

Dataset: “Metacritic best tv shows of all time”

Alumna: Génesis Álvarez Gamboa
Profesor: Miguel Guevara
Asignatura: Taller Integrado 2022

Python es uno de los lenguajes de programación dinámicos más populares que existen entre los que se encuentran Java, Javascript, Go y C#. Aunque es considerado a menudo como un lenguaje "scripting", es realmente un lenguaje de propósito general.

#	Column	Count	Non-Null	Dtype
0	title	3084	Non-null	object
1	release	3084	Non-null	object
2	Summary	3084	Non-null	object
3	metascore	3084	Non-null	int64
4	userscore	3084	Non-null	object
5	link	3084	Non-null	object
6	Img_url	3084	Non-null	object

El dataset original tiene :

- 7 columnas
- 3084 filas
- 1 atributo entero
- 6 strings u objects

```
df.info(verbose=True) #nos da información False nos dará un resumen corto de los datos es conveniente si tu dataframe es pequeño
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 3084 entries, 1 to 3084
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   title                 3084 non-null  object
1   release              3084 non-null  datetime64[ns]
2   summary              3084 non-null  object
3   metascore            3084 non-null  int64
4   userscore            3084 non-null  object
5   link                 3084 non-null  object
6   img_url              3084 non-null  object
7   month                3084 non-null  object
8   year                 3084 non-null  object
9   Indicadores_Metascore 3084 non-null  object
dtypes: datetime64[ns](1), int64(1), object(8)
memory usage: 265.0+ KB
```

Este método imprime información sobre un DataFrame, incluidos el dtype de índice y las columnas, los valores no nulos y el uso de memoria.


```
[4]: df.describe()#Se utiliza para ver algunos detalles estadísticos básicos como percentil, media, std, etc. de un marco de datos o una serie de valores numéricos.  
#nos dice que count es 3084 el cual es la cantidad total de filas  
#mean es el promedio de metascor(e) (puntuaje del 1 al 99) en este caso es 64,8  
#std desviación estándar de 15  
#min el metascor(e) mas bajo pertenece a 13  
#el 25% pertenece al metascor(e) 55  
#el 50% se encuentra en el metascor(e) 66  
#y el máximo metascor(e) es de 99.
```

```
[4]:
```

	metascor(e)
count	3084.000000
mean	64.877108
std	15.012487
min	13.000000
25%	55.000000
50%	66.000000
75%	76.000000
max	99.000000

`df.describe()` : Genera estadísticas descriptivas, las que incluyen aquellas que resumen la tendencia central, la dispersión y la forma de la distribución de un conjunto de datos, excluyendo los valores NaN.

```
[5]: print(df.dtypes) #nos muestra solo los tipos de datos en el marco de datos, object string datetime fecha e int entero.
```

```
title                object
release             datetime64[ns]
summary             object
metascore            int64
userscore           object
link               object
img_url            object
month              object
year               object
Indicadores_Metascore object
dtype: object
```

```
[6]: #pd.set_option('max_rows', None) #nos muestra todos los datos, es bueno sobre todo cuando tenemos un dataset pequeño
```

```
[7]: df.drop(df.index[df['userscore'] == 'tbd'], inplace = True) #eliminamos filas con puntuación de usuario no definida
```

```
[8]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2820 entries, 1 to 3084
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   title                 2820 non-null  object
1   release               2820 non-null  datetime64[ns]
2   summary              2820 non-null  object
3   metascore            2820 non-null  int64
4   userscore            2820 non-null  object
5   link                 2820 non-null  object
6   img_url              2820 non-null  object
7   month                2820 non-null  object
8   year                 2820 non-null  object
9   Indicadores_Metascore 2820 non-null  object
dtypes: datetime64[ns](1), int64(1), object(8)
memory usage: 242.3+ KB
```

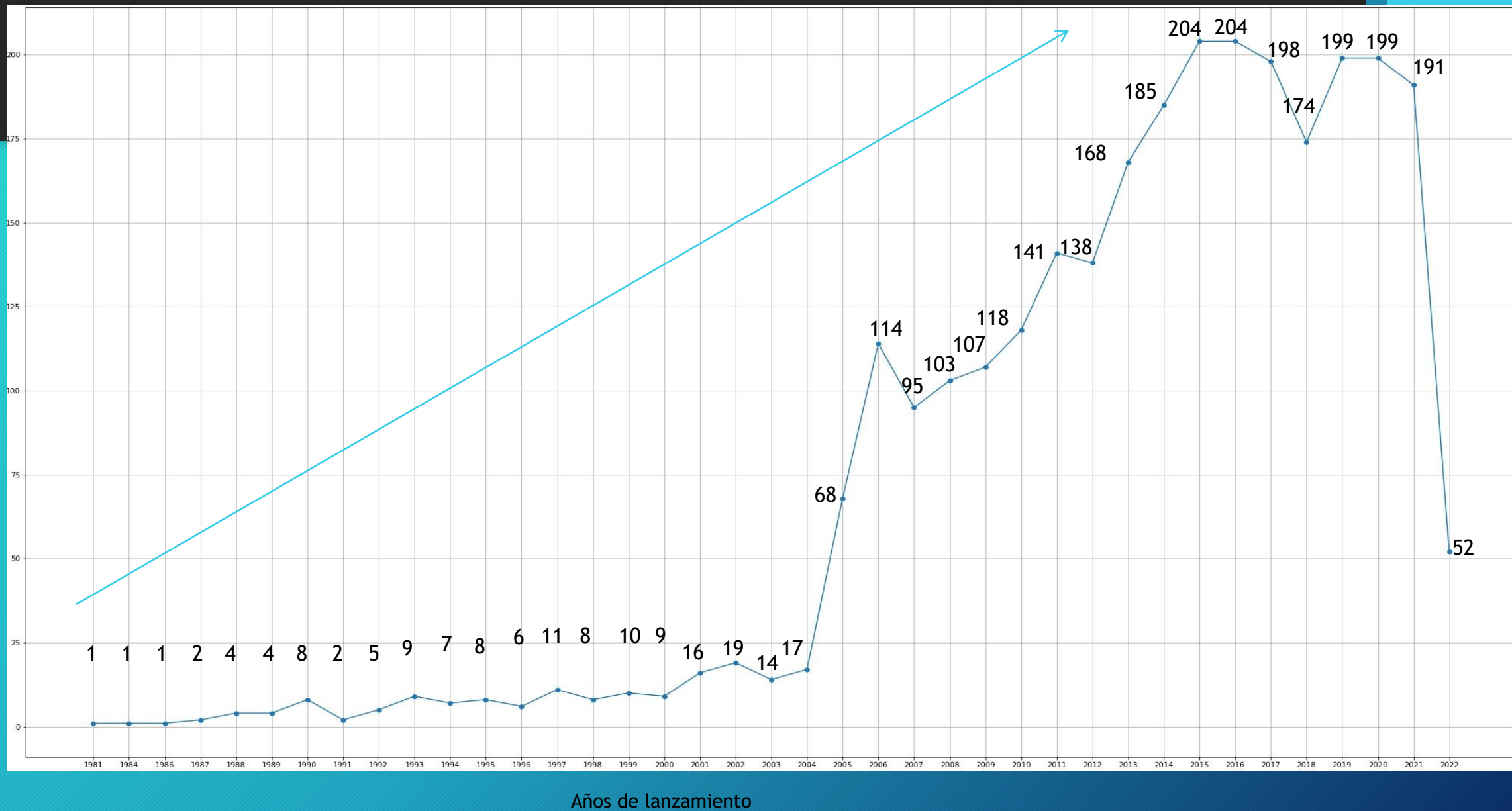
```
[7]: df.describe()#Se utiliza para ver algunos detalles estadísticos básicos como percentil, media, std, etc. de un marco de datos o una serie de valores numéricos.  
#nos dice que count es 2820 el cual es la cantidad total de filas  
#mean es el promedio de metascor (puntuaje del 1 al 99) en este caso es 65,1  
#std desviación estándar de 15  
#min el metascor mas bajo pertenece a 13  
#el25% pertenece al metascor 55  
#el 50% se encuentra en el metascor 66  
#el 75% se encuentra en el metascor 76  
#y el máximo metascor es de 99.
```

```
[7]:
```

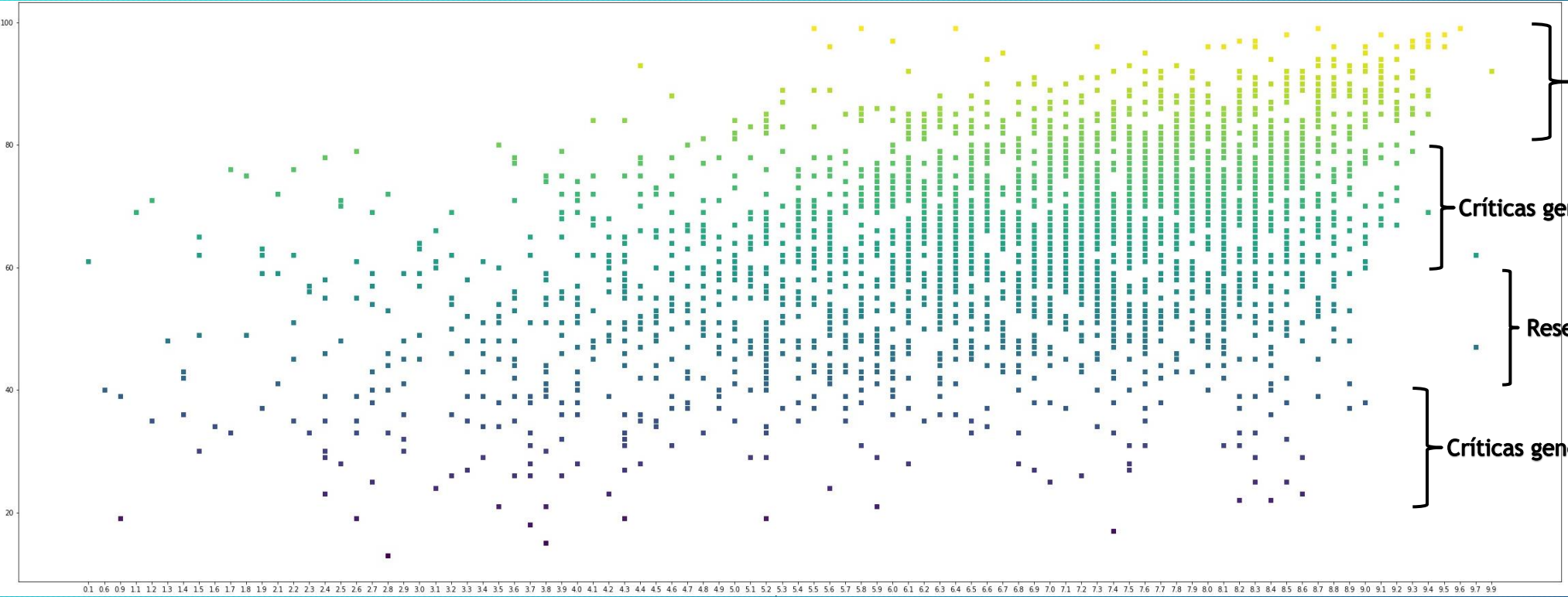
	metascor
count	2820.00000
significar	65.15922
ets	15.03053
min	13.00000
25%	55.00000
50%	66.00000
75%	76.00000
máximo	99.00000

Cantidad de votaciones que han hecho los usuarios por años

Cantidad de votaciones que han hecho los usuarios

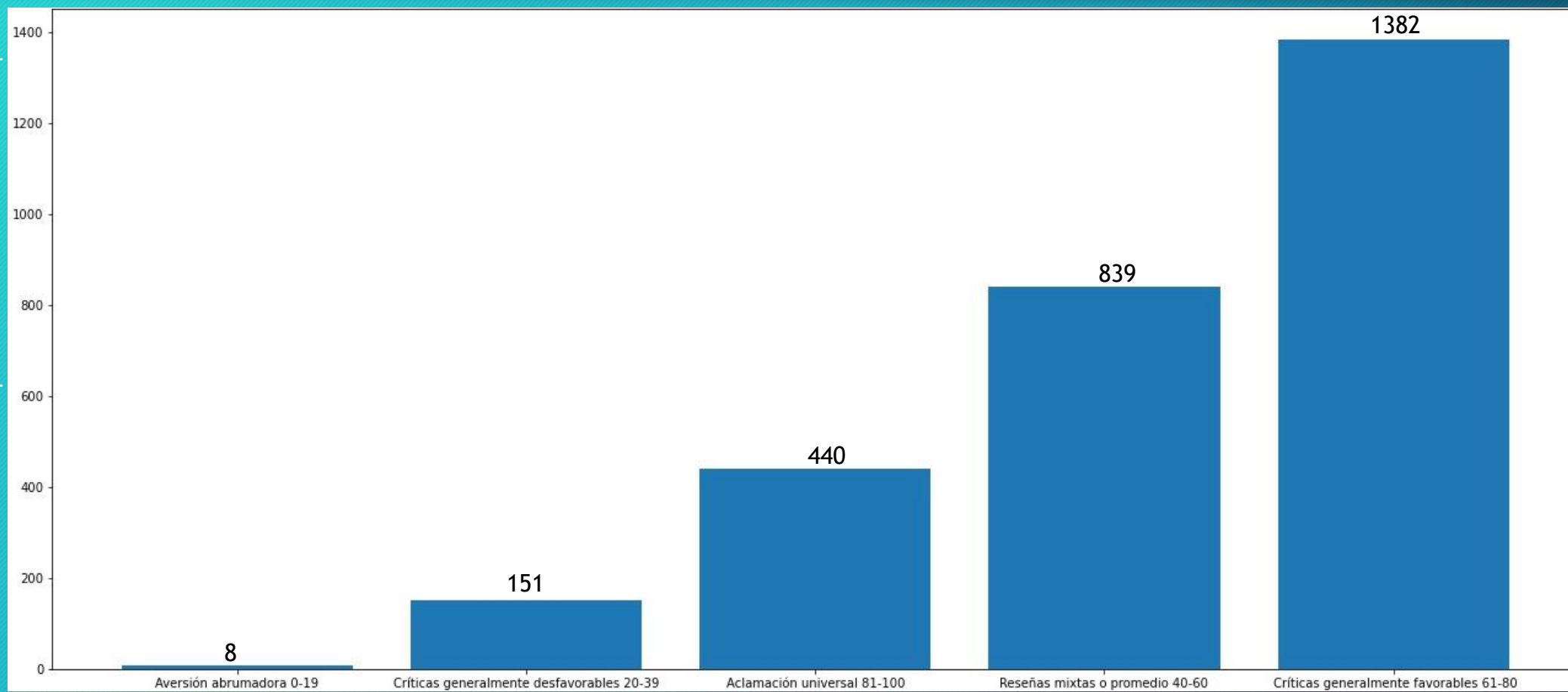


Muestra la relación entre la puntuación de los usuarios y el metascoring



Puntuación de usuarios

Muestra la cantidad de métricas total por categoría



Indicadores metascoring