

SAE 2.04 - Statistiques descriptives

Franceus-Cointrel Milwenn, Renaud Julien, Deschanel Alexandre, Rigaux H  lo  se

Partie 1 : Album

Variable type

type d  signe les diff  rentes cat  gories de sorties de chanson¹. Il en existe trois types : single, album, compilation. Il s'agit donc d'une variable **qualitative nominale**, car elle n'exprime pas de quantit   et ne peut pas   tre ordonn  e dans une suite logique. Le graphe le plus appropri   est un **diagramme en secteur**². Ce diagramme permet de repr  senter les proportions des diff  rents types.

Observation : L'album est la cat  gorie de sortie la plus populaire, repr  sentant    lui seul 56,9% de toutes les sorties enregistr  es dans le graphe. Vient ensuite les singles, repr  sentant pr  s de 42,6% de sorties. En tout dernier se trouve la compilation, avec un   cart de 42,1% aux singles : il repr  sente 0,5% de toutes les sorties.

Variable label

label d  signe les diff  rents noms de maisons de disques. Les modalit  s   tant importantes et n'  tant pas des nombres op  rables, il s'agit d'une variable **qualitative nominale**.

Observation :    la lecture de ce tableau³, nous ne pouvons pas d  terminer sur combien d'effectif le mode appara  t. Un nombre de label pouvant appara  tre plusieurs fois, il n'est pas possible de conna  tre directement l'effectif total.

Variable year

year d  signe les diff  rentes ann  es de sorties de titres de musique. Les modalit  s   tant connues et n'  tant pas des nombres op  rables mais ordonnables, il s'agit d'une variable **qualitative ordinale**.

Observation : Pour conna  tre la proportion des albums sortis apr  s x ann  es inclus, il faut ajouter la colonne "fi \searrow " (fr  quence d  croissante)⁴. Selon nos r  sultats, parmi la p  riode 2010-2023, il y a 82,45% des titres qui sont sortis apr  s 2021, dans lequel environ 60% sont sortis en 2023⁵.

Variable total_tracks

total_tracks désigne le nombre de titres appartenant à un album sorti après 2021 (inclus). Les modalités étant connues et étant des nombres opérables, il s'agit d'une variable **quantitative discrète**.

Observation : Le tableau⁶ nous indique les variables de positions (moyenne, médiane, quartiles) et de dispersion (écart type, écart interquartile). En effet, on peut observer qu'une large étendue de la variable de 82. Cependant les effectifs sont peu dispersés et sont plus groupés autour de la médiane avec la médiane située à 16. C'est ce que nous indiquent les variables IQ (écart interquartile) et écart type.

Partie 2 : Pistes

Variable tempo

tempo désigne la vitesse moyenne d'exécution de la musique⁷, mesurée en bpm. Les modalités étant importantes et opérables, il s'agit d'une variable **quantitative continue**. Dans cette partie, nous étudions la distribution des musiques en fonction du tempo.

Observation : Nous pouvons en déduire sur le graphique⁸ qu'une grande proportion des musiques se situent entre environ 120 et 140 bpm. La classe modale se situe entre 140 et 145 bpm.

Variable energy

energy désigne la nuance d'une musique, représentée par un réel compris entre 0 et 1 (0 signifiant une nuance minimale, et 1 une nuance maximale). Il s'agit d'une variable **quantitative continue**. Dans notre étude², nous nous contentons de seulement trois genres : "pop", "rock", "chanson" et "adult standards".

Observation : Le genre de musique "rock" se démarque par sa médiane élevée de 0,740, la plus élevée parmi tous les genres. De plus, il affiche le plus grand écart interquartile de 0,36, indiquant une grande dispersion des valeurs, bien qu'il n'y ait aucune valeur aberrante¹⁰. En revanche, le genre "pop" a la deuxième médiane la plus élevée à 0,659, mais il est plus dispersé avec une plage de valeurs jusqu'à 0,962 et quelques valeurs aberrantes, relativement proches de la limite inférieure. Les genres "chanson" et "adulte" ont des médianes beaucoup plus basses, respectivement à 0,213 et 0,201. Le genre "chanson" est moins dispersé, avec un écart interquartile de 0,089 et une plage de valeurs de 0,333, mais compte trois valeurs aberrantes, situées à environ 0,55 de la limite supérieure de la boîte à moustaches. Enfin, le genre "adulte" présente une dispersion plus importante, avec une plage de 0,7 et six valeurs aberrantes, la plus éloignée étant à 1,4 de la limite supérieure de la boîte à moustaches.

Annexe

Partie 1 : Albums

Tableau 1 : Fréquences des sorties de chansons selon le type

Modalités	effectifs	fréquences
single	235	42,57%
album	314	56,88%
compilation	3	0,54%
Total	552	100,00%

Diagramme 1 (diagramme en secteur) :

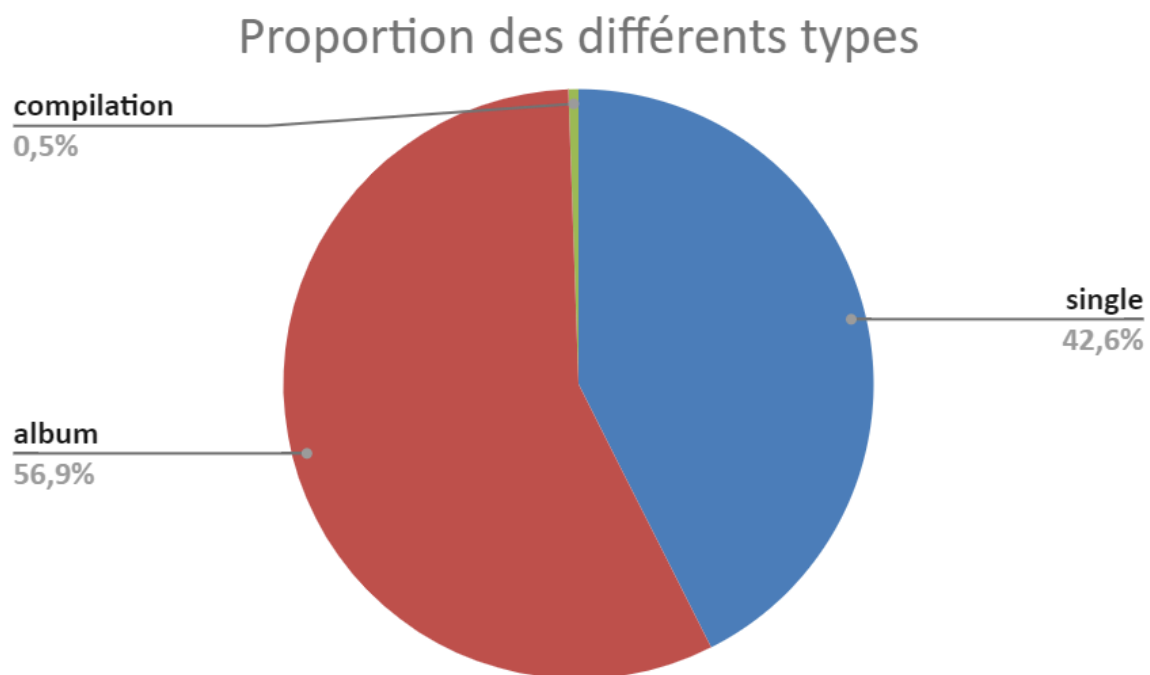


Tableau 2 : Labels ayant sorti le plus de musique selon l'année

année	mode	effectif du mode	nb de labels	effectif total
2021	Universal Music Division Capitol Music France	4	23	28
2022	Columbia	6	72	102
2023	Play Two	14	148	288

Tableau 3 : Fréquences de parution de titres entre 2010 et 2023

xi	ni	fi	fi ↘
2010	7	1,38%	100,00%
2011	7	1,38%	98,62%
2012	3	0,59%	97,24%
2013	3	0,59%	96,65%
2014	3	0,59%	96,06%
2015	5	0,99%	95,46%
2016	10	1,97%	94,48%
2017	8	1,58%	92,50%
2018	7	1,38%	90,93%
2019	17	3,35%	89,55%
2020	19	3,75%	86,19%
2021	28	5,52%	82,45%
2022	102	20,12%	76,92%
2023	288	56,80%	56,80%
Total	507	100,00%	-

Diagramme 2 (diagramme en bâtons) :

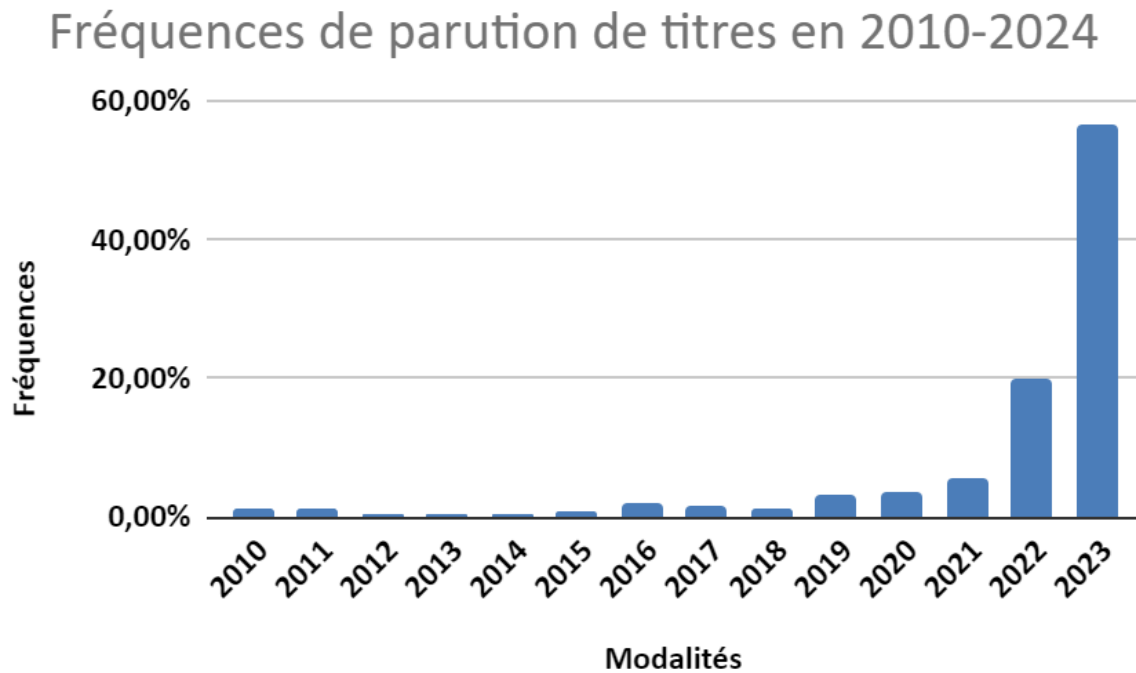


Tableau 4 : Variables de dispersion et de position de l'ensemble des titres

	min	max	Q1	médiane	Q3	moyenne	écart type	IQ
total_tracks	7	89	14	16	19	17,4	7,9	5

Partie 2 : Pistes

Tableau 5 : Valeurs fréquentielles et d'amplitudes selon le tempo

classe	tempo	effectif	fréquence	amplitude	hi × 1000
	40	-	-	-	-
[40 50[50	1	0,04%	10	0,04
[50 65[65	9	0,34%	15	0,22
[65 75[75	57	2,13%	10	2,13
[75 85[85	147	5,49%	10	5,49
[85 95[95	280	10,46%	10	10,46
[95 105[105	335	12,51%	10	12,51
[105 115[115	161	6,01%	10	6,01
[115 120[120	157	5,86%	5	11,73
[120 130[130	438	16,36%	10	16,36
[130 140[140	462	17,25%	10	17,25
[140 145[145	250	9,34%	5	18,67
[145 150[150	114	4,26%	5	8,51
[150 160[160	108	4,03%	10	4,03
[160 165[165	28	1,05%	5	2,09
[165 170[170	15	0,56%	5	1,12
[170 180[180	53	1,98%	10	1,98
[180 195[195	40	1,49%	15	1,00
[195 210[210	23	0,86%	15	0,57
	Total	2678			

Diagramme 3 (histogramme) :

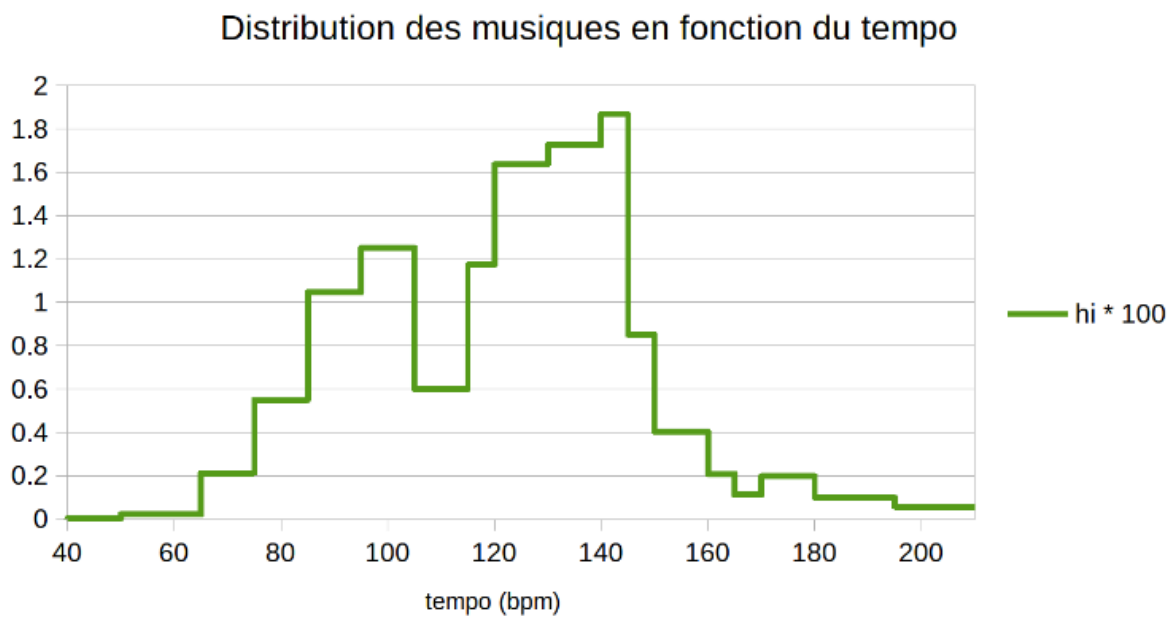


Tableau 6 : Détermination des valeurs aberrantes selon le genre

name_genre	min	max	moyenne	Q1	Q2	Q3	écart type	variance	IQ
adult standards	0,0103	0,793	0,256	0,134	0,201	0,343	0,165	0,027	0,209
chanson	0,114	0,447	0,221	0,163	0,213	0,252	0,079	0,006	0,089
pop	0,0234	0,985	0,629	0,505	0,659	0,784	0,198	0,039	0,279
rock	0,000	0,996	0,680	0,517	0,740	0,877	0,244	0,059	0,36

Boxplots :

