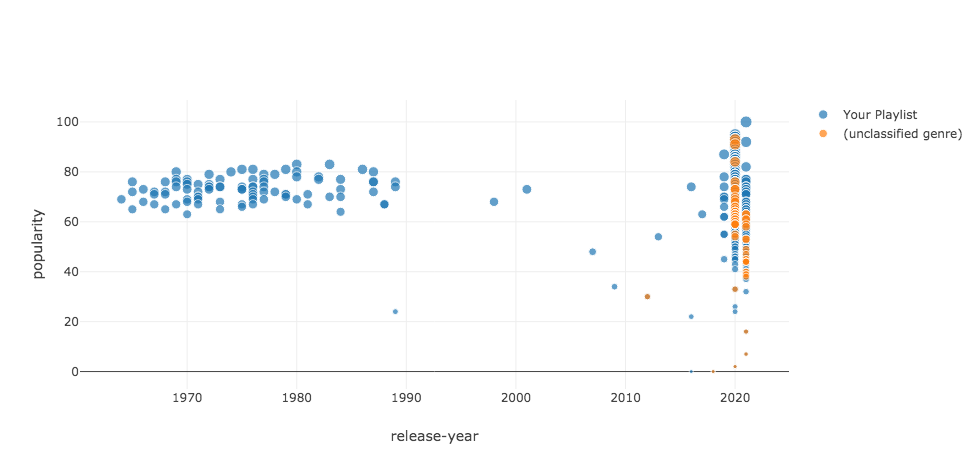
**Results**

For the first part, I moved from a previous research’s statement: “Spotify is the most listened user with his playlist system”. I started to question about this as we are not talking about albums, neither collections. We are talking about algorithmically selected contents.

This point of view is really far from the “analogical era” model of cultural education, it delivers the content (it feels to be) the most appropriate. Taking a step back from the *Accuracy Level* of these learning methods I wondered about the model itself and its limits. It is true that Data Science and Machine learning gives us the best way to empower these tasks but is still important that humans tells them *where* and *how* to look at things and their regularities. For this reason, I started a process of Data retrieving via the Spotify API, asking for the playlists Spotify produced. From this large set of playlists (Is already easy to see that there are some national-markets oriented and global. A company’s choice of course, but how valid in terms of cultural diffusion in globalized era?), I asked to get back just the first 10 because my research was oriented to obtain simple answers and not to train a new model or find patterns.

Just this list gives the idea of the variety of requested music. Next step was to retrieve information about every track in each playlist, to obtain the set of available metadata of the songs. Now a look at the variables reveals (part of) the intercurrent factors in this clustering model. In this early step of research, my interest was not to find patterns or different degrees of correlation, for mere exploratory purpose, I just wanted to verify a condition: Spotify’s model tends to give back to the user results in terms of their popularity. I merged the csv files I get from the playlists to create one dataset and analyze it more clearly. After different tries for a better visualization I choose to transform my raw csv files into a playlist back on Spotify. This allowed me to use an online tool that plots Spotify’s data by its *density* with assignable x and y among available variables. Now it has been possible to look at data about the condition to verify and get a better look of how this model reflects to daily use. It is interesting to see, a first answer to the question about the accuracy of the parameters, the dataset should be cleaned and fixed firstly. One error, for example is clear searching by *Release-year* the 00’s decade is present, but 2 of the 3 items considered are repress versions of 70’s and 80’s tracks. Other thing to notice is the presence of 194 different genres that cluster these 646 elements, some of them are categorized under *unclassified genre.*



As we can see, *unclassified genre* songs are not a relevant number (67, to be exact) but this shows that the definition of items is faulty under different aspects. To answer the initial question, we can notice that effectively this model prefers Most Popular items as 599 have a popularity ≥50 in a dataset of 646 items.

