

Analysing Spotify Data

Learning Level: First (P2 to P4), Second (P5 to P7)

PPDAC Framework: Data | Analysis | Conclusion

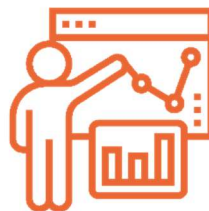
Theme: Data & Analysis | Searching, processing and managing information responsibly

Curricular Area: Maths & Numeracy, Technologies

Duration: 1 hour

Materials:

- Spotify dataset appropriate to your learners. [Top 20](#) or [Top 50](#).
- Learner worksheet that corresponds to the data set in use
- Paper, pencil, ruler
- Coloured paper & coloured pencils (optional)
- Access to spreadsheet software (optional)



Introduction

This task makes use of secondary data provided by Spotify. Learners will use the data to identify patterns and relationships between different characteristics. The activity will support learners in developing their ability to review and interpret a dataset. The activity starts by encouraging learners to think of questions that the dataset might answer and thus develop their “Problem” skills. Once the dataset has been analysed learners will have the opportunity to reorganise and restructure data to help them answer their questions (draw conclusions)

Curriculum Links

This activity can be completed with or without a computer. The curriculum links for both options are shown below.

Curriculum Links for Offline Task

- **MTH 1-21a:** Using technology and other methods, I can display data simply, clearly and accurately by creating tables, charts and diagrams, using simple labelling and scale.
- **MNU 1-20a:** I have explored a variety of ways in which data is presented and can ask and answer questions about the information it contains.
- **MNU 2-20a:** Having discussed the variety of ways and range of media used to present data, I can interpret and draw conclusions from the information displayed, recognising that the presentation may be misleading.
- **MTH 2-21a:** I can display data in a clear way using a suitable scale, by choosing appropriately from an extended range of tables, charts, diagrams and graphs, making effective use of technology.

Curriculum Links for Online Task

- **TCH 1-02a:** Using digital technologies responsibly I can access, retrieve and use information to support, enrich or extend learning in different contexts.
- **TCH 2-01a:** I can extend and enhance my knowledge of digital technologies to collect, analyse ideas, relevant information and organise these in an appropriate way.
- **TCH 1-13a:** I can explore and comment on processes in the world around me making use of core computational thinking concepts and can organise information in a logical way.

Differentiation

This task can be tailored between first and second level by varying the dataset used.

Pre-Requisites

This activity can be used to introduce learners to elements of PPDAC or to provide a further opportunity to develop their understanding.

Discussion

Hold a discussion with the group about music. Start with favourite songs and artist and support the group to share their own favourites. Progress to asking what makes a good song and why it might be popular. Finally start to identify characteristics of a song and note them either on a board or encourage the learners to write them on a piece of paper.

Characteristics can include.

- Song name
- Artist name
- Genre/Type
- Length
- Rank (popularity)

Exploring the dataset (suitable for home)

This step is focused on the “**Problem**” step of PPDAC. We are encouraging learners to think about interesting questions about the data they have.

Look through the printable [‘Top 20’](#) or [‘Top 50’](#) Spotify songs on the ‘Spotify Analysis’ resource page of the Data Education in Schools website or at the end of this document, or access the data via Kaggle at www.kaggle.com.

What do you notice about the data, what can you find out? Can you think of any questions you would like to know the answer to?

Write down:

- 3 different genres of music in your dataset
- 2 songs with the same popularity score
- The song that will come first and last alphabetically

Organising & Analysing the dataset

This task is focused on the “**Analysis**” step of PPDAC. We are encouraging learners to review the dataset and identify useful information that may help them understand the questions they thought about in step 1.

If completing this task offline then it can be easier to use the smaller dataset (20 songs) even if the learners are at second level. The task is split into two, one set of questions for each dataset. Worksheets are available for each dataset at the end of this document.

Smaller Dataset

Use your dataset to complete the following tasks.

1. What is the most popular song?
2. What is the least popular song?
3. Group the songs by genre
 - a. How many songs are there in the pop category?
 - b. What categories only have 1 song in them?
4. Rewrite the table starting at the most popular song and ending with the least popular song.

Larger Dataset

Use your dataset to complete the following tasks.

1. What is the most popular song?
2. What is the least popular song?
3. Group the songs by genre
 - a. How many songs are there in the pop category?
 - b. What categories only have 1 song in them?
4. Which song is the best to dance to?
5. Which artist sings the song with the most beats per minute?
6. Identify 3 artists with more than one song in the dataset

The above tasks can be made easier if the data is sorted in Excel (or similar spreadsheet package). Teachers or Parents can encourage children to learn how to sort the data by following the instructions here: <https://youtu.be/KS9N4yAjuYQ>

Visualise your dataset

This task mixes the “**Analysis**” and “**Conclusions**” part of PPDAC (Problem, Plan, Data, Analysis, Conclusion) cycle. Creating a summary table continues the analysis phase and seeks to show how a large amount of data can be presented in a different way. Visualising the data can support both analysis and drawing conclusions. Learners should be encouraged to connect questions they asked at the start to the analysis they are doing the information they can extract from the visualisation.

We are now going to learn to create a visual representation of the data. Again this can be done with or without a computer.

Smaller Dataset

1. Add up the total score for each genre of music. Complete this table.

Genre	Total Score
Canadian hip hop	
Canadian pop	
country rap	
dance pop	
dfw rap	
electropop	
escape room	
Latin	
Panamanian pop	
pop	
reggaeton	
reggaeton flow	
trap music	

2. Create a bar chart of genre and total score. The chart can be created in a variety of different ways.
 - a. Draw it on paper
 - b. Use Lego, or similar blocks to create it on a table
 - c. Use sticks and other natural materials to create it outside
 - d. Use Excel to create a chart in your spreadsheet.

The group can have an optional discussion about whether the highest scoring genre is therefore the most popular. This is an opportunity consider the ideas behind averages.

Learners could be asked to:

- Review what are the 3 most popular songs, are these in the most popular genre?
- What might cause the most popular genre to be different to the genre of the most popular song?

Larger Dataset

The following tasks are more suited to using a spreadsheet.

1. Create a table with 4 columns, Genre, Number of songs, Total Popularity Score and Average score
2. Fill in your table
 - a. Put each genre in the first column
 - b. Find out how many songs are in each genre and enter it in column 2
 - c. Add up total popularity score for each genre and enter it in column 3
3. Calculate the average score by dividing the total popularity score by the number of songs and enter it in column 4.
4. Create a bar chart of genre and average score in your spreadsheet software.

Dataset Top 20

[Excel version of Top 20 dataset](#)

	Track.Name	Artist	Genre	Popularity
15	Money In The Grave (Drake ft. Rick Ross)	Drake	canadian hip hop	92
19	Lalala	Y2K	canadian hip hop	88
1	Señorita	Shawn Mendes	canadian pop	79
9	Old Town Road - Remix	Lil Nas X	country rap	87
12	Loco Contigo (feat. J. Balvin & Tyga)	DJ Snake	dance pop	86
3	boyfriend (with Social House)	Ariana Grande	dance pop	85
16	No Guidance (feat. Drake)	Chris Brown	dance pop	82
5	Goodbyes (Feat. Young Thug)	Post Malone	dfw rap	94
18	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone	dfw rap	91
10	bad guy	Billie Eilish	electropop	95
20	Truth Hurts	Lizzo	escape room	91
17	LA CANCIÓN	J Balvin	latin	90
14	Otro Trago - Remix	Sech	panamanian pop	87
8	How Do You Sleep?	Sam Smith	pop	90
13	Someone You Loved	Lewis Capaldi	pop	88
4	Beautiful People (feat. Khalid)	Ed Sheeran	pop	86
6	I Don't Care (with Justin Bieber)	Ed Sheeran	pop	84
11	Callaita	Bad Bunny	reggaeton	93
2	China	Anuel AA	reggaeton flow	92
7	Ransom	Lil Tecca	trap music	92

Dataset Top 50

[Excel version of Top 50 dataset](#)

[PDF printable](#)

	Track.Name	Artist.Name	Genre	Beats.Per. Minute	Energy	Danceability	Loudness..dB..	Liveness	Variance	Length.	Acousticness..	Speechiness.	Popularity
1	Señorita	Shawn Mendes	canadian pop	117	55	76	-6	8	75	191	4	3	79
2	China	Anuel AA	reggaeton flow	105	81	79	-4	8	61	302	8	9	92
3	boyfriend (with Social House)	Ariana Grande	dance pop	190	80	40	-4	16	70	186	12	46	85
4	Beautiful People (feat. Khalid)	Ed Sheeran	pop	93	65	64	-8	8	55	198	12	19	86
5	Goodbyes (Feat. Young Thug)	Post Malone	dfw rap	150	65	58	-4	11	18	175	45	7	94
6	I Don't Care (with Justin Bieber)	Ed Sheeran	pop	102	68	80	-5	9	84	220	9	4	84
7	Ransom	Lil Tecca	trap music	180	64	75	-6	7	23	131	2	29	92
8	How Do You Sleep?	Sam Smith	pop	111	68	48	-5	8	35	202	15	9	90
9	Old Town Road - Remix	Lil Nas X	country rap	136	62	88	-6	11	64	157	5	10	87
10	bad guy	Billie Eilish	electropop	135	43	70	-11	10	56	194	33	38	95
11	Callaita	Bad Bunny	reggaeton	176	62	61	-5	24	24	251	60	31	93
12	Loco Contigo (feat. J. Balvin & Tyga)	DJ Snake	dance pop	96	71	82	-4	15	38	185	28	7	86

	Track.Name	Artist.Name	Genre	Beats.Per. Minute	Energy	Danceability	Loudness..dB..	Liveness	Variance	Length.	Acousticness..	Speechiness.	Popularity
28	It's You	Ali Gatie	canadian hip hop	96	46	73	-7	19	40	213	37	3	89
29	Con Calma	Daddy Yankee	latin	94	86	74	-3	6	66	193	11	6	91
30	QUE PRETENDES	J Balvin	latin	93	79	64	-4	36	94	222	3	25	89
31	Takeaway	The Chainsmokers	edm	85	51	29	-8	10	36	210	12	4	84
32	7 rings	Ariana Grande	dance pop	140	32	78	-11	9	33	179	59	33	89
33	0.9583333333	Maluma	reggaeton	96	71	78	-5	9	68	176	22	28	89
34	The London (feat. J. Cole & Travis Scott)	Young Thug	atl hip hop	98	59	80	-7	13	18	200	2	15	89
35	Never Really Over	Katy Perry	dance pop	100	88	77	-5	32	39	224	19	6	89
36	Summer Days (feat. Macklemore & Patrick Stump of Fall Out Boy)	Martin Garrix	big room	114	72	66	-7	14	32	164	18	6	89
37	Otro Trago	Sech	panamanian pop	176	70	75	-5	11	62	226	14	34	91
38	Antisocial (with Travis Scott)	Ed Sheeran	pop	152	82	72	-5	36	91	162	13	5	87
39	Sucker	Jonas Brothers	boy band	138	73	84	-5	11	95	181	4	6	80
40	i'm lonely (with Anne-Marie) - from "13 Reasons Why: Season 3"	Lauv	dance pop	95	56	81	-6	6	68	199	48	7	78
41	Higher Love	Kygo	edm	104	68	69	-7	10	40	228	2	3	88
42	You Need To Calm Down	Taylor Swift	dance pop	85	68	77	-6	7	73	171	1	5	90
43	Shallow	Lady Gaga	dance pop	96	39	57	-6	23	32	216	37	3	87

	Track.Name	Artist.Name	Genre	Beats.Per. Minute	Energy	Danceability	Loudness..dB..	Liveness	Variance	Length.	Acousticness..	Speechiness.	Popularity
44	Talk	Khalid	pop	136	40	90	-9	6	35	198	5	13	84
45	Con Altura	ROSALÍA	r&b en espanol	98	69	88	-4	5	75	162	39	12	88
46	One Thing Right	Marshmello	brostep	88	62	66	-2	58	44	182	7	5	88
47	Te Robaré	Nicky Jam	latin	176	75	67	-4	8	80	202	24	6	88
48	Happier	Marshmello	brostep	100	79	69	-3	17	67	214	19	5	88
49	Call You Mine	The Chainsmokers	edm	104	70	59	-6	41	50	218	23	3	88
50	Cross Me (feat. Chance the Rapper & PnB Rock)	Ed Sheeran	pop	95	79	75	-6	7	61	206	21	12	82

Learner Worksheet – Top 20 Songs

Exploring the dataset

Look at the table of songs.

What kind of information do you see in the data?

Can you think of any questions that you might ask about the songs Write down:

- 3 different genres of music in your dataset
 -
 -
 -
- 2 songs with the same popularity score
 -
 -
- The song that will come first and last alphabetically
 - First alphabetically:
 - Last alphabetically:

Organising & Analysing the dataset

Use your dataset to complete the following tasks.

1. What is the most popular song? _____
2. What is the least popular song? _____
3. Group the songs by genre
 - a. How many songs are there in the pop category? _____
 - b. What categories only have 1 song in them? _____
4. Rewrite the table starting at the most popular song and ending with the least popular song. You can do this on the back of this page.

Visualise your dataset (without a computer)

We are now going to learn to create a visual representation of the data.

1. Add up the total score for each genre of music. Complete this table.

Genre	Total Score
canadian hip hop	
canadian pop	
country rap	
dance pop	
dfw rap	
electropop	
escape room	
latin	
panamanian pop	
pop	
reggaeton	
reggaeton flow	
trap music	

2. Create a bar chart of genre and total score. You can create your chart in different ways.
 - a. Draw it on paper
 - b. Use Lego, or similar blocks to create it on a table
 - c. Use sticks and other natural materials to create it outside
 - d. Use excel to create a chart in your spreadsheet.
3. From your chart, which is the most popular genre? _____
4. Think about the results. Do you agree with them or do you think another genre is more popular? Write down what you think.

Learner Worksheet – Top 50 Songs

Exploring the dataset

Look at the table of songs. What kind of information do you see in the data?

Can you think of any questions that you might ask about the songs. Write down:

- 3 different genres of music in your dataset
 -
 -
 -
- 2 songs with the same popularity score
 -
 -
- The song that will come first and last alphabetically
 - First alphabetically:
 - Last alphabetically:

Organising & Analysing the dataset

Use your dataset to complete the following tasks.

5. What is the most popular song? _____
6. What is the least popular song? _____
7. Group the songs by genre
 - a. How many songs are there in the pop category? _____
 - b. What categories only have 1 song in them? _____
8. Which song is best to dance to? _____
9. Which artist sings the song with the most beats per minute? _____
10. Identify 3 artists with more than one song in the dataset.
 - 1.
 - 2.
 - 3.

Visualise your dataset (with a computer)

Use your spreadsheet to visualise your dataset.

1. Create a new tab
2. Rename your tab "Summary Data"
3. Create a new table with 4 columns, Genre, number of songs, Total Popularity Score and average score
4. Fill in your table
 - a. Put each genre in the first column
 - b. Find out how many songs are in each genre and enter it in column 2
 - c. Add up total popularity score for each genre and enter it in column 3
 - d. Calculate the average score by dividing the total popularity score by the number of songs and enter it in column 4
5. Create a bar chart of genre and average score in your spreadsheet software
6. From your chart, which is the most popular genre? _____
7. Think about the results. Do you agree with them or do you think another genre is more popular? Write down what you think.

8. Pick any two types of information and create a new table or chart.
9. Write down something interesting you have discovered in your data.

If you require this document in an alternative format, such as large print or a coloured background, please contact

dataschools@ed.ac.uk



CITY REGION DEAL
Edinburgh & South East Scotland

