Is Mozart Really That Popular? An Analysis of New York Philharmonic's Performance History and Concert Programming

Daniel Liu

Spring 2025 | Digital Humanities Independent Work Seminar

Professor Brian Kernighan

Motivation and Goal

In every form of art, there exists the concept of the "cliché", the overrated, and in the case of classical music, the overplayed. One hears the opening to Mozart's Eine Kleine Nachtmusik or the dot-dot-dot-dash pattern of Beethoven's Fifth Symphony and maybe immediately rolls their eyes—but to what degree is such criticism valid? Are certain composers or pieces really played that frequently in concert halls, and how does that correlate to audience attendance at the symphony? On the one hand, pieces such as those may be overrated to the point of audience boredom and disinterest. On the other, maybe the pieces are cliché from overprogramming due to them being crowd pleasers; even orchestras care about profit and continuously including certain composers are a method of selling more tickets. The goal of this project is to quantify such claims, analyzing in particular the New York Philharmonic's performance history, in part due to their extensive and thorough documentation of every concert from the very first on December 7th, 1842 until today. Is it true that specific pieces are reprogrammed time and again? Did a composer fall out of fashion in the 1800s but come back frequently in the late 1900s? Are there certain pieces of a composer that are more programmed than others? These questions and more can be answered with an exploratory data analysis of the NY Philharmonic's archives, and an additional correlation can be drawn to popularity due to the assumed revenue motive of the institution. This might clue into larger musical trends throughout the years since it spans from middle Romanticism to contemporary modernism, and it would be helpful for any musicologist looking through a historical lens at classical music. Amongst the ongoing discourse of featuring more contemporary or underplayed and historically underrepresented composers such as women and people of color, this may lend insight into how to structure a concert program and how best to create interest for and expose audiences to non-canonical classical music.

Problem Background and Related Work

Little quantitative work has been done on the issue of concert programming. There are guide-esque papers such as Mark Gotham's "Coherence in Concert Programming" discuss the strategy and social consciousness behind programming choices, acting both as a study

and a didactic on how to best program a concert. Samuel Gilmore's "Tradition and Novelty in Concert Programming" performs a cultural analysis of classical music concerts, asking the question, how can music directors best balance the old and familiar with the new? What my project hopes to accomplish is to fill in the quantitative gap between these studies, providing a verifiable analysis of an open-source database and performing large-scale trend analysis that will hopefully help musicologists and music programmers alike.

Approach

As an exploratory data analysis project, I will begin by retrieving all of the program data from the NY Philharmonic archive and processing that into a dataframe. From there on, I plan on using Python to search for and analyze specific patterns in the data, in addition to modeling them in ways that are accessible for a non-technical person, for example a musicologist.

Plan

Some consideration will have to be given with regards to how exactly to represent the data, given that each concert has a variable number of pieces, and if I want to be thorough for additional potential research questions later, it would be nice to include information such as the length of the piece, the type of the piece (symphony, concerto, overture, etc.), or the instrumentation, which may require cross-referencing data from other sources such as imslp.org. That may be slightly risky, as I would most likely have to web-scrape that data, but it is doable. Besides that, my plan is to create a custom class or data structure that is able to store all of this information, while also being easily iterable and sortable. I foresee figuring out exactly how to structure the data to be the most time-consuming part of the project.

Time allowing, I would love to create an interactive website of some sort that would allow musicologists to see trends for themselves and filter/compare/contrast the data as they desire.

Evaluation

As an EDA project, especially one with a dataset that has been meticulously organized and cleaned, I do not anticipate error or accuracy statistics being of particular use to the analysis. The most interesting part will be after data processing: what questions about concert programming are most useful for a musicologist to be answered? Perhaps the sheer number of concerts does not matter as much as the number of concerts featuring an underrepresented composer. I plan on polling (informally) peers and faculty in the music

department as to what information would be most interesting for them. If I have time to make the interactive website as mentioned previously, I would also be sure to ask what features would be relevant or useful and what models would be most comprehensible for someone with no technological background.

Gilmore, Samuel. “Tradition and Novelty in Concert Programming: Bringing the Artist Back into Cultural Analysis.” Sociological Forum, vol. 8, no. 2, 1993, pp. 221–42. JSTOR, http://www.jstor.org/stable/684636.

Gotham, Mark. “Coherence in Concert Programming: A View from the U.K.” International Review of the Aesthetics and Sociology of Music, vol. 45, no. 2, 2014, pp. 293–309. JSTOR, http://www.jstor.org/stable/43198649.

This proposal represents my work according to University Guidelines. /s/ Daniel Liu /s/