



Python project



The state of the American music industry in recent years

Teacher : Olivier CROCE

Introduction:

The United States is the biggest Music Market in the world, as listed by the International Federation of the Phonographic Industry (IFPI), based on the retail value it generates each year. As relatively new streaming platforms (such as Spotify, Apple music etc) keep gaining new listeners, it is mandatory for labels to know the current industry trends in order to know how much is worth investing where.

That is why we will focus on trends, evolutions and the real prevalence those streaming platforms have had in recent years (2017-2019) in the United States, according to Billboard's numbers.

In this analysis we will focus on Total Album Consumption (calculated thanks to **Album sales**, **On-Demand Streams**, and Total Song consumption).

Album sales themselves are sorted by whether they are Digital or Physical (Physical contains CD, Vinyl and Cassettes).

A "**sale**" per se, means there was the act of buying a specific song or album.

Every consumption of songs and albums through streaming platforms, (that does not include specific payments for individual songs/albums), is considered as an **On-Demand stream**, whether it is through a paid subscription or an ad-supported service.

For every .txt file provided, each element of each line is separated by the "\t" tabulation, and each line separated by "\n".

Units of the values in .txt files are always millions.

Write one script for each part

PART 1 – Overall view of the US music industry market from 2017 to 2019

Whether you chose to use the tabulation characteristic of the .txt file, or just copy the values you need, is up to you in this part.

Using A.txt:

- Calculate evolution (in %) of Total Album Consumption and Total Song Consumption from 2017 to 2019, display that one number for each Total Consumption.
- Then write a function that calculates the evolution (in %) from one year to another (2017 to 2018, and 2018 to 2019) of the Total Album Consumption. Display the result for each year in a new line in a .txt file named B.
- Then calculate the coefficient between the last two evolution percentages, in order to model the evolution percentage that would occur from 2019 to 2020; and finally, calculate what the number for Total Album Consumption would be in 2020 according to this rate. Display that number at the end of the B file in red.

PART 2 – Focus on Physical Sales in the United States: paradox between biggest physical format seller compared to the current sale trends

Using C.txt:

- Write a script that goes through the C.txt file, reads each line as a new separate list with its own 4 elements (1st- being format and 2nd-3rd-4th- being percentages), focuses only on 2019 then finds and displays the format of physical sales that is most consumed out of the three that year.
- Continue script and add a widget, so that when you click on the button “Which format of physical sales is most consumed in 2019?”, it displays the answer previously found.
- Display graph showing 3 trends, the evolution of the percentage of Cassettes (1st), Vinyl (2nd) and CD (3rd) sales- amongst Total Physical Album Sales from 2017 to 2019 (with a one-year interval). Highlight the Vinyl trend line in red.
- Write a function that, for each three formats, finds out whether there was an increase for its total percentage’s portion from 2018 to 2019. It should display “Increase in **[format]** sales in from 2017 to 2019”, or display the decrease, or display there has been a relatively insignificant evolution.
- Is the evolution for one of these three format’s portions surprising considering the world’s ever-growing digitalization? If so, it should be displayed with a sentence.

PART 3 – The prevalence of On-Demand streams in the American music industry

Using D.txt:

To count the Total Album Consumption, it is considered that **1500 On-Demand streams** are worth as much as **1 album sale**.

The “D.txt” file contains the Total Album Consumption (TAC), and the number of On-Demand streams (OD str) in 2017, 2018 and 2019 (units in millions).

- Write script that adds a new column named “OD to AS” to the D.txt file, (use “\t” tabulation). In this column will be calculated the equivalent of On-Demand streams into Album Sales for each of the three years.
- We know that $\text{Total Album Consumption} = \text{Album Sales} + (\text{On-Demand Streams}/1500) + (\text{Song Sales}/10)$

Thanks to the newly found data in D.txt file, find the percentage “a” that On-Demand streams represents out of Total Album Consumption (TAC) in 2019, display that result.

Make a pie chart of that percentage (with example given in PieChart.txt file if needed; the *ONLY* variable that has to be set is “a”).

To go further: Add an option button called “2018” to the canvas, so that when you click on it, it displays a similar pie chart but for the 2018 percentage.

When clicking the button, the title of the pie chart should also change to “Total Album Consumption in **2018**”

Reminder: `canvas.delete("all")` is the command to delete what was previously drawn/written on the canvas

Good luck!

Sources:

https://www.musikindustrie.de/fileadmin/bvmi/upload/06_Publikationen/GMR/GMR2017_press.pdf

<https://www.musicbusinessworldwide.com/files/2019/01/BuzzAngle-Music-2018-US-Report.pdf>

<https://www.musicbusinessworldwide.com/files/2020/01/2019-U.S.-Music-Consumption-Year-End-Report-1.pdf>

https://www.musicbusinessworldwide.com/files/2021/01/MRC_Billboard_YEAR_END_2020_US-Final.pdf