# MINI PROYECTO

Equipo 12 - Deep Learning

# Descripción breve

- El presente informe aborda la exploración de los datos para su entendimiento dentro del contexto organizacional, la preparación de los datos para poder utilizarlos como entrada para modelos predictivos, el análisis preliminar de selección de modelos relevantes para responder a la pregunta, el desarrollo y calibración de modelos y la visualización de resultados.

# **INTEGRANTES**

-Christian Beraún Chamorro - D'sharlie Sánchez Rozo - William Alexander Morales Valera

# 1. INTRODUCCIÓN

El presente miniproyecto se basa en las características musicales de una canción y si estas influyen o son determinantes respecto del año de publicación o lanzamiento. En resumen, se busca responder a la pregunta ¿Existe una relación entre las características musicales de una canción y el año en que fue publicada/lanzada? Una respuesta positiva a esta pregunta revelaría una profunda comprensión de la naturaleza de una composición musical y, lo que es más importante, esta comprensión se demostraría matemáticamente.

En este proyecto, se utilizan datos reales para predecir el año en que fue lanzada una canción a partir de sus características del timbre en la grabación. En total, son 90 atributos predictores: los 12 primeros corresponden al timbre promedio y los 78 siguientes a la covarianza. Originalmente, este tipo de datos fue recolectado en un proyecto llamado Million Song Dataset de la Universidad de Columbia (http://millionsongdataset.com/). Los datos para trabajar tienen los mismos predictores pero con canciones no consideradas en la base de datos original.

En general, se deben reportar los resultados del modelamiento predictivo siguiendo los pasos que se muestran a continuación:

- Exploración de los datos para su entendimiento dentro del contexto organizacional.
- Preparación de los datos para poder utilizarlos como entrada para modelos predictivos.
- Análisis preliminar de selección de modelos relevantes para responder a la pregunta.
- Desarrollo y calibración de modelos.
- Visualización de resultados.

#### **DATASETS PROPORCIONADOS**

- sampleSubmission.csv
- testReg.csv
- trainReg.csv

Disponibles en: https://www.kaggle.com/competitions/music-year-prediction/data

# 2. PARTE I: EXPLORACIÓN DE LOS DATOS Y CONTEXTO ORGANIZACIONAL

Para empezar, necesitamos mapear la descripción, cantidad de variables y calidad de los datos de ambos datasets, los cuales se realizó a continuación:

Train Data: 77779 rows × 92 columns

Tabla 1. Análisis Descriptivo de los datos

Туре	ID	Y	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
count	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779
mean	48244.2383	2002.30813	43.425185	-0.13671979	0	3.75531308	-2.33976807	-1.64309543	-6.81499697	-9.58726436	3.74957225	1.94851205	8.67485686	1.97787842	33.6375361	2419.77863	1960.66285
std	27846.2837	10.811038	6.12886878	4.37046639	0	17.6091827	14.4839754	7.89857439	22.980515	12.9119143	10.6385807	6.43009341	35.0960495	52.1597874	22.6040562	1734.20162	1266.74071
min	1	1926	4,83688	-69.68087	0	-165.22161	-121.47534	-72.50385	-152,40755	-70.69342	-119.48753	-41.63166	-249.10745	-329.57138	0.5874	59.01231	71.92873
25%	24079.5	1998	40.060315	-2.612435	0	-7.01624	-10.685075	-6.30069	-21.265305	-18.583065	-2.2894	-2.270755	-11.115595	-25.00927	18.04011	1298.63774	1098.55515
50%	48357	2006	44,32385	-0.063	0	2.0221	-2.05456	-1.58624	-6.30761	-11.23833	3.92844	1.89768	10.89415	9.3701	28.84442	1995,79785	1679.37829
75%	72340.5	2010	47,90008	2,46595	0	12,77645	6.4239	3.077055	7.617995	-2.32933	10.039325	6.163805	29.76277	37.11787	43.343385	3030.43891	2484.72071
max	96435	2014	60.03401	23.81526	0	274.65858	160.81522	68,44796	262.06887	112.97141	80.78712	39.97683	262.28304	304.26639	396.91851	43921.6472	28231.0998
Type	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	V31	V32
count	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779
mean	1518.44301	908.683312	882.374897	602.121692	515.486995	391.864407	324.662498	289.652977	291.749852	42.6531646	45.1744792	-51.7059202	-27.2903774	16.3361058	45.3634331	5.10268845	24.1629371
std	1146.29573	476.064042	590.64115	319.122294	314,465741	212.942125	169,499811	193,684396	155.259611	122.34173	705.925653	565,245898	219.284056	164,130994	137.895841	101.773912	72.0424704
min	68.11035	60.89138	23.90388	43.6845	9.15615	22.12135	18.35053	10.92598	10.83169	-2214,38045	-7645.54417	-12017.0889	-3163.0211	-2243.27056	-1121.16473	-3805.66617	-1328.84733
25%	810.625395	602.9576	484,929745	389,49676	315.046905	255.597385	211.5009	167.39054	197.37992	-11.42276	-282.901785	-255.459435	-117.066795	-58.5	-24,470705	-38.45101	-6.26207
50%	1236.65707	812.65969	739.569	539.12575	441.19302	348.22422	290.62639	241.30872	263.54078	27.66455	-282.901785	-41.89223	-22.16499	8.90315	30.62104	4.20457	23.2163
75%	1881.72451	1099.89625	1118.72944	737.833145	623.834955	474,79821	395,4166	356.105855	349.54909	88.85793	286.10921	165.780825	66,428495	85.818855	97.90438	48.68869	55.569455
max	31792.7148	10787.6512	16831.949	6360,24268	9569,77809	9616.6156	3365.4697	6737.1215	5036.61707	1432.56131	16337.8676	14505.3422	3410.61556	2963,43419	3553.18488	1584,70058	1954.35548
	V33	V34				V38	V39	V40	V41	V42	V43	V44	V45	V46	V47	V48	1954.35548 V49
Туре			V35	V36	V37												
count	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779
mean	9.45009103	-3.79119002	0.56498076	72.4339041	-52.1220811	115.195914	-191.301896	19.8639784	0.17065014	17.4710593	-52.3960796	3.10810553	-2.10795885	5.8277606	75.6339744	145.50257	-88.6308129
std	75.0767429	56.2198568	43.4797702	108.840815	420.088149	450.650852	275.823815	205.132499	122.914826	120.701322	76.5278512	39.5468929	42.1739681	55.8935541	468.856902	265.295702	220.503087
min	-1679.11832	-1590.63713	-765.99379	-1711.484	-7298.32896	-5924.44473	-9752.74109	-3919.39572	-4434.06853	-4175.41268	-4071.55347	-1072.95552	-1021.28921	-1329.95974	-13731.1786	-3992.68866	-6642.39958
25%	-24.06061	-28.466095	-16.70903	14.397875	-223.686765	-96.732745	-289.399435	-71.23451	-51.736725	-27.09117	-78.900325	-14.606395	-21.06588	-17.78936	-113.413005	13.063705	-158.43634
50%	8.12595	-9.53697	3.47933	55.35636	-73.57698	80.60075	-159.15431	19.10876	-2.66917	26.01287	-45.36519	3.09825	-1.28065	3.62073	71.60452	107.81089	-63.27651
75%	40.85454	12.379805	20.87695	116.15152	80.962925	304.055655	-60.162905	111.54419	48.462315	75.45868	-18.1711	21.5464	18.549035	27.342455	278.486235	243.07607	11.28586
max	2868.35173	2330.33366	691.12904	1636.96095	14148.9979	6011.87812	6065.05481	3542.43713	3537.50359	1301.4294	934.82714	1830.54468	568.99828	943.93652	6664.22793	4550.38543	2792.32447
Туре	V50	V51	V52	V53	V54	V55	V56	V57	V58	V59	V60	V61	V62	V63	V64	V65	V66
count	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779	77779
count	77779 26.2429905	77779 6.30456856	77779 28.7691004	77779 12.1550628	77779 1.05757721	77779 -10.7076745	77779 62.2000638	77779 104.668247	77779 2.92360621	77779 36.7551425	77779 -25.9475077	77779 3.74109316	77779 0.33110399	77779 -0.6415598	77779 -138.292931	77779 -2.66932844	77779 0.82408989
count mean std	77779 26.2429905 122.029293	77779 6.30456856 93.4899766	77779 28.7691004 75.9489943	77779 12.1550628 72.5352572	77779 1.05757721 83.9874145	77779 -10.7076745 58.1211171	77779 62.2000638 273.554783	77779 104.668247 310.125825	77779 2.92360621 279.538596	77779 36.7551425 165.606073	77779 -25.9475077 147.081674	77779 3.74109316 61.3342389	77779 0.33110399 49.6714256	77779 -0.6415598 38.4150885	77779 -138.292931 305.029857	77779 -2.66932844 225.204798	77779 0.82408989 130.350883
count mean std min	77779 26.2429905	77779 6.30456856	77779 28.7691004	77779 12.1550628	77779 1.05757721	77779 -10.7076745	77779 62.2000638	77779 104.668247	77779 2.92360621	77779 36.7551425	77779 -25.9475077	77779 3.74109316	77779 0.33110399	77779 -0.6415598	77779 -138.292931	77779 -2.66932844	77779 0.82408989
count mean std	77779 26.2429905 122.029293	77779 6.30456856 93.4899766	77779 28.7691004 75.9489943	77779 12.1550628 72.5352572	77779 1.05757721 83.9874145	77779 -10.7076745 58.1211171	77779 62.2000638 273.554783	77779 104.668247 310.125825	77779 2.92360621 279.538596	77779 36.7551425 165.606073	77779 -25.9475077 147.081674	77779 3.74109316 61.3342389	77779 0.33110399 49.6714256	77779 -0.6415598 38.4150885	77779 -138.292931 305.029857	77779 -2.66932844 225.204798	77779 0.82408989 130.350883
count mean std min	77779 26.2429905 122.029293 -2075.49185	77779 6.30456856 93.4899766 -2270.81107	77779 28.7691004 75.9489943 -1089.50545	77779 12.1550628 72.5352572 -3188.17738	77779 1.05757721 83.9874145 -2199.78221	77779 -10.7076745 58.1211171 -847.6934	77779 62.2000638 273.554783 -4536.69953	77779 104.668247 310.125825 -3494.01692	77779 2.92360621 279.538596 -4730.5991	77779 36.7551425 165.606073 -2327.42228	77779 -25.9475077 147.081674 -1601.64462	77779 3.74109316 61.3342389 -1900.1048	77779 0.33110399 49.6714256 -1129.51344	77779 -0.6415598 38.4150885 -583.20098	77779 -138.292931 305.029857 -5365.13856	77779 -2.66932844 225.204798 -7375.97744	77779 0.82408989 130.350883 -3896.27522
count mean std min 25%	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035	77779 2.92360621 279.538596 -4730.5991 -115.189255	77779 36.7551425 165.606073 -2327.42228 -45.623365	77779 -25.9475077 147.081674 -1601.64462 -93.068185	77779 3.74109316 61.3342389 -1900.1048 -24.852835	77779 0.33110399 49.6714256 -1129.51344 -21.619025	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436	77779 -138.292931 305.029857 -5365.13856 -261.40949	77779 -2.66932844 225.204798 -7375.97744 -86.45431	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545
count mean std min 25% 50%	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545
count mean std min 25% 50% 75%	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545
count mean std min 25% 50% 75% max	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473
count mean std min 25% 50% 75% max Type	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489 <b>V74</b>	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535 V82	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83
count mean std min 25% 50% 75% max Type count	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489 V74 77779	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 77779	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535 V82 77779	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779
count mean std min 25% 50% 75% max Type count mean	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779 28.1411986	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 2460.43343 V69 77779 32.0184738	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489 V74 77779 -11.8669472	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779 -5.47216709	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 77779 -23.2969748	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09635 V82 77779 -71.7330078	77779 0.82408989 130.350883 -3696.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.6792868
count mean std min 25% 50% 75% max Type count mean std	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.1226656	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779 28.1411986 118.435311	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2990.52018 V71 77779 -4.96434133 249.503583	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.706924	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039 172.780891	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779 -5.47216709 26.8709447	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 77779 -23.2969748 269.101216	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535 V82 77779 -71.7330078	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.6792868 123.523393
count mean std min 25% 50% 75% max Type count mean std min	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.1226656 -1048.04155	77779 6.30456856 93.4899766 92.70.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779 28.1411986 118.435311 -2564.78812	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -7057.71245	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.708924 -6953.35736	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039 172.780891 -6395.81177	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779 -5.47216709 26.8703447 -495.54551	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 77779 -23.2969748 269.101216	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09635 V82 77779 -71.7330078 176.893351 -4402.37644	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.6792868 123.52393 -1733.72211
count mean std min 25% 50% 75% max Type count mean std min 25%	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.126656 -1048.04155 -45.54368	77779 6.30456856 93.4899766 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V88 77779 28.1411986 118.435311 -2564.78812 -19.128765	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769 -16.079025	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 90.52018 V71 77779 -4.96434133 249.503583 -7057.71245 -92.167525	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.706924 -6953.35736 -89.70404	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 161.617165 773 77779 -30.4072039 172.780891 -6395.81177 -87.59399	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 77779 -21.2431762 64.5118109 -1312.45848 -50.531905	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779 -5.47216709 26.8709447 -495.54551 -14.907965	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 V77 77779 -23.2969748 -269.101216 -9812.57578 -150.129885	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255	77779 -0.6415598 38.415085 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156	77779 -138.292931 305.029857 -5395.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265	77779 -2.68932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095	77779 0.82408899 130.350883 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 W83 77779 41.6792668 123.523993 127.523193 -1733.72211 -21.467975
count mean std min 25% 50% 75% max Type count mean std min 25% 50%	77779 26.2429905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.1226656 -1048.04155 -45.54368 2.83274	77779 6.30456856 93.4899766 93.4899766 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779 28.1411986 118.435311 -2564.78812 -19.128765 18.20059	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078 22.21309	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769 -16.079025 -0.72087	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -7057.71245 192.167525 17.17829	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.706924 -6953.35736 -89.70404 12.88197	77779 62 2000638 273.554783 -70.45673 27.89036 161.617165 5734.17435 77779 -30.4072039 172.780891 -6395.81177 -87.59399 -21.7561	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5646 223.208035 5489.79489 774 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -11.91485	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 77779 -5.47216709 26.8709447 -495.54551 -14.907965 -2.22658	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.848855 5690.29165 V77 77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251	77779 3.74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 778 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022	77779 -0.6415598 38.415085 -583.20098 -19.022475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.54156 25.72258	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998	77779 -2.66932844 225.204798 -7375.97744 -86.45431 6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.6792868 123.523393 -1733.72211 -21.467975 28.21756
count mean std min 25% 50% 75% max Type count mean std min 25% 50% 50% max	77779 26.242905 122.029293 -2075.49185 -23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.1226656 -1048.04155 -45.54368 2.83274 51.79019	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V68 77779 28.1411986 118.435311 -2564.78812 -19.128765 18.20059 61.58248	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078 22.21309 71.772825	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.1576 -16.079025 -0.72087 17.72087 411.36502	77779 1.05757771 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -7057.71245 -92.167525 17.17829 112.53175 3958.07011	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 506.60846 V72 77779 4.55240036 229.706924 -6953.35736 -89.70404 12.88197	77779 62.2000638 273.554783 -4536.69953 -70.45673 27.89036 5734.17435 V73 77779 -30.4072039 172.780891 -6395.81177 -87.59399 -21.7561 39.601015	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% 50% 75% max Type count mean std min 25% 50% 75% 75%	77779 26.2429905 122.029293 -2075.49185 -23.57538 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.1226656 -1048.04155 -45.54368 2.83274 51.79019 1290.30838	77779 6.30456856 93.4899766 -2270.81107 -38.51442 51.3511 50.610915 1426.84804 V68 77779 28.1411986 118.435311 -2564.78812 -19.128765 18.20059 61.58248 4779.80027	77779 28.7691004 75.9489943 -1089.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078 22.21309 77.772825 2696.88075	77779 12.1550628 72.5352572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769 -16.079025 -0.72087 14.743075	77779 1.05757721 83.9874145 -2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -7057.71245 -92.167525 17.17829 112.53175	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.706924 -6953.35736 -89.70404 12.88197 103.427945 2907.04705	77779 62.2000638 273.554783 4556.69953 -70.45673 27.89036 161.61765 V73 77779 -30.4072039 172.780891 -6395.81177 -87.59399 -21.7561 39.601015 1395.24174	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% 50% 75% max Type count mean std min 25% 50% 75% Type Type Type Type Type Type Type	77779 26.242990 26.242990 275.49185 22.075.49185 23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.122656 -1048.04155 -45.54368 2.83274 51.79019 1290.30838 V84	77779 6.30456856 9.34899766 -2270.81107 -38.51442 5.13511 5.06.10915 1426.84804 V88 77779 28.1411986 118.435311 -2564.78812 -191.128765 18.20059 61.58248 4779.80027 V85	77779 28.7691004 775-9489943 -1089.50545 -8.26689 26.22024 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078 22.21309 71.772825 2686.88075	77779 12.1550628 72.535572 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 -77779 -0.93296169 36.9420541 -549.15769 -16.079025 -0.72087 14.743075 411.36502 V87	77779 1.05757721 83.9874145 2-2199.78221 -34.089275 0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -705.7.1245 -92.167525 17.17829 112.53175 3958.07011 V88	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.60846 V72 77779 4.55240036 229.706924 -6953.35736 -89.70404 12.88197 103.427945 2997.04705 V89	77779 62.2000638 777.554783 -4558.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039 172.780891 -6395.81177 -87.593399 -21.7561 39.601015 1395.24174 V90	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% 50% 75% max Type count mean std min 25% 50% Type count mean std min 25% 50% Type Type count mean std min 25% 50% Type Type Count max	77779 26.242905 26.242905 2075.49185 22.57536 22.57536 23.13812 86.83728 1739.28042 V67 77779 3.53222347 99.1226566 -1048.04155 -45.54368 2.83274 51.79019 1290.30838 V84 77779	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.84804 V88 77779 28.1411986 118.435311 -2564.76812 -19.128765 18.20059 61.58248 4779.80027 V85 77779	77779 28.7691004 75.9489943 -1089.50545 -8.26889 -8.26827 62.2027 62.2027 62.20244 2460.43343 V69 77779 32.018478,06641 -19.16078 22.21309 71.772825 2898.86075 V86 777779	77779 12.1550628 72.5355272 -3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769 14.743075 411.36502 V87 777779	77779 1.05757721 83.9874145 -2199.78221 -34.089275 -0.23035 36.61764 2900.52018 V71 77779 -4.96434133 -7057.71245 -92.167525 112.53175 3958.07011 V88 777779	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.0046 V72 77779 4.55240934 -6953.35736 -89.70404 12.88197 103.427945 2907.04705 V89 77779	77779 62.2000638 2273.554783 -4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.40720391 172.780891 -6395.81177 -87.59399 -21.7561 39.601015 1395.24174 V90 777779	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% 50% 50% 50% max Type count mean std min 25% 600 to 50% 50% 50% 50% 50% max System std min 500 to 50% max System std mean std std 50% max System std 50% max Syste	77779 26.2429905 26.2429905 275.49185 275.536 22.57536 22.13812 86.83728 1739.29042 V67 77779 3.53222347 51.79019 1290.30838 V84 77779 77779 377639 59.1226565 45.45368 2.83274 51.79019 290.30838 V84 77779 95.33532567 96.3353567	77779 6.30456856 93.4899766 -2270.81107 -38.51442 5.13511 50.610915 1426.4804 V88 77779 28.141198 118.43531 -2564.78812 -19.128765 61.58248 4779.80027 V85 77779 0.328212883	77779 28.7691004 77.59489943 -1089.50545 -8.26689 -8.26689 -8.2027 62.2027 62.2027 62.2027 63.2016 -10.555083 -1578.06641 -19.16078 -22.21309 71.772825 -2696.86075 -78667 -78779 -78667	77779 12.1550628 72.5352572 3188.17738 -18.690025 8.00347 38.64616 2394.66234 V70 77779 -0.93296169 36.9420541 -549.15769 -16.079025 -0.72087 14.743075 411.36502 V87 77779 -25.6288567	77779 1.05757721 83.9874145 -2199.78221 -34.089275 -0.23035 36.61764 2900.52018 V71 77779 -4.96434133 249.503583 -7057.71245 -92.167525 17.17829 112.53175 3958.07011 V88 777779 4.46348419 13.5264139	77779 -10.7076745 58.1211171 -847.6934 -29.20602 -2.10108 18.41268 506.66846 V72 77779 4.5524003 225.706924 -6953.35736 -89.70404 12.88197 103.427945 2907.04705 V88 77779 18.6648955 18.666895183	77779 62.200638 273.554783 4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039 172.780936 6395.81177 -87.59399 -21.7561 39.601015 1395.24174 V90 77779 1.240194 22.3796535	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% max Type count mean std min 25% max Type Type Type Type Type Type Type Type	77779 26.242905 122.0729293 2075.49185 2-205.49185 32.13812 86.83728 1739.29042 799.122656 1048.04155 45.54368 2.83274 51.79019 1290.30838 V84 77779 37.8163527 37.8163527 1848.70225	77779 6.3045856 83.4899766 93.4899766 93.4899766 9270.81107 93.851442 5.13511 50.610915 1426.34804 V68 77779 828.1411986 118.435311 -19.128765 18.20059 61.58248 4779.80027 V85 77779 0.32821288	77779 28.769104 75.9489943 1088.50545 -8.26689 26.22027 62.20244 2460.43343 V69 77779 32.0184738 105.55083 -1578.06641 -19.16078 22.21399 77779 17.5750966 115.243644 -3188.92457	77779 12.1550628 72.5352572 -3188.17738 -38.64616 2394,66234 -70 -77779 -0.93296169 -16.079025 -0.72087 -14.749075 -411.36502 -0.72087 -25.6288567 -25.6288567 -25.6288567 -231.93024	77779 1.05757721 83.9874145 2199.78221 -34.089275 -0.23035 36.61764 2990.52018 V71 77779 -4.9634131 249.50383 249.50383 112.5317 3958.07011 V88 77779 4.46348419 1.35264139	77779 -10.7076745 -58.1211171 -847.6934 -29.20602 -2.10108 -18.41268 -506.60846 -777 -77779 -4.5524003 -229.706924 -6953.35736 -89.70404 -12.88197 -103.427945 -2907.04705 -788 -748.68855 -7486.87815 -7486.87815 -7486.87815	77779 62.200638 273.554783 4536.69953 -70.45673 27.89036 161.617165 5734.17435 77779 37.7779 37.7779 37.7779 37.7760891 12.780891 21.7561 39.601015 1395.24174 v90 77779 1.240194 22.3796535 -281.1506	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% max Type count mean std min 25% max Type count mean std min 25% count mean std min 25% max Type count mean std min std min std min std min 25%	77779 26.242905 122.029293 2075.49185 22.357536 32.13812 66.83728 1739.29042 V67 77779 9.12686 1048.04155 445.4388 2.83274 51.79019 1290.30838 V84 777779 9.37.8163827 95.3383867 1848.70226	77779 6.3045895766 93.4699766 93.4699766 9270.81107 93.851442 5.13511 50.610915 90.610915 90.610915 91.426.84804 988 77779 28.1411986 118.435311 12.9564.78812 19.128765 18.20059 61.58248 4779.80027 985 777779 0.32812188 16.2728833 1-238.38673 -238.38673	77779 28.769104 28.769104 28.769104 75.5488943 -1088.50545 -8.26689 26.22027 62.20244 -2460.45334 V69 77779 32.0184738 105.55083 -11578.06641 -19.16078 22.21309 71.772825 2696.86075 V86 77779 17.7570966 115.243644 -3168.92457 -31.56939	77779 12.155028 72.5352572 3188.17738 -18.690025 8.00347 38.64616 2394,66234 V70 77779 0.33296169 16.079025 -0.72087 14.743075 411.86502 V87 77779 777779 777779 725.6288567 173.310304 -43319.99232	77779 1.05757721 83.9874145 2199.78221 2199.78221 230.52018 V1 77779 4.96434133 249.503583 V21 249.503583 1717789 112.53175 3958.07011 V88 777779 4.46348419 113.5264139 -236.03926 -236.03926	77779 -10.7076745 -58.1211171 -847.6934 -29.20602 -2.10108 -18.41258 -506.60846 -72 -77779 -4.55240036 -229.706924 -6953.35736 -89.70404 -12.88197 -103.427945 -89.704705 -77779 -18.6648958 -77779 -7779 -77	77779 62.200638 273.554783 4538.69953 -70.45673 27.89038 161.617165 5734.17435 V73 77779 -30.4072039 172.780891 6395.81177 -87.59399 -21.7561 39.601015 1395.24174 V80 777779 1.240194 22.3796535 -261.1506 -8.90012	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% max Type count mean std min 25% some 50% max Type count mean std min 25% max Type count mean std min 25% max Type count mean std min 50% max Type 50% max Type 50% max Type 50% max Std min 55% max Std	77779 26.242905 122.029293 2075.49185 2-23.57536 32.13812 86.83728 1739.29042 V67 77779 3.53222347 99.122656 1.048.04155 4.554368 2.83274 5.7799 37.8165527 99.95.3353567 1.848.70226 4.6.04295 3.441737	77779 6.3045859 83.4899766 93.4899766 93.4899766 9270.81107 93.851442 5.13511 50.610915 1426.84804 V88 77779 28.1411986 118.435311 2-2584.78812 -19.128765 18.20059 61.58248 4779.80027 V85 77779 0.32821288 16.2726893 -238.38673 -6.70237 0.78898	77779 28.769104 75.9489943 -1098.50545 -8.26689 26.22027 62.20244 -2460.43343 -989 32.0184738 105.55083 -11578.06641 -19.16078 -22.21309 -77779 17.172825 2896.88075 -78779 115.5750966 115.24754 -31.86939 -31.56939	77779 12.1550628 72.5352572 -3188.17738 -38.64616 2394.6629 -0.93296169 36.9420541 -549.15769 -0.70207 -1.549.15769 -0.70207 -1.549.15769 -1.57779 -0.93296169 -0.72087	77779 1.05757721 83.9874162 2199.78221 2199.78221 0.23035 36.61764 22900.52018 V71 77779 -4.96434133 249.50358 -92.167525 17.17829 112.53175 3958.07011 V88 77779 4.46348419 13.5264139 236.03926 -2.56936	77779 -10,70°F6745 -58.1211171 -847.6934 -29.20602 -2.10108 -18.41268 -506.60846 -722 -77779 -4.55240036 -89.70404 -12.88197 -103.427945 -2907.04705 -89.7	77779 62.200638 273.554783 4536.69953 -70.45673 27.89036 161.617165 5734.17435 V73 77779 -30.4072039 172.780891 6385.81177 -87.59399 -21.7561 339.601015 339.624174 V90 77779 1.240194 22.3796535 -281.1506 -8.90012	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3607.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475
count mean std min 25% count mean 25% count mean std min 25% count m	77779 26.242905 122.029293 2075.49185 22.357536 32.13812 66.83728 1739.29042 V67 77779 9.12686 1048.04155 445.4388 2.83274 51.79019 1290.30838 V84 777779 9.37.8163827 95.3383867 1848.70226	77779 6.3045895766 93.4699766 93.4699766 9270.81107 93.851442 5.13511 50.610915 90.610915 90.610915 91.426.84804 988 77779 28.1411986 118.435311 12.9564.78812 19.128765 18.20059 61.58248 4779.80027 985 777779 0.32812188 16.2728833 1-238.38673 -238.38673	77779 28.769104 28.769104 28.769104 75.5488943 -1088.50545 -8.26689 26.22027 62.20244 -2460.45334 V69 77779 32.0184738 105.55083 -11578.06641 -19.16078 22.21309 71.772825 2696.86075 V86 77779 17.7570966 115.243644 -3168.92457 -31.56939	77779 12.155028 72.5352572 3188.17738 -18.690025 8.00347 38.64616 2394,66234 V70 77779 0.33296169 16.079025 -0.72087 14.743075 411.86502 V87 77779 777779 777779 725.6288567 173.310304 -43319.99232	77779 1.05757721 83.9874145 2199.78221 2199.78221 230.52018 V1 77779 4.96434133 249.503583 V21 249.503583 1717789 112.53175 3958.07011 V88 777779 4.46348419 113.5264139 -236.03926 -236.03926	77779 -10.7076745 -58.1211171 -847.6934 -29.20602 -2.10108 -18.41258 -506.60846 -72 -77779 -4.55240036 -229.706924 -6953.35736 -89.70404 -12.88197 -103.427945 -89.704705 -77779 -18.6648958 -77779 -7779 -77	77779 62.200638 273.554783 4538.69953 -70.45673 27.89038 161.617165 5734.17435 V73 77779 -30.4072039 172.780891 6395.81177 -87.59399 -21.7561 39.601015 1395.24174 V80 777779 1.240194 22.3796535 -261.1506 -8.90012	77779 104.668247 310.125825 -3494.01692 -38.60207 84.5546 223.208035 5489.79489 V74 77779 -11.8669472 64.0142015 -889.47244 -42.55214 -41.91485 17.31348	77779 2.92360621 279.538596 -4730.5991 -115.189255 -14.28362 88.179795 11971.1799 V75 77779 -21.2431762 64.5118109 -1312.45848 -50.531905 -16.55181 11.458575	77779 36.7551425 165.606073 -2327.42228 -45.623365 26.60129 108.99046 3807.75017 V76 -77779 -5.472216709 26.8709447 -495.54551 -14.907965 -2.22658 7.68559	77779 -25.9475077 147.081674 -1601.64462 -93.068185 -30.27674 31.846855 5690.29165 V77 -77779 -23.2969748 269.101216 -9812.57578 -150.129885 -57.01251 61.996475	77779 3,74109316 61.3342389 -1900.1048 -24.852835 0.09307 27.21505 1811.22866 V78 77779 31.9069743 144.142312 -1628.81769 -30.22832 38.54099 104.181415	77779 0.33110399 49.6714256 -1129.51344 -21.619025 1.64608 23.779375 906.46898 V79 77779 -105.995059 209.971584 -8390.03545 -162.72255 -71.00022 -7.34313	77779 -0.6415598 38.4150885 -583.20098 -19.032475 -3.45574 14.7436 647.15692 V80 77779 25.663725 124.016038 -2760.59631 -29.04156 25.72258 84.283625	77779 -138.292931 305.029857 -5365.13856 -261.40949 -114.62037 1.72641 5100.98473 V81 77779 15.7172149 32.2119947 -424.51757 -1.968265 9.07998 26.2117	77779 -2.66932844 -225.204798 -7375.97744 -86.45431 -6.52295 104.23111 2313.09535 V82 77779 -71.7330078 176.893351 -4402.37644 -136.511095 -51.32257 13.99987	77779 0.82408989 130.350883 -3896.27522 -51.019845 5.6545 54.864545 3127.04473 V83 77779 41.679288 123.523393 -1733.72211 -21.467975 89.190475

Se realiza un primer análisis descriptivos de las variables revisando la medida, mínimos, máximos y desviación estándar, a simple vista se puede evidenciar que la columna V3 esta presente en todos los registros pero su valor es cero, por otra lado en las primeras 12 variables se pueden evidenciar valores mas acordes según su desviación estándar, siendo que son 91 variables, se revisó en el notebook de una forma más amplia en histogramas, para efectos de observar el comportamiento y distribución de las variables.

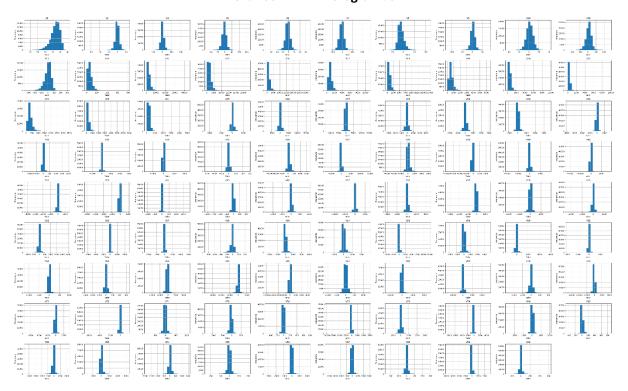


Gráfico n.º 1 - Histogramas

Luego se realizó un análisis de la distribución de la variable a predecir:

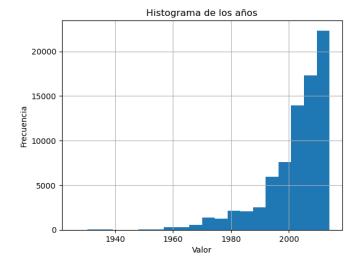


Gráfico n.º 2 - Distribución de la variable a predecir

Como podemos observar, existe una distribución mayor para los años más cercanos a la fecha; ello podría deberse a que los datos se encuentran disponibles en diversas plataformas streaming de música siendo que estas surgieron en los últimos 20 años, tal como lo precisa en la página web de los datasets:

- SecondHandSongs dataset -> cover songs
- musiXmatch dataset -> lyrics
- Last.fm dataset -> song-level tags and similarity
- Taste Profile subset -> user data
- thisismyjam-to-MSD mapping -> more user data

- tagtraum genre annotations -> genre labels
- Top MAGD dataset -> more genre labels

Ahora bien, tengamos en cuenta este dato importante "los 12 primeros (V1 A V12) corresponden al timbre promedio y los 78 siguientes a la covarianza (V13 a V90)". Analizaremos la correlación que existe entre las variables predictoras con la variable de interés:

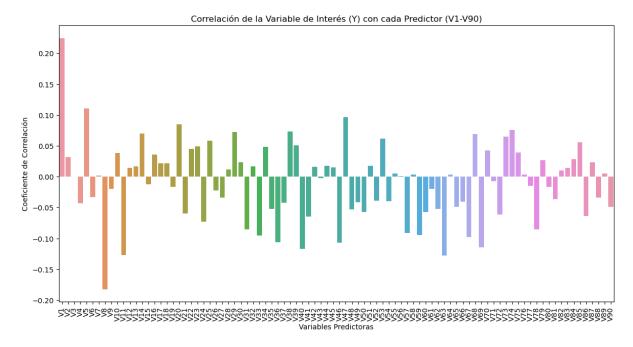


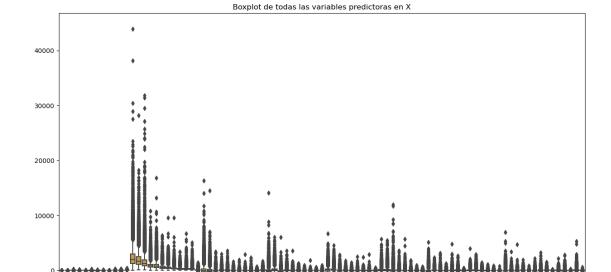
Gráfico n.º 3 - Matriz de Correlación

Realizando el análisis de correlación de las variables predictoras vs la variable objetivo, podemos evidenciar que hay una correlación de -20 a 20, lo que indica que hay muy poca correlación entre la variable objetivo y las variables predictoras.

### 3. PARTE II: PREPARACIÓN DE LOS DATOS PARA MODELOS PREDICTIVOS

Como pudimos observar, las variables predictoras van de V1 a V90, pero se debe eliminar V3 y ID, siendo que no agregan valor al modelo, por ser valores vacíos o el índice de los registros. Ahora procedemos a analizar la dispersión de los datos

Gráfico n.º 4 - Determinación de outliers



Ahora bien, si observamos las variables V14 en adelante, podemos ver que hay mucha variabilidad. Mientras que de V1 a V12, no hay mucha.

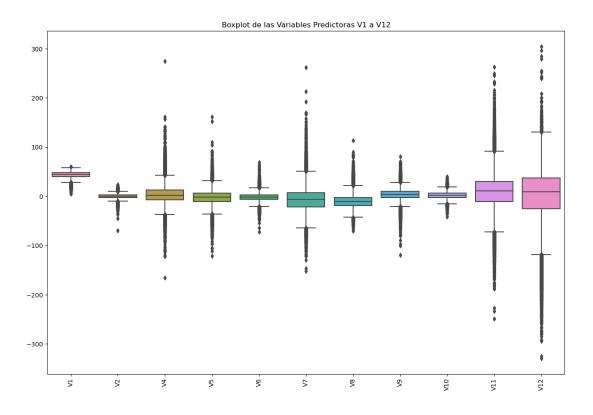


Gráfico n.º 5 – Boxplot de las variables V1 a V12

-10000

Considerando todo lo anterior, se vio por conveniente crear variables que capturen los datos de V1 a V12 y de V13 a V90, usando métricas como suma, promedio, mediana, producto, entre otras. Conforme se detalla a continuación:

```
# Función para crear características
train_data = train_data.drop(columns=['ID','V3'])
test_data = test_data.drop(columns=['ID'])
def create_features(data):
      # Adjusted range to exclude 'V3'
timbre_cols = [f"V{i}" for i in range(1, 13) if i != 3]
         Check if all required timbre columns except V3 exist
      if all(col in data.columns for col in timbre_cols):
            data['mean_timbre'] = data[timbre_cols].mean(axis=1)
data['median_timbre'] = data[timbre_cols].median(axis=1)
            data['std_timbre'] = data[timbre_cols].std(axis=1)
data['sum_timbre'] = data[timbre_cols].sum(axis=1)
            data['max_timbre'] = data[timbre_cols].max(axis=1)
data['min_timbre'] = data[timbre_cols].min(axis=1)
             data['prod_timbre'] = data[timbre_cols].prod(axis=1)
      print("Algunas columnas necesarias (V1, V2, V4 a V12) no existen en el DataFrame")
# Check if all covariance columns (V13 to V90) exist
if all(f"V{i}" in data.columns for i in range(13, 91)):
    covariance_cols = [f"V{i}" for i in range(13, 91)]
            data['mean_covariance'] = data[covariance_cols].mean(axis=1)
data['median_covariance'] = data[covariance_cols].median(axis=1)
data['std_covariance'] = data[covariance_cols].std(axis=1)
data['sum_covariance'] = data[covariance_cols].sum(axis=1)
data['max_covariance'] = data[covariance_cols].max(axis=1)
data['min_covariance'] = data[covariance_cols].min(axis=1)
             # Using logarithmic addition to prevent overflow in product
            data['prod_covariance'] = np.exp(data[covariance_cols].apply(np.log).sum(axis=1))
      else:
           print("Algunas columnas de V13 a V90 no existen en el DataFrame")
      return data
# Aplicar la creación de características
train_data = create_features(train_data)
test_data = create_features(test_data)
```

Como se puede evidenciar en el fragmento de código anterior, se valida la existencia de columnas en el set de datos con el objetivo de manejar posibles errores al momento de realizar la ejecución de proceso.

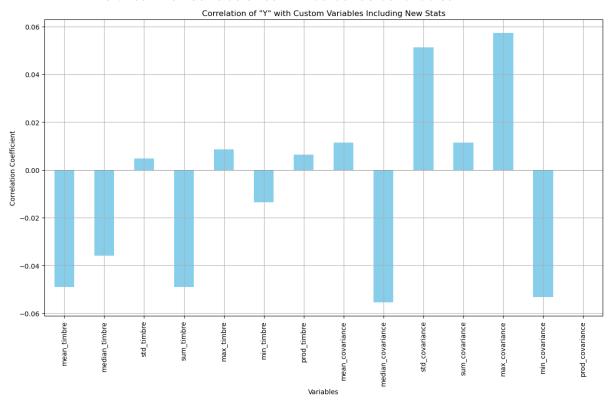


Gráfico n.º 6 - Correlación con "Y" de las nuevas variables

En la gráfico anterior podemos observar que fue buena decisión agregar estas características, especialmente aquellas que muestran una correlación más alta de 0.02. Estas no solo capturan los datos de los dos grupos mencionados, sino que también están más estrechamente relacionadas con la variable objetivo

# 4. PARTE III: ANÁLISIS PRELIMINAR DE SELECCIÓN DE MODELOS RELEVANTES

#### \*Modelo 1: Una sola neurona \*

```
In [21]: # Modelo 1: Simplificado con una sola neurona
           model1 = Sequential([
    Dense(1, input_dim=X_train_scaled.shape[1], activation='linear') # Una sola salida para regresión
           # Compilación del modelo con el optimizador Adam

optimizer = Adam(learning_rate=0.2)
model1.compile(optimizer=optimizer, loss='mean_squared_error')
           # Callback de EarlyStopping para evitar el sobreajuste
early_stopping = EarlyStopping(monitor='val_loss', patience=5, verbose=1, restore_best_weights=True)
           # Entrenamiento del modelo con EarlyStopping model1.fit(X_train_scaled, y_train, epochs=10, batch_size=32, validation_data=(X_val_scaled, y_val), callbacks=[early_stopping])
            # Evaluación del modelo
           # Lucation |
y pred = model1.predict(X_val_scaleu)
if np.isnan(y_pred).any():
    raise ValueError("Las predicciones contienen NaN")
            # Cálculo del RMSE para evaluar el rendimiento del modelo
           rmse = sqrt(mean_squared_error(y_val, y_pred))
print(f'RMSE: {rmse}')
           Epoch 1/10
1945/1945 -
                                                - 2s 640us/step - loss: 3646274.7500 - val_loss: 2653277.7500
            Epoch 2/10
            1945/1945 -
                                          ----- 1s 602us/step - loss: 2378656.0000 - val_loss: 1631743.1250
           Epoch 3/10
1945/1945 -
                                                - 1s 597us/step - loss: 1429066.5000 - val loss: 888608.5000
            Epoch 4/10
1945/1945 -
                                   ______ is 621us/step - loss: 748726.3125 - val_loss: 392420.5625
            Epoch 5/10
                                            ---- 1s 616us/step - loss: 309997.9375 - val loss: 116953.0547
            1945/1945 -
            Epoch 6/10
1945/1945 —
                                       ----- 1s 602us/step - loss: 81577.9453 - val_loss: 14191.3877
            Epoch 7/10
            1945/1945 -
                                                - 1s 611us/step - loss: 7712.5869 - val loss: 218.8185
                                                - 1s 613us/step - loss: 135.9781 - val loss: 102.4660
            1945/1945 -
            Epoch 9/10
            1945/1945
                                                -- 1s 599us/step - loss: 95.4779 - val_loss: 97.0490
            Epoch 10/10
            15 605us/step - loss: 102.8793 - val_loss: 108.8134
Restoring model weights from the end of the best epoch: 9.
487/487 — 0s 491us/step
            RMSE: 9.851345353532812
```

\*Modelo 2: Múltiples Capas \*

```
model2 = Sequential([
    Dense(200, input_dim=X_train_scaled.shape[1], activation='relu'), # Primera capa oculta con 100 neuronas
    Dense(50, activation='relu'), # Segunda capa oculta con 20 neuronas
Dense(1, activation='linear') # Capa de salida para regresión
1)
# Compilación del modelo con el optimizador Adam
             Adam(learning_rate=0.001)
model2.compile(optimizer=optimizer, loss='mean_squared_error')
# Callback de EarlyStopping para evitar el sobreajuste
early_stopping = EarlyStopping(monitor='val_loss', patience=5, verbose=1, restore_best_weights=True)
# Entrenamiento del modelo con EarlyStoppina
model2.fit(X_train_scaled, y_train, epochs=10, batch_size=32, validation_data=(X_val_scaled, y_val), callbacks=[early_stopping])
# Evaluación del modelo
  pred = model2.predict(X_val_scaled)
raise ValueError("Las predicciones contienen NaN")
# Cálculo del RMSE para evaluar el rendimiento del modelo
rmse = sqrt(mean_squared_error(y_val, y_pred))
print(f'RMSE: {rmse}')
Epoch 1/10
C:\Users\CHRISTIAN\anaconda3\lib\site-packages\keras\src\layers\core\dense.py:86: UserWarning: Do not pass an `input_shape`/`in put_dim` argument to a layer. When using Sequential models, prefer using an `Input(shape)` object as the first layer in the mod
  super().__init__(activity_regularizer=activity_regularizer, **kwargs)
                                2s 754us/step - loss: 1203407.5000 - val_loss: 51299.4648
Epoch 2/10
2066/2066
                                — 1s 716us/step - loss: 37058.0742 - val loss: 11845.1768
Epoch 3/10
2066/2066
                                — 2s 719us/step - loss: 7206.3481 - val_loss: 1883.6832
Epoch 4/10
                                 - 2s 722us/step - loss: 1215.3699 - val loss: 666.9042
2066/2066
Epoch 5/10
                                2s 737us/step - loss: 395.8275 - val loss: 492.5071
2066/2066
Epoch 6/10
2066/2066

    2s 725us/step - loss: 283.0558 - val loss: 659.5401

2066/2066
                                — 1s 713us/step - loss: 341.4096 - val_loss: 618.0717
Epoch 8/10
2066/2066
                                — 2s 727us/step - loss: 418.1860 - val_loss: 289.3042
Epoch 9/10
2066/2066
                                — 2s 722us/step - loss: 258.7532 - val_loss: 192.1150
Epoch 10/10
2066/2066
                                 - 2s 721us/step - loss: 308.6641 - val_loss: 233.8360
Restoring model weights from the end of the best epoch: 9.
                                0s 606us/step
RMSE: 13.86055419151385
```

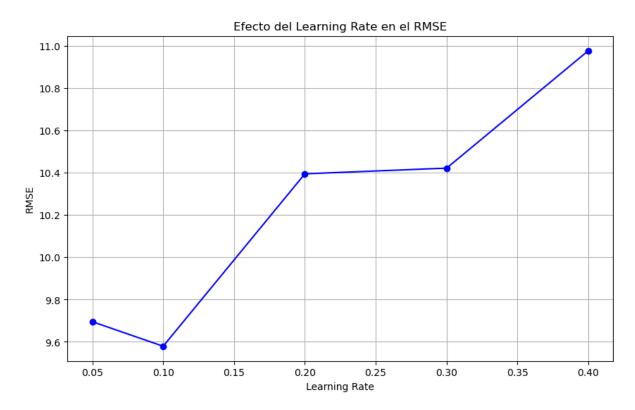
Cabe precisar que existirá un modelo 3, pero será resultante de la calibración del modelo 2 lo cual se observará en las siguientes páginas.

# 5. PARTE IV: DESARROLLO Y CALIBRACIÓN DE MODELOS

Se realizó la calibración de ambos modelos, por ejemplo a nivel de learning rate:

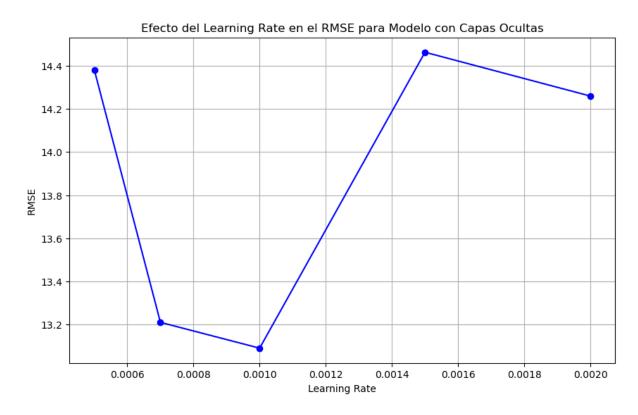
```
In [80]: learning_rates = [0.05, 0.1, 0.2, 0.3, 0.4]
         rmse_results = []
         for lr in learning_rates:
             print(f"Entrenando modelo con learning rate: {lr}")
             model1 = Sequential([
                 Dense(1, input dim=X train scaled.shape[1], activation='linear')
             1)
             model1.compile(optimizer=Adam(learning_rate=lr), loss='mean_squared_error')
             model1.fit(X\_train\_scaled, y\_train, epochs=50, batch\_size=32, validation\_data=(X\_val\_scaled, y\_val), verbose=0)
             y_pred = model1.predict(X_val_scaled)
             rmse = sqrt(mean_squared_error(y_val, y_pred))
             rmse results.append(rmse)
         plt.figure(figsize=(10, 6))
         plt.plot(learning_rates, rmse_results, marker='o', linestyle='-', color='b')
         plt.title('Efecto del Learning Rate en el RMSE')
         plt.xlabel('Learning Rate')
         plt.ylabel('RMSE')
         plt.grid(True)
         plt.show()
```

Gráfico n.º 7 - Calibración Learning Rate Modelo 1



Y para el modelo 2 con capas ocultas (múltiples capas):

Gráfico n.º 78- Calibración Learning Rate Modelo 2



Como podemos observar, en ambos casos el óptimo del Leargnin Rate es 0.001; ahora procedemos calibrar aún más el modelo 2:

```
# Configuraciones de neuronas para cada modelo
     'Config 1': [160, 80, 45, 25, 20], # Configuración original
     'Config 2': [165, 80, 45, 25, 20], # Ligeramente más grande
'Config 3': [170, 80, 45, 25, 20], # Ligeramente más pequeño
'Config 4': [175, 80, 45, 25, 20], # Configuración original
     'Config 5': [180, 80, 45, 25, 20], # Ligeramente más grande
     'Config 6': [160, 85, 45, 25, 20], # Ligeramente más pequeño 'Config 7': [160, 75, 45, 25, 20], # Configuración original
     'Config 8': [160, 70, 45, 25, 20], # Ligeramente más grande
     'Config 9': [160, 80, 45, 20, 20], # Ligeramente más pequeño 'Config 10': [160, 80, 45, 20, 15], # Configuración original 'Config 11': [160, 80, 45, 25, 10], # Ligeramente más grande 'Config 12': [160, 90, 45, 25, 20] # Ligeramente más pequeño
learning_rate = 0.001
results = []
# Pruebas de cada configuración
for config_name, layers in configurations.items():
    print(f"Probando modelo {config_name} con configuración de capas: {layers}")
    model = Sequential()
    model.add(Dense(layers[0], input_dim=X_train_scaled.shape[1], activation='elu'))
    model.add(BatchNormalization())
    for neurons in layers[1:-1]:
          model.add(Dense(neurons, activation='elu'))
          model.add(BatchNormalization())
    # Última capa antes de la salida
     model.add(Dense(layers[-1], activation='relu'))
    model.add(BatchNormalization())
     model.add(Dense(1, activation='linear')) # Salida para regresión
    optimizer = Adam(learning_rate=learning_rate)
    model.compile(optimizer-optimizer, loss='mean_squared_error')
early_stopping = EarlyStopping(monitor='val_loss', patience=10, verbose=1, restore_best_weights=True)
    model.fit(X_train_scaled, y_train, epochs=50, batch_size=32, validation_data=(X_val_scaled, y_val), callbacks=[early_stopping
     y_pred = model.predict(X_val_scaled)
     if np.isnan(y_pred).any():
          raise ValueError("Las predicciones contienen NaN")
     rmse = sqrt(mean_squared_error(y_val, y_pred))
     results.append((config_name, layers, rmse))
    print(f"RMSE para {config name}: {rmse}")
# Encontrar la mejor configuración
best_result = min(results, key=lambda x: x[2])
print(f"Mejor configuración: {best_result[0]} con capas {best_result[1]}, RMSE: {best_result[2]}")
```

El código anterior fue diseñado para explorar la configuración óptima de una red neuronal multicapa. Una vez establecida la estructura básica, se procedió a calibrar ciertos parámetros y a incorporar diversas técnicas para mejorar el rendimiento y la generalización del modelo. Entre estas técnicas se incluyeron el uso de dropout y normalización por lotes (batch normalization) para regularizar el modelo, la implementación de paradas tempranas (early stopping) para evitar el sobreajuste, la optimización mediante el algoritmo Adam, y la selección cuidadosa de las funciones de activación. Estas estrategias fueron aplicadas con el objetivo de afinar el modelo y alcanzar los mejores resultados posibles en términos de precisión y eficiencia durante el entrenamiento.

En el modelo final calibrado (Modelo 3), tenemos lo siguiente:

```
In [30]: from tensorflow.keras.models import Sequential
           from tensorflow.keras.layers import Dense, BatchNormalization, Dropout from tensorflow.keras.optimizers import Adam
           from tensorflow.keras.callbacks import EarlyStopping
           from sklearn.metrics import mean_squared_error
           from math import sqrt
          # Modelo modifica
          model = Sequential([
               Dense(360, input_dim=X_train_scaled.shape[1], activation='relu'),
               Dropout(0.3), # Agregar dropout para regularización
               BatchNormalization(),
Dense(180, activation='elu'),
               Dropout(0.3), # Agregar dropout para regularización
BatchNormalization(),
Dense(60, activation='relu'),
               BatchNormalization().
               Dense(25, activation='relu'),
               BatchNormalization(),
Dense(20, activation='relu'),
              BatchNormalization(),
Dense(1, activation='linear') # Salida para regresión
           # Compilación con Adam
                        Adam(learning_rate=0.001)
           model.compile(optimizer=optimizer, loss='mean_squared_error')
           # Callback de EarlyStopping modificado
           early_stopping = EarlyStopping(monitor='val_loss', patience=10, verbose=1, restore_best_weights=True)
           # Entrenamiento con EarlyStoppina
          model.fit(X_train_scaled, y_train, epochs=50, batch_size=50, validation_data=(X_val_scaled, y_val), callbacks=[early_stopping])
           # Evaluación
            _pred = model.predict(X_val_scaled)
          if np.isnan(y_pred).any():
raise ValueError("Las predicciones contienen NaN")
                 sqrt(mean_squared_error(y_val, y_pred))
          print(f'RMSE: {rmse}')
```

#### Sobre el modelo:

- **Dense:** Representa una capa densamente conectada (o completamente conectada). Cada neurona en una capa densa recibe entrada de todas las neuronas de la capa anterior, lo cual es una característica de las redes neuronales tradicionales.
- La primera capa densa con 360 neuronas, donde input\_dim especifica el número de características de entrada del modelo. La función de activación 'relu' (Rectified Linear Unit) es comúnmente utilizada por su eficiencia y efectividad en redes neuronales profundas.
- **Dropout(0.3):** Una técnica de regularización donde aleatoriamente se "apagan" un porcentaje de las neuronas durante el entrenamiento para prevenir el sobreajuste. 0.3 significa que el 30% de las neuronas en la capa anterior se desactivarán aleatoriamente en cada paso durante el entrenamiento.
- **BatchNormalization():** Normaliza las activaciones de la capa anterior al restar la media del batch y dividir por la desviación estándar, lo que ayuda a mejorar la velocidad, rendimiento y estabilidad del entrenamiento.

Considerando que tiene normalización, aplicación técnicas y ajustes, se considero a este modelo 2 calibrado, como el modelo 3 "Múltiples capas Adam y calibración de hiperparámetros"

Siendo el mejor modelo obtenido, se generaron los predicts

#### Hito:

- RMSE EN LAB LOCAL 8.41185
- RMSE EN KAGGLE 8.64070
- RANKING EN EL PUBLIC LEADERBOARD 5 DE 21 EQUIPOS.
- DIFERENCIA CON EL RANKING 1: 8.64070 8.4901 = 0.1505

# 6. PARTE V: VISUALIZACIÓN DE RESULTADOS

A continuación se presentan los resultados obtenidos por cada tipo de modelo empleado:

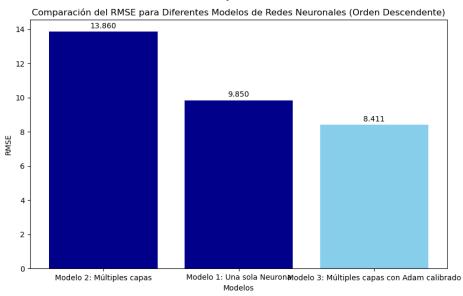


Gráfico n.º X- Comparación de modelos

Como podemos observar, inicialmente daba la impresión de que el modelo de una sola neurona sería mejor que de múltiples capas, pero a lo largo de las calibraciones realizadas, empleando batchnormalization, técnica de dropout y búsqueda intensiva de parámetros óptimos que incluye el learning rate, se obtuvieron los siguientes resultados:

- 1. Modelo 1: Una sola Neurona / RMSE OBTENIDO: 9.85
- 2. Modelo 2: Múltiples capas / RMSE OBTENIDO: 13.86
- 3. Modelo 3: Múltiples capas Adam y calibración de hiperparámetros / RMSE OBTENIDO: 8.411

En ese sentido, el modelo 3 al tener el mejor RMSE, fue utilizado para generar las predicciones y así obtener el puntaje de 8.64070, con lo cual nuestro equipo se colocó en el ranking 5 de la competencia.

Table 0	Modelos		hadaa ah	+
i avia z.	Modelos v	i los resuli	เสนบร บบ	temuos

Detalle del Modelo	RMSE JupyterNotebook	Puntuación Kaggle
Modelo 01: Red neuronal – Una sola neurona	9.85	10.11
Modelo 02: Red neuronal – Multicapa	13.86	15.86
Modelo 03: Red neuronal – Multicapa Adam calibrado	8.411	8.64

Recomendaríamos utilizar el modelo de redes neuronales Adam, calibrando los parámetros de número de capas, tasa de aprendizaje, neuronas por capa, tipo de función de activación, implementando estrategias de batchnormalization, dropout y earlystopping, a efectos de mejorar aún más el modelo.

Ahora bien, respondiendo a la pregunta de interés... sí, es posible siendo que las variables indican que existiría una relación entre las características musicales de una canción y el año en que fue publicada/lanzada.

# Anexos.

Anexo 1. Resultado en Competencia (pdf)

Anexo 2. HTML Notebook

Anexo 3. Notebook en Jupyter