Lucy Boyle

2036340B  December 2018

Music curation and analytics project

Please find all relevant data and evidence of lab tasks at:

<https://github.com/LucyBoylex/MCA>

Week 1:

The week 1 task was an introduction to the three types of data relating to music curation and analytics: notated, acoustic and descriptive. This report will focus on these three types of data and how the insights gained from the curation techniques and analyses in the lab tasks relate to these three data types.

The first task focused on choosing a theme for the dataset, as well as analysing how the theme related to the three data types. I chose Swedish pop group Abba to concentrate on throughout the dataset as they have a wide range of music and are readily available on streaming and music score websites. It has been widely curated and stored and it can be found in notated, descriptive and acoustic formats easily online. If we look at Apple Music for example, we can see it holds acoustic and descriptive data on the music it has. Descriptive data being the way the music is displayed on screen, showing the artist, song, duration, album, producer etc. Acoustic data being the actual sound file and the way we hear the songs played. Apple music does not, as far as I can see, hold any notated data. However notated data is easy to find online for the works of Abba. Looking at sheetsdaily.com I was able to find scores of Abba music for several instruments and songs.

Week 2:

Week 2’s task required us to find a piece of music and transcribe it using the Musescore2 software. I spent a significant amount of time reading and watching tutorials[[1]](#footnote-1) online before properly starting to transcribe my piece. I decided to transcribe Dancing Queen from a score I found on sheetsdaily.com[[2]](#footnote-2). I chose this piece as it is complex enough to be able to use several features in Musescore2 whilst transcribing it. One of the insights I gained from this task, still being an introduction to music curation, is how labour intensive and time consuming it is to digitise a score originally in pdf form. I also gained a fair understanding of how Musescore2 functions and after using it for a few days was able to confidently use the keyboard shortcuts. The limitations I found with Musescore2 included; how long it took to transcribe a piece of music, as well as the difficulties I encountered as a novice trying to add things like a slur – which looks the same as a tie, however is added to the score in a completely different way. Other things; such as adding an extra line as the score had three lines of music, and changing a treble clef to a bass clef proved difficult and time consuming. I also had great difficulty figuring out how to have 3 notes in the same chord, with one note being longer than the others as the programme didn’t allow this. Only after looking at online forums did I discover the feature of having voices 1-4 which allows this action[[3]](#footnote-3).

Overall I found Musescore2 to be an excellent piece of software for the digitisation of music scores, however it does have the potential for improvement. As a music novice I found it very complicated to use, and looking at the guide online only helped to a certain extent as it is evidently geared to those already knowledgeable about music. I would improve this service by further simplifying their tutorials to include a guide on what each term actually means. In terms of the actual software I would improve it by adding a corrections feature which explains *why* a certain action is disallowed. For example when I was (mistakenly-due to lack of musical knowledge) trying to use a tie in place of a slur; it would change my second note to be the same as the first. A useful option would be an optional pop-up explanation of why the programme has done this.

Reflecting on the use of Musescore2 in this project, I enjoyed using the software and it gave a good insight into the digitisation of notated data, as well as displaying some acoustic and descriptive data.

Week 3:

In week 3 the lab tasks involved exporting the musescore files to XML and MEI format. The Music Encoding Initiative (MEI) is a community-driven effort to define a system for encoding musical documents in a machine-readable structure[[4]](#footnote-4). This was useful in that it allowed a further understanding of the metadata behind the descriptive data. I was able to see within the MEI-encoded file the metadata contained within the <meiHead> tag[[5]](#footnote-5). This metadata includes the title of the piece as well as information such as how it was converted to MEI – in this case using Verovio. The MEI file allowed me to see how music is encoded, from an acoustic, descriptive dataset to a notated form which is machine readable. By looking at the <music> tag I was able to see how each note is represented in the code.

As this task mostly focused on metadata and encoding, the most prevalent of the three data types was notated data, although I was able to see how the metadata could be interpreted as descriptive data when looking at the code.

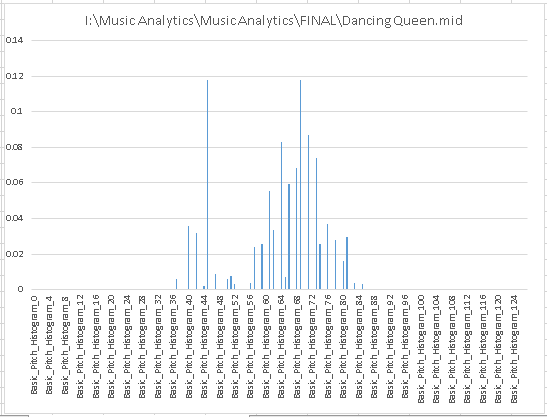
Week 4:

In week 4’s lab tasks we looked further into the basics of encoded music by first comparing MusicXML to MEI using Oxygen. In a group we were able to conclude that the strengths in MEI are the readability of the file, to both computers and humans and that there is longevity should the technology become obsolete as we can still read the notes and transcribe to relevant technology. Weaknesses are the limitations that still exist, as the technology is still not fully developed. We were also able to make a decision about when the two formats are best to be used. We concluded that XML is best for storing data if it only needs to be read by a computer and not actually played, whereas MEI is better for being read by both humans and computers. This was another example of notated data and how it is used in the curation of music.

In week 4 I also created my Github repository and began to add metadata to my index-meta.html page. This task was a useful introduction into how Github works.

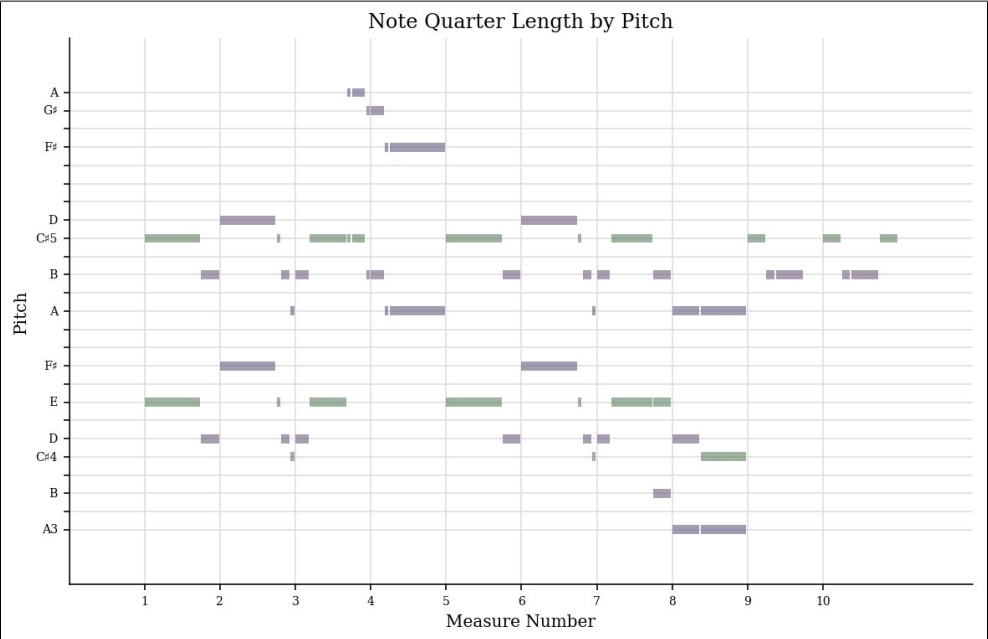
Week 5:

Week 5’s lab tasks were an introduction to the analytics part of the course. I used jSymbolic, Music21 and Jupyter Notebook to find and analyse different parts of the music score I created in Musescore2. In jSymbolic I selected features I thought were most relevant to my piece, such as pitch frequency. The software allowed me to extract this information as a CSV file, which I opened in Excel. I produced this graph from the data:



The graph provides a visual aid in determining which pitches are most often played in this piece. There are two significant spikes in the histogram; indicating the two pitches most frequently played.

Also using jSymbolic and Music21 I was able to create a piano roll:



This was another way of looking analytically at the makeup of the piece of music, this time focusing on the note quarter length by pitch for the first ten measures of the piece. The piano roll is an effective way of visualising notated data, once again making it ‘human-readable’. The limitation with this visualisation is the fact it only shows the first ten measures. I was able to extend this to the first twenty measures however, I could not personally find a way to visualise the whole piece as a piano roll. This could evidently be improved by making the piano roll show all the measures in the piece by default.

Week 7:

Week 7’s tasks were centred on descriptive data and some of the basic methods for describing digital content. I added important metadata to my MEI files created in Verovio. I chose to include the song title, the date of recording, the songwriters and producers as well as the release date. I found working with metadata to be an important part of this project and an important part of music curation and analytics as a whole, as it gives a deeper understanding of the concept of descriptive data as well as providing valuable data information in the files themselves. The use of MEI as a medium for displaying metadata is appropriate as it ‘enables the integration of detailed metadata with the full music text, including variants and emendations within the same file in a format that is interchangeable with other software such as a graphical note editor’[[6]](#footnote-6). This ability to interchange the metadata between different file formats is one of the main strengths of using MEI for the integration of metadata into encoded music.

In my index-meta.html I was only able to get the first ten measures of the transcribed piece to show, despite having the whole piece as the MEI file, which I consider to be a limitation in my final project as I would have liked to have been able to display the whole piece. I have however included the whole mscz file in my data repository. Another difficulty I encountered was trying to get my metadata elements to display correctly. It took several attempts and some editing of the code to make them work. Unfortunately, in my index-meta.html page there is a closing tag displayed after the recording dates, which I could not manage to find and remove. Due to time constraints this is not a good example of a page containing good metadata and I would improve it by editing the format of the displayed metadata to make it more user-friendly and readable, as well as removing the visible closing tag and making sure the whole score was visible.

Week 8:

In week 8 the task focused primarily on acoustic data and how to extract important technical metadata from audio tracks. To do this I used the Sonic Visualizer software, as well as the ‘properties’ viewer in Windows. This gave me an insight into the kinds of technical metadata that are generally stored alongside an audio file. I sourced three tracks on freemusicarchive[[7]](#footnote-7):

* Driving Home by DeeYanKee
* Frogs and Stoned Funghi by KieLoBot
* Driven to Success by Scott Holmes

All of these tracks were of different genres; to allow an effective comparison of different types of metadata. Using the Sonic Visualizer allowed me to extract a waveform and spectrogram visualisation of my three pieces of music. The waveform displays the amplitude or level changes throughout the song[[8]](#footnote-8) and the spectrogram displays much the same information; showing the frequencies which make up sound, from low to high[[9]](#footnote-9). The difference lies in their appearance and interpretability. I believe the time-frequency spectrogram based analysis has a visual advantage over the waveform based analysis as it is easier to view the frequency data it contains.

The Sonic Visualizer software was overall very efficient and displayed the data well however I found that it had some limitations. I found that it doesn’t seem to be possible to export more than one layer/pane at a time, whether it be exporting as a CSV or an image file. In the interests of comparison and analysis, it would be useful to have this feature, even if just to export two or three panes as one image file.

Week 9:

In week 9 we continued to look at audio features and acoustic data, again using Sonic Visualizer. We also used Python to create histograms representing features from the three tracks we had chosen.

In Sonic Visualizer I was able to extract the Spectrogram, Chromagram and Mel Frequency Cepstral Coefficients. The Spectrogram made a visualisation of the frequencies and levels which made up the songs, the chromagram displays the tonal data in the song and the Mel Frequency Cepstral Coefficient (MFCC) displays the timbre, or tone quality of each of the tracks. This was another example of extracting important metadata from acoustic data. I saved the three layers for each of the three songs as both an image file and a csv. Unfortunately the CSV files for the Spectrograms were around 630,000kb and so I couldn’t add these to my repository.

The second part of the task involved computing histograms of the three tracks in Python. I chose to compute the histogram for the MFCC based analysis. It was interesting to see the variance in the three histogram groups. However, I once again encountered problems. I could only get half of the Histograms for the Dee Yan-Kee track to show. This is an example of a limitation I found while visualising this data. In looking at the three sets of Histograms side-by-side it was clear to see that they had come from three different styles of music.

Week 10:

Week 10’s task was an interesting insight into the limitations of music curation software. This task also allowed me to develop further confidence in Python. The first task allowed me to create and view a similarity matrix, to which I added my three chosen tracks. I chose to use the same three tracks I had used for the histogram and data visualisation tasks as it allowed continuity of the information I was gathering. I encountered difficulty in rewriting the code to include the three tracks with different genres and it took me several attempts to correct this. This task was very insightful as it demonstrated again, the variety of ways in which acoustic data can be visualised as metadata.

The second task particularly interested me as it showed clearly the limitations that still exist in digital music curation. I found there to be a huge difference between the original composition and the generated transcribed version. It lost almost all semblance to the original and was essentially unreadable. This is evidence of how music curation and analytics is still very much a work in progress.

Conclusion:

In conclusion, this project was very insightful and challenging. I feel I have developed an understanding of the different ways in which notated, descriptive and acoustic data can be viewed and analysed. There are several limitation with the software encountered during this project, as there are with most pieces of software, however overall I believe each piece of software worked well for its intended use and I learned a great deal.

Bibliography:

* Niels Krabbe and Axel Teich Geertinger, “MEI (Music Encoding Initiative) as a Basis for Thematic Catalogues: Thoughts, Experiences, and Preliminary Results,” Presentation for the 2012 RISM Conference (June 2012), available online: http://music-encoding.org/downloads/TeichGeertinger\_Final.pdf.
* Musescore tutorials, last accessed 07/12/18, available: https://musescore.org/en/tutorials
* Sheets Daily – Free Sheet Music, last accessed 07/12/18, available: http://www.sheetsdaily.com/piano/sheets/44065/ABBA\_Dancing\_Queen.html
* Musescore User Forum, last accessed 07/12/18, available: https://musescore.org/en/node/17111
* Music Encoding Initiative, last accessed 07/12/18, available: https://music-encoding.org/
* Music Encoding Initiative, last accessed 07/12/18, available: https://music-encoding.org/resources/introduction.html
* Free Music Archive, last accessed 07/12/18, available: http://freemusicarchive.org/
* What are waveforms and how do they work? Last accessed 07/12/18, available: https://soundbridge.io/what-are-waveforms-how-they-work/
* Spectrogram, last accessed 07/12/18, available: https://musiclab.chromeexperiments.com/Spectrogram/

1. Musescore tutorials, last accessed 07/12/18, available: https://musescore.org/en/tutorials [↑](#footnote-ref-1)
2. Sheets Daily – Free Sheet Music, last accessed 07/12/18, available: http://www.sheetsdaily.com/piano/sheets/44065/ABBA\_Dancing\_Queen.html [↑](#footnote-ref-2)
3. Musescore User Forum, last accessed 07/12/18, available: https://musescore.org/en/node/17111 [↑](#footnote-ref-3)
4. Music Encoding Initiative, last accessed 07/12/18, available: https://music-encoding.org/ [↑](#footnote-ref-4)
5. Music Encoding Initiative, last accessed 07/12/18, available: https://music-encoding.org/resources/introduction.html [↑](#footnote-ref-5)
6. Niels Krabbe and Axel Teich Geertinger, “*MEI (Music Encoding Initiative) as a Basis for Thematic Catalogues: Thoughts, Experiences, and Preliminary Results*,” Presentation for the 2012 RISM Conference (June 2012), available online: http://music-encoding.org/downloads/TeichGeertinger\_Final.pdf. [↑](#footnote-ref-6)
7. Free Music Archive, last accessed 07/12/18, available: http://freemusicarchive.org/ [↑](#footnote-ref-7)
8. What are waveforms and how do they work? Last accessed 07/12/18, available: https://soundbridge.io/what-are-waveforms-how-they-work/ [↑](#footnote-ref-8)
9. Spectrogram, last accessed 07/12/18, available: https://musiclab.chromeexperiments.com/Spectrogram/ [↑](#footnote-ref-9)