

EVENTPEAK



Carlota Salazar, Yushan Yang, Julia Yi

Índice

1. Extracción de datos
2. Transformación de datos
3. Limpieza de datos
4. Carga en Parquet
5. Exploración de datos

EXTRACCIÓN

1 - EXTRACCIÓN

Ticketmaster API Explorer V2.0

Make live API calls without writing a single line of code! The API Explorer currently works with the [Discovery API](#)

Discovery API v2

Please enter your API key

Methods

All Get

Event Search

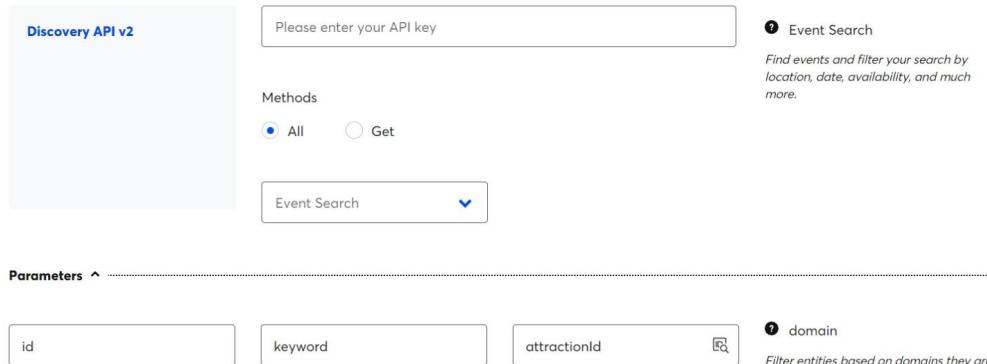
Find events and filter your search by location, date, availability, and much more.

Parameters ^

id keyword attractionId

domain

Filter entities based on domains they are available on

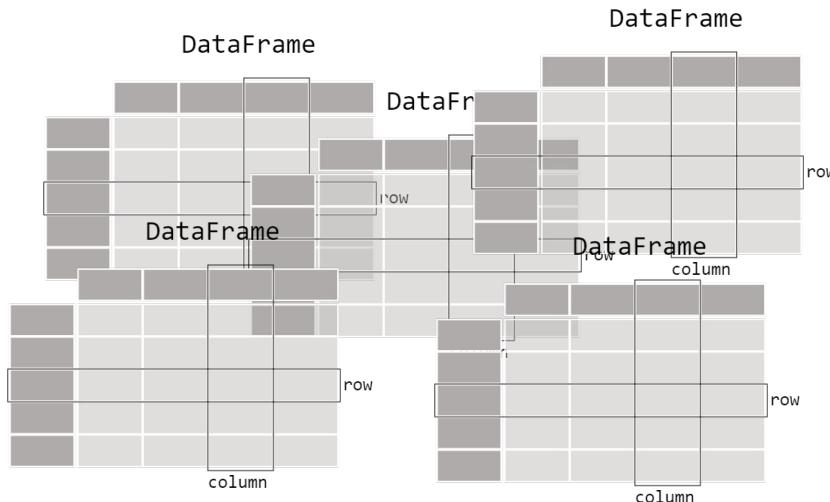


Para filtrar

```
import requests
import pandas as pd
import time
import math
from pathlib import Path
from pandas.io.parsers import read_csv
from google.colab import drive
import json
```

Para extraer

1.1. DATOS “EN BRUTO”



5129

Filas duplicadas 2637

Num filas sin duplicados 2492



Luego veremos esto en la
parte de limpieza.

TRANSFORMACIÓN

2 - VARIABLES DESCARTADAS

	name	type	id	test	url	locale	images	sales	dates	classifications	promoter	promoters	priceRanges	_links	_embedded	seatmap
0	Servicio de Autobús - BRESH - Grand Opening ...	event	Z698xZ2qZaaM	False	https://www.ticketmaster.es/event/servicio-de-...	es-es	[{"ratio": "16_9", "url": "https://s1.ticketm..."}]	{"public": [{"start": "2024-01-24T14:00...", "end": "2024-06-28T00:00..."}]}	{"start": "2024-06-28T00:00...", "end": "2024-06-29T14:00..."}]	[{"primary": true, "segment": {"id": "KZFzniwn..."}]}	{"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2024-06-29T14:00...", "end": "2024-06-29T14:00..."}, {"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2024-06-29T14:00...", "end": "2024-06-29T14:00..."}]	[{"type": "standard", "currency": "EUR", "min": "..."}]	{"self": {"href": "/discovery/v2/events/Z698xZaaM"}, "venues": [{"name": "ServicioBus ConcertMusic..."}]}	NaN		
1	Manolo García - Concert Music Festival	event	Z698xZ2qZaaC4	False	https://www.ticketmaster.es/event/manolo-garcia...	es-es	[{"ratio": "3_2", "url": "https://s1.ticketm..."}]	{"public": [{"start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}]}	{"start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}, {"start": "2024-06-29T11:00...", "end": "2024-06-29T11:00..."}]	[{"primary": true, "segment": {"id": "KZFzniwn..."}]}	{"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}, {"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2024-06-29T11:00...", "end": "2024-06-29T11:00..."}]	[{"type": "standard", "currency": "EUR", "min": "..."}]	{"self": {"href": "/discovery/v2/events/Z698xZaaC4"}, "venues": [{"name": "Concert Music Festival..."}]}	NaN		
2	Servicio de Autobús - Manolo García - Concert...	event	Z698xZ2qZaaC9	False	https://www.ticketmaster.es/event/servicio-de-...	es-es	[{"ratio": "16_9", "url": "https://s1.ticketm..."}]	{"public": [{"start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}]}	{"start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}, {"start": "2024-06-29T11:00...", "end": "2024-06-29T11:00..."}]	[{"primary": true, "segment": {"id": "KZFzniwn..."}]}	{"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2023-11-30T11:00...", "end": "2024-06-29T11:00..."}, {"id": "5875", "name": "CONCERT TOUR Y EN VIVO...", "segment": {"id": "KZFzniwn..."}, "start": "2024-06-29T11:00...", "end": "2024-06-29T11:00..."}]	[{"type": "standard", "currency": "EUR", "min": "..."}]	{"self": {"href": "/discovery/v2/events/Z698xZaaC9"}, "venues": [{"name": "ServicioBus ConcertMusic..."}]}	NaN		



3 - VARIABLES QUE SE QUEDAN

		name	sales	dates	classifications	promoter	priceRanges	_embedded
1671	Homenaje a John Williams	{'public': {'startDateTime': '2024-01-30T17:21...}}	{'access': {'startDateTime': '2024-12-07T18:45...}}	[{'primary': True, 'segment': {'id': 'KZFzniwn...}}	{'id': '5703', 'name': 'Overlook Events SAS'}	[{'type': 'standard including fees', 'currency...']}	{'venues': [{"name": "Auditorio Forum CCIB", "type": "..."}]}	
4212	Black Coffee	{'public': {'startDateTime': '2024-02-23T16:30...}}	{'start': {'localDate': '2024-10-05', 'localTi...}}	[{'primary': True, 'segment': {'id': 'KZFzniwn...}}	{'id': '5895', 'name': 'Only Tickets, S.L.,'}	[{'type': 'standard including fees', 'currency...']}	{'venues': [{"name": "Hí Ibiza", "type": "..."}]}	
3171	David Guetta & Morten present FUTURE RAVE	{'public': {'startDateTime': '2024-02-26T11:00...}}	{'start': {'localDate': '2024-09-27', 'localTi...}}	[{'primary': True, 'segment': {'id': 'KZFzniwn...}}	{'id': '5895', 'name': 'Only Tickets, S.L.,'}	[{'type': 'standard including fees', 'currency...']}	{'venues': [{"name": "Hí Ibiza", "type": "..."}]}	
2438	Armin van Buuren at Ushuaïa Ibiza	{'public': {'startDateTime': '2024-02-14T17:30...}}	{'start': {'localDate': '2024-09-29', 'localTi...}}	[{'primary': True, 'segment': {'id': 'KZFzniwn...}}	{'id': '5895', 'name': 'Only Tickets, S.L.,'}	[{'type': 'standard', 'currency': 'EUR', 'min'...}]	{'venues': [{"name": "Ushuaïa Ibiza", "type": "..."}]}	
4439	Marlon	{'public': {'startDateTime': '2023-11-07T11:00...}}	{'access': {'startDateTime': '2024-05-17T19:30...}}	[{'primary': True, 'segment': {'id': 'KZFzniwn...}}	{'id': '2771', 'name': 'Promociones Musicales ...'}	[{'type': 'standard', 'currency': 'EUR', 'min'...}]	{'venues': [{"name": "Sala Razzmatazz 1", "type": "..."}]}	

Son variables con mucha información

True, False → true, false

string → diccionario

4 - PRIMER ANÁLISIS DE VARIABLES

```
sales
{'public': {'startDateTime: '2024-01-24T14:00:00Z',
  'startTBD': False,
  'startTBA': False,
  'endDateTime
```

Sales → ventas de entradas.

Name → nombre del evento.

```
name
'Servicio de Autobús - BRESH - Grand Opening - Concert Music Festival'
```

```
dates
{'start': {'localDate': '2024-06-28',
  'dateTBD': False,
  'dateTBA': False,
  'timeTBA': True,
  'noSpecificTime': False},
  'timezone': 'Europe/Madrid',
  'status': {'code': 'onsale'},
  'spanMultipleDays': False}
```

Dates → realización del evento.

```
classifications
[{'primary': True,
 'segment': {'id': 'KZFzniwnSyZfZ7v7nn', 'name': 'Film'},
 'genre': {'id': 'KnvZfZ7vAkJ', 'name': 'Music'},
 'subGenre': {'id': 'KZazBElonSMnZfZ7vave', 'name': 'Music'},
 'family': False}]
```

Classifications → clasificación del evento

```
promoter
{'id': '5875', 'name': 'CONCERT TOUR Y EN VIVO AIE'}
```

Promoter → promotor

```
priceRanges
[{'type': 'standard', 'currency': 'EUR', 'min': 17.5, 'max': 27.5},
 {'type': 'standard including fees',
  'currency': 'EUR',
  'min': 19.0,
  'max': 29.0}]
```

priceRanges → Precios del evento

```

_embedded
{'venues': {
    'name': 'Concert Music Festival',
    'type': 'venue',
    'id': 'Z598xZ2qZdkaF',
    'test': False,
    'url': 'https://www.ticketmaster.es/venue/concert-music-festival-chiclana-de-la-frontera-entradas/rccsancpchi/114',
    'locale': 'en-us',
    'postalCode': '111139',
    'timezone': 'Europe/Madrid',
    'city': {'name': 'Chiclana de la Frontera'},
    'state': {'name': 'Cádiz'},
    'country': {'name': 'Spain', 'countryCode': 'ES'},
    'address': {'line1': 'Calle Avenida (Poblado Sancti Petri)'},
    'location': {'longitude': '-6.20735', 'latitude': '36.39397'},
    'upcomingEvents': {'mfx-es': 37, '_total': 37, '_filtered': 0},
    '_links': {'self': {'href': '/discovery/v2/venues/Z598xZ2qZdkaF?locale=en-us'}}},
    'attractions': [
        {'name': 'Manolo Garcia',
        'type': 'attraction',
        'id': 'K8vZ917G40f',
        'test': False,
        'url': 'https://www.ticketmaster.com/manolo-garcia-tickets/artist/725914',
        'locale': 'en-us',
        'externalLinks': {'youtube': [{'url': 'https://www.youtube.com/channel/UCP9yQtx8Za8wlfnuXt33iGA'}]}]
}

```

_embedded → contenedor de lugar y otros

```

'classifications': [
    {'primary': True,
    'segment': {'id': 'KZFzniwnSyZfZ7v7nJ', 'name': 'Music'},
    'genre': {'id': 'KnvZfZ7vAeA', 'name': 'Rock'},
    'subGenre': {'id': 'KZazBEonSMnZfZ7v6dt', 'name': 'Alternative Rock'},
    'type': {'id': 'KZAyXgnZfZ7v7nI', 'name': 'Undefined'},
    'subType': {'id': 'KZFzBErXgnZfZ7v7lJ', 'name': 'Undefined'},
    'family': False}]

```

LIMPIEZA

5 - ELIMINACIÓN DE FILAS

5129

Filas duplicadas 2637

Num filas sin duplicados 2492

```
generos = df_1['genero'].value_counts()  
generos
```

```
Dance/Electronic    712  
Rock                692  
Other                198  
Undefined            107  
Alternative          93  
Hip-Hop/Rap          34  
Music                25  
World                25  
Classical              11  
Theatre                  6  
Blues                    4  
Pop                      3  
Fairs & Festivals        1  
Name: genero, dtype: int64
```

	name
0	The Music of Hans Zimmer & others - A Celebrat...
1	The Music of Hans Zimmer & others - A Celebrat...
2	The Music of Hans Zimmer & others - A Celebrat...
3	Servicio de Autobús - BRESH - Grand Opening ...
4	Servicio de Autobús - Manolo García - Concert...
...	...
796	Alice Cooper - Diamond Meet and Greet Package
797	Take That - VIP PACKAGE - Alma Festival 2024 ...
798	ATRÉVETE, el musical

```
['film', 'servicio de autobús', 'festival', 'vip',  
'meet', 'meet & greet', 'ticket', 'package', 'musical']
```

6 - VARIABLES FINALES DE TICKETMASTER

name	promoter	startDateTime	end_datetime	min_price	max_price	genero	subgnero	nameArtist	genre	subgenre	type	subtype	VenueName	VenueCity	VenueState
Jamie Cullum	Tickex Experience, S.L.	2024-06-14T20:00:00Z	2024-06-12T21:59:00Z	30.01	59.29	Music	Music	Jamie Cullum	Jazz	Jazz	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga
Sheryl Crow	Tickex Experience, S.L.	2024-06-21T20:00:00Z	2024-06-19T21:59:00Z	30.01	80.00	Music	Music	Sheryl Crow	Rock	Alternative Rock	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga
Aitana	Tickex Experience, S.L.	2024-06-22T20:00:00Z	2024-06-20T21:59:00Z	65.18	146.43	Music	Music	Aitana	Pop	Pop Vocal	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga

6 - INCORPORACIÓN VARIABLES SECUNDARIAS

	name	promoter	startDateTime	endSalesTime	min_price	max_price	genero	subgenero	nameArtist	genre	subgenre	type	subtype	VenueName	VenueCity	VenueState	Spotify_Mensuales	Last.fm_Oyentes
0	Jamie Cullum	Tickex Experience, S.L.	2024-06-14T20:00:00Z	2024-06-12T21:59:00Z	30.01	59.29	Music	Music	Jamie Cullum	Jazz	Jazz	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga	514905	1210983
1	Sheryl Crow	Tickex Experience, S.L.	2024-06-21T20:00:00Z	2024-06-19T21:59:00Z	30.01	80.00	Music	Music	Sheryl Crow	Rock	Alternative Rock	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga	1119957	1596151
2	Aitana	Tickex Experience, S.L.	2024-06-22T20:00:00Z	2024-06-20T21:59:00Z	65.18	146.43	Music	Music	Aitana	Pop	Pop Vocal	Undefined	Undefined	Auditorio Starlite	Marbella	Málaga	2898099	151211

Resumen:

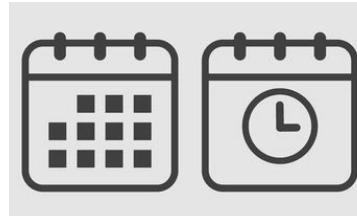
- 1000 observaciones x 18 variables
- 4 numéricas, 2 fechas, resto categóricas

CARGA EN PARQUET

7 - GESTIÓN DATOS PARA PARQUET

PD1	
	EventPeak.parquet
	PD1.ipynb
	codigo_entrega_1.i...
	codigo_entrega_2.i...
	col.xlsx
	col_em.csv
	d_17_02.csv
	d_1_03.csv
	d_21_02.csv
	d_26_02.csv
	d_28_02.csv
	datos_4_03.csv
	datos_5_03.csv

Dividir por:



Año o mes



Ubicación

Distribución de datos por VenueState
(comunidad autónoma)

Madrid	326
Barcelona	225
Málaga	129
Valencia	38
Biscay	37
Seville	29
La Corunna	29
Murcia	20
Cantabria	18
Granada	16
Navarre	15
Saragossa	14
Alacant	12
Salamanca	11
Cádiz	10
Asturias	9
Pontevdra	9
Balearic Islands	8
Córdoba	8
Badajoz	7
Valladolid	6
Almeria	6
Guipuscoa	6
Santa Cruz de Tenerife	6
Zamora	5
León	4
Araba	3
Burgos	3
Cáceres	3
Palencia	3
Jáén	3
Castellón	2
Albacete	2
La Rioja	2
Huelva	1
Guadalajara	1
Toledo	1

EXPLORACIÓN

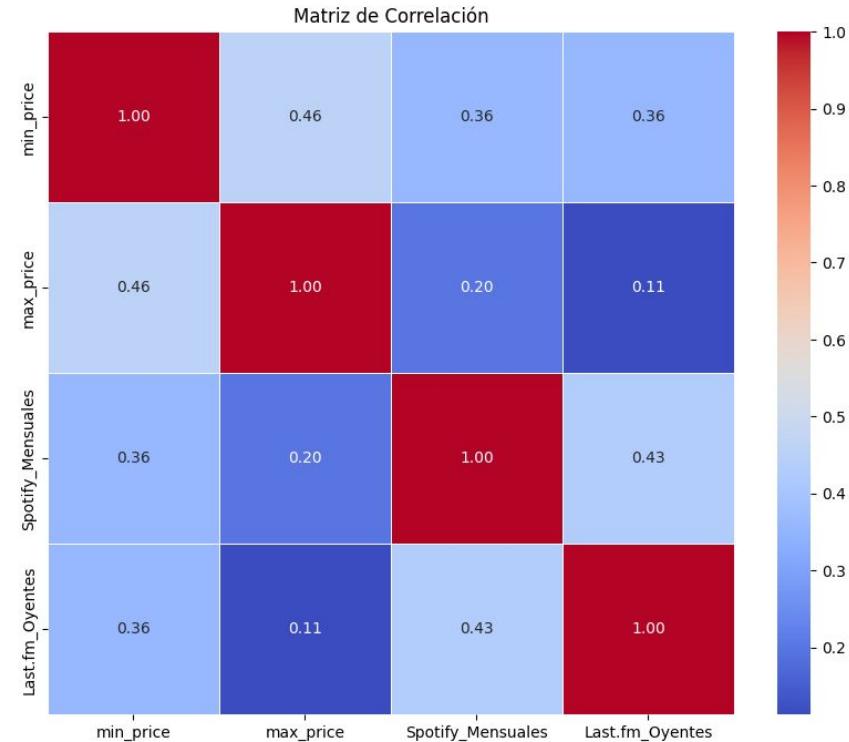
8 - ANÁLISIS DE DATOS

Estadísticas Descriptivas para Variables Numéricas:

	min_price	max_price	Spotify_Mensuales	Last.fm_Oyentes
count	1024.000000	1024.000000	1.027000e+03	1.027000e+03
mean	37.096162	103.724238	3.923214e+06	4.753290e+05
std	22.381562	348.913542	9.272530e+06	8.537190e+05
min	10.000000	10.000000	1.000000e+00	1.000000e+00
25%	25.000000	25.000000	8.274700e+04	2.750100e+04
50%	31.625000	40.000000	4.954890e+05	1.166390e+05
75%	43.000000	90.000000	3.705543e+06	5.063450e+05
max	179.460000	6000.000000	1.031583e+08	4.517917e+06

Moda para Variables Categóricas:

	name	promoter	startDateTime	\			
0	Camilo	Live Nation España S.A.U.	2024-06-21T20:00:00Z				
1	Nan	Nan	2024-07-16T20:00:00Z				
	endSalesTime	min_price	max_price	genero	subgnero	nameArtist	\
0	2024-04-05T19:00:00Z	25.0	25.0	Rock	Pop	Camilo	
1	Nan	Nan	Nan	Nan	Nan	Nan	
	genre	subgenre	type	subtype	VenueName	VenueCity	\
0	Rock	Pop	Undefined	Undefined	Auditorio Starlite	Madrid	
1	Nan	Nan	Nan	Nan	Nan	Nan	
	VenueState	Spotify_Mensuales	Last.fm_Oyentes				
0	Madrid	18544604.0	257632.0				
1	Nan	Nan	Nan				



9 - PRUEBAS CON VARIABLES NO NUMÉRICAS

