

HONG KONG INSTITUTE OF INFORMATION TECHNOLOGY

DEPARTMENT OF AI AND DATA

Higher Diploma in Data Science and AI

ITP4864 Data Visualization

EA Assignment

Group Project

You are required to form a group of 2~ 3 classmates or even work individually. Each group or individual will receive two sets of data. All groups or individuals have to process the dataset and generate some findings. The data is about the Top 50 Popular Songs in 73 Countries. All groups and individuals use the provided data and finish the following tasks: Data preparation, Mandatory task, Advanced technique task, and Finding demonstrations. All tasks and demonstrations are required to utilise the features of Tableau Public and R for data pre-processing.

To Conduct the project smoothly, The project topic should not be limited to the following suggestions but should included in the theme of “**Music with You**”:

1. Review of song style in different country?
2. Relationship between artist and song?
3. Which kind of song will become the next trend?
4. The style of different artists.

Main Dataset Description

Popular Music Data

This is public data from the popular Music Platform¹. The data contain the details of each transaction, from 1995 to 2024. The following is a sample of the dataset.

2RkZ5LklOrdinary	Alex Warren	1	1	0	2025/6/11	95	FALSE	186964 You'll Be	2024/9/26	0.368	0.694	2	-6.141	1	0.06
42UBPzR Manchild	Sabrina Carpenter	2	-1	48	2025/6/11	89	TRUE	213645 Manchild	2025/6/5	0.731	0.685	7	-5.087	1	0.0572
0FTmksd; back to friends	sombr	3	0	1	2025/6/11	98	FALSE	199032 back to fri	2024/12/27	0.436	0.723	1	-2.291	1	0.0301
7so0lqd0z Die With A Smile	Lady Gaga, Bruno	4	0	-1	2025/6/11	91	FALSE	251667 MAYHEM	2025/3/7	0.519	0.601	6	-7.727	0	0.0317
6d0rVTD; BIRDS OF A FEAT	Billie Eilish	5	1	0	2025/6/11	100	FALSE	210373 HIT ME F	2024/5/17	0.747	0.507	2	-10.171	1	0.0358
27xkOIEFDon? Say You LcJin		6	-1	-4	2025/6/11	93	FALSE	180716 Echo	2025/5/16	0.729	0.562	8	-5.49	1	0.0304
6iOndD4C; La Plena - W Sound	W Sound, Be; 蒙le,	7	0	0	2025/6/11	95	TRUE	150001 La Plena (2025/2/19	0.894	0.643	5	-3.485	1	0.132
4AajxC; Evundressed	sombr	8	0	-2	2025/6/11	96	FALSE	182088 undressed	2025/3/21	0.642	0.887	0	-3.883	1	0.0406
4wJ5Qq0j APT.	ROS?, Bruno Mars	9	1	-1	2025/6/11	89	FALSE	169917 rosie	2024/12/6	0.777	0.783	0	-4.477	0	0.26
4WFGvKVL; Love Me Not	Ravyn Lenae	10	-1	9	2025/6/11	91	FALSE	213466 Bird's Eye	2024/8/9	0.779	0.731	0	-6.176	1	0.0306
3QaPy1K; WILDFLOWER	Billie Eilish	11	1	1	2025/6/11	96	FALSE	261466 HIT ME F	2024/5/17	0.467	0.247	6	-12.002	0	0.0431
6eLQXa6; TU SANCHO	Fuerza Regida	12	2	-1	2025/6/11	90	TRUE	177994 111XPAN	2025/5/2	0.689	0.84	8	-6.423	1	0.0457
2yWIGeg; Just Keep Watching	Tate McRae, F1 TI	13	0	-3	2025/6/11	91	FALSE	142550 Just Keep	2025/5/30	0.656	0.796	9	-5.856	0	0.113
3xkHsm; Beautiful Things	Benson Boone	14	3	-1	2025/6/11	93	FALSE	180304 Fireworks	2024/4/5	0.472	0.471	10	-5.692	1	0.0603
2HRqTp; Espresso	Sabrina Carpenter	15	1	6	2025/6/11	91	TRUE	175459 Short n' S'	2024/8/23	0.699	0.776	0	-5.282	1	0.0293
21IYMdz; Sailor Song	Gigi Perez	16	-1	-2	2025/6/11	89	FALSE	211978 At The Be	2025/4/25	0.499	0.392	11	-10.441	1	0.0262
3sK8wGT; D; MF	Bad Bunny	17	2	-2	2025/6/11	97	TRUE	237117 DeB? TiR	2025/1/5	0.625	0.131	7	-27.405	0	0.0717
7tl8dRuH; Who	Jimin	18	-7	-9	2025/6/11	90	FALSE	170887 MUSE	2024/7/19	0.66	0.756	0	-3.743	0	0.032
7ne4VBA; That? So True	Gracie Abrams	19	2	-1	2025/6/11	96	TRUE	166300 The Secre	2024/10/18	0.554	0.808	1	-4.169	1	0.0368
0QC1pQV; Shake It To The Ma	MOLIIY, Silent Ad	20	-2	-4	2025/6/11	93	TRUE	178775 Shake It T	2025/2/21	0.856	0.548	0	-6.609	0	0.123

¹ Data source website: <https://www.kaggle.com/datasets/asaniczka/top-spotify-songs-in-73-countries-daily-updated>

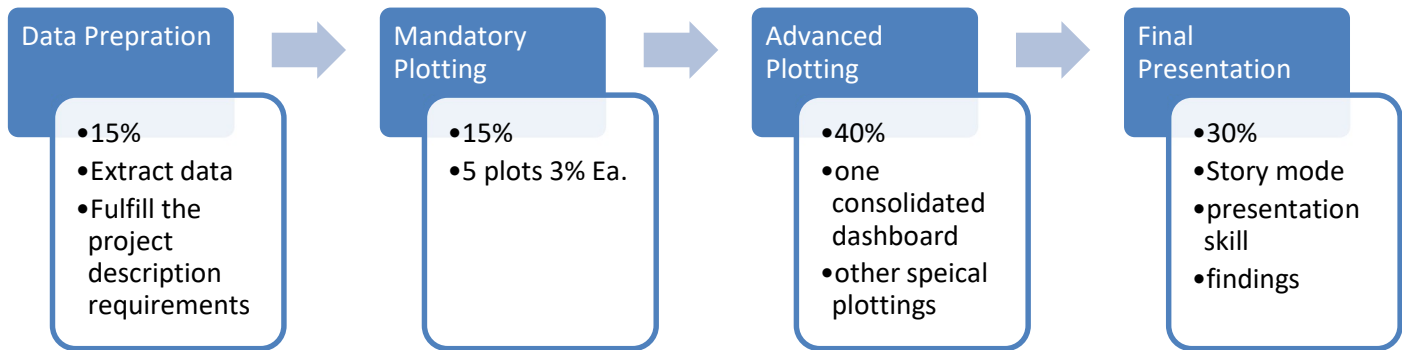
Data Description

The following is the initial Data Description of the dataset provided by Kaggle platform, and this information may help you summarise the finalized data dictionary for this project.

Column	Explanation (where appropriate)
spotify_id	The unique identifier for the song in the Spotify database
name	The title of the song
daily_rank	The daily rank of the song in the top 50 list
artists	The name(s) of the artist(s) associated with the song
daily_movement	The change in rankings compared to the previous day
weekly_movement	The change in rankings compared to the previous week
country	The ISO code of the country of the Top 50 Playlist. If Null, then the playlist is 'Global Top 50'
snapshot_date	The date on which the data was collected from the Spotify API
popularity	A measure of the song's current popularity on Spotify
is_explicit	Indicates whether the song contains explicit lyrics
duration_ms	The duration of the song in milliseconds
album_name	The title of the album the song belongs to
album_release_date	The release date of the album the song belongs to
danceability	A measure of how suitable the song is for dancing based on various musical elements
energy	A measure of the intensity and activity level of the song.
key	The key of the song
loudness	The overall loudness of the song in decibels
mode	Indicates whether the song is in a major or minor key
speechiness	A measure of the presence of spoken words in the song
acousticness	A measure of the acoustic quality of the song
instrumentalness	A measure of the likelihood that the song does not contain vocals
liveness	A measure of the presence of a live audience in the recording

Column	Explanation (where appropriate)
valence	A measure of the musical positiveness conveyed by the song
tempo	The tempo of the song in beats per minute
time_signature	The estimated overall time signature of the song

Task



Totally: 100%

Data preparation

You will receive 2,110,316 records for your project, and you have to ETL before pouring the data into the Tableau workbook. Therefore, create an R notebook to:

1. Extract the full name of ISO Country code from <https://help.adjust.com/en/article/countries-by-region>
2. convert the country to the fill name
3. Label the valence and speechiness value by the following table

Category/Range	< 1 st quartile	1 st quartile ~ Median	Median ~ 3 rd quartile	> 3 rd quartile
speechiness	Pure music	Song and words balaced	Word leads music	Words oriented
valence	darkness	Gray feeling	positive	brightness

4. Extract the Managed data to csv format and named Music_Final.csv

Deliverable:

1. Several dataset files in a CSV format:
 - I. Main Data should be in one file (Music_Final.csv)

Hint:

1. Filtrate all unnecessary columns (up to you)
2. The below is just for your reference (in a CSV format)

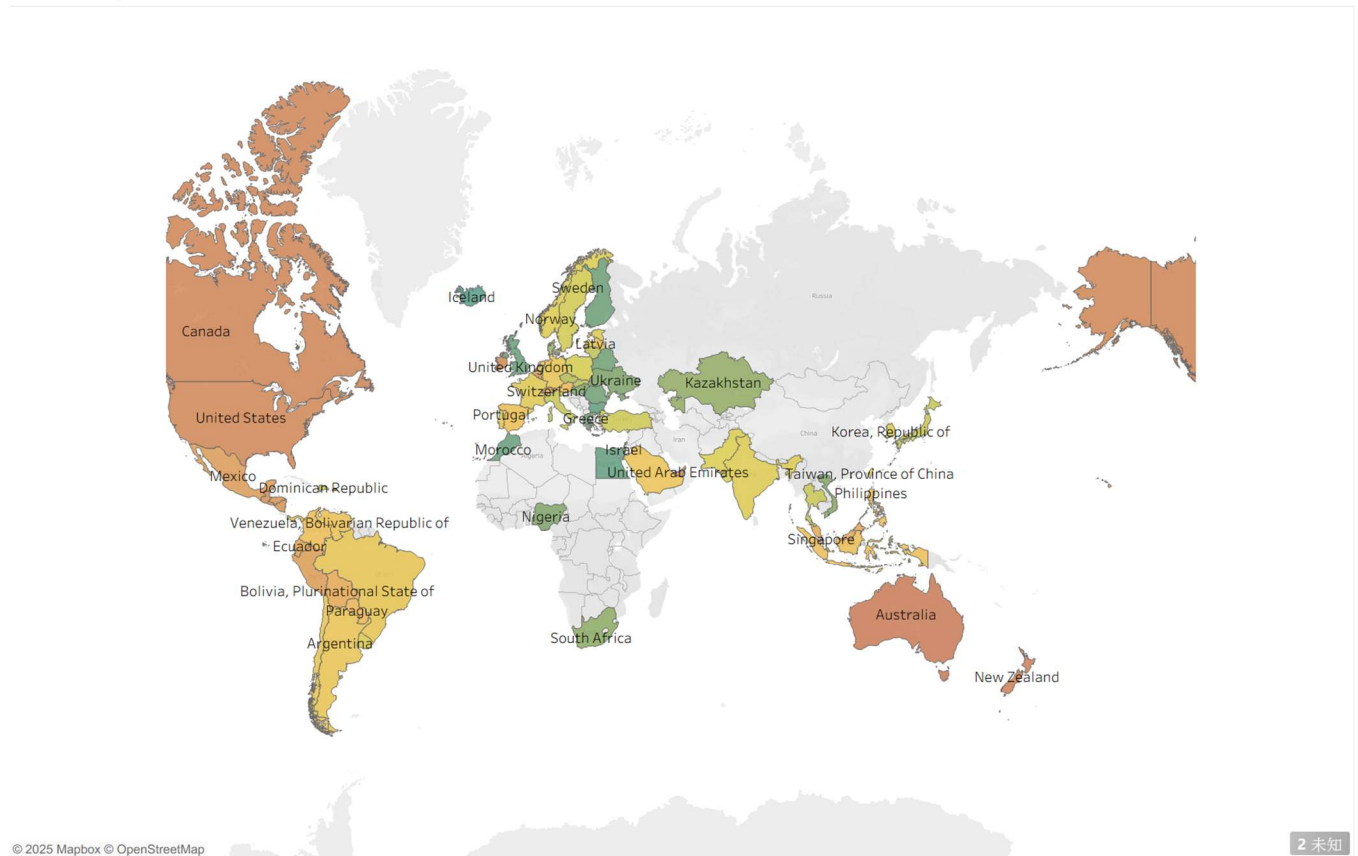
Name	country	spotify_id	name	artists	daily_rank	daily_mov	weekly	m_snapshot	c_popularity	is_explicit	duration	album	par_album	album_releases	danceability	energy	key	loudness	mode	speechiness	speechiness	acousticness	instrumentalness	liveness	valence	valence_la	tempo	time_signature	Region_name
Argentina AR	5AKLEnor	Si Un D	Big One, I		41	1	-7	2025/3/1	69	FALSE	164021	Si Un D		2024/9/3	0.63	0.507	1	-4.865	1	0.0647	word leads	0.0386	0	0.0666	0.744	brightness	183.896	4	Latin America
Argentina AR	6MuUSY	Ese Maldit	No Te Va		36	5	0	2025/6/9	73	FALSE	185866	El Calor C		2012/10/18	0.643	0.616	11	-6.358	0	0.0266	pure music	0.469	5.74e-05	0.0886	0.652	positive	139.931	4	Latin America
Argentina AR	38BE4th8	Con otra	Cazzu		6	0	0	2025/4/17	85	FALSE	231016	Con otra		2025/9/19	0.581	0.61	10	-5.787	0	0.0492	song and v	0.274	0	0.949	0.55	positive	168.162	4	Latin America
Argentina AR	7d6yK3vG	VeLD?	Bad Bunny		9	-2	-1	2025/4/17	93	TRUE	235336	DeB? TIR.		2025/1/5	0.613	0.981	1	-18.638	1	0.149	words orie	0.29	0.0023	0.437	0.0337	darkness	101.852	3	Latin America
Argentina AR	6WafBL	'Si Antes T	KAROL C		12	-1	-6	2025/1/16	92	FALSE	195824	Si Antes T		2024/6/21	0.924	0.668	11	-6.795	1	0.0469	song and v	0.446	0.000594	0.0678	0.787	brightness	128.027	4	Latin America
Argentina AR	4xXvGEI	COLOCA	Renzo ED		23	4	-1	2025/3/31	73	FALSE	176250	COLOCA		2024/10/24	0.857	0.701	0	-3.483	0	0.063	word leads	0.434	0	0.0395	0.812	brightness	96.046	4	Latin America
Argentina AR	5MKbW	'Amor de V	La T y La		7	1	0	2025/4/17	81	FALSE	188423	Amor de V		2024/8/23	0.65	0.507	2	-7.27	1	0.111	word leads	0.706	2.47e-06	0.0725	0.603	positive	179.984	4	Latin America
Argentina AR	5DSMmO	DOCTOR	Luck Ra, I		43	4	-1	2025/3/1	69	FALSE	144148	DOCTOR		2024/8/29	0.651	0.762	9	-2.969	1	0.0315	pure music	0.0834	0	0.0701	0.721	positive	90.989	4	Latin America
Argentina AR	5QjmUqq	UWAIE	Kapo		26	-1	-3	2024/10/1	89	FALSE	172427	UWAIE		2024/8/15	0.705	0.783	9	-4.783	0	0.0403	song and v	0.138	0	0.0984	0.454	gray Feelli	103.001	4	Latin America
Argentina AR	4YnmQG	WYA REI	Abdiel, I		30	-1	-2	2024/10/1	83	TRUE	321741	WYA REI		2024/5/10	0.733	0.527	0	-5.965	0	0.0486	song and v	0.364	0	0.0999	0.371	gray Feelli	110.045	4	Latin America
Argentina AR	0qelFGR	DEGENEIM	Myke Tow		43	-1	-6	2025/3/21	84	TRUE	130880	LA PANT		2024/10/17	0.744	0.709	6	-5.778	0	0.284	words orie	0.0501	0	0.0847	0.712	positive	194.055	4	Latin America
Argentina AR	6Es8Sk3e	La_Origins	Emilia, TI		25	-1	0	2024/10/1	76	FALSE	140625	.mp3		2023/11/3	0.833	0.809	5	-2.751	0	0.0381	pure music	0.0836	0	0.112	0.905	brightness	128.055	4	Latin America
Argentina AR	0SoBfYQ	Hoy	Valentino		3	0	1	2024/6/25	71	FALSE	198495	Hoy		2024/5/22	0.451	0.711	11	-4.612	1	0.0426	song and v	0.24	0	0.154	0.661	positive	176.161	4	Latin America
Argentina AR	2RmIEc2Z	Ojos Verde	Nicki Nicc		4	0	2	2024/6/25	77	FALSE	142410	Ojos Verde		2024/4/24	0.582	0.834	4	-2.654	1	0.0394	song and v	0.224	2.67e-06	0.145	0.95	brightness	167.945	4	Latin America
Argentina AR	3sa6dEW	IMAN	(TV)Maria Bec		5	5	4	2024/6/25	75	FALSE	122405	IMAN (TV)		2024/5/16	0.8	0.679	3	-5.01	0	0.0485	song and v	0.154	2.59e-05	0.205	0.974	brightness	106.045	4	Latin America
Argentina AR	4dLLdini	Perdonarte	Los Angeles		2	0	0	2024/6/25	80	FALSE	193933	Perdonarte		2024/5/23	0.672	0.702	11	-4.662	0	0.0314	pure music	0.168	0	0.0445	0.615	positive	90.039	4	Latin America
Argentina AR	5rQSQJ2X	Un Besito	Salasibbroi		7	-1	-2	2024/6/25	74	FALSE	160750	Un Besito		2024/9/14	0.693	0.641	2	-5.775	0	0.0359	pure music	0.368	0	0.294	0.514	gray Feelli	92.01	4	Latin America
Argentina AR	5AOnkbW	Luck Ra: I	Bizarrap, I		6	0	-2	2025/1/30	81	FALSE	170953	Luck Ra: I		2024/12/27	0.704	0.89	0	-4.649	0	0.13	words orie	0.0623	0	0.0982	0.638	positive	159.927	4	Latin America
Argentina AR	2x6H0Cu	SINVERC	Emanero,		27	0	-3	2024/10/1	72	FALSE	217490	SINVERC		2023/12/5	0.764	0.738	0	-4.774	1	0.0336	pure music	0.182	0	0.173	0.647	positive	93.963	4	Latin America
Argentina AR	3cWo1621	Tu Foto	I-Q' Lokura,		6	-1	1	2024/6/25	72	FALSE	180385	Tu Foto		2024/5/16	0.561	0.803	11	-5.132	0	0.073	word leads	0.146	0	0.188	0.338	darkness	153.669	4	Latin America

Mandatory task (2% Ea.)

Visualise the following requirements

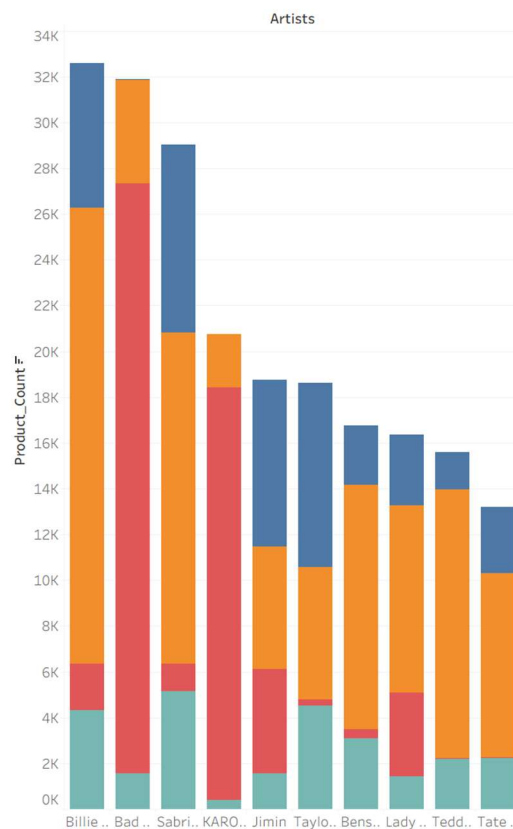
1. The global population coloured by value of population value.

Global Population



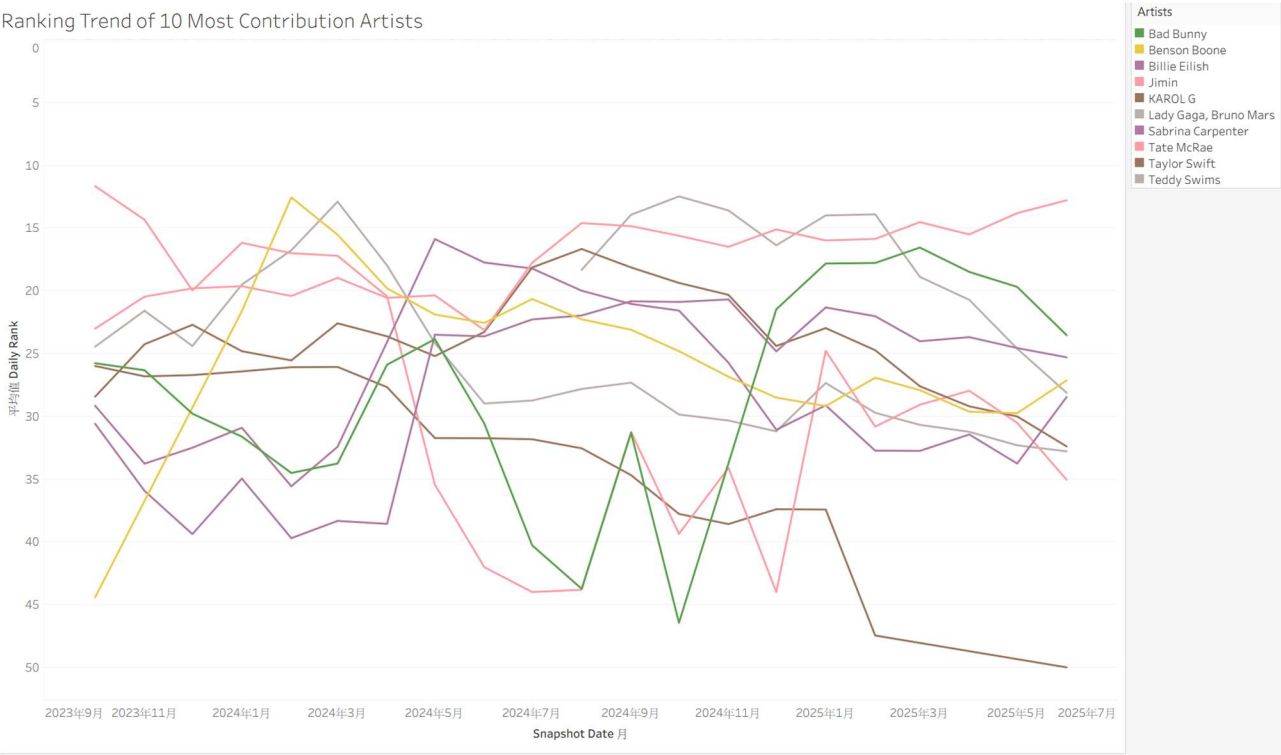
2. The 10 most contribution artists and bar should display the component of regions.

The 10 Most Contribution Artists



3. The rank trend of 10 most contribution artists.

Ranking Trend of 10 Most Contribution Artists



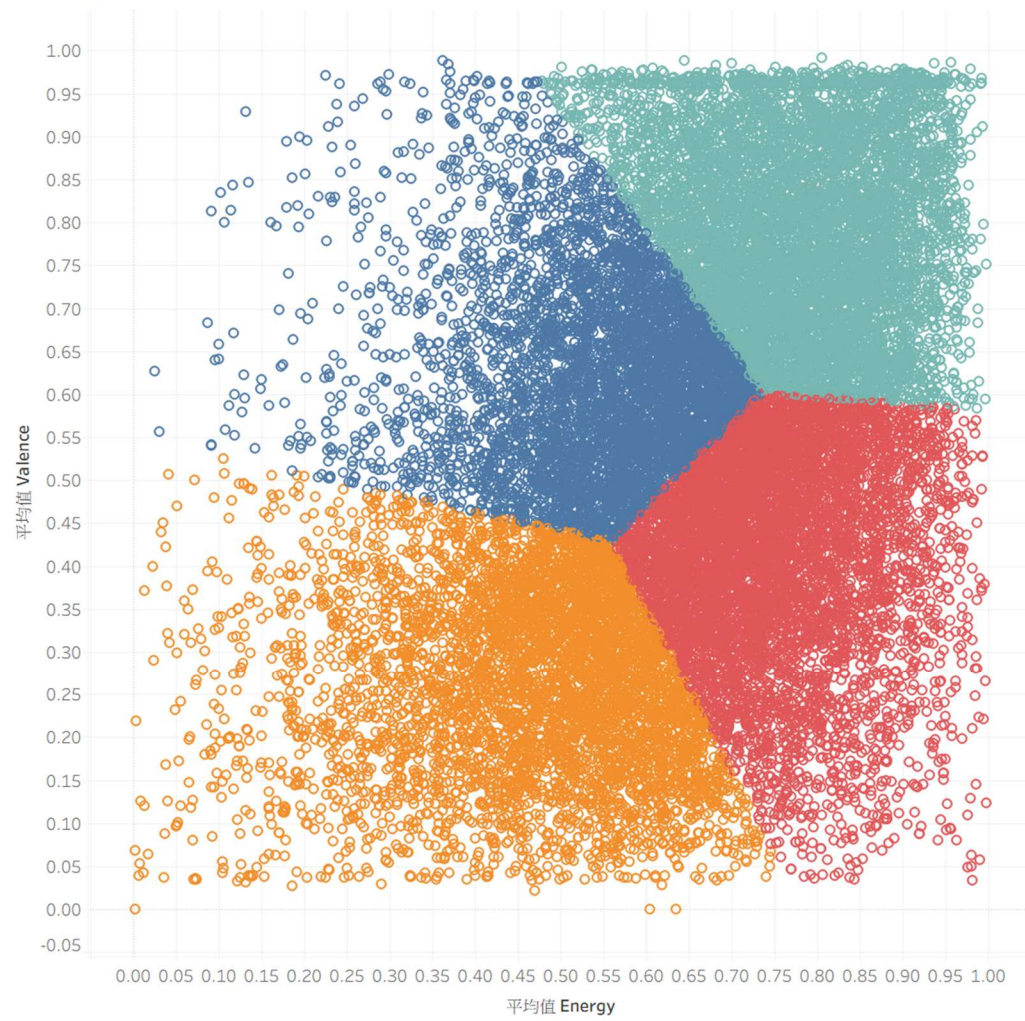
4. The difference music of artists

Artist Style

Artists	Album Name	
22 Zos Pithikos, Dani Gambino, BoyPan..	Still Standing	words oriented
13 Killoloki, Şehinşah	Pembe Yalanlar	song and words balanced
13K, Gleb	BA Connect	words oriented
14 Casper, Bon Nghiêm	SỐ KHÔNG	song and words balanced
14 Casper, Bon Nghiêm, buitruonglinh	SỐ KHÔNG	pure music
21 Savage	american dream	words oriented
21 Savage, Brent Faiyaz	american dream	word leads music
21 Savage, Burna Boy, Metro Boomin	american dream	word leads music
21 Savage, Doja Cat	american dream	word leads music
21 Savage, Lil Durk, Metro Boomin	american dream	words oriented
21 Savage, Mariah the Scientist	american dream	words oriented
21 Savage, Mikky Ekko, Tommy Newp..	american dream	word leads music
21 Savage, Summer Walker	american dream	words oriented
21 Savage, Travis Scott, Metro Boomin	american dream	words oriented
21 Savage, Young Thug, Metro Boomin	american dream	word leads music
23	Bland Sagor & Va..	words oriented
	Förstå om du in..	words oriented
	HOLD UP	words oriented
	Lucky Luciano	words oriented
	PSG	song and words balanced
23, 01an	Förstå om du in..	words oriented
	Kontroversiell	words oriented
	Kontroversiell (L..	words oriented
23, C. Gambino	M.O.B	words oriented
23, Mala	Förstå om du in..	words oriented
23, Roc Boyz	Feeling Myself (..	words oriented
23, Shenzi Beats	Drama	words oriented
	En gång till	words oriented
	Kod Röd	words oriented
	Säg Mig	words oriented
24K, Asme, ROMANOS	Med oss	song and words balanced
25, ADAAM, Kayen	BONANZA	words oriented
25, Greekazo, Swayway	BUNTAR O SÄNT	word leads music

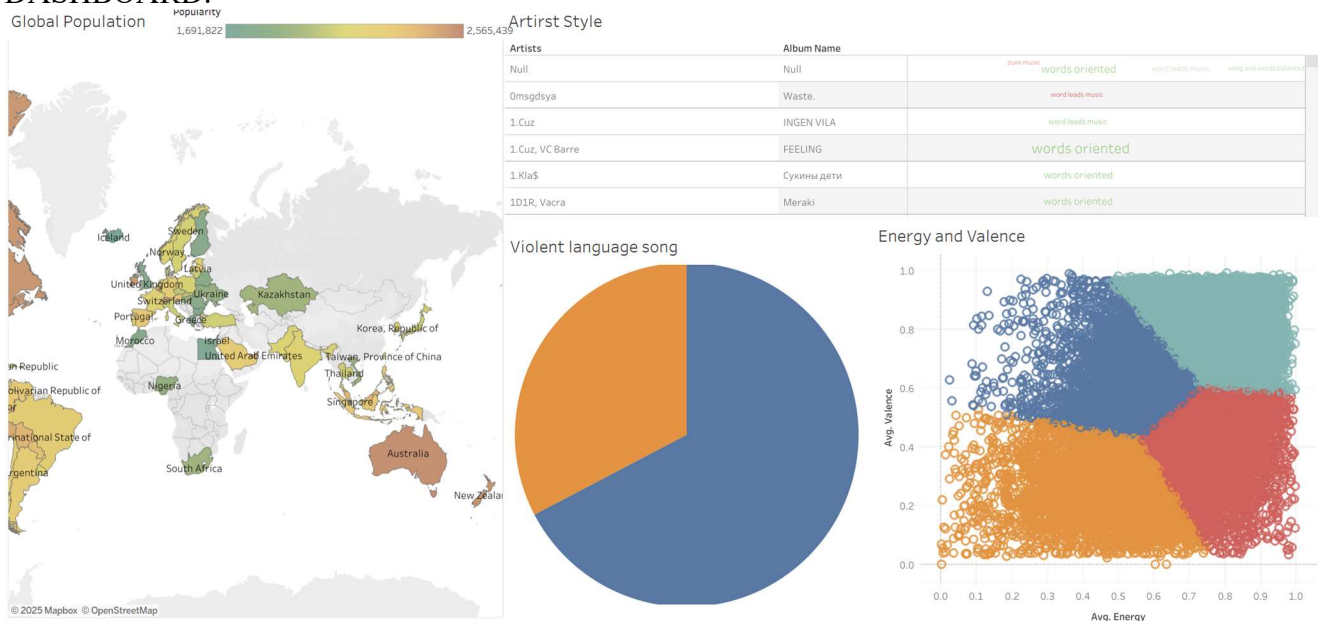
5. The clustering of Music energy and valence.

Energy and Valence



Advanced technique task

This task is a freestyle from your idea. First, you have to create dashboards for your project purpose. Bear in mind that you must complete one of the dashboards by using visualized plottings from the last section and drilling some valuable findings. Furthermore, you still have to plot some graphs not mentioned in the previous section and consolidate them as a story-telling dashboard. The below is a dashboard for your reference. You may follow the guidelines of Steve Few informative dashboards for designing a good dashboard. Remember, a good dashboard is not an ornament; it should be a tool for assisting your findings. PLEASE DON'T USE A SINGLE TYPE OF GRAPH FOR YOUR DASHBOARD.



Finding demonstrations

You should present your idea/findings through the story mode in 10 minutes. No PowerPoint or other presentation software is allowed. The presentation should make use of the following Tableau features:

1. Story mode
2. Dashboard
3. Filtering data
4. Drill-in data
5. Story-telling
6. Presentation Mode
7. Single Graph sheet

Furthermore, you also have prepared a report for our record. You may find the report template in Moodle. The Report has to include the following parts:

1. Introduction of project
2. Data Dictionary
3. Findings
4. Suggestions
5. Reference (if any)

Instructions to Students

1. Backup your initial dataset for Data recovery
2. The main dataset in a single dataset file is good practice
3. Think of one of the topics if there is no direction for the Advance and demonstration part.

Submission Day

Section	Date	Deliverable
Data Preparation	Check Moodle	Zip/Compress all the files in one folder includes: 1. Music_Final.csv 2. R Notebook
Final Demonstration	Check Moodle	Zip/Compress all the files in one folder includes: 1. Single Tableau Workbook (Sheet + Dashboard + Story) 2. Written Report