Monash University: Assessment Cover Sheet

Student name	Goh	Kai Yuan	
School/Campus		**Student's I.D. number**	30881919
Unit name	FIT3179 Data visualisation - S2 2022 MUM		
Lecturer's name		**Tutor's name**	
Assignment name	Data Visualisation I Report	**Group Assignment: No** **Note, each student must attach a coversheet**	
Lab/Tute Class:	**Lab/Tute Time:**	**Word Count:**	
Due date: 05-09-2022	**Submit Date:**	**Extension granted** ☐	

If an extension of work is granted, specify date and provide the signature of the lecturer/tutor. Alternatively, attach an email printout or handwritten and signed notice from your lecturer/tutor verifying an extension has been granted.

Extension granted until (date): ...../...../..... Signature of lecturer/tutor: .....

Late submissions policy	**Days late**	**Penalty applied**
Penalties apply to late submissions and may vary between faculties. Please refer to your faculty's late assessment policy for details.		

Patient/client confidentiality: Where a patient/client case study is undertaken a signed Consent Form must be obtained.

Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations

Plagiarism: Plagiarism means to take and use another person's ideas and or manner of expressing them and to pass these off as one's own by failing to give appropriate acknowledgement. This includes material from any source, staff, students or the Internet - published and unpublished works.

Collusion: Collusion means unauthorised collaboration on assessable written, oral or practical work with another person. Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or nominee, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing.

Student Statement:

- I have read the university's Student Academic Integrity Policy and Procedures
- I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) Regulations (academic misconduct).
- I have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied.
- No part of this assignment has been previously submitted as part of another unit/course.
- I acknowledge and agree that the assessor of this assignment may, for the purposes of assessment, reproduce the assignment and:

  i. provide it to another member of faculty and any external marker; and/or

  ii. submit to a text matching/originality checking software; and/or

  iii. submit it to a text matching/originality checking software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.
- I certify that I have not plagiarised the work of others or participated in unauthorised collaboration or otherwise breached the academic integrity requirements in the Student Academic Integrity Policy.

Date: ...../...../..... Signature:.. KAIYUAN . *

Privacy Statement:

For information about how the University deals with your personal information go to http://privacy.monash.edu.au/guidelines/collection-personal-information.html#enrol

Sound of Top Spotify Songs from 2010-2019

Assignment 1 Report

Name: Goh Kai Yuan

ID: 30881919

Tutorial Class: 4

URL:
https://public.tableau.com/app/profile/ervin.goh/viz/Assignment1_16618347499680/Dashboard2?publish=yes

Number of words: 995

b. A brief description of the domain, Why and Who

The domain that I have chosen is Music. For music to work, music is built up by different components of sound. These components of sound determine how loud, lively, positive and etc the music is. With that, I want to visualize these components of sound by creating relationship between them.

The main intention or the main target for this visualization is to provide insights for the future company music producer or self-produced producer. This will make sure they are able to grasp the interest of the public in their taste of music to produce a high-quality song. Furthermore, the trend of genre in a song can be viewed easily so that music producer can avoid the genre that less people are interested on.

c. What: A brief description of the data

The link to the dataset is: https://www.kaggle.com/datasets/leonardopena/top-spotify-songs-from-20102019-by-year

The author is Leonardo Henrique (Henrique, 2019), and it has been upvoted for 342 times by the time of writing this report. Moreover, it has a usability score of 10 and thus, I know that it is a trusted source of dataset. As per download, the data is already cleaned. However, I manually checked again in Excel form and found that there are no null values. To ensure that it really has no NA values, I have used RStudio which is a platform to perform R codes to check if there are no null values again.

Dataset Type: Table

It has items and attributes

Qualitative:

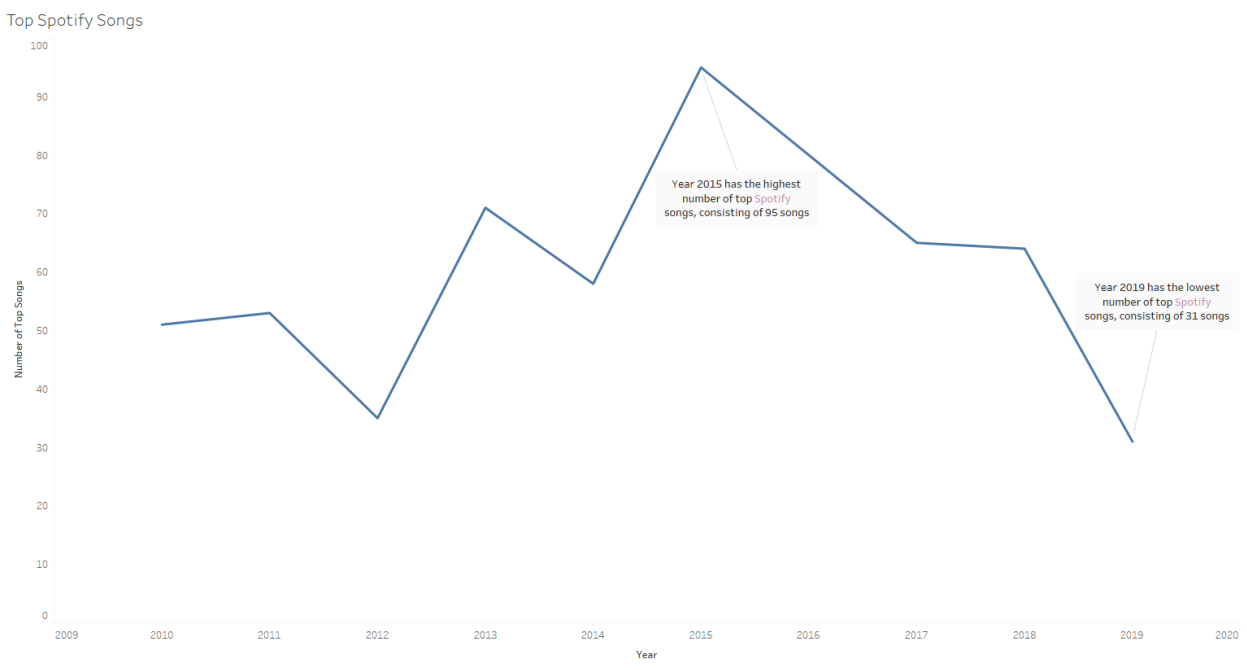
1. Title
2. Artist

Quantitative:

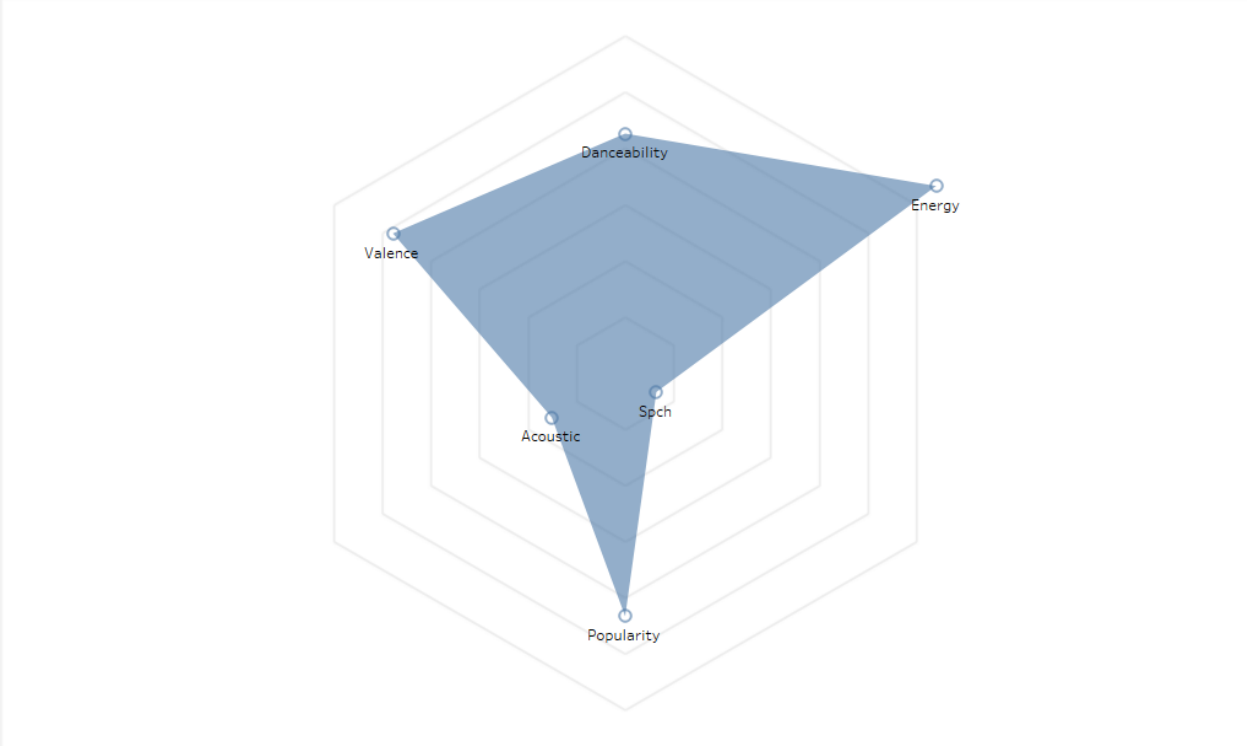
1. Year (ordinal)
2. top genre
3. bpm
4. nrgy
5. dnce
6. dB
7. live
8. val
9. dur

10. acous
11. spch
12. pop

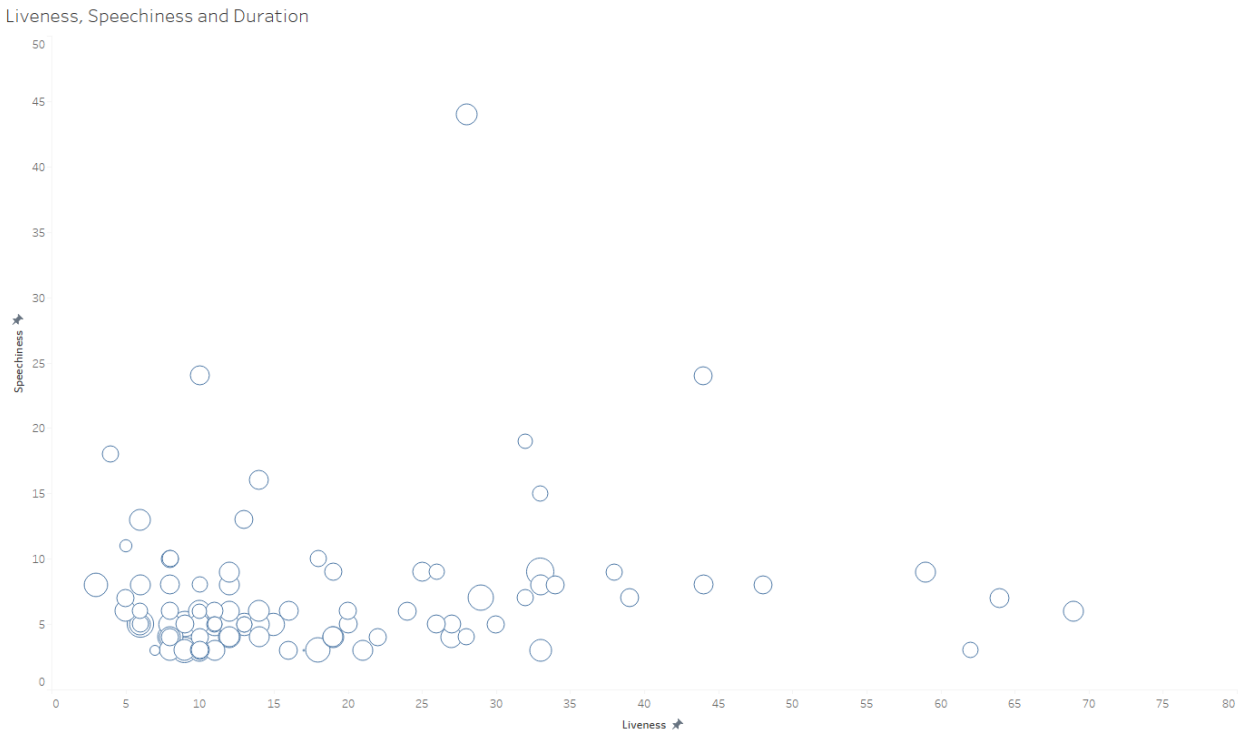
d. Why and How:



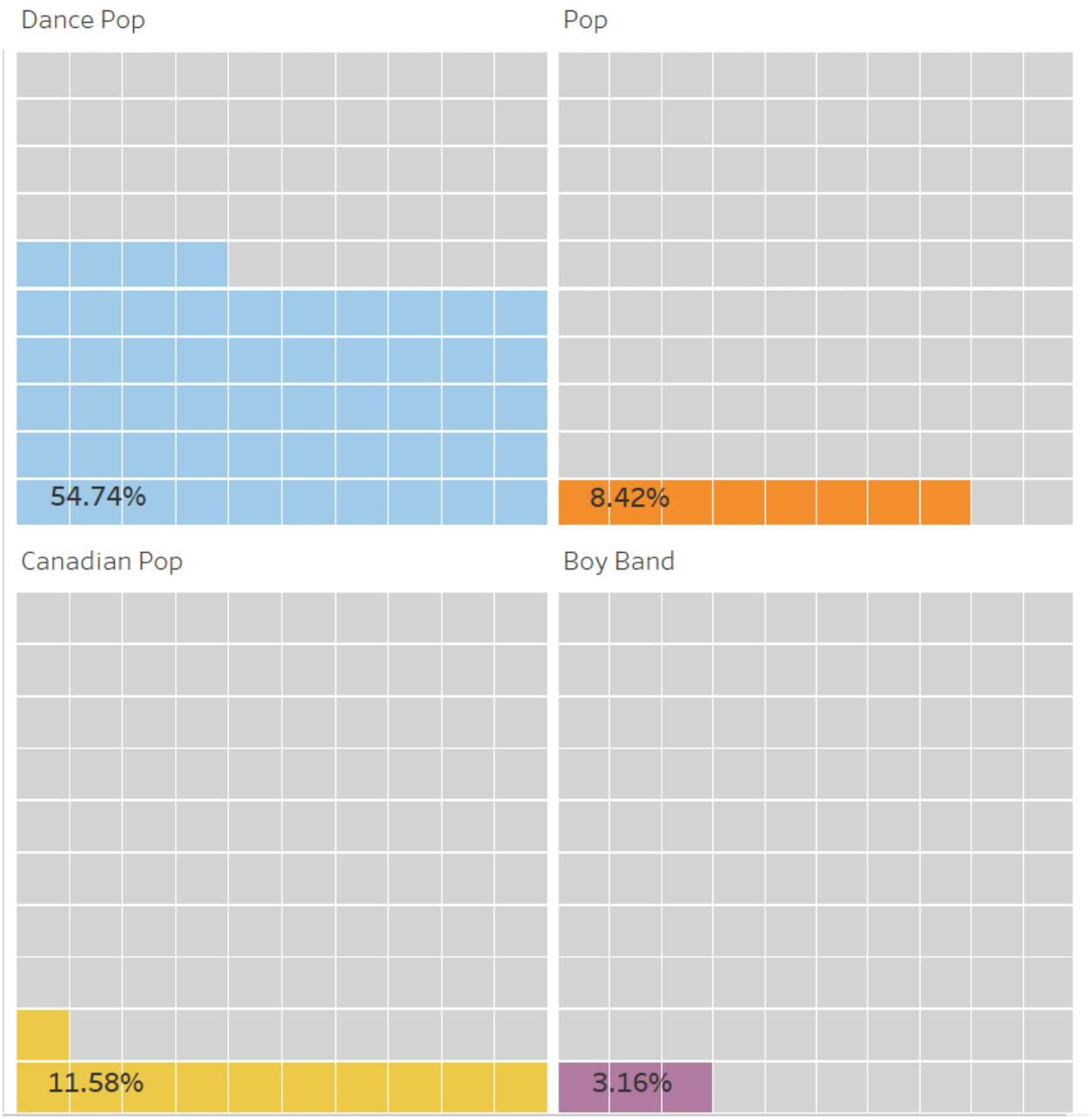
I am choosing this specific idiom to visualize the trendline of Top Spotify Song over the years. Line graph is suitable to view the trend of a product over the year with 1 quantitative attribute and 1 ordinal attribute. User can instantly know that if the top Spotify songs is increasing or decreasing these few years. Annotations have been made to describe the highest number of top Spotify songs and lowest number of top Spotify songs.



I am choosing this specific idiom which is a radar chart as it helps me to visualize the relationship of 6 sound components of the average of top Spotify songs throughout the span of 9 years. (2010-2019). I have learnt to build the radar chart by reviewing the website “Use radar charts to compare dimensions over several metrics” (Trajkovic, n.d.). Radar chart is suitable when I want to compare between different components of attributes. This will help me get an overall relationship between those attributes. For users, they can use this radar chart as a guideline when they are composing a song and implement the suitable of sound component into their music. This can be controlled by the year parameter and show the result of each year.



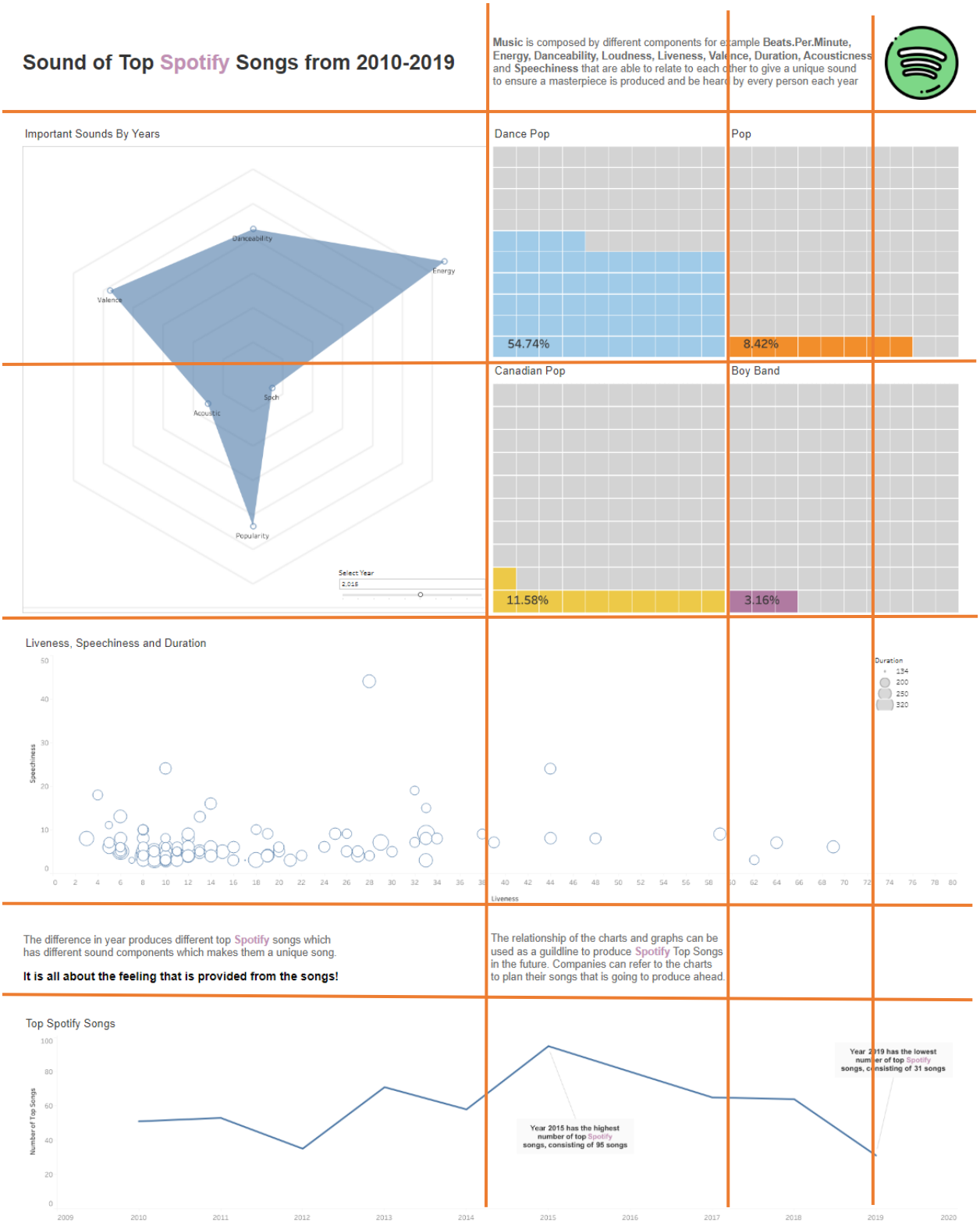
I am choosing this specific idiom which is a bubble plot to help me visualize the relationship of liveness, speechiness and duration. Three of these dataset attribute types are quantitative data type. The x axis shows the liveness and y-axis shows the speechiness. For the duration, it will be determined by the size of the bubble in the graph. The larger the bubble, the longer the duration. For users, they can see a relationship between 3 attributes and that when the whole bubble plot is plotted out, if there is a majority of bubble exists in one corner, it means that it is better to produce a song that follows that corner's attribute value. This can be controlled by the year parameter and show the result of each year.



I am choosing this specific idiom which is a waffle chart to visualize the top 4 top genre of the music produced that are on the top Spotify songs. I have learnt this from watching YouTube called "How to Create Waffle Charts in Tableau" (Kriebel, 2017). This format of visualization makes users understand the visualization better about each genre percentage and they can compose music according to the genre that Is most popular. This can be controlled by the year parameter and show the result of each year.

e. Design

Layout



My radar chart is the Visual Center and I have used the minimum amount of sight lines to build my visualization. There is a symmetry if we cut the visualization into left and right. The visualisation is very balanced arranged. There are no sudden big or small elements that mess up the balance of the whole visualisation.

Colour

The colour I have chosen is colours that are available to be seen by vision-impaired person. The waffle chart is being coloured with different colours and I did not use any red or green colour for those charts. The word Spotify is being bold with purple colour to tell the users that Spotify is the topic here. Blue colour is the main colour that I have chosen for the radar chart, bubble chart and line chart to keep the consistency of colour.

Figure-ground

I create a visual hierarchy by having bolder text for the important words in my visualization. Firstly, Spotify is being bold to show that it is a data from Spotify. Secondly, I have bolded all the sounds components to emphasize that these are the important components to compose a good music out. Title is being bold, and it is bigger than other texts to show that this is the title.

Typography

For the typefaces, I have used Arial for my entire visualization. As Arial is part of Sans Serif, it has better readability on diagrams and maps.

Storytelling

The genre of my visualization is annotated chart. For my visualization, it is mainly about the relationship between each component of sounds. I do have a line chart that has annotations explaining the highest and lowest number of top Spotify songs during 2010 to 2019.

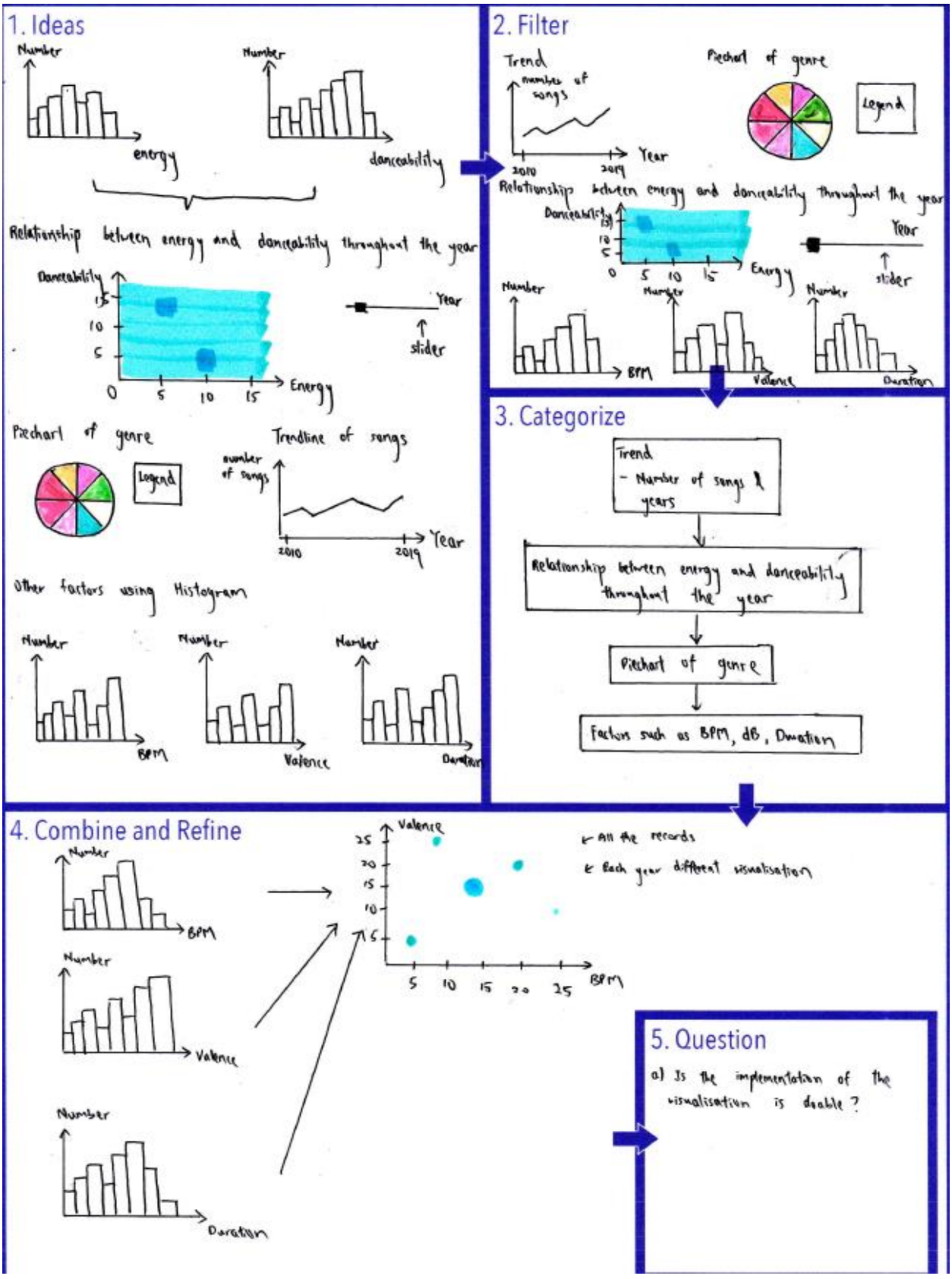
f. References

Henrique, L. (2019). Top Spotify songs from 2010-2019 - By Year. Retrieved from https://www.kaggle.com/datasets/leonardopena/top-spotify-songs-from-20102019-by-year

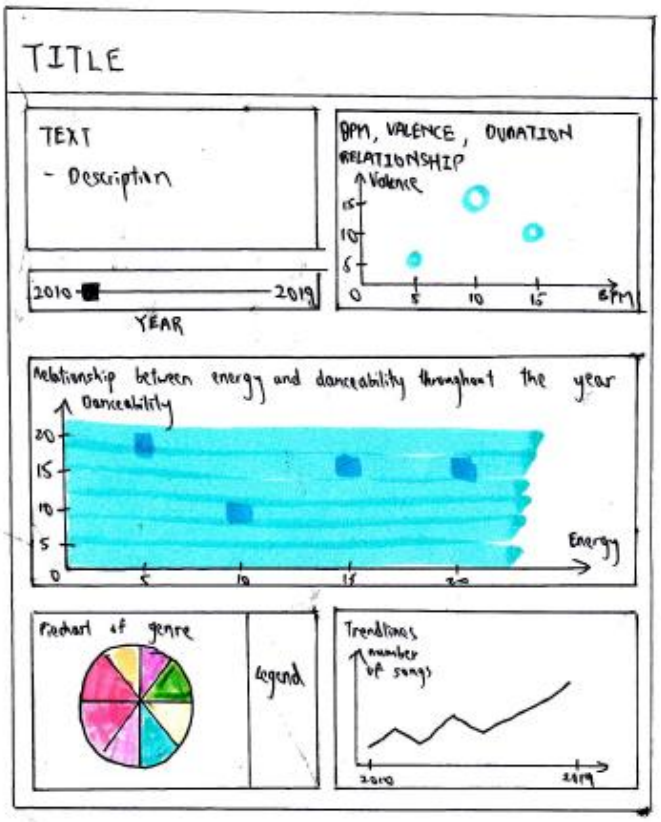
Kriebel, A. (2017, January 17). *How to create waffle charts in Tableau* [Video file]. Retrieved from https://www.youtube.com/watch?v=wRpBkl6-uhU

Trajkovic, J. (n.d.). Use radar charts to compare dimensions over several metrics. Retrieved from https://www.tableau.com/about/blog/2015/7/use-radar-charts-compare-dimensions-over-several-metrics-41592

Appendix

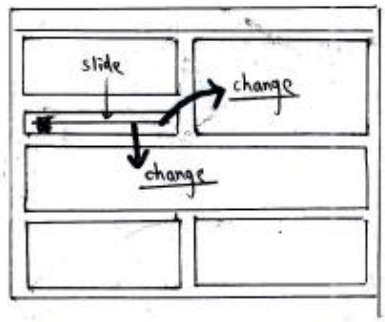

1. Ideas
Number
energy
Number
danceability
Relationship between energy and danceability throughout the year
Danceability
Energy
Year
slider
Piechart of genre
Legend
Trendline of songs
number of songs
Year
2010
2019
Other factors using Histogram
Number
BPM
Valence
Duration
2. Filter
Trend
number of songs
Year
Piechart of genre
Legend
Relationship between energy and danceability throughout the year
Danceability
Energy
Year
slider
Number
BPM
Valence
Duration
3. Categorize
Trend
- Number of songs & years
Relationship between energy and danceability throughout the year
Piechart of genre
Factors such as BPM, dB, Duration
4. Combine and Refine
Number
BPM
Valence
Duration
Valence
BPM
All the records
Each year different visualisation
5. Question
a) Is the implementation of the visualisation is doable ?

Layout



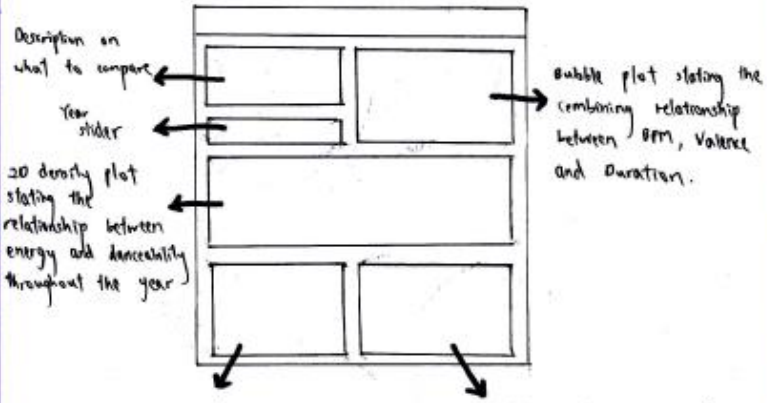
Title: Dashboard view
Auther: Goh Kai Yuan
Date: 24/8/2022
Sheet: 2
Task: Top Spotify Songs from 2010 - 2019 visualizations for Host

Operations



Focus

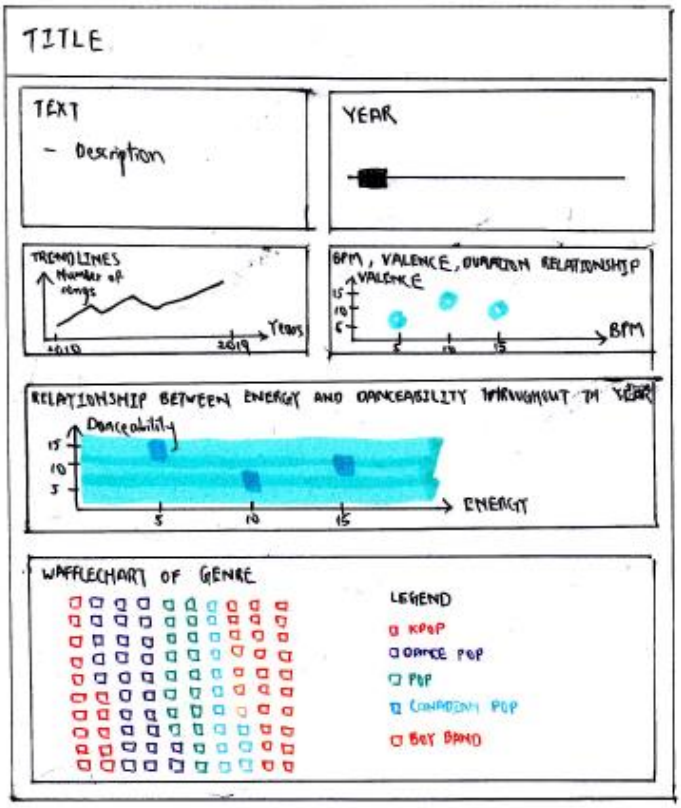
In dashboard, everything is being focused



Discussion

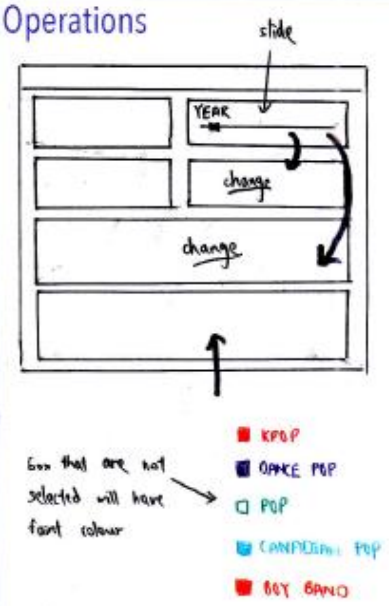
- Less engagement with user
- No symmetry
- No visual center
- Whitespace is used up

Layout

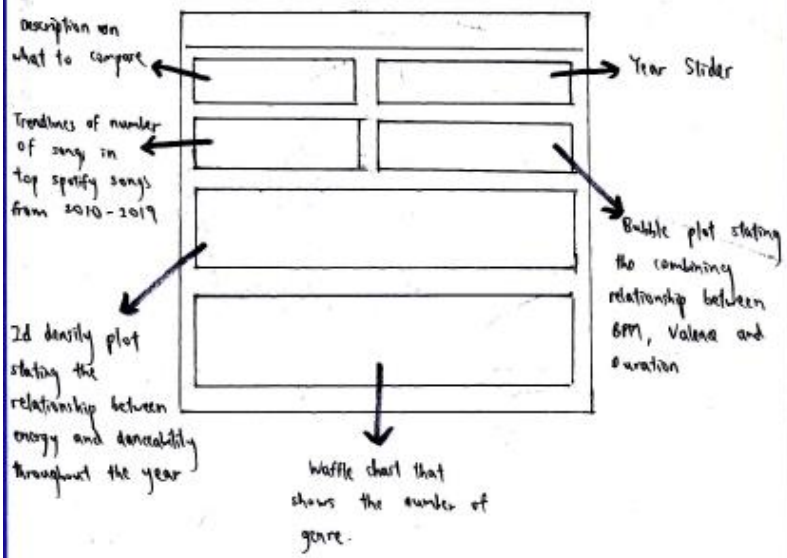


Title: Dashboard - view
Auther: Goh Kai Yuan
Date: 24/8/2022
Sheet: 3
Task: Narrative Visualization for Host

Operations



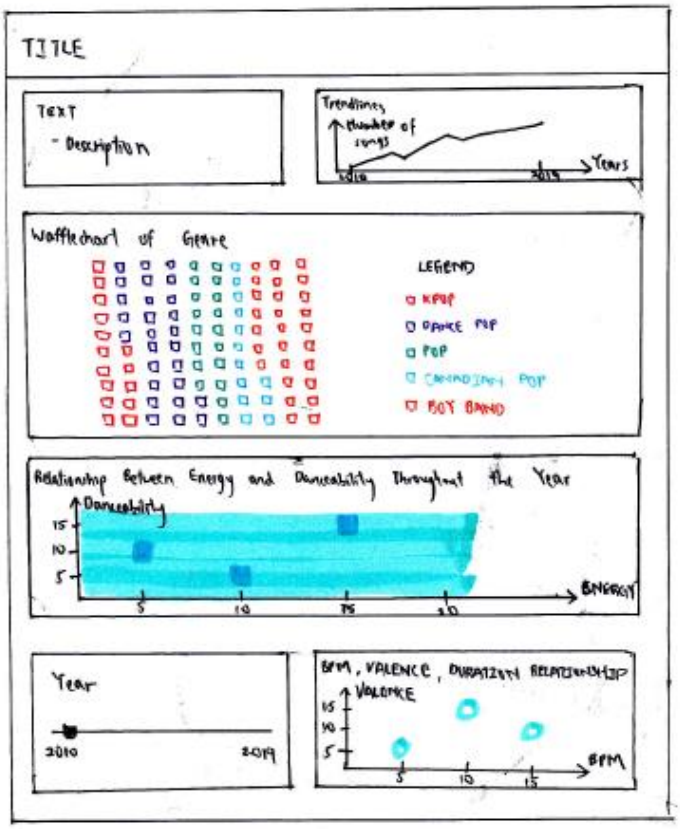
Focus



Discussion

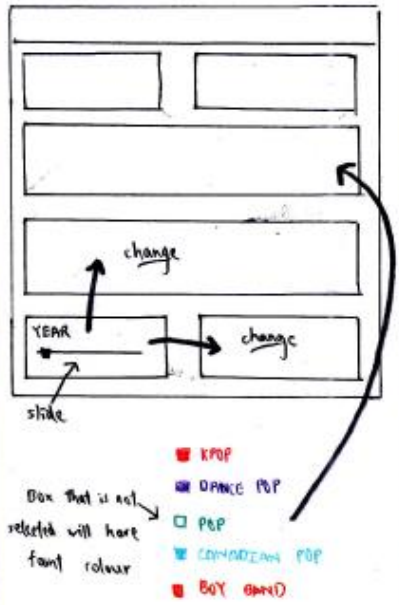
- Symmetry
- Whitespace is used up
- Waffle chart better visualisation than pie chart
- No visual center

Layout

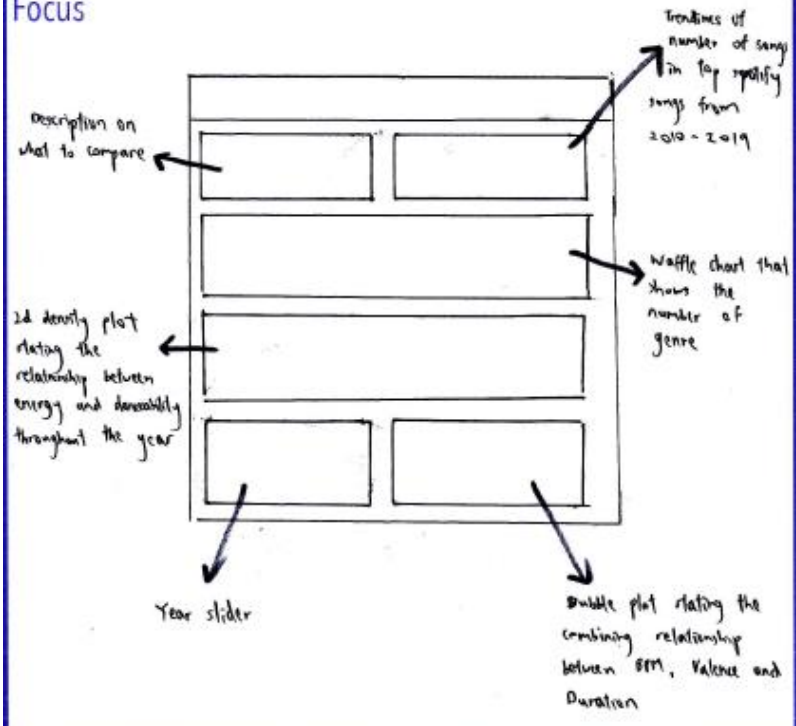


Title: Dashboard view
Auther: Goh Kai Yuan
Date: 24/8/2022
Sheet: 4
Task: Visualization of Spotify

Operations



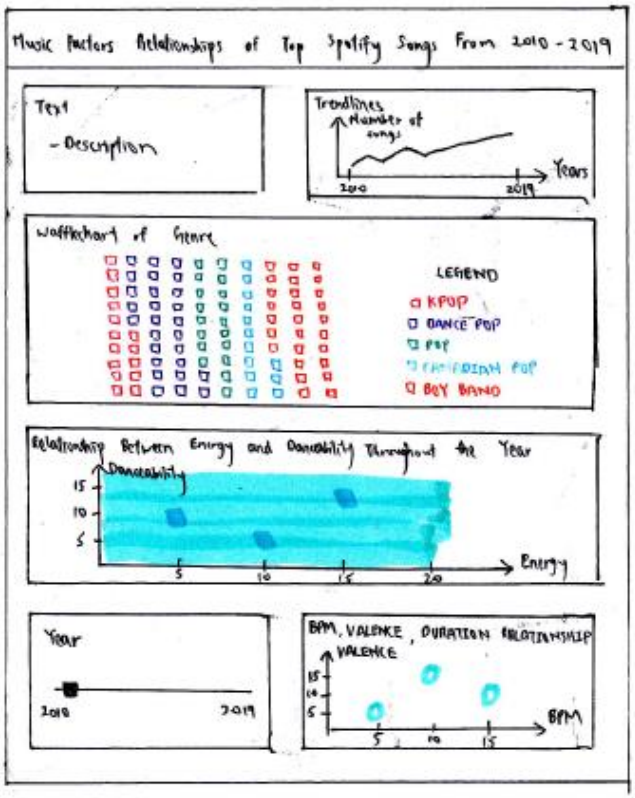
Focus



Discussion

- Have visual center, waffleschart
- Symmetry
- No whitespace

Layout



Title: Final Design Sheet

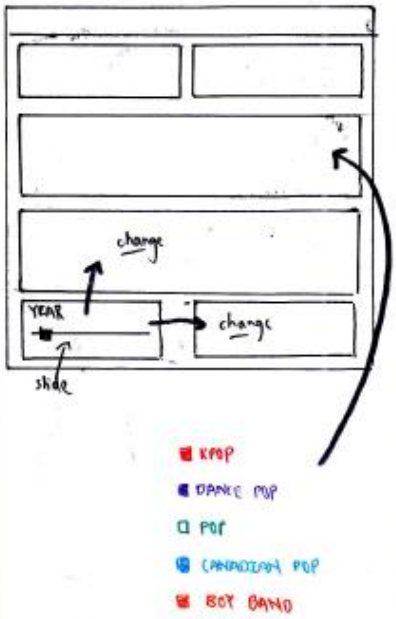
Auther: Goh Kai Yuan

Date: 24/8/2022

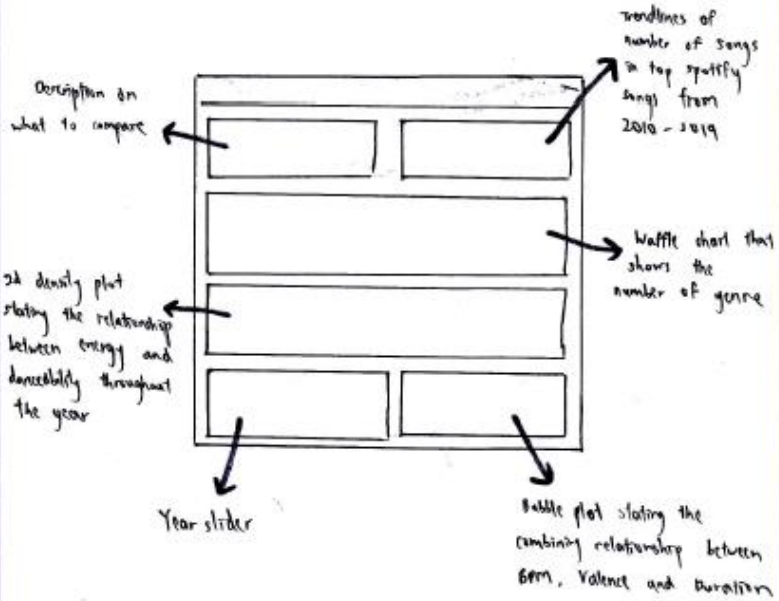
Sheet: 5

Task: Final Implementation Design

Operations



Focus



Detail

- Database implemented using CSV file
- Clean data
- Time to build : 1 week