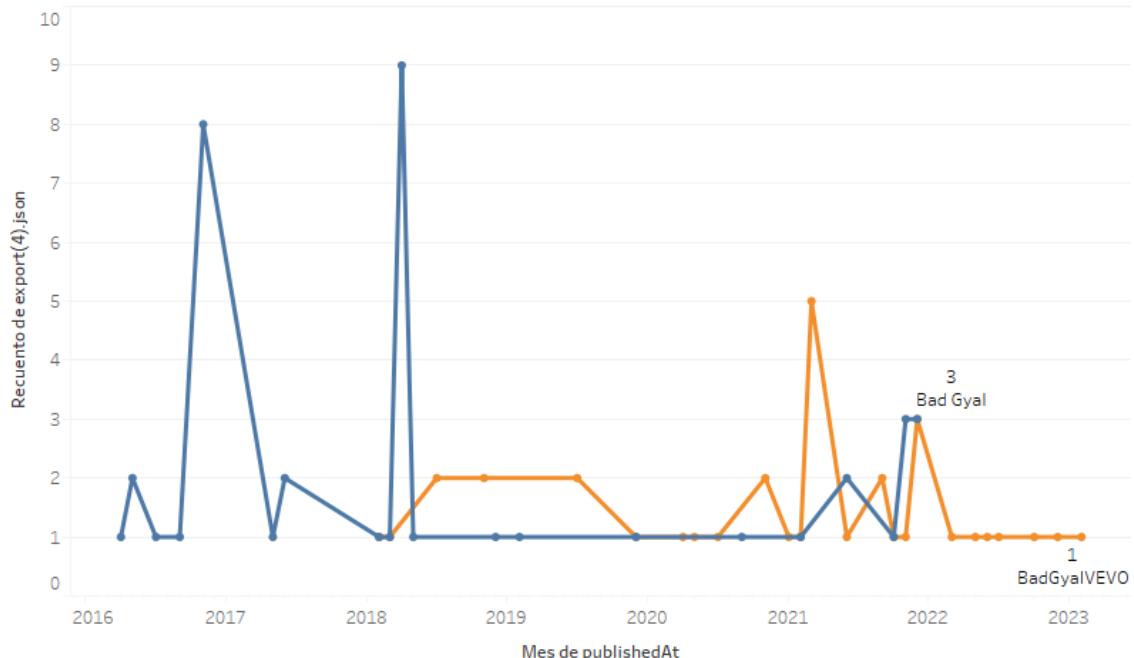
El segon pic que es visualitza és el més important. Els dijous hi ha la màxima concentració de publicacions, ja que arriba fins a un recompte de 31 vídeos. Aquesta concentració es deu al fet que el cap de setmana els vídeos han d'estar accessibles, inclos el divendres, ja que és el dia que s'actualitzen les llistes de novetats musicals. Si ens fixem en la quantitat que representa cada any, hi ha un augment significatiu de les publicacions en dijous en els últims anys, mentre que les publicacions dels primers vídeos estan més repartits al llarg de tota la setmana sense que es vegi un patró clar.



Il·lustració 18: Nombre de publicacions de Bad Gyal per mes i canal a YouTube. Font: Elaboració pròpia.

A la vegada, també es pot observar a la il·lustració 18, que la majoria de publicacions de cançons en dijous es deuen al canal BadGyalVEVO (en color taronja), i no al canal original de l'artista (en color blau). VEVO és “la xarxa de vídeos musicals líder del món, connectant una audiència global en constant creixement amb contingut de vídeo musical d'alta qualitat durant més d'una dècada. [...] Des de grans estrelles fins a nous talents en ascens, Vevo ofereix un suport de promoció interconnectada incomparable a artistes de tot l'espectre musical, en cada etapa de la seva carrera.” (VEVO, 2023).

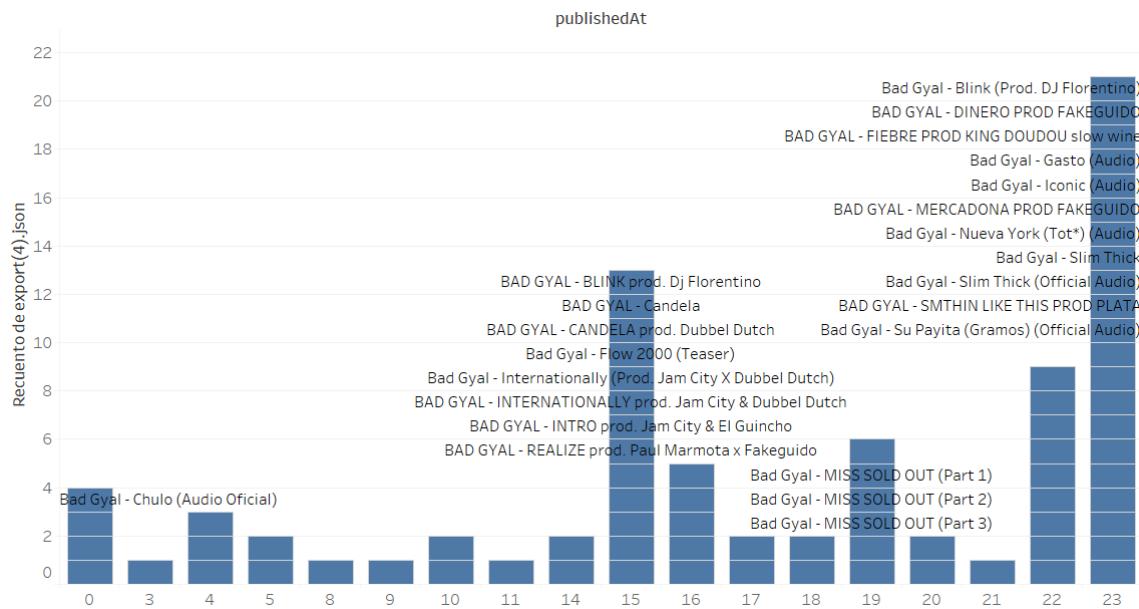
### Anàlisi per canal de publicació



Il·lustració 19: Evolució temporal de publicacions per canal a YouTube. Font: Elaboració pròpia.

Tant el canvi en el dia de publicació, com la publicació dels vídeos a través de Vevo, es deuen a una professionalització del contingut en els últims anys. El 2018, tal com es veu en la il·lustració 19, Bad Gyal comença a treballar amb Vevo per publicar els seus vídeos de cançons. I aquesta plataforma cada vegada agafa més pes, fins al punt en què només es publica contingut secundari des del canal inicial de l'artista i totes les cançons es publiquen exclusivament a Vevo. Fins i tot, des de desembre de 2021, només hi ha hagut publicacions en el canal de Vevo. Tot i que durant els últims anys ja s'anava observant aquest canvi.

Cal destacar que aquests dos canals són els propietaris dels vídeos de manera interna per a YouTube, però des del canal inicial de Bad Gyal apareixen tots els vídeos publicats de manera conjunta, indiferentment del propietari. Cal que l'usuari es fixi en el nom del canal una vegada ha obert el vídeo per poder-se adonar del propietari.

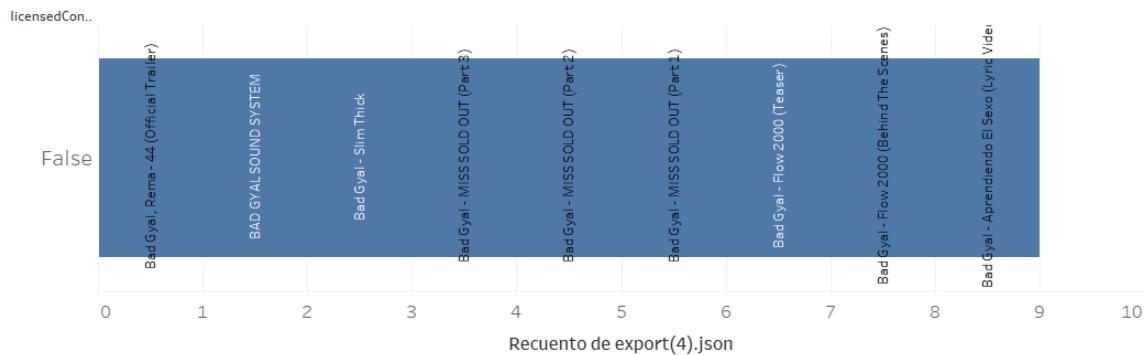


II·lustració 20: Nombre de publicacions de Bad Gyal per hora del dia a YouTube. Font: Elaboració pròpia.

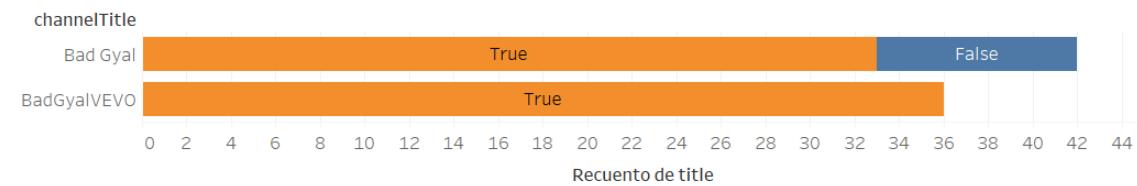
Quant a l'hora de publicació, la majoria es concentra a la nit, des de les 22h, però especialment a les 23h. Els primers vídeos del canal tenien menys planificació, perquè es publicaven a qualsevol hora i minut. En canvi, els últims llançaments es pot observar que han estat programats perquè es publiquin a l'hora en punt, és a dir, a les 22:00h o 23:00h.

També trobem que a les 15h hi ha un pic de publicacions. Això es deu a la publicació dels vídeos de l'àlbum Worldwide Angel que representen 9 dels 13 vídeos publicats a les 15h. Aquestes publicacions també van ser preparades amb antelació perquè es van publicar els nou vídeos al mateix minut.

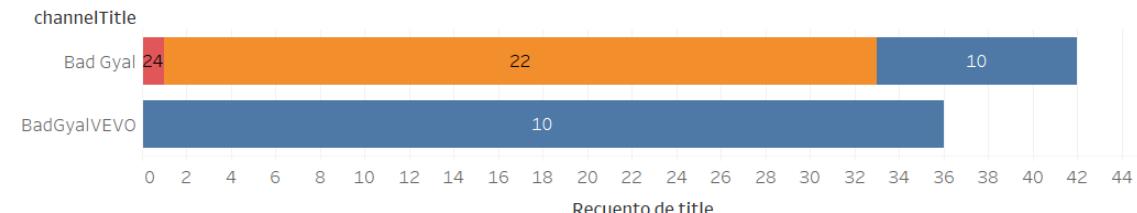
Finalment, respecte a les hores del dia en les quals es publica, caldria destacar que les cançons més noves es concentren en el pic de la nit, i que el contingut extra, com els directes o els reportatges, són publicats durant la tarda i el vespre.



II·lustració 21: Vídeos sense llicència. Font: Elaboració pròpia.



II·lustració 22: Llicència per canal. Font: Elaboració pròpia.



II·lustració 23: Categoria per canal. Font: Elaboració pròpia.

Seguint amb la professionalització del contingut i l'augment del pes del canal Bad Gyal Vevo, s'observa que tot el contingut d'aquest canal està sota llicència (II·lustració 22) i classificat correctament com a “música” a la categoria 10 (II·lustració 23). En canvi, en el canal principal (Bad Gyal), el contingut no està tot sota llicència ni està correctament classificat amb les categories.

És cert que el contingut del canal principal conté vídeos secundaris com *teasers* o resums d'esdeveniments que no necessiten tanta protecció com aquells vídeos musicals. Però, com es pot veure en l'ampliació (II·lustració 21), hi ha dues cançons que no estan protegides amb *licensed*: Bad Gyal - Aprendiendo El Sexo (Lyric Video) i Bad Gyal - Slim Thick. Pel que fa a la classificació dels continguts, gairebé tots estan classificats amb la categoria 22 (People & Blogs), quan la majoria del contingut es

tracta de vídeos musicals, i per això s'haurien de classificar com a 10 (Music) (Mixed Analytics, 2023).

En relació amb el tipus de vídeo, Bad Gyal no ha aprofitat les noves eines que proporciona YouTube per crear altres tipus de continguts com directes, vídeos en vertical i vídeos curts. Tots els vídeos, independentment del canal propietari, són en format horitzontal i gravats prèviament a excepció d'un directe.

### 5.3.5. Anàlisi de les interaccions de Bad Gyal a YouTube

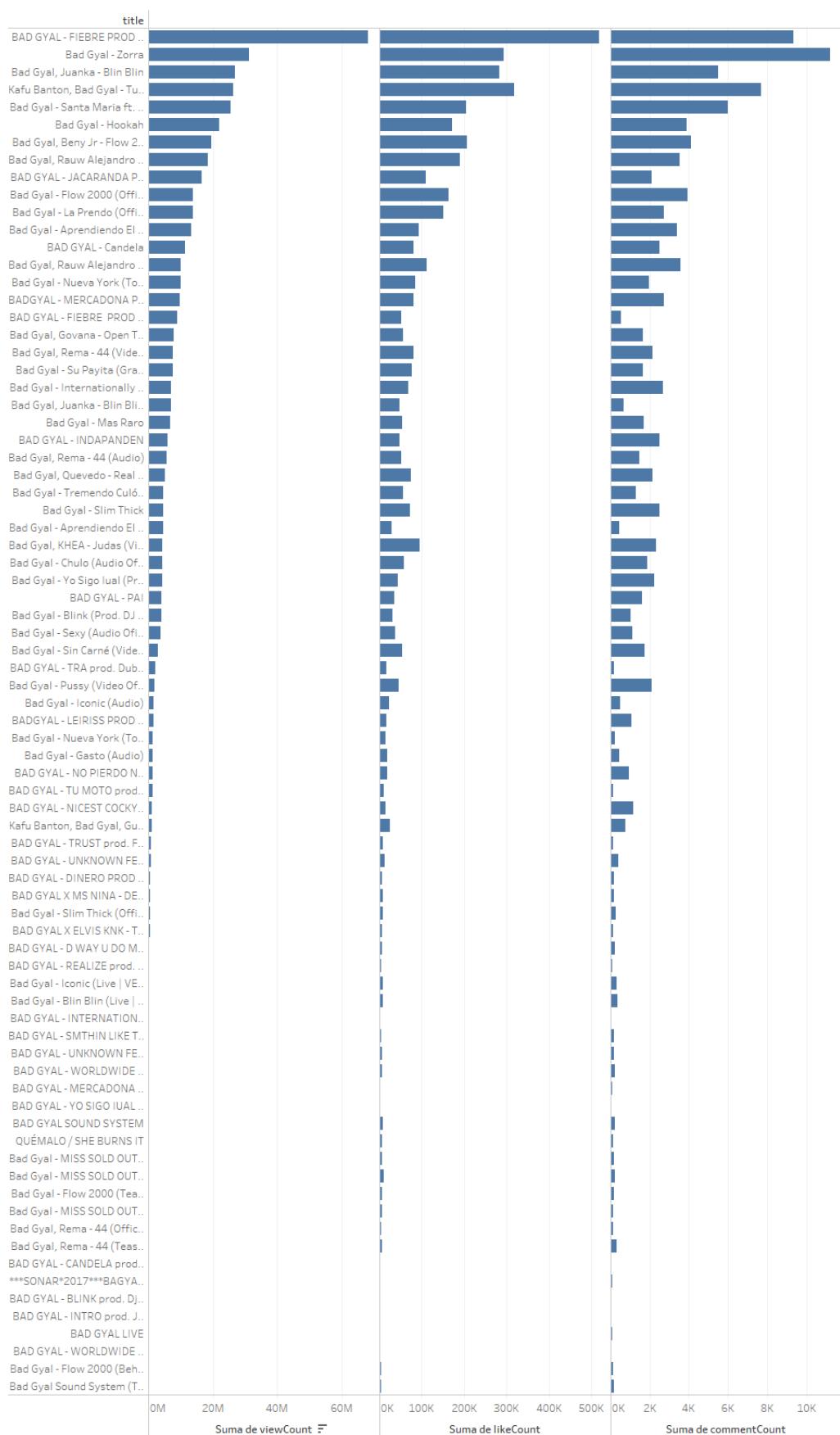
La mesura de les interaccions de Bad Gyal és molt important perquè suposa la visualització de la resposta dels usuaris, mostrant si ha agrat o no un contingut de manera concreta i quantificable perquè mostra en quina mesura la resta d'usuaris també hi estan d'acord. A YouTube hi ha tres mètriques públiques relacionades amb la interacció o *engagement*: les visualitzacions (*view*), els m'agrades (*like*) i els comentaris (*comment*).

En relació amb aquestes mètriques en nivells generals, Bad Gyal es troba de la següent manera.

Taula 3: Màxim, mínim i valor mitjà de view, like i comment de Bad Gyal a YouTube. Font: Elaboració pròpia.

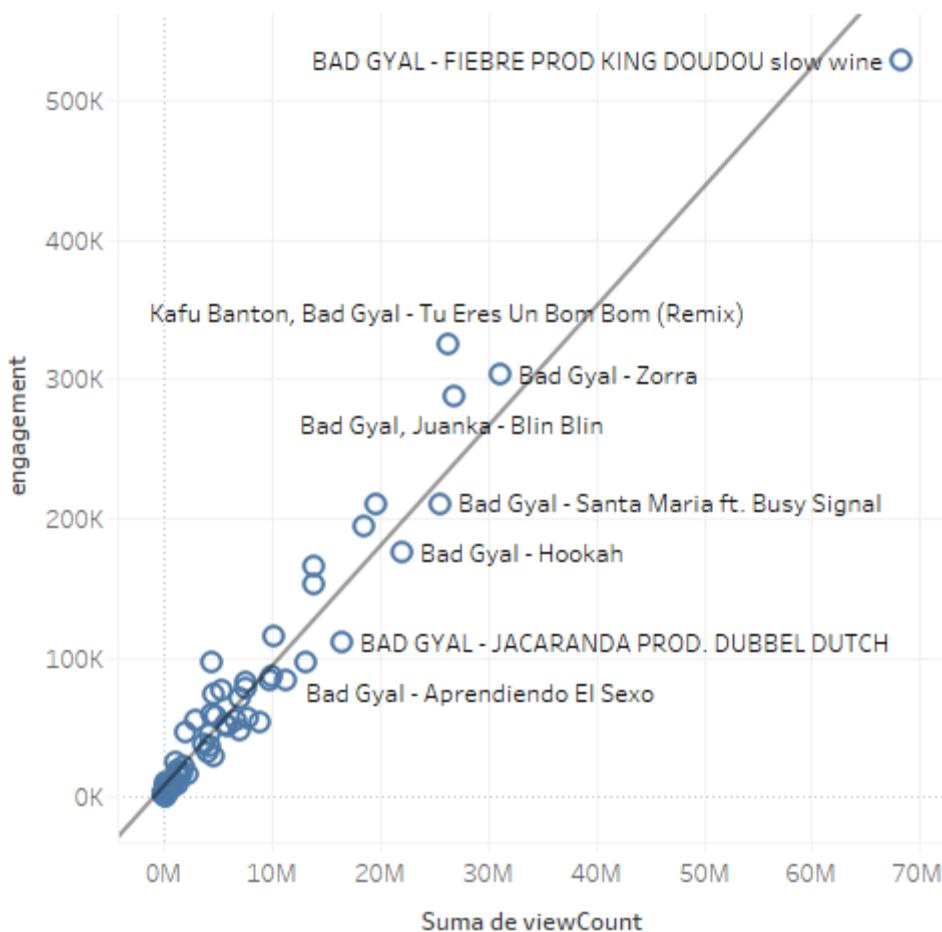
|              | <b>View</b>       | <b>Like</b>    | <b>Comment</b>   |
|--------------|-------------------|----------------|------------------|
| <b>Màxim</b> | <b>68.192.939</b> | <b>518.779</b> | <b>11.249</b>    |
|              | BAD GYAL -        | BAD GYAL -     | Bad Gyal – Zorra |
|              | FIEBRE PROD       | FIEBRE PROD    |                  |
|              | KING DOUDOU       | KING DOUDOU    |                  |
|              | slow wine         | slow wine      |                  |
| <b>Mínim</b> | <b>42.238</b>     | <b>1.125</b>   | <b>20</b>        |
|              | Bad Gyal Sound    | BAD GYAL -     | BAD GYAL -       |
|              | System (Tour      | BLINK prod. Dj | BLINK prod. Dj   |
|              | 2021)             | Florentino     | Florentino       |
| <b>Mitja</b> | <b>6.047.883</b>  | <b>58.187</b>  | <b>1.541</b>     |

Més concretament, per poder visualitzar fàcilment tots els continguts s'ha fet la següent comparativa entre les tres mètriques. És interessant observar que en la majoria de vídeos, va en consonància el nombre de visites, m'agrada i comentaris. Però no sempre és així. A continuació, en la il·lustració estan ordenats els valors pel nombre de visualitzacions i apareixen tots els vídeos de Bad Gyal.



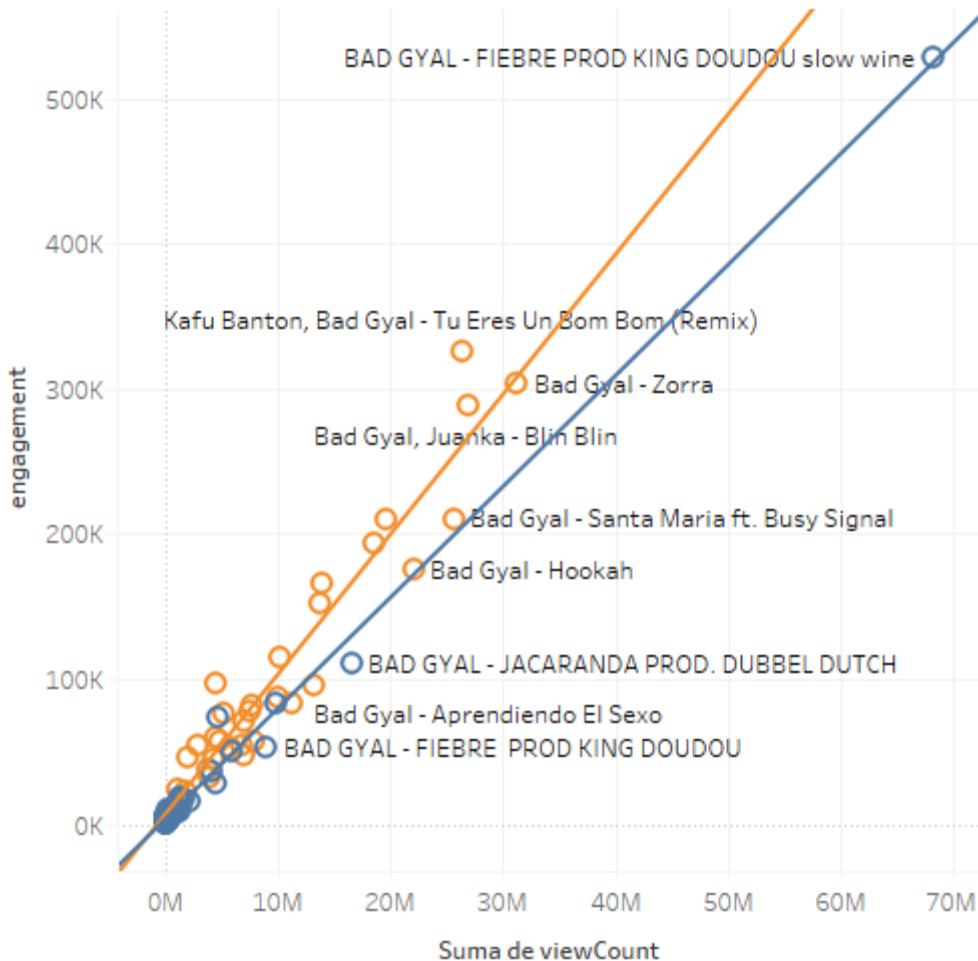
Il·lustració 24: Comparativa del nombre de views, like i comment dels vídeos de Bad Gyal a YouTube. Font: Elaboració pròpia.

El gràfic anterior ens permet comparar els valors màxims i mínims, així com relacionar les mètriques aplicades a l'artista dins de la plataforma YouTube. Però per poder veure millor com afecta individualment a cada vídeo i com afecta les mètriques, s'ha creat un gràfic de dispersió que compara aquelles mesures, mostrant quins vídeos es desvien més de la tendència general. Per generar la taula s'han sumat les interaccions (*engagement*), i a l'altre eix s'ha situat el nombre d'impressions que ha tingut el contingut. El nombre d'interaccions és la suma de *likes* i *comment* per a cada publicació. La línia diagonal representa la mitjana dels vídeos de Bad Gyal.



Il·lustració 25: Nombre d'interaccions respecte de les visualitzacions per vídeo de Bad Gyal a YouTube. Font: Elaboració pròpia.

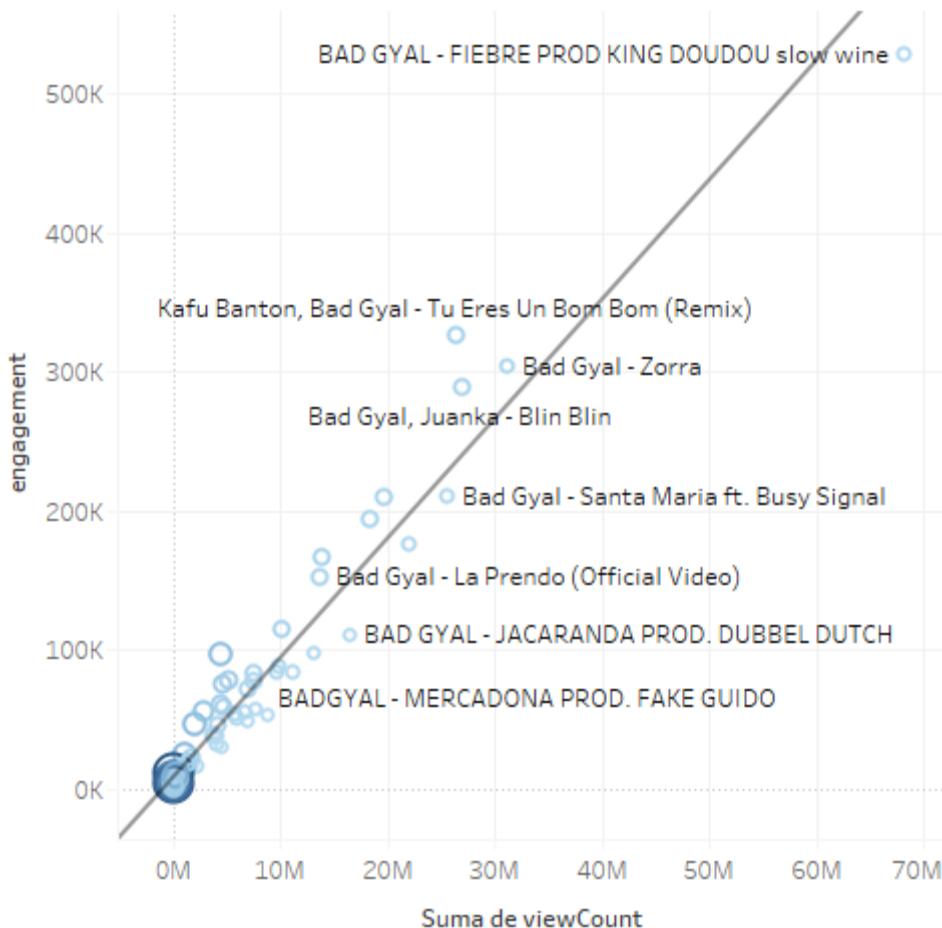
En general, es pot observar una concentració majoritària per sota de les 100.000 interaccions i dels 10 milions de visualitzacions, i després es troba una minoria que n'ha tingut més. Principalment, podem veure que la majoria de vídeos segueixen la mateixa tendència i que els vídeos més vistos també són aquells que tenen més *engagement*. Els vídeos més rellevants no separen gaire de la línia de tendència.



Il·lustració 26: Nombre d'interaccions respecte de les visualitzacions per vídeo i canal de Bad Gyal a YouTube.  
Font: Elaboració pròpria.

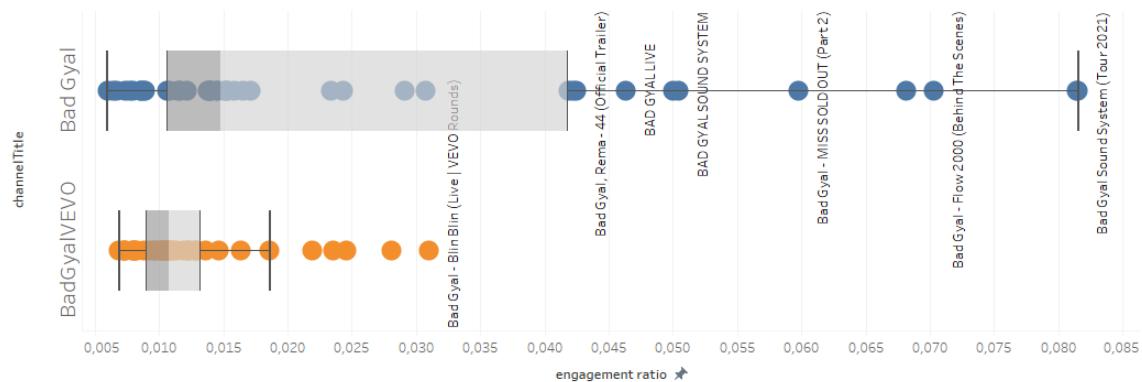
Si separem el contingut per canals, es pot observar que el canal BadGyalVeo (en taronja) té un nombre més elevat d'interaccions en termes absoluts, i que el canal original obté el millor resultat en visualitzacions. També cal destacar que els nombres màxims de cada canal són molt diferents, el canal Veo obté uns nombres d'*engagement* i de visualització molt més grans que el canal original, a excepció d'un vídeo, el videoclip de Fiebre (slow wine). Aquest vídeo té les millor mètriques en

termes absoluts, és el vídeo més vist i també amb més likes, tot i que en nombre de comentaris és el segon.



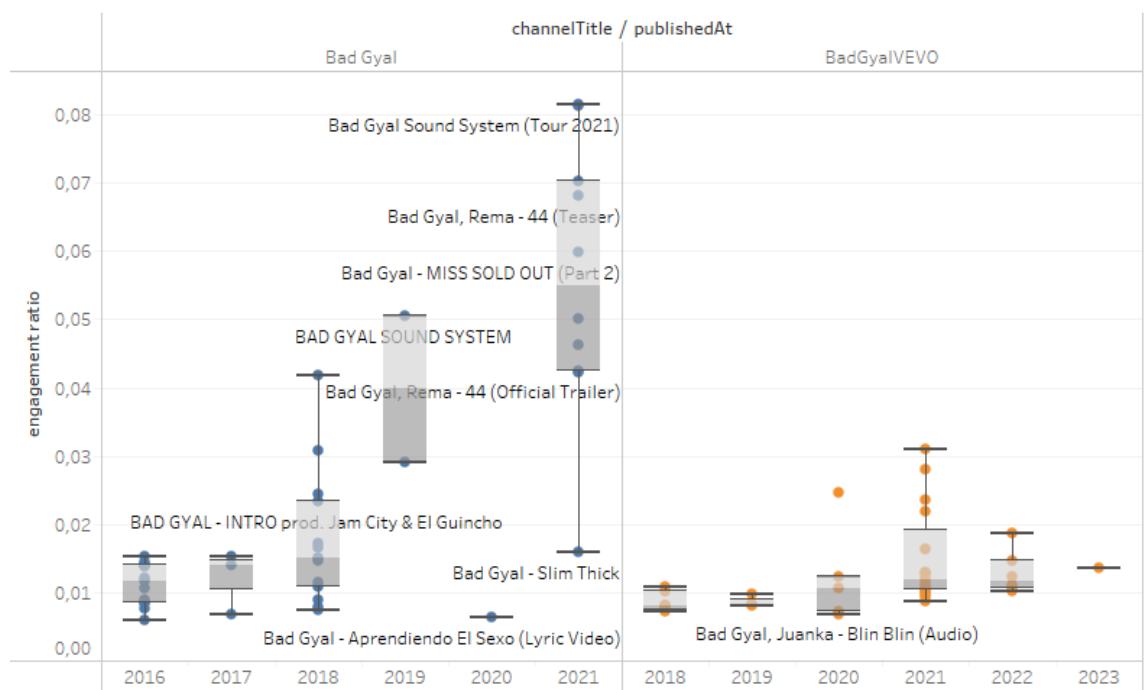
II·Il·lustració 27: Ràtio d'interacció respecte de les visualitzacions per vídeo de Bad Gyal a YouTube. Font: Elaboració pròpia.

Però cal tenir en compte el nombre de vídeos que suposa cada franja, ja que la majoria de vídeos tenen unes xifres d'interaccions i visualitzacions molt inferiors. Si tenim en compte el nombre d'interaccions respecte del nombre total de visualitzacions del vídeo (II·Il·lustració 27), obtenim l'*engagement rate*, un valor que ens mostra el percentatge de manera més realista. D'aquesta manera, trobem que els vídeos que tenen una major ràtio d'interacció (representada amb la mida i la intensitat del color), són aquells que ha vist menys gent. Aquest és un fet que acostuma a passar a les plataformes, ja que com més petita sigui la comunitat més s'interacciona perquè hi ha més sentiment de pertinença.



II·lustració 28: Engagement rate per vídeo i canal de Bad Gyal a YouTube. Font: Elaboració pròpia.

Per aquest motiu, si observem la ràtio d'interacció respecte a les visualitzacions, el canal original de l'artista és molt superior que el de Vevo. Això també es deu a altres fets, i és que al tenir una comunitat tan gran com en l'actualitat hi ha molts fans que demanen contingut secundari, dirigit molt més als fans més adeptes. Aquest contingut es puja exclusivament al canal original, com es pot observar en el títol dels vídeos que tenen una ràtio més elevada per canal, però especialment en la il·lustració 29, que mostra l'evolució dels dos canals per any.



II·lustració 29: Evolució temporal de la interacció per any, vídeo i canal a YouTube. Font: Elaboració pròpia.

## 5.3. Twitter

Twitter és una de les plataformes més populars a Espanya. Destaca per la capacitat d'adaptar-se a l'actualitat i per formar-se espais de crítica i discussió de temes molt variats. És una xarxa social on la gent comparteix moltes opinions públicament i es generen moltes interaccions entre aquestes. Per aquest motiu, és una font d'informació valuosa, ja que ajuda a saber com es desenvolupa la conversa al voltant de Bad Gyal.

Històricament, Twitter ha sigut una xarxa social molt oberta en el sentit que la informació és pública i fàcilment accessible. Però amb els últims canvis que s'han aplicat pel canvi d'equip directiu de l'empresa, s'estan perdent algunes d'aquestes característiques. Per exemple, s'han posat més murs de pagament i altres traves per accedir a la informació de manera massiva des de l'API. És per aquest motiu que no he pogut accedir a l'API personalment. Però s'ha pogut obtenir les dades de tots els *tweets* en els quals apareix Bad Gyal en català i en espanyol a partir d'una altra persona amb un compte de pagament. Per aquest motiu, s'analitzarà directament les dades obtingudes, fent la visualització d'aquestes amb Tableau.

Els tuits s'han obtingut a través de les dues consultes a l'API que s'exposen a continuació.

“Bad Gyal lang:es -is:retweet”

“Bad Gyal lang:ca -is:retweet”

L'objectiu de treballar amb dades de Twitter és obtenir informació sobre la conversa que es forma al voltant de Bad Gyal. Més concretament, aquesta font de dades ens permet observar quan i com evoluciona la conversa; quina mena de contingut es crea; com i quin valor té cada interacció; el tipus d'opinió que s'exposa i quines relacions es generen entre els usuaris.

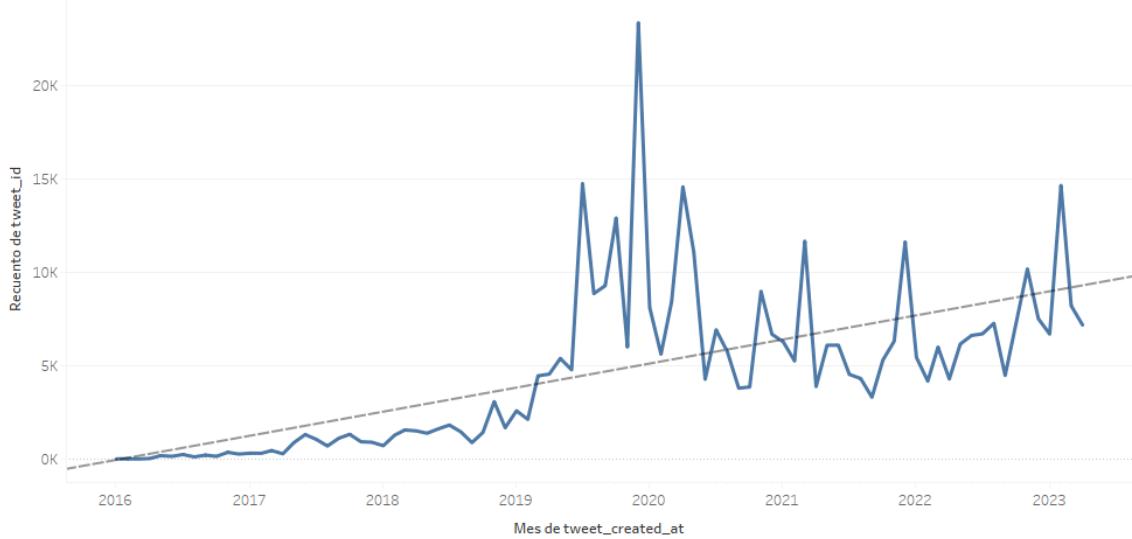
A continuació es farà una anàlisi de cada punt anterior per observar l'evolució temporal, les mètriques dels *tweets*, el contingut i la interrelació d'usuaris.

### 5.3.1. Anàlisi de l'evolució temporal a Twitter

Per entendre l'entorn de Bad Gyal és molt important entendre quines publicacions i altres accions ha dut a terme l'artista al llarg de la seva carrera. Per això, també cal comprovar la seva evolució en aquesta plataforma.

#### *Evolució temporal del nombre de tweets*

Per aconseguir-ho, primer de tot cal obtenir una visió global de les dades recollides. El primer tuit recollit és del març de 2010, època en la qual l'artista encara no havia iniciat la seva carrera musical, ja que tenia 13 anys. I l'última piulada recollida és del dia 1 de maig a les 23:40h. Per poder obtenir una visió més realista sobre els *tweets* que realment estan parlant sobre Bad Gyal, cal filtrar les dades. S'han agafat exclusivament els tuits a partir de gener de 2016, l'any que comença a fer les seves primeres publicacions musicals. D'altra banda, per obtenir una visió més real de les dades, també s'ha limitat els *tweets* fins al dia 30 d'abril de 2023 a les 23:59:59.



II·Il·lustració 30: Evolució temporal de nombre de tweets per mes i tendència (2016-2023). Font: Elaboració pròpia.

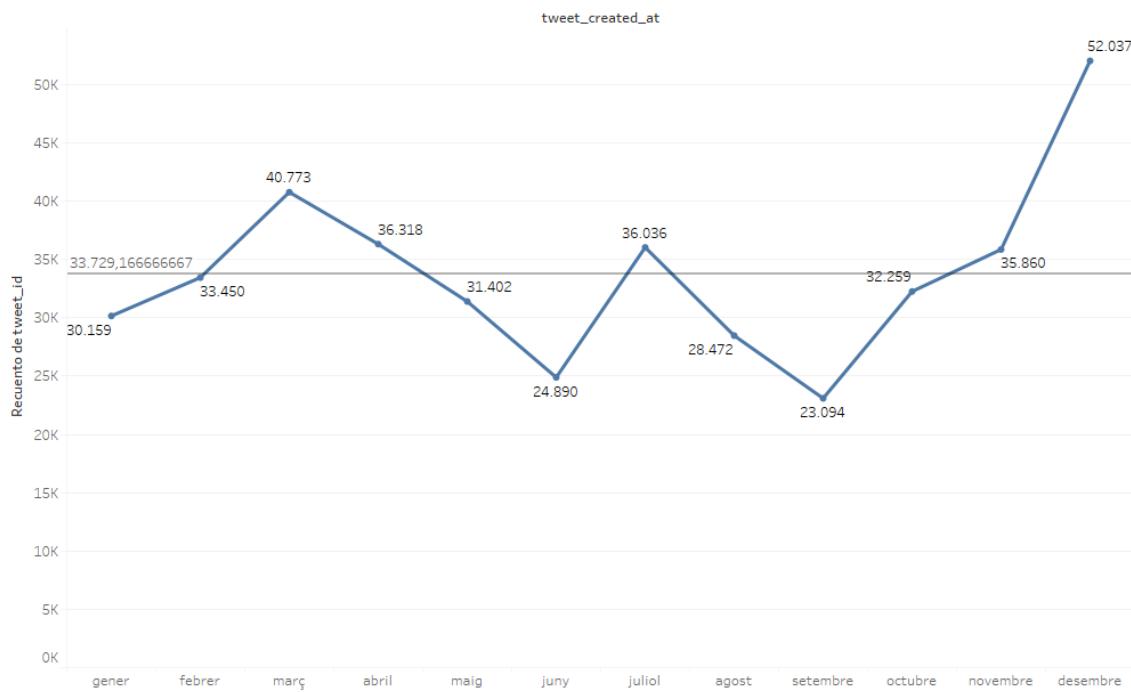
A partir dels filtres anteriors i de la suma de tuit per cada mes s'ha obtingut la il·lustració 30. Es pot observar una tendència a l'alça del nombre de *tweets*, així com una acumulació de pics en la part central, i una consolidació del nombre de publicacions posteriors.

Entre juliol de 2019 i abril de 2020 es produeix aquest període d'especial acumulació de piulades. El primer pic, al juliol de 2019 es produeix en consonància del llançament de dues cançons, Santa Maria i Hookah. El segon pic el trobem a octubre de 2019, moment en el qual comparteix la seva col·laboració amb Omar Montes, Alocao. Però el gran pic el trobem al desembre de 2019, moment en el qual aquesta última cançó es troba en les posicions més altes de les llistes musicals. El dia 10 de desembre es fa trending tòpic a Twitter, i la plataforma social s'omple de fans, però especialment *haters* que parlen d'ella. També coincideix al desembre, el concert a Razzmatazz i publicació de la seva gira “Bad Gyal Sound System”. Per bo o per dolent, aquest període va canviar completament la trajectòria de l'artista, entrant al *mainstream* d'Espanya i de la música en urbana mundial.

La resta de pics que apareixen més endavant es deuen a llançaments musicals i destaquen al març de 2021, amb la publicació del seu tercer àlbum, i al desembre de 2021 amb la publicació de les cançons restants de la gira. Cal destacar l'últim pic, al febrer de 2023, moment en el qual publica Chulo. Però aquest pic, que coincideix amb Chulo, no es deu realment a aquesta cançó, sinó al concert de Barcelona de la seva nova gira “La Joia”. Principalment hi ha dos motius perquè es creessin tants tweets, i és que el concert va ser al Palau Sant Jordi amb Sold Out, un èxit importantíssim per a qualsevol artista català. I, en segon lloc, l'artista va estrenar noves cançons, així com va anunciar futures col·laboracions i el nom del seu últim àlbum “La Joia”.

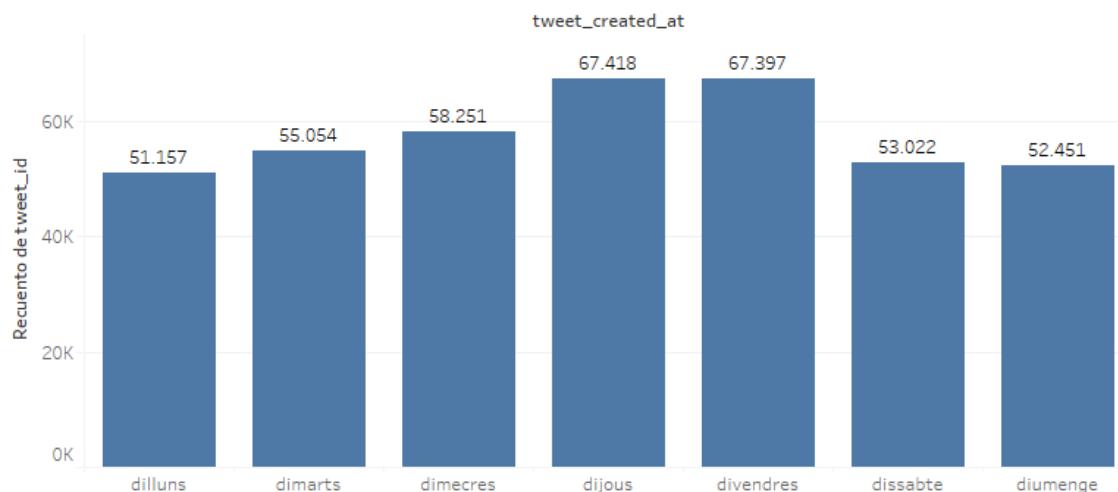
#### *Períodes destacats d'acumulació de publicacions*

De la mateixa manera que en les plataformes anteriors, també cal analitzar en quin moment es fan les publicacions per saber quins són els períodes més destacats. Amb aquest objectiu, s'analitzaran les dades segons el mes de l'any, el dia de la setmana i l'hora del dia.



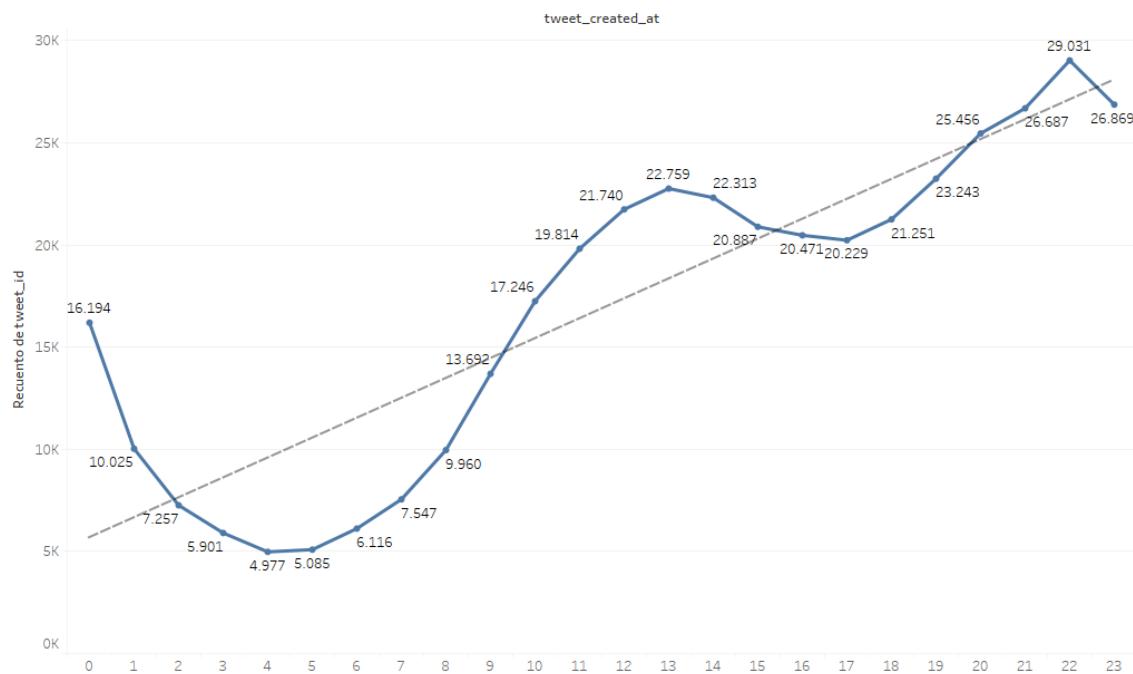
II·Il·lustració 31: Nombre de tweets per mes de l'any. Font: Elaboració pròpia.

En relació amb el mes de l'any, es pot observar a la il·lustració 31, que durant els mesos de març, abril, juliol, novembre i desembre la creació de *tweets* es troba per sobre de la mitjana mensual (33.729 piulades) i la resta de mesos per sota d'aquest valor. Destaca especialment el període de finals i inici d'any, moment en el qual es produeix el pic de tuits al desembre (52.037 *tweets*) i després cau fins als 30.159, sent la davallada més forta que hi ha. I el mes en el qual es creen menys tuits és setembre, amb 23.094 publicacions.



II·lustració 32: Nombre de tweets per dia de la setmana. Font: Elaboració pròpia.

L'acumulació de *tweets* per dia de la setmana no varia gaire i es troba entre els 51.000 i els 68.000 tuits diaris. Com és d'esperar, i en consonància amb els llançaments musicals de l'artista, els dijous i divendres hi ha un nombre més elevat de publicacions. El cap de setmana es manté en el nivell més baix (mínim de 52.451 *tweets* el diumenge), i posteriorment va augmentant el nombre de piulades fins a arribar al dijous (67.418 *tweets*) i al divendres (67.397 *tweets*).



II·lustració 33: Nombre de tweets per hora del dia. Font: Elaboració pròpria.

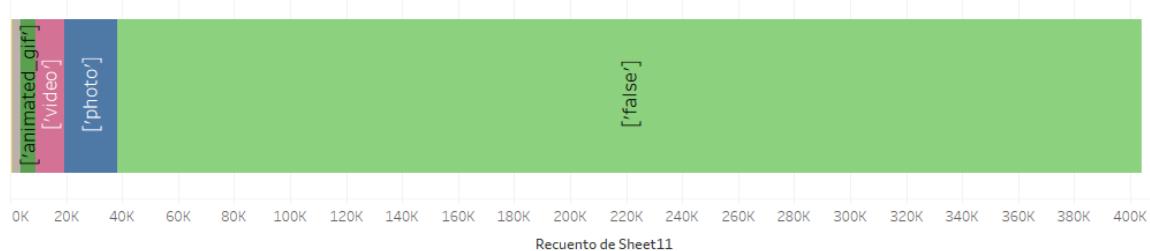
Respecte a les hores del dia en les quals es fan més *tweets*, es pot observar que durant la matinada hi ha el període vall, amb el mínim de 4.977 tuits acumulats, després va creixent aquest nombre fins a arribar al pic màxim de les 22 hores amb 29.031 tuits acumulats. Cal tenir en compte que entre les 13 hores i les 19 hores hi ha un petit període vall d'estancament al voltant dels 20.000 i 23.000 tweets acumulats. Per aquest motiu, el període en el qual es produeix més activitat és abans de dinar (13h) i a la nit (19-23h). Cal destacar que si l'augment de *tweets* durant el dia és progressiu, després de les 23 hores, la xifra cau ràpidament.

### 5.3.2. Anàlisi del tipus de contingut a Twitter

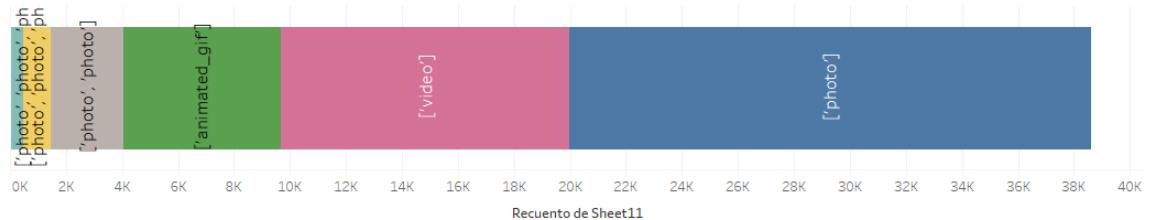
Twitter és una plataforma especialitzada en les publicacions de text, però també ofereix un gran ventall de possibilitats de publicar d'altres maneres, ja sigui per la inclusió de contingut multimèdia complementari o per la direccionalitat del *tweet*.

### Contingut complementari

A Twitter és molt comú trobar contingut multimèdia juntament amb el text de la piulada. Per aquest motiu cal analitzar si aquesta tendència també se segueix en la conversa de Bad Gyal.

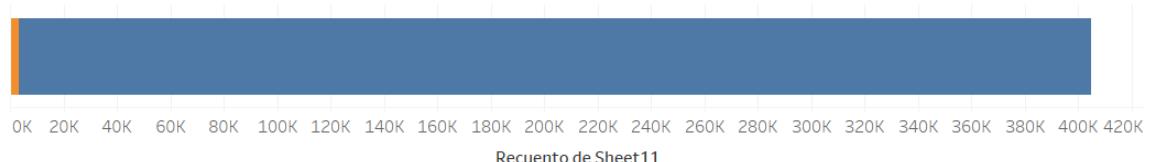


Il·lustració 34: Tweets amb contingut multimèdia. Font: Elaboració pròpia.



Il·lustració 35: Tipus de contingut multimèdia. Font: Elaboració pròpia.

Com es pot observar a la il·lustració 34, la majoria dels tweets extrets no tenen cap contingut multimèdia complementari (part verda). Tot i això, representen una part del total que no es pot deixar de banda, ja que com es veurà més endavant, el contingut complementari no és nombrós, però acostuma a oferir millors resultats d'interacció. Així, en la il·lustració 35 es pot observar que els tipus de contingut multimèdia més comuns són una imatge, un vídeo, un gif, i finalment trobem els tweets amb més d'un fitxer multimèdia, que principalment també són vàries imatges.

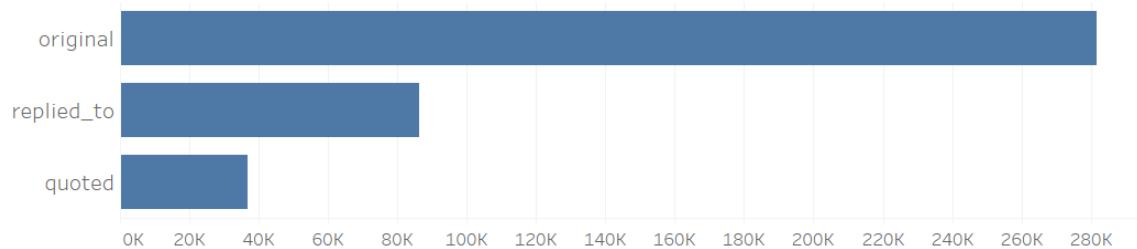


Il·lustració 36: Nombre de tweets amb contingut sensible (taronja) i sense contingut sensible (blau). Font: Elaboració pròpia.

En relació amb el tipus de contingut, aquesta és una xarxa social molt gran en la que es parlen de molts temes, alguns d'ells tenen contingut sensible que en certes ocasions es bloqueja. Aquesta mètrica també facilita la detecció de comunitats radicals i contingut que fomenta el discurs d'odi.

Com es pot observar, la quantitat de *tweets* de Bad Gyal que incorporen contingut sensible és molt petita, inferior al 0,001%.

#### *Direccionalitat de la publicació*



Il·lustració 37: Nombre de tweets per direccionalitat de la publicació. Font: Elaboració pròpia.

És important observar a qui es dirigeixen els tuits en els quals es parla de Bad Gyal, ja que també ens permet saber si els usuaris actuen de manera proactiva o reactiva.

Una de les mètriques més útils per saber això és la naturalesa del *tweet*. Poden ser contingut original o en resposta a un *tweet* anterior, el que pot ser un “*replied\_to*” o un “*quoted*”. Més endavant, també es valoraran les interaccions entre els usuaris representant les interaccions amb un graf.

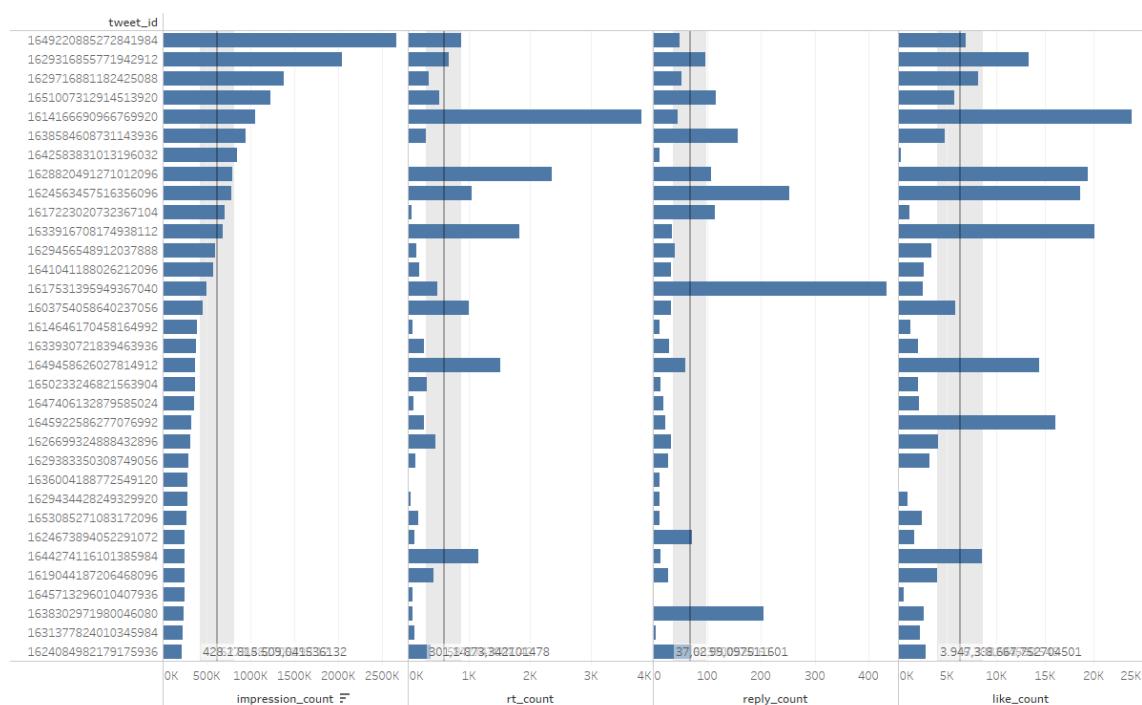
Com es pot observar en la il·lustració 37, la majoria de piulades són originals, però hi ha un nombre molt important que corresponen a respostes. Aproximadament representen un 30% del total i aquesta xifra es divideix entre *replies* i *quotes*. Com és majoritari en aquesta xarxa social, el *reply* requereix menys esforç i té menys importància, per la qual cosa hi ha més quantitat que el *quoted*, que s'utilitza en ocasions més importants.

En qualsevol cas que un 70% dels *tweets* siguin originals i un 30% respostes, vol dir que hi ha una comunitat força interactiva entre si que es contesta als tuits i opina. A

l'apartat final de Twitter es podrà analitzar la direccionalitat i la vinculació d'aquesta comunitat mitjançant un graf dels *tweets* extrets.

### 5.3.3. Anàlisi del tipus d'interaccions a Twitter

Per a cada *tweet* trobem quatre mètriques bàsiques, primer el recompte de les impressions i després les tres possibles interaccions: *like*, *retweet* i *reply*. El nombre d'impressions és el recompte d'usuaris als quals arriba el tuit, i ens permet saber a quantes persones els hi ha aparegut. Mentre que el recompte d'interaccions ens permet saber com s'han sentit aquests usuaris i si els hi ha destacat aquesta publicació.



Il·lustració 38: Comparativa de mètriques per a cada tweet (ordenat per impressions). Font: Elaboració pròpia.

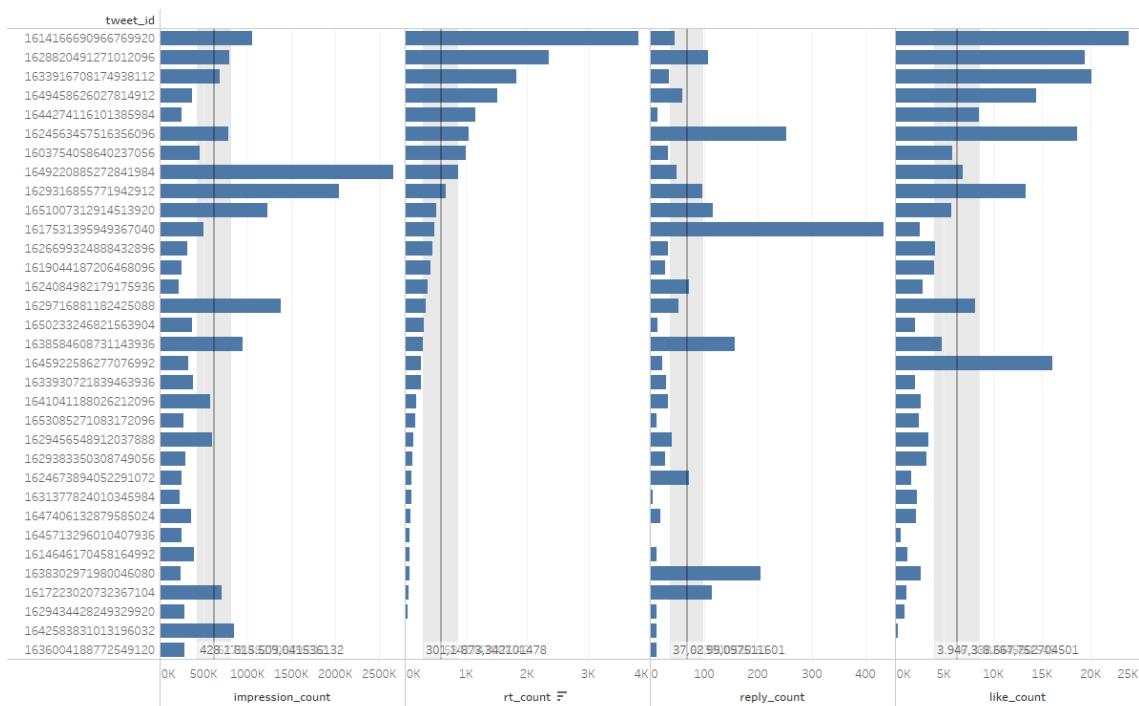
Per comparar aquestes mètriques s'ha creat la taula que es pot veure a il·lustració 38. A primer cop d'ull es pot veure que les mètriques no són proporcionals. És a dir, a diferència del que es podria esperar, els *tweets* amb més impressions no es corresponen amb els que tenen més interaccions. Així, analitzant el nombre

d'interaccions, s'obtenen els següents valors en els que es mouen els tweets que mencionen a Bad Gyal.

Taula 4: Comparació de les mètriques dels principals tweets. Font: Elaboració pròpia.

|              | <b>Impressions</b>  | <b>Like</b>   | <b>Retweet</b>  | <b>Reply</b>  |
|--------------|---|---|---|---|
| <b>Màxim</b> | <b>2.656.225</b>  | <b>45.614</b>                                       | <b>12. 575</b>  | <b>1. 857</b>   |
|              | @BadGyalU<br>pdates: Bad<br>Gyal<br>presentando<br>“Chulo” en los<br>#LatinAMAs.<br><a href="https://t.co/ZWPZiMb23M">https://t.co/ZWPZiMb23M</a> | @kheayf: se<br>viene temita<br>con la bad<br>gyal 😭 | @estresaita:<br>RT A LA BAD<br>GYAL DE LA<br>SUERTE<br>PARA QUE<br>EMPECÉIS<br>2020 A LO<br>GRANDE Y<br>NO OS PASE<br>NADA MALO<br>ESE AÑO<br><a href="https://t.co/gWFJmXsVzy">https://t.co/gWFJmXsVzy</a> | @<br><b>Havana7_ES:</b><br>"¡¡¡SORTEAZ<br>O!!!<br>🔥🔥🔥¿Quier<br>es ir a la<br>fiesta privada<br>de Bad Gyal<br>de este<br>jueves?<br>🔥🔥🔥<br>🔥 Síguenos.<br>🔥 Responde<br>con tu<br>canción<br>favorita,<br>usando el<br>#HAVANACL<br>UBxBADGYA<br>L<br>🔥 Etiqueta a<br>un amigo/a<br>en el tweet |

|                  |                  |                  |                  |  |
|------------------|------------------|------------------|------------------|--|
|                  |                  |                  |                  | Diremos el ganador/a de una entrada doble el jueves a las 12:00. |
| <b>Mínim</b>     | <b>0</b>         | <b>0</b>         | <b>0</b>         | <b>0</b>   |
| Diferents tweets   |
| <b>Mitja</b>     | <b>621.841</b>   | <b>6.307</b>     | <b>587</b>       | <b>68</b>  |



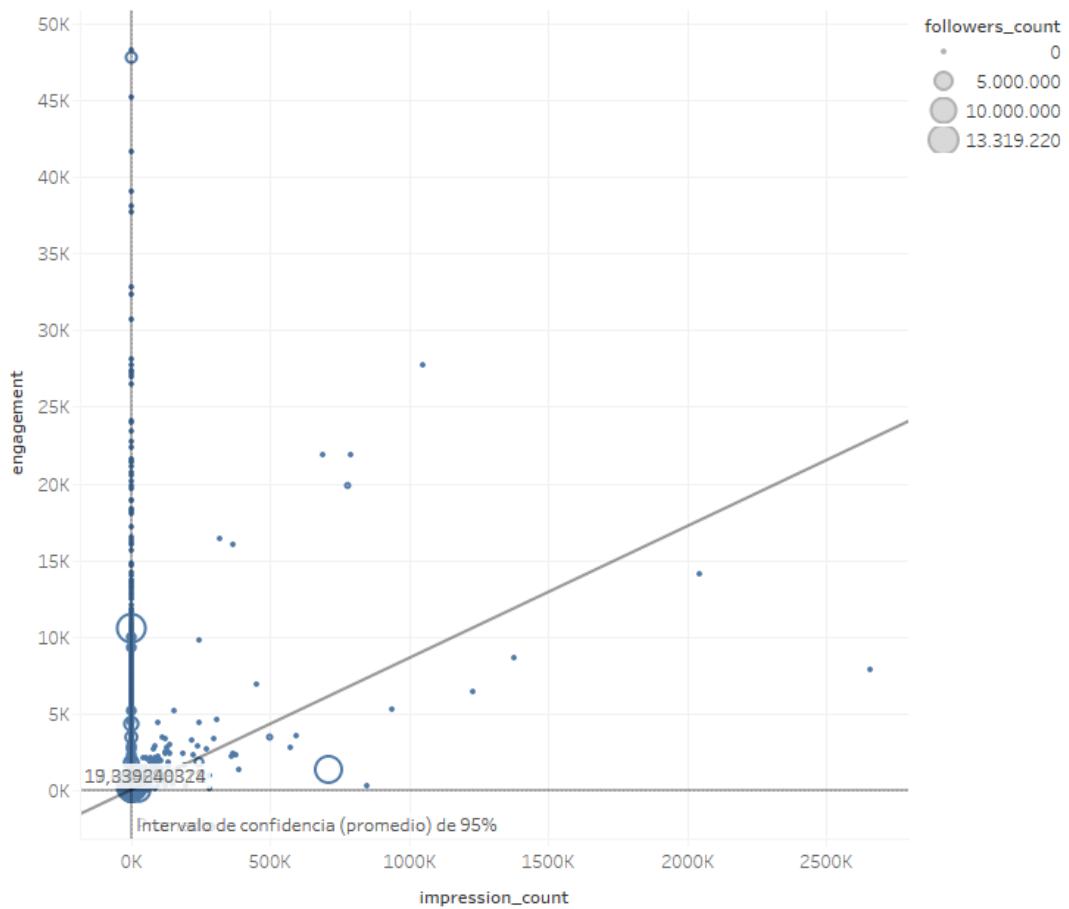
II·lustració 39: Comparativa de mètriques per a cada tweet (ordenat per retweets). Font: Elaboració pròpia.

Ampliant la informació anterior a la il·lustració 39, pel que fa a les reaccions dels usuaris, els *retweets* i els *likes* es corresponen força entre si, però el nombre de *replies*

no es vincula a cap altra mètrica. Aquesta situació sorprèn força, ja que a la resta de plataformes analitzades, aquestes dades són molt més proporcionals.

### 5.3.4. Anàlisi de l'engagement rate a Twitter

Per entendre millor les xifres màximes anteriors, cal comparar les interaccions entre si, però especialment cal mostrar quantes es produeixen respecte de les impressions que tenen. Això s'analitza per saber si el contingut és molt nombrós i s'interacciona, o la xarxa de tweets més densa ocupa la part amb unes dades d'impressions més escasses.



Il·lustració 40: Interaccions respecte del nombre d'impressions i el nombre de seguidors de l'usuari. Font: Elaboració pròpia.

A partir de la visualització de les dades s'ha pogut extreure la mitjana d'interaccions per tweet (19,34), així com la mitjana d'impressions per piulada (174,20). La tendència

que segueixen aquests valors entre un i l'altre, queda representada per la línia diagonal a il·lustració 40.

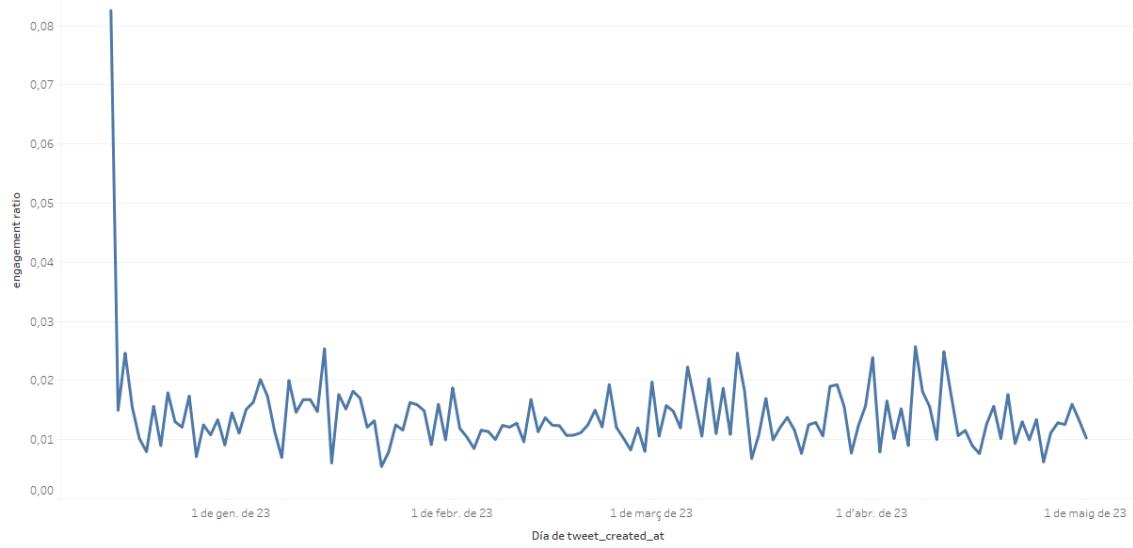
La gran majoria de *tweets* es troben per sota de les 5.000 interaccions i de les 500.000 impressions. Però el valor màxim arriba fins a les 48.308 interaccions i les 2.656.225 impressions. Destaquen punts molt més grans que d'altres perquè tenen un gran nombre de seguidors. És el cas de LaLiga (11M), IbaiLlanos (13M) o comptes de premsa com el\_pais (8M).

#### *Evolució de l'engagement rate*

L'*engagement rate* es tracta del càlcul del percentatge d'interaccions que es fan respecte de les impressions que s'obtenen. Així, a través de la suma de *l'engagement* (*like\_count*, *rt\_count* i *reply\_count*), s'ha obtingut el nombre total d'interaccions. Posteriorment s'ha dividit pel nombre d'impressions per obtenir el valor final. La fórmula utilitzada per calcular els valors s'ha aplicat a Tableau de la manera següent.

$$\frac{\text{SUM}([\text{like\_count}]+[\text{rt\_count}]+[\text{reply\_count}])}{\text{SUM}([\text{impression\_count}])}$$

D'aquesta manera s'ha obtingut un valor per a cada publicació, que permet valorar de forma molt més realista la quantitat d'interaccions.

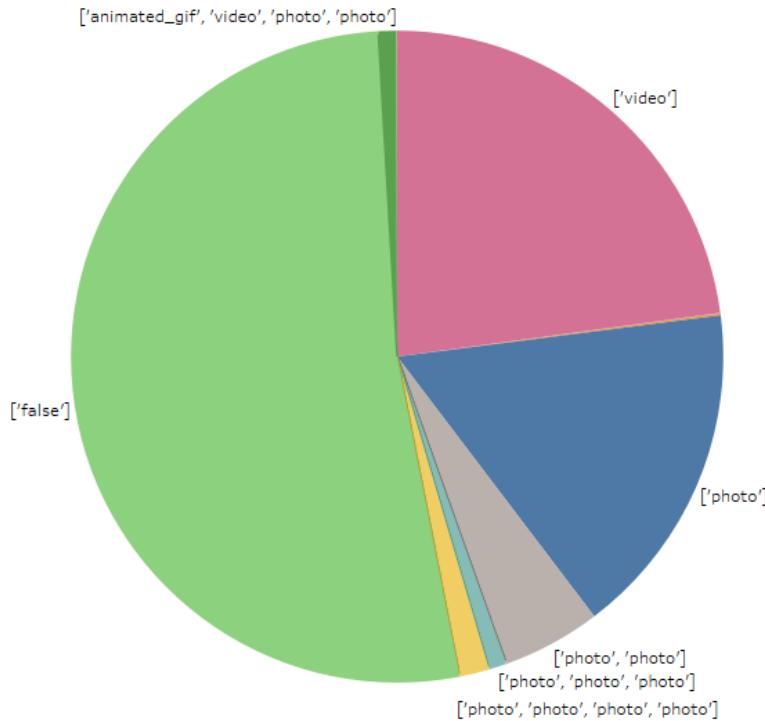


Il·lustració 41: Evolució de l'engagement rate (2022-2023). Font: Elaboració pròpia.

Com es pot observar en la il·lustració 41, el nombre d'interaccions respecte el nombre de visualitzacions pateix variacions, però es manté per sota del 3% i amb una mitjana de poc més d'un 1% durant l'últim any. El primer pic apareix per una qüestió tècnica de Tableau.

#### *Engagement per tipus de contingut*

Com s'ha analitzat en els apartats anteriors, segons el tipus de tweet es genera més o menys engagement. En la il·lustració 42, es mostra el percentatge d'interaccions sobre el total per tipus de contingut multimèdia.



Il·lustració 42: Nombre d'interaccions per tipus de contingut multimèdia. Font: Elaboració pròpia.

Així, podem observar a la il·lustració 42 que la majoria d'interaccions continuen estant relacionades a les publicacions sense contingut multimèdia (color verd), però aquestes perden molt de terreny. I les que contenen algun element com imatges o vídeos, augmenten significativament. Això vol dir que el contingut de Bad Gyal acompañat d'elements audiovisuals atrau més que no aquell simplement format per text. Així, afegir un vídeo (color rosa) és l'element que més interaccions aconsegueix, seguit d'afegir una foto (color blau), en tercer lloc posar un parell de fotos (color gris), i finalment tota la resta de combinacions.

### 5.3.5. Anàlisi de les relacions entre usuaris a Twitter

L'objectiu d'aquest apartat és trobar les relacions que s'estableixen entre els usuaris que mencionen a Bad Gyal. Per visualitzar les 50.030 mensions que s'han extret, es requereixen eines de visualització massives. Per fer-ho, s'utilitzarà el mateix procediment que en les analisis de relacions fetes anteriorment, primer es preparà un

document CSV mitjançant un programa de Python, després es farà la visualització amb Gephi i finalment s'analitzaran els resultats.

### Codi d'extracció

Per crear el document amb les columnes Source, Target i Timestamp per a Gephi, s'ha creat un programa de Python per extreure aquells tweets que tenen mencions des del fitxer original en format Excel. El codi utilitzat es troba a continuació a Codi 7.

```

import pandas as pd
from tqdm import tqdm
import openpyxl

# cream els valors necessaris previs
count = 0
llista_prov = []
total = 500000
chunksize = 100000

# iterem pels valors per extreure cada dada del tweet
for skip in tqdm(range(0, total, chunksize)): # per no sobrecarregar l'ordinador apliquem range i
    # chunksize. Per saber com va el procés apliquem tqdm
    df = pd.read_excel('datasets/dataset-Bad Gyal-lang-es-cat(3).xlsx', skiprows=skip, nrows=chunksize,
usecols="K,X,C")
    for a, tweet in df.iterrows():
        try:
            username = tweet['username']
            created_at = tweet['tweet_created_at']
            # print(username)
        except KeyError:
            pass

        try: # agafem les mencions
            txt = tweet['ent_mentions'].split(';')
            llista_mencions = txt # cream una llista amb les mencions del tweet
            for element in llista_mencions:
                if element == 'false' or element == 'FALS': # treiem els tweets sense menció
                    pass
                else:
                    rel = (username, element, created_at) # agrupem els tres elements
                    llista_prov.append(rel)

        except KeyError: # per evitar que es pari el programa si no hi ha mencions
            pass
        count += 1
        # print(count)

# cream el df i l'exportem a csv
df_final = pd.DataFrame(llista_prov, columns=['Source', 'Target', 'Timestamp'])
# print(df_final)
df_final.to_csv('datasets/mencions(3).csv', index=False)

```

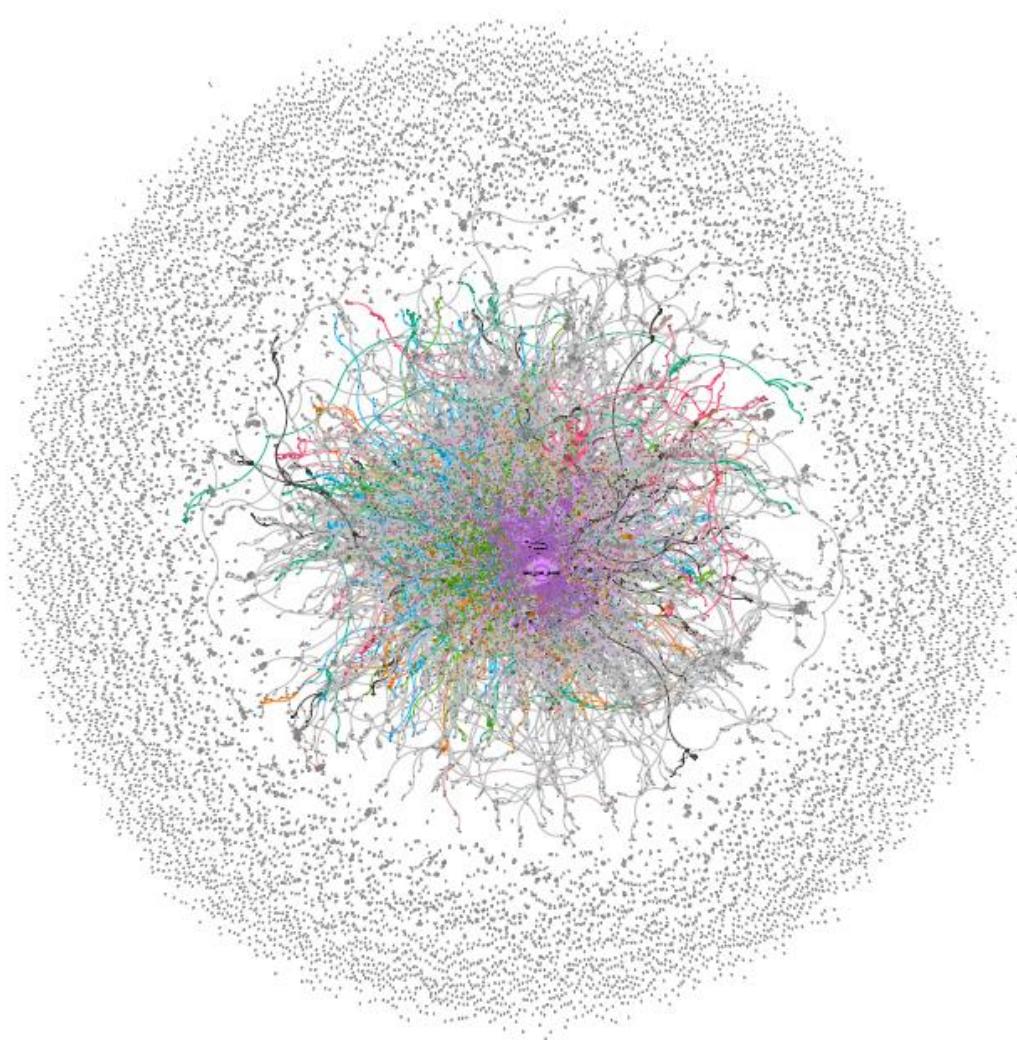
Codi 7: Extracció de les mencions a Twitter per a Gephi. Font: Elaboració pròpia.

Primer de tot s'han importat les llibreries necessàries i s'han creat els valors previs a la iteració. A continuació s'inicia l'extracció amb els bucles. Degut a la mida del fitxer i les limitacions de l'ordinador en el qual s'executa, ha sigut necessari tractar-lo per parts. Per carregar-lo per parts s'ha utilitzat *range* i s'ha establert que carregui les files en grups de 100.000 per tractar-les. Es carreguen aquestes files en un *Dataframe* únicament amb les columnes que ens interessen (K, X i C). Després es creen les variables *username*, *created\_at* i les mencions. Com que no tots els *tweets* tenen mencions, s'utilitza *try* per evitar els errors. I perquè es carreguin totes les mencions es crea una relació per a cada menció del *tweet*.

Es carreguen totes les dades a *rel* i després s'afegeix a la llista *llista\_prov*. A partir d'aquesta llista es crea un *Dataframe* anomenat *df\_final* amb les tres columnes Source, Target i Timestamp. Finalment s'exporta el *Dataframe* a un document CSV que s'obrirà amb Gephi per tractar-lo.

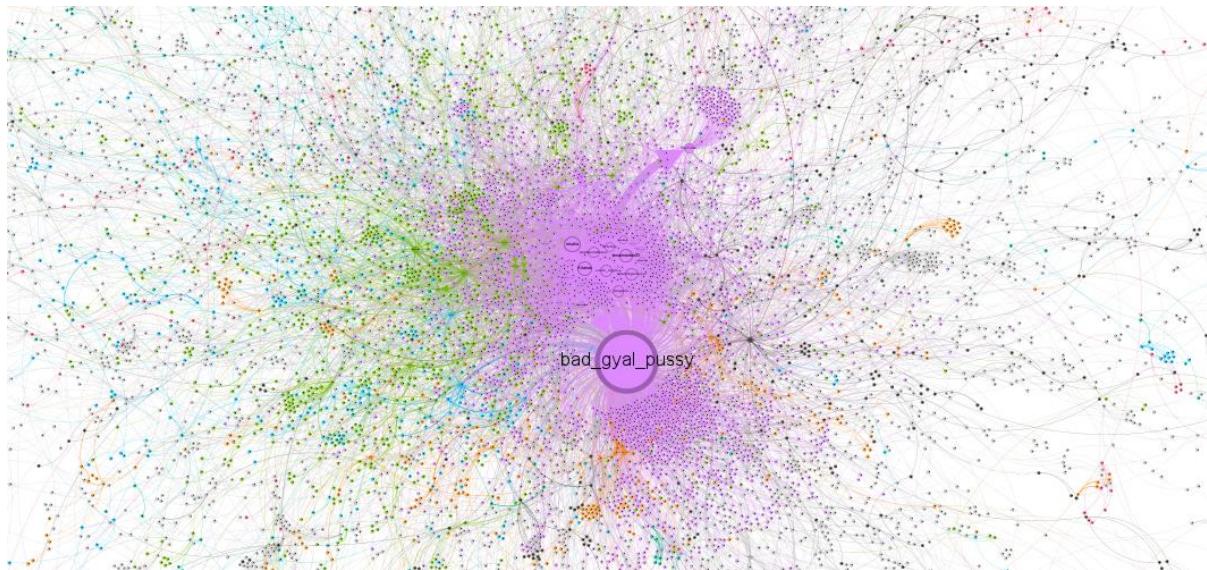
#### *Anàlisi del graf general*

A partir del graf creat amb Gephi, s'ha pogut investigar i extreure informació sobre les mencions que apareixen als *tweets* que mencionen “Bad Gyal”. A continuació s'exposa l'anàlisi i les il·lustracions que s'han generat.



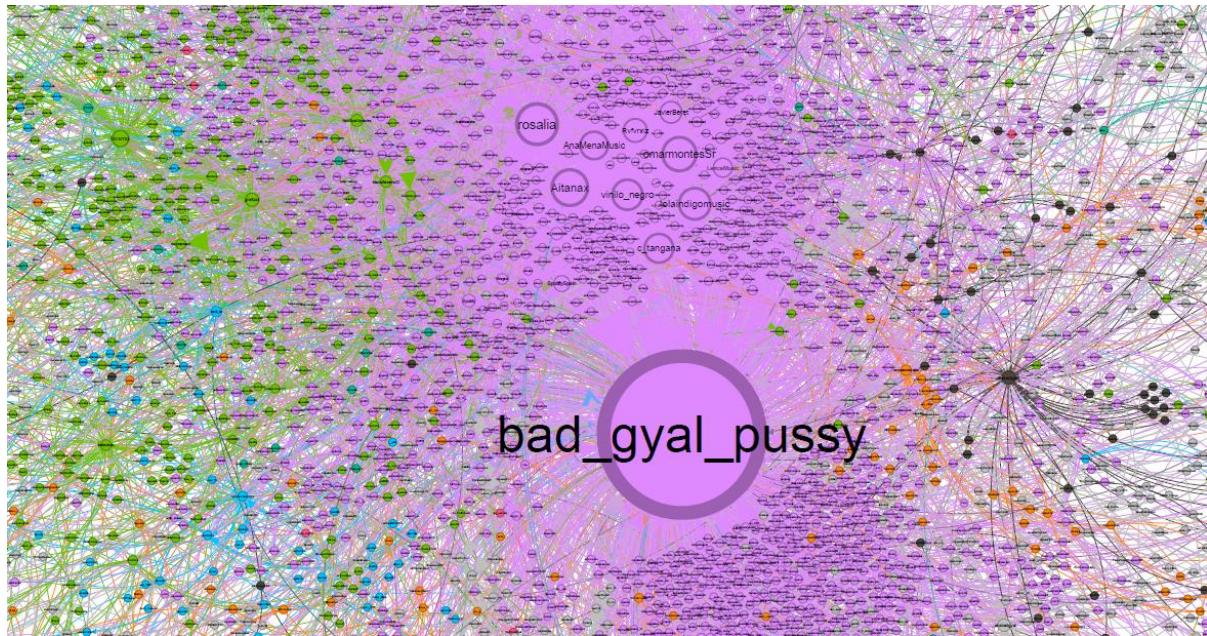
II·lustració 43: Graf amb les mencions de Bad Gyal a Twitter segons el grau d'entrada. Font: Elaboració pròpia.

A primera vista es poden observar dos grups principals. El centre, on els nodes estan molt propers i interrelacionats, i, en segon lloc, l'extraradi amb els nodes aïllats. La comunitat central és més important i té molt més valor, ja que està molt més interrelacionada, tant amb el nombre de perfils, com el nombre de mencions. En canvi, l'extraradi apareix majoritàriament en una única menció i no es mencionen entre altres usuaris. Majoritàriament, es troben desvinculats de la resta d'usuaris.



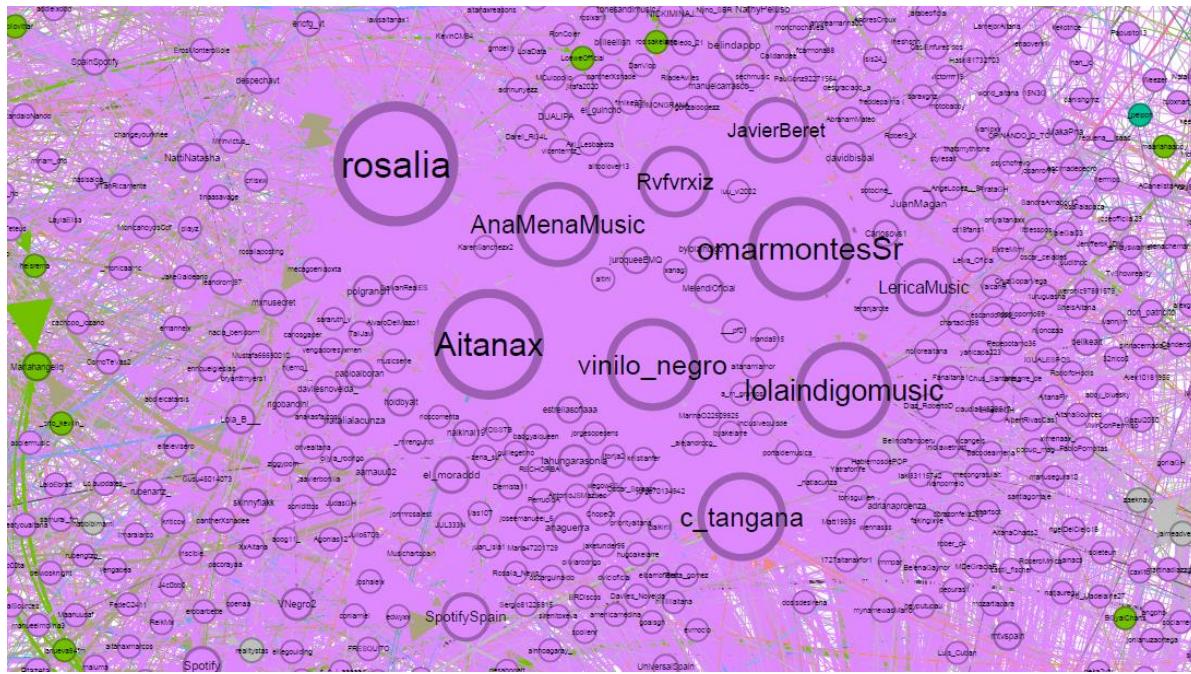
Il·lustració 44: Ampliació de les comunitats al grup central del graf. Font: Elaboració pròpia.

Ampliant el grup central, es poden observar diferents comunitats que s'han representat per colors. Hi destaca una comunitat per sobre de les altres, la de color lila. Aquesta inclou un node que destaca per sobre dels altres, ja que és molt més gran, el que significa que rep moltes més mencions. Aquest node es tracta del compte “bad\_gyal\_pussy”. Destaca perquè aquest és el perfil de l'artista, però no té cap mena de verificació per part de Twitter (*blue checkmark*), ni es recomana en les seves altres xarxes socials o webs. Aquest fet pot ser degut a restriccions de contingut i de drets amb la discogràfica, entre altres motius.

*Anàlisi de la comunitat lila*

Il·lustració 45: Ampliació de la comunitat lila. Font: Elaboració pròpria.

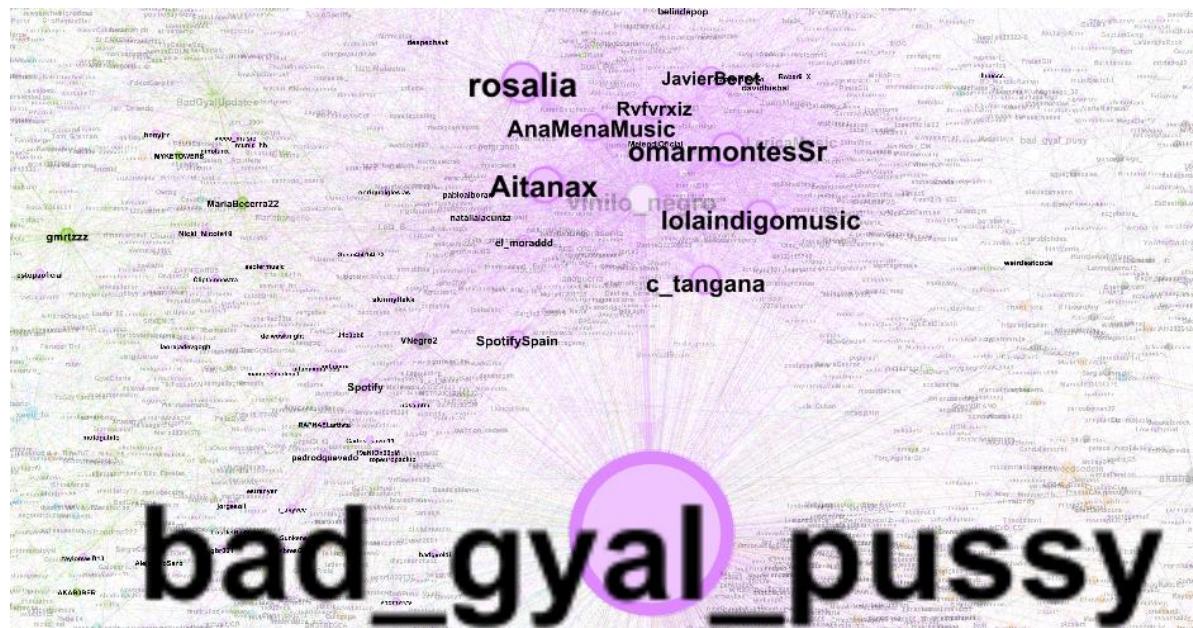
La comunitat de color lila és la més gran i important. Està relacionada amb molts grups i comunitats diferents. Generalment, aquesta comunitat menciona a grans artistes relacionats amb Bad Gyal. En primer lloc, es troba el compte “bad\_gyal\_pussy” el compte més mencionat per tota la comunitat. Però aquests usuaris també fan tweets mencionant a molts altres usuaris relacionats amb el món de la música, com es pot veure en l’ampliació de la il·lustració 46.



II·lustració 46: Mencions destacades comunitat lila. Font: Elaboració pròpia.

En les mencions a altres artistes en un mateix tweet on es menciona a Bad Gyal, destaquen Rosalia, Aitana, Omar Montes i Lola Indigo. Aquests artistes són molt pròxims en l'escena musical espanyola en la qual es troba l'artista Bad Gyal i generalment també són seguits pels fans de l'artista.

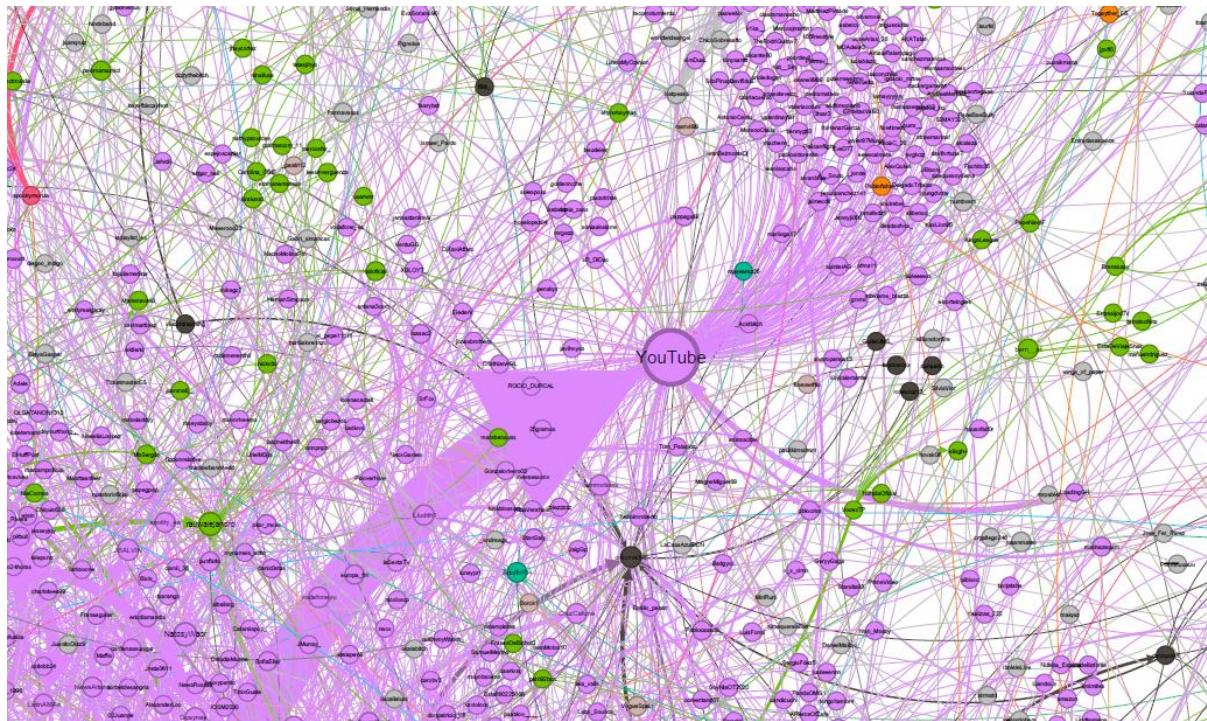
Fora dels comptes dels artistes, també destaquen d'altres relacionats amb el món de la música com el portal de música Vinilo Negro o Spotify. Vinilo Negro es troba molt vinculat a aquests artistes, perquè el seu contingut es basa a publicar actualitzacions de les llistes musicals a Espanya cada vegada que s'actualitzen. Destaca també que el compte vinilo\_negro que apareix amb un node d'una mida molt gran va deixar de *tuitar* el 22 de maig de 2022, però el compte encara segueix actiu. Fent recerca, el portal web segueix actualitzat, però, en canvi, el compte de Twitter associat ha canviat a VNegro2, que sí que segueix publicant diàriament. Aquest fet és destacable perquè vol dir que la importància del node es restringeix a un període de temps més limitat, però, així i tot, aconsegueix ser molt més gran que la resta.



Il·lustració 47: Mencions de l'usuari VNegro2. Font: Elaboració pròpia.

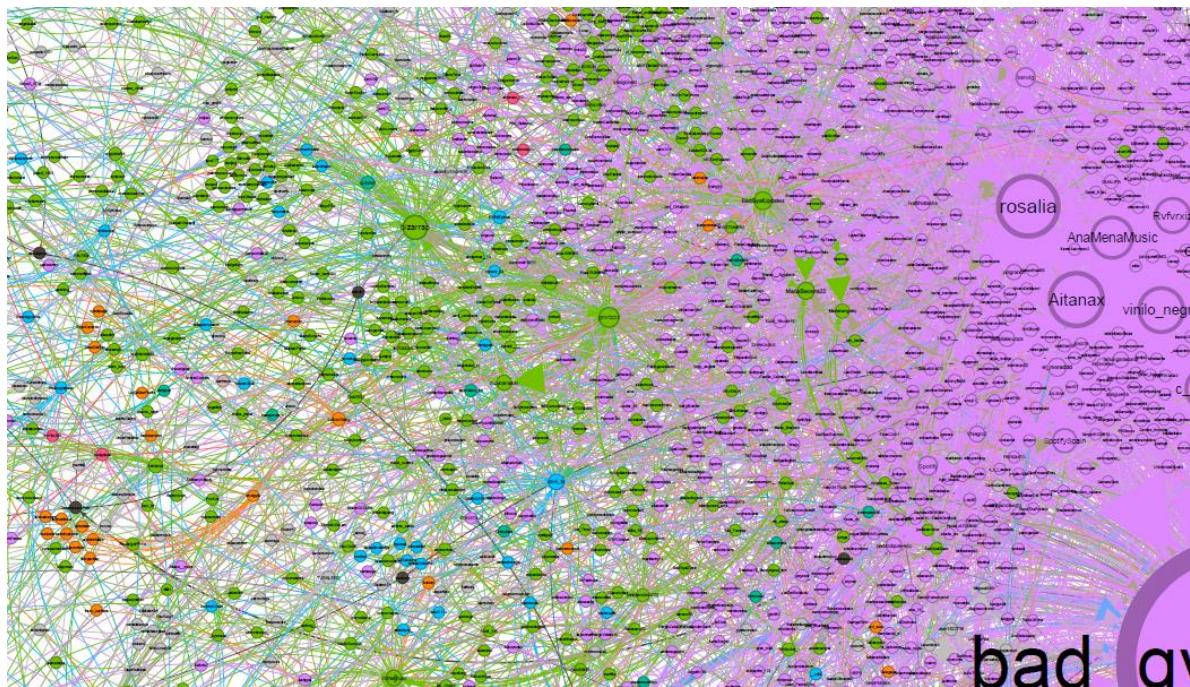
Una vegada que el compte vinilo\_negro va quedar inactiu, va aparèixer el compte VNegro2, el qual es pot observar a la part central-esquerra de la il·lustració 47. Presenta un comportament i unes relacions similars a l'anterior compte, però destaca que no s'han volgut vincular en cap moment amb el compte anterior i així no tenen cap menció en comú.

Respecte al cas de Spotify, els tweets es dirigeixen a dos comptes diferents: SpotifySpain i Spotify, ambdós es troben en la part inferior esquerra de la il·lustració 46. Aquests comptes són oficials i pertanyen a la plataforma líder analitzada anteriorment en aquest treball. S'observa que el compte representant de la plataforma a Espanya té més mencions que el general de tot el món, com pot ser d'esperar, ja que Bad Gyal i la seva comunitat d'usuaris està més vinculada a l'entorn espanyol que al global. Els comptes de Spotify són continus i professionals, però també s'adapten al contingut que funciona millor a Twitter. També cal destacar que les vinculacions del compte SpotifySpain estan molt vinculades a artistes espanyols d'àmbit estatal com Omar Montes, i el compte global Spotify apareix en mencions d'artistes que també poden ser espanyols, però tenen una visió molt més internacional, com és el cas de Rosalía.



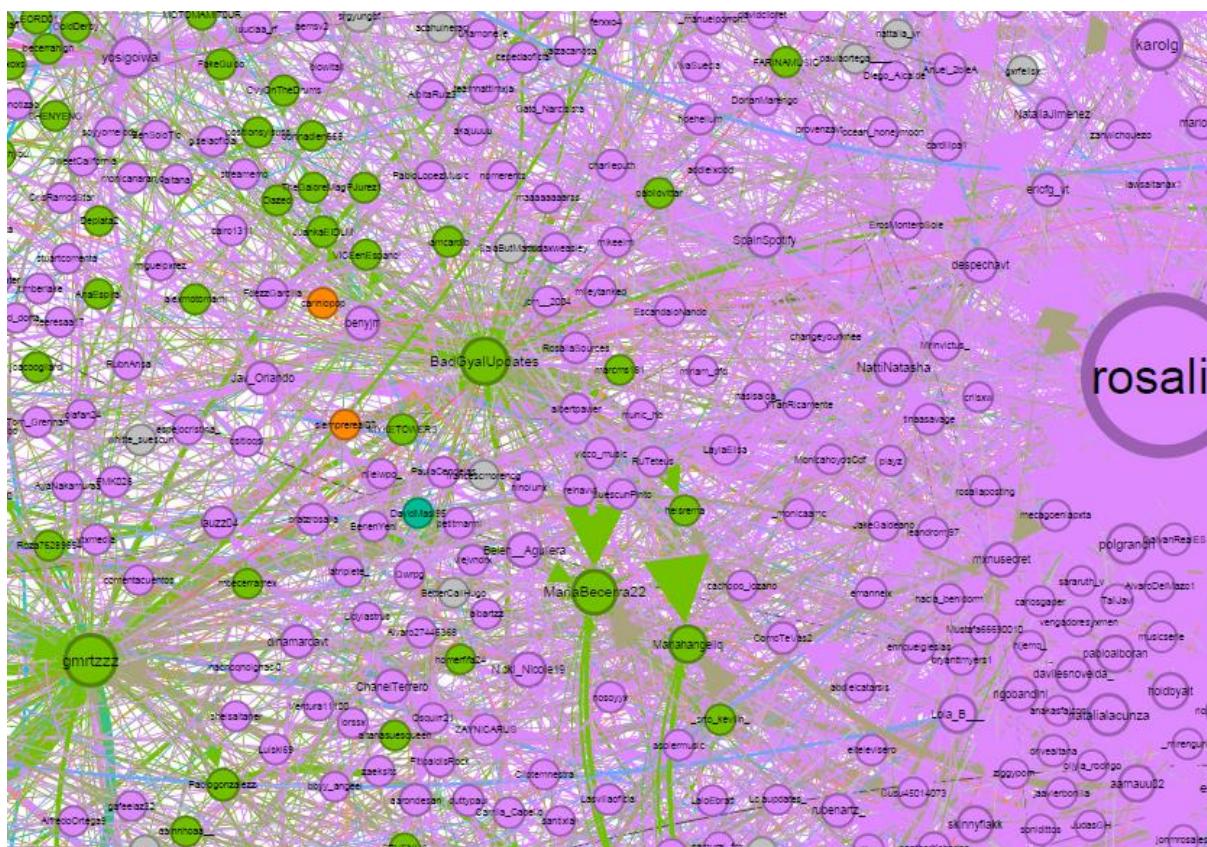
Il·lustració 48: Representació del compte YouTube al graf. Font: Elaboració pròpia.

Finalment, cal destacar una última dada de la comunitat lila, la vinculació amb YouTube. Si bé la mida del node que representa el compte de YouTube oficial a Twitter és d'una mida mitjana, destaca una aresta molt gran d'entrada al node, com s'observa a il·lustració 48. Això vol dir que un compte de Twitter ha mencionat moltes vegades a l'usuari YouTube. Per destacar en una xarxa tan gran, cal que sigui un bot que automatitzi algun tipus de menció. El compte que ha generat tantes mensions ha sigut *vinilo\_negro*. El comportament d'aquest compte està sempre automatitzat i crea diferents *tweets* cada dia amb un nombre molt elevat de mensions, especialment entre els comptes que apareixen amb nodes grans en la comunitat lila, com es pot veure a il·lustració 46 i 48.

*Anàlisi de la comunitat verda*

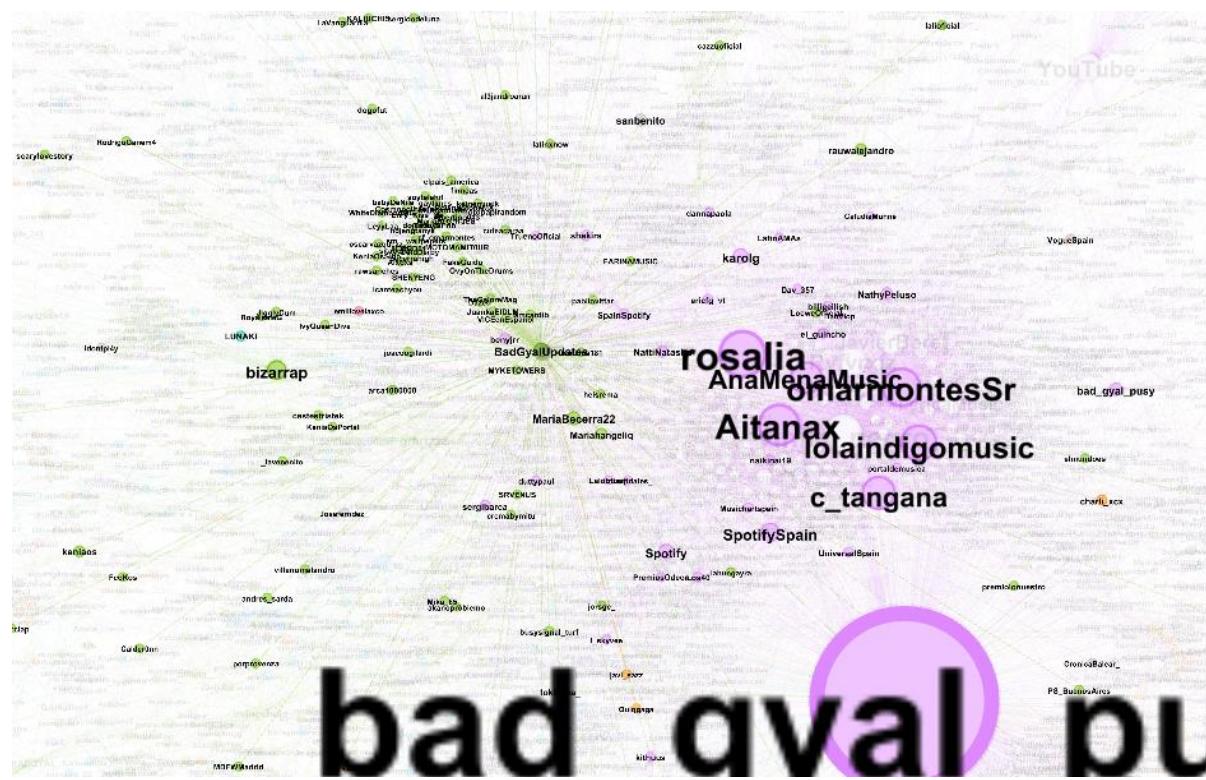
Il·lustració 49: Visió general de la comunitat verda. Font: Elaboració pròpia.

Quant a la segona comunitat més visible, la de color verd, principalment es troben comptes d'actualització del món musical o de fans de Bad Gyal. A causa de la naturalesa dels comptes i a la publicació continua d'aquests tipus de posts destaquen certs nodes i certes arestes que els connecten.



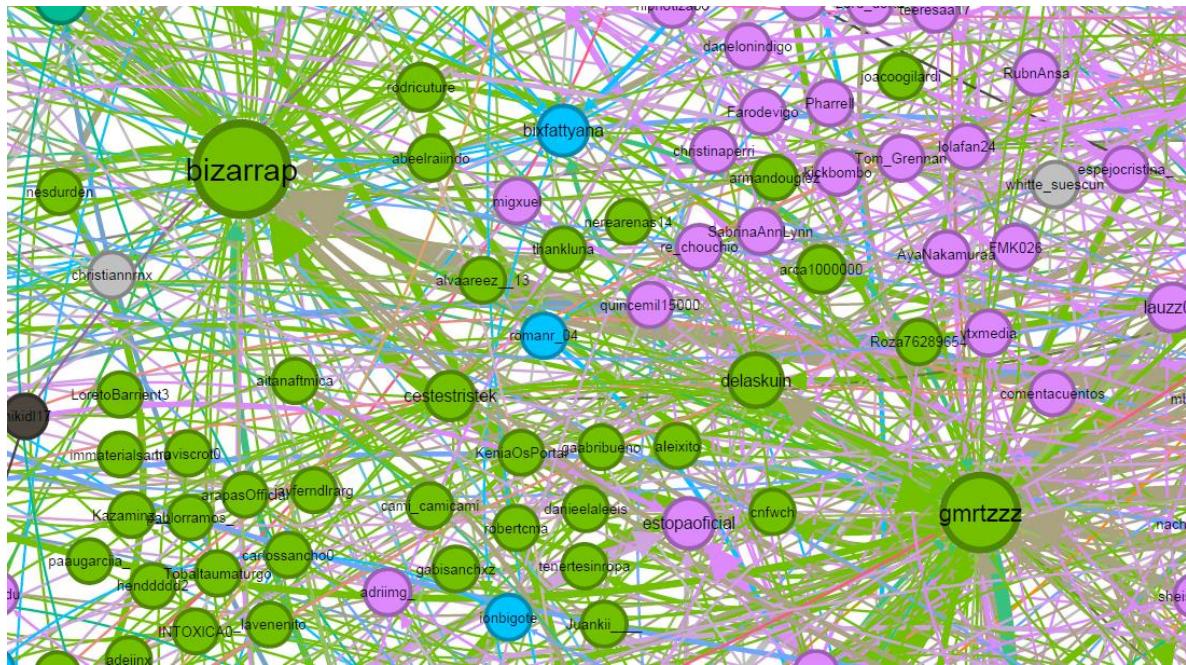
Il·lustració 50: Vinculacions de la comunitat verda i lila. Font: Elaboració pròpia.

Així, les principals vinculacions que apareixen entre les comunitats verda i lila són BadGyalUpdates, MariaBecerra22 i Mariahangeliq. Cal destacar que un gran nombre de les mencions de les dues últimes artistes provenen del primer compte, BadGyalUpdates.



Il·lustració 51: Ampliació de les vinculacions de BadGyalUpdates. Font: Elaboració pròpia.

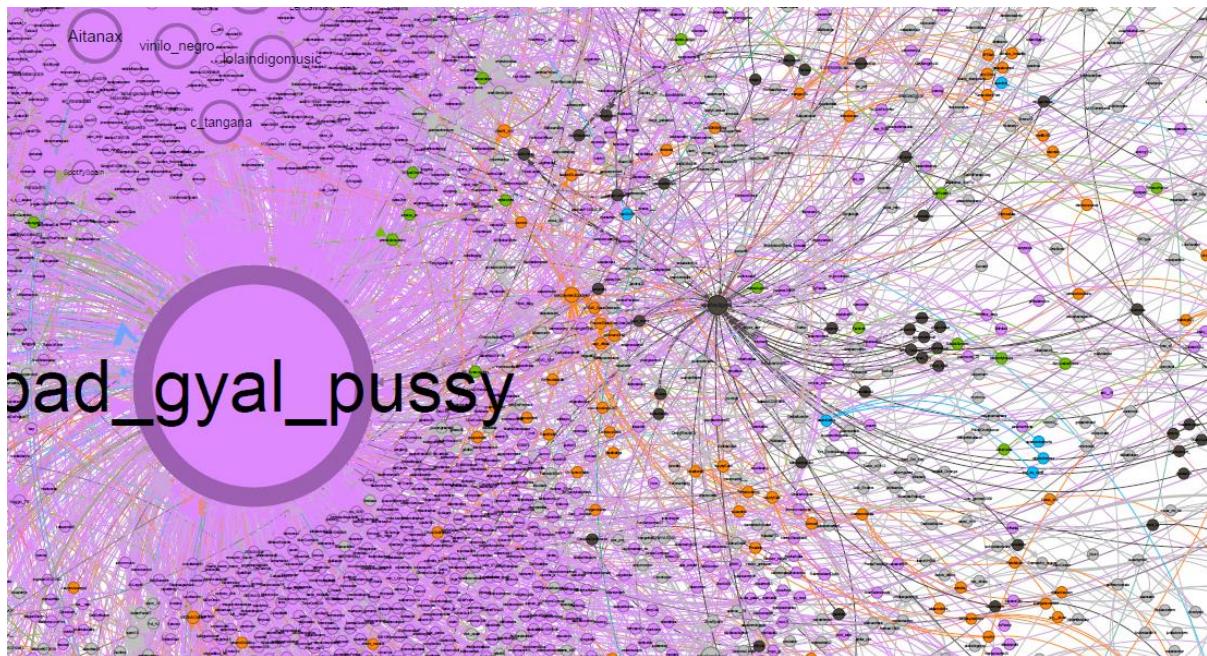
Aquest compte és el principal referent per a les actualitzacions i la informació extraoficial per a fans de Bad Gyal. Com es pot observar a il·lustració 51, fa moltes mencions però aquestes estan diversificades, relacionant-se amb altres grups d'usuaris a Twitter. D'altra banda, segueix la tendència de relacionar-se amb les dues comunitats predominants.



Il·lustració 52: Ampliació de bizarrap i gmrtzzz. Font: Elaboració pròpia.

Els altres dos nodes que destaquen en la comunitat verda, són bizarrap i gmrtzzz. Bizarrap és un artista argentí molt famós actualment a escala internacional, i és conegut per fer col·laboracions amb artistes urbans que arriben als primers llocs de les llistes internacionals. Aquest artista destaca en una comunitat diferent dels altres perquè hi ha hagut rumors que col·laboraria amb Bad Gyal. Fins i tot, ella l'ha anomenat en entrevistes deixant caure que hi havia hagut passos per dur-la a terme. Però especialment, aquests rumors han fet parlar molt entre la comunitat de fans de Bad Gyal, i obviament entre els principals usuaris d'aquesta, que es troben en la comunitat de color verd.

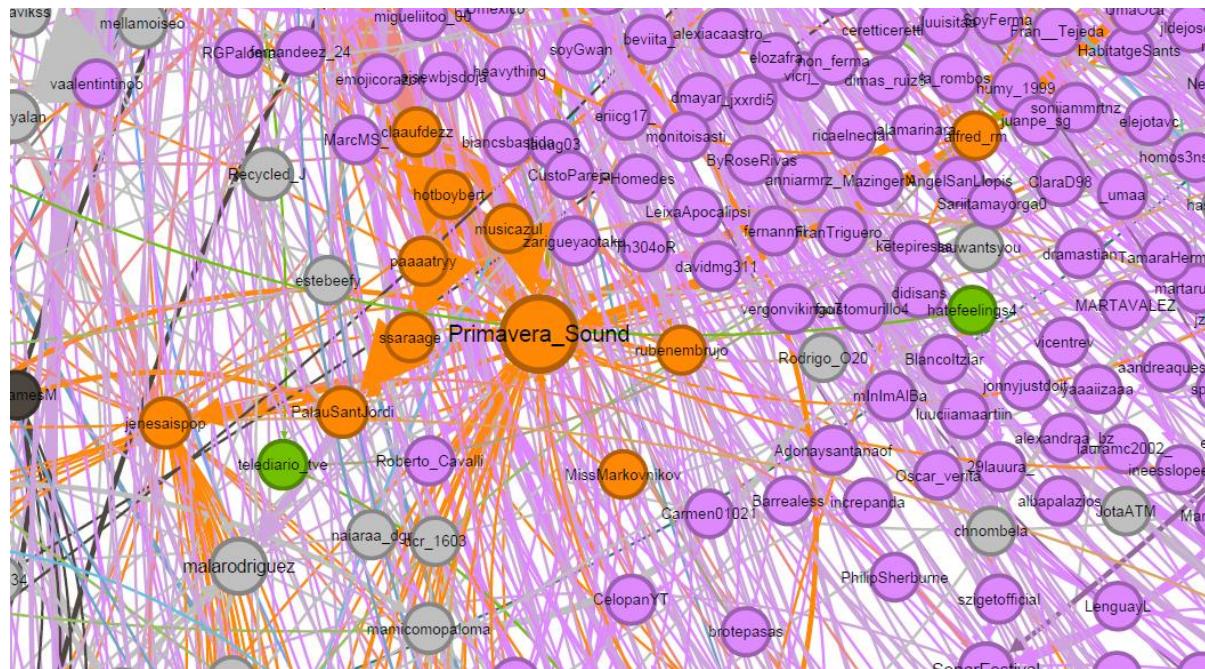
El segon node que s'observa a la il·lustració 52 és gmrtzzz, un compte de fans de les artistes catalanes Bad Gyal, Rosalía i Aitana. Aquest compte destaca pel nombre de mencions que se li fan (mida del node), però també per la quantitat de mencions que fa a certs usuaris (gruix de l'aresta). Menciona en repetides ocasions als artistes Bizarrap i Maria Becerra. Però també estableix moltes relacions amb altres comptes de fans com BadGyalUpdates o muycsansaoo.

*Anàlisi d'altres comunitats (taronja i negra)*

Il·lustració 53: Comunitat negra. Font: Elaboració pròpia.

Si bé el compte controlat per l'artista a Twitter és `bad_gyal_pussy`, a la resta de xarxes socials el seu nom d'usuari és `akabadgyal`. Això pot generar confusions, ja que els artistes, especialment aquells que tenen un equip professional més al darrere, acostumen a unificar la seva imatge a les plataformes i crear una estratègia comuna. En canvi, a Twitter no és així. El compte `akabadgyal` era un compte no controlat per l'artista, tot i que estava dedicat a ella. Actualment cal destacar que el compte apareix com a “Compte suspès” i el següent missatge: “El Twitter suspèn els comptes que infringeixen les normes d'ús.”. Això podria indicar que aquest compte no estava sent prou específic indicant que no es tractava del perfil de l'artista.

Sabent aquesta informació s'entén que el compte real de Bad Gyal a Twitter tingui moltes més mencions, però això no impedeix que altres comptes s'equivoquin mencionant a `akabadgyal` en lloc del compte oficial `bad_gyal_pussy`. De fet, en la xarxa de color negre, no només apareixen comptes personals, sinó que perfils professionals com la ràdio “Los40”, entre d'altres, també s'equivoca.



Il·lustració 54: Comunitat taronja. Font: Elaboració pròpia.

Finalment, es vol destacar la comunitat taronja, que si bé està molt repartida, destaca l'usuari PrimaveraSound, ubicat a sota del de Bad Gyal. Destaca perquè fa moltes mencions a l'artista, però també perquè es troba relacionat amb la xarxa d'organitzadors i sales de concerts, com els comptes PalauSantJordi o WiZinkCenter.

Tant la comunitat de color negre com la de color taronja i la resta que no s'han mencionat explícitament, estan molt espaiades entre si, així com, estan repartides i relacionades per tot l'entorn del graf. Generalment, els usuaris d'aquestes comunitats apareixen en poques mencions i tampoc generen un nombre elevat de mencions.

### 5.3.6. Anàlisi del contingut dels tweets

Per avaluar qualitativament els *tweets* és molt important saber què diuen. Aquesta és una tasca molt laboriosa i lenta que s'hauria de fer tenint a vàries personnes llegint *tweet* per *tweet* el seu contingut. Però a causa de la quantitat de publicacions recollides, gairebé mig milió, aquest mètode no és factible. Per aquest motiu s'ha optat per analitzar el seu contingut de manera automatitzada amb Intel·ligència Artificial (IA) i Machine Learning (ML).

A causa de la novetat d'aquestes eines i que estan en fase de millora, la fiabilitat d'aquestes pot ser dubtosa. Per aquest motiu, s'ha decidit utilitzar tres models diferents per comparar-los posteriorment i obtenir una visió més realista. Els models triats són "xlm-roberta-base-sa-spanish" (MMG, 2023), "spanish\_hate\_speech" (jorgeortizfuentes, 2023) i "distilbert-base-uncased-finetuned-with-spanish-tweets-clf" (francisco-perez-sorrosal, 2023). S'ha investigat per trobar quines eines podrien adaptar-se millor per a ser utilitzades des de codi Python, així com que siguin accessibles i que hagin obtingut bons resultats en altres estudis. Per aconseguir aquests models i obtenir-ne tota la informació, s'ha utilitzat la plataforma Hugging Face. Aquesta és una plataforma que recull molts models d'IA de manera transparent i en la que s'hi poden fer valoracions d'aquests, en comparació amb la majoria de desenvolupadors d'IA que són grans empreses molt opaques (Hugging Face, 2023).

Amb la llista de totes les piulades de Bad Gyal que s'ha utilitzat anteriorment, s'ha extret l'ID i s'ha analitzat el text de cada *tweet* amb el primer, el segon i el tercer model i finalment s'han unit els resultats en un únic document per poder ser utilitzat amb Excel i Tableau. L'automatització s'ha dut a terme a través de Python amb el codi que s'exposa a continuació. Després s'analitzaran els resultats i es valorarà la fiabilitat d'aquests.

```

import glob
from tqdm import tqdm
import pandas as pd
from transformers import AutoTokenizer, AutoModelForSequenceClassification, pipeline
import openpyxl

# Models que s'utilitzaran extrets de Hugging Face
models = ['MMG/xlm-roberta-base-sa-spanish',
          "francisco-perez-sorrosal/distilbert-base-uncased-finetuned-with-spanish-tweets-clf",
          "jorgeortizfuentes/spanish_hate_speech"]
# El dataset en el que estan tots els tweets amb les dades
dataset = "datasets/dataset-BadGyal-lang-es-dataset.xlsx"

# es defineix el procés que es farà per obtenir els resultats
def proc(dataset):
    for model in models: # S'itera per a cada model
        model_name = model.split("/")[1]
        print(model_name)
        dataset_new_name = f"datasets/procesado-{model_name}.xlsx" # Es crea el nom del fitxer amb les
        # respostes del model
        tup_list = []
        df = pd.read_excel(dataset, usecols='E,M') # Per optimitzar les dades s'agafa només les
        # columnes necessàries
        tweets = df["text"].to_list()
        tweet_id = df["tweet_id"].to_list()

        t = AutoTokenizer.from_pretrained(model)
        m = AutoModelForSequenceClassification.from_pretrained(model)

        for tweet, tid in tqdm(zip(tweets, tweet_id)): # Per a cada tweet s'nalitza amb el codi
            seguent
            try:
                pipe = pipeline("text-classification", model=m, tokenizer=t)
                result = pipe(tweet)
                content = result[0]
                label = content["label"]
                score = content["score"]
                tupla = (str(tid), tweet, label, score) # Es desa la resposta en una tupla
                tup_list.append(tupla) # S'afegeix a la llista
            except: # Es crea una excepció per tal que no s'aturi el procés si hi ha cap error
                label = 'null'
                score = 0
                tupla = (str(tid), tweet, label, score)
                tup_list.append(tupla) # S'afegeix el valor a la llista igualment per tenir-lo en
                compte

            # Es crea el data frame a partir de la llista de tuples
            data = pd.DataFrame.from_records(tup_list, columns=[f"tweet_id", "text", f"label-{model}",
            f"score-{model}"])
            # S'exporta la resposta del model
            data.to_excel(dataset_new_name, index=False)

        # S'executa el procediment definit anteriorment
        proc(dataset)

        # Es recull la resposta de cada model
        datasets = glob.glob("datasets/procesado-*.xlsx")
        print(datasets)
        dataset_madre = pd.read_excel(datasets[0])

        # Per a cada model s'afegeixen les respostes al dataset final "dataset_madre"
        for d in datasets[1:]:
            df = pd.read_excel(d)
            dataset_madre = dataset_madre.merge(df, on=["tweet_id", "text"])
            print(dataset_madre)

        # S'exporta el dataset final amb tots els valors en un document d'Excel
        dataset_madre.to_excel('final.xlsx', index=False)

```

Codi 8: Anàlisi del contingut dels comentaris a Twitter amb Intel·ligència Artificial

En el codi 8, primer s'han importat les llibreries necessàries i s'han creat les variables, com la llista de models que es voldrà utilitzar. Posteriorment s'ha definit el procés que es vol que automatitzi per a cada model i per a cada tuit. En aquest procés principalment agafa les columnes d'identificació del *tweet* i del text del *tweet* per tal de no carregar tot el document i optimitzar-ho. Posteriorment s'agafa el *string* del text i s'analitza amb el model. I després s'ordenen les dades per afegir-les en una *tupla* per tal que es puguin exportar en un *Dataframe*. Aquest es desarà en un fitxer XLSX per a cada model.

Una vegada acabat el procés d'anàlisi, per facilitar la visualització amb Tableau, el que es fa és obrir els tres fitxers exportats abans i s'uneixen per tal que no es repeteixin les columnes.

Cal destacar que el codi s'ha hagut d'executar més vegades que els anteriors, ja que han aparegut nombrosos errors. Per optimitzar el temps de creació del codi, s'ha creat el codi amb 1.000 fileres, i una vegada s'ha afegit tot el codi necessari, s'ha utilitzat el *dataset* sencer amb tots els milers de *tweets*. Especialment, han aparegut errors per certes fileres que no estaven correctament completades o que estaven buides. Per això s'ha depurat els principals errors del fitxer d'Excel inicial, ja que eren pocs. De totes maneres, s'han afegit excepcions per tal d'evitar que es pari el codi si hi ha cap altre error.

Establir les excepcions d'errors ha sigut molt important per no haver de repetir l'anàlisi des de l'inici una vegada ja feia hores que s'executava. Cal tenir en compte que l'execució del codi ha durat aproximadament 8 hores per model, és a dir, un dia sencer en total. És una quantitat de temps molt important, i encara més tenint en compte que per executar aquest codi s'ha canviat d'ordinador, a un de més potent per escurçar el temps.

### *Anàlisi de resultats*

Una vegada obtingudes les valoracions dels models d'anàlisi per a cada *tweet*, s'han explorat mitjançant les visualitzacions amb Tableau. Així, a continuació s'exposen els fets més destacats que s'han trobat juntament amb les il·lustracions oportunes.

## Jorge-Hate Speech



## Jorge-hate Speech (2)



II·lustració 55: Resposta del model Hate Speech. Font: Elaboració pròpia.

El primer model que s'ha aplicat ha estat el de “hate speech”. És un model més genèric que es dedica a valorar el discurs d'odi (*hate*) en el *tweet* que se li passa. Simplement classifica els *tweets* en “hate” o “no\_hate”. Com es pot observar a II·lustració 55, el nombre de comentaris amb discurs d'odi (1,15%) és molt inferior als que s'han classificat com a “no hate” (98,85%).

## Francisco-Distilbert



## Francisco-Distilbert (2)



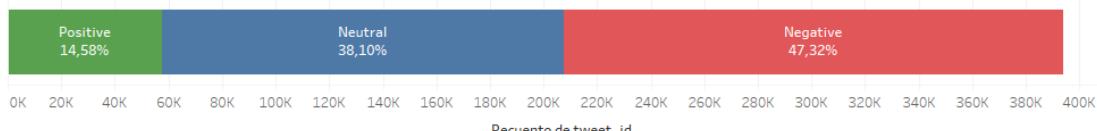
II·lustració 56: Resposta del model Distilbert. Font: Elaboració pròpia.

En el segon model, es pot observar un altre tipus de classificació, segons si les paraules dels usuaris són positives, neutrals o negatives per a la imatge de Bad Gyal. S'observa que l'etiqueta amb més *twets* (207,985) és la neutral, un entorn que inclou aquells *tweets* en els quals no s'exposa ni una opinió a favor ni en contra de manera explícita. D'altra banda, un 39,22% es correspon a tuits negatius i únicament un 7,96% de les piulades són amb contingut clarament a favor de l'artista.

Label-Roberta



Label-Roberta (2)



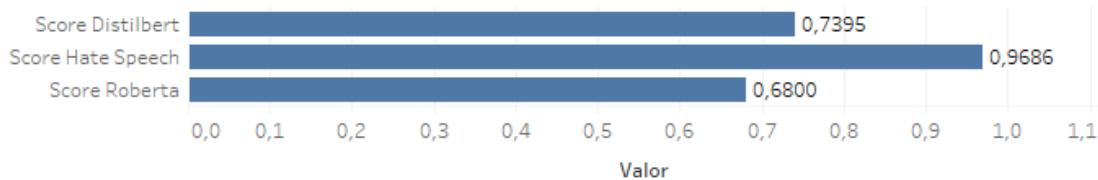
Il·lustració 57: Resposta del model Roberta. Font: Elaboració pròpia.

En el tercer model, s'ha trobat més polarització que en les anteriors anàlisis. Aquest fet destaca perquè pot significar un funcionament més concret o que aquest model tingui menys factors en compte, com per exemple la ironia, i llavors sigui més radical.

Tant el nombre de tweets positius (14,58%) com negatius (47,32%) ha augmentat, fent disminuir a l'etiqueta neutral (38,10%). Tot i això, els resultats de Roberta i de Distilbert són força semblants, a diferència de Hate Speech, que ha obtingut un valor molt baix de comentaris amb discurs d'odi. És probable que els valors variïn perquè el discurs d'odi és un comentari negatiu molt més radical i explícit, per la qual cosa es pot esperar que siguin molt menors.

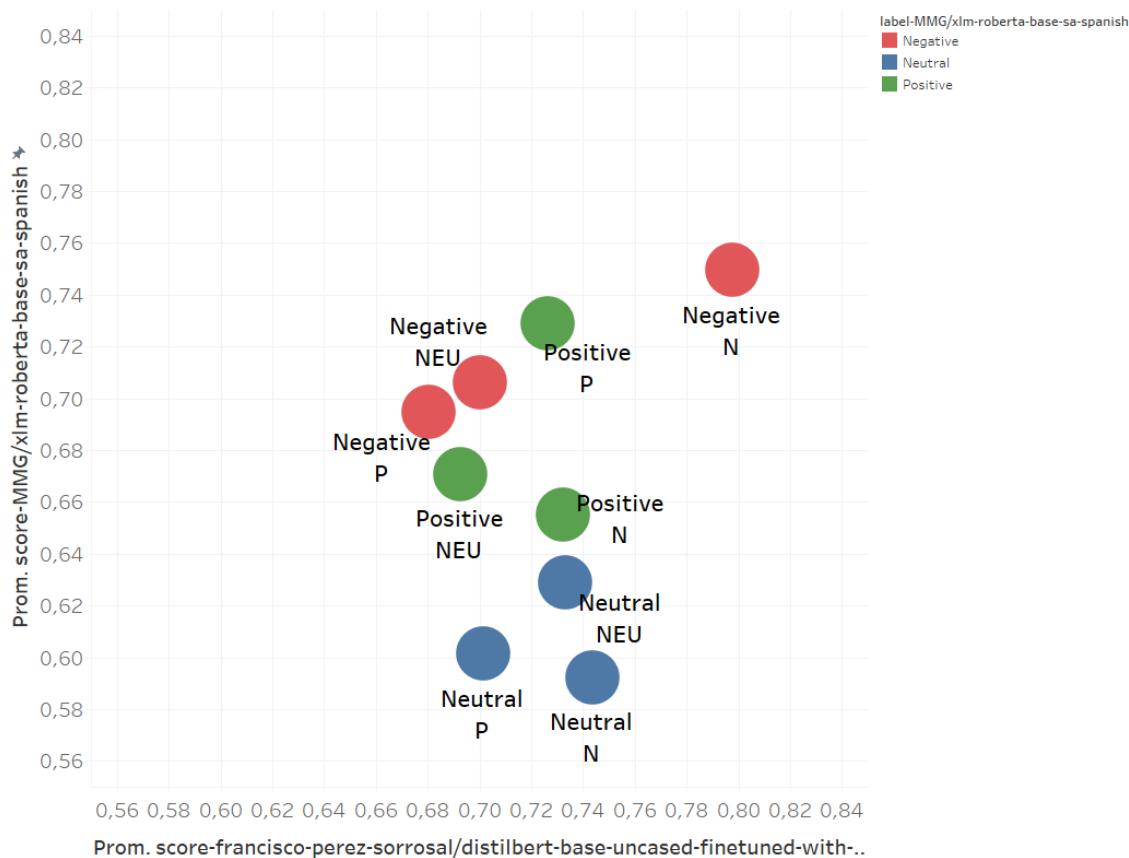
### *Fiabilitat dels resultats*

Com s'ha esmentat a l'inici del punt 5.3.6, aquest tipus d'anàlisi a través d'eines que incorporen intel·ligència artificial i *machine learning* encara requereixen moltes millores per disminuir notablement el nombre d'errors. Per això, quan el model retorna la resposta de quina etiqueta se li posa al tweet, també retorna el percentatge de seguretat amb el que assegura aquesta decisió. Així, a continuació s'analitzaran els resultats de fiabilitat a través de diferents visualitzacions de Tableau.



Il·lustració 58: Comparació de la mitjana de fiabilitat per a cada model. Font: Elaboració pròpia.

En la il·lustració 58 es mostra la mitjana del valor de seguretat amb que es retorna la resposta per a cada model. Així el model que afirma amb més seguretat la resposta és el Hate Speech, seguit del model Distilbert i finalment el model Roberta. Sembla que si hi ha discurs d'odi (*hate speech*) és detectat més fàcilment que la classificació en tres etiquetes que es fa en els altres models.



Il·lustració 59: Comparació de la fiabilitat dels models Roberta (eix y i colors) i Distilbert (eix x)

A la il·lustració 59 es pot observar la comparació de la fiabilitat segons el resultat dels models Roberta i Distilbert. En general, Distilbert ofereix més fiabilitat que Roberta en

els seus resultats, però segons la resposta aquesta fiabilitat també varia. Si coincideixen els resultats dels dos models és quan hi ha la màxima fiabilitat de resposta (Negative-N o Positive-P). En canvi, destaca que els pitjors resultats, en general, no s'obtenen quan un model aporta una resposta contrària a l'altre, sinó quan apareix un resultat que és neutral. Els trets d'un *tweet* positiu o negatiu són més fàcils d'identificar que un de neutral, per això pot ser que tinguin menys fiabilitat les piulades classificades com a neutrals. Així es pot observar a la il·lustració que les respostes en que coincideixen els models es troben a la part superior dreta i en les que hi ha l'element neutral es troben a la part inferior esquerra.

# 6. Conclusions

Aquest apartat ha de permetre tancar el treball de fi de grau recopilant tota la informació tractada fins ara per tal d'extreure conclusions. Aquestes, cal que estiguin vinculades als objectius del treball esmentats en el punt 3. Per aconseguir-ho, les conclusions es dividiran en dues parts, la primera farà una valoració punt per punt dels objectius secundaris respecte la informació que s'ha extret durant l'anàlisi, i en la segona part es farà un resum per conoure el treball.

## 6.1. Anàlisi dels objectius

A continuació es recuperaran els objectius i les preguntes de recerca per analitzar els resultats obtinguts. Els cinc primers objectius corresponen a l'anàlisi quantitativa i els dos últims a l'anàlisi qualitativa.

### 6.1.1. Observar l'evolució temporal de l'artista a les plataformes i serveis d'àudio sota demanda.

Bad Gyal inicia la seva carrera musical l'any 2016 i actualment segueix activa. En relació amb les tres plataformes analitzades, la primera a la qual s'uneix és YouTube. El dia 8 de març 2016 es crea el compte i la seva primera aparició és la publicació del vídeo “BAD GYAL – PAI” l’11 d’abril de 2016. Després, l’11 de novembre de 2016 fa la seva primera aparició a Spotify amb el primer àlbum Slow Wine Mixtape. I la tercera plataforma a la qual s’uneix és Twitter al desembre de 2016.

En relació amb la tendència, es pot observar que el nombre de cançons a Spotify per any, ha augmentat. Destaca també per com es fan les publicacions en aquesta plataforma, ja que fins al 2020, pràcticament només es va publicar dos àlbums (2016 i 2018). I posteriorment es van promocionar més els singles, encara que també es publiqués un àlbum de vuit cançons (2021). En canvi, la tendència de vídeos a YouTube, no ha augmentat, ja que ha sigut molt desigual al llarg del temps. Als primers

anys destaca la publicació de singles exclusivament a YouTube i que no s'han traspassat als altres serveis musicals professionals. És per aquest motiu que als primers anys hi va haver moltes publicacions a YouTube, gairebé les mateixes que en els anys més actuals. Finalment, la tendència de Twitter és clarament a l'alça, ja que la seva comunitat ha augmentat any rere any.

En les publicacions, s'observa un pic per temporada, tant a les publicacions de cançons com al Twitter. La temporada en la qual es publica més és a final d'any (desembre) i a la primavera (abril). En canvi, en els mesos d'agost i gener, l'activitat disminueix notablement en totes les plataformes.

També s'observa un comportament cíclic en totes les plataformes en relació amb el nombre de publicacions per dia de la setmana. Els dijous és el dia en el qual es fan més publicacions, seguit del divendres amb un nombre de publicacions molt pròxim. Destaca que al llarg dels anys, aquesta tendència s'ha consolidat, i cada vegada es publica més en aquests dies des del perfil oficial. El motiu principal és la professionalització de l'artista, que ha incorporat especialistes per assessorar-la en aquests aspectes, principalment per obtenir uns bons resultats en cada publicació.

S'observen períodes pics i vall de manera diferent en totes les plataformes, però en general destaca que durant les hores de la tarda i nit corresponents a Barcelona hi ha un pic de publicacions, respecte a la matinada i el matí que és el període vall.

La relació entre els llançaments musicals (YouTube i Spotify) i el nombre de tuits, acostumen a incrementar de manera simultània.

### 6.1.2. Analitzar quin tipus de contingut comparteix l'artista a les plataformes YouTube i Twitter i al servei d'àudio sota demanda Spotify.

A partir de l'anàlisi es pot concloure que en les plataformes i serveis analitzats trobem diferents tipus de continguts. A Spotify trobem únicament 47 cançons i en alguns casos tenen la lletra inclosa i el vídeo de fons. A YouTube es troben 78 vídeos, tot i que també hi ha un directe i publicacions de “comunitat” cada vegada que lanza un vídeo

o fa una publicació important. Cal destacar que els vídeos són propietat de dos canals diferents. El primer, “Bad Gyal”, té 42 vídeos, generalment amb contingut extra, però també amb les cançons més antigues. El segon canal, “Bad Gyal Vevo”, té les cançons més modernes i suposa una professionalització molt important del contingut de l'artista, juntament amb la col-laboració amb el segell discogràfic Universal Musica.

Finalment, a Twitter s'han extret 400.000 *tweets* amb el text “Bad Gyal” des de 2016 fins al 2023, tant en català (11.276) com en castellà (392.252). La majoria de les piulades (90%) contenen únicament text, mentre que un 10% contenen elements multimèdia, la majoria d'aquestes amb imatges.

### 6.1.3. Detectar el contingut amb més repercussió a YouTube i Twitter.

Primer cal recordar que Spotify és un servei d'àudio sota demanda i que no permet les interaccions socials amb el contingut. D'altra banda, Youtube i Twitter sí que permeten aquestes accions però de manera diferent.

Les publicacions de YouTube que reben més interaccions són els vídeos musicals, especialment els que s'han publicat més recentment des de Vevo. Però la cançó estrella de Bad Gyal, Fiebre, continua amb les millors dades amb més de 500.000 d'interaccions i 65 milions de visualitzacions. D'altra banda, el contingut amb una ràtio d'interaccions més alta és el contrari, el contingut menys musical del canal Bad Gyal.

A Twitter, hi ha més possibilitats d'interacció que en les altres plataformes. D'una banda, la naturalesa del tuit és majoritàriament original (69%), però també hi ha una part important de *reply* (21%) i *quoted* (9%). D'altra banda, la interacció mitjana dels usuaris amb els botons és a través de m'agrada (6.307), *retweet* (587) i finalment *reply* (68). Destaca que les piulades que tenen contingut multimèdia reben moltes més interaccions que aquelles que no en tenen.

### 6.1.4. Quantificar la resposta del públic segons l'usuari que publica el contingut.

A Spotify i YouTube s'ha analitzat exclusivament el contingut publicat de manera oficial perquè agrupa gairebé tot l'espai al voltant de l'artista. En canvi, a Twitter és molt més important el contingut d'altres usuaris fora del perfil de l'artista, ja que Bad Gyal gairebé només fa publicacions quan es llança una nova cançó. De totes maneres, des del perfil oficial es fan moltes repulades, mentre que el text que s'insereix és curt i l'artista no es menciona a ella mateixa. Per aquest motiu, no queden recollits els *tweets* en el contingut analitzat massivament.

En relació amb la quantitat de publicacions, s'ha observat que els usuaris que publiquen més són comptes que opinen i publiquen informació sobre música de diferents artistes com gmrtzzz (1.938 tuits) o yosigoiwal (960 tuits). En canvi, pel recompte d'interaccions destaca el compte BadGyalUpdates (255.180 interaccions acumulades) amb contingut d'actualitat sobre l'artista Bad Gyal exclusivament.

En relació amb l'evolució temporal, s'observa que a partir de 2016, quan Bad Gyal comença la seva carrera musical, hi ha un increment del nombre de tuits, posteriorment, l'any 2019, quan fa la col·laboració amb Omar Montes i signa amb la discogràfica torna a haver-hi un increment. Finalment, destaca un pic del compte gmrtzzz que no és seguit per la resta de comptes, des de mitjans de 2021 fins a inicis de 2022.

### 6.1.5. Realitzar una anàlisi comparativa entre YouTube, Twitter i Spotify.

L'evolució temporal del nombre de publicacions als diferents llocs ha sigut desigual. Spotify ha tingut un nombre de publicacions desigual i la seva tendència de publicació ha disminuït significativament. YouTube també ha tingut un nombre de publicacions desigual per anys i té una tendència lleugerament a la baixa, però es manté pràcticament estable. D'altra banda, Twitter incrementa significativament la seva

tendència de publicació, ja que ha crescut el nombre de publicacions en els últims anys que mencionen a l'artista.

Spotify no té interaccions, mentre que YouTube i Twitter en tenen moltes. A YouTube destaca el nombre d'interaccions en els vídeos oficials (més de 6 milions de mitjana) amb poques publicacions (78 en total), mentre que a Twitter hi ha moltes més publicacions (més de 400.000) però no hi ha tantes interaccions per a cada publicació (19,34 interaccions de mitjana). El nombre d'interaccions per impressió és d'1,94% a YouTube i d'1,17% a Twitter, uns valors força elevats.

### 6.1.6. Saber si la conversa és positiva, negativa o neutra per a Bad Gyal a Twitter.

En l'anàlisi a Twitter a través d'Intel·ligència artificial, s'ha observat que hi ha un gran gruix de contingut detectat com a negatiu, aproximadament un 40%. Mentre que únicament un 10% és positiu. Analitzant el contingut de manera manual s'observa que no hi ha tant de contingut negatiu. Això podria ser degut al fet que els fans de Bad Gyal utilitzen moltes paraules malsonants de manera positiua, així com un llenguatge molt explícit, per la qual cosa el model d'anàlisi podria donar resultats erronis.

S'ha automatitzat l'anàlisi del contingut perquè és impossible fer-ho de manera individual per a cada *tweet* amb una sola persona. Els resultats obtinguts no són els esperats, ja que llegint tuits aleatoris s'observa molt més contingut informatiu i de celebració que no negatiu. Això podria ser perquè aquests models d'anàlisi amb ML són generalistes i encara s'estan revisant. De fet, la fiabilitat mitjana de les respostes és d'entre un 60% i un 70%. En estudis futurs es podria ampliar aquesta anàlisi amb IA de les piulades estudiant més els models i comparant-lo amb els resultats d'altres plataformes, com podrien ser els comentaris de YouTube.

### 6.1.7. Analitzar les relacions del públic de Bad Gyal amb artistes musicals i altres temes.

Els vincles entre persones que es generen a partir de l'artista Bad Gyal s'ha creat a partir de Spotify i de Twitter. A Spotify s'han extret els 20 artistes recomanats a partir del perfil de Bad Gyal, i després s'ha extret el mateix amb aquests (420 recomanacions en total). Els artistes més pròxims a Bad Gyal són: Albany, Sticky M.A., Soto Asa, La zowy, La Mafia del Amor, Kaydy Cain, Yung Beef o Dellafuente. D'altra banda, si s'observen les recomanacions que es fan des dels altres artistes cap a Bad Gyal destaca que és una de les menys recomanades del seu entorn més proper.

Respecte a la relació dels usuaris de Twitter, s'observa que el compte oficial de Bad Gyal és el més central i el més mencionat amb diferència. Hi ha quatre tipus d'usuari amb els que es relaciona: altres artistes, professionals de la música, comptes de fans i altres usuaris. Els comptes que més es mencionen juntament amb el de Bad Gyal són els artistes Rosalia, Aitana, Omar Montes, Bizarrap i Lola Indigo, així com Vinilo Negro, una revista musical que fa publicacions automatitzades. En un segon grau es troben els comptes de fans com gmrtzzz o BadGyalUpdates i també els comptes de professionals de la música que fan de promotores com Primavera\_Sound o plataformes com SpotifySpain i YouTube. Finalment, trobem la majoria d'altres usuaris que han mencionat o són mencionats poques vegades.

Entre els artistes més propers a Bad Gyal no apareixen gaires amb els que ja hagi fet col·laboracions, almenys de manera destacada, sinó que apareixen artistes que també agraden als fans de Bad Gyal. D'altra banda, a Twitter els artistes que han col·laborat amb Bad Gyal són més mencionats i apareixen més propers que a Spotify.

En relació amb els gèneres vinculats amb l'artista, destaca que són força acurats. Alguns d'aquests són *urbano espanol*, *urbano latino*, *rap latina*, *trap latino*, *reggaeton*, *trap catala*, *rap catala*, *neoperreo*, *puerto rican pop* o *reggaeton flow*. En general, Spotify vincula les cançons de l'artista amb els gèneres *reggaeton*, *urban*, *rap* i *trap*; una classificació força correcta amb la realitat. Cal destacar també que es vincula a gèneres i subgèneres molt concrets en algunes cançons, especialment en les que són més antigues o en les col·laboracions.

Per a futurs estudis seria interessant analitzar altres artistes vinculats amb Bad Gyal per poder-los comparar i tenir una visió més exhaustiva d'aquests gèneres musicals.

## 6.2. Conclusions generals

Finalment, en aquest segon apartat se sintetitzen les principals conclusions del treball de fi de grau. En relació amb l'objectiu principal d'"Analitzar la conversa a les plataformes i serveis sota demanda al voltant de l'artista Bad Gyal", crec que s'ha complert amb escreix, tal com s'ha pogut veure en l'apartat anterior en la valoració dels objectius secundaris i les preguntes de recerca associades. **S'ha fet una anàlisi en profunditat de Spotify, YouTube i Twitter que ha permès obtenir una visió holística de l'entorn de l'artista.**

Especialment, crec que **té molt de valor el mètode que s'ha seguit**, ja que hi ha pàgines web que extreuen certes mètriques però no el conjunt de dades de la plataforma o servei. En canvi, seguint el mètode que he triat, amb la programació i extracció de dades a través de les API, s'obté unes dades molt més completes. Es pot veure la importància del mètode d'extracció de dades a través de l'API, per exemple, en el nombre de tweets analitzats, més de 400.000.

També voldria destacar que **la implantació d'Intel·ligència Artificial amb Machine Learning ha estat una iniciativa innovadora** molt útil perquè ha permès analitzar aspectes qualitatius molt importants que d'altra manera no haguessin sigut possibles d'analitzar. És interessant aplicar aquestes metodologies innovadores perquè l'entorn de les plataformes i serveis sota demanda s'actualitza molt sovint, implementant tecnologies en prova també.

Al llarg d'aquest treball s'ha fet un esforç per mostrar tot el codi creat i els problemes que han aparegut. Però cal indicar que **per a cada codi funcional, s'han fet moltes hores d'investigació, i resolució** que no és adequat de posar en aquesta redacció del treball perquè s'estendria massa.

Tot i que l'anàlisi ha sigut molt complet, es podria ampliar en estudis futurs amb més recursos. Principalment, es podria analitzar altres xarxes socials com TikTok o Instagram, així com comparar les dades de perfils d'artistes similars. Però també es podria ampliar l'anàlisi amb Intel·ligència Artificial en altres apartats com els

comentaris de YouTube i amb altres objectius com trobar temes o vocabulari recurrents.

Crec necessari apuntar que aquest treball de final de grau ha tingut certes limitacions per les característiques d'entrega del treball. En primer lloc, hi ha una **limitació de format** marcada per la normativa de la facultat. El format d'entrega com a memòria en DIN A-4 limita la capacitat de mostrar certes dades, especialment quan es tracta de Big Data. És per aquest motiu que en algun moment cal fer un esforç per entendre que s'ha realitzat una exploració de les dades amb el programari de *big data* mencionat. D'altra banda, el treball també es troba **limitat temporalment** per un període de certs mesos inferior a l'any universitari. El treball es podria dividir en d'altres o investigar més a fons amb més temps per a fer-ho. L'última que cal destacar és la **limitació tècnica**. Aquest treball s'emmarca dins del grau de comunicació interactiva, una formació propera a aquesta investigació però no especialitzada en aquest entorn. És per aquest motiu que s'ha hagut de fer un esforç d'aprenentatge de codi Python i comunicació amb API. D'altra banda els recursos tècnics com l'ordinador, el processador o la memòria, entre d'altres, no són els més punters del mercat, perquè són els meus recursos personals que tenia disponibles.

Personalment, faig una valoració positiva del treball, ja que **he après molt tècnicament però també de manera transversal** en àrees que m'interessen com són les plataformes digitals, l'artista Bad Gyal o els processos de recerca i investigació. En l'àmbit pràctic he ampliat els meus coneixements en programació amb Python, l'ús d'APIs i l'extracció i el tractament de Big Data. D'altra banda, també he adquirit noves habilitats per gestionar les tasques i el temps, així com avaluar i sintetitzar les dades millor.

Aquest treball ha investigat amb profunditat i de manera complexa l'entorn de les plataformes i serveis sota demanda de Bad Gyal. És per això que en l'entorn acadèmic **pot ser un precedent de recerca innovador i valuós** per a futures investigacions en el camp de les plataformes i serveis sota demanda.

També voldria destacar el valor d'aquest treball pel fet que qualsevol empresa, persona o marca necessita saber com s'està comportant el seu públic. Concretament,

aquest estudi aporta molt de valor per a les entitats que donen suport als artistes musicals com productores o discogràfiques, entre d'altres. Aquest estudi **és especialment valuós per als col-laboradors de Bad Gyal**, per això es podria oferir a alguna de les següents entitats: Universal Group, Interscope Records, Aftercluv o Canada Videos.

# Bibliografia

20 minutos. (2022). *La artista Halsey destapa la obsesión de las discográficas con TikTok: «No me dejan sacar mi próxima canción si no se hace viral».*

<https://www.20minutos.es/noticia/5004607/0/la-artista-halsey-destapa-la-obsesion-de-las-discograficas-con-tiktok-no-me-dejan-sacar-mi-proxima-cancion-si-no-se-hace-viral/>

Bad Gyal. (2023a). *Biografia Bad Gyal a Spotify*.

Bad Gyal. (2023b). *Bad Gyal - YouTube*. <https://www.youtube.com/@BadGyal>

Billboard. (s.d.). *Bad Gyal Signs With Interscope and Aftercluv*. Recuperat 14 maig 2023, de <https://www.billboard.com/pro/bad-gyal-signs-with-interscope-and-aftercluv-exclusive/>

Brent Barnhart. (2023). Social listening: Your launchpad to success on social media. *Sprout Social*. <https://sproutsocial.com/insights/social-media-listening/#tools>

francisco-perez-sorrosal. (2023). *distilbert-base-uncased-finetuned-with-spanish-tweets-clf · Hugging Face*. <https://huggingface.co/francisco-perez-sorrosal/distilbert-base-uncased-finetuned-with-spanish-tweets-clf>

Gen Playz - RTVE. (2023). *TikTok e industria musical: ¿Son los artistas esclavos del algoritmo? - YouTube*. [https://www.youtube.com/watch?v=V-TbLPH2Mrk&t=1675s&ab\\_channel=playz](https://www.youtube.com/watch?v=V-TbLPH2Mrk&t=1675s&ab_channel=playz)

Gephi. (2023). *Gephi - The Open Graph Viz Platform*. <https://gephi.org/>

Google. (2023). *YouTube Data API | Google Developers*.  
<https://developers.google.com/youtube/v3>

Google Cloud console. (2023). *Google Cloud console*.  
<https://console.cloud.google.com/getting-started>

Hugging Face. (2023). *Hugging Face – The AI community building the future*.  
<https://huggingface.co/>

IAB Spain. (2022). *Estudio Anual de Redes Sociales 2022*.

IP Mark. (2021). *La audiencia diaria de música en streaming llega al 18% de la población*. <https://ipmark.com/la-audiencia-diaria-de-musica-en-streaming-llega-al-18-de-la-poblacion/>

jorgeortizfuentes. (2023). *spanish\_hate\_speech · Hugging Face*.  
[https://huggingface.co/jorgeortizfuentes/spanish\\_hate\\_speech](https://huggingface.co/jorgeortizfuentes/spanish_hate_speech)

Kantar. (2023). *The power of TikTok*.  
<https://www.kantar.com/uki/inspiration/advertising-media/the-power-of-tiktok>

La Vanguardia. (2022). *TikTok ya es la mejor plataforma para descubrir música*.  
<https://www.lavanguardia.com/tecnologia/aplicaciones/20220602/8308148/tiktok-mejor-plataforma-descubrir-musica-artistas-les-guste-pmv.html>

Lil Nas X Talks «Old Town Road» and the Billboard Controversy | Time. (s.d.).  
Recuperat 6 abril 2023, de <https://time.com/5561466/lil-nas-x-old-town-road-billboard/>

LOS 40. (2021). *Bad Gyal anuncia nueva música: “Voy a lanzar temas que lleváis pidiendo tres años”*.  
[https://los40.com/los40/2021/11/19/los40urban/1637334109\\_212598.html](https://los40.com/los40/2021/11/19/los40urban/1637334109_212598.html)

LOS 40. (2023). *Bad Gyal aclara si hay sesión con Bizarrap: “Me ha llamado”*.  
[https://los40.com/los40/2023/03/01/los40urban/1677682085\\_345132.html](https://los40.com/los40/2023/03/01/los40urban/1677682085_345132.html)

Mata Valldaura, A. (2022). *Spotify com a plataforma digital: anàlisi de la posició de la figura de l'artista en el nou context social. Pot ser Spotify una opció viable per reproduir contingut musical, essent una eina potencial i d'importància pels artistes com a plataforma digital en streaming que reproduceix contingut musical?*  
<https://dugi-doc.udg.edu/handle/10256/22198>

Mixed Analytics. (2023). *List of YouTube Category IDs*.  
<https://mixedanalytics.com/blog/list-of-youtube-video-category-ids/>

MMG. (2023). *xlm-roberta-base-sa-spanish · Hugging Face*.

<https://huggingface.co/MMG/xlm-roberta-base-sa-spanish?text=Te+quiero.+Te+amo.>

Pandas. (2023). *Pandas 2.0.1 documentation*. <https://pandas.pydata.org/pandas-docs/stable/#>

PROMUSICAE. (2023). *MERCADO MUSICA ESPAÑA 2022*.

<https://www.promusicae.es/descarga-informes/mercado-de-la-musica-grabada-en-espana-2022-n252/>

Rosalía. (2022). *ROSALÍA - CHICKEN TERIYAKI (Official Video) - YouTube*.

[https://www.youtube.com/watch?v=OG4gq9fCoRE&t=133s&ab\\_channel=RosalíaVEVO](https://www.youtube.com/watch?v=OG4gq9fCoRE&t=133s&ab_channel=RosalíaVEVO)

Sala de prensa de TikTok. (2021). *Un nuevo estudio revela cómo TikTok está cambiando la industria de la música para las marcas, los artistas, el público y las compañías*. <https://newsroom.tiktok.com/es-es/estudio-revela-como-tiktok-cambiando-industria-musica-marcas-artistas-publico-companias>

Spotify. (2023a). *¿Qué es Spotify?* <https://support.spotify.com/es/article/what-is-spotify/>

Spotify. (2023b). *Web API | Spotify for Developers*.

<https://developer.spotify.com/documentation/web-api>

Spotipy. (2023). *Welcome to Spotipy! — spotipy 2.0 documentation*.

<https://spotipy.readthedocs.io/en/2.22.1/#>

Stack Overflow. (2023). *Stack Overflow - Where Developers Learn, Share, & Build Careers*. <https://stackoverflow.com/>

Statista. (2022). *Spain: social network penetration 2010-2021*.

<https://www.statista.com/statistics/312935/social-network-penetration-in-spain/>

Statista. (2023a). *Global music streaming subscribers 2022*.

<https://www.statista.com/statistics/653926/music-streaming-service-subscriber-share/>

Statista. (2023b). *Industria de música: ingresos anuales en España 2009-2026.*

<https://es.statista.com/estadisticas/472604/prevision-de-facturacion-de-la-industria-de-la-musica-en-espana/>

Statista. (2023c). *Spotify: number of premium subscribers worldwide 2022.*

<https://www.statista.com/statistics/244995/number-of-paying-spotify-subscribers/>

Tableau. (2023). *What Is Tableau?* <https://www.tableau.com/why-tableau/what-is-tableau>

UAB Barcelona. (2023). *La indústria del videoclip debat el seu futur davant de la irrupció de TikTok - Facultat de Ciències de la Comunicació - UAB Barcelona.*

<https://www.uab.cat/web/detall-de-noticia/la-industria-del-videoclip-debat-el-seu-futur-davant-de-la-irrupcio-de-tiktok-1345702086369.html?noticiaid=1345882515874>

Universal Musica. (2023). *BAD GYAL - Universal Musica.*

<https://www.universalmusica.com/artist/bad-gyal/>

VEVO. (2023). *The World's Leading Music Video Network | About.*

<https://www.hq.vevo.com/about>

YouTube. (2023). *YouTube.* <https://www.youtube.com/>

YouTube Data Tools. (2023). *YouTube Data Tools.*

[https://labs.polsys.net/tools/youtube/mod\\_videos\\_list.php](https://labs.polsys.net/tools/youtube/mod_videos_list.php)

# Annexos

A causa de la naturalesa dels annexos utilitzats en aquest treball de fi de grau, podeu consultar els fitxers informàtics d'emmagatzematge de dades i el codi dels programes al meu *repositori* de GitHub personal.

<https://github.com/Jaume14/TFG-BadGyal>

