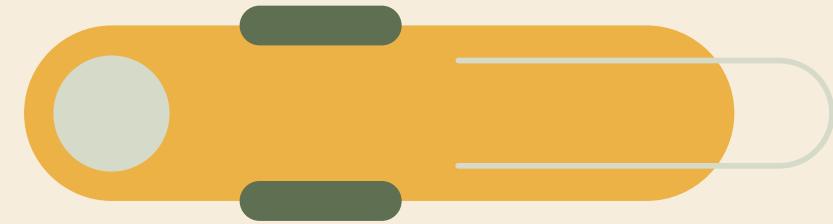




## § Nhóm 5

# Thị trường âm nhạc Việt Nam

2025



# Mục lục

01

Thành viên nhóm

02

Tổng quan dữ liệu

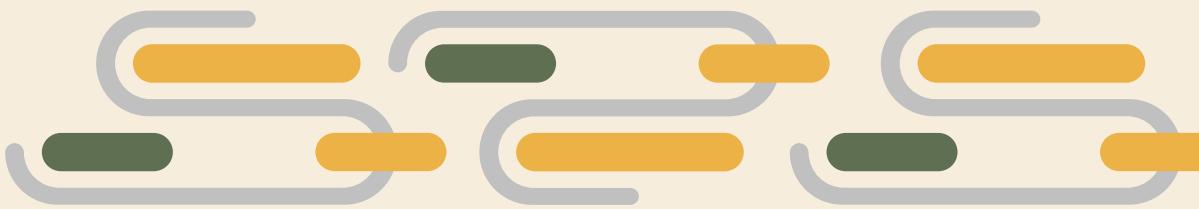
03

Khám phá dữ liệu

04

Dashboard

# Thành viên nhóm



**22127216 - Nguyễn Đình Kiên**

**22127257 - Phạm Minh Mẫn**

**22127148 - Dương Nhật Huy**

**22127224 - Trương Thuận Kiệt (Leader)**

**22127492 - Hồ Đăng Phúc**

# Tổng quan dữ liệu



S

# Nguồn gốc dữ liệu



# Vietnamese\_only\_songs\_cleaned dataset

Cột	Ý nghĩa
<b>Id</b>	Mã Spotify của mỗi bài hát.
<b>Album</b>	Tên album mà bài hát xuất hiện
<b>Album_type</b>	Album_type Loại album. Giá trị cho phép: "album", "single", "compilation".
<b>Album_popularity</b>	Điểm đánh giá trung bình của phim (0 - 100)
<b>Artists</b>	Tên các nghệ sĩ biểu diễn bài hát
<b>Disc_number</b>	Disc_number

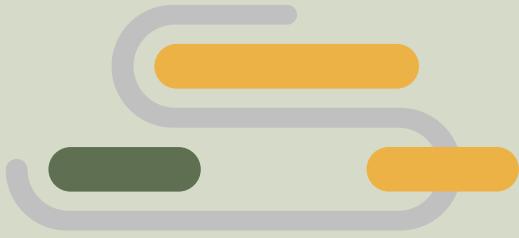
# Vietnamese\_only\_songs\_cleaned dataset

Cột	Ý nghĩa
Duration_ms	Độ dài của bài hát.
Explicit	Bài hát có chứa lời bài hát tục tĩu hay không.
Name	Tên bài hát.
Popularity	Điểm đánh giá trung bình của phim (0 - 100)
vote_count	Độ phổ biến của bài hát (0 - 100)
Track_number	Số thứ tự của bài hát trong album. Nếu album có nhiều đĩa, số thứ tự bài hát là số trên đĩa được chỉ định

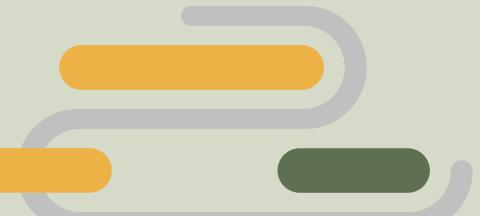
# Vietnamese\_only\_songs\_cleaned dataset

Cột	Ý nghĩa
<b>Release_date</b>	Ngày album được phát hành lần đầu tiên.
<b>Genres</b>	Thể loại của bài hát.
<b>Stream_counts</b>	Tổng số lần bài hát được phát trên Spotify

# **position\_streams\_by\_time\_daily\_total\_peak\_cleaned & position\_streams\_by\_time\_weekly\_total\_peak\_cleaned dataset**



Cột	Ý nghĩa
<b>Id</b>	Mã Spotify của từng bài hát.
<b>TotalStream</b>	Tổng số lượt phát theo tuần/ngày.
<b>PeakStream</b>	Lượt phát cao nhất theo tuần/ngày.
<b>PeakPosition</b>	Vị trí cao nhất theo tuần/ngày.



# Vietnamese\_artists dataset

Cột	Ý nghĩa
<b>Id</b>	Mã Spotify của từng nghệ sĩ.
<b>Name</b>	Tên của nghệ sĩ.
<b>Followers</b>	Số lượt theo dõi của nghệ sĩ.
<b>Genres</b>	Thể loại âm nhạc của nghệ sĩ.

# Khám phá dữ liệu



S

# Vietnamese\_only\_songs\_cleaned dataset

Cột	Số dòng không trống	Kiểu dữ liệu
album	4846	object
album_type	4846	object
album_popularity	4846	int
artists	4846	object
disc_number	4846	int
duration_ms	4846	int
explicit	4846	bool
id	4846	object

# Vietnamese\_only\_songs\_cleaned dataset

Cột	Số dòng không trống	Kiểu dữ liệu
name	4846	object
popularity	4846	int
track_number	4846	int
release_date	4846	object
genres	4846	object
stream_counts	4817	float

# Vietnamese\_only\_songs\_cleaned dataset

	album_popularity	disc_number	duration_ms	popularity	track_number	stream_counts
count	4846.00	4846.000000	4.846000e+03	4846.000000	4846.000000	4.817000e+03
mean	25.919934	1.003714	2.381399e+05	27.985349	3.120718	4.121705e+06
std	15.573992	0.060839	6.362212e+04	14.618266	3.771611	1.772154e+07
min	0.000000	1.000000	1.952800e+04	0.000000	1.000000	1.027000e+03
25%	14.000000	1.000000	1.990928e+05	18.000000	1.000000	1.545430e+05
50%	25.000000	1.000000	2.328505e+05	27.000000	1.000000	6.609920e+05
75%	37.000000	1.000000	2.725415e+05	38.000000	4.000000	2.770964e+06
max	83.000000	2.000000	1.813776e+06	80.000000	41.000000	8.532404e+08

# position\_streams\_by\_time\_daily\_total\_peak\_cleaned

Cột	Số dòng không trống	Kiểu dữ liệu
id	772	object
TotalStreamCounts	771	float
PeakStreamCounts	771	float
PeakPosition	771	float

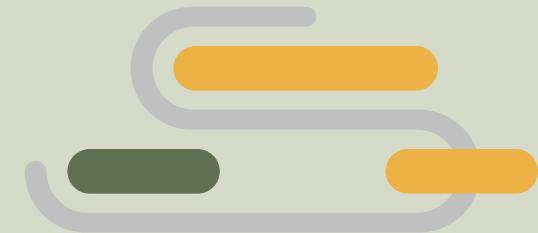
# position\_streams\_by\_time\_weekly\_total\_peak\_cleaned

Cột	Số dòng không trống	Kiểu dữ liệu
id	726	object
TotalStreamCounts	725	float
PeakStreamCounts	725	float
PeakPosition	724	float

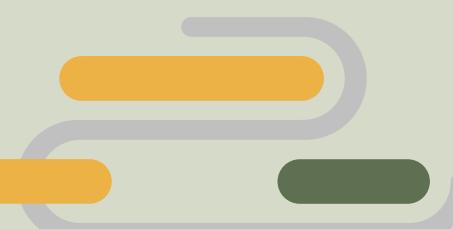
# position\_streams\_by\_time\_weekly\_total\_peak\_cleaned

	<b>TotalStreamCounts</b>	<b>PeakStreamCounts</b>	<b>PeakPosition</b>
count	7.710000e+02	7.710000e+02	771.000000
mean	5.851733e+06	2.680038e+05	66.584955
std	9.287612e+06	3.178791e+05	58.811201
min	4.698000e+03	4.698000e+03	1.000000
25%	2.867365e+05	9.259600e+04	12.000000
50%	1.499057e+06	1.797870e+05	51.000000
75%	6.552288e+06	3.087205e+05	111.000000
max	5.269765e+07	3.445288e+06	200.000000

# position\_streams\_by\_time\_daily\_total\_peak\_cleaned



	<b>TotalStreamCounts</b>	<b>PeakStreamCounts</b>	<b>PeakPosition</b>
count	7.250000e+02	725.000000	725.000000
mean	5.583042e+06	43500.244138	50.513103
std	9.160756e+06	47062.807284	45.455326
min	1.704000e+03	1364.000000	1.000000
25%	3.157060e+05	16963.000000	9.000000
50%	1.464363e+06	31830.000000	39.000000
75%	6.013064e+06	50816.000000	83.000000
max	5.294350e+07	455604.000000	189.000000



# Vietnamese\_artists

Cột	Số dòng không trống	Kiểu dữ liệu
Id	1115	object
Name	1115	object
Followers	1115	int
Genres	1115	object

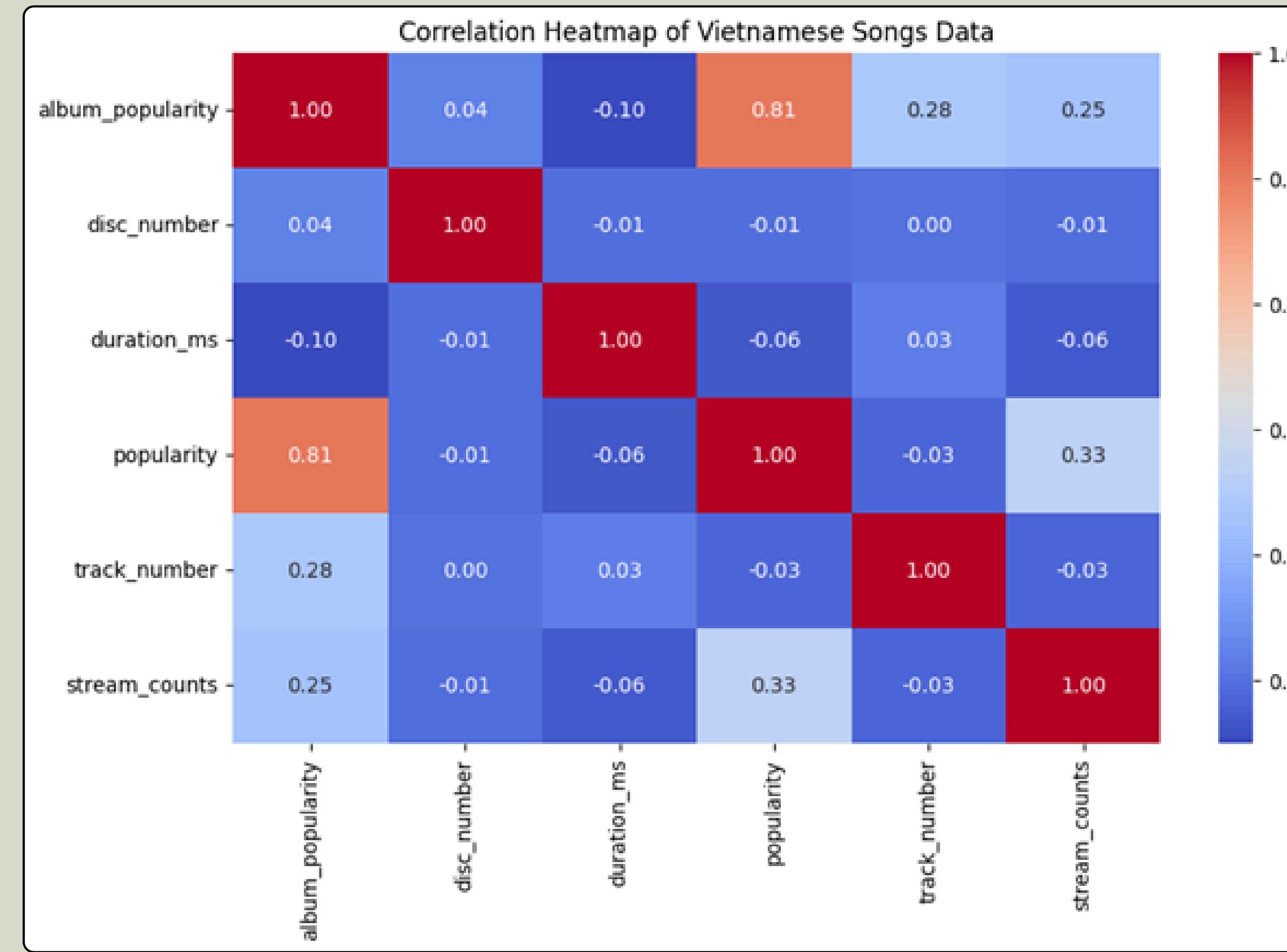
# Vietnamese\_artists

	<b>popularity</b>	<b>followers</b>
count	1115.000000	1.115000e+03
mean	30.951570	8.933268e+04
std	14.979501	3.273770e+05
min	0.000000	0.000000e+00
25%	20.000000	6.185000e+02
50%	32.000000	4.320000e+03
75%	42.000000	4.008950e+04
max	68.000000	6.549380e+06

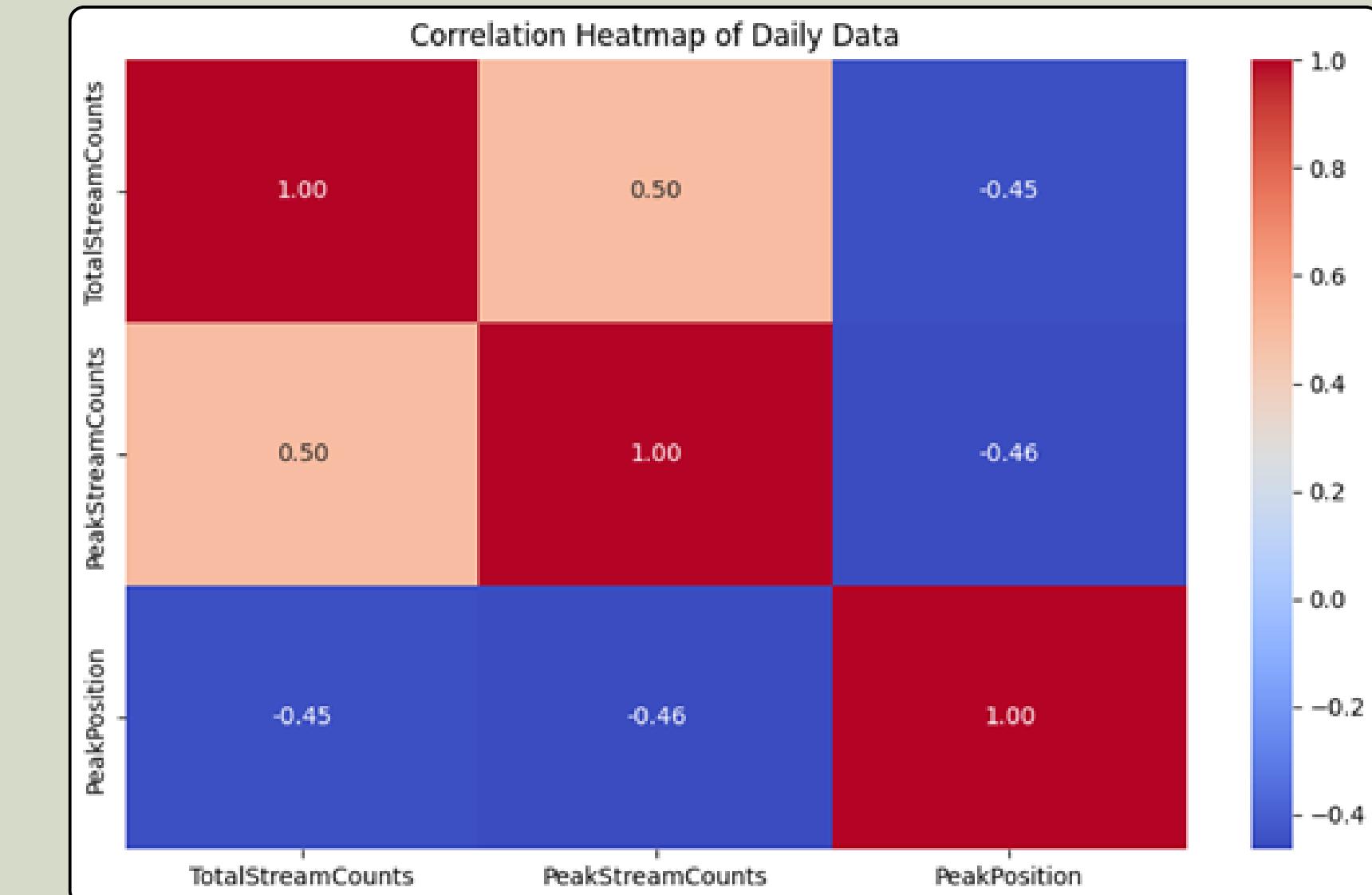
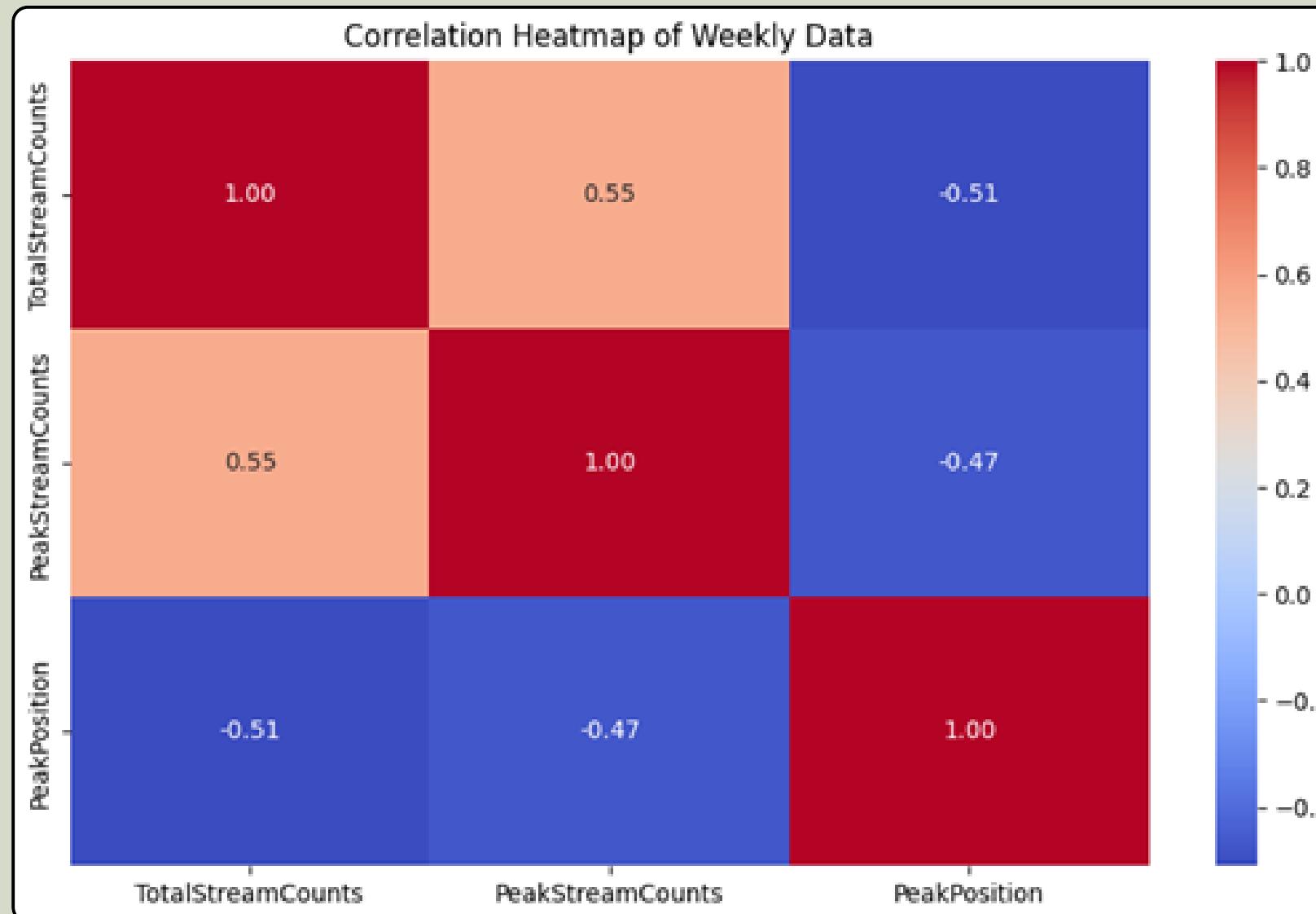
# Mối tương quan giữa các biến



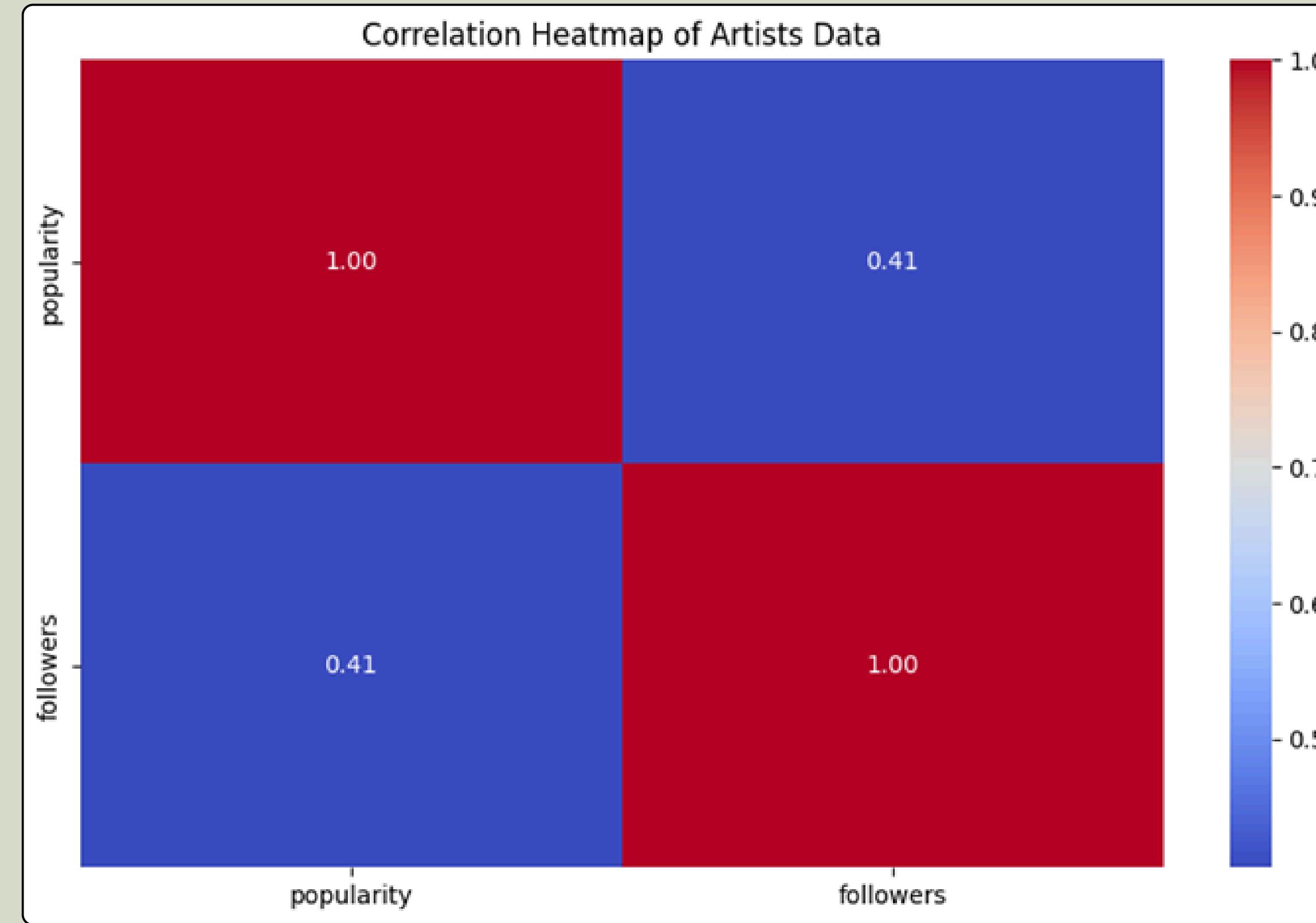
# Vietnamese\_only\_songs\_cleaned



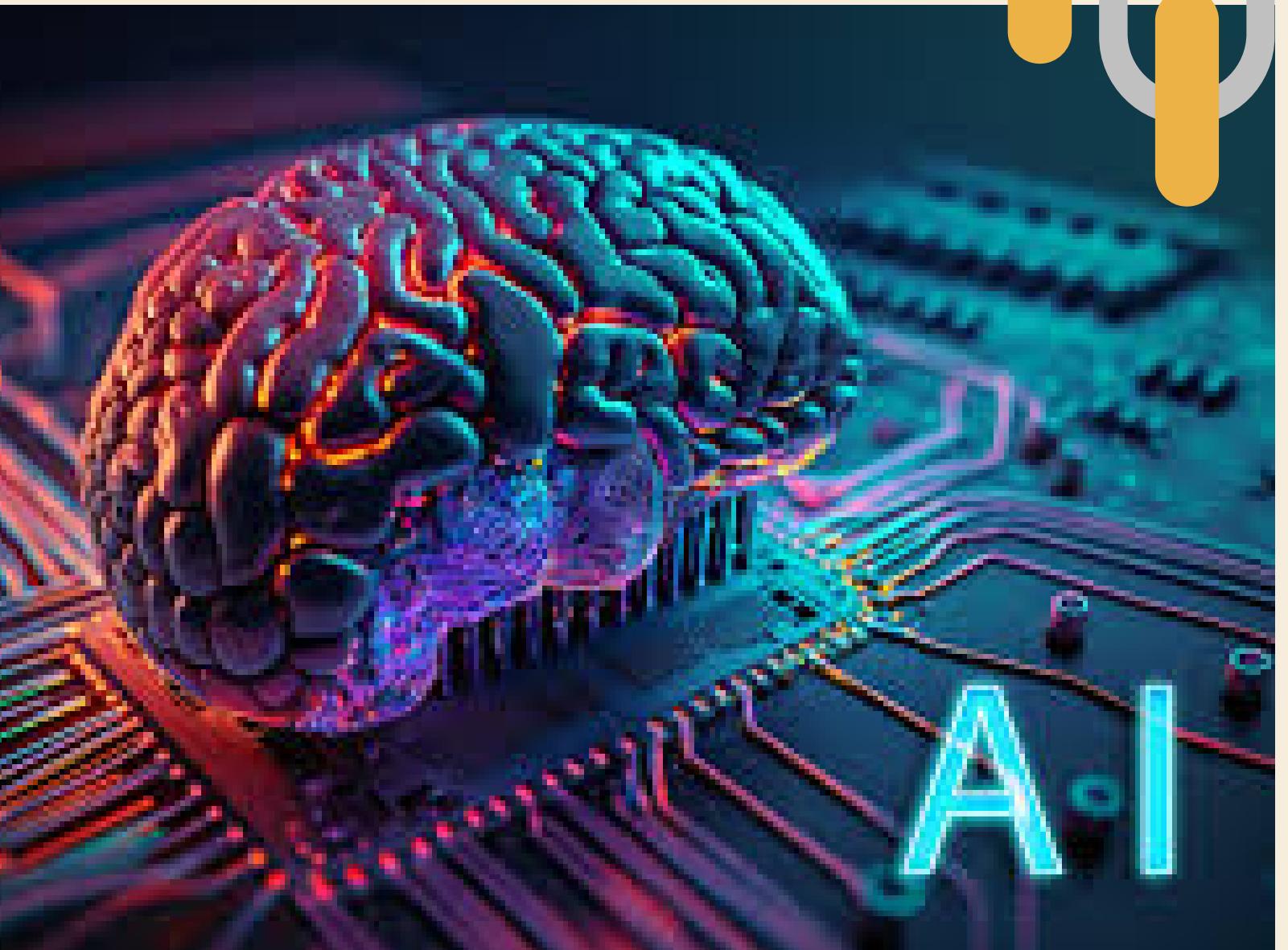
# position\_streams\_by\_time\_daily\_total\_peak\_cleaned và position\_streams\_by\_time\_weekly\_total\_peak\_cleaned



# Vietnamese\_artist



# Tích hợp AI



S



i need to build dashboard on these dataset can you recommend how many pages should i have and each page should have what

vietnamese\_artists.csv

vietnamese\_songs\_full.csv

#### Q Thoughts

Expand for details

To build an effective dashboard based on the provided datasets, which include artist data (e.g., name, popularity, followers, genres) and track data (e.g., album, artists, popularity, stream counts, genres), I recommend structuring it across **four distinct pages**. Each page will focus on a specific aspect of the data, providing clear insights for different user needs, such as understanding artist popularity, track performance, genre trends, and comparative analyses. Below, I outline the recommended pages and their content.

How can Grok help?



DeepSearch



Think

Grok 3



## Page 1: Artist Overview

**Purpose:** Provide a high-level summary of artists, their popularity, and their audience reach.

**Content:**

- **Total Artists:** A single-value card displaying the total number of unique artists in the dataset (e.g., count of distinct `id` values in the artist dataset).
- **Top Artists by Popularity:** A bar chart or table listing the top 10 artists by `popularity`, including their `name`, `popularity` score, and `followers`. For example, artists like Sơn Tùng M-TP (popularity 64) or HIEUTHUHAI (popularity 67) could appear here.
- **Follower Distribution:** A histogram or box plot showing the distribution of `followers` across all artists, highlighting the range (e.g., from 51 for Thien Dung to 6,425,018 for Sơn Tùng M-TP) and identifying outliers.
- **Genre Breakdown:** A pie chart or treemap displaying the frequency of genres across artists (e.g., `v-pop`, `vinahouse`, `vietnamese hip hop`), derived from the `genres` column. This helps identify dominant music styles.
- **Filters:** Interactive filters for `popularity` range, `followers` range, and `genres` to allow users to drill down into specific artist segments.

**Why:** This page serves as the entry point, giving users a quick snapshot of the artist landscape and enabling them to identify key players and trends.

How can Grok help?



DeepSearch



Think

Grok 3



## Page 2: Track Performance

Purpose: Analyze individual tracks, their popularity, and streaming success.

### Content:

- **Total Tracks:** A single-value card showing the total number of unique tracks (e.g., count of distinct `id` values in the track dataset).
- **Top Tracks by Streams:** A bar chart or table of the top 10 tracks by `stream_counts`, including `name`, `artists`, `popularity`, and `stream_counts`. For instance, "Blinding Lights" by The Weeknd with 4,293,744,430 streams would rank high.
- **Popularity vs. Streams Scatter Plot:** A scatter plot with `popularity` on the x-axis and `stream_counts` on the y-axis, with points sized by `duration_ms` or colored by `album_type` (e.g., single, album). This visualizes the relationship between popularity and streaming success.
- **Release Date Trends:** A line chart plotting the number of tracks or total `stream_counts` over time based on `release_date`, showing peaks (e.g., 2020 releases like "Dynamite" by BTS).
- **Filters:** Options to filter by `artists`, `album_type`, `release_date` range, and `genres` for focused analysis.

Why: This page focuses on track-level insights, crucial for understanding what content drives engagement and how performance varies across releases.

How can Grok help?



DeepSearch



Think

Grok 3



## Page 3: Genre Insights

Purpose: Explore the distribution and performance of music genres across artists and tracks.

### Content:

- **Genre Popularity:** A bar chart showing average `popularity` per genre, calculated by aggregating artist and track `popularity` scores linked to each genre (e.g., averaging `v-pop` artists like Vũ. and AMEE).
- **Genre Stream Counts:** A stacked bar chart or table displaying total `stream_counts` by genre, derived from tracks, with contributions from top tracks highlighted (e.g., `k-pop` streams from BTS's "Dynamite").
- **Artist Count per Genre:** A bar chart or heatmap showing the number of artists associated with each genre, revealing genre saturation (e.g., many artists in `vinahouse` like Orinn and Anh Quân Idol).
- **Genre Overlap:** A Sankey diagram or chord diagram illustrating how artists and tracks span multiple genres (e.g., artists with both `v-pop` and `vietnamese hip hop`), based on the `genres` lists.
- **Filters:** Filters for specific genres and a toggle to switch between artist-based and track-based views.

Why: Genres are a key dimension in music data, and this page helps users understand their prevalence and impact, useful for market analysis or content planning.

How can Grok help?



DeepSearch



Think

Grok 3



## **Page 4: Comparative Analysis**

**Purpose:** Enable side-by-side comparisons of artists and tracks for deeper insights.

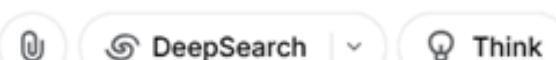
## Content:

- **Artist Comparison:** A dual-axis chart or parallel coordinates plot comparing selected artists across metrics like `popularity`, `followers`, and count of tracks (linked via the `artists` field in the track dataset). Users can select artists (e.g., Đen vs. Sơn Tùng M-TP).
  - **Track Comparison:** A table or radar chart comparing selected tracks on `popularity`, `stream_counts`, `duration_ms`, and `explicit` status. For example, compare "Đi Về Nhà" by Đen and JustaTee with "CUA" by MANBO and HIEUTHUHAI.
  - **Artist-Track Linkage:** A network graph or table showing which artists contribute to which tracks, with metrics like average track `popularity` per artist, leveraging the overlap between datasets.
  - **Custom Filters:** Dropdowns to select specific artists and tracks, with dynamic updates to visuals.

**Why:** This page caters to users needing detailed comparisons, such as marketers or analysts benchmarking performance across entities.

### **Summary of Recommendation**

## How can Grok help?



strok 3 ✓



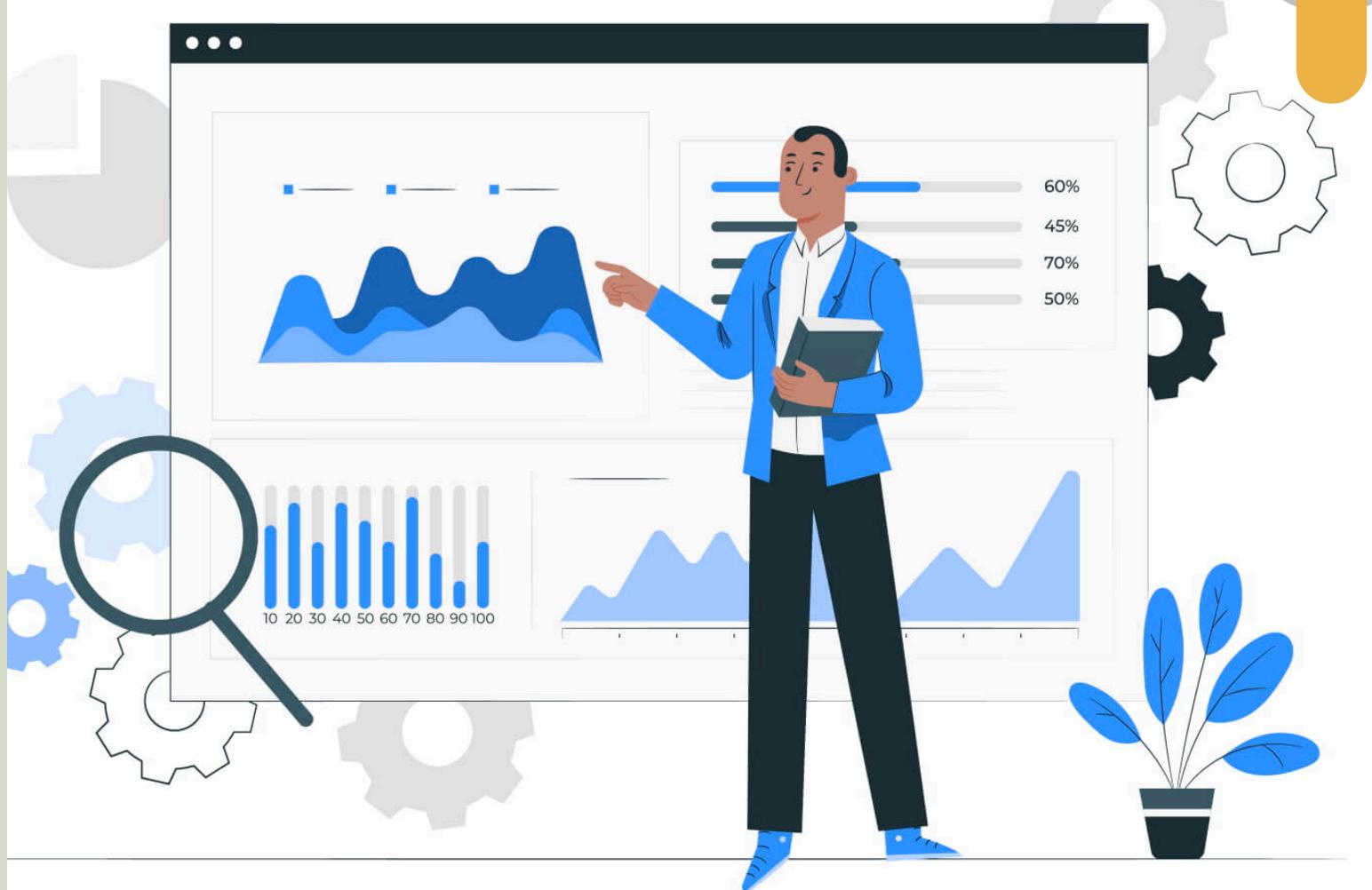
## **Summary of Recommendation**

- **Number of Pages:** 4
    - 1. **Artist Overview** - Summary of artists, popularity, and genres.
    - 2. **Track Performance** - Focus on track metrics and streaming trends.
    - 3. **Genre Insights** - Genre distribution and performance analysis.
    - 4. **Comparative Analysis** - Detailed artist and track comparisons.

**Rationale:** Four pages strike a balance between comprehensive coverage and usability, aligning with common dashboard goals: overview, detail, segmentation, and comparison. Each page leverages the datasets' key fields (`popularity`, `followers`, `genres`, `stream_counts`, etc.) and supports interactivity for exploration, making the dashboard versatile for various stakeholders like music analysts, marketers, or fans.

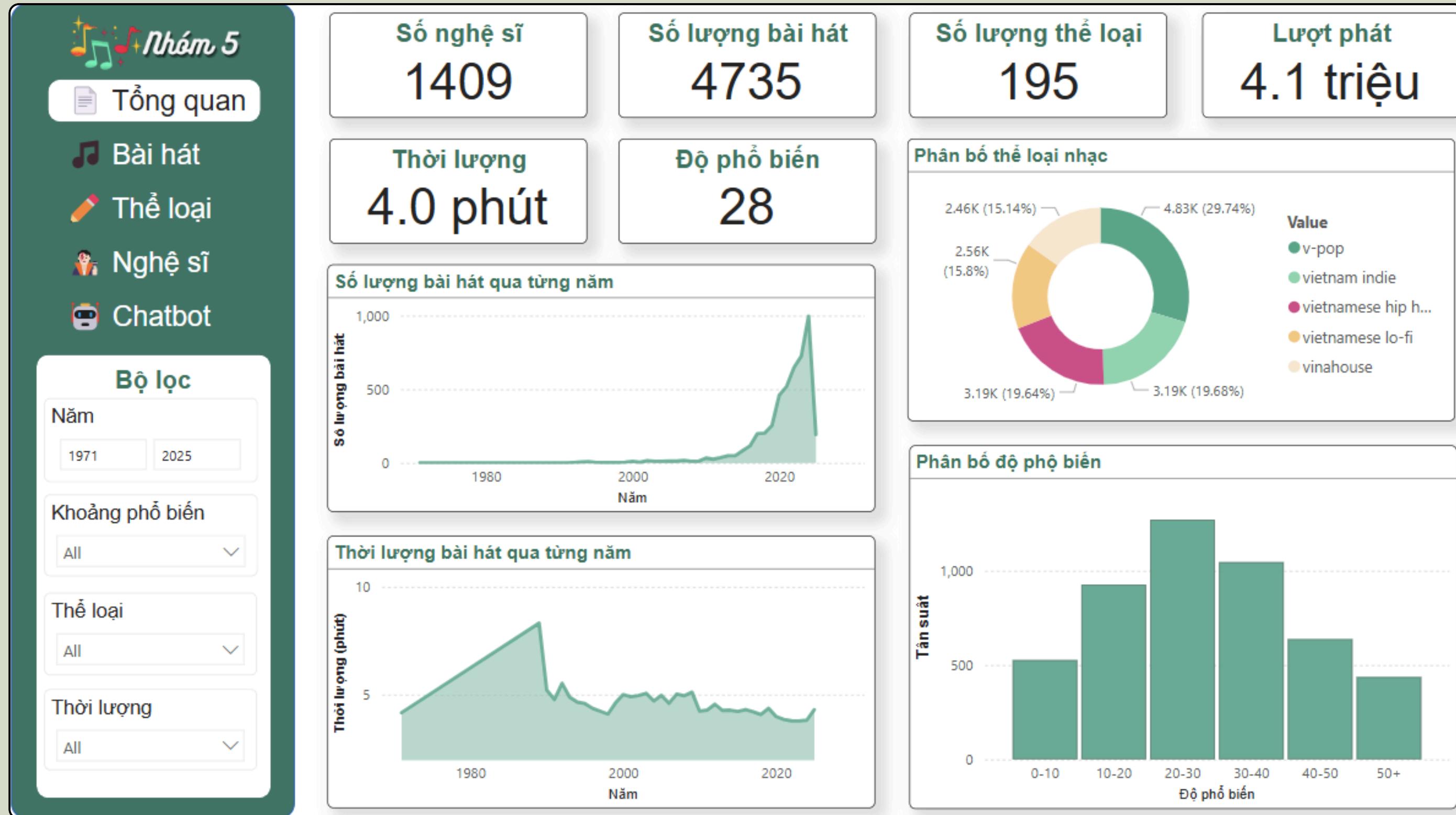
This structure ensures clarity, avoids overwhelming users with too much data on a single page, and provides actionable insights tailored to different aspects of the music ecosystem represented in the datasets.

# Dashboards



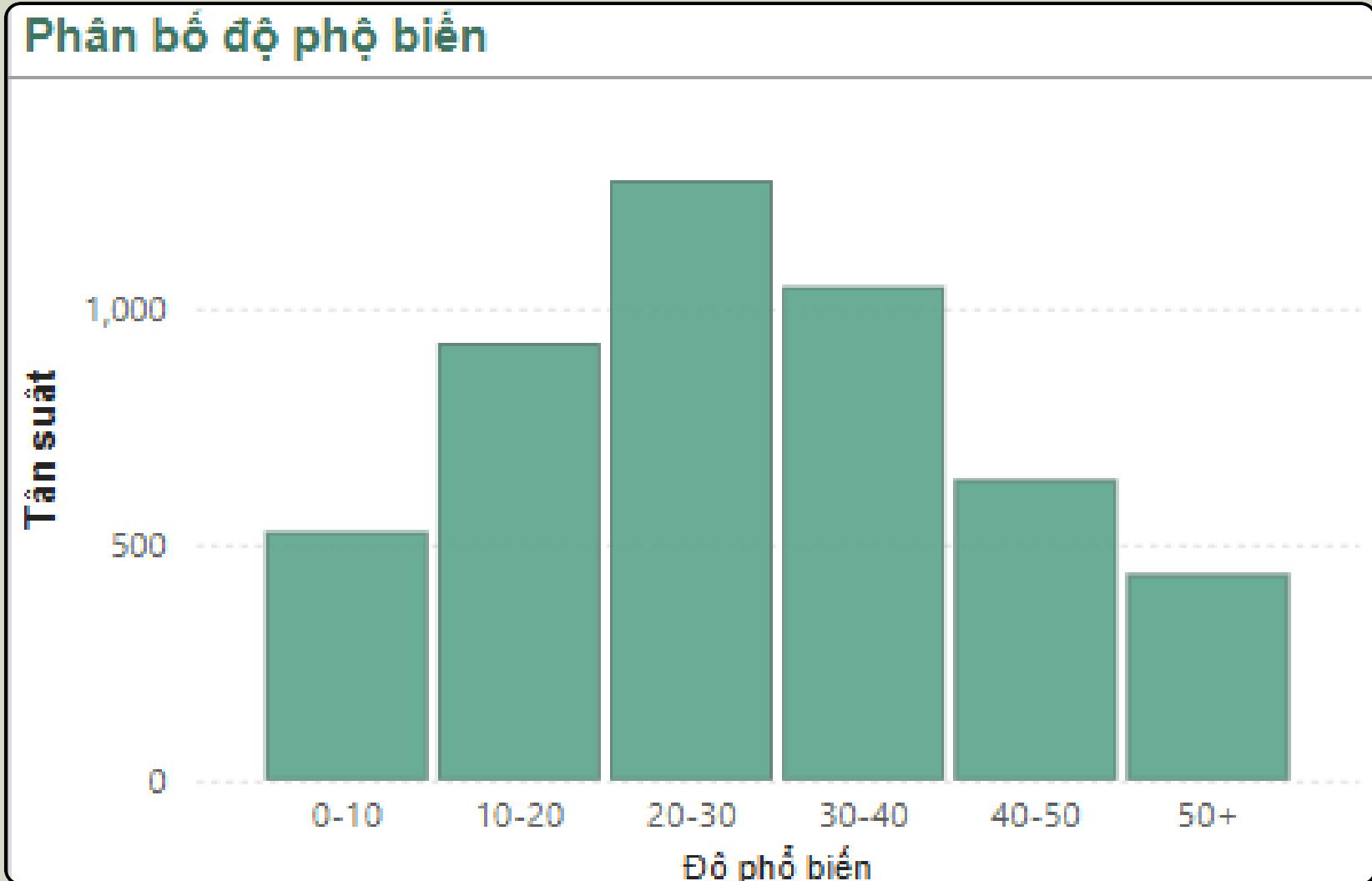
S

# Giao diện Dashboard tổng quan

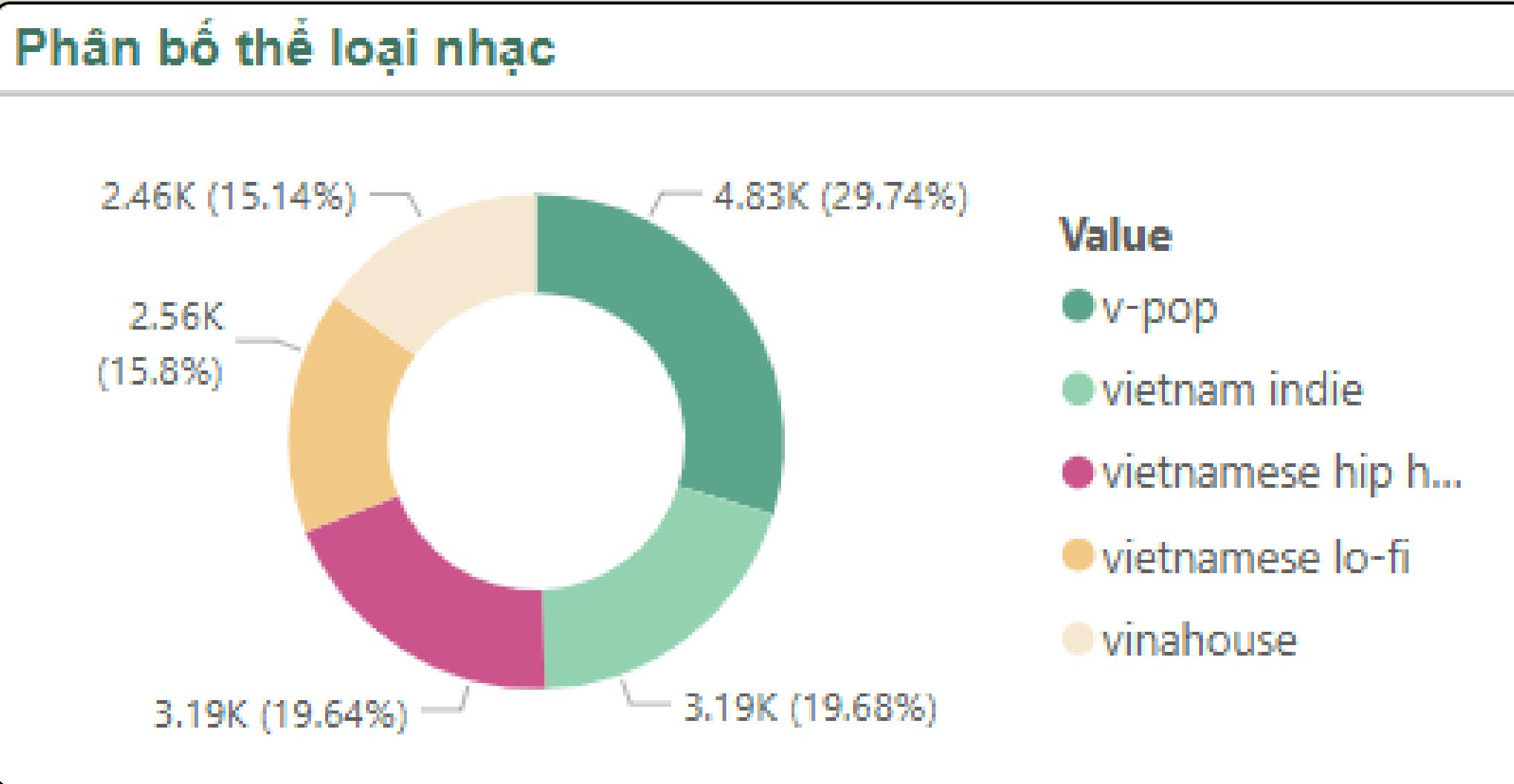


# Các phân bố

Phân bố độ phổ biến

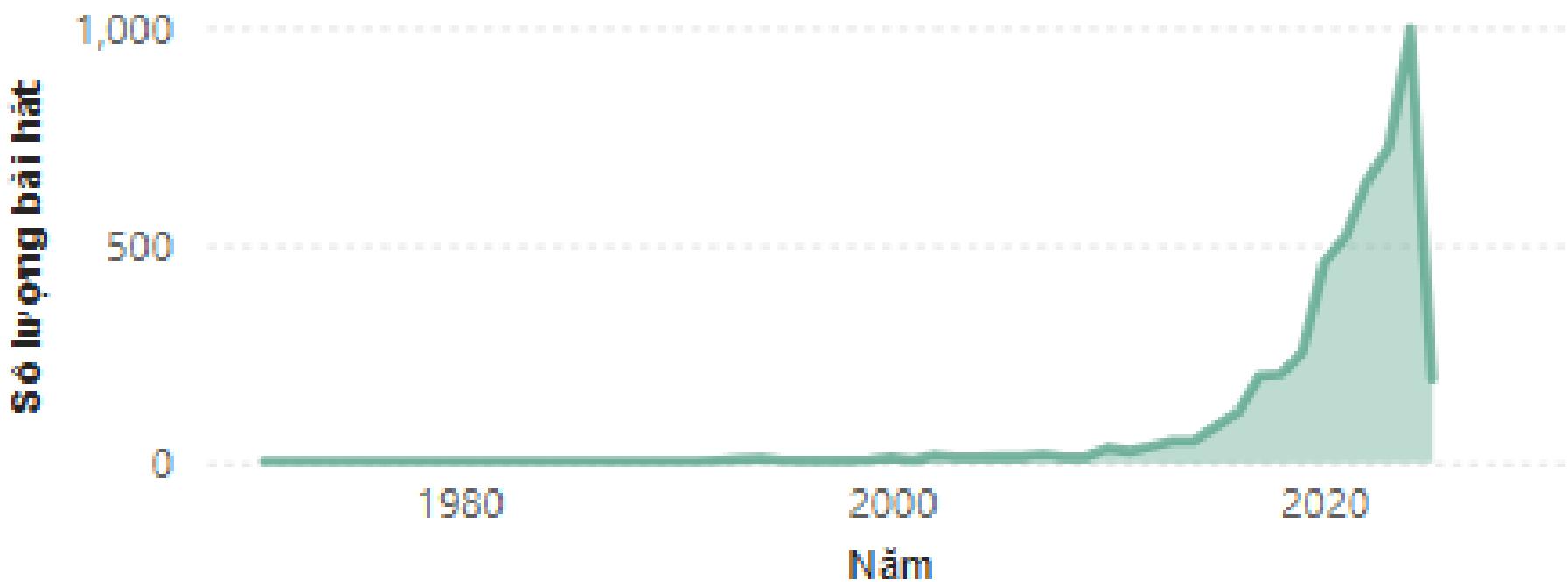


Phân bố thể loại nhạc

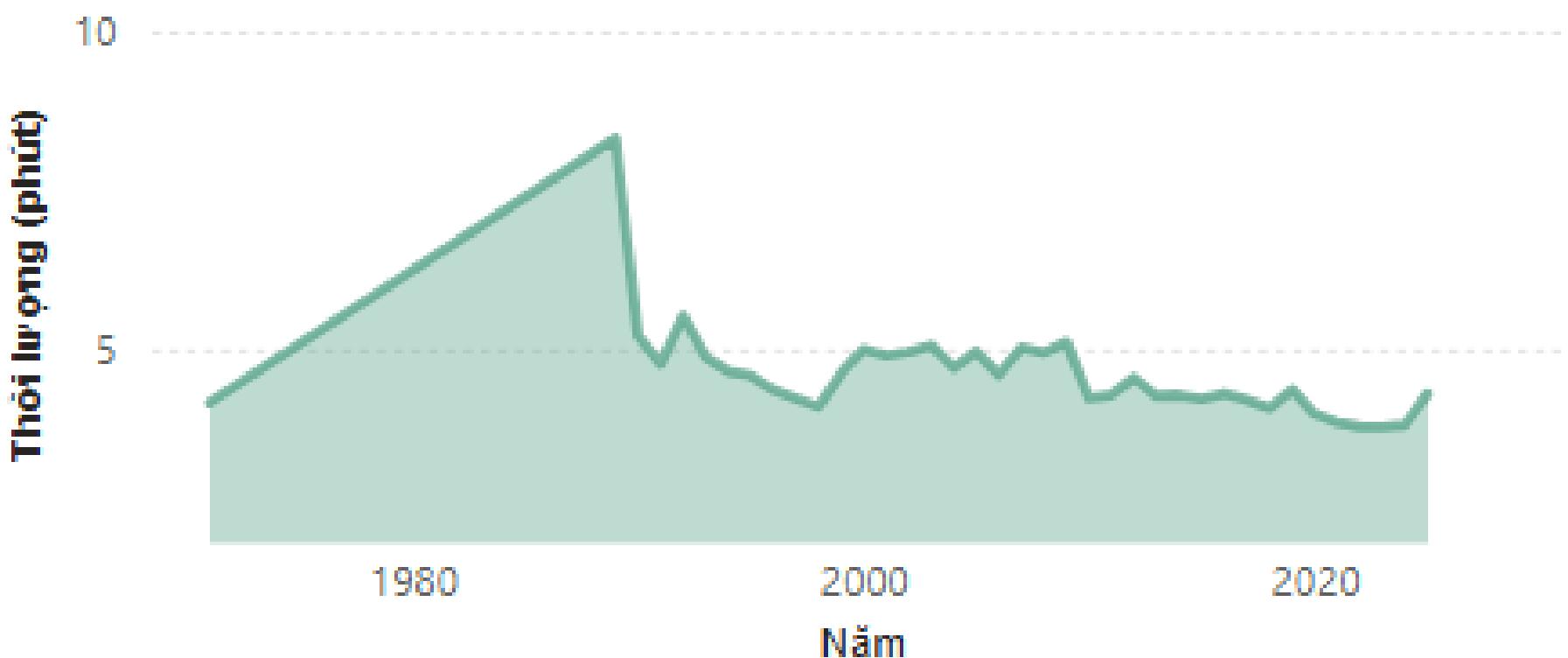


# Các xu hướng...

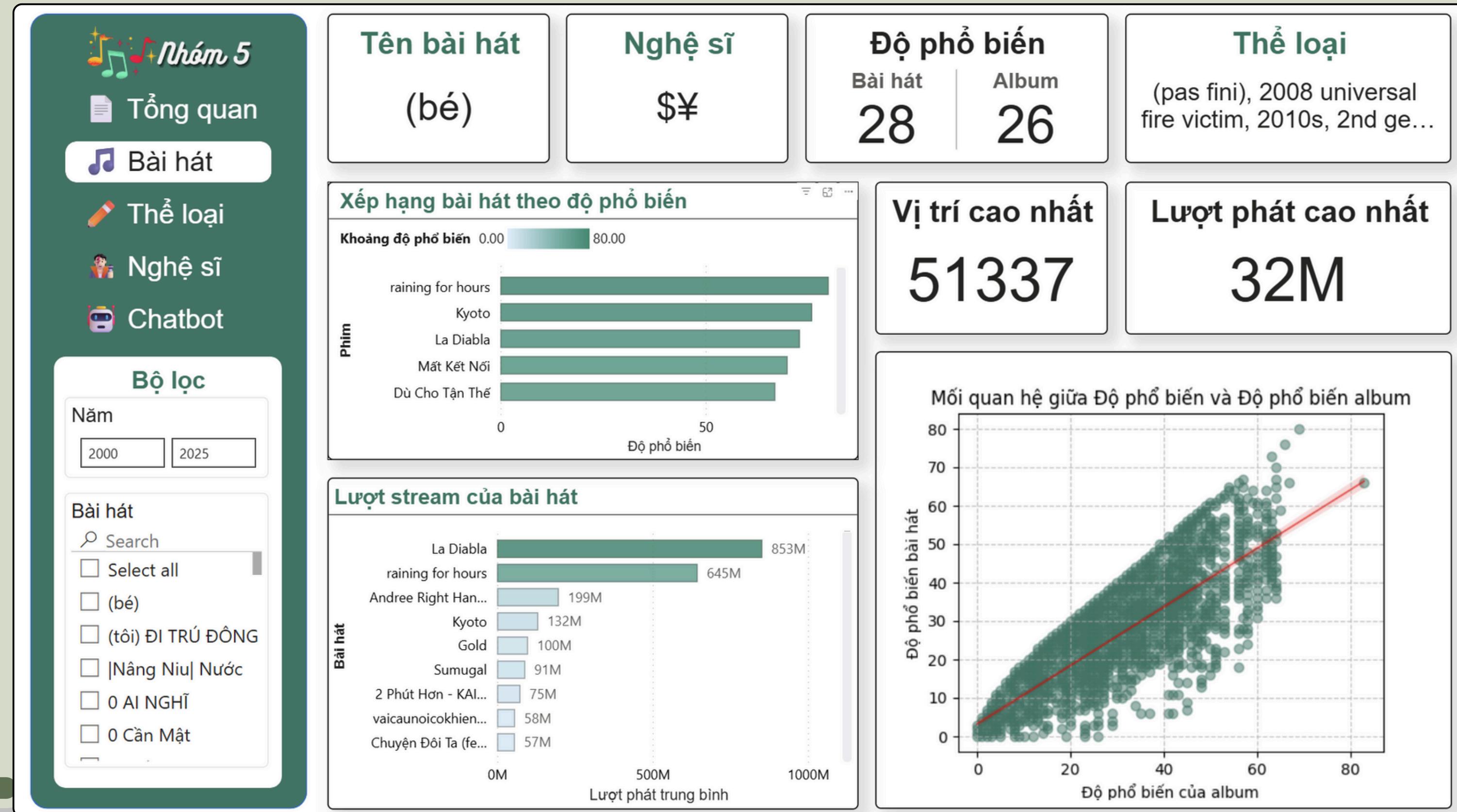
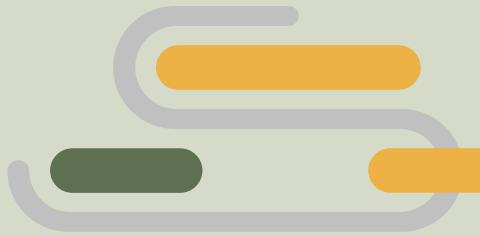
Số lượng bài hát qua từng năm



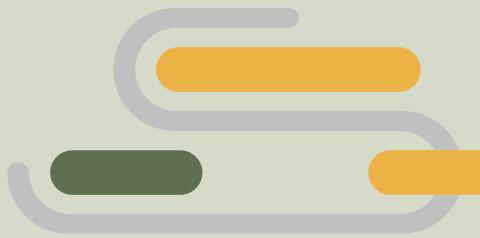
Thời lượng bài hát qua từng năm



# Dashboard bài hát



# “Dù cho tận thế”



**Nhóm 5**

- Tổng quan
- Bài hát
- Thể loại
- Nghệ sĩ
- Chatbot

**Bộ lọc**

Năm

2000 2025

Bài hát

- Dù Cho
- Dù Cho Mai Về S...
- Dù Cho Tận Thế**

**Tên bài hát**  
Dù Cho Tận Thế

**Nghệ sĩ**  
ERIK

**Độ phổ biến**

Bài hát	67
Album	57

**Thể loại**  
vietnam indie, vietnamese lo-fi, vinahouse, v-pop

**Xếp hạng bài hát theo độ phổ biến**

Phim	Độ phổ biến
raining for hours	80.00
Kyoto	~75
La Diabla	~70
Mất Kết Nối	~65
Dù Cho Tận Thế	~60

**Vị trí cao nhất**  
1

**Lượt phát cao nhất**  
347K

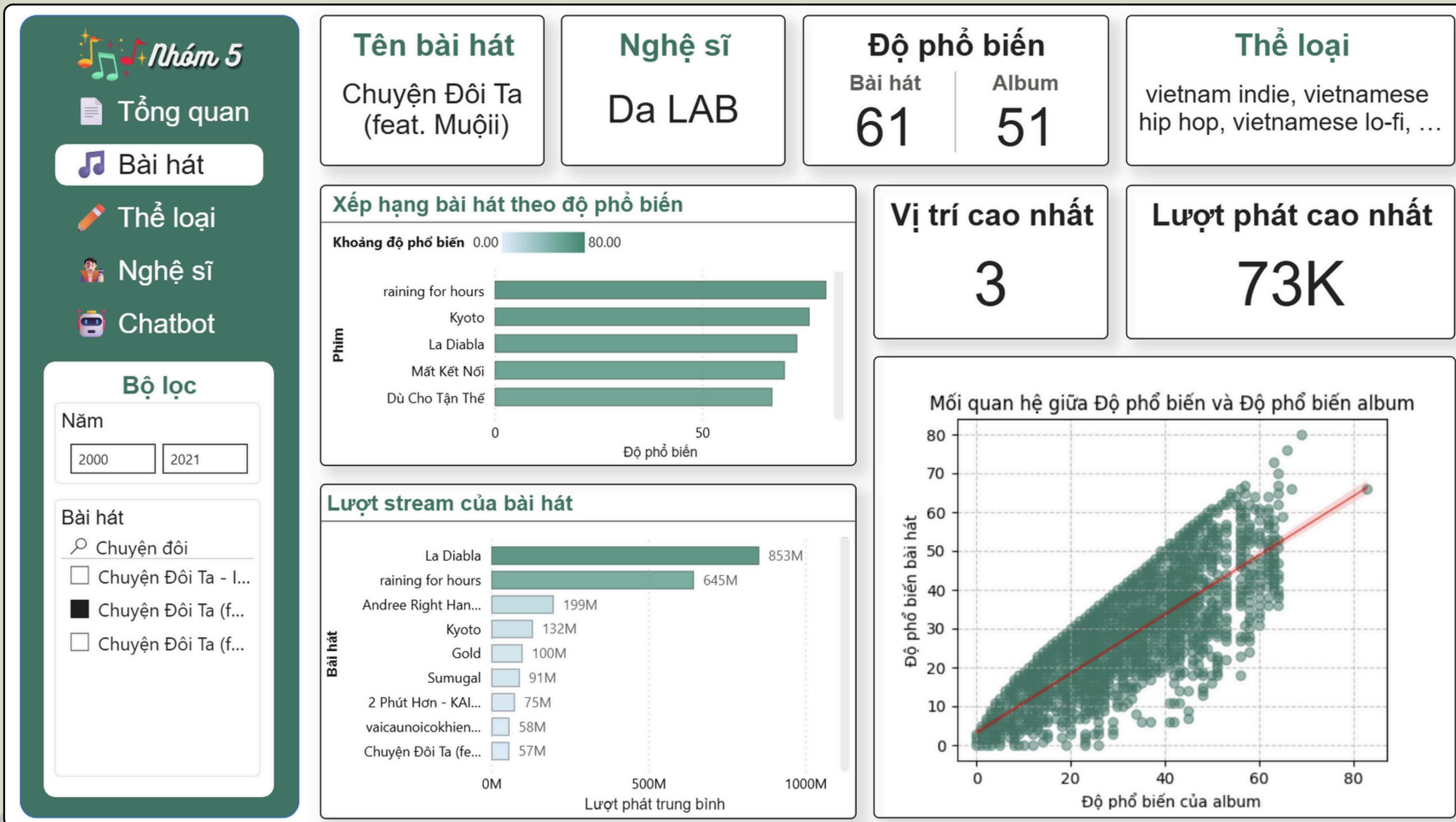
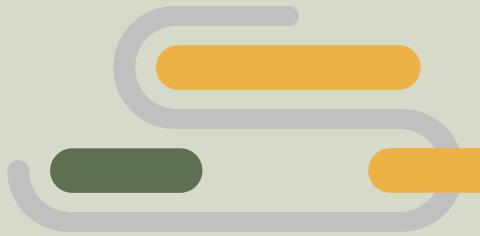
**Lượt stream của bài hát**

Bài hát	Lượt phát trung bình
La Diabla	853M
raining for hours	645M
Andree Right Han...	199M
Kyoto	132M
Gold	100M
Sumugal	91M
2 Phút Hơn - KAI...	75M
vaicaunoicokhien...	58M
Chuyện Đôi Ta (fe...	57M

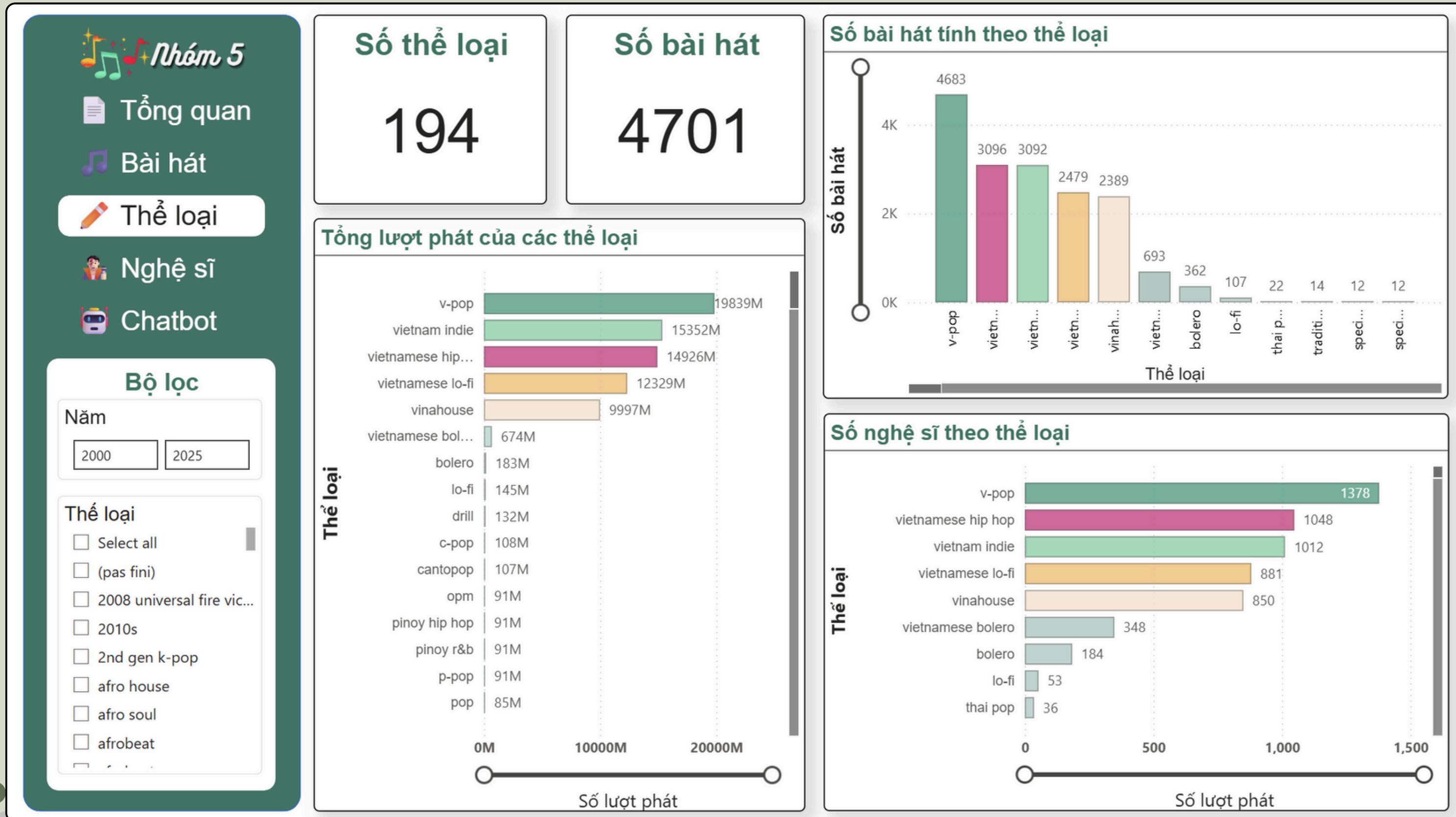
**Mối quan hệ giữa Độ phổ biến và Độ phổ biến album**

Độ phổ biến của album	Độ phổ biến bài hát
0	0
10	10
20	20
30	30
40	40
50	50
60	60
70	70
80	65
10	5
20	15
30	25
40	35
50	45
60	55
70	65
80	75
10	10
20	20
30	30
40	40
50	50
60	60
70	70
80	80

# “Chuyện đôi ta”

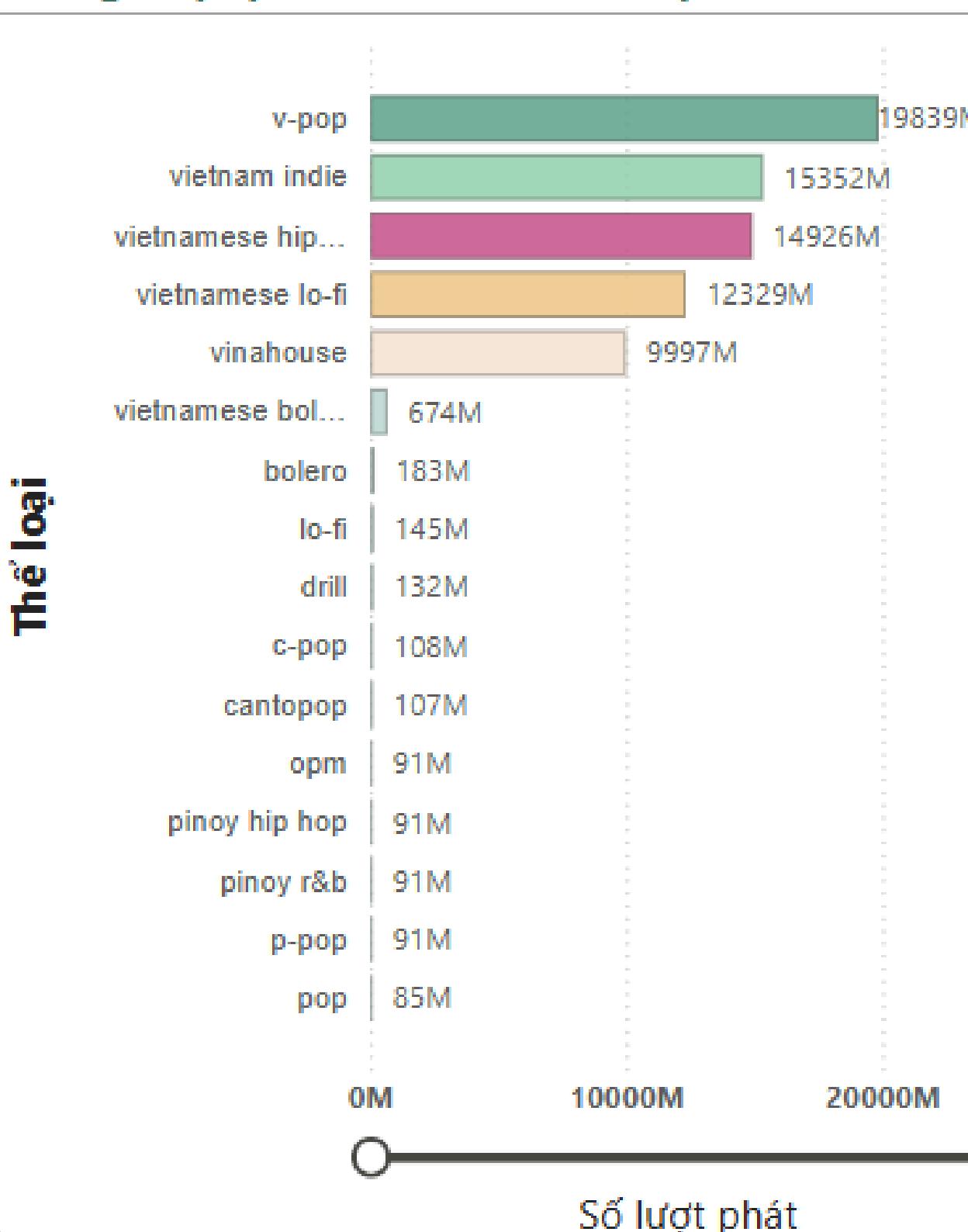


# Dashboard thể loại

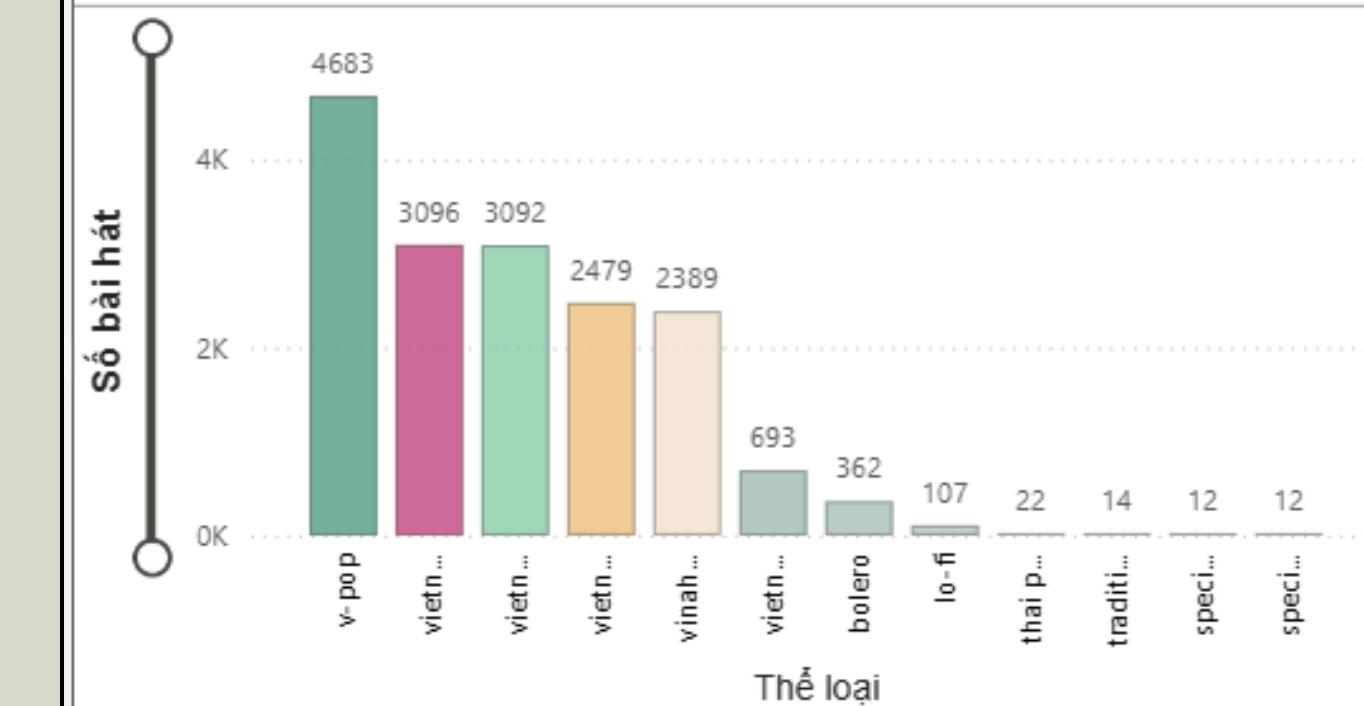


# Dashboard thể loại

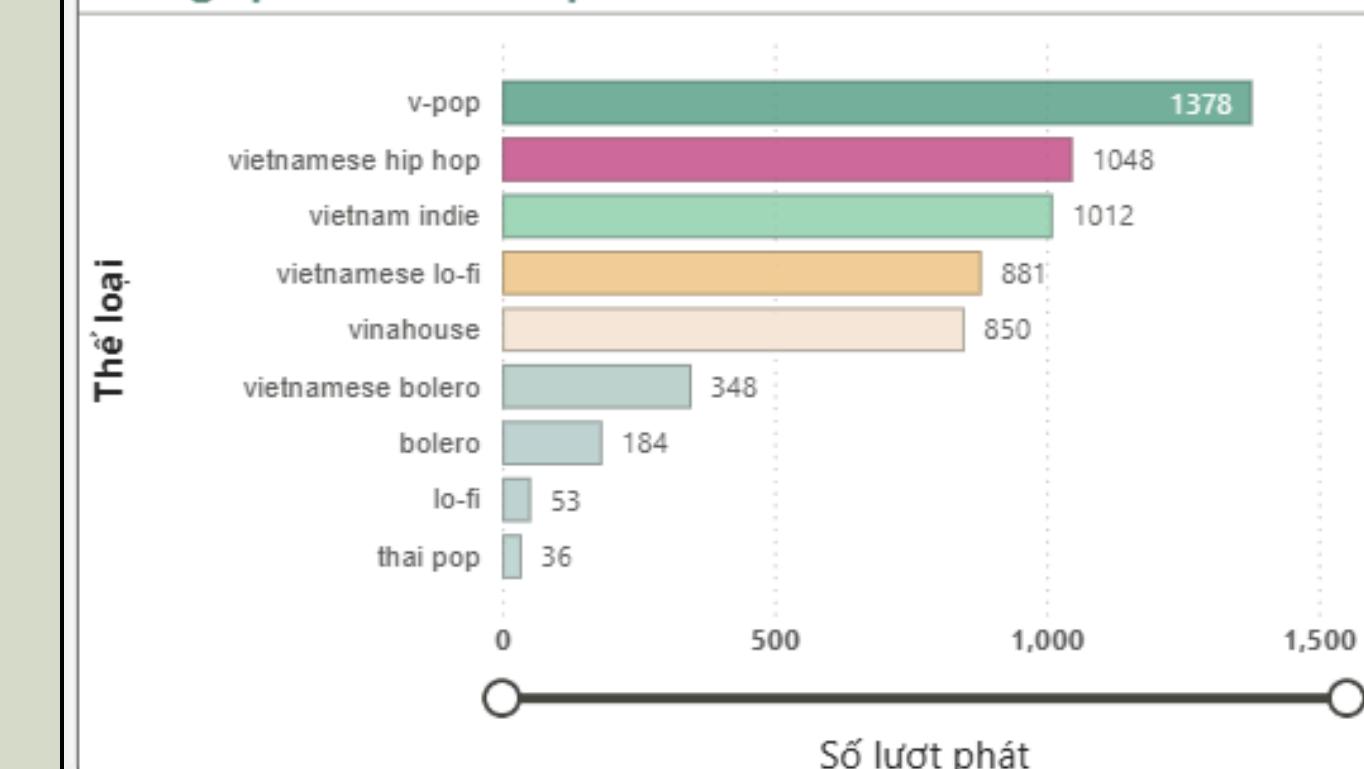
Tổng lượt phát của các thể loại



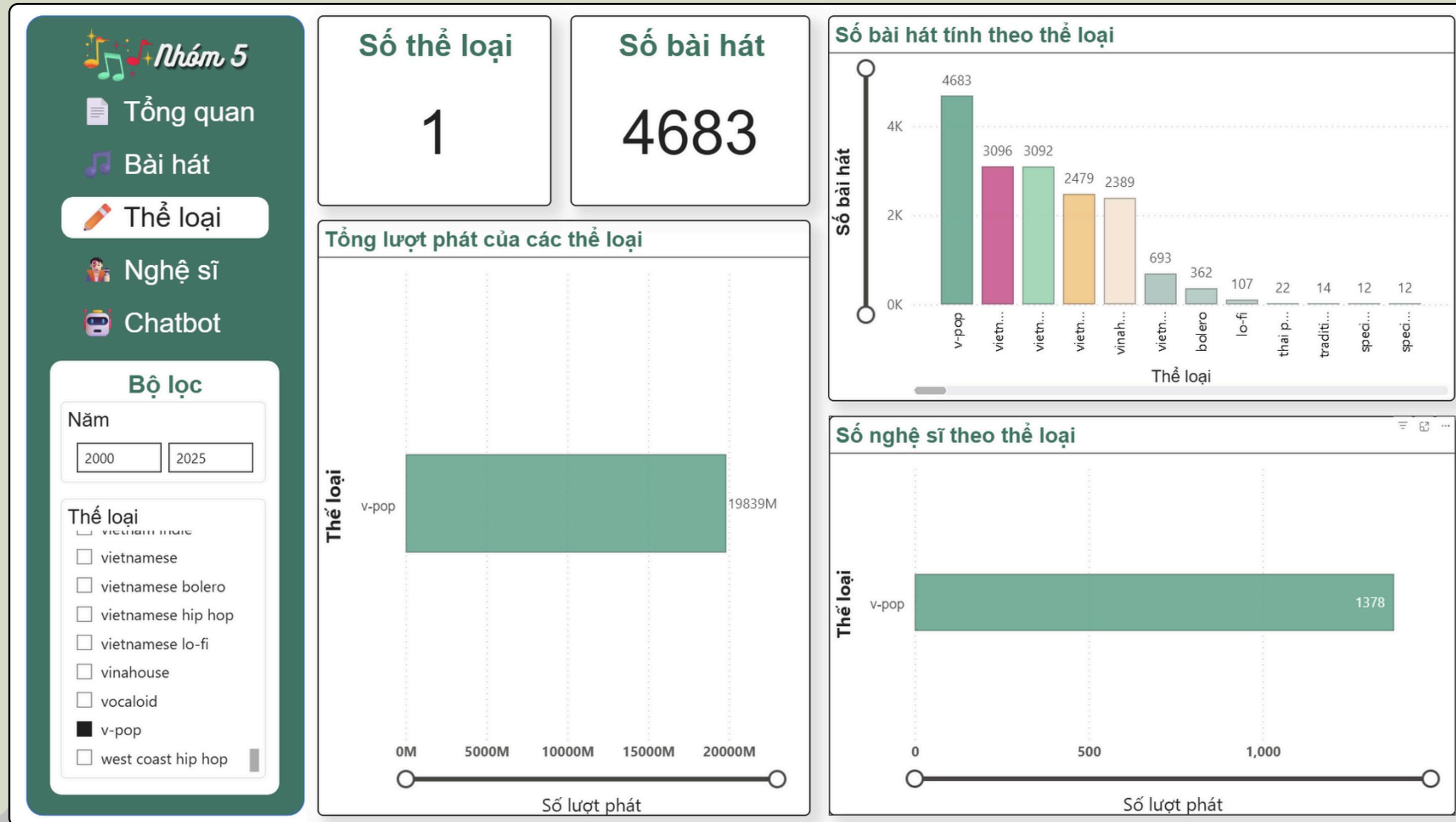
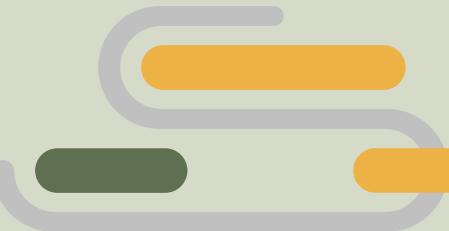
Số bài hát tính theo thể loại



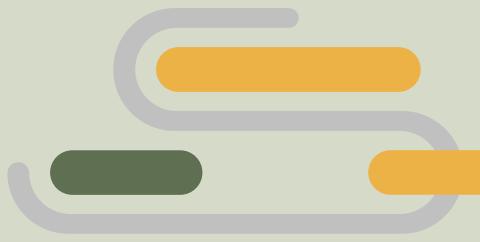
Số nghệ sĩ theo thể loại



# "v-pop"



# Dashboard nghệ sĩ



Nhóm 5

Tổng quan

Bài hát

Thể loại

Nghệ sĩ

Chatbot

**Bộ lọc**

Nghệ sĩ

Search

Select all

\$¥

\$A Milo

\$ativa

\$u

(S)TRONG

1 9 6 7

14 Casper

16 BeanCd

...

Tên  
\$A Milo

Độ nổi tiếng  
31

Lượt theo dõi  
99.6M

Lượt phát  
20bn

Thể loại âm nhạc  
bolero, vietnamese  
bolero, vinhouse, vi...

Số lượng bài hát và album

Tên nghệ sĩ	Số bài hát và album
Oriinn	~50
JustaTee	~50
Binz	~100
RAP VIỆT	~100
tlinh	~50
Đen	~50
ĐỨC PHẠM	~50
HIEUTHUHAI	~50
HUSTLAR	~50
B Ray	~50
Gill	~50
Masew	~50
ERIK	~50
HURRY...	~50
Andiez	~50
Low G	~50
Andree...	~50
Orange	~50
Karik	~50

Xếp hạng độ nổi tiếng

Tên nghệ sĩ	Độ nổi tiếng
HIEUTHUHAI	~55
ANH TRUNG	~50
Dangra...	~45
tlinh	~40
Dương...	~35
Sơn Tùng...	~30
buitruong...	~25
Wrdie	~20
ERIK	~18
Low G	~15
Obito	~12
RHYDER	~10
Shiki	~8
SOOBIN	~5
Vũ.	~3
Da LAB	~2
Negav	~1
RAP VIỆT	~1
RPT MCK	~1
Đen	~1

Tên bài hát

Lượt phát

Tên bài hát	Lượt phát
La Diabla	853240373
raining for hours	645233450
Tùng Quen	207097026
Andree Right Hand xin chào	198549794
vaicaunoicokhiennguithaydoi	173890637
nếu lúc đó	172030540
Chìm Sâu	170234048

Những nghệ sĩ có nhiều lượt theo dõi

Artist	Lượt theo dõi
Sơn Tùng M-TP	6.5M
Vũ.	2.4M
RAP VIỆT	1.9M
Low G	1.6M
W/N	1.6M
HIEUTHUHAI	2.2M
AMEE	1.5M
Đen	4.3M
Mỹ Tâm	2.0M
tlinh	1.4M
ĐỨC PHẠM	1.3M
Hà Anh	1.3M
Bích Phương	1.3M

# “Sơn Tùng MTP”

Nhóm 5

Tổng quan

Bài hát

Thể loại

Nghệ sĩ

Chatbot

**Bộ lọc**

Nghệ sĩ  
 Sơn Tùng  
 Sơn Tùng M-TP

Tên	Độ nổi tiếng	Lượt theo dõi	Lượt phát	Thể loại âm nhạc
Sơn Tùng M-TP	65	6.5M	865M	v-pop, vinahouse, vietnamese hip hop

**Số lượng bài hát và album**

**Xếp hạng độ nổi tiếng**

**Tên bài hát**

Tên bài hát	Lượt phát
Chúng Ta Của Hiện Tại	116794698
Muộn Rồi Mà Sao Còn	99870709
Đừng Làm Trái Tim Anh Đau	79947388
Âm Thầm Bên Em	66112333
Nơi Nay Có Anh	63986346
Có Chắc Yêu Là Đây	62462409
Chúng Ta Của Tương Lai	59022893

**Những nghệ sĩ có nhiều lượt theo dõi**

Nghệ sĩ	Lượt theo dõi
Sơn Tùng M-TP	6.5M
Vũ.	2.4M
HIEUTHUHAI	2.2M
RAP VIỆT	1.9M
Low G	1.6M
W/N	1.6M
AMEE	1.5M
Đức Phúc	1.3M
Hà Anh	1.3M
tlinh	1.4M
Mỹ Tâm	2.0M
Bích Phương	1.3M

# "HIEUTHUHAI"

Nhóm 5

Tổng quan

Bài hát

Thể loại

Nghệ sĩ

Chatbot

Bộ lọc

Nghệ sĩ  
HIEUTHUHAI

HIEUTHUHAI

Tên	Độ nổi tiếng	Lượt theo dõi	Lượt phát	Thể loại âm nhạc
HIEUTHUHAI	68	2.2M	840M	v-pop, vietnamese hip hop, vietnamese...

Số lượng bài hát và album

Xếp hạng độ nổi tiếng

Số bài hát và album

Tên bài hát	Lượt phát
Không Thể Say	87184891
Exit Sign	86860398
Hẹn Gặp Em Dưới Ánh Trăng	80301683
ngủ một mình (tình rất tình)	68880726
Nghe Như Tình Yêu	67048963
KIM PHÚT, KIM GIỜ (feat. HURRYKNG, HIEUTHUHAI, Negav, Pháp Kiều & Isaac)	53783518

Những nghệ sĩ có nhiều lượt theo dõi

Nghệ sĩ	Lượt theo dõi
Sơn Tùng M-TP	6.5M
Vũ.	2.4M
HIEUTHUHAI	2.2M
Đen	4.3M
AMEE	1.5M
tlinh	1.4M
Đức Phúc	1.3M
Hà Anh	1.3M
Mỹ Tâm	2.0M
Bích Phương	1.6M

# Chatbot

## Nhóm 5

- Tổng quan
- Bài hát
- Thể loại
- Nghệ sĩ
- Chatbot

### Mô tả

AI Lens là một visual cho Power BI tích hợp OpenAI, cho phép bạn đặt câu hỏi bằng ngôn ngữ tự nhiên và nhận insight trực quan tức thì ngay trên dashboard. Giúp tối ưu hóa quá trình ra quyết định nhờ khả năng phân tích tự động và trực quan hóa linh hoạt.

Hỗ trợ bởi OpenAI GPT-4.  
© Mọi quyền được bảo lưu.

## Hỏi đáp

Ask a question about your data

Try one of these to get started

- what is the duration minute by album type
- top values by số thể loại
- what is the số thể loại by value
- top popularity bins by số bài hát

Show all suggestions

### Chatbot bài hát

Xin chào, tôi là Manbell, tôi có thể giúp gì cho bạn?

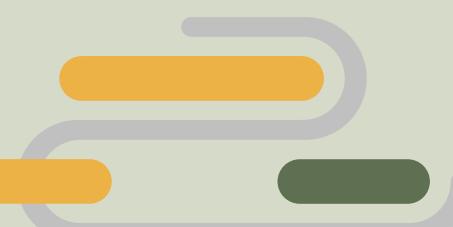
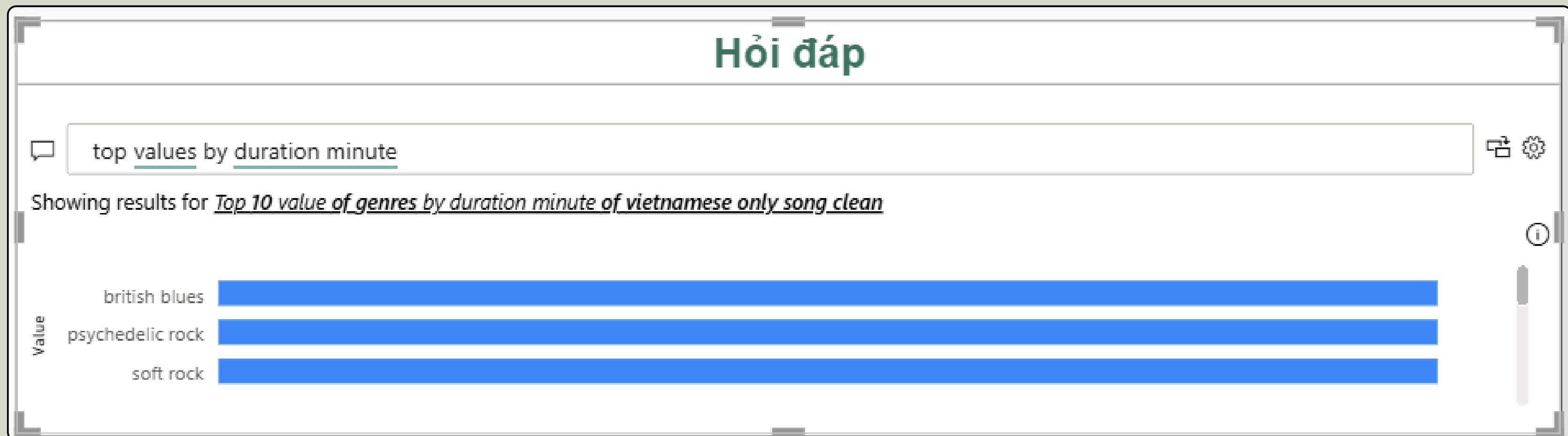
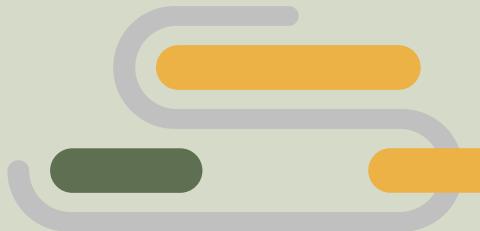
Send a message. Send

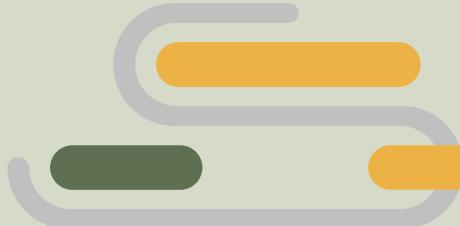
### Chatbot nghệ sĩ

Xin chào, tôi là Manbell, tôi có thể giúp gì cho bạn?

Send a message. Send

# Q&A có sẵn trong PowerBI





## Chatbot bài hát



Xin chào, tôi là Manbell, tôi có thể giúp gì cho bạn?



Bài hát nào nổi tiếng nhất



Một trong những bài hát nổi tiếng nhất trong làng nhạc Việt Nam hiện nay là "Chúng Ta Của Hiện Tại" của Sơn Tùng M-TP. Bài hát này đã thu hút được rất nhiều lượt nghe và sự yêu thích từ khán giả. Ngoài ra, các bài hát khác như "Hãy Trao Cho Anh" cũng của Sơn Tùng M-TP, hay "See Tình" của Hoàng Thùy Linh cũng rất được yêu thích và nổi tiếng.

Send a message.

Send

## Chatbot nghệ sĩ



Xin chào, tôi là Manbell, tôi có thể giúp gì cho bạn?



Nghệ sĩ nào có lượt theo dõi và độ nổi tiếng nhất



Nghệ sĩ có lượt theo dõi và độ nổi tiếng nhất hiện tại là **Sơn Tùng M-TP** với 6,549,380 lượt theo dõi và độ nổi tiếng là 65.

Send a message.

Send



**Thank You**