

Análisis de comentarios en redes sociales

Realizado por:

Ing. Jorge Sebastian Cortés Mejía



Introducción

El presente proyecto tiene como objetivo procesar y presentar la información contenida en un dataset que compila comentarios de diversas redes sociales, utilizando consultas SQL. El propósito es transformar los datos en crudo en información comprensible y valiosa, revelando insights y patrones que no son evidentes a simple vista. Este procesamiento incluirá limpieza de datos, análisis exploratorio y visualización de datos, para extraer y comunicar información significativa que permita una comprensión más profunda del comportamiento y las interacciones de los usuarios en las redes sociales.



¿Cuál es el objetivo de este estudio?

Obtener información relevante y útil para responder a preguntas tales como :

- ¿Cuántos usuarios hay por país en el dataset?
- ¿Cuántos usuarios hay por cada red social del dataset?
- ¿Cuántos comentarios hay registrados por año en el dataset?
- ¿Cómo se encuentran distribuidos los sentimientos por comentario?
- ¿ Cuáles tipos de comentarios/sentimientos generan más reacciones?
- ¿Como están distribuidos los sentimientos por plataforma?

Base de datos

	Unnamed: 0	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes	Country	Year	Month	Day	Hour
0	0	Enjoying a beautiful day at the park!	Positive	15/01/2023 12:30	User123	Twitter	#Nature #Park	15	30	USA	2023	1	15	12
1	1	Traffic was terrible this morning.	Negative	15/01/2023 08:45	CommuterX	Twitter	#Traffic #Morning	5	10	Canada	2023	1	15	8
2	2	Just finished an amazing workout! ðŸ”ð	Positive	15/01/2023 15:45	FitnessFan	Instagram	#Fitness #Workout	20	40	USA	2023	1	15	15
3	3	Excited about the upcoming weekend getawa	Positive	15/01/2023 18:20	AdventureX	Facebook	#Travel #Adventure	8	15	UK	2023	1	15	18
4	4	Trying out a new recipe for dinner tonight.	Neutral	15/01/2023 19:55	ChefCook	Instagram	#Cooking #Food	12	25	Australia	2023	1	15	19
5	5	Feeling grateful for the little things in life.	Positive	16/01/2023 09:10	GratitudeNow	Twitter	#Gratitude #PositiveVibes	25	50	India	2023	1	16	9
6	6	Rainy days call for cozy blankets and hot coc	Positive	16/01/2023 14:45	RainyDays	Facebook	#RainyDays #Cozy	10	20	Canada	2023	1	16	14
7	7	The new movie release is a must-watch!	Positive	16/01/2023 19:30	MovieBuff	Instagram	#MovieNight #MustWatch	15	30	USA	2023	1	16	19
8	8	Political discussions heating up on the timel	Negative	17/01/2023 08:00	DebateTalk	Twitter	#Politics #Debate	30	60	USA	2023	1	17	8
9	9	Missing summer vibes and beach days.	Neutral	17/01/2023 12:20	BeachLover	Facebook	#Summer #BeachDays	18	35	Australia	2023	1	17	12
10	10	Just published a new blog post. Check it out!	Positive	17/01/2023 15:15	BloggerX	Instagram	#Bloggng #NewPost	22	45	USA	2023	1	17	15
11	11	Feeling a bit under the weather today.	Negative	18/01/2023 10:30	WellnessCheck	Twitter	#SickDay #Health	7	15	Canada	2023	1	18	10
12	12	Exploring the city's hidden gems.	Positive	18/01/2023 14:50	UrbanExplorer	Facebook	#CityExplore #HiddenGems	12	25	UK	2023	1	18	14
13	13	New year, new fitness goals! ðŸ”ð	Positive	18/01/2023 18:00	FitJourney	Instagram	#NewYear #FitnessGoals	28	55	USA	2023	1	18	18
14	14	Technology is changing the way we live.	Neutral	19/01/2023 09:45	TechEnthusiast	Twitter	#Tech #Innovation	15	30	India	2023	1	19	9
15	15	Reflecting on the past and looking ahead.	Positive	19/01/2023 13:20	Reflections	Facebook	#Reflection #Future	20	40	USA	2023	1	19	13
16	16	Just adopted a cute furry friend! ðŸ”ð	Positive	19/01/2023 17:10	PetAdopter	Instagram	#PetAdoption #FurryFriend	15	30	Canada	2023	1	19	17
17	17	Late-night gaming session with friends.	Positive	20/01/2023 00:05	GamerX	Twitter	#Gaming #LateNight	18	35	UK	2023	1	20	0
18	18	Attending a virtual conference on AI.	Neutral	20/01/2023 11:30	TechConference	Facebook	#AI #TechConference	25	50	USA	2023	1	20	11
19	19	Winter blues got me feeling low.	Negative	20/01/2023 15:15	WinterBlues	Instagram	#WinterBlues #Mood	8	15	USA	2023	1	20	15
20	20	Sipping coffee and enjoying a good book	Positive	21/01/2023 08:40	Bookworm	Twitter	#Reading #CoffeeTime	22	45	India	2023	1	21	8

[Ver en Kaggle](#)

Base de datos

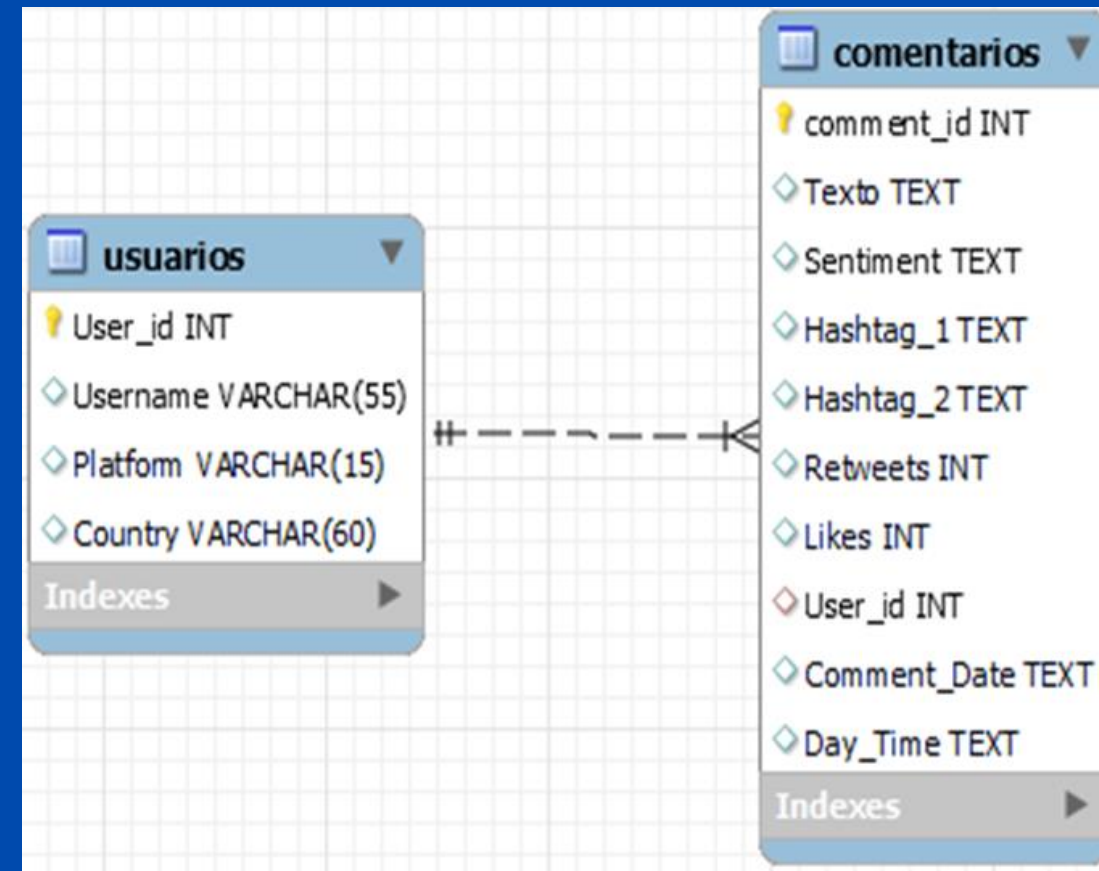
comment_id	Text	Sentiment	User_id	Username	Platform	Hashtag_1	Hashtag_2	Retweets	Likes	Country	date	exact_time	Hour_24
1	Enjoying a beautiful day at the park!	Positive	0	User123	Twitter	#Nature	#Park	15	30	USA	2023-01-15	12:30	12
2	Traffic was terrible this morning.	Negative	1	CommuterX	Twitter	#Traffic	#Morning	5	10	Canada	2023-01-15	8:45	8
3	Just finished an amazing workout! 💪🔥	Positive	2	FitnessFan	Instagram	#Fitness	#Workout	20	40	USA	2023-01-15	15:45	15
4	Excited about the upcoming weekend getaway!	Positive	3	AdventureX	Facebook	#Travel	#Adventure	8	15	UK	2023-01-15	18:20	18
5	Trying out a new recipe for dinner tonight.	Neutral	4	ChefCook	Instagram	#Cooking	#Food	12	25	Australia	2023-01-15	19:55	19
6	Feeling grateful for the little things in life.	Positive	5	GratitudeNow	Twitter	#Gratitude	#PositiveVibes	25	50	India	2023-01-16	9:10	9
7	Rainy days call for cozy blankets and hot cocoa.	Positive	6	RainyDays	Facebook	#RainyDays	#Cozy	10	20	Canada	2023-01-16	14:45	14
8	The new movie release is a must-watch!	Positive	7	MovieBuff	Instagram	#MovieNight	#MustWatch	15	30	USA	2023-01-16	19:30	19
9	Political discussions heating up on the timeline.	Negative	8	DebateTalk	Twitter	#Politics	#Debate	30	60	USA	2023-01-17	8:00	8
10	Missing summer vibes and beach days.	Neutral	9	BeachLover	Facebook	#Summer	#BeachDays	18	35	Australia	2023-01-17	12:20	12
11	Just published a new blog post. Check it out!	Positive	10	BloggerX	Instagram	#Bloggng	#NewPost	22	45	USA	2023-01-17	15:15	15
12	Feeling a bit under the weather today.	Negative	11	WellnessCheck	Twitter	#SickDay	#Health	7	15	Canada	2023-01-18	10:30	10
13	Exploring the city's hidden gems.	Positive	12	UrbanExplorer	Facebook	#CityExplore	#HiddenGems	12	25	UK	2023-01-18	14:50	14
14	New year, new fitness goals! 💪🔥	Positive	13	FitJourney	Instagram	#NewYear	#FitnessGoals	28	55	USA	2023-01-18	18:00	18
15	Technology is changing the way we live.	Neutral	14	TechEnthusiast	Twitter	#Tech	#Innovation	15	30	India	2023-01-19	9:45	9
16	Reflecting on the past and looking ahead.	Positive	15	Reflections	Facebook	#Reflection	#Future	20	40	USA	2023-01-19	13:20	13
17	Just adopted a cute furry friend! 🐾💕	Positive	16	PetAdopter	Instagram	#PetAdoption	#FurryFriend	15	30	Canada	2023-01-19	17:10	17
18	Late-night gaming session with friends.	Positive	17	GamerX	Twitter	#Gaming	#LateNight	18	35	UK	2023-01-20	0:05	0
19	Attending a virtual conference on AI.	Neutral	18	TechConference	Facebook	#AI	#TechConference	25	50	USA	2023-01-20	11:30	11
20	Winter blues got me feeling low.	Negative	19	WinterBlues	Instagram	#WinterBlues	#Mood	8	15	USA	2023-01-20	15:15	15
21	Sipping coffee and enjoying a good book.	Positive	20	Bookworm	Twitter	#Reading	#CoffeeTime	22	45	India	2023-01-21	8:40	8
22	Exploring the world of virtual reality.	Positive	21	VRExplorer	Facebook	#VR	#VirtualReality	15	30	USA	2023-01-21	13:20	13
23	Productive day ticking off my to-do list.	Positive	22	ProductivityPro	Instagram	#Productivity	#WorkFromHome	30	60	USA	2023-01-21	16:45	16

Esquema de la base de datos

1

```
CREATE TABLE Usuarios (  
  User_id INT PRIMARY KEY,  
  Username VARCHAR(55),  
  Platform VARCHAR(15),  
  Country VARCHAR(60)  
);
```

2



3

```
CREATE TABLE Comentarios (  
  Comment_id INT PRIMARY KEY,  
  Texto VARCHAR(255),  
  Sentiment VARCHAR(15),  
  Hashtag_1 VARCHAR(50),  
  Hashtag_2 VARCHAR(50),  
  Retweets INT,  
  Likes INT,  
  User_id INT,  
  Comment_Date DATE,  
  Day_Time TIME,  
  FOREIGN KEY (User_id) REFERENCES Usuarios(User_id)  
);
```

1. Creación de la Tabla usuarios
2. Diagrama Entidad-Relación
3. Creación de la Tabla comentarios

Primeras consultas

```
-- Muestra la cantidad de sentimientos almacenados  
SELECT COUNT(DISTINCT Sentiment) AS Cantidad_sentimientos  
FROM Comentarios;
```

	Cantidad_sentimientos
▶	191

```
-- Muestra la cantidad de usuarios registrados  
SELECT COUNT(DISTINCT User_id) AS Cantidad_de_Usuarios  
FROM Usuarios;
```

	Cantidad_de_Usuarios
▶	685

```
-- Muestra la cantidad de paises que aparecen en el dataset  
SELECT COUNT(DISTINCT Country) AS Cantidad_de_Paises  
FROM Usuarios;
```

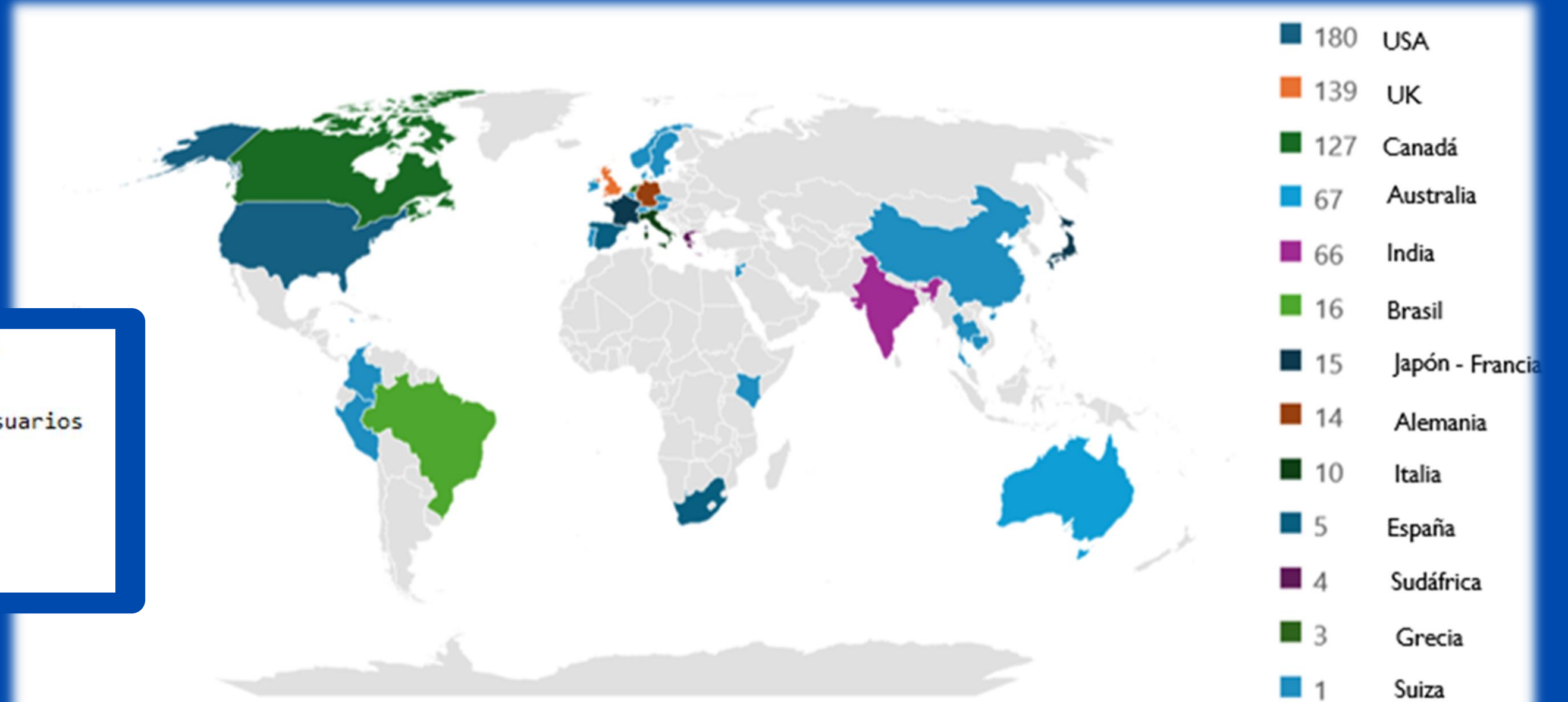
	Cantidad_de_Paises
▶	33

```
-- Muestra la cantidad de comentarios registrados  
SELECT COUNT(comment_id) AS Cantidad_de_Comentarios  
FROM Comentarios;
```

	Cantidad_de_Comentarios
▶	712

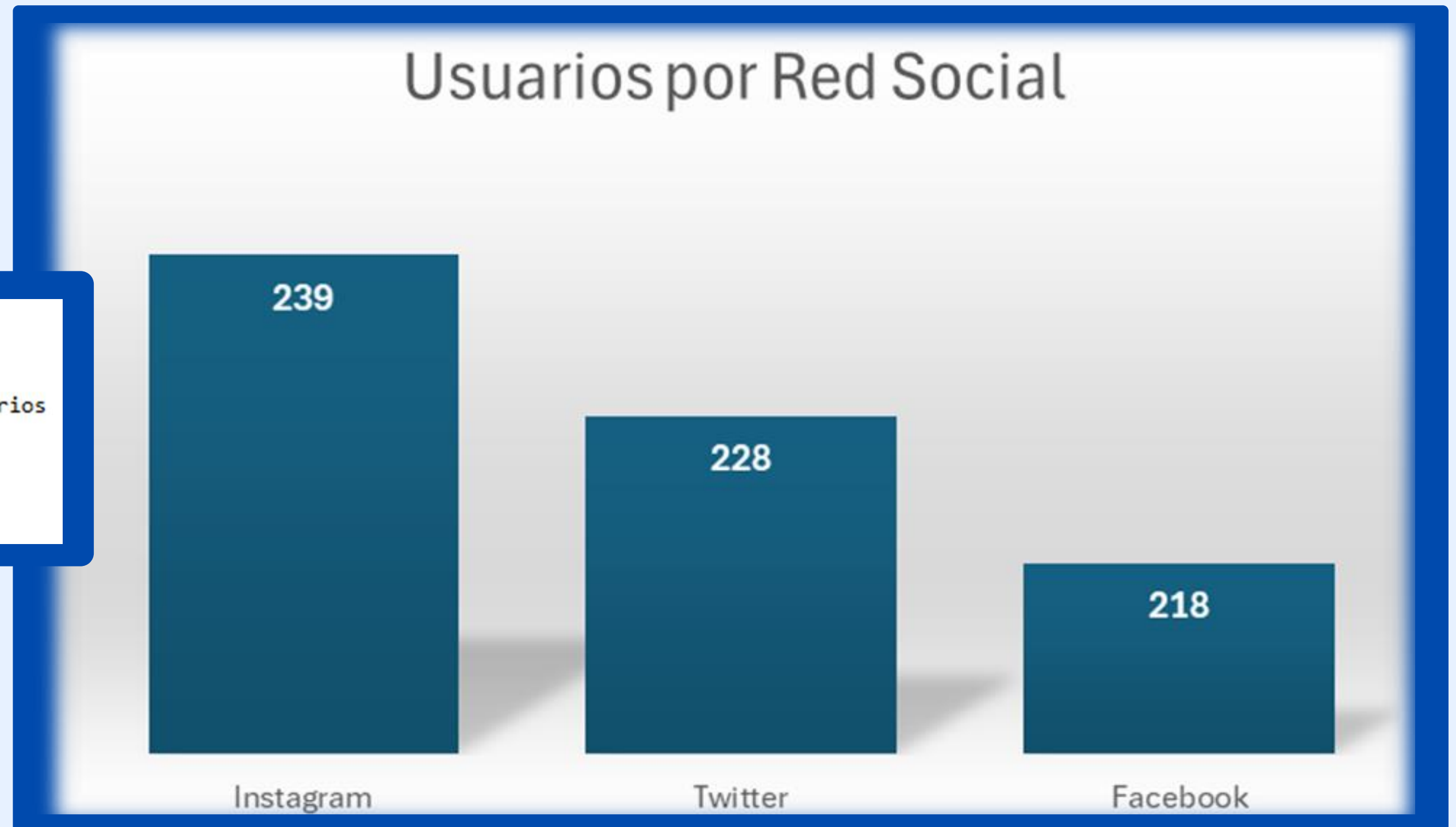
Usuarios por país

```
-- Muestra la cantidad de usuarios  
-- por país  
SELECT Country,Count(*) AS Total_Usuarios  
FROM usuarios  
GROUP BY Country  
Order by Total_Usuarios DESC  
LIMIT 13;
```



Usuarios por red social

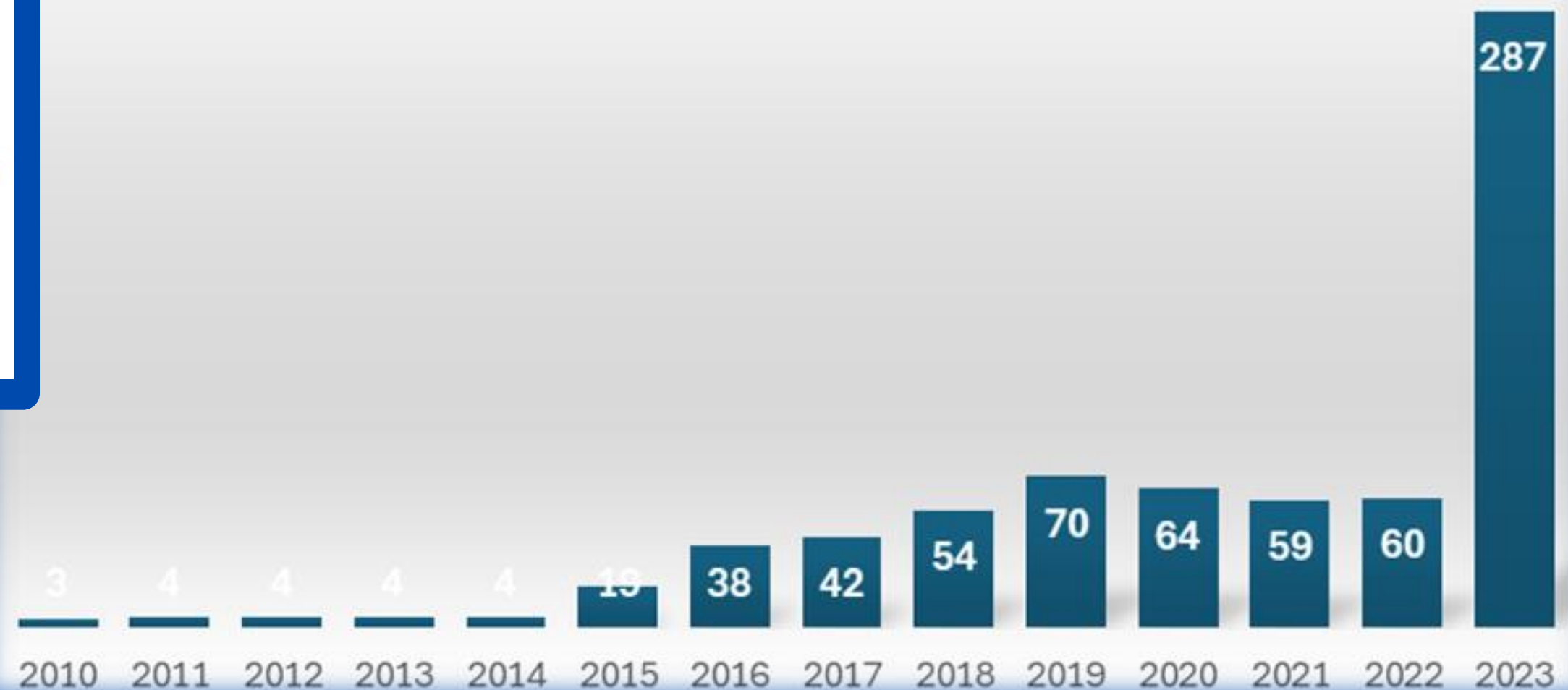
```
-- Muestra la distribucion de los  
-- usuarios por red social  
SELECT Platform,Count(*) AS Total_Usuarios  
FROM usuarios  
GROUP BY Platform  
Order by Total_Usuarios DESC;
```



Distribución de los comentarios por año

```
-- Muestra la distribución de los  
-- comentarios registrados por año  
SELECT year(Comment_Date) AS anio,  
       COUNT(*) Cantidad_comentarios  
FROM comentarios  
GROUP BY anio  
ORDER BY anio ASC;
```

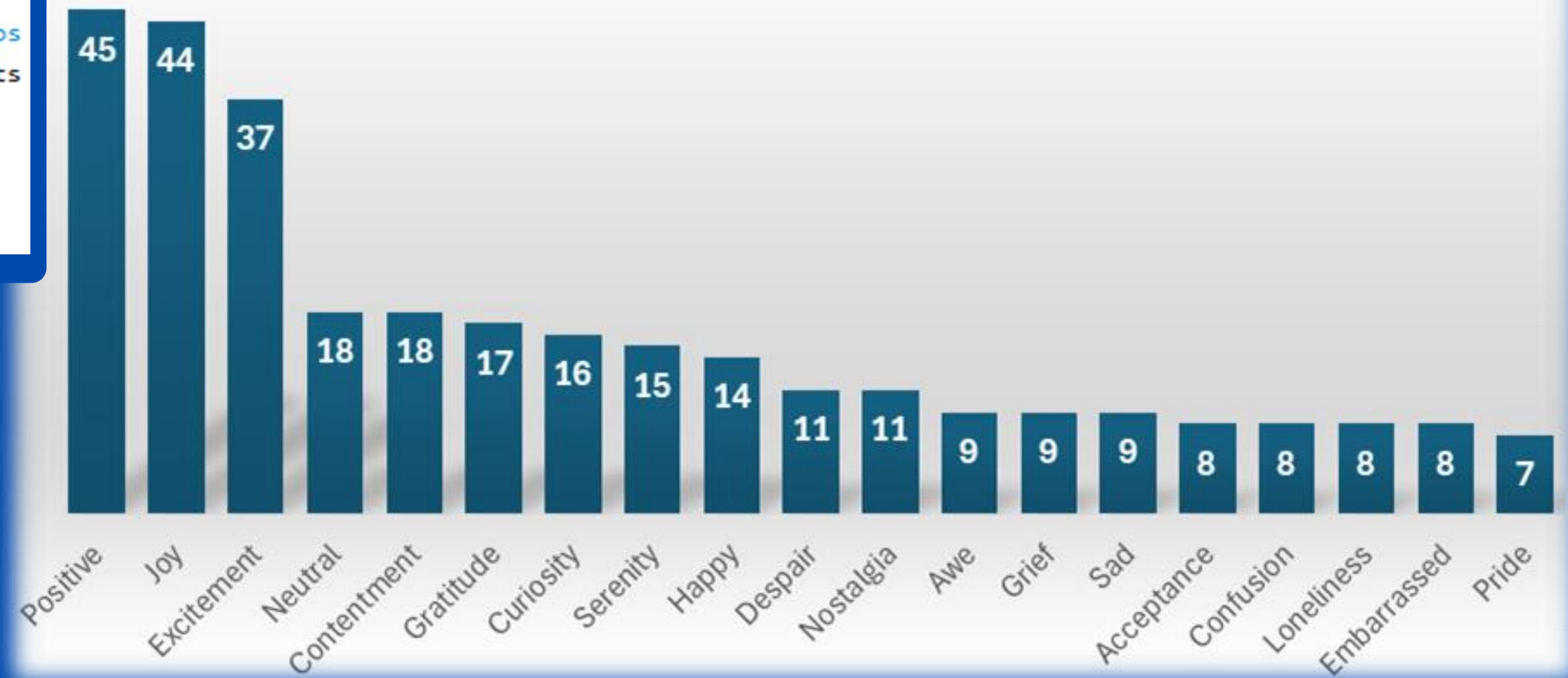
Cantidad de comentarios por año



Sentimientos más frecuentes

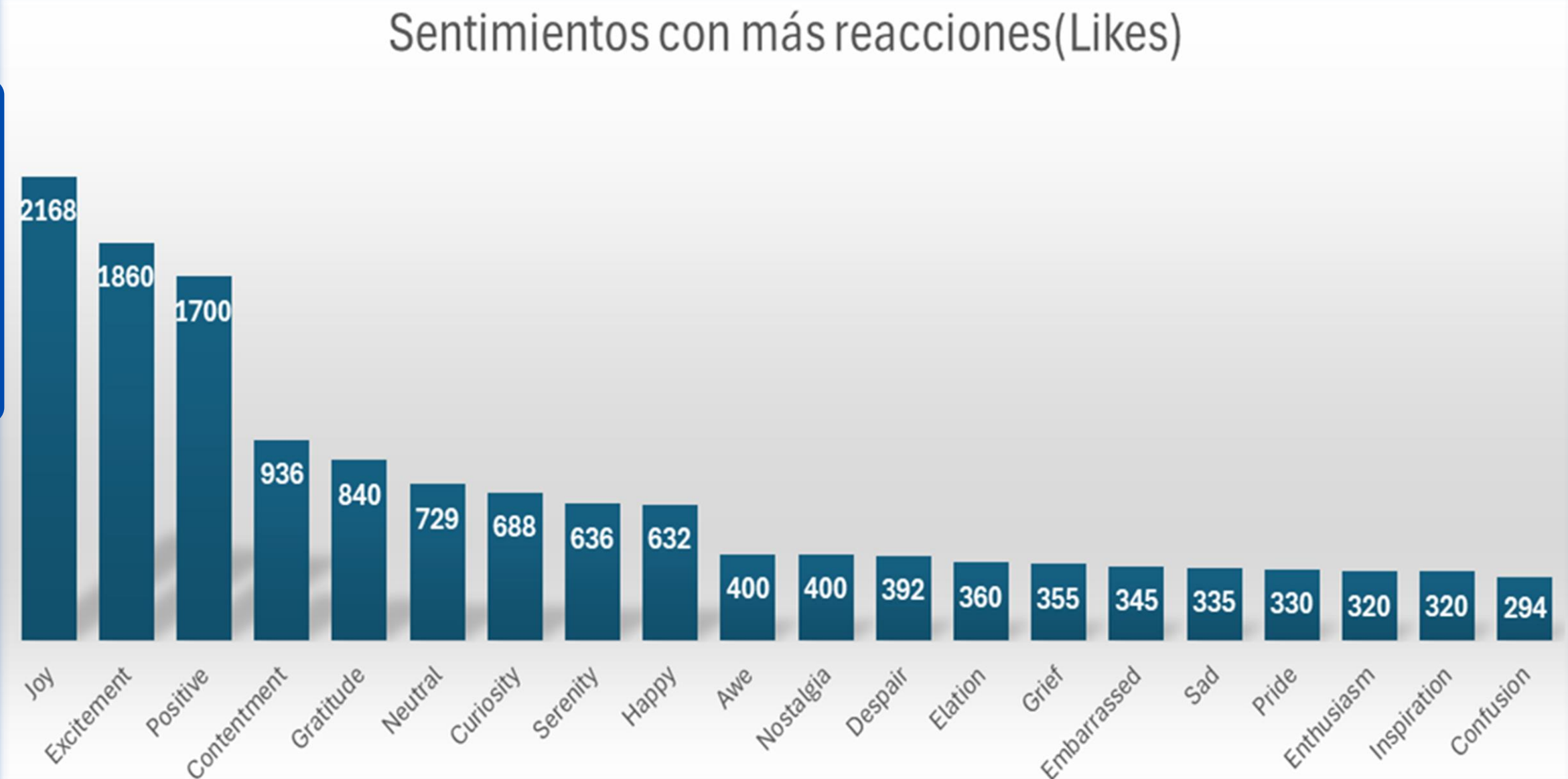
```
-- Muestra los primeros 20 sentimientos  
-- con mayor número comentarios registrados  
SELECT Sentiment,Count(*) AS Total_comments  
FROM comentarios  
GROUP BY Sentiment  
ORDER BY Total_comments DESC  
LIMIT 20;
```

Sentimientos con más registros



Tipo de sentimiento con mayor número de likes

```
-- Muestra sentimientos con más reacciones por encima  
-- del promedio  
SELECT Sentiment, SUM(Likes) Total_Likes  
FROM comentarios  
GROUP BY Sentiment  
HAVING Total_Likes > AVG(Likes)  
ORDER BY Total_Likes DESC  
LIMIT 20;
```



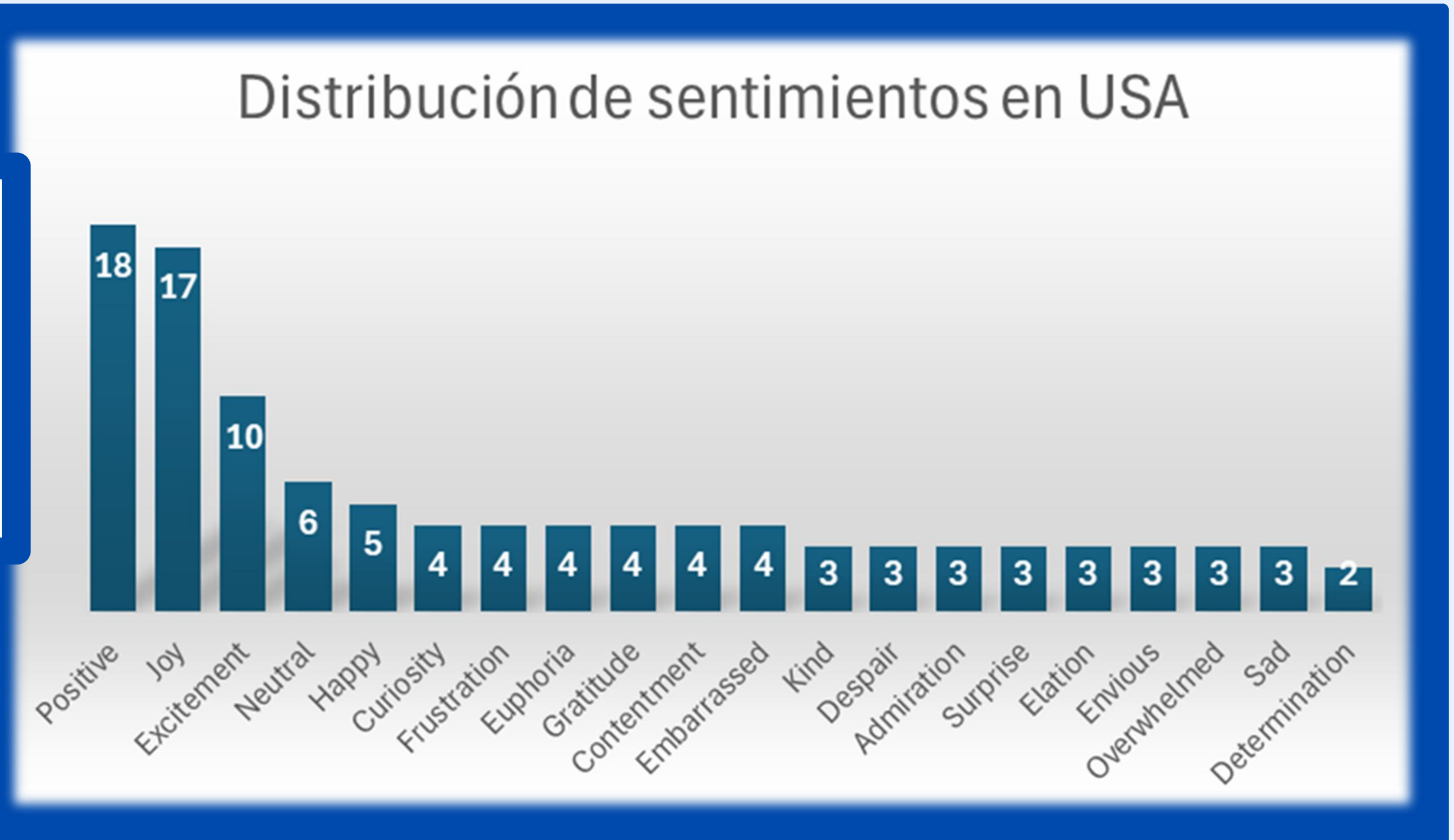
Tipo de sentimiento con mayor número de Retweets

```
-- Muestra sentimientos con más Retweets por encima  
-- del promedio  
SELECT Sentiment, SUM(Retweets) Total_Retweets  
FROM comentarios  
GROUP BY Sentiment  
HAVING Total_Retweets > AVG(Retweets)  
ORDER BY Total_Retweets DESC  
LIMIT 20;
```



Sentimientos más frecuentes en USA

```
-- Muestra la distribución de sentimientos en USA
-- primeros 20 registros
SELECT u.Country,
       c.Sentiment,
       COUNT(*) as Sentiment_Count
FROM   Comentarios AS c
JOIN   Usuarios AS u ON c.User_id = u.User_id
GROUP BY u.Country, c.Sentiment
HAVING u.Country = 'USA'
ORDER BY Sentiment_Count DESC
LIMIT 20;
```



Conclusiones

Este proyecto me permitió comprender la relevancia de las herramientas de ciencia y análisis de datos en la era digital. Los datos obtenidos a través de estas herramientas proporcionan una base sólida para la toma de decisiones informadas y precisas