1. Identify a theme for your dataset. This will be the central, cohesive strand that will bring your data together. It can be an artist, composer, songwriter, performer, album, larger work (i.e. opera, musical, etc.), time period, genre, archival grouping, etc. You will want to be sure that you can gather all three types of data based on your selected theme.

Musicals

notated, acoustic, or descriptive data

1. Describe the current manifestations of data relating to your selected theme. Has it already been curated somewhere? How is the data presented, described, and analysed?

There are many curations of musical theatre music. The New York Public Library has a curation of acoustic music data which is notoriously difficult to gain access too. It is contained within the Theatre on Film and Tape Archive which films Broadway shows including musicals. The music is put into the context of the show. To access the data you have to be a student, academic or have access to the library through location. You can only have access to one show at a time and can only ever watch the show once. There is only one recording made and you have to be in a special room. This makes it incredibly difficult to compare shows as you cannot view more than one show at a time.

Another curating site Internet Broadway Database (IBDB) has descriptive data. It has a record of many Broadway shows, the songs performed, information about the cast recording (such as producers, artists, number of songs and date of release) and composers. This database tries to give all the descriptive information needed to be able to find the music within any particular show.

The library at the University of Illinois at Urbana-Campaign has many notated scores for musical shows.

New York Public Library (Theatre on Film and Tape Archive) – Acoustic

IBDB (Internet Broadway Database) – Descriptive

CastAlbums.org – Descriptive

University of Illinois at Urbana-Campaign (University Library) – Notated

Hal Leonard - Notated