**Project report – Data Viz Group 3**

*Answer these questions concisely below each other according to the given numbering. Use the Cambria font, no Bold. (length guideline: max 3 pages)*

**Date: 31.05.2020**

**Name of students:**

**Tsz Kin Chau tszkin.chau@student.kuleuven.be**

**Fengan Li fengan.li@student.kuleuven.be**

**Hugo Enrique Montaño Castillo hugoenrique.montanocastillo@student.kuleuven.be**

**Lars Magne Tungland larsmagne.tungland@student.kuleuven.be**

**Martina Verna martina.verna@student.kuleuven.be**

**Xin Wang xin.wang@student.kuleuven.be**

1. **Motivation.**

What is the problem that the visualization tries to explain or solve? Why is this problem important? Why is this problem complex?

Problem: What are songs about and how are they connected

Why: The song lyrics is a key component, what can we discover when ignoring melodies and examining solely lyrics

Complexity:

- big dataset

- word frequency count is on word stems

- lyrics haven’t been analyzed in detail

1. **Data.**

Which datasets did you use? How difficult was it to use the data?

Title dataset 1: Million song dataset

Source: http://millionsongdataset.com/

Data attributes: 46 fields: song name, artist, duration, hotness

License: CCA - Proceedings of the 12th International Conference on Music Information Retrieval {ISMIR} 2011

Title dataset 2: MusixMatch

Source: http://millionsongdataset.com/musixmatch/

Data attributes: ~20 fields: word stem frequency of lyrics

License: CCA

Title dataset 3: TagTraum

Source: http://www.tagtraum.com/msd\_genre\_datasets.html

Data attributes: compilation of tabs from last.fm, musiXmatch, MSD

License: Non-commercial only - Hendrik Schreiber. Improving Genre Annotations for the Million Song Dataset. In Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR), pages 241-247, Málaga, Spain, Oct. 2015.

1. **Inspiration.**

Which existing visualizations inspired you? Which aspects specifically did you find interesting? Why?

(Fill in the table below, or create your own table; Give at least 3 different examples)

Title example 1: Connected bubble chart

Short description:

Source: https://www.conceptdraw.com/How-To-Guide/picture/Bubble-chart-sample.png

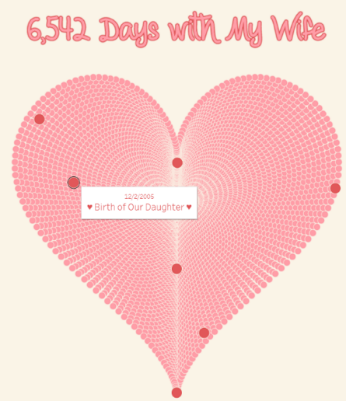
Inspiration aspect + argumentation: natural way to show connections. Show importance among connections by bubble size.

Relevant screenshots (copy / paste image (s)):

Title example 2: Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.

Short description:

Source: https://en.wikipedia.org/wiki/Charles\_Joseph\_Minard#/media/File:Minard.png

Inspiration aspect + argumentation: dense but accessible. Benchmark for good visualization.

Relevant screenshots (copy / paste image (s)):

Title example 3: Hearts

Short description:

Source: <https://www.flerlagetwins.com/2017/> 02/creating-heart-in-tableau\_43.html

Inspiration aspect + argumentation: create visualizations that people can remember / make figurative impact

Relevant screenshots (copy / paste image (s)):

1. **Visualisation design and interaction.**

What does your visualization look like? What are the interactive possibilities? What is the narrative with which the user can use the visualization and understand the insights?

Screenshots (copy / paste image (s)):

Short description:

Interaction:

Narrative:

Insight:

1. **Effort.**

How much effort did it take to implement the project? How did you distribute it in your group?

It required substantial effort and collaboration to deliver this project. We spent significant effort cleaning the data and cutting it into usable slices, which was repeated as we found which preliminary visualizations did or did not work. Additionally, we were required to learn new tools to execute the visualizations, all of which through many rounds of revision.

High-dimensional data. Balance between preserving insights and overloading charts.

**Xin** – Proposed this MSD/MM dataset and initial avenues of exploration. Data cleanup and prep on song title data as rest of group used song lyric data. Learned D3 wrapper for chord (ref: <https://shahinrostami.com/> and <http://holoviews.org/>). Implemented chord diagram.

[Lars, I also want to recognize your efforts in organizing the meetings with agenda/notes/actions. – Jim]

1. **Reflection.**

How is your work related to the topics taught in the course? If you would start over, what would you do differently?

[Do we have the checklist of key points from lectures and such? – Jim]