

# **¿Existirá algunos factores de éxito en los videojuegos?**

## **INDICE**

- **Objetivo del proyecto:** pág. 2
- **Descripción del problema:** pág 2
- **Recursos disponibles:** págs 2-3
- **Hipótesis iniciales:** page 3-4
- **Definición de Stakeholders:** págs. 4-7
- **Preguntas claves:** pág 8
- **Fuentes de datos identificadas:** págs 8-10
- **Justificación del proyecto:** pág 10
- **Cantidad de datos y tipos de datos:** págs 10-11

**Objetivo del proyecto:** El objetivo principal de este proyecto es identificar y analizar los factores que influyen en el éxito de un videojuego, entendiendo “éxito” no solo como ventas elevadas, sino también como la capacidad de mantener una base activa de jugadores, obtener valoraciones positivas y generar una comunidad alrededor del título. Para lograr esto, se estudiarán variables como el precio, la duración de la campaña, el modelo de negocio (juego gratuito o de pago), el número de jugadores activos a lo largo del tiempo y la relación entre la calidad percibida y la inversión realizada por los usuarios.

**Descripción del problema:** La industria de los videojuegos ha crecido de manera exponencial en las últimas décadas y actualmente constituye uno de los sectores más rentables del entretenimiento. Sin embargo, este crecimiento viene acompañado de un aumento en la competencia y en los costos de producción. Desarrollar un videojuego hoy en día implica inversiones significativas en tecnología, marketing y talento humano. A pesar de ello, no existe garantía de que un título se convierta en un éxito comercial o cultural.

La incertidumbre es un problema central: muchos videojuegos fracasan porque no logran atraer a un público amplio, otros se estancan por un mal modelo de monetización, mientras que algunos triunfan gracias a factores inesperados como la viralidad en redes sociales o la recomendación boca a boca. Por ello, se vuelve fundamental encontrar patrones que permitan anticipar qué características aumentan la probabilidad de éxito en el mercado.

## **Recursos disponibles:**

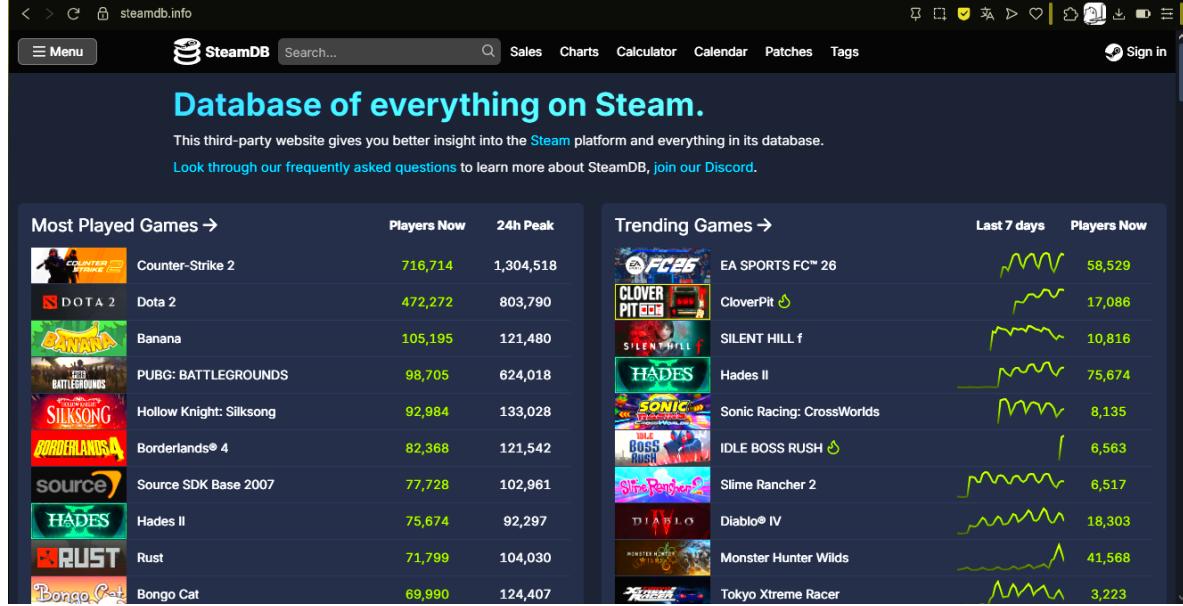
**Tecnología y herramientas:** se utilizará principalmente Python como lenguaje principal de programación y Pandas como biblioteca para almacenar los datos y poder manipularlos, además de Matplotlib como visualizador de los datos

**Datos:** la base de datos manejada en este proyecto tiene unos 60,000 registros de juegos, su fecha de lanzamiento, precio, puntuación positiva y negativa, la id del juego, el mínimo y el máximo de personas que compraron o tiene el videojuego y por último un pequeño contador de cuantos dura la campaña base del juego

También se usará SteamDB ofrece información detallada sobre la evolución del número de jugadores por título, lo que permite estudiar tendencias de crecimiento o decadencia. También se tomará en cuenta la comunidad de Steam como referencia de reseñas y calificaciones.

A	B	C	D	E	F	G	H	I	J	K	L
name	release_date	price	positive	negative	app_id	min_owners	max_owners	hltb_single			
Train Bandit	Oct 12, 2017	0.99	53	5	655370	0	20000				
Henosis&c	Jul 23, 2020	5.99	3	0	1355720	0	20000				
Two Weeks II	Feb 3, 2020	0	50	8	1139950	0	20000				
Wartune Reb	Feb 26, 2021	0	87	49	1469160	50000	100000				
TD Worlds	Jan 9, 2022	10.99	21	7	1659180	0	20000				
MazM: Jekyll	Apr 2, 2020	14.99	76	6	1178150	0	20000	7			
Deadlings: R	Nov 11, 2014	3.99	225	45	320150	50000	100000	4			
WARSAW	Oct 2, 2019	23.99	589	212	1026420	20000	50000	5			
Cthulhu Real	Jul 1, 2016	0	147	58	485000	50000	100000	3			
Clockwork D	Aug 27, 2021	1.99	5	0	1620060	0	20000				
Royal Battles	Apr 6, 2018	2.99	36	11	825930	20000	50000				
Diary of Luci	Nov 25, 2020	12.99	100	5	1454010	0	20000				
Hunting Unti	Nov 6, 2020	9.99	17	0	1330820	0	20000				
Hero of the K	Feb 20, 2015	7.99	2046	120	346560	100000	200000	5			
KHIO	Jul 24, 2020	9.99	3	1	1362670	0	20000				
Steam Squac	Jul 28, 2016	17.99	61	29	290870	20000	50000				
Project: R.E.E	Jan 27, 2017	2.99	29	28	575760	20000	50000	6			
Aerofly FS 2 F	Nov 20, 2017	37.49	1490	408	434030	100000	200000				
Turtle Lu	Mar 22, 2018	2.99	4	10	810740	20000	50000				
Kanjozoku G	Jul 16, 2022	5.99	392	57	2073470	20000	50000				

Lista Accesibilidad: No disponible



## Hipótesis iniciales:

**Hipótesis 1:** Un juego será exitoso si es gratuito y no posee una campaña lineal tradicional, priorizando la modalidad de servicio en línea (ej. juegos como servicio o “live service”), que fomentan la interacción continua.

**Hipótesis 2:** Los juegos de pago cuyo precio oscile entre 10 y 20 dólares y cuya campaña tenga al menos 2 horas de duración logran una mejor

relación costo-beneficio para los usuarios, lo cual incrementa su rentabilidad y aceptación.

**Hipótesis 3:** Los títulos con precios elevados (por encima de los 60 dólares) tienden a tener menos jugadores activos, lo que reduce su rentabilidad futura salvo que se trate de franquicias muy reconocidas.

## Definición de Stakeholders clave:

Definición de Stakeholders clave

En este proyecto, los stakeholders son aquellos actores directa o indirectamente involucrados en la industria de los videojuegos, que pueden verse beneficiados o impactados por los resultados del análisis. Identificar a los principales interesados y comprender sus necesidades es fundamental para asegurar que los hallazgos tengan utilidad práctica y aporten valor al mercado.

- **Compañías de videojuegos Indies:** Las desarrolladoras independientes representan una parte esencial de la industria actual. Muchas de ellas carecen de grandes presupuestos de marketing y producción, por lo que sus posibilidades de éxito dependen en gran medida de decisiones estratégicas bien fundamentadas.
- **Beneficios del estudio para ellas:**
- Obtener información sobre el rango de precios que resulta más atractivo para los jugadores.
- Identificar qué duración de campaña o modalidad de juego genera mayor retención de usuarios.
- Evaluar si conviene apostar por un modelo gratuito con micropagos, un precio bajo con buena relación costo-beneficio, o una estrategia más arriesgada con precios altos.
- Conocer los géneros y estilos más exitosos en el mercado actual, lo que les permitirá orientar su diseño hacia propuestas más competitivas.
- Con estos datos, las compañías indie podrán reducir la incertidumbre, tomar decisiones basadas en evidencia y mejorar sus probabilidades de lograr visibilidad en una industria altamente saturada.

COLABORADORES



- **Publishers o distribuidoras de videojuegos:** Los publishers son compañías que financian, publican y distribuyen títulos de diferentes desarrolladores. Su papel es fundamental en el mercado porque deciden qué proyectos apoyar y en cuáles invertir recursos de promoción.
- **Beneficios del estudio para ellos:**
- Contar con información más confiable para seleccionar juegos con mayor potencial de éxito.
- Evaluar la relación entre precio, duración y retención de jugadores para determinar qué propuestas resultan rentables.
- Prever riesgos al publicar juegos con características poco atractivas para los consumidores actuales.
- Diseñar campañas de marketing más precisas, dirigidas a segmentos de jugadores que realmente muestran interés en el producto.
- En resumen, este estudio les permite optimizar su portafolio de títulos y aumentar la probabilidad de que sus inversiones generen retornos significativos.

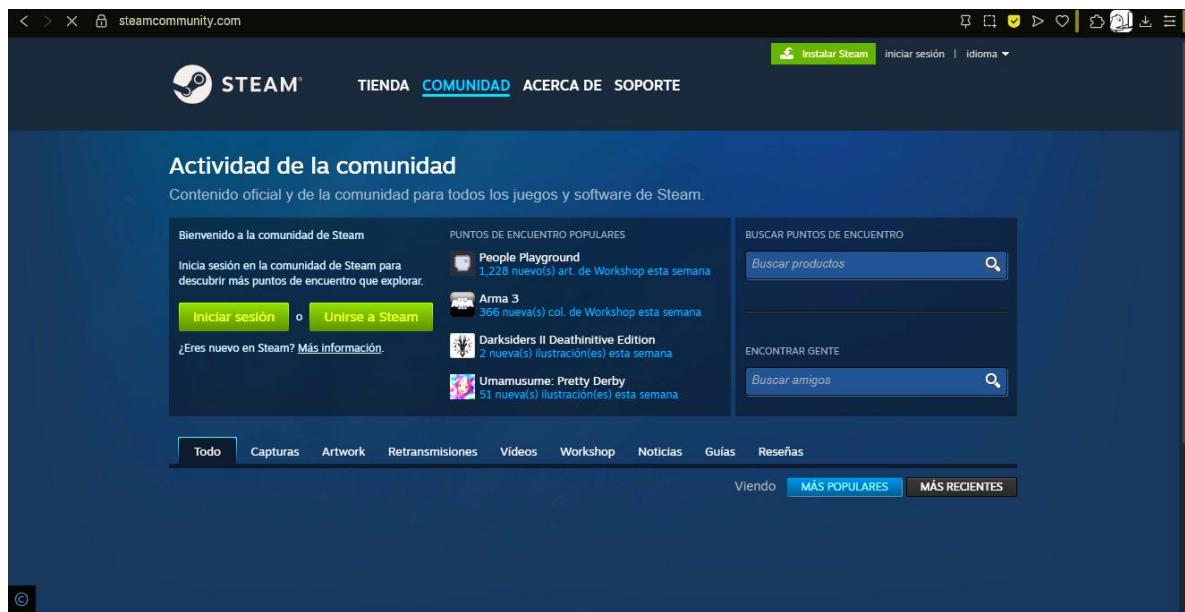


- **El mismo steam:** Steam, como la principal plataforma de venta y distribución digital de videojuegos para PC, juega un papel central en el ecosistema. Al permitir la publicación y promoción de miles de juegos cada año, necesita herramientas que le ayuden a gestionar mejor qué títulos recomendar o destacar en su escaparate.
- **Beneficios del estudio para Steam:**
- Analizar si existe una relación directa entre el modelo de precios y la popularidad de un juego.
- Ajustar su sistema de recomendaciones y algoritmos de visibilidad en la tienda, favoreciendo a títulos con mayor probabilidad de éxito.
- Entender si los juegos gratuitos mantienen un flujo más alto de jugadores activos o si se trata de una percepción parcial.
- Mejorar la experiencia de los usuarios finales al sugerirles títulos que se alineen mejor con sus preferencias de juego.
- Incrementar la rentabilidad general de la plataforma al promover juegos que mantengan una comunidad activa por más tiempo.



- **Comunidad de jugadores:** Aunque no sean actores productivos en el desarrollo, los jugadores constituyen la base de toda la industria, ya que su nivel de aceptación y compromiso define el éxito de los juegos.
- **Beneficios del estudio para ellos:**
- Acceder a títulos que se ajusten más a sus gustos y expectativas, ya que los desarrolladores tendrán mejores referencias para diseñarlos.
- Evitar gastar dinero en experiencias que no correspondan a la relación calidad-precio que esperan.

- Disfrutar de una mayor diversidad de juegos bien orientados al mercado, en lugar de propuestas repetitivas o mal diseñadas.



- **Inversionistas y analistas de mercado:** Los inversionistas buscan oportunidades dentro de la industria para generar ganancias, mientras que los analistas estudian tendencias para asesorar a empresas o diseñar estrategias de mercado.
- **Beneficios del estudio para ellos:**
  - Identificar segmentos de mercado con mayor potencial de crecimiento.
  - Comprender qué factores garantizan mayor rentabilidad en la industria de los videojuegos.
  - Disminuir riesgos al invertir en compañías o proyectos que presenten características alineadas con los factores de éxito identificados.



## Preguntas claves del proyecto

- 1. ¿Tendrán relación el precio con el aumento o disminución de jugadores de un videojuego?**
- 2. ¿Qué tendencia muestran los juegos mas largos y los más cortos?**
- 3. ¿Los juegos gratis siempre son los que más jugadores tienen?**
- 4. ¿El numero de jugadores afectara a las calificaciones positivas?**
- 5. ¿Existirá alguna relación entre el numero de opiniones positivas y el precio de un juego?**
- 6. ¿La fecha afecta al precio de un juego?**
- 7. ¿Cuál será el rango de precios para un videojuego?**
- 8. ¿Los videojuegos indie tendrán más éxito en precios bajo o en altos?**
- 9. ¿Tendrá algo que ver la duración de la campaña con el número de jugadore mínimos de un videojuego?**
- 10. ¿Qué característica presentan los juegos mas populares?**
- 11. ¿Tendrá algo que ver el nombre de un videojuego con su popularidad dentro de un género?**

## Fuentes de datos

- **Steam:** Steam es la plataforma de distribución digital de videojuegos más grande para PC y representa un punto de partida fundamental para el análisis.
- **Tipo de datos disponibles:**
  - Catálogo de videojuegos con información detallada como nombre, género, precio, fecha de lanzamiento y desarrollador.
  - Valoraciones de la comunidad en forma de reseñas positivas y negativas.

- Estadísticas sobre ventas y popularidad de los títulos más recientes y clásicos.
- **Relevancia:** Steam es la principal vitrina de la industria del PC gaming. Analizar los datos disponibles en su plataforma permite identificar patrones de consumo, conocer la percepción que tienen los usuarios sobre los juegos y establecer correlaciones entre variables como precio, género y calificaciones.
- **SteamDB (Steam database):** SteamDB es una base de datos no oficial pero altamente confiable que recopila información pública de la API de Steam. Se ha convertido en una de las herramientas más utilizadas por analistas, jugadores y desarrolladores.
- **Tipo de datos disponibles:**
  - Número de jugadores concurrentes de cada juego a lo largo del tiempo, lo que permite ver tendencias de crecimiento o decrecimiento.
  - Historial de precios, incluyendo descuentos, promociones y variaciones según región.
  - Información sobre actualizaciones y cambios realizados en los juegos, lo que ayuda a estudiar su impacto en la retención de usuarios.
- **Relevancia:** Esta fuente es clave para estudiar la **evolución temporal** del éxito de un videojuego. Gracias a los datos de jugadores concurrentes y al historial de precios, es posible evaluar la efectividad de estrategias de marketing, eventos especiales y actualizaciones, así como analizar cómo afectan estos factores a la permanencia de un juego en el mercado.
- **La comunidad de Steam (Reseñas y foros):** La comunidad activa de usuarios de Steam constituye una fuente de información cualitativa y cuantitativa muy valiosa.
- **Tipo de datos disponibles:**
  - Opiniones de jugadores a través de reseñas escritas.
  - Calificaciones numéricas o porcentuales de satisfacción.
  - Discusiones en foros sobre puntos fuertes y débiles de cada título.
- **Relevancia:** Analizar los comentarios y reseñas permite obtener información más subjetiva pero igualmente importante, ya que refleja la percepción del jugador sobre la calidad del producto. Estos datos pueden complementar el análisis estadístico, ayudando a explicar por qué ciertos juegos alcanzan la popularidad y otros no, incluso cuando presentan características similares en precio o duración.

**Bases de datos internas del proyecto:** La base de datos ya disponible para el proyecto, con más de 60,000 registros de videojuegos, constituye el núcleo del análisis.

- **Tipo de datos incluidos:**
  - Fecha de lanzamiento.
  - Precio (con posibilidad de estudiar su variación en el tiempo).
  - Puntuaciones positivas y negativas.
  - Número mínimo y máximo de jugadores que han adquirido el título.
  - Duración de la campaña principal.
- **Relevancia:**

Con esta información se pueden realizar análisis exploratorios, visualizaciones y modelos predictivos iniciales. Al estar estructurada en un formato compatible con herramientas como Python y Pandas, facilita la limpieza, manipulación y el cruce con otras fuentes externas.

**Fuentes complementarias externas (opcional):** Aunque el foco principal está en Steam y SteamDB, también existen fuentes adicionales que podrían enriquecer el estudio:

- **Metacritic:** para obtener puntuaciones profesionales y comparar la percepción de críticos con la de jugadores.
- **Redes sociales (Twitter, Reddit, YouTube, TikTok):** útiles para medir viralidad, tendencias y menciones de juegos populares.
- **Informes de la industria (Newzoo, Statista):** que proporcionan datos macroeconómicos y tendencias generales del mercado de videojuegos.

## Justificación del proyecto

La creciente relevancia de la industria de los videojuegos, tanto en términos económicos como culturales, justifica la necesidad de contar con estudios que reduzcan la incertidumbre en el desarrollo y lanzamiento de nuevos títulos.

Comprender qué factores contribuyen al éxito de un videojuego permitirá a los desarrolladores y publishers tomar decisiones basadas en evidencia, en lugar de depender únicamente de la intuición o de tendencias pasajeras.

Analizar más de 60,000 registros de Steam y complementarlos con información de SteamDB ofrece la oportunidad de encontrar patrones significativos: desde cómo influye el modelo de monetización en la retención de jugadores hasta qué relación existe entre la duración de una campaña y la percepción de valor por parte de los usuarios.

Además, este proyecto aportará herramientas prácticas para:

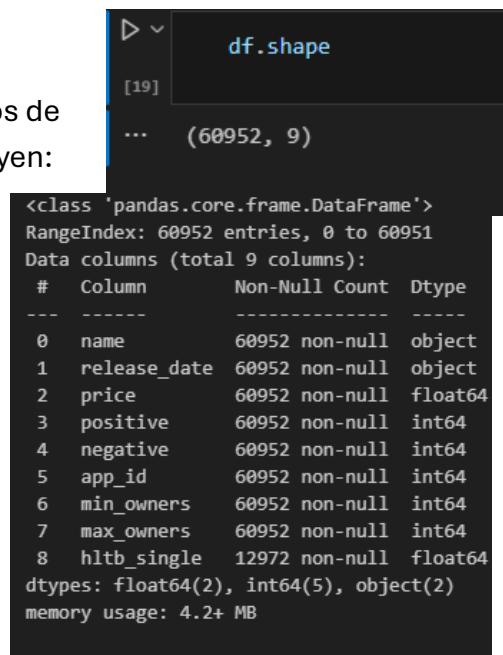
- Reducir riesgos financieros al momento de fijar precios o elegir estrategias de mercado.
- Mejorar la competitividad de las empresas independientes.
- Optimizar las recomendaciones y la publicidad en plataformas como Steam.
- Satisfacer de mejor manera los intereses y expectativas de los jugadores.

En conclusión, este estudio representa una oportunidad estratégica para todos los actores involucrados en la industria, pues permitirá anticipar con mayor precisión los elementos que determinan el éxito de un videojuego y, al mismo tiempo, mejorar la experiencia de los usuarios finales.

## ¿Cuántos datos son y qué tipo son?

La base de datos utilizada contiene más de 60,000 registros de videojuegos publicados en Steam. Los tipos de datos incluyen:

- **Enteros (int):** número de jugadores mínimos y máximos, identificador del juego, conteo de reseñas positivas y negativas.
- **Flotantes (float):** precio del videojuego, duración de la campaña principal, métricas promedio de jugadores.
- **Objetos (object):** nombre del juego, género, fecha de lanzamiento, etiquetas adicionales proporcionadas por la comunidad.



The screenshot shows a Jupyter Notebook cell with the following output:

```
df.shape  
[19] ... (60952, 9)
```

Below the output, the DataFrame structure is displayed:

#	Column	Non-Null Count	Dtype
0	name	60952 non-null	object
1	release_date	60952 non-null	object
2	price	60952 non-null	float64
3	positive	60952 non-null	int64
4	negative	60952 non-null	int64
5	app_id	60952 non-null	int64
6	min_owners	60952 non-null	int64
7	max_owners	60952 non-null	int64
8	hltb_single	12972 non-null	float64

dtypes: float64(2), int64(5), object(2)  
memory usage: 4.2+ MB

La diversidad de tipos de datos permite un análisis cuantitativo y cualitativo. Gracias a la combinación de variables numéricas y categóricas, será posible aplicar técnicas de análisis estadístico, visualización de tendencias y, potencialmente, modelos predictivos de machine learning para identificar patrones de éxito.