



**Bootcamp Data Science**  
**Promoción Febrero 2025**  
**EDA Google Play Store**

**Autor: David García García**  
**Mayo 2025**

# ÍNDICE

1.	INTRODUCCIÓN.....	3
2.	HIPÓTESIS .....	4
3.	ANÁLISIS UNIVARIANTE .....	5
3.1	Rating (puntuación media) .....	5
3.2	Installs (número de descargas) .....	6
3.3	Type (gratuita vs de pago) .....	7
3.4	Price (precio en USD).....	8
3.5	Category (categoría) .....	9
4.	ANÁLISIS BIVARIANTE .....	10
4.1	Ratings vs Type .....	10
4.2	Rating vs Installs .....	11
4.3	Rating vs Reviews .....	12
4.4	Rating vs Category .....	13
4.5	Rating vs Price.....	13
5.	EJEMPLOS Y EXTRAS.....	15
5.1	Junk Apps vs Category .....	15
5.2	Rankings.....	16
6.	CONCLUSIONES.....	18
7.	ANEXO: CLAVES PARA LANZAR UNA APP EXITOSA .....	19
7.1	Fórmula del éxito al lanzar una app.....	19
7.2	Modelo de monetización: <i>freemium</i> , anuncios o pago único .....	19

## 1. INTRODUCCIÓN

En pleno 2025 y con el desarrollo tecnológico avanzando a pasos agigantados, el teléfono móvil se ha convertido en una parte indispensable de nuestro día a día. Tanto como herramienta de comunicación, fotografía, entretenimiento, o incluso trabajo... recurrir a nuestro smartphone es algo que ya tenemos tan interiorizado que cuesta imaginarse como sería el mundo si estos no existieran.

En este contexto, resulta vital el software con el que llevamos todo a cabo, y es que, si podemos hacer tantas cosas con un pequeño aparato, es gracias a la infinidad de aplicaciones de todo tipo que tenemos al alcance de un par de clics. Pero...

¿Alguna vez te has frustrado por la gran cantidad de anuncios y spam que tienen algunas de estas aplicaciones? O quizás quieres lanzarte al mercado con una idea revolucionaria y no sabes si para triunfar hay una “fórmula mágica” o, al menos, ciertos factores comunes entre los grandes gigantes de éxito.

Estas son algunas de las preguntas que trataremos de responder con este EDA, de manera que, al final de todo el análisis, podamos entender un poco mejor algo que tenemos tan al alcance pero que, a la vez, puede parecer tan desconocido.

## 2. HIPÓTESIS

Nuestra base de trabajo será un CSV publicado en Kaggle, obtenido mediante *scrapping*, que recoge datos de más de 10000 aplicaciones de Google Play Store. Los datos aquí recogidos son del año 2018, aunque esto no nos impide hacer nuestro análisis, puesto que muchas de las grandes aplicaciones ya existían entonces y, además, no centraremos nuestras hipótesis en ejemplos de apps en concreto, sino en el comportamiento general según ciertas variables para resolver, después del análisis pertinente, las cuestiones que nos hayamos planteado.

Vamos a plasmar aquí algunas de ellas para entender mejor por qué se hacen los análisis estadísticos mostrados durante el desarrollo del EDA:

¿Qué tipo de apps tienen mejores valoraciones? ¿Hay relación entre el número de descargas y la valoración media? ¿Qué proporción de apps de pago hay? ¿Influyen estas en los factores que acabamos de mencionar? ¿Qué categorías concentran más apps “basura”? ¿Qué características comparten las apps más exitosas del mercado?

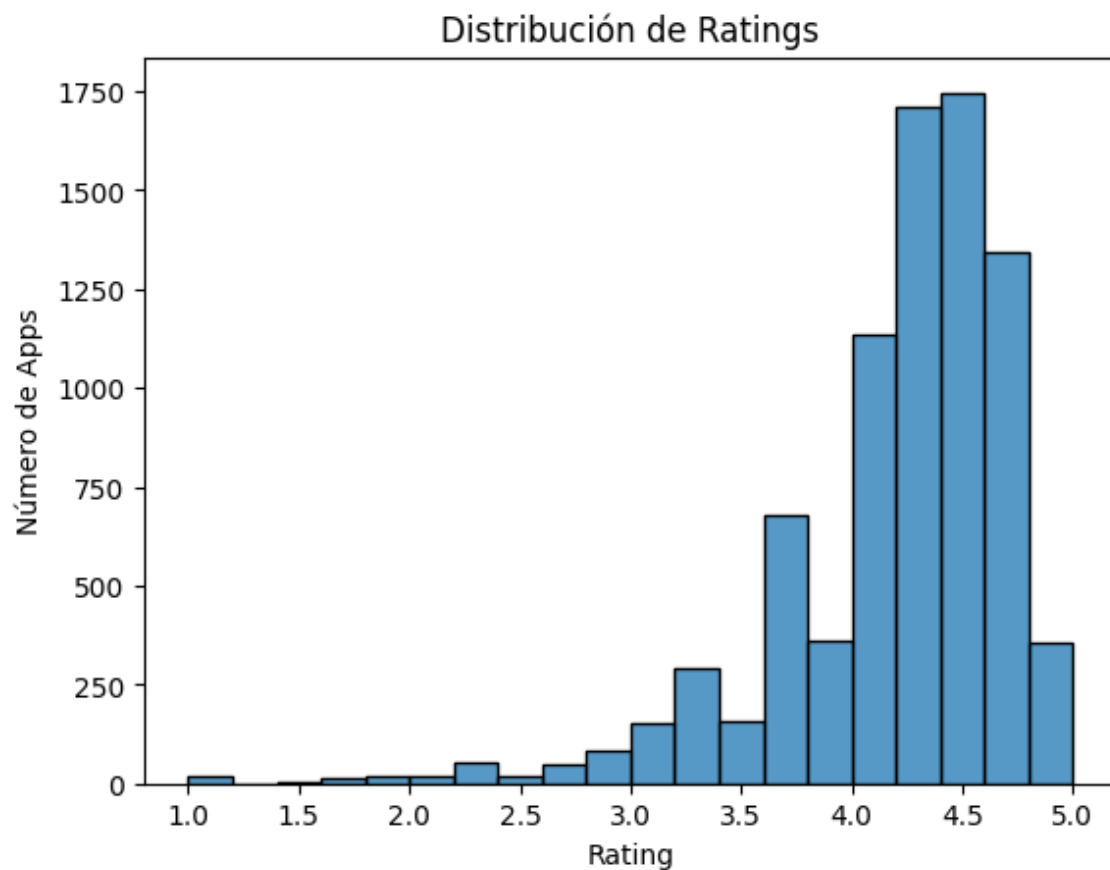
Contestaremos todas estas preguntas para que, finalmente, podamos llegar a la conclusión del trabajo:

Si queremos lanzar una aplicación al mercado, ¿cuáles son las directrices que podríamos seguir para tener una mayor probabilidad de triunfar?

### 3. ANÁLISIS UNIVARIANTE

En primer lugar, vamos a hacer un primer análisis en profundidad variable por variable para ver como se comportan y sacar nuestras primeras conclusiones:

#### 3.1 Rating (puntuación media)



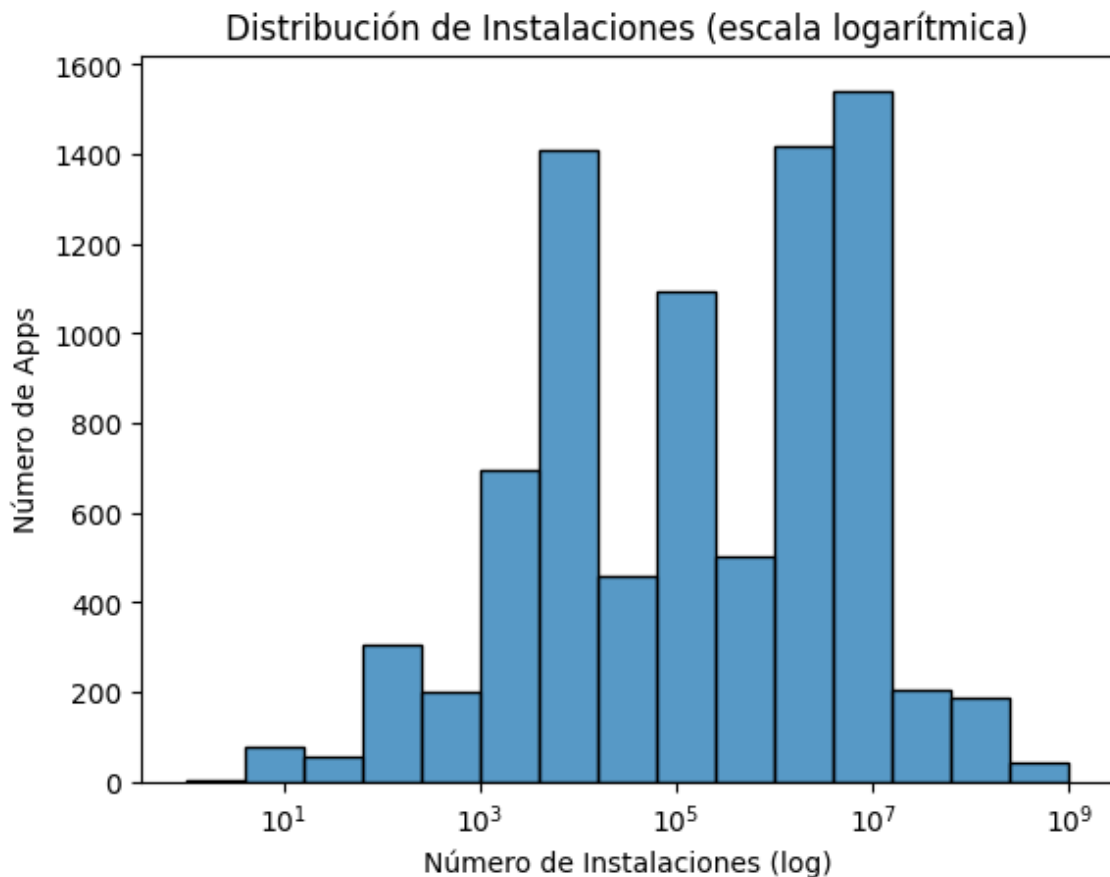
La mayoría de apps tienen valoraciones entre 4,0 y 4,5, y hay muy pocas con rating por debajo de 3,0. Podemos apreciar claramente la forma acampanada de la distribución, asentándose la mayor parte de las valoraciones entre 3.8 y 4.6 más o menos. Hay en comparación muchas más apps con un 5 que con valoración muy baja (por debajo de 2.5).

Podríamos decir que Google Play parece tener un sesgo positivo en valoraciones, y es que seguramente las apps mal valoradas acaben desapareciendo, dejándose abandonadas o incluso siendo ocultadas por la plataforma.

Es un buen primer análisis, aunque debemos compararlo con otras variables, como el número de descargas, ya que, por ejemplo, apps con valoraciones extremadamente altas y sin muchas descargas no resultarían demasiado significativas.

Es decir, el rating por sí solo no es suficiente para determinar el éxito, ya que la mayoría de las aplicaciones tienen una puntuación media bastante alta.

### 3.2 Installs (número de descargas)



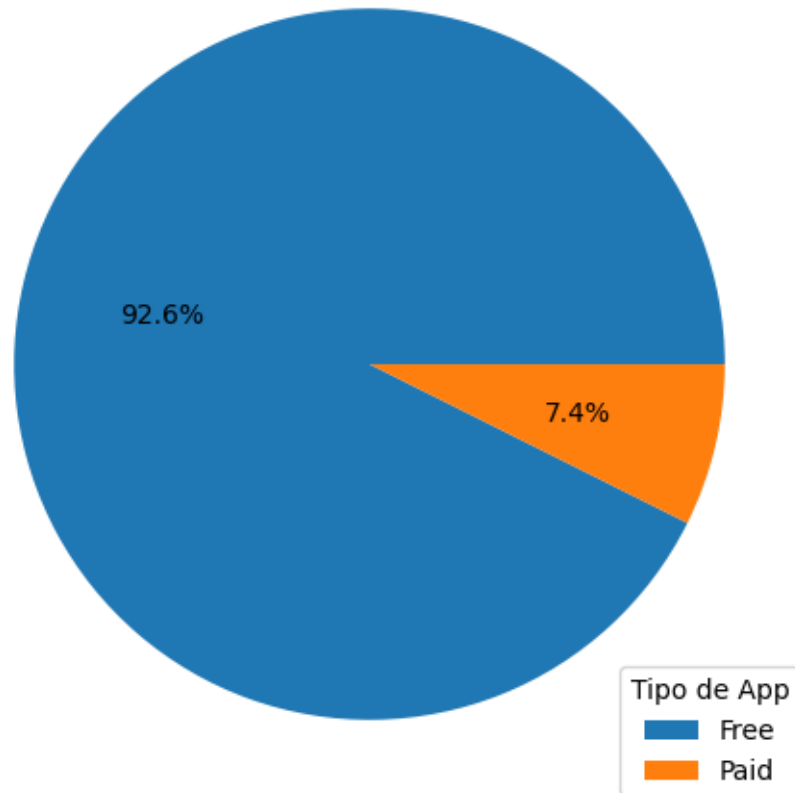
Installs es la variable que mide la popularidad objetiva de una aplicación. Nos permite detectar lo difícil que puede ser viralizar una app, qué proporción del mercado dominan las apps top y si dicho éxito está repartido o concentrado.

Es importante usar la escala logarítmica, pues así se pueden visualizar tanto los grandes gigantes, como Youtube o Instagram, como también las pequeñas *startups*, recién salidas al mercado.

En este caso la distribución está algo más centrada (en esta escala claro). La mayoría de apps tienen entre 10.000 y 10 millones de instalaciones, habiendo muy pocas sobre todo por debajo de 1000. Del mismo modo, solo unas pocas superan los 100 millones.

### 3.3 Type (gratuita vs de pago)

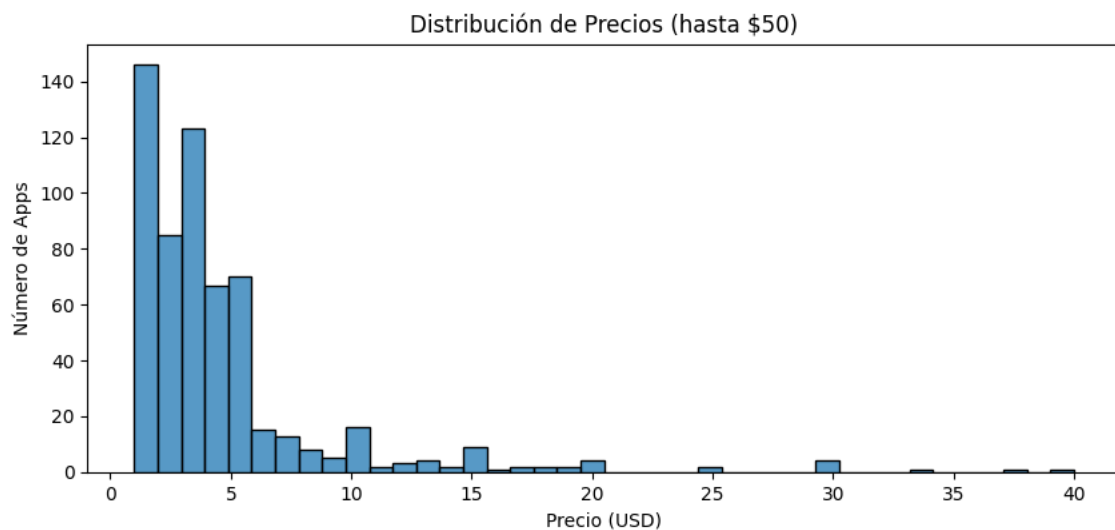
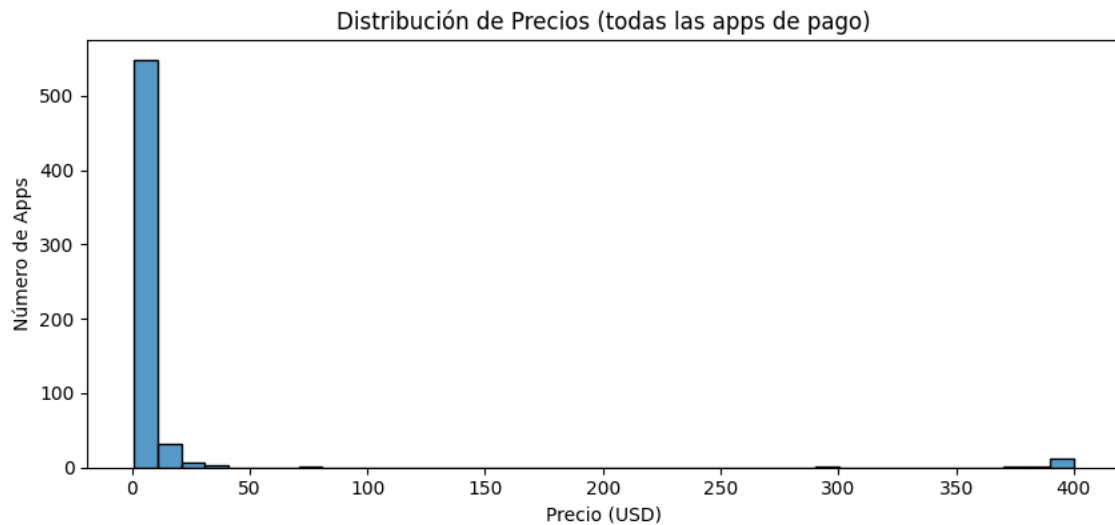
Distribución de Apps: Gratuitas vs de Pago



Esto confirma que el ecosistema de apps en Google Play está dominado por el modelo gratuito (contando dentro obviamente las apps freemium, micropagos, anuncios implícitos, etc). Es decir, si quieres que tu app triunfe, ha de ser gratuita preferiblemente.

Las apps de pago son una minoría, más tarde analizaremos si, siendo de pago, tienden a triunfar más, ofrecer contenido de valor...

### 3.4 Price (precio en USD)



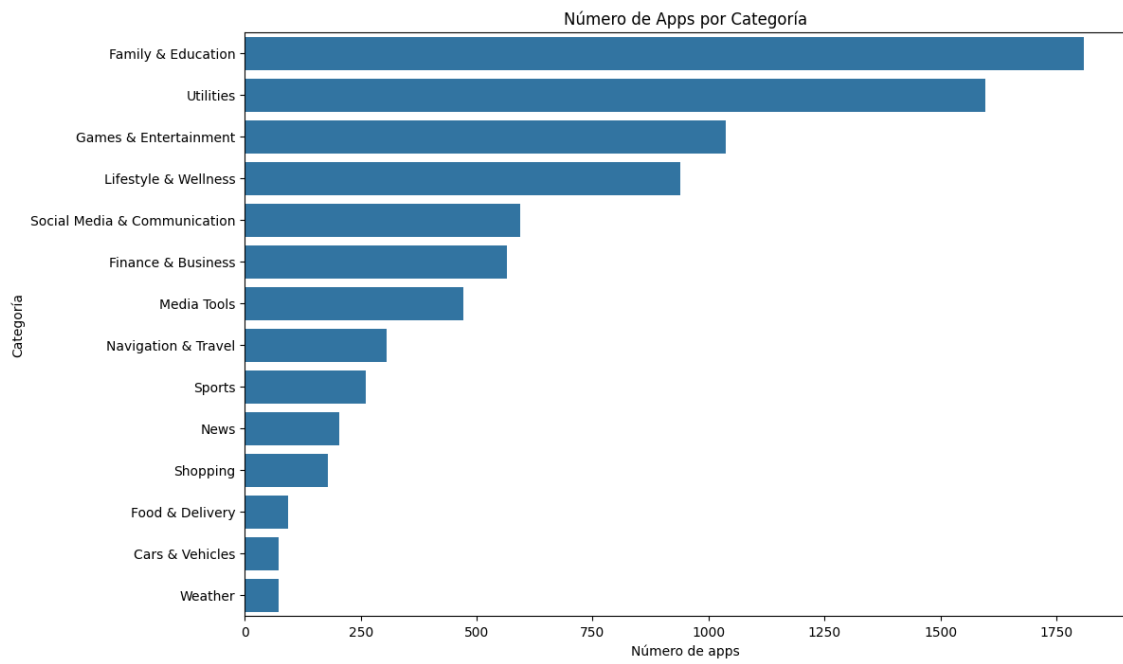
Claramente se ve que casi todas las apps de pago tienen un precio bajo.

Aparecen algunos outliers extremos (hay menos de 10 aplicaciones que valgan más de \$40, en general a partir de \$20 podríamos decir que son todo outliers), por eso hacemos la segunda gráfica, para la visualización del rango de precios más común.

La distribución es, como era de esperar, decreciente: a medida que sube el precio, disminuye el número de apps, lo cual tiene sentido pues, en un contexto en que la mayoría de apps son gratuitas, si el usuario elige una de pago no va a estar dispuesto a gastar en exceso.



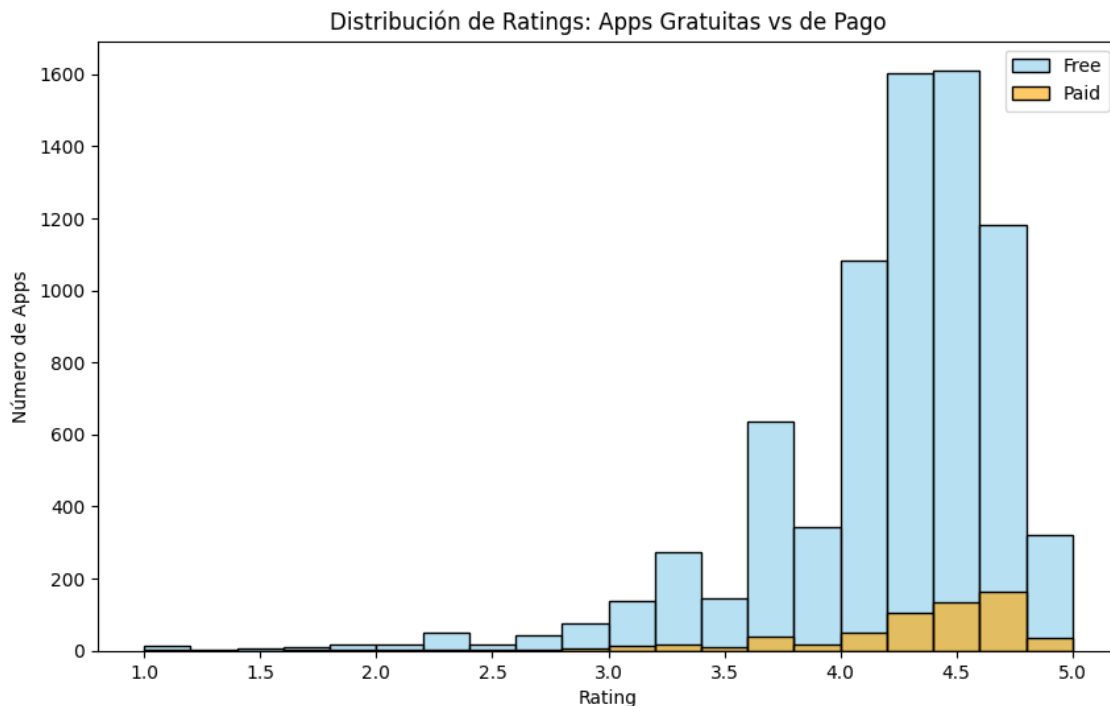
### 3.5 Category (categoría)



## 4. ANÁLISIS BIVARIANTE

Comencemos a sacar conclusiones más interesantes cruzando variables:

### 4.1 Ratings vs Type

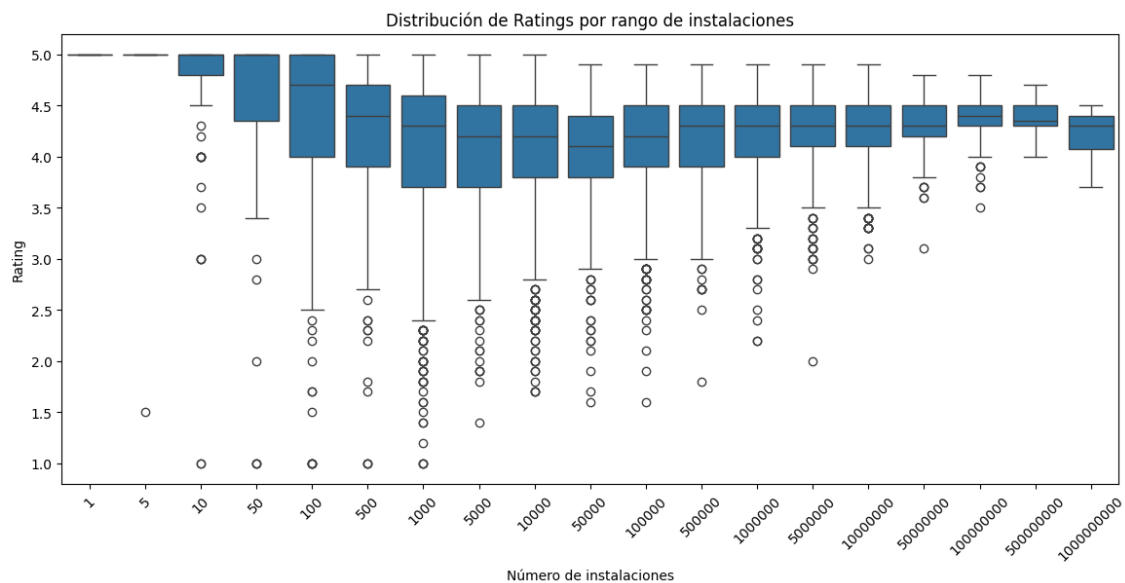


Ambas distribuciones están centradas en valoraciones altas (entre 4 y 5), como habíamos analizado antes. Además, podemos ver que la distribución de las de pago está un poco más desplazada a la derecha (valoración media más alta), aunque no es del todo significativo. Lo que si apreciamos es que hay muchas menos apps de pago con ratings bajos (por debajo de 3.5)

Quizás los usuarios que pagan por una app tienen más predisposición a valorarla bien, por el efecto psicológico de "justificar la compra", o quizás es que, al ser de pago, están más pulidas y por eso están mejor valoradas. Sobre todo, y algo que puede ser diferencial en las valoraciones, es que las apps de pago no suelen (casi nunca), tener anuncios. De hecho, gran parte del modelo *freemium* se basa en aplicaciones gratuitas plagadas de anuncios, que te ofrecen eliminarlos pagando una cierta cantidad. Es decir, acabarían siendo una aplicación de pago, a pesar de estar categorizadas como gratuitas en Google Play.

En cuanto al test de hipótesis correspondiente, el p-valor es muy bajo ( $0.00008 < 0.05$ ), es decir, podríamos decir que las apps de pago presentan, en promedio, una valoración significativamente más alta que las gratuitas. Las gratuitas compiten en volumen, y aunque la valoración suele ser alta, muchas también pueden ser “apps basura”.

## 4.2 Rating vs Installs



Interpretemos este boxplot por tramos, puesto que es como están categorizadas las descargas:

En cuanto a las apps con un número bajo de instalaciones, los ratings son sospechosamente altos, muchos clavados en 5, lo cual puede hacernos sospechar que sean autovaloraciones del desarrollador o de conocidos. La dispersión es muy pequeña, casi nula (las cajas son prácticamente una línea) y los valores están concentrados casi en 5, está muy sesgado.

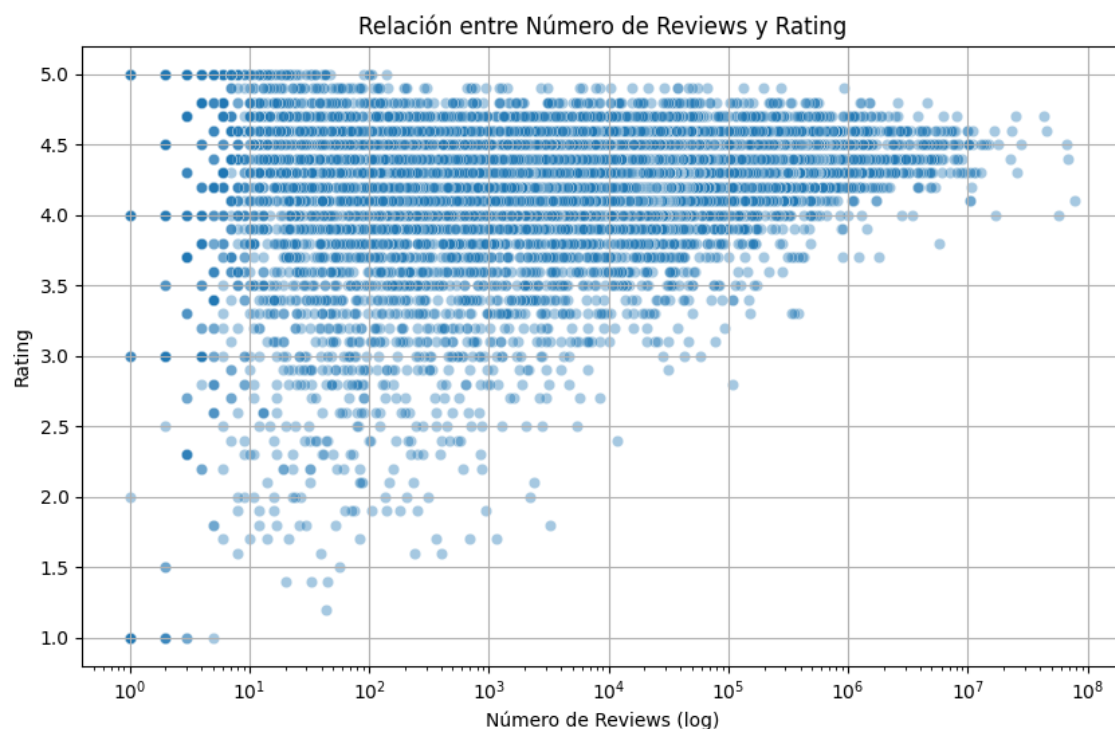
En cuanto a los tramos intermedios (100 a 10.000), la mediana baja y aparece más variabilidad, hay tanto buenos como malos ratings, con algún outlier con puntuaciones muy bajas. Es decir, cuando llegamos a un número ya decente de instalaciones, es raro que la gente siga dando uso a la aplicación si no es suficientemente buena, por eso las puntuaciones tan bajas son casos extremos. En este rango, ya se empiezan a consolidar y dispersar más las valoraciones, no hay tanto sesgo.

En tramos ya altos, a partir de 100.000, la mediana se estabiliza alrededor de 4.3–4.4 y las cajas se estrechan, hay menos dispersión. Esto refleja que las apps muy descargadas

y ya consolidadas seguramente sean ya conocidas, estén bien optimizadas y, por tanto, asentadas en un rating medio bastante alto.

En resumen, las apps con muy pocas instalaciones tienden a estar infladas por valoraciones extremas. A medida que una app se vuelve más popular, su rating se estabiliza y se vuelve más representativo del juicio de la comunidad. Finalmente, las apps más descargadas tienden a consolidarse en un rango de rating alto, pero no extremo (4.2–4.4).

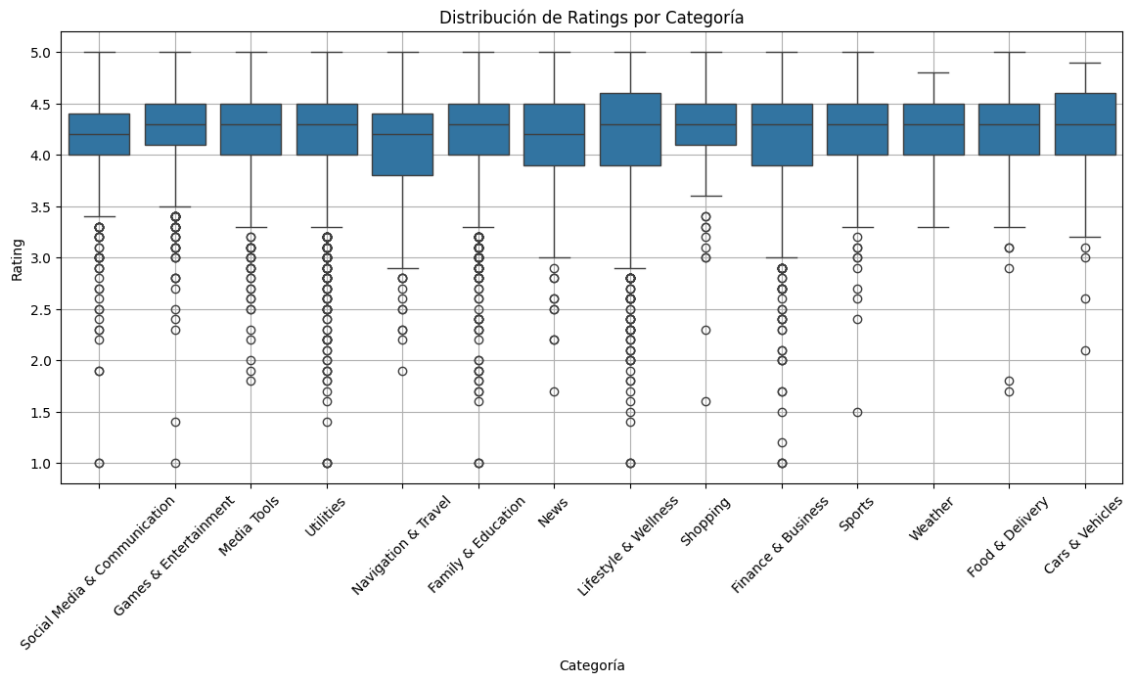
### 4.3 Rating vs Reviews



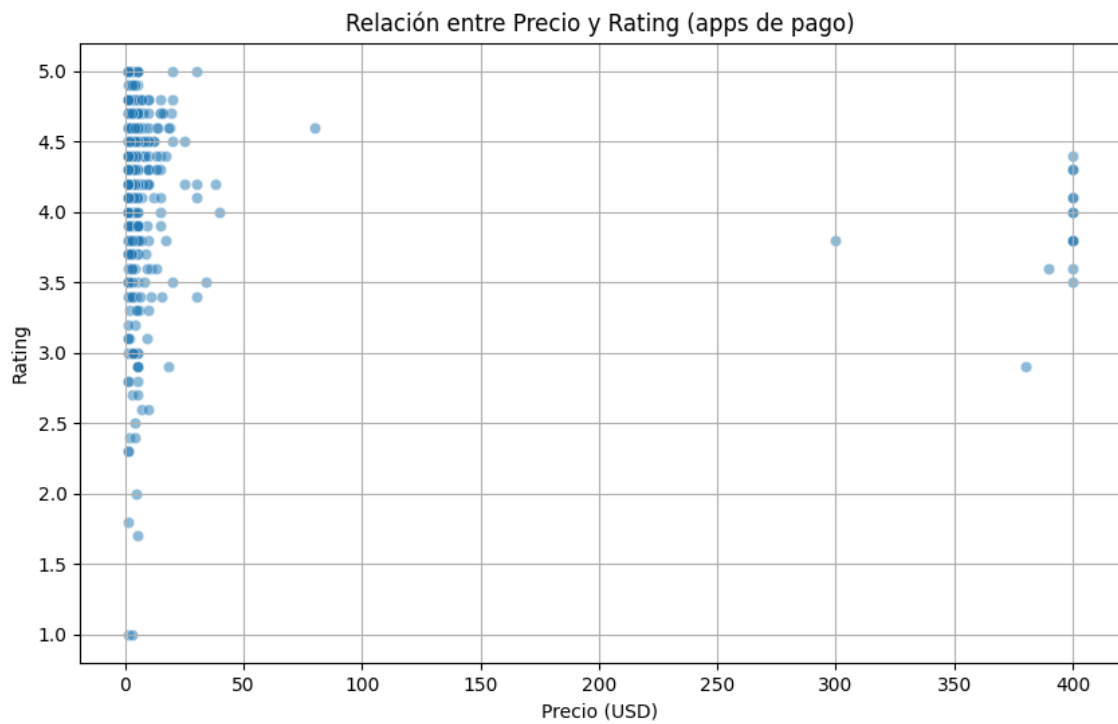
El gráfico muestra una nube densa de puntos en la zona de rating 4–5 y sobre todo a partir de las 100.000 reviews (ya sabíamos que era el rango de instalaciones, y por tanto reseñas y de rating dónde se concentran casi todas las apps). Por debajo de 10000 hay una enorme variabilidad de rating entre ellas.

Las apps con muchas reviews suelen ser más conocidas, con más comunidad, más soporte y mejoras continuas, resolviendo los errores que pueden señalarse en estas mismas reseñas. En cambio, las apps con pocas reseñas son más volátiles (claramente el número de reseñas es proporcional al de instalaciones).

#### 4.4 Rating vs Category



#### 4.5 Rating vs Price



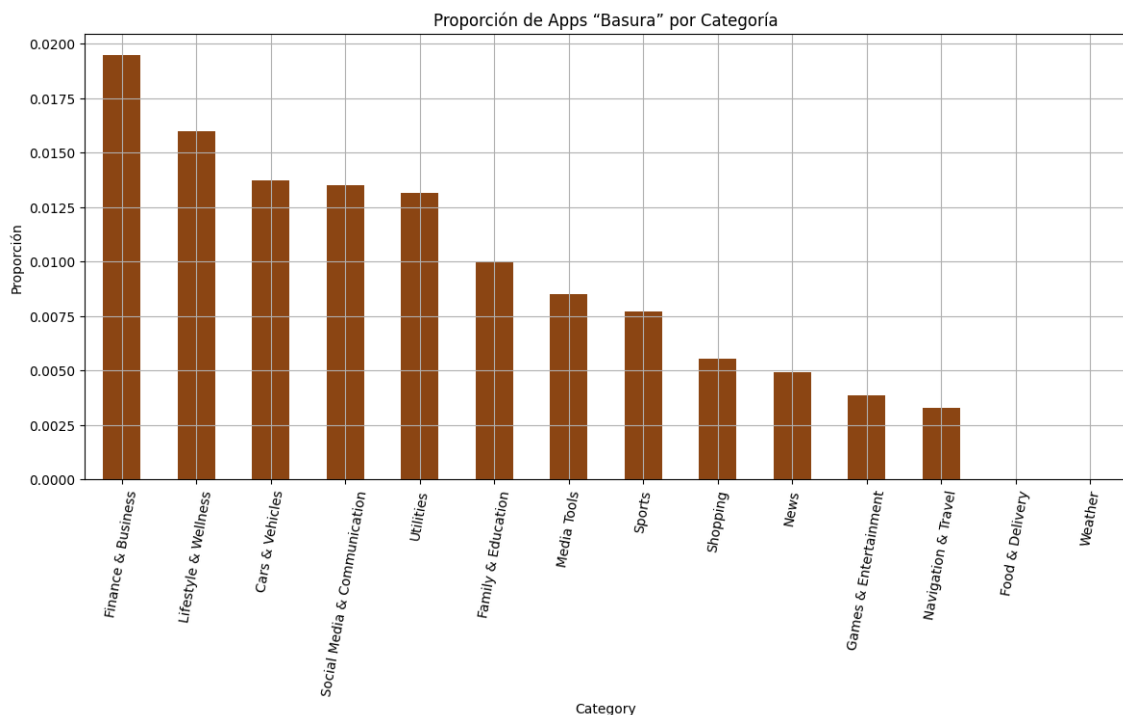
Lo más significativo es que las pocas apps de más de \$100 no tienen mejores ratings, de hecho, bajan un poco en general. Quizás puede ser debido a las altas expectativas: si una app es cara, el usuario espera mucho más (como podría ser lógico) y va a ser más duro en sus críticas y valoraciones en caso de que haya errores o algo no le guste. Además, algunas de las apps más caras tienden a ser de nicho profesional o incluso estafas en algunos casos. Luego, una vez más, queda claro que, en caso de descargar una app de pago, el usuario medio prefiere precios contenidos.

En conclusión, aunque la relación entre precio y valoración es débil, los datos sí que muestran una leve tendencia a que las apps más caras estén de hecho peor valoradas. Es decir, precio no garantiza calidad ni satisfacción, e incluso puede generar frustración si el producto no cumple las expectativas.

## 5. EJEMPLOS Y EXTRAS

En esta sección entraremos a definir y valorar lo que consideramos una aplicación basura, haremos algún que otro análisis centrándonos en ese término, y mostraremos algunos rankings de aplicaciones según diversos criterios que ejemplifiquen de manera mas tangible lo que hemos estado desarrollando:

### 5.1 Junk Apps vs Category



Diremos que una app es “basura” si tiene una valoración media inferior a 3,5 y al menos 1000 instalaciones (para que sea representativo).

Teniendo esto en cuenta, hemos calculado la proporción de *Junk Apps* por categoría y hemos llegado a las siguientes conclusiones:

Finance & Business y Lifestyle & Wellness son las categorías con más proporción de apps basura. Curiosamente, Social Media & Communication también tiene una tasa bastante alta.

En el otro extremo, Weather y Food & Delivery no tienen ningún caso de apps de mala calidad.

Podríamos decir que las apps funcionales y de servicios (finanzas, salud, coches) tienen mayor riesgo de frustrar al usuario, en cambio, aplicaciones de entretenimiento, meteorología, comida o navegación, tienden a cumplir las expectativas, quizás porque

se les da un uso menos intensivo, por ser más simples de desarrollar correctamente o, simplemente, por cubrir una necesidad mucho más concreta.

## 5.2 Rankings

Estas son las 20 aplicaciones que, en 2018, habían superado los mil millones de descargas a nivel global:

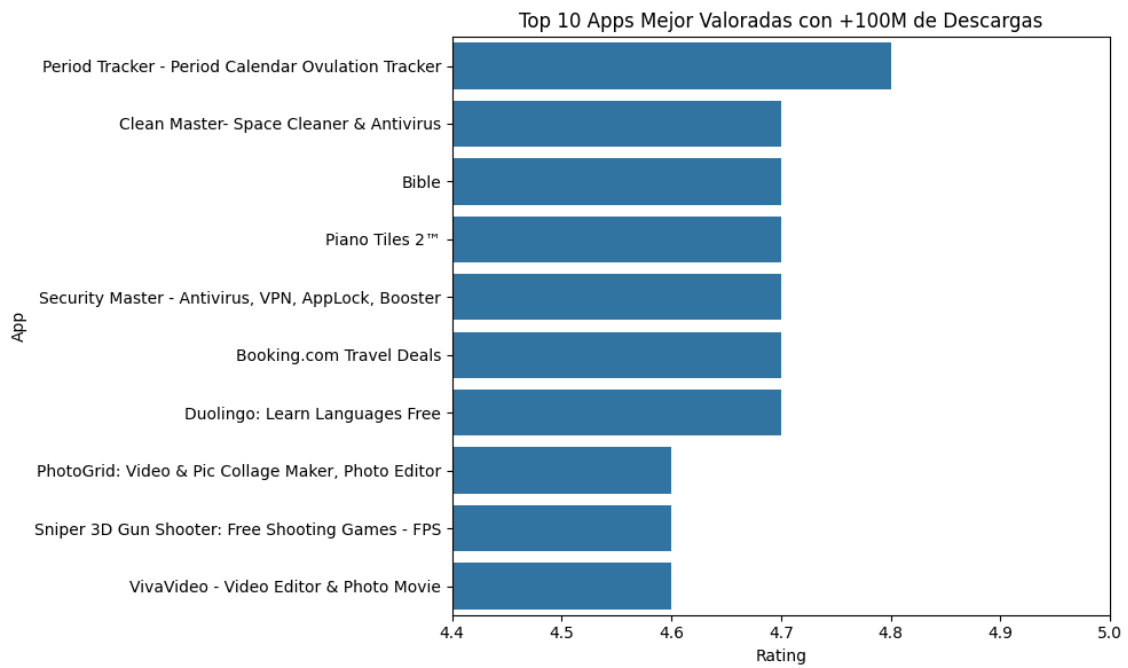
	App	Category	Rating	Reviews	Size	Installs	Type	Price	Is Junk
0	Instagram	Social Media & Communication	4.5	66577446.0	Varies with device	1000000000	Free	0.0	False
1	Subway Surfers	Games & Entertainment	4.5	27725352.0	76M	1000000000	Free	0.0	False
2	Google Photos	Media Tools	4.5	10859051.0	Varies with device	1000000000	Free	0.0	False
3	Google	Utilities	4.4	8033493.0	Varies with device	1000000000	Free	0.0	False
4	WhatsApp Messenger	Social Media & Communication	4.4	69119316.0	Varies with device	1000000000	Free	0.0	False
5	Google Drive	Utilities	4.4	2731211.0	Varies with device	1000000000	Free	0.0	False
6	Google Chrome: Fast & Secure	Social Media & Communication	4.3	9643041.0	Varies with device	1000000000	Free	0.0	False
7	Maps - Navigate & Explore	Navigation & Travel	4.3	9235373.0	Varies with device	1000000000	Free	0.0	False
8	Google Play Games	Family & Education	4.3	7168735.0	Varies with device	1000000000	Free	0.0	False
9	Gmail	Social Media & Communication	4.3	4604483.0	Varies with device	1000000000	Free	0.0	False
10	YouTube	Media Tools	4.3	25655305.0	Varies with device	1000000000	Free	0.0	False
11	Google+	Social Media & Communication	4.2	4831125.0	Varies with device	1000000000	Free	0.0	False
12	Google Street View	Navigation & Travel	4.2	2129707.0	Varies with device	1000000000	Free	0.0	False
13	Facebook	Social Media & Communication	4.1	78158306.0	Varies with device	1000000000	Free	0.0	False
14	Skype - free IM & video calls	Social Media & Communication	4.1	10484169.0	Varies with device	1000000000	Free	0.0	False
15	Messenger - Text and Video Chat for Free	Social Media & Communication	4.0	56646578.0	Varies with device	1000000000	Free	0.0	False
16	Hangouts	Social Media & Communication	4.0	3419513.0	Varies with device	1000000000	Free	0.0	False
17	Google News	News	3.9	878065.0	13M	1000000000	Free	0.0	False
18	Google Play Books	Utilities	3.9	1433233.0	Varies with device	1000000000	Free	0.0	False
19	Google Play Movies & TV	Media Tools	3.7	906384.0	Varies with device	1000000000	Free	0.0	False

Estas son las 20 aplicaciones mejor valoradas que han conseguido superar el umbral del millón de descargas:

	App	Category	Rating	Reviews
232	DU Recorder – Screen Recorder, Video Editor, Live	Media Tools	4.8	2588730.0
453	PewDiePie's Tuber Simulator	Family & Education	4.8	1499466.0
454	Vlogger Go Viral - Tuber Game	Family & Education	4.8	1304467.0
452	Eternium	Family & Education	4.8	1506783.0
455	Amino: Communities and Chats	Social Media & Communication	4.8	1264084.0
44	Period Tracker - Period Calendar Ovulation Tra...	Lifestyle & Wellness	4.8	4559407.0
497	Domino's Pizza USA	Food & Delivery	4.7	1032935.0
46	Duolingo: Learn Languages Free	Family & Education	4.7	6297590.0
236	Video Editor Music,Cut,No Crop	Media Tools	4.7	2163282.0
238	Bubble Witch 3 Saga	Games & Entertainment	4.7	1732263.0
47	Bible	Utilities	4.7	2440695.0
48	Booking.com Travel Deals	Navigation & Travel	4.7	1830388.0
45	Piano Tiles 2™	Games & Entertainment	4.7	8118880.0
21	Security Master - Antivirus, VPN, AppLock, Boo...	Utilities	4.7	24900999.0
20	Clean Master- Space Cleaner & Antivirus	Utilities	4.7	42916526.0
494	Muslim Pro - Prayer Times, Azan, Quran & Qibla	Lifestyle & Wellness	4.7	1133393.0
490	Toon Blast	Games & Entertainment	4.7	1351771.0
242	ZenUI Launcher	Utilities	4.7	1141545.0
241	Brightest Flashlight Free ®	Utilities	4.7	1335799.0
240	Bowmasters	Games & Entertainment	4.7	1536349.0



Por último, las mejores entre las mejores. El top 10 de aplicaciones con más de 100 millones de descargas y, al menos, una valoración media de 4,5:



## 6. CONCLUSIONES

Tras haber explorado en profundidad el dataset de Google Play Store con el que hemos trabajado, podemos responder con bastante contundencia a las hipótesis planteadas:

Las aplicaciones de pago tienden, por lo general, a estar mejor valoradas que las gratuitas, aunque estas representan una proporción bajísima del total de la oferta que encontramos en la Play Store. Sin embargo, una puntuación media mayor no garantiza el éxito, y es que las aplicaciones más populares siguen siendo, casi exclusivamente, gratuitas.

Por otro lado, aunque pueda parecer contradictorio, tener un mayor número de descargas no garantiza una buena valoración. A pesar de que si una app va cogiendo fama y descargas es porque tiene que ser lo suficientemente buena, también hay casos con millones de instalaciones que obtienen puntuaciones bajas, lo cual desmonta la idea de que popularidad implica directamente calidad. Por otro lado, también existen apps con excelente valoración, pero muy poco alcance, probablemente por falta de visibilidad o marketing.

En cuanto a los precios de las aplicaciones de pago, mayor precio no garantiza más descargas ni mayor valoración, más bien al contrario: la correlación entre precio y rating es ligeramente negativa. Esto sugiere que, salvo en nichos muy concretos, precios muy elevados generan expectativas difíciles de satisfacer.

También hemos visto que la categoría de la aplicación también influye. Y es que algunas, como Juegos o Entretenimiento, presentan valoraciones más altas en promedio. Otras, como Finanzas, Salud o Negocios, acumulan más casos de apps mal valoradas, lo que puede deberse, seguramente, a una mayor complejidad técnica. Si una app consigue satisfacer una necesidad concreta, esa app triunfa, cuantas más variables o tecnicismos involucre, más problemas suele dar. No hay mejor ejemplo que las propias apps de banca, o las de la Seguridad Social y demás aplicaciones de gestiones online.

Para concluir, podemos afirmar que los factores que definen el éxito de una app no son solo técnicos. La utilidad, una experiencia de usuario sencilla y cuidada y una buena accesibilidad influyen tanto como su categoría o su precio.

## 7. ANEXO: CLAVES PARA LANZAR UNA APP EXITOSA

### 7.1 Fórmula del éxito al lanzar una app

Aunque no hay una receta infalible, para sacar una aplicación al mercado y que triunfe, las conclusiones obtenidas en este estudio nos dan al menos una serie de requisitos o recomendaciones que convendría tener en cuenta si queremos que nuestra futura aplicación tenga más oportunidades de triunfar:

Por un lado, que tenga una utilidad real y cubra un nicho de mercado que no esté saturado. Las apps más valoradas son aquellas que resuelven una necesidad clara o aportan un valor funcional al usuario desde el primer uso. Además, aunque es más probable que un juego o alguna app de entretenimiento sencilla triunfe, si conseguimos desarrollar un producto de calidad en el nicho de las finanzas, la banca o la productividad, es muy probable que destaquemos por encima del resto.

Por otro lado, debemos tener como prioridad máxima una buena experiencia de usuario: tiempos de carga rápidos, interfaz limpia e intuitiva, fluidez en la navegación y resolución rápida de problemas y bugs.

Además, como en todo negocio que acaba de empezar, cuidar a los primeros usuarios es clave. Sus valoraciones iniciales pueden determinar la trayectoria de la app en el algoritmo de visibilidad y en la percepción general de nuestro producto.

Por último, pero no menos importante, hemos de tener en cuenta que debemos lanzar actualizaciones frecuentes y estar pendientes al mantenimiento de la aplicación: las apps que no se actualizan tienden a perder rating y visibilidad, y corregir errores y añadir mejoras demuestra compromiso con el usuario.

### 7.2 Modelo de monetización: *freemium*, anuncios o pago único

Uno de los aspectos clave en el diseño de una app es cómo monetizarla y sacar rédito de nuestro duro trabajo. Podemos distinguir tres modelos básicos cuyo análisis exhaustivo requeriría un EDA mucho mas extenso que este, pero, a grandes rasgos, tenemos:

- Modelo *freemium*: es el modelo más recomendable para viralizar una app. Consiste en ofrecer acceso gratuito a las funcionalidades básicas, con opciones de pago para desbloquear mejoras o eliminar limitaciones. Permite captar usuarios fácilmente con el gancho de una app gratuita, pero sin renunciar a ganar dinero. Además, si implementamos los micropagos con sentido y cuidado, estos

son una fuente de ingresos muy efectiva. Funcionan especialmente bien en juegos y apps de personalización o que tengan un desarrollo progresivo que llegado cierto momento se vuelve limitado.

- Publicidad: el modelo basado en anuncios puede ser rentable si se evita una intrusividad excesiva. Los usuarios penalizan mucho los anuncios a pantalla completa, con sonido o que interrumpen un uso óptimo de la aplicación. Si se implementen, es mejor optar por banners discretos o publicidad contextual opcional.
- Pago único: como hemos visto en el EDA, son una minoría. Este modelo funciona solo si el valor percibido de la app es muy alto y está claramente justificado. En general, las apps de pago reciben mejores valoraciones, pero tienen una barrera de entrada mucho mayor y menor alcance.