

# POPULARIDAD EN CANCIONES

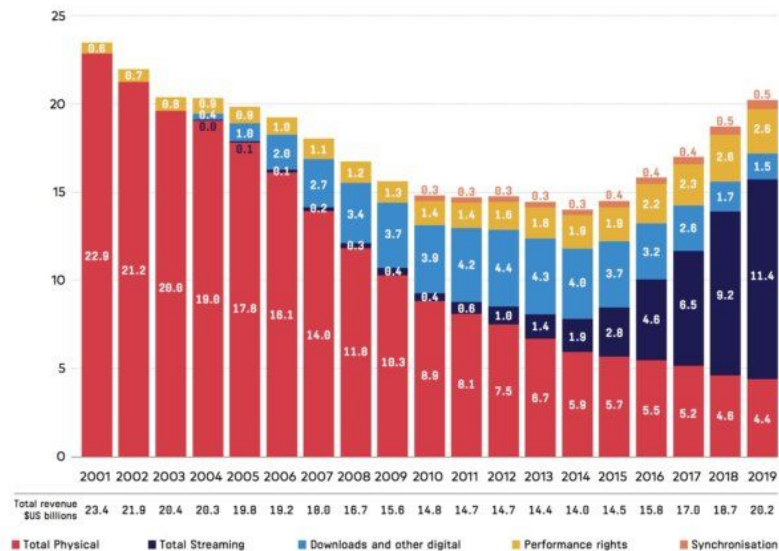
Alvarado Morán Óscar  
Bermúdez Marbán Dante  
García Avendaño Martín  
Licenciatura en Ciencia de Datos, IIMAS, UNAM

# Contenido

- Introducción
- Datos
- Metodología
- Resultados
- Conclusiones
- Referencias

# Introducción

Global Recorded Music Industry Revenues 2001-2019 (US\$ Billions)



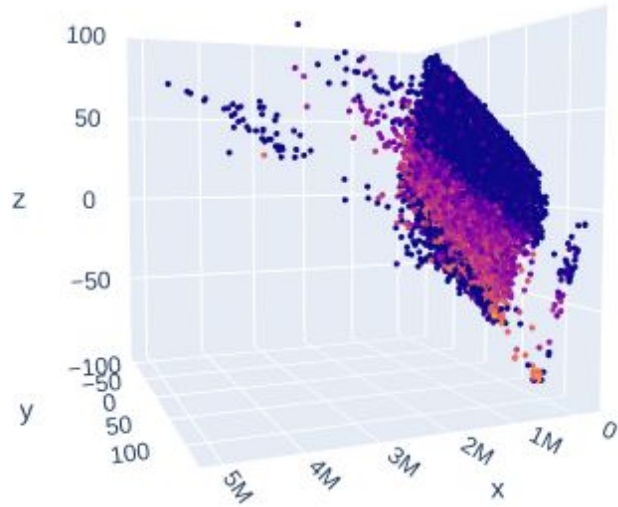
[Statista](#)

# Datos

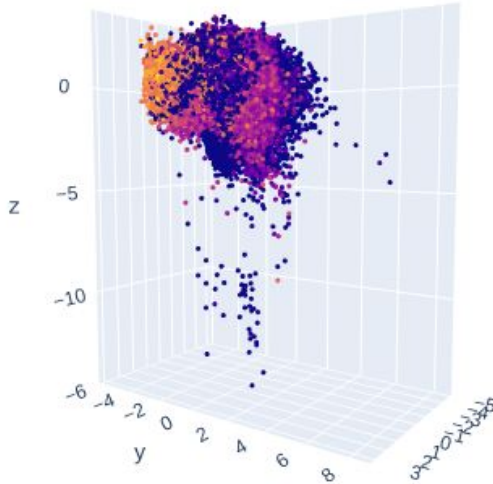
- acousticness
- artists
- danceability
- duration\_ms
- energy
- explicit
- id
- instrumentalness
- key
- liveness
- loudness
- mode
- name
- popularity
- release\_date
- speechiness
- tempo
- valence
- year

# Datos: Reducción

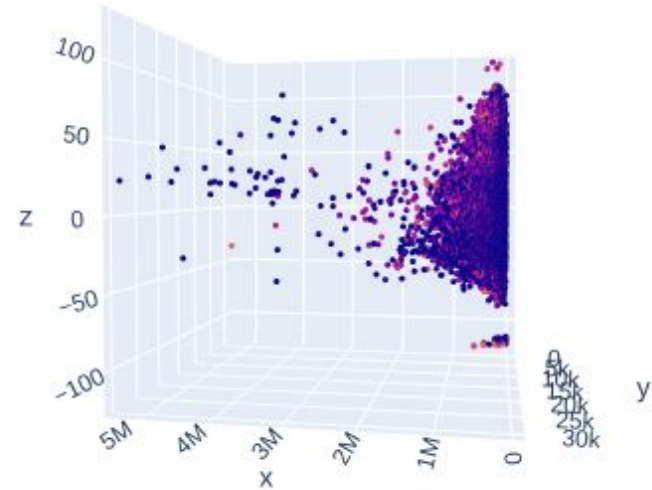
PCA



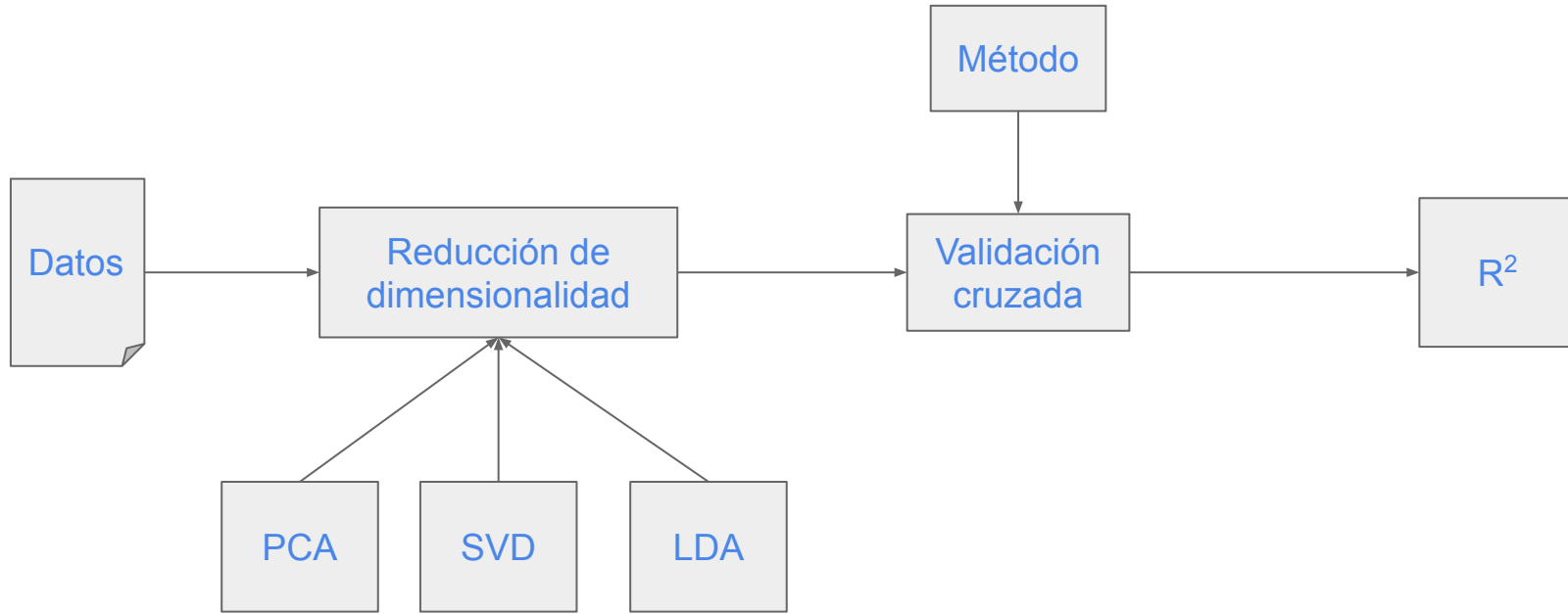
LDA



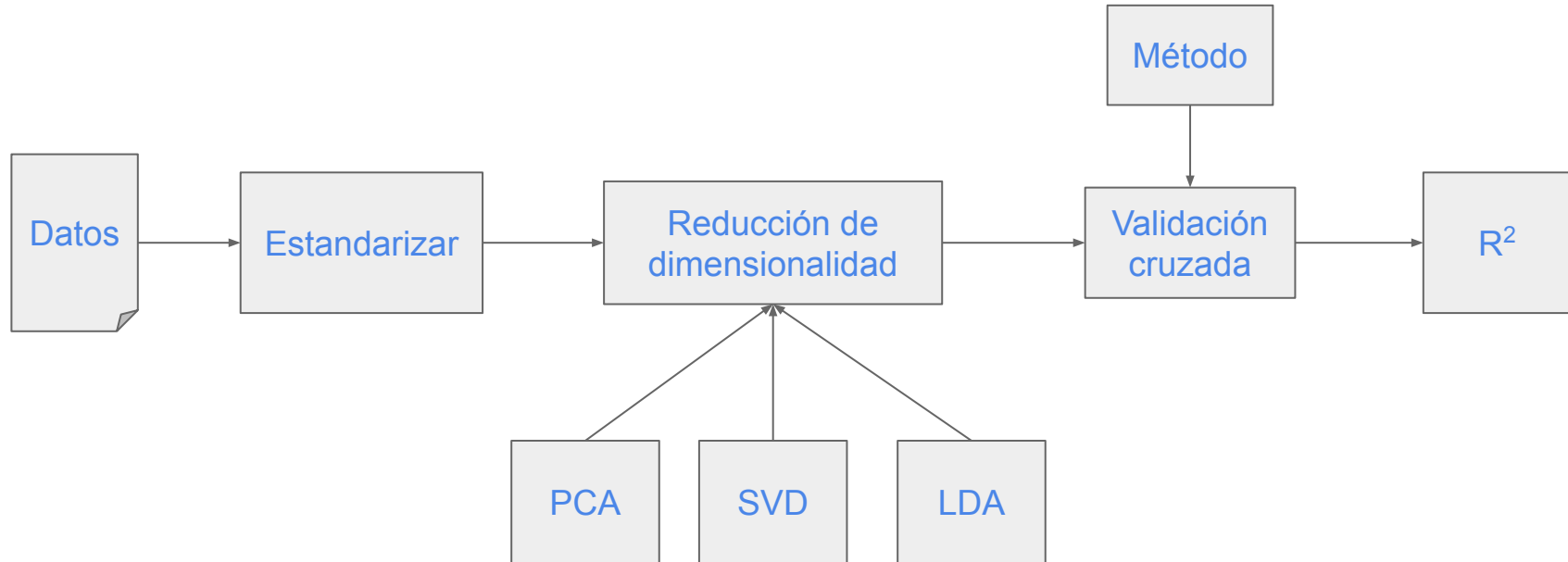
SVD



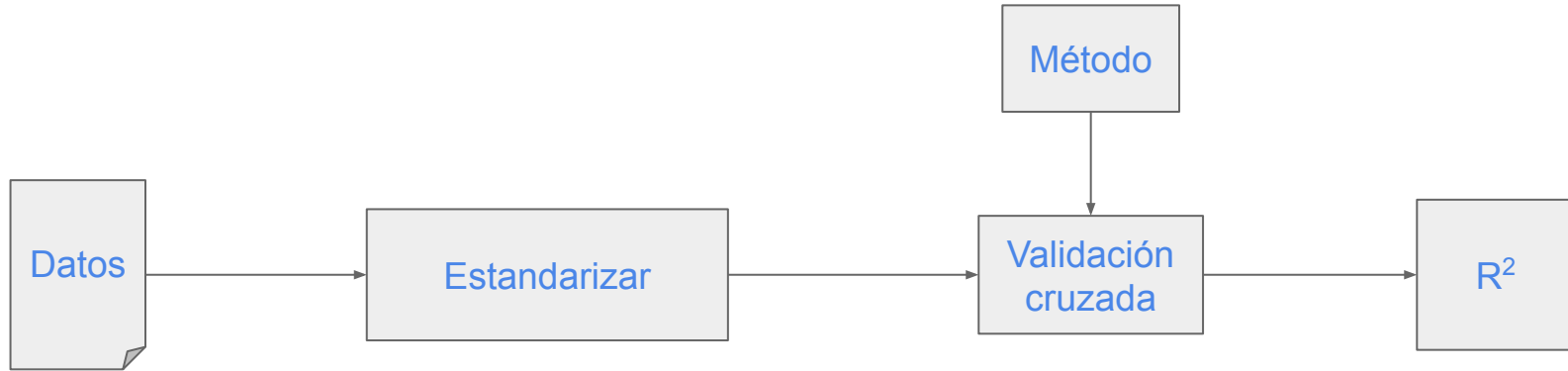
# Metodología



# Metodología



# Metodología





# Resultados: $R^2$

Sin estandarizar	PCA	SVD	LDA
Regresión lineal	0.273	0.271	0.368
Regresión poli	0.304	0.395	0.407
K - vecinos (10)	0.283	0.282	0.455
Árbol de decisión	0.49	0.302	0.464
Bosque aleatorio	0.525	0.48	0.473
GBT	0.471	0.27	0.448
LGBM	0.05	-6.169	0.038

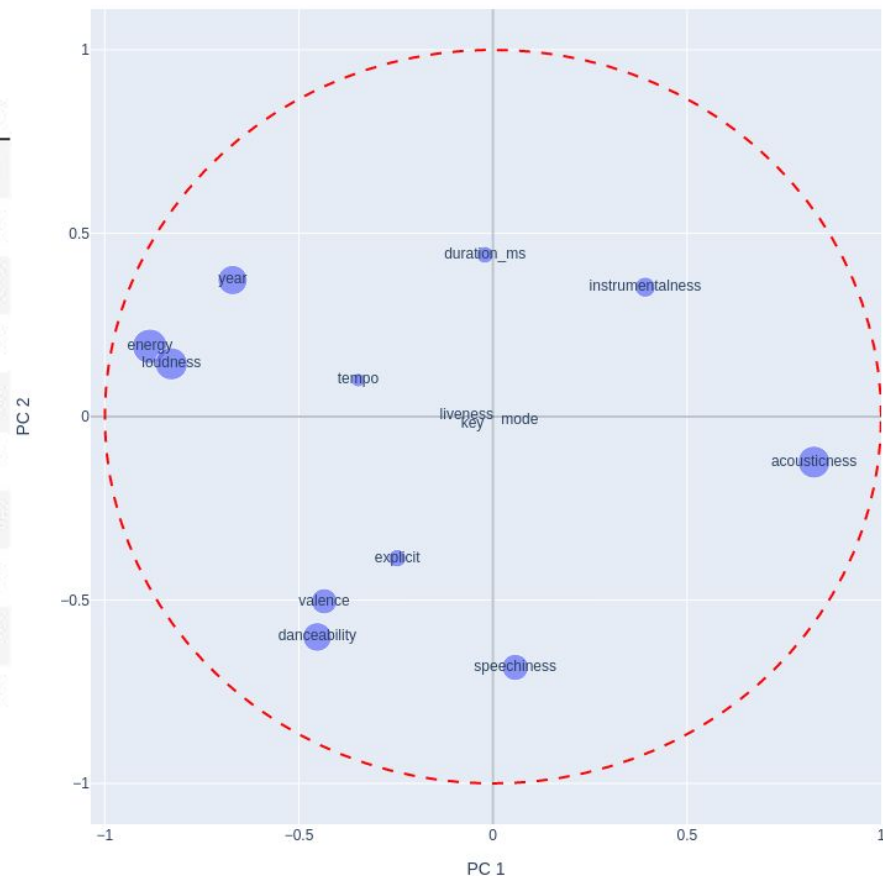
Estandarizar	PCA	SVD	LDA
Regresión lineal	0.23	0.23	0.368
Regresión poli	0.33	0.326	0.407
K - vecinos (10)	0.296	0.296	0.455
Árbol de decisión	0.349	0.312	0.464
Bosque aleatorio	0.314	0.317	0.473
GBT	0.38	0.261	0.448
LGBM	0.027	-6.767	0.038

Estandarizar	$R^2$	Tiempo [s]
Regresión lineal	0.368	0.957
Regresión poli	0.591	13.3
K - vecinos (10)	0.539	165
Árbol de decisión	0.601	2.98
Bosque aleatorio	0.671	111
GBT	0.603	18.1
LGBM	0.636	1.56

# Conclusiones

	importance
<b>year</b>	0.534975
<b>instrumentalness</b>	0.078800
<b>loudness</b>	0.050508
<b>duration_ms</b>	0.047704
<b>energy</b>	0.038958
<b>acousticness</b>	0.037943
<b>danceability</b>	0.036465
<b>liveness</b>	0.036304
<b>valence</b>	0.035960
<b>speechiness</b>	0.034949
<b>tempo</b>	0.034516
<b>key</b>	0.015117
<b>explicit</b>	0.014658
<b>mode</b>	0.003144

name	popularity	year
drivers license	100	2021
Mood (feat. iann dior)	96	2020
positions	96	2020
DÁKITI	95	2020
BICHOTA	95	2020
34+35	94	2020
Whoopty	94	2020
WITHOUT YOU	94	2020
Therefore I Am	94	2020
LA NOCHE DE ANOCHE	94	2020



# Referencias

- [Datos](#)
- [Datos spotify](#)
- [API spotify](#)
- [PCA](#)
- [LDA](#)
- [SVD](#)
- [Regresión Lineal](#)
- [Polinomial](#)
- [K-nn](#)
- [Árboles de decisión](#)
- [Bosque Aleatorio](#)
- [Gradient Boosting](#)
- [Light Gradient Boosting Machine](#)
- [Grid Search](#)
- [Random Search](#)
- [Github](#)