

# Proyecto final de Análisis de datos

Integrantes: Chumania Nestor, Daniel Quishpe, Joel Tate, Dany Vinueza.

**Resumen**— El análisis de datos es el proceso de examinar, limpiar y transformar datos para obtener información útil y significativa. Este proceso implica aplicar técnicas estadísticas y matemáticas para identificar patrones, tendencias y relaciones entre variables en un conjunto de datos. El análisis de datos se utiliza en una amplia gama de campos, desde la investigación científica hasta el marketing y los negocios, y puede ayudar a tomar decisiones informadas y optimizar procesos y resultados.

**Abstract**—Data analysis is the process of examining, cleaning, and transforming data to obtain useful and meaningful information. This process involves applying statistical and mathematical techniques to identify patterns, trends, and relationships between variables in a data set. Data analytics is used in a wide range of fields, from scientific research to marketing and business, and can help make informed decisions and optimize processes and results.

## Introducción

Los temas a analizar son los siguientes:

- Estado de covid a nivel mundial Web Scrapping.
- Top ventas en Amazon.
- Lista de millonarios a nivel mundial CSV.
- Top Streamers Twitch Kaggle.
- Top musicas en Spotify.



figura 1 Datos

## I. DEFINICIÓN DEL CASO COVID

Estado de covid a nivel mundial en base a esto nos referimos a organizar o desplegar información por países y número de casos que existe, los nuevos contagios y el total de muertes que existen hasta la actualidad, total de casos recuperados y casos activos Figura 2.

Unnamed: 0	#	Country, Other	Total Cases	New Cases	Total Deaths	\
0	0	NaN North America	124713684	1375.0	1612527.0	
1	1	NaN Asia	214474070	34365.0	1537319.0	
2	2	NaN Europe	246559618	28857.0	2019348.0	
3	3	NaN South America	68045724	3031.0	1350925.0	
4	4	NaN Oceania	14015805	NaN	26203.0	

New Deaths	Total Recovered	New Recovered	Active Cases	...	Total Tests	\
0	5.0	120000633.0	4075.0	3100524.0	...	NaN
1	172.0	199108565.0	53192.0	13828186.0	...	NaN
2	65.0	242096217.0	28694.0	2444053.0	...	NaN
3	19.0	66293552.0	90.0	401247.0	...	NaN
4	NaN	13852354.0	NaN	137248.0	...	NaN

Figura 2 Covid

## II. DEFINICIÓN DEL CASO LISTA DE MILLONARIOS A NIVEL MUNDIAL

La lista de millonarios a nivel mundial es una fuente útil para entender la distribución de la riqueza y las tendencias económicas globales. La lista destaca la creciente brecha entre los ricos y los pobres, así como la concentración de la riqueza en manos de unas pocas personas.

Esta información esta filtrada por nombre, ranking, año, fundación de empresa, nombre de la empresa etc. Por la cual será de fácil entender para poder continuar con la definición Figura 3.

	name	rank	year	company.founded	company.name	company.relationship	company.sector	company.type	demographics.age
0	Bill Gates	1	1996	1975	Microsoft	founder	Software	new	40
1	Bill Gates	1	2001	1975	Microsoft	founder	Software	new	45
2	Bill Gates	1	2014	1975	Microsoft	founder	Software	new	58

Figura 3 Millonarios

## III. DEFINICIÓN DEL CASO TOP VENTAS AMAZON

El análisis de las ventas de los productos más populares en Amazon puede ser una herramienta útil para comprender las tendencias del mercado y los hábitos de compra de los consumidores. La lista de los productos más vendidos en Amazon puede cambiar rápidamente debido a factores como la disponibilidad de productos, las promociones, la temporada y la demanda de los consumidores Figura 4.

	Title	Rating
0	TECLAST Tablet de 10 pulgadas Android 12 Table...	4.5 de 5 estrellas
1	Tablet Android, tablet Android 12 de 10.1 pulg...	5.0 de 5 estrellas
2	Tablet Android de 10 pulgadas, tablet Android ...	3.7 de 5 estrellas
3	TECLAST Android 12 Tablet 10 inch Tablets, M40...	4.4 de 5 estrellas
4	KYASTER Tablet para niños, 7 pulgadas, 5G, WiF...	3.6 de 5 estrellas

**Figura 4 Amazon**

#### IV. DEFINICIÓN DEL CASO DE TOP STREAMERS TWICH

Es una forma útil de entender las tendencias y preferencias de los espectadores en cuanto a la transmisión de videojuegos en línea. Aunque la lista puede cambiar con el tiempo debido a la competencia y la evolución del mercado de los juegos en línea Figura 5.

(11000, 11)	Channel	Watch time(Minutes)	Stream time(minutes)	Peak viewers
0	xQcOW	6196161750	215250	222720
1	summit1g	6091677300	211845	310998
2	Gaules	5644590915	515280	387315
3	ESL_CSGO	3970318140	517740	300575
4	Tfue	3671000070	123660	285644

	Average viewers	Followers	Followers gained	Views gained	Partnered
0	27716	3246298	1734810	93036735	True
1	25610	5310163	1370184	89705964	True
2	10976	1767635	1023779	102611607	True
3	7714	3944850	703986	106546942	True
4	29602	8938903	2068424	78998587	True

**Figura 5 Streamers**

#### V. DECISIÓN CASO DE TOP MUSICAS EN SPOTIFY

Los Útil para comprender las tendencias musicales de la actualidad en la cual nos sirve para entender cuál es el ritmo más escuchado por la sociedad y el artista con su música más destacada del momento. Esta lista cambia constantemente por la popularidad de artistas, factores de discografías, genero, categoría y dependiendo su región Figura 6.

	Name	Artists	Album
0	TQG	KAROL G,Shakira	MAÑANA SERÁ BONITO
1	Die For You - Remix	The Weeknd,Ariana Grande	Die For You (Remix)
2	Flowers	Miley Cyrus	Flowers
3	Kill Bill	SZA	SOS
4	Boy's a liar Pt. 2	PinkPantheress,Ice Spice	Boy's a liar Pt. 2

**Figura 6 Spotify**

#### VI. OBJETIVOS GENERALES Y ESPECÍFICOS

- Comprender la situación del mundo en la actualidad en diferentes tipos de casos o temas a través de estadísticas concretas del mundo entero.
- Probar y sacar nuestra teoría sobre los casos estudiados con el fin de conocer las interacciones con el mundo en general.
- Identificar tendencias y preferencias de las personas a nivel mundial.
- Analizar el impacto en la sociedad de los diferentes casos a tratar.
- Comparar temas tratados en diferentes países del mundo para comprender su estado sobre salud, música, capital y las redes sociales.

#### VII. DESCRIPCIÓN DEL EQUIPO DE TRABAJO

Nestor Chumania: Ventas y Millonarios a nivel mundial extracción y guardar datos csv.

Daniel Quishpe: Top Streamers y casos covid extracción datos y guardar archivo csv.

Dany Vinueza: Top musicas Spotify e Informe, video.

Joel Tates: Visualizaciones, Conexión, Importaciones, Dashboard y video.

#### VIII. - RECURSO Y HERRAMIENTAS UTILIZADAS

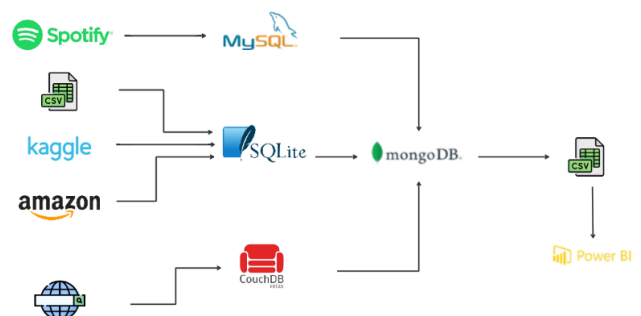
- Spotify: Tener una cuenta en Spotify o conseguir una cuenta abierta.
- Kaggle: Para obtener datos desde esta web es necesario descargar con nuestras credenciales.
- Amazon: Extracción de datos a través de iniciar sesión con tu cuenta.
- Casos covid: Utilización del web scripting para extracción de datos de la web de diferentes sitios.
- Millonarios a nivel mundial: extracción de datos a través de un archive CSV desde la web.

Herramientas:

- MySQL: Optimización de consultas y de almacenamiento de datos.
- SQL: Se ejecuta como una biblioteca de enlace dinámico dentro de la aplicación cliente, lo que significa que no requiere un servidor separado.
- CouchBD:Base de datos de documentos NoSQL en la cual nos almacena los archivos json en lugar de tablas relacionales como una base de datos SQL tradicional.
- MongoDB: Base de datos NoSQL de documentos, que utiliza un formato de documento JSON BSON.
- Power BI:Visualizacion y análisis de datos en tiempo real de una interfaz de usuario intuitiva.Conexion a las diferentes bases de datos para la extracción de datos y presentación de los mismos.

#### IX. ARQUITECTURA DE LA SOLUCIÓN

Arquitectura utilizada en el Proyecto Figura 7.



**Figura 7 Arquitectura utilizada**

#### X. EXTRACCIÓN DE DATOS

Extracción de datos de las diferentes plataformas de la web se necesitaron varios pasos como diferentes credenciales y

librerías dentro del editor de código a utilizar en nuestro caso es Jupyter notebook Python.

- Importación de librerías a utilizar Figura 8.

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import missingno as msno
```

**Figura 8 Librerías**

- Extracción de datos desde un archivo CSV de los casos de COVID Figura 9.

```
: data = pd.read_csv('./data_original/covid_table.csv')
print('Original data')
print((data.size, 4))
print(data.head(5))
```

**Figura 9 Covid datos**

- Extracción de datos de los productos mas vendidos en amazon en Tablet **Figura 10.**

```
: data = pd.read_csv('./data_original/catalogo_tablets.csv')
print('Original data')
print((data.size, 4))
print(data.head(5))
```

**Figura 10 Amazon datos**

- Extracción de datos Top Streamers Twitch Kaggle Figura 11.

- En este caso vamos a trabajar con datos de kaggle así que lo descargamos de kaggle, con nuestros credenciales

```
dataset = 'https://www.kaggle.com/brendan45774/countries-life-expectancy'
od.download(dataset)
```

```
data = pd.read_csv('./data_original/twitchdata-update.csv')
print('Original data')
print((data.size, 11))
print(data.head(5))
```

**Figura 11 Datos Streamers**

- Extracción datos Top musicas Spotify Figura 12 .

```
data = pd.read_csv('./datos_originales/top_songs_spotify.csv')
print('Original data')
print((data.size, 4))
print(data.head(5))
```

**Figura 12 Datos Spotify**

## XI. ANÁLISIS DE INFORMACIÓN

- Casos de covid Figura 13.

El análisis de los casos de COVID-19 es una de las principales preocupaciones a nivel mundial debido a la propagación continua de la enfermedad. Los datos son recopilados por las autoridades sanitarias de cada país y organizaciones como la Organización Mundial de la Salud (OMS), y se analizan para entender mejor la propagación de la enfermedad y cómo se puede prevenir su propagación.

En los siguientes indicadores Podemos definir que:  
Total, cases: personas que dieron positivo a una prueba covid.

Total Debates: Personas fallecidas por casos de covid.

Total Recovered: Personas recuperadas después de dar positivo a la prueba covid.

Active cases: Casos nuevos o personas con covid con peligrosidad de contagio.

Serious Critical: Personas con dificultades de salud como dificultades al respirar.

Top cases 1m pop: Casos de covid que superaron el millón en los diferentes países.

Country, Other	Total Cases	Total Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop	Deaths/ 1M pop	Total Tests
USA	105393864	1146728.0	102848904.0	1398232.0	2726.0	314792.0	3425.0	1.167633e+09
France	39638159	165030.0	39399753.0	73376.0	869.0	604383.0	2516.0	2.714902e+08
Japan	33251497	72721.0	21692695.0	11486081.0	131.0	264773.0	579.0	9.558908e+07
S. Korea	30555102	34020.0	30350687.0	170395.0	143.0	595269.0	663.0	1.580406e+07
Vietnam	11526962	43186.0	10614788.0	868988.0	3.0	116489.0	436.0	8.582655e+07
Poland	6431888	118952.0	5335940.0	976996.0	696.0	170427.0	3152.0	3.830693e+07
Austria	5936666	21921.0	5854585.0	60160.0	72.0	654776.0	2418.0	2.112735e+08
Chile	5177770	64222.0	5101501.0	12047.0	83.0	268972.0	3336.0	4.850162e+07
Malaysia	5043626	36965.0	4997422.0	9239.0	4.0	152003.0	1114.0	6.825567e+07
Israel	4801230	12303.0	4781252.0	7675.0	57.0	514822.0	1319.0	4.137336e+07
Czechia	4613882	42464.0	4561632.0	9786.0	45.0	429727.0	3955.0	5.701771e+07
Peru	4486665	219485.0	4265173.0	2007.0	34.0	133198.0	6516.0	3.807850e+07

**Figura 13 Análisis casos covid**

Se puede determinar el virus covid-19 no ha terminado por su prolongación rápida, el impacto que tiene cada país con nuevos contagios, y el alza o disminución de casos en diferentes países. Gracias a estos datos es esencial reconocer que cada país busca la manera de sobresalir de estos contagios que se les presentan, con medidas de bioseguridad y poder controlarlos a tiempo.

- Análisis lista de millonarios a nivel mundial.

Revelación de patrones y tendencias de la distribución de la riqueza en grupos pequeños de personas o empresas con alto prestigio dentro del país.

	name	rank	year	company.founded	company.name	company.relationship	company.sector	com
0	Bill Gates	1	1996	1975	Microsoft	founder	Software	
1	Bill Gates	1	2001	1975	Microsoft	founder	Software	
2	Bill Gates	1	2014	1975	Microsoft	founder	Software	
3	Warren Buffett	2	1996	1962	Berkshire Hathaway	founder	Finance	
4	Warren Buffett	2	2001	1962	Berkshire Hathaway	founder	Finance	

- Análisis de top ventas Amazon

El top de ventas de Tablet en Amazon podemos analizar que los sistemas operativos son las tablets Android. Los consumidores confían más en este sistema operativo antes que otro. Además de los tamaños de pantallas los mas populares son de 10 pulgadas y 12 pulgadas para hacer el trabajo en ellas más fácil. Tiene una buena calificación y un buen precio para el consumidor Figura 14.

	Link \	
0	<a href="https://m.media-amazon.com/images/I/71t+pN7w6H...">https://m.media-amazon.com/images/I/71t+pN7w6H...</a>	
1	<a href="https://m.media-amazon.com/images/I/81TtgFNjRg...">https://m.media-amazon.com/images/I/81TtgFNjRg...</a>	
2	<a href="https://m.media-amazon.com/images/I/81huX80Xrc...">https://m.media-amazon.com/images/I/81huX80Xrc...</a>	
3	<a href="https://m.media-amazon.com/images/I/71s1mbLNa...">https://m.media-amazon.com/images/I/71s1mbLNa...</a>	
4	<a href="https://m.media-amazon.com/images/I/71iq-Tv0l...">https://m.media-amazon.com/images/I/71iq-Tv0l...</a>	
	Title	Rating \
0	TECLAST Tablet de 10 pulgadas Android 12 Table...	4.5 de 5 estrellas
1	Tablet Android, tablet Android 12 de 10.1 pulg...	5.0 de 5 estrellas
2	Tablet Android de 10 pulgadas, tablet Android ...	3.7 de 5 estrellas
3	TECLAST Android 12 Tablet 10 inch Tablets, M40...	4.4 de 5 estrellas
4	KYASTER Tablet para niños, 7 pulgadas, 5G, WiF...	3.6 de 5 estrellas
	Price	
0	84.0	
1	89.0	
2	82.0	
3	143.0	
4	59.0	

Figura 14 Análisis Amazon

- Análisis top musicas en Spotify

La popularidad del género que está en tendencia hoy día en la actualidad, el artista del momento y el álbum del momento además de que su fecha de lanzamiento ya fue hace varios días atrás. Estas estadísticas nos ayudan a entender la popularidad del género, artista más popular, las colaboraciones, el éxito global que de muchas de estas canciones sugiere que la música es medio para unir a las personas de diferentes países y culturas Figura 15.

Original data (200, 4)			
	Name	Artists	Album \
0	TQG	KAROL G, Shakira	MAÑANA SERÁ BONITO
1	Die For You - Remix	The Weeknd, Ariana Grande	Die For You (Remix)
2	Flowers	Miley Cyrus	Flowers
3	Kill Bill	SZA	SOS
4	Boy's a liar Pt. 2	PinkPantheress, Ice Spice	Boy's a liar Pt. 2
Release Date			
0	2023-02-24		
1	2023-02-24		
2	2023-01-13		
3	2022-12-08		
4	2023-02-03		

Figura 15 Análisis Spotify

- Análisis Top Streamers Twitch

Los streamers de Twitch por canal, tiempo de transmisión y vistas muestra que la popularidad del juego es un factor importante que afecta el número de vistas y la duración de las transmisiones. Cada streamer tiene una audiencia objetivo específica y la duración de la transmisión varía según el streamer y la popularidad del juego. Además, la plataforma de streaming más popular entre los streamers es Twitch, aunque algunos también utilizan otras plataformas. Por lo cual podemos concluir que cada streamer tiene sus maneras de hacer entretener al publico y poder ganar mas audiencia con sus formas de videojuego y dependiendo de la plataforma utilizada Figura 16.

	Channel	Watch time(Minutes)	Stream time(minutes)	Peak viewers	Average viewers	Followers	Followers gained	Views gained	Partnered	Mature	Language
0	xQcOW	6196181750	215250	222720	27716	3246298	1734810.0	93036735.0	True	False	English
1	summit1g	6091677300	211845	310998	25610	5310163	1370184.0	88705964.0	True	False	English
2	Gaules	5644590915	515280	387315	10976	1767635	1023779.0	102611807.0	True	True	Portuguese
3	ESL_CSGO	3970318140	517740	300575	7714	3944850	703986.0	106549942.0	True	False	English
4	Tfue	3671000070	123660	265644	29602	8938903	2068424.0	78998587.0	True	False	English
5	Azmongold	3668799075	82260	263720	42414	1563438	554201.0	61715781.0	True	False	English
6	NICKMERCS	3386075195	136275	115633	24181	4074287	1088624.0	46084211.0	True	False	English
7	Foxtraite	3301867485	147885	88795	10985	508816	425466.0	670137548.0	True	False	English
8	loky1er1	2928356940	122490	89367	22381	3530767	951730.0	51348926.0	True	False	English
9	Anomaly	2865429915	92880	125408	12377	2607076	1532689.0	36350662.0	True	False	English
10	TenTheTatman	2834436990	108780	142067	25664	5265659	1244341.0	50119786.0	True	True	English
11	LIRIK	2832930285	128490	89170	21739	2666382	199077.0	50504526.0	True	False	English

Figura 16 Análisis Twitch

## XII. VISUALIZACIÓN DE INFORMACIÓN

- Visualización de los casos de covid

Unnamed: 0	#	Country, Other	Total Cases	Total Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop	Deaths/ 1M pop	Total Tests	Tests/ 1M pop	Population	
8	8	1.0	USA	105393864	1146728.0	102848904.0	1398232.0	2726.0	314792.0	3425.0	1.167633e+09	3487500.0	334805269.0
10	10	3.0	France	39638159	165030.0	39399753.0	73376.0	869.0	604383.0	2516.0	2.714902e+08	4139547.0	65584518.0
13	13	6.0	Japan	33251497	72721.0	21692695.0	11486081.0	131.0	264773.0	579.0	9.558908e+07	761151.0	125584838.0
14	14	7.0	S. Korea	30555102	34020.0	30350687.0	170395.0	143.0	595269.0	663.0	1.580406e+07	307892.0	51329899.0
20	20	13.0	Vietnam	11526962	43186.0	10614788.0	868988.0	3.0	116489.0	436.0	8.582655e+07	867342.0	98953541.0
28	28	21.0	Poland	6431888	118952.0	5335940.0	976996.0	696.0	170427.0	3152.0	3.830693e+07	1015028.0	37739785.0
30	30	23.0	Austria	5936666	21921.0	5854585.0	60160.0	72.0	654776.0	2418.0	2.112735e+08	23302116.0	9066710.0
34	34	27.0	Chile	5177770	64222.0	5101501.0	12047.0	83.0	268972.0	3336.0	4.850162e+07	2519539.0	19250195.0
35	35	28.0	Malaysia	5043626	36965.0	4997422.0	9239.0	4.0	152003.0	1114.0	6.825567e+07	2057067.0	33181072.0
36	36	29.0	Israel	4801230	12303.0	4781252.0	7675.0	57.0	514822.0	1319.0	4.137336e+07	4436346.0	9326000.0
40	40	33.0	Czechia	4613882	42464.0	4561632.0	9786.0	45.0	429727.0	3955.0	5.701771e+07	5310502.0	10736784.0
42	42	35.0	Peru	4486665	219485.0	4265173.0	2007.0	34.0	133198.0	6516.0	3.807850e+07	1130455.0	33684208.0
44	44	37.0	Philippines	4076866	66143.0	4001665.0	9058.0	162.0	36236.0	588.0	3.462961e+07	307794.0	112508994.0
47	47	40.0	Denmark	3175474	8277.0	3166219.0	978.0	19.0	544216.0	1419.0	1.292018e+08	22142752.0	5834950.0
48	48	41.0	Hong Kong	2883655	13462.0	2855882.0	14311.0	28.0	379214.0	1770.0	7.612772e+07	10011143.0	7604299.0
50	50	43.0	Serbia	2495808	17847.0	2463552.0	14409.0	10.0	288432.0	2063.0	1.242815e+07	1436279.0	8653016.0
56	56	49.0	Slovakia	1863377	21030.0	1841083.0	1264.0	20.0	341266.0	3852.0	7.405734e+06	1356314.0	5460193.0
60	60	53.0	Pakistan	1577119	30643.0	1538689.0	7787.0	7.0	6872.0	134.0	3.057086e+07	133213.0	229488994.0
64	64	57.0	Slovenia	1329775	7074.0	1318364.0	4337.0	8.0	639920.0	3404.0	2.839092e+06	1366239.0	2078034.0
65	65	58.0	Lithuania	1306273	9594.0	1290412.0	6267.0	7.0	490765.0	3604.0	1.038185e+07	3900446.0	2661708.0
67	67	60.0	Morocco	1272486	16296.0	1256124.0	66.0	3.0	33688.0	431.0	1.298991e+07	343896.0	37772756.0
74	74	67.0	Cuba	1112622	8530.0	1104004.0	88.0	2.0	98413.0	754.0	1.431960e+07	1266588.0	11305652.0

- Visualización datos Top ventas en Amazon.

(900, 4)	Link \	Title	Rating \
0	<a href="https://m.media-amazon.com/images/I/71t+pN7w6H...">https://m.media-amazon.com/images/I/71t+pN7w6H...</a>	TECLAST Tablet de 10 pulgadas Android 12 Table...	4.5 de 5 estrellas
1	<a href="https://m.media-amazon.com/images/I/81TtgFNJrG...">https://m.media-amazon.com/images/I/81TtgFNJrG...</a>	Tablet Android, tablet Android 12 de 10.1 pulg...	5.0 de 5 estrellas
2	<a href="https://m.media-amazon.com/images/I/81huX80Xrc...">https://m.media-amazon.com/images/I/81huX80Xrc...</a>	Tablet Android de 10 pulgadas, tablet Android ...	3.7 de 5 estrellas
3	<a href="https://m.media-amazon.com/images/I/71s1mlbLNa...">https://m.media-amazon.com/images/I/71s1mlbLNa...</a>	TECLAST Android 12 Tablet 10 inch Tablets, M40...	4.4 de 5 estrellas
4	<a href="https://m.media-amazon.com/images/I/71iq-TvOl-...">https://m.media-amazon.com/images/I/71iq-TvOl-...</a>	KYASTER Tablet para niños, 7 pulgadas, 5G, WiF...	3.6 de 5 estrellas
Price			
0	84.0		
1	89.0		
2	82.0		
3	143.0		
4	59.0		

- Visualización datos Lista de millonarios a nivel mundial CSV.



name	rank	year	company.founded	company.name	company.relationship	company.sector	company.type	demographics.age	demographics.gender	...	location.country code
Bill Gates	1	1996	1975	Microsoft	founder	Software	new	40	male	...	USA
Bill Gates	1	2001	1975	Microsoft	founder	Software	new	45	male	...	USA
Bill Gates	1	2014	1975	Microsoft	founder	Software	new	58	male	...	USA
Warren Buffett	2	1996	1962	Berkshire Hathaway	founder	Finance	new	65	male	...	USA
Warren Buffett	2	2001	1962	Berkshire Hathaway	founder	Finance	new	70	male	...	USA
Carlos Slim Helu	2	2014	1990	Telmex	founder	Communications	privatization	74	male	...	MEX

- Visualización datos Top Streamers Twitch Kaggle.

	Channel	Watch time(Minutes)	Stream time(minutes)	Peak viewers	Average viewers	Followers	Followers gained	Views gained	Partnered	Mature	Language
0	xQcOW	6196161750	215250	222720	27716	3246298	1734810.0	93036735.0	True	False	English
1	summit1g	6091677300	211845	310998	25610	5310163	1370184.0	89705964.0	True	False	English
2	Gaules	5644590915	515280	387315	10976	1767635	1023779.0	102611607.0	True	True	Portuguese
3	ESL_CSGO	3970318140	517740	300575	7714	3944850	703986.0	106546942.0	True	False	English
4	Tfue	3671000070	123660	285644	29602	8938903	2068424.0	78998587.0	True	False	English
5	Asmongold	3668799075	82260	263720	42414	1563438	554201.0	61715781.0	True	False	English
6	NICKMERCS	3360675195	136275	115633	24181	4074287	1089824.0	46084211.0	True	False	English
7	Fextralife	3301867485	147885	68795	18985	508816	425468.0	670137548.0	True	False	English
8	lolyer1	2928356940	122490	89387	22381	3530767	951730.0	51349926.0	True	False	English
9	Anomaly	2865429915	92880	125408	12377	2607076	1532689.0	36350662.0	True	False	English
10	TimTheTatman	2834436990	108780	142067	25664	5265659	1244341.0	50119786.0	True	True	English
11	LIRIK	2832930285	128490	89170	21739	2666382	199077.0	50504526.0	True	False	English
12	Riot Games (riotgames)	2674646715	80820	639375	20960	4487489	497678.0	56855694.0	True	False	English
13	Rubius	2588632635	58275	240096	42948	5751354	3820532.0	58599449.0	True	False	Spanish
14	auronplay	2410022550	40575	170115	53986	3983847	3966525.0	41514854.0	True	False	Spanish
15	MontanaBlack88	2408460990	67740	181600	33514	2911316	1101093.0	37189666.0	True	True	German
16	sodapoppin	2329440420	115305	107833	19659	2786162	236169.0	39334821.0	True	True	English
17	동월랑 (hanryang1125)	2186662470	181230	26999	12201	494445	92205.0	34405975.0	True	False	Korean
18	alanzoka	2055003870	103770	89153	19560	3445134	1325075.0	46515698.0	True	False	Portuguese
19	CohhCarnage	2029212570	175230	43615	11343	1264808	124242.0	38718674.0	True	False	English
20	Lord_Kebun	1943299035	153720	34830	12367	434200	137215.0	19645967.0	True	False	English

- Visualización datos Top musicas en Spotify.

Original data  
(200, 4)

	Name	Artists	Album \
0	TQG	KAROL G,Shakira	MAÑANA SERÁ BONITO
1	Die For You - Remix	The Weeknd,Ariana Grande	Die For You (Remix)
2	Flowers	Miley Cyrus	Flowers
3	Kill Bill	SZA	SOS
4	Boy's a liar Pt. 2	PinkPantheress,Ice Spice	Boy's a liar Pt. 2

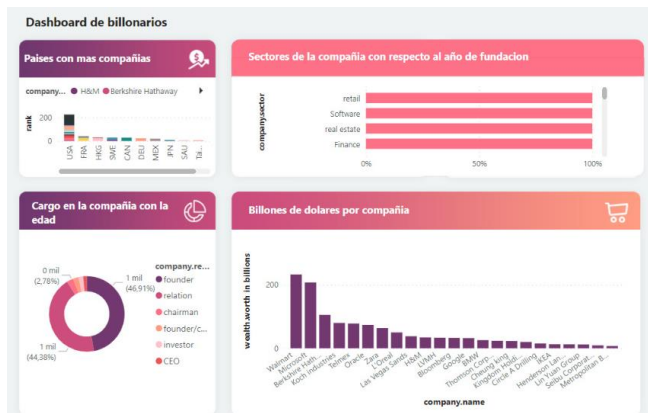
	Release Date
0	2023-02-24
1	2023-02-24
2	2023-01-13
3	2022-12-08
4	2023-02-03

### XIII. – DASHBOARD

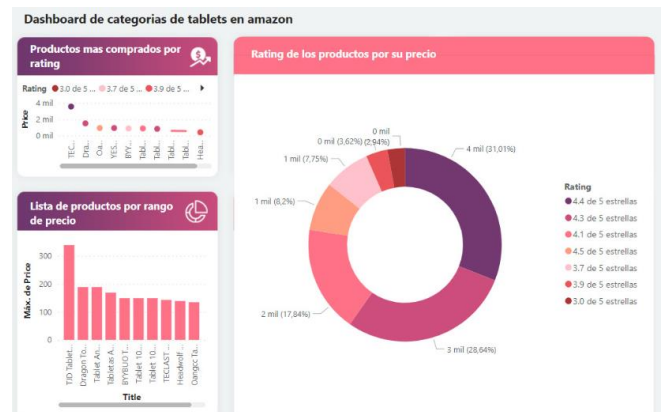
- Dashboard de casos de covid



- Dashboard de millonarios



- Dashboard de top ventas Amazon



- Dashboard de top musicas en Spotify

**Dashboard de spotify**

**Top 50 canciones de spotify a nivel mundial**

Artists	Name	Album
AgroPlay, Ana Castela	Nosso Quadro	Nosso Quadro
Arcaangel, Bad Bunny	La Jumpa	Sr. Santos
Arctic Monkeys	I Wanna Be Yours	AM
Bad Bunny, Chencho Corleone	Me Porto Bonito	Un Verano Sin Ti
Bizarrap, Quevedo	Quevedo: Bzrp Music Sessions, Vol. 52	Quevedo: Bzrp Music Sessions, Vol. 52
Bizarrap, Shakira	Shakira: Bzrp Music Sessions, Vol. 53	Shakira: Bzrp Music Sessions, Vol. 53
Coi Leray	Players	Players
dávid	Here With Me	Here With Me
David Guetta, Bebe Rexha	I'm Good (Blue)	I'm Good (Blue)
Drake, 21 Savage	Rich Flex	Her Loss
Eminem	Mockingbird	Encore (Deluxe Version)
Feid	CHORRITO PA LAS ANIMAS	SIXDO
Feid	Feliz Cumpleaños Feroxo	FELIZ CUMPLEAÑOS FERROXO TE PIRATEAMOS EL A
Fuerza Regida, Grupo Frontera	Bebe Dame	Sigan Hablando
Gonizaz, Bad Bunny	Tormenta (feat. Bad Bunny)	Cracker Island
Harry Styles	As It Was	Harry's House
Imagine Dragons	Bones	Mercury - Acts 1 & 2
J-hope, J. Cole	on the street (with J. Cole)	this is what ____ feels like (Vol. 1-4)
JVKE	golden hour	MAÑANA SERÁ BONITO
KAROL G, Romeo Santos	X SI VOLVEMOS	MAÑANA SERÁ BONITO
KAROL G, Shakira	TQG	MAÑANA SERÁ BONITO

- Dashboard de top streamers en twitch.

### XIV. IMPORTACIONES DE DATOS.

```

In [1]: from BDDs.Couch import Introducir_csv
        Introducir_csv('./Covid/data_clean/covid.csv')

In [2]: from BDDs.MySQL import Introducir_mysql
        Introducir_mysql('./Musicas/datos_limpios/songs_spotify.csv')

In [3]: from BDDs.SQLite import *
        createDB()
        createTablecantantes()
        createTablemillonarios()
        createTablestreamers()

In [2]: Introducir_streamers('./Streamers/data_clean/topstreamers.csv')

In [3]: Introducir_ventas('./Ventas/data_clean/amazon.csv')

In [4]: Introducir_billonarios('./Millonarios/data_clean/billonarios.csv')

In [1]: from BDDs.Mongo import ejecutar_importaciones
        ejecutar_importaciones('./Datos.db')
  
```

### XV. RESULTADOS OBTENIDOS

La aplicación de análisis de datos a través de datos abiertos o plataformas que requieren autorización al momento de realizar las respectivas estadísticas podemos conocer varios datos sorprendentes o datos importantes ya que si se aplica correctamente podemos conocer varios parámetros del mundo que nos rodea como es: como está el virus despues de la pandemia, los artículos con mayor demanda dentro de una plataforma de compra, personas con alto capital dentro del mundo entero, música, género y artista preferido por las

personas a nivel mundial y por ultimo tenemos a los streamer que no son menos importantes sino que también es un dato importante con el cual nos damos cuenta a que está orientado nuestro entretenimiento ya sea con videojuegos, interacciones por plataformas o simplemente viendo videos.

#### XVI. CONCLUSIONES Y RECOMENDACIONES

Cada análisis de los temas a tratados tiene su propio objetivo y sus propias conclusiones, pero todos tienen el mismo fin que es buscar estadísticas concretas del tema a tratar y sacar conclusiones.

- En la tabla de casos de covid podemos concluir el gran impacto que tuvo en diferentes países y regiones en los últimos años.
- En la tabla de millonarios del mundo entendemos la distribución de la riqueza del mundo, las empresas y los países más pobres o lucrativos. es decir la desigualdad económica y los factores a la creación y acumulación de riqueza de cada país.
- En la tabla de Streamer en Twitch la oportunidad de darse a conocer en una plataforma muy usada por las personas hoy en día en la cual puede tener éxito o fracaso dentro de la misma.
- En la tabla de Top música Spotify la comprensión de la industria musical a través de la plataforma y la oportunidad de dar al artista una mejor trayectoria con sus nuevos temas o álbumes.
- En la tabla de Top ventas Amazon la estrategia de venta de diferentes productos que se adaptan a las necesidades de los consumidores y así adaptar el precio la oferta y la demanda del producto.

Como punto final se puede decir que gracias a estos datos que nos brindan podemos saber un poco más del mundo que nos rodea con el único fin de mantenernos informados de lo que pasa a nuestros alrededores y en todos los rincones del mundo.

#### XVII. - DESAFÍOS Y PROBLEMAS ENCONTRADOS

- Extracción de datos de las diferentes plataformas y exportarlos a las diferentes BD NoSQL y SQL.
- Establecer conexión con las BD.
- Al momento de exportar los datos que se exporten vacíos.
- Limpiar los datos ya que contenían valores nulos.

#### XVIII. ANEXOS

##### **Link del github:**

<https://github.com/ItsgabrielJT/AnalisisProyectos.git>