
Recopilación y análisis de datos de las tendencias de la industria de videojuegos a nivel global y local, enfocado en Guatemala

José Rodolfo Pérez García



UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Recopilación y análisis de datos de las tendencias de
la industria de videojuegos a nivel global y local,
enfocado en Guatemala**

Trabajo de graduación en modalidad de Tesis presentado por
José Rodolfo Pérez García
para optar al grado académico de Licenciado en Ingeniería en Ciencia de
la Computación y Tecnologías de la Información

Guatemala,
2021

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Recopilación y análisis de datos de las tendencias de
la industria de videojuegos a nivel global y local,
enfocado en Guatemala**

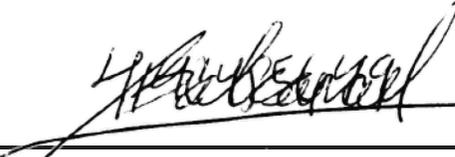
Trabajo de graduación en modalidad de Tesis presentado por
José Rodolfo Pérez García
para optar al grado académico de Licenciado en Ingeniería en Ciencia de
la Computación y Tecnologías de la Información

Guatemala,
2021

Vo.Bo.:

(f) 
Ing. Alhvi Romancina Balcarcel Rodas

Tribunal Examinador:

(f) 
Ing. Alhvi Romancina Balcarcel Rodas

(f) 
Ing. Douglas Barrios

(f) 
Lic. Carmen Matia Lopez Juarez

Fecha de aprobación: Guatemala, 21 de octubre de 2021.

Agradecimientos

Quisiera agradecer a mis padres Brenda García y Rodolfo Pérez por acompañarme durante los momentos más duros de mi carrera y también en los momentos de paz y tranquilidad. Por compartir su tiempo y su infinita paciencia. Agradezco a mi asesora Alhvi Balcarcel por ayudarme en la dirección de este proyecto.

Agradecimientos	III
Lista de figuras	VI
Lista de cuadros	VII
Resumen	VIII
1. Introducción	1
2. Objetivos	2
2.1. Objetivo general	2
2.2. Objetivos específicos	2
3. Justificación	3
4. Marco teórico	4
4.1. Herramientas para recolección de datos	4
4.1.1. Entrevista	4
4.1.2. Grupo focal	4
4.1.3. Encuesta	4
4.1.4. Observación	5
4.1.5. Análisis de contenido	5
4.1.6. Entrevista de profundidad	5
4.2. Tipos de información	5
4.2.1. Información primaria	5
4.2.2. Información secundaria	5
4.3. Jugabilidad	5
4.3.1. Juegos de acción	6
4.3.2. Juegos de aventura	6
4.3.3. Juegos de carreras	6
4.3.4. Juegos de pelea	6
4.3.5. Rompecabezas	6
4.3.6. Juegos de rol	6
4.3.7. Juegos de disparos	6
4.3.8. Juegos de simulación	6
4.3.9. Juegos de deportes	6

4.3.10. Juegos de estrategia	7
4.4. R	7
4.4.1. Definición de R	7
4.4.2. Programa de R	7
4.4.3. Evaluador de R	8
4.4.4. Información en lenguaje S	8
4.5. Python:	8
4.5.1. Definición de Python:	8
4.5.2. Bibliotecas para análisis de datos	8
4.6. Almacenamiento de información	9
4.6.1. Formato JSON	9
4.6.2. Valores separados por comas	9
4.6.3. Bases de datos	10
4.7. Web Scrapping	10
4.7.1. API's	10
4.7.2. Steam API	10
4.8. Limpieza de datos	11
4.8.1. Detección de valores atípicos	11
4.8.2. Deduplicación de datos	11
4.8.3. Limpieza de datos basada en reglas	11
4.9. Predictores del mercado a futuro	12
4.9.1. Google Trends	12
4.9.2. Método de segmentación	12
4.10. Métricas utilizadas en estudios de mercado	12
4.10.1. Tipos de juegos para la industria	12
4.10.2. Análisis de Delphi	13
5. Marco metodológico	14
5.1. Resultados de las entrevistas de profundidad	14
5.2. Perspectivas y opiniones de los entrevistados	17
5.3. Información demográfica de los expertos en videojuegos seleccionados	18
5.4. Cálculo del tamaño de la muestra en investigación	18
5.5. Generación de herramienta de recolección de datos de la muestra	19
5.6. Minado de datos de la plataforma Steam	20
6. Resultados	21
6.1. Descripción de la muestra poblacional, datos generales	21
6.2. Descripción de la muestra poblacional, plataformas usadas	24
6.3. Descripción de la muestra poblacional, dispositivos usados	26
6.4. Población guatemalteca que utiliza la plataforma STEAM	28
6.5. Población global que utiliza la plataforma STEAM	30
7. Discusión	34
7.1. Delphi expertos	34
7.2. Limitaciones del área rural	34
7.3. Generalización de la muestra	35
7.4. Resultados globales, comparación con resultados locales	35
8. Conclusiones	36
9. Recomendaciones	37
10. Bibliografía	38
11. Glosario	40

Lista de figuras

6.1. Descripción de casos segmentados por género seleccionado.	21
6.2. Número de casos del datagrama, segmentados por edad.	22
6.3. Número de casos segmentados por gasto mensual.	22
6.4. Número de casos segmentados por género más jugado.	23
6.5. Matriz de relación entre género más jugado y gasto mensual en videojuegos.	23
6.6. Gasto mensual segmentado por ámbito de más gasto.	23
6.7. Plataformas utilizadas segmentadas por gasto mensual.	24
6.8. Plataformas utilizadas dividido por género.	24
6.9. Plataformas utilizadas dividido por nivel académico.	25
6.10. Plataformas utilizadas dividido por gasto mensual.	25
6.11. Número de casos segmentados por dispositivos más utilizados.	26
6.12. Ámbito de más gasto dividido por dispositivo utilizado.	26
6.13. Género más jugado dividido por dispositivo utilizado.	27
6.14. Gasto mensual dividido por dispositivo utilizado.	27
6.15. Estado de las cuentas de los jugadores que frecuentan la plataforma.	28
6.16. Juegos más utilizados últimamente.	28
6.17. Géneros de juegos utilizados últimamente.	29
6.18. Agrupación de nivel de plataforma.	29
6.19. Estado de las cuentas de los jugadores estadounidenses que frecuentan la plataforma.	30
6.20. Estado de las cuentas de los jugadores franceses que frecuentan la plataforma.	30
6.21. Estado de las cuentas de los jugadores japoneses que frecuentan la plataforma.	31
6.22. Juegos más utilizados últimamente por jugadores estadounidenses.	31
6.23. Juegos más utilizados últimamente por jugadores franceses.	31
6.24. Juegos más utilizados últimamente por jugadores japoneses.	32
6.25. Géneros de juegos utilizados últimamente por jugadores estadounidenses.	32
6.26. Géneros de juegos utilizados últimamente por jugadores franceses.	33
6.27. Géneros de juegos utilizados últimamente por jugadores japoneses.	33

Lista de cuadros

4.1. Tabla con tipos de juego agrupados por género. [2].	7
4.2. Dispositivo de videojuegos, tipo y desarrollador	7
5.1. Entrevista de profundidad 1: Género masculino 20 años	15
5.2. Entrevista de profundidad 2: Género masculino 28 años	15
5.3. Entrevista de profundidad 3: Género masculino 19 años	15
5.4. Entrevista de profundidad 4: Género masculino 18 años	15
5.5. Entrevista de profundidad 5: Género femenino 24 años	15
5.6. Entrevista de profundidad 6: Género femenino 22 años	16
5.7. Entrevista de profundidad 7: Género femenino 21 años	16
5.8. Entrevista de profundidad 8: Género femenino 23 años	16
5.9. Entrevista de profundidad 9: Género masculino 20 años	16
5.10. Entrevista de profundidad 10: Género masculino 28 años	16
5.11. Valores Alpha	19
5.12. Resultados de géneros populares de videojuegos	20

El objetivo general de este proyecto consistió en determinar las tendencias del mercado de videojuegos en Guatemala mediante entrevistas de profundidad sobre población joven y un análisis de Delphi en expertos del área. Se hizo posible mediante múltiples iteraciones con el objetivo de determinar cuáles eran las características principales y hábitos de consumo, que se obtuvieron mediante encuestas de profundidad y minado de datos. El resultado del trabajo fueron gráficos de relación entre variables descriptivas de la muestra, que se puede extender a la población. Se determinaron los dispositivos más populares, los hábitos de consumo monetario, y los géneros más jugados. Entre los descubrimientos se detectó un abandono hacia el dispositivo Xbox, se detectó un declive en el uso de dispositivos móviles vinculados al nivel educativo y se determinó que los géneros aventura y disparos son los más rentables para los desarrolladores. Se recomendó realizar un estudio complementario, que abarque un rango más amplio en edades ya que según lo determinado por los expertos, contiene un porcentaje significativo de la población de estudio.

CAPÍTULO 1

Introducción

El presente trabajo tiene el objetivo de proponer una perspectiva al estado actual del mercado de videojuegos en Guatemala. Dicha población fue representada por jóvenes entre 18 y 29 años de edad, los cuales fueron entrevistados durante múltiples ocasiones para concretar sus hábitos de consumo de videojuegos.

El trabajo se divide en dos tipos de información: la obtenida mediante preguntas a los individuos de estudio y la generada por expertos del tema ubicados dentro de la región.

La perspectiva de los expertos se utilizó para delimitar la población de estudio, generar perspectivas para evaluar a la población y señalar áreas de mejora en la investigación. Se utilizó la metodología Delphi, en la que se entrevistó a los expertos y luego se encuestó para concretar sus aserciones.

El trabajo realizado concretó 10 entrevistas de profundidad para conocer perspectivas directamente de los individuos y generar sus perspectivas. Estas entrevistas se realizaron entre el 3/8/2021 y 29/08/2021. Luego se realizó una encuesta más cerrada con el objetivo de concretar estas perspectivas y analizar la información. Estas encuestas se realizaron entre el 12/10/2021 y el 15/10/2021 y se lograron encuestar 80 personas, de las cuales 72 aplicaron para el estudio.

Se filtró y limpió la información generada por la encuesta, eliminando los casos que no cumplían con el parámetro de edad. Se analizaron las dimensiones generadas por la entrevista y se presentaron los gráficos que definen sus características únicas.

Utilizando la técnica de minado de datos se generó un conjunto de información de usuarios de múltiples regiones de jugadores de videojuegos, entre ellas Estados Unidos, Francia, Japón y Guatemala. Se limpió y filtró la información obtenida y se generaron gráficos con sus hábitos de consumo y utilización de la plataforma.

2.1. Objetivo general

Determinar las tendencias del mercado de videojuegos en Guatemala.

2.2. Objetivos específicos

- Conocer los géneros de juegos más populares en la región.
- Conocer los hábitos de consumo de los jugadores de la región.
- Determinar los perfiles demográficos de los jugadores de la región.
- Recopilar las características de consumo de carácter global para comparar con los resultados locales.

Debido a que la industria de videojuegos en Guatemala es una industria emergente, es necesario un estudio de mercado para tomar decisiones estratégicas, para poder desarrollar productos orientados hacia este mercado. Previamente, esta creación de contenido estaba en las manos de los desarrolladores del hardware utilizado para jugar, pero actualmente este trabajo lo realizan también millones de desarrolladores externos.[17]

Actualmente se clasifica los tipos de videojuegos utilizando los géneros más populares, como acción, estrategia, deportes, aventura y carreras. Y la industria de videojuegos es clasificada por plataforma utilizada para ejecutar los videojuegos, con un crecimiento acelerado asociado a los dispositivos portátiles. [15]

Este proyecto busca ofrecer una herramienta de detección de tendencias de mercado enfocado en la industria local, además de información sobre los perfiles demográficos de la población de manera gratuita, para generar información relevante para la industria de desarrolladores externos que buscan abarcar a la región en sus futuros proyectos.

4.1. Herramientas para recolección de datos

4.1.1. Entrevista

Las entrevistas se usan para recolectar información de individuos utilizando un conjunto de preguntas en un área de interés. La información tomada de entrevistas se utiliza para generar teorías y modelos. Muchas de las preguntas de las entrevistas pueden ser resueltas mediante encuestas, pero la información obtenida mediante la entrevista da como resultado información más profunda, sin embargo la realización de entrevistas es más laboriosa y requiere más tiempo. [11] Entre los tipos de entrevistas se encuentran las entrevistas estructuradas, donde el entrevistador pregunta por preguntas previamente definidas y entrevistas no estructuradas, en donde se permite que ciertas de las preguntas se basen en las respuestas del entrevistado. [10]

4.1.2. Grupo focal

Los grupos focales son usados para obtener información en un ambiente grupal mediante un conjunto de preguntas predeterminadas, con el objetivo de simular conversaciones grupales. Esta herramienta permite a los investigadores capturar reacciones a los puntos de vista de los participantes. El número óptimo de personas en un grupo focal varía entre 8 y 10 personas. [11]

4.1.3. Encuesta

Las encuestas se utilizan para documentar percepciones, actitudes y creencias dentro de un grupo predeterminado de individuos. Las preguntas abiertas generan un amplio espectro de respuestas del individuo. [11]

La forma en que se analiza la información de la encuesta dicta el tipo de encuesta que se usa en la investigación. Las preguntas cerradas proveen información fácil de codificar, mientras que las preguntas abiertas dan más libertad para responder, pero la información es más difícil de codificar. [10]

4.1.4. Observación

Las observaciones se usan para obtener información mediante los sentidos. El objetivo es comprender qué hacen y por qué lo hacen en lugar de enfocarse en sus propias percepciones. Son ideales para documentar, explorar y entender fenómenos. [11]

Los estudios observacionales ayudan a comprender problemas complejos, pero son caros y requieren largos períodos de observación, además de presentar sesgo por parte del observador. [10]

4.1.5. Análisis de contenido

Se usa para investigar cambios en un tópico en específico. También como el método principal en una investigación para contextualizar resultados de otros estudios. El número de documentos que se utilizan como base de investigación y el tipo de los documentos son dictados por la pregunta de investigación. [11]

4.1.6. Entrevista de profundidad

A diferencia de las entrevistas comunes, las entrevistas de profundidad carecen de preguntas predeterminadas y se asemejan a las conversaciones más casuales. El entrevistador no busca utilizar un conjunto de preguntas sino que busca entablar una conversación sobre un tema determinado con el entrevistado. Dentro de las aproximaciones para desarrollar este tipo de recopilación de información cualitativa se encuentran diferentes pasos, por ejemplo desarrollar una lista de temas donde realizar preguntas, indagar en temas de interés y grabado de las respuestas para su posterior análisis. [?]

4.2. Tipos de información

4.2.1. Información primaria

Definida como información colectada directamente de la fuente que genera la información. Originalmente mediante individuos utilizando grupos de enfoque, paneles de correspondientes creados por el investigador con el objetivo de buscar información sobre un problema. [10]

4.2.2. Información secundaria

Un proyecto con información secundaria utiliza información que ya fue previamente recogida. En esta categoría se encuentran bases de datos, archivos de empresas, publicaciones del gobierno y análisis de las redes sociales. [10]

4.3. Jugabilidad

Se define jugabilidad como la naturaleza de la experiencia del jugador definido por el patrón de interacciones y de reglas del juego además de los desafíos y la manera en que los superan. Está vinculado a las mecánicas del juego y a la experiencia del jugador, junto al diseño del juego. [2]

4.3.1. Juegos de acción

Juegos con énfasis profundo en series de acciones realizadas por el jugador para alcanzar una serie de objetivos. [2]

4.3.2. Juegos de aventura

Juegos localizados en un mundo donde el jugador puede explorar y completar una serie de objetivos a través de una serie de acciones. [2]

4.3.3. Juegos de carreras

Juegos que involucran conducción de vehículos como acción principal, en algunos casos con el objetivo de ganar carreras contra oponentes. [2]

4.3.4. Juegos de pelea

Juegos que involucran al jugador con un personaje controlable para enfrascarse en combate contra oponentes. [2]

4.3.5. Rompecabezas

Juegos con el objetivo de solucionar enigmas, navegación, manipulación y reconfiguración de objetos para alcanzar objetivos. [2]

4.3.6. Juegos de rol

Juegos con énfasis en el desarrollo del personaje jugable y los componentes narrativos. [2]

4.3.7. Juegos de disparos

Juegos que involucran disparar a objetivos u oponentes usualmente para su destrucción. [2]

4.3.8. Juegos de simulación

Juegos que buscan recrear una experiencia de una actividad del mundo real en el mundo de los videojuegos. [2]

4.3.9. Juegos de deportes

Juegos que presentan una simulación de juegos físicos en el mundo de los videojuegos. [2]

4.3.10. Juegos de estrategia

Juegos caracterizados por las decisiones estratégicas de los jugadores y sus intervenciones para obtener un resultado deseado. [2]

Categoría de juego	Ejemplos
Acción	Super Mario Bros, Metal Gear
Aventura	The Legend of Zelda, Prince of Persia
Juegos de carreras	Mario Kart, Grand Turismo
Juegos de pelea	Street Fighter, Mortal Kombat
Rompecabezas	Tetris, Buscaminas
Juegos de rol	Baldour's Gate, Final Fantasy, World of Warcraft
Juegos de disparos	Doom, Duck Hunt
Juegos de simulación	Sims, Grand Turismo
Juegos de deportes	FIFA; Wii Sports
Juegos de estrategia	Civilization, Command and Conquer

Tabla 4.1: Tabla con tipos de juego agrupados por género. [2].

Nombre de dispositivo	Tipo de dispositivo	Empresa desarrolladora
Computadora (PC)	Fijo y portátil	No aplica
Nintendo Switch	Fijo y portátil	Nintendo
Playstation	Fijo	Sony
Xbox	Fijo	Microsoft
Teléfono	Portátil	No aplica
Oculus	Fijo	Facebook
Sega Genesis	Fijo	SEGA
Gameboy	Portátil	Nintendo

Tabla 4.2: Dispositivo de videojuegos, tipo y desarrollador

4.4. R

4.4.1. Definición de R

R es un software de código abierto respaldado y mantenido por un grupo de voluntarios de múltiples países. El sistema base provee un lenguaje interactivo para computación numérica, manejo de información, graficación y calculaciones relacionadas. La mayoría del software de R está escrito en el mismo lenguaje, un dialecto del lenguaje S. [3]

La unidad básica de programación es la función, y las llamadas a Funcion:es usualmente introducen un objeto que es una función de objetos como argumentos. Otras iteraciones de R introdujeron métodos y clases. [3]

4.4.2. Programa de R

Para la utilización de R, inicialmente se invocaba desde una consola interactiva y el usuario introducía expresiones de R o comandos del programa. Actualmente se utilizan consolas interactivas pero el corazón de estas consolas aún tienen actualmente una ventana para evaluar y componer

expresiones de R. El proceso de evaluación de expresiones se basa en el ciclo de lectura, análisis y evaluación de cada cadena del lenguaje. [3]

El concepto central del lenguaje S es que cada una de las expresiones que se están evaluando es un objeto. Esto permite a las computaciones en R la generación y modificación de expresiones para luego evaluar el resultado de las expresiones. [3]

4.4.3. Evaluador de R

La evaluación de expresiones en el programa R consiste en la obtención de un objeto producto del análisis de la expresión sobre el programa de R. El evaluador no es parte del programa gráfico, pero existe como la función `eval()`. El mismo se encarga de calificar dentro de las definiciones del lenguaje, lo que significa que se están evaluando llamadas a Funcion:es. Ya que casi todas las expresiones del lenguaje son similares a llamadas de Funcion:es se pueden calificar dentro del sistema. [3]

4.4.4. Información en lenguaje S

El lenguaje S fue creado en 1976 y ha pasado por múltiples reestructuraciones en las técnicas de representación de información y estructuras de datos. Se divide en cuatro épocas, las cuales perduraron y aportan diferentes fortalezas al lenguaje hasta el día de hoy.

Tipos de objetos: Se refieren a los tipos definidos en las implementaciones de C, originalmente llamados modos de S.

Estructuras vectoriales: Se refieren a objetos indexables de tipo vectorial con atributos.

Clases S3: Se refiere a los objetos con definiciones de atributos que no tienen definiciones formales de sus clases.

Clases formales: Se refiere a los objetos con definición de clases y que corresponden a Funcion:es genéricas y métodos generales. [3]

4.5. Python:

4.5.1. Definición de Python:

Python: aparece en 1991 como un lenguaje de programación tipo intérprete con una larga comunidad científica computacional. Con un incremento significativo de usuarios desde el año 2000. Entre sus aplicaciones se encuentran: análisis de datos, visualización y computación exploratoria. Debido a la fortaleza de las Bibliotecas de análisis de datos creadas para este lenguaje como pandas, se ha posicionado como una de las mejores opciones para el desarrollo de aplicaciones centradas en manipulación de datos. [7]

4.5.2. Bibliotecas para análisis de datos

NumPy

Es la abreviatura de Numerical Python: una Biblioteca fundamental para computación científica en Python. Provee entre otras Funcion:es: Arreglo:s multidimensionales, Funcion:es operando sobre

los objetos de los Arreglo:s, herramientas de escritura y lectura al disco, operaciones de álgebra lineal y conexión con otros lenguajes de programación como C, C++ y Fortran. [7]

Debido a la velocidad de procesamiento de las operaciones sobre Arreglo:s de NumPy hacen que Python: se convierta en uno de los principales lenguajes para manipulación de datos intermedios sin necesidad de copiar la información durante las manipulaciones. [7]

pandas

Esta Biblioteca: de Python: provee estructuras de datos con Funcion:es para diseñadas para su manipulación con el objetivo de análisis de datos rápido y eficiente. La estructura de datos más utilizada de la Biblioteca: es el DataFrame, un objeto bidimensional con filas y columnas similar a una matriz. Combinado con la capacidad de manipulación de bases de datos Relacional:es como SQL provee indexación, capacidad de modificación, agregación y segregación. [7]

matplotlib

Es una de las Biblioteca: más populares de Python: para producir gráficos y representaciones bidimensionales de información. Bajo continuo mantenimiento por una comunidad voluntaria de programadores. Provee un entorno interactivo durante la exploración de la información. [7]

Scipy

Scipy es la colección de múltiples paquetes de Python:, es el paquete estándar de computación científica que incluye solucionadores de ecuaciones, optimizadores de Funcion:es, herramientas de proceso de señales, distribuciones de probabilidad estándar y herramientas de uso de C + + en computaciones de Arreglo:s. [7]

4.6. Almacenamiento de información

4.6.1. Formato JSON

El formato JSON, una abreviatura de *JavaScript Object Notation*, debido a su velocidad en serialización y deserialización de datos es bastante utilizado como contenedor en análisis de datos. Es más flexible que la información en formato tabular. Es bastante utilizado en Python: ya que comparten tipos con los objetos de Python: a excepción del tipo null. [7]

4.6.2. Valores separados por comas

Este tipo de archivos delimita los valores dentro del archivo mediante comas. Cada línea es un registro de datos con uno o más valores separados por comas. Usualmente al procesarlos se transforman en la estructura de datos utilizada, como el Data Frame, que permite indexación, inferencia de tipos, conversiones, iteratividad y limpieza de datos. [7]

4.6.3. Bases de datos

Debido a que el proceso de la información proveniente de archivos de texto es ineficiente para mucha información se utilizan bases de datos Relacionales como SQL Server, PostgreSQL y MySQL. Existen algunas alternativas como NoSQL que no son Relacionales. [7]

4.7. Web Scrapping

Se define como el proceso de extracción y combinación de los contenidos de un sitio Web de una forma sistemática. En el proceso el programa de automatización busca repetir la interacción entre el servidor y el humano paso a paso, con el objetivo de extraer los contenidos deseados. [4]

Se divide en acceso al sitio, donde el programa accede el contenido del sitio web mediante el protocolo HTTP que coordina la comunicación entre el servidor y el cliente mediante una interfaz como un navegador web. En estos métodos se utilizan solicitudes HTTP como GET y POST. El siguiente paso busca recuperar los documentos y los contenidos de interés de las respuestas HTTP. Para este proceso se requieren bibliotecas de análisis de HTML. El último paso se enfoca en estructurar y transformar la información para representaciones más apropiadas para el análisis de datos. Con el producto final en archivos tipo XML o CSV. [4]

4.7.1. API's

Es el acrónimo de interfaz de aplicación de programación, por sus siglas en inglés. Son el mecanismo clave para acceder a servicios por internet. Las interfaces que aplican este sistema proveen un mejor servicio con un costo reducido a los usuarios. Se siguen múltiples protocolos al momento de implementar este sistema, como protocolos SOAP, API's REST, entre otros. Se utilizan para conectar secciones de plataformas independientes. [13]

La documentación de los API web es la fuente principal de información para los desarrolladores e integradores de la interfaz en distintas plataformas. La comunicación y las listas de cambios y actualizaciones son clave para la evolución e implementación de este sistema, por lo que es necesario una página de aterrizaje, configuración de encabezados que indiquen el tipo de petición y autenticación para poder realizar estas peticiones. [13]

4.7.2. Steam API

Steam es la tienda de videojuegos más popular para computadoras. SteamAPI es la herramienta de conexión a la base de datos e infraestructura de la tienda para desarrolladores de conectores y herramientas de comunicación a la tienda.

El Steam API se refiere a la interfaz de desarrollador que provee Valve a como método de acceso a la información disponible en sus servidores. Permite a los desarrolladores consultar Steam para información que puedan presentar en sus propios sitios web. Las respuestas varían entre JSON, XML y VDF. [1]

Se divide en: Noticias de Steam, Estadísticas de usuario, características de usuario y artículos del usuario. Se puede acceder mediante una llave de desarrollador de Steam, después de verificación de la cuenta y llenado de un formulario donde se aceptan los términos de uso de la información. [1]

4.8. Limpieza de datos

4.8.1. Detección de valores atípicos

La detección de valores atípicos a menudo se enfoca en anomalías de datos con respecto a la definición de los valores normales para la implementación. Puede ayudar a detectar posibles intrusiones y actividades maliciosas dentro de la información. [5]

Entre los desafíos de la detección de valores atípicos se encuentra la búsqueda del patrón normal. Además de esto, muchas técnicas de detección de valores atípicos pierden efectividad cuando el número de atributos en el conjunto de datos es extensa. [5]

4.8.2. Deduplicación de datos

La duplicación de datos puede ocurrir por muchas razones, como por ejemplo tener múltiples representaciones después de integración de datos o ingreso de datos repetidos. La deduplicación de datos se refiere a al proceso de identificación de tuplas que se refieren a la misma entidad del mundo real. Usualmente seguido de una consolidación de entidades para unificar las representaciones duplicadas para que representen de mejor manera a la entidad del mundo real. [5]

Transformación de datos

La transformación de datos se refiere a la preparación de la información mediante a programas definidos que buscan convertir los datos en información que se adecue a un formato o estructura. Se dividen en dos categorías, transformaciones sintácticas y transformaciones semánticas. [5]

Transformaciones sintácticas

Tienen como objetivo transformar una tabla de información desplegada de una forma en otra representación diferente, usualmente no requieren conocimientos previos de la información que se está transformando. [5]

Transformaciones semánticas

Este tipo de transformaciones requieren conocimientos de la información que se transforma, ya que requiere entendimiento del significado de la información. Este tipo de transformación requiere referenciar fuentes externas de información. [5]

4.8.3. Limpieza de datos basada en reglas

La limpieza de datos basada en reglas son técnicas de estructuración utilizadas para la limpieza de una base de datos sucia, utilizando el esquema de la base de datos y el conjunto de reglas por las que se rige. Entre las técnicas utilizadas para este tipo de detección se encuentran: detección de violaciones y reparación de errores. [5]

4.9. Predictores del mercado a futuro

4.9.1. Google Trends

Entre las herramientas de investigación de tendencias de mercado centrado en un área se encuentra Google Trends. Esta herramienta permite a los investigadores rastrear múltiples términos y frases utilizadas en el servicio de búsqueda de Google. Se tiene acceso relativo a la información sobre el tráfico de usuarios desde el año 2004, con restricción de términos de búsqueda menores a 100 caracteres. [6]

Entre los estudios más reconocidos utilizando la herramienta se encuentra el rastreo del desarrollo de influenza basado en el historial de búsqueda por región. Este estudio, uno de los más citados sobre la herramienta, demostró su utilidad para detección de fenómenos. Probando ser una herramienta excelente para monitoreo de eventos por internet. [6]

4.9.2. Método de segmentación

Entre los métodos de estratificación de resultados de encuestas sobre el uso y apoyo de los usuarios hacia las empresas que manufacturan consolas de videojuegos se encuentran tres métodos de segmentación. Originalmente creados por Gameage RI Co., Ltd. para sus encuestas sobre la población Japonesa. Estos estratos se crearon como ejes para determinar puntos de mejora y fortalezas en la penetración de los dispositivos en el mercado Japonés. [12]

Segmento de energía de innovación (IPS): Se refiere a un índice de segmentación de usuarios exclusivo para consolas de videojuegos. [12]

Segmento de comportamiento móvil (MBS): Se refiere a un índice de segmentación de usuarios exclusivo para dispositivos portátiles. [12]

Medición de lealtad de a la marca (BLM): Se refiere a una herramienta de segmentación para cada marca y el manejo de cada uno de sus productos. [12]

4.10. Métricas utilizadas en estudios de mercado

4.10.1. Tipos de juegos para la industria

En la industria de videojuegos se encuentran tres grandes ramas. La primera, videojuegos de computadora personal. La segunda rama son los videojuegos para consola, que se definen como los videojuegos solamente disponibles en dispositivos específicos como los producidos por Nintendo, Sony o Microsoft. La tercera y última rama son los videojuegos disponibles en dispositivos móviles, especialmente los desarrollados con fines comunicativos. Entre sus características se encuentran los modelos “freemium”, gacha y pago por objetos dentro del juego. [16]

Entre las conclusiones del estudio de mercado de videojuegos en Japón, entre los años 2014 y 2015, se descubrió que los videojuegos en dispositivos móviles son pioneros en un mercado nuevo que coexiste con el mercado de videojuegos para consola y computadora personal. [16]

4.10.2. Análisis de Delphi

El análisis de Delphi es un método utilizado por su confiabilidad para la exploración de ideas mediante expertos independientes con el objetivo de generar argumentos para la toma de decisiones. [8]

El objetivo del método de Delphi es generar una técnica de pronóstico crítico mediante un intercambio de opiniones informadas que generan un consenso sobre el futuro más probable mediante iteración. [8]

Las ventajas de utilizar este método que inicialmente fue desarrollado para capturar el conocimiento y la experiencia en un campo relativamente nuevo incluye que ha demostrado que es útil para resolver problemas de una sola dimensión y mediante el soporte de información de otras fuentes puede ser utilizada para resolver problemas complejos. [8]

Marco metodológico

Se utilizó el método Delphi para obtener información de fuentes expertas en el mercado de videojuegos de Guatemala. Luego se generó una herramienta de recolección de información enfocada en la población objetivo.

La herramienta de recolección de información se hizo mediante dos partes. La primera parte implicó entrevistas de profundidad sobre representantes de la población de estudio, jugadores de videojuegos entre 18 y 29 años de edad, de manera aleatoria. La segunda parte implicó la generación de una encuesta con preguntas cerradas, producto de los resultados de las entrevistas de profundidad.

Se realizó un minado de información de la plataforma Steam, de videojuegos para computadoras. El objetivo de este minado de información fue generar una estructura gráfica para comparar las distintas regiones y sus hábitos de consumo de videojuegos.

5.1. Resultados de las entrevistas de profundidad

Se realizaron 10 entrevistas de profundidad sobre individuos que pertenecían, al momento del estudio, a la población de jugadores de videojuegos en el área de Guatemala, entre 18 y 29 años.

El resultado de estas entrevistas, a 6 personas que se identificaron con el género masculino y 4 con el género femenino se tabularon e identificaron ciertas perspectivas que ayudaron a generar una herramienta más precisa para identificación masiva de características relevantes de la población.

Géneros populares	Shooters, deportes
Ejemplos	League of legends, FIFA, Call of Duty
Dispositivo	Computadora
Gasto / Consumo	No microtransacciones, compra directa de videojuegos
Hábitos temporales	Varía, entre 1 y 8 horas
Videojuegos móviles	No juega
Nivel de estudios	Secundaria

Tabla 5.1: Entrevista de profundidad 1: Género masculino 20 años

Géneros populares	Deportes
Ejemplos	FIFA, NBA, NFL, League of legends
Dispositivo	Computadora
Gasto / Consumo	Compra videojuegos
Hábitos temporales	Entre 5 y 10 horas
Videojuegos móviles	No juega
Nivel de estudios	Universitario

Tabla 5.2: Entrevista de profundidad 2: Género masculino 28 años

Géneros populares	Estrategia
Ejemplos	Crash Royale, League of legends, Brawlhalla
Dispositivo	Computadora
Gasto / Consumo	Paga suscripciones, no compra videojuegos
Hábitos temporales	Hasta 12 horas por semana
Videojuegos móviles	Si juega
Nivel de estudios	Secundaria

Tabla 5.3: Entrevista de profundidad 3: Género masculino 19 años

Géneros populares	Disparos, acción, estrategia
Ejemplos	Valorant, League of legends, DOTA
Dispositivo	Computadora
Gasto / Consumo	Paga suscripciones, compra videojuegos
Hábitos temporales	14 horas a la semana
Videojuegos móviles	No juega
Nivel de estudios	Secundaria

Tabla 5.4: Entrevista de profundidad 4: Género masculino 18 años

Géneros populares	Mundo abierto, aventura
Ejemplos	Legend of zelda, Spiderman, Last of us
Dispositivo	Nintendo Switch, DS
Gasto / Consumo	Compra videojuegos, paga suscripciones
Hábitos temporales	1 - 2 horas
Videojuegos móviles	No juega
Nivel de estudios	Universitario

Tabla 5.5: Entrevista de profundidad 5: Género femenino 24 años

Géneros populares	MMORPG, Battle Royale, Estrategia, Aventura
Ejemplos	League of Legends, Fortnite, PUBG, Brawlhalla
Dispositivo	Teléfono, Computadora, Nintendo Switch
Gasto / Consumo	Compra videojuegos, no paga suscripciones
Hábitos temporales	Hasta 2 horas a la semana
Videojuegos móviles	Si juega
Nivel de estudios	Universitario

Tabla 5.6: Entrevista de profundidad 6: Género femenino 22 años

Géneros populares	Acción
Ejemplos	Genshin Impact, League of legends, Valorant
Dispositivo	Computadora
Gasto / Consumo	No compra videojuegos, no paga suscripciones
Hábitos temporales	7 a 10 horas a la semana
Videojuegos móviles	No juega
Nivel de estudios	Secundaria

Tabla 5.7: Entrevista de profundidad 7: Género femenino 21 años

Géneros populares	Acción
Ejemplos	Fortnite, Free Fire, Mortal Kombat
Dispositivo	Nintendo Switch
Gasto / Consumo	Compra videojuegos
Hábitos temporales	4 horas a la semana
Videojuegos móviles	No juega
Nivel de estudios	Universitario

Tabla 5.8: Entrevista de profundidad 8: Género femenino 23 años

Géneros populares	Disparos, deportes
Ejemplos	Call of Duty, FIFA, The Last of Us
Dispositivo	Playstation 4, Playstation 5, Teléfono
Gasto / Consumo	Compra Videojuegos y paga suscripciones
Hábitos temporales	7 horas a la semana
Videojuegos móviles	No juega
Nivel de estudios	Secundaria

Tabla 5.9: Entrevista de profundidad 9: Género masculino 20 años

Géneros populares	Disparos, deportes
Ejemplos	Fortnite, Minecraft, Free Fire
Dispositivo	Computadora
Gasto / Consumo	Compra videojuegos y paga suscripciones
Hábitos temporales	10 horas semanalmente
Videojuegos móviles	No juega
Nivel de estudios	Secundaria

Tabla 5.10: Entrevista de profundidad 10: Género masculino 28 años

5.2. Perspectivas y opiniones de los entrevistados

- “Las microtransacciones son las que generan dinero” - masculino, 20 años
- “Los más comunes son los niños entre 5 y 13 años son los jugadores móviles” - masculino, 20 años
- “En cualquier teléfono simple se puede jugar cualquier juego” - masculino, 20 años
- “Los adolescentes y adultos jóvenes entre 20 y 30 juegan juegos móviles” - masculino, 28 años
- “Juego juegos que no utilicen internet, para hacerlos más accesibles” - masculino, 19 años
- “No conozco a mucha gente de más de 25 que jueguen juegos móviles” - masculino, 18 años
- “No conozco mucha gente que juega videojuegos, solamente opinaré sobre lo que me gusta a mi - femenino, 24 años
- “Lo estereotipado sería decir que las mamás son las que juegan juegos móviles” - femenino, 24 años
- “Rara vez compro un videojuego, y si lo compro esta en oferta” - femenino, 22 años
- “Me da más seguridad tener la copia de videojuego yo” - femenino, 22 años
- “De mi círculo nadie paga suscripción” - femenino, 22 años
- “Los videojuegos móviles son para gente más adulta, juegos no en línea son para niños” - femenino, 22 años
- “Compra cosméticos en lugar de comprar videojuegos” - femenino, 21 años
- “Las personas que se mantienen menos en su casa son los jugadores móviles” - femenino, 21 años
- “He visto paquetes de internet para jugar videojuegos” - femenino, 23 años
- “Tengo Mac y eso no me ayuda mucho” - femenino, 23 años
- “He visto personas obsesionadas que juegan 8 horas diarias en juegos móviles” - femenino, 23 años
- “Procuro no pasar demasiado tiempo jugando videojuegos, aunque si me gusta” - masculino, 20 años
- “Con los asuntos que tengo que hacer a lo largo del día, no tengo el tiempo para jugar” - masculino, 20 años
- “Normalmente le hacen tiempo para poder jugar, normalmente cuando se juega con amigos” - masculino, 27 años
- “Ahora que los juegos son portados a celular es más sencillo jugar en su celular” - masculino, 27 años

5.3. Información demográfica de los expertos en videojuegos seleccionados

Se encuestó a 6 expertos de videojuegos del área de Guatemala. La edad de los expertos seleccionados para liderar la primera ronda de Delphi se encuentra en promedio de 33 años, teniendo el más joven 24 años y el mayor 35 respectivamente. Sus niveles educativos están divididos en: 3 personas con título universitario, 2 con título diversificado y uno con título técnico. Tienen en conjunto 45 años de experiencia en desarrollo de videojuegos, programación y creación de recursos 3D. Además de aportar 12 años en conjunto mediante la docencia de nuevos programadores, diseñadores y emprendedores del área.

Entre las preguntas principales se les informó que la población de estudio, jóvenes guatemaltecos entre 18 y 29 años con un estimado de 3.9 millones de personas. [9] Ya con esta información se les pidió que estimaran la cantidad de jóvenes pertenecientes a esta población que jugaran videojuegos frecuentemente. El consenso fue que 61 % de los jóvenes pertenece a este subgrupo, el cual se utilizó para calcular el número de de entrevistas para alcanzar 90 % de confianza y 10 % de margen de error.

Otra pregunta que se realizó sobre los expertos se enfoca en que otros segmentos son importantes para un estudio similar a este, se llegó al consenso de que otra población importante será la ubicada entre los 30 y los 40 años, con razones como capacidad de pago e inclusión al mercado como primera generación en la capacidad de jugar videojuegos. Una consideración que se tomó en cuenta es su opinión sobre el género de videojuegos más jugado por los jóvenes, siendo Acción el género más votado por los expertos y seguido por un empate entre Aventura, Disparos, Deportes y Estrategia/Rompecabezas.

En el área de dispositivos utilizados para jugar videojuegos el consenso unánime es que los jóvenes utilizan en alguna capacidad las consolas móviles, entre las cuales se encuentran los teléfonos celulares y otros dispositivos como el Nintendo Switch. Un porcentaje bastante interesante es la cantidad de personas que compran videojuegos, según los expertos menos del 20 % de los jóvenes guatemaltecos entre 18 y 29 años pagan por videojuegos. De esta misma manera argumentan que más del doble de los mismos, un 42 % incurre a la piratería como medio de adquisición de nuevos títulos.

La opinión de los expertos en el ámbito temporal nos indica que se cree que los jóvenes pasan en promedio entre 2 horas y media y 3 horas y media diariamente jugando videojuegos.

5.4. Cálculo del tamaño de la muestra en investigación

Debido a que el estudio realizado fue de carácter mixto, es decir con características cualitativas y cuantitativas y de tipo binario, las preguntas que se presentaron requerían la selección única de una opción entre un set limitado a dos opciones. Gracias a esto se pudieron tabular y analizar posteriormente.

Valores Alpha de significancia estadística:

Alpha	Z/2
0.20	1.28
0.15	1.44
0.10	1.64
0.05	1.96
0.01	2.58
0.001	3.29

Tabla 5.11: Valores Alpha

Se decidió utilizar una distribución cuadrada, con un alpha de 0.1, es decir que el nivel de confianza buscado por los resultados es del 90 % y el valor $Z * \alpha / 2$ utilizado fue de 1.64.

Una meta común de un estudio de investigación es la obtención de información representativa de una población. El investigador luego utiliza la información de la encuesta para hacer generalizaciones sobre la población, dentro de los rangos de error previamente establecidos. Una regla general es que el error aceptable para un estudio basado en encuestas está entre el 5 y el 10 %.[14] El error aceptable que se utilizó en esta investigación fue del 10 %.

Ecuación de tamaño de muestra aleatoria:

$$\frac{Z^2 P(1 - P)}{E^2} / (1 + \frac{Z^2 P(1 - P)}{E^2 N})$$

Donde el valor Z es la significancia estadística seleccionada, en este caso 1.64. E es el error máximo tolerado en este estudio, en este caso 10 %. P es la prevalencia del evento de interés del estudio, en este caso utilizando los resultados del análisis de Delphi aplicado sobre la población experta se determinó un 61 % de que un joven entre 18-29 años se encuentre dentro de la población de jugadores de videojuegos. Para el objetivo de este estudio se utilizó un valor P de 0.61. El valor N es el tamaño de la población, en este caso 3.9M de jóvenes. [9]

$$\frac{1.64^2 0.61(1 - 0.61)}{0.1^2} / (1 + \frac{1.64^2 0.61(1 - 0.61)}{0.1^2 3900000}) \approx 64$$

Por lo mismo se determinó que el tamaño de la muestra debía de ser mayor o igual a 64 personas.

5.5. Generación de herramienta de recolección de datos de la muestra

Entre las características obtenidas de las entrevistas de profundidad se encontraron ciertos géneros de videojuegos que los entrevistados repetían regularmente, debido a que alguno de sus conocidos o hasta los mismos entrevistados los jugaban regularmente. Entre los géneros que se mencionaron se encontraban:

Género	Número de menciones
Disparos	4
Deportes	4
Estrategia	3
Acción	3
Aventura	2
Mundo abierto	1
Battle Royale	1
MMORPG	1

Tabla 5.12: Resultados de géneros populares de videojuegos

Se encuestó a 80 personas con la herramienta de recolección de datos final. Este número se redujo a 72 individuos ya que 8 no cumplían con la edad para pertenecer a la población de estudio. Se eliminaron los duplicados de las respuestas de opción múltiple con capacidad de respuesta abierta. Y se generaron dimensiones de datos vinculados a un identificador único de caso. Este identificador único de caso fue generado aleatoriamente y vincula la base de datos de respuestas con la dimensión de individuos mediante una relación unidireccional (identificador a respuestas).

Se utilizó la herramienta de Python: para la limpieza de datos debido a su velocidad y eficiencia utilizando las bibliotecas de análisis de datos mencionadas previamente. Además de ese factor la interfaz gráfica utilizada para representar los resultados permite su utilización y ciertos análisis dentro de la misma lo requieren. La herramienta visual utilizada en este paso fue Power Bi Desktop, debido a que presenta la funcionalidad de enlazar múltiples dimensiones de datos utilizando relaciones presentes en bases de datos relacionales.

5.6. Minado de datos de la plataforma Steam

Se obtuvo una lista de usuarios de distintos países utilizando la características de grupos virtuales dentro de la plataforma Steam. Se identificaron los grupos correspondientes a los usuarios de distintas regiones como Guatemala, Estados Unidos, Francia y Japón, y se generó una lista de usuarios y sus identificadores únicos.

En el proceso de minado de datos se programó un autómata web de raspado de información. Este programa generó una base de datos de códigos únicos de identificación de los usuarios de la plataforma. La herramienta que se utilizó fue webscraper.io en la que se generó un mapa de selectores de información y se utilizó para cada una de las regiones seleccionadas.

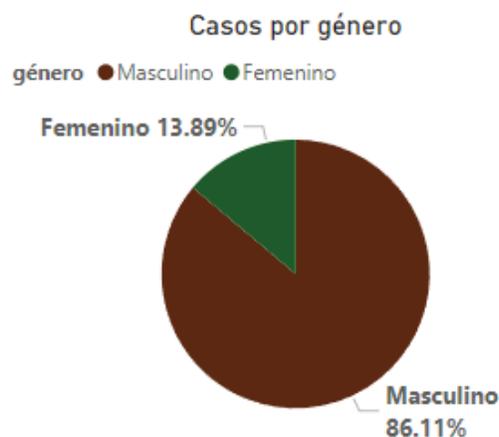
El flujo de minado de datos luego pasó a una etapa de recolección de datos de la plataforma, en los cuales se obtuvo información en específico para cada uno de los usuarios que se recolectaron mediante el proceso de raspado web. Entre las características se encuentran: nombre, nombre real, código de país, nivel y una lista de juegos recientemente utilizados, y un arreglo de juegos de los cuales es propietario. Las peticiones al punto de información proporcionado por la plataforma se realizaron utilizando el API de Steam de desarrollador al cual se puede aplicar con el objetivo de programar interfaces, productos e investigaciones utilizando esta información. Se utilizó una librería de Python llamada Steam API bajo la licencia MIT.

Utilizando la información que se obtuvo mediante el proceso de raspado de datos, se generaron distintas herramientas visuales utilizando la plataforma de visualización y presentación de reportes e informes de Microsoft, PowerBI.

6.1. Descripción de la muestra poblacional, datos generales

Los siguientes resultados son producto de la segunda encuesta a jóvenes, un total de 80 personas participaron en esta encuesta, de los cuales solamente 72 cumplían con estar dentro de las edades 18 y 29 años. En la misma se utilizaron los hallazgos generados por las entrevistas de profundidad.

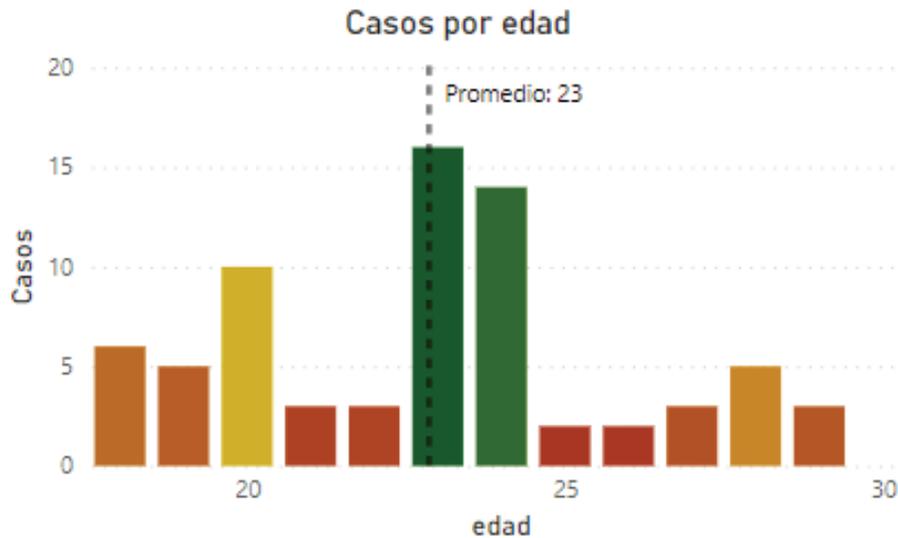
Figura 6.1: Descripción de casos segmentados por género seleccionado.



En el gráfico anterior se puede observar una predominancia de individuos que se identifican con el género masculino (86.11%) sobre el 13.89% que se identifican con el género femenino. Lo que se extrajo de esta información es que las futuras comparaciones del estudio basadas en este tipo de género no serán representativas de la realidad poblacional debido a el desbalance en casos. Si se llegara a necesitar un análisis utilizando este campo tendría que hacerse usando un análisis

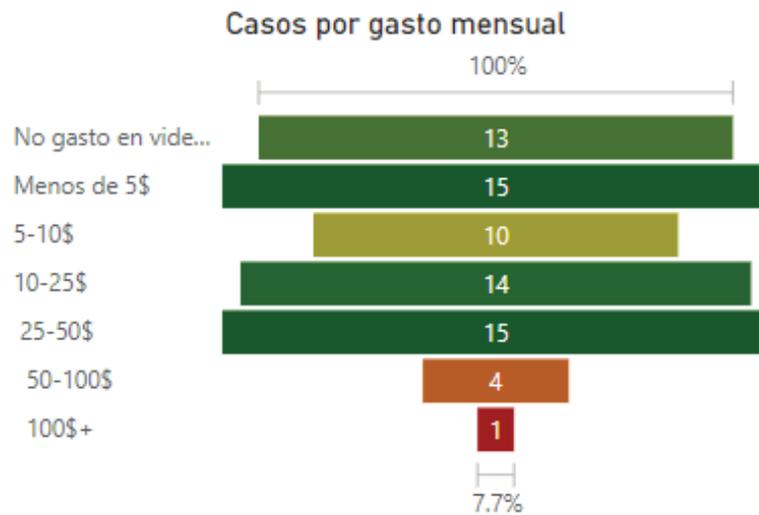
porcentual.

Figura 6.2: Número de casos del datagrama, segmentados por edad.



La anterior gráfica representa la distribución de edades de la muestra poblacional. De la misma se puede obtener que aún cuando hay predominancia en ciertas edades (23 y 24 años) la distribución de casos con forma seminormal, con un promedio de 23 años ubicado a la mitad de la muestra (18-29 años) .

Figura 6.3: Número de casos segmentados por gasto mensual.



Se observa en el gráfico anterior que los individuos se encuentran ubicados bajo el rango de 50\$ mensuales representan el 75 % de la muestra, mientras que el 25 % restante se encuentra ubicado en las categorías de ningún gasto mensual y gasto superior a los 50\$.

Figura 6.4: Número de casos segmentados por género más jugado.

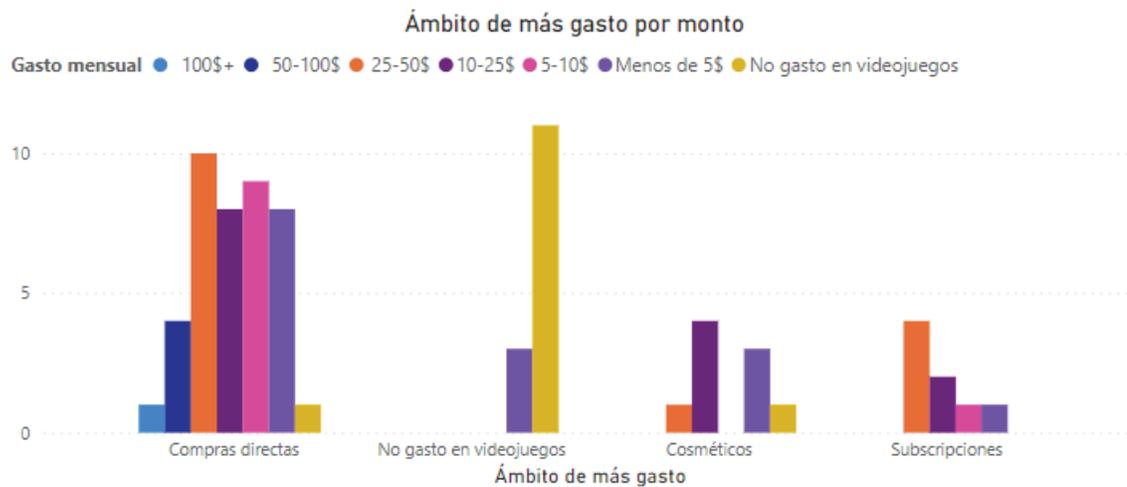


Figura 6.5: Matriz de relación entre género más jugado y gasto mensual en videojuegos.

Género más jugado	100\$+	50-100\$	25-50\$	10-25\$	5-10\$	Menos de 5\$	No gasto en videojuegos	Total
Aventura	1	2	9	3	3	2	4	24
Estrategia		1	2	5	2	5	5	20
Shooters		1	2	3	3	7	2	18
Acción			1	3	1	1		6
Deportes			1		1		2	4
Total	1	4	15	14	10	15	13	72

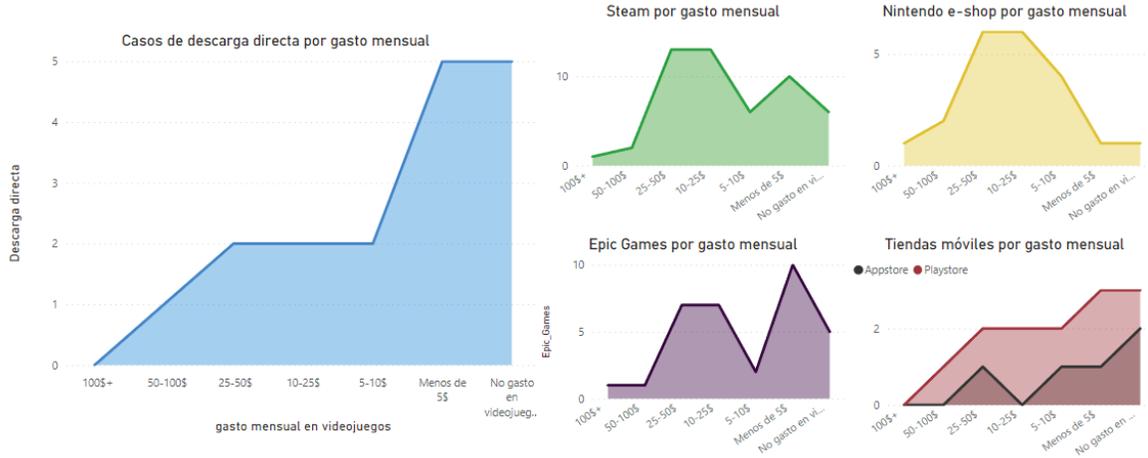
En la matriz de relación entre los géneros de videojuegos más jugados y el rango de gasto mensual se puede observar que el género más rentable, con un total de 20 casos es el género aventura, seguido por disparos (shooters) con 16 casos y estrategia con 15 casos. Los géneros menos rentables son acción y deportes con 6 y 2 casos respectivamente.

Figura 6.6: Gasto mensual segmentado por ámbito de más gasto.



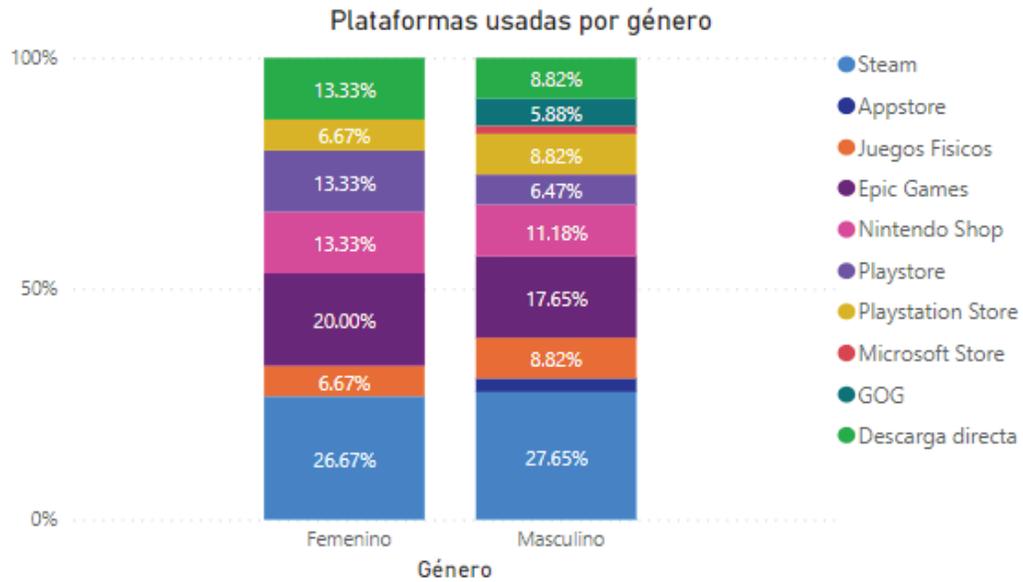
6.2. Descripción de la muestra poblacional, plataformas usadas

Figura 6.7: Plataformas utilizadas segmentadas por gasto mensual.



El conjunto de gráficos anterior nos enseña que existe una relación directa entre las variables de gasto mensual en videojuegos y la variable de plataforma en la que se adquieren los mismos. La relación es inversamente proporcional en el caso de descarga directa, es decir entre menos dinero es gastado en videojuegos mas casos de descarga directa existen, y directamente proporcional en el caso de Steam, Nintendo Eshop, Epic-games y tiendas móviles.

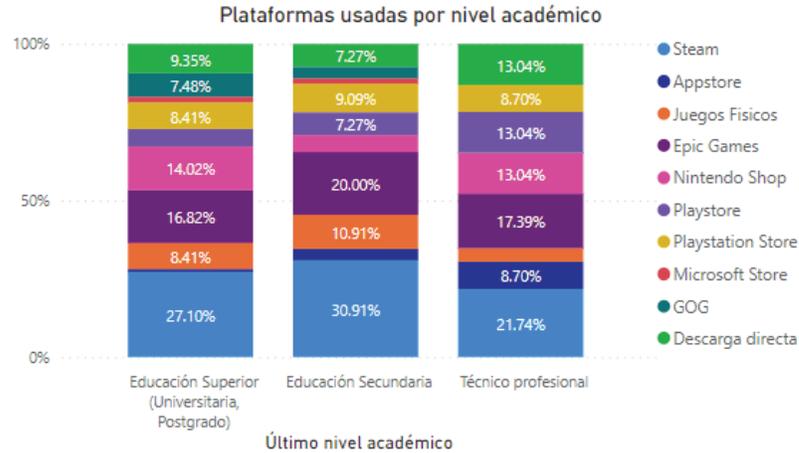
Figura 6.8: Plataformas utilizadas dividido por género.



Según lo presentado en la primera gráfica de la sección descripción de la muestra el indicador género de la población no nos entrega una representación de la población debido al desbalance de

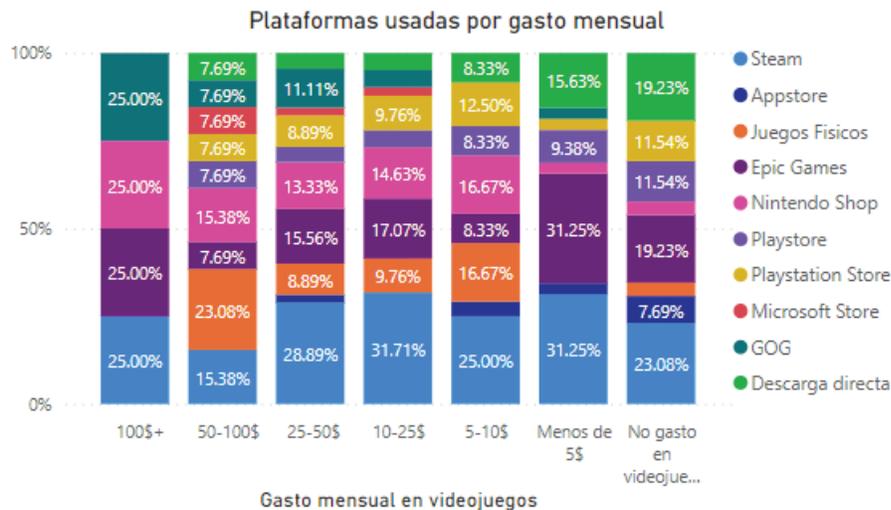
casos, debido a esto se generó este gráfico porcentual que nos presenta una versión de las plataformas de más uso dentro de la muestra, que se pueden extender a la población.

Figura 6.9: Plataformas utilizadas dividido por nivel académico.



Un hallazgo significativo en el área de plataformas más utilizadas por nivel académico es la reducción en el uso de plataformas móviles, existe una relación entre el nivel educativo y el gasto en plataformas móviles. Entre más avanzado el nivel educativo menos consumo en tiendas móviles. La excepción es la tienda Nintendo Eshop que permite portabilidad y aumenta con el nivel educativo del individuo.

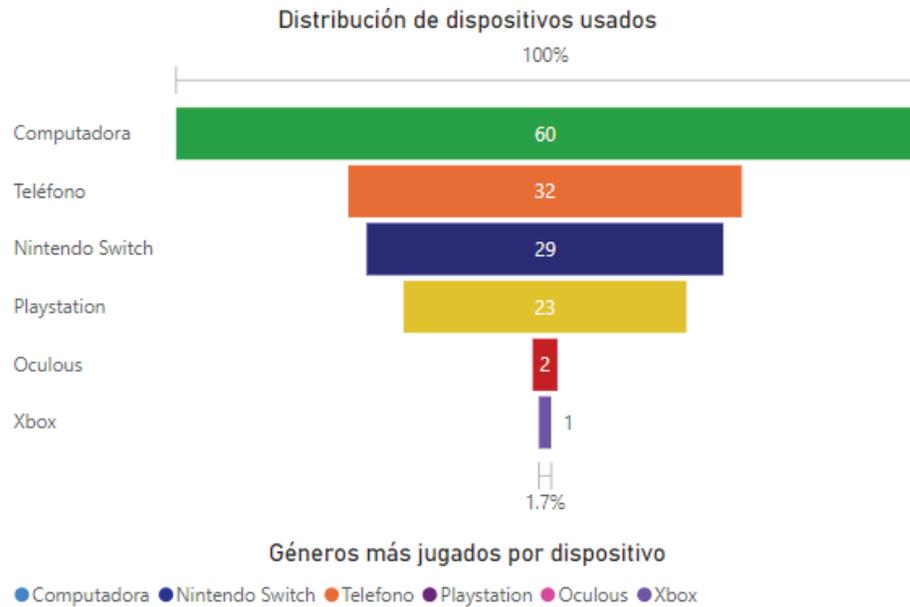
Figura 6.10: Plataformas utilizadas dividido por gasto mensual.



En el campo de gasto mensual podemos observar que los individuos que obtienen sus videojuegos mediante copias físicas incrementa la cantidad de dinero que gasta mensualmente, al igual que los individuos que utilizan la Nintendo Eshop y la tienda virtual GoG. Como se presentó en el análisis del gráfico anterior una disminución en las plataformas móviles Appstore y Playstore.

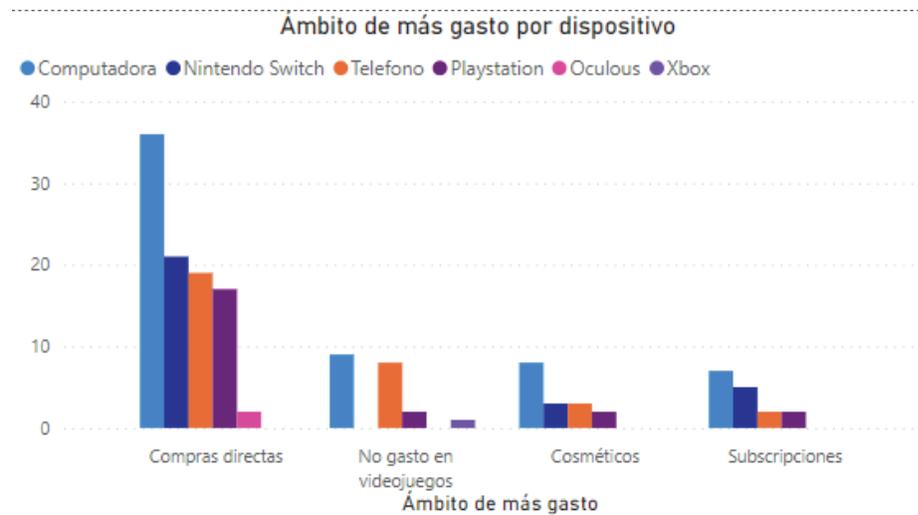
6.3. Descripción de la muestra poblacional, dispositivos usados

Figura 6.11: Número de casos segmentados por dispositivos más utilizados.



El gráfico anterior nos indica que las computadoras son las más populares, solamente un 8.64 % de los encuestados no las utiliza para jugar videojuegos. Otra característica importante es la ausencia de jugadores de Xbox, solamente un individuo lo utiliza actualmente y representa menos del 1 % de la muestra.

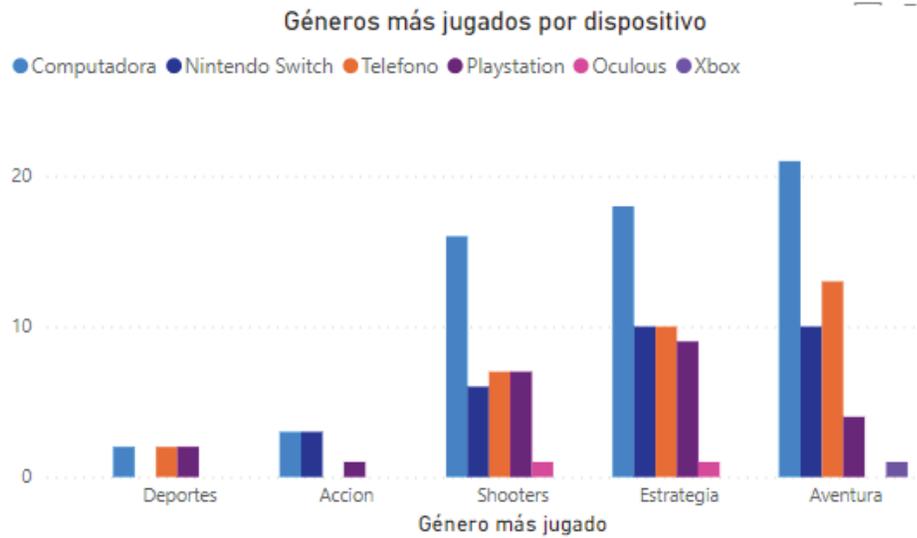
Figura 6.12: Ámbito de más gasto dividido por dispositivo utilizado.



El ámbito de gasto en el área de videojuegos nos indica que la muestra prefiere compras directas de títulos que pago de suscripciones y cosméticos. Los juegos de computadora son más comunes en

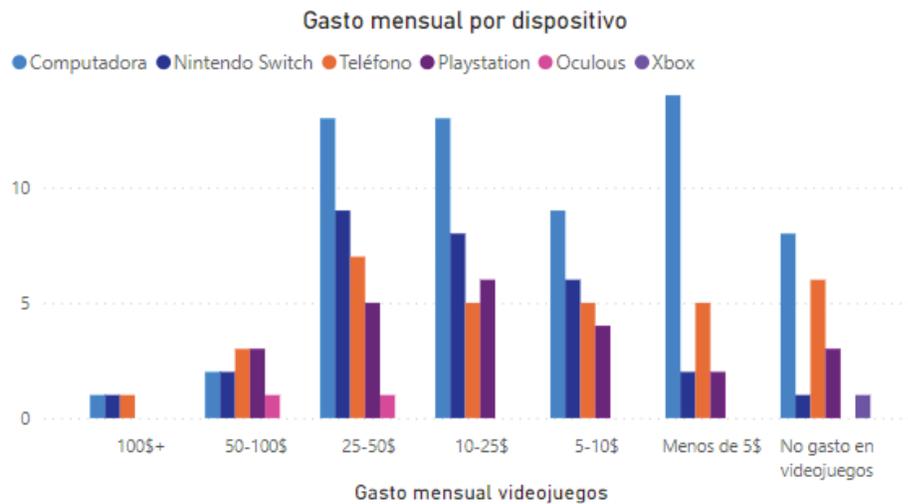
todos los ámbitos seguido por Nintendo Switch y teléfono/Playstation.

Figura 6.13: Género más jugado dividido por dispositivo utilizado.



El gráfico anterior nos reafirma lo presentado en descripción de la muestra, que género de videojuego es más popular. La segregación utilizando dispositivos nos indica que los más populares son computadora y teléfono, para Aventura, el género más jugado.

Figura 6.14: Gasto mensual dividido por dispositivo utilizado.

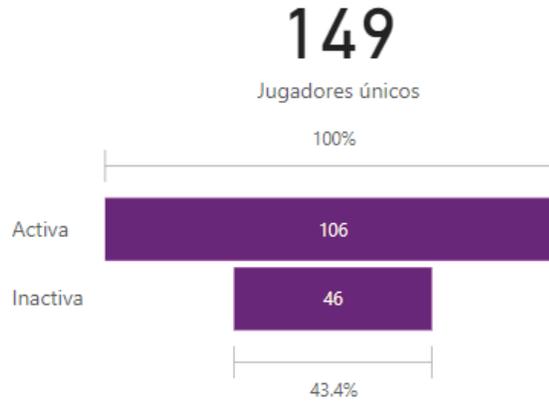


En el rubro de gasto mensual por dispositivo, se observa que los dispositivos más usados se encuentran en el rango de 5\$ a 50\$ mensuales, los jugadores de teléfono mantienen un constante. El dispositivo más utilizado en el rango bajo los 5\$ es la computadora.

6.4. Población guatemalteca que utiliza la plataforma STEAM

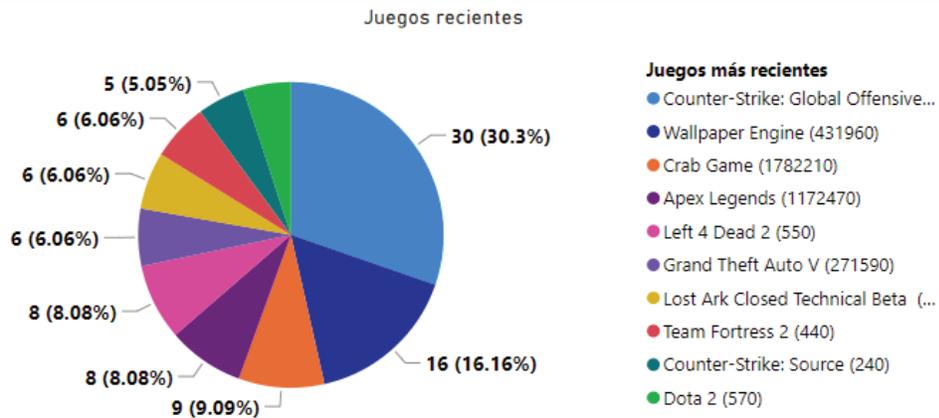
Después de un periodo de inactividad de 2 años o más, las cuentas de la plataforma pasan a un estado de inactividad.

Figura 6.15: Estado de las cuentas de los jugadores que frecuentan la plataforma.



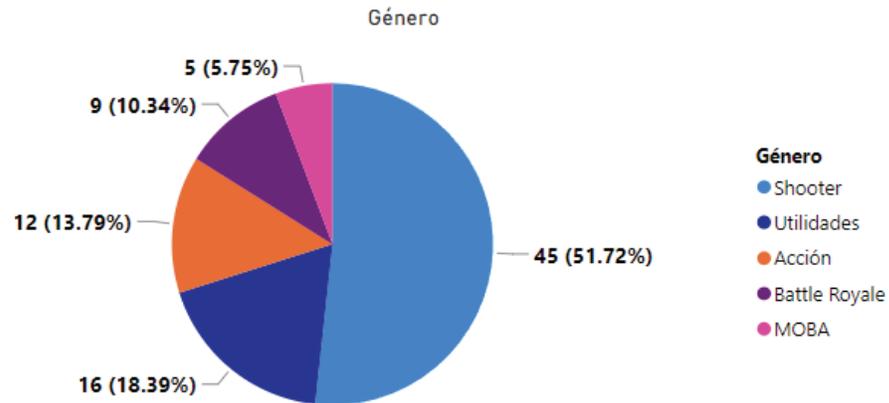
Resumen de los individuos que tenían su perfil en público y fue posible extraer su información de la plataforma Steam, segmentado por cuentas activas e inactivas, basado en la última vez que se ejecuto un programa utilizando el lanzador de la plataforma. Podemos observar que, al menos, un 30 % abandona la plataforma o no frecuenta el uso de la misma regularmente.

Figura 6.16: Juegos más utilizados últimamente.



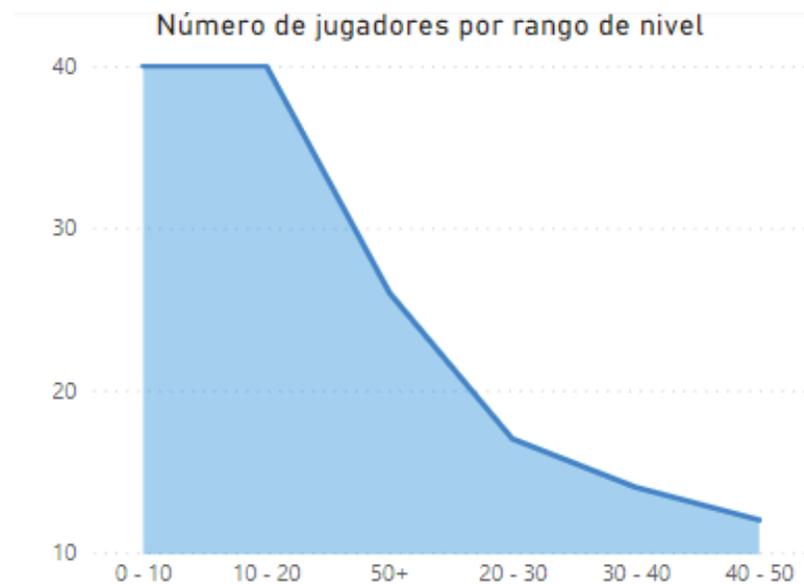
Segmentación de últimos juegos lanzados utilizando la plataforma de Steam, categorizados en los 10 juegos más populares para todos los individuos de la muestra. En el gráfico se pueden observar múltiples juegos que utilizan el ratón, como los juegos de disparos, que son populares entre los jugadores de computadora.

Figura 6.17: Géneros de juegos utilizados últimamente.



En el siguiente gráfico se segmentaron los 10 juegos más jugados de la población en su género. Se obtuvo una dominancia de los juegos tipo shooter con 45 % seguido por juegos de utilidades con 16 % y acción con 14 %. Nos indica que ya que la computadora permite el uso de ratón como periférico de entrada de información empuja a los jugadores de este dispositivo a jugar videojuegos que utilicen al máximo sus capacidades.

Figura 6.18: Agrupación de nivel de plataforma.



El nivel de jugador en la plataforma es medido basado en el número de medallas y participaciones en eventos de la plataforma, cada 100 puntos se otorga un nivel.

La última gráfica muestra el nivel de jugador, se puede observar que los individuos se mantienen constantemente en los rangos de 0-10 y 10-20 mientras que los siguientes niveles empiezan a decaer debido a la dificultad y la inversión de tiempo necesaria para llegar a estos niveles.

6.5. Población global que utiliza la plataforma STEAM

Estados de las cuentas de los grupos globales de jugadores de Steam de Estados Unidos, Francia y Japón

Figura 6.19: Estado de las cuentas de los jugadores estadounidenses que frecuentan la plataforma.

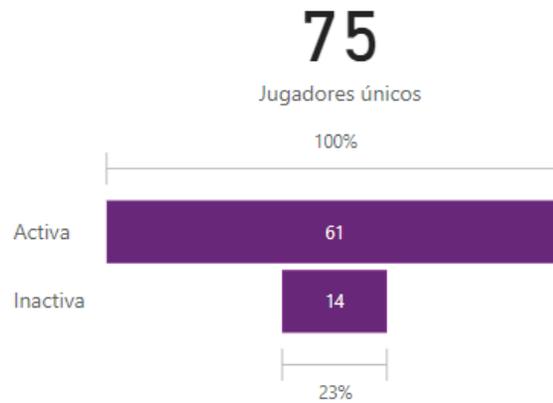
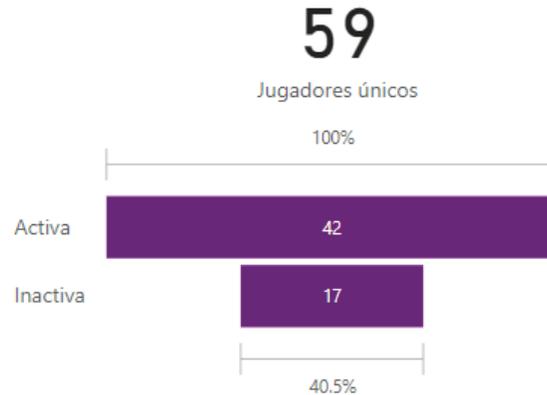


Figura 6.20: Estado de las cuentas de los jugadores franceses que frecuentan la plataforma.



Figura 6.21: Estado de las cuentas de los jugadores japoneses que frecuentan la plataforma.



Juegos más utilizados recientemente por grupos globales de jugadores de Estados Unidos, Francia y Japón

Figura 6.22: Juegos más utilizados últimamente por jugadores estadounidenses.

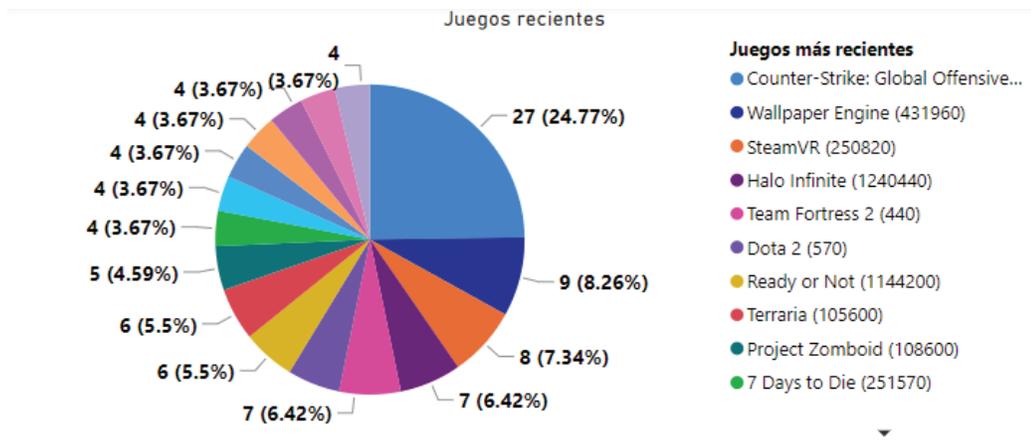


Figura 6.23: Juegos más utilizados últimamente por jugadores franceses.

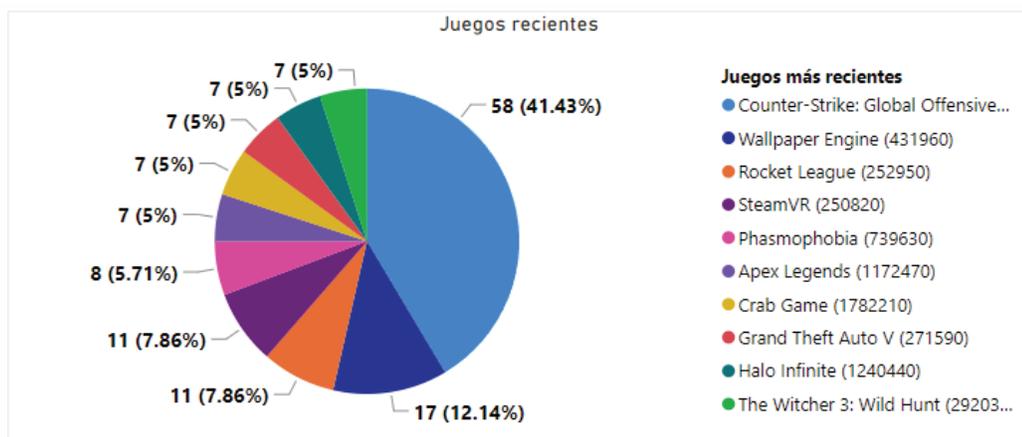
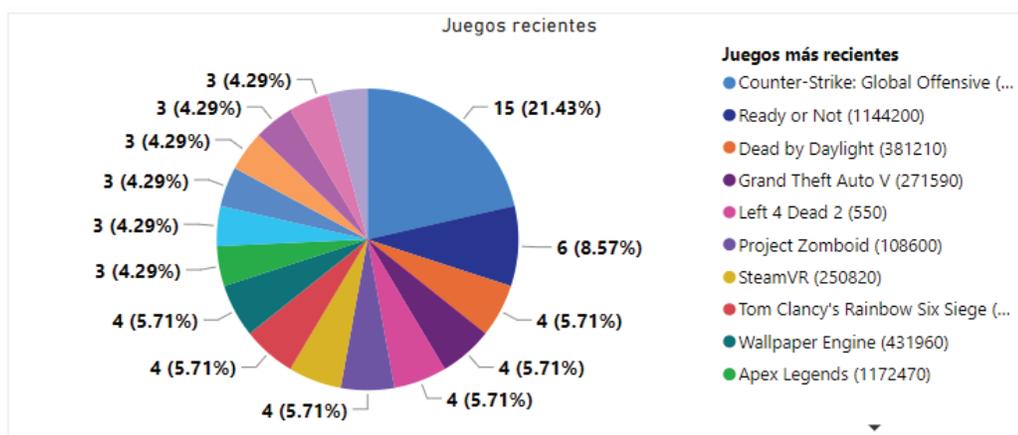


Figura 6.24: Juegos más utilizados últimamente por jugadores japoneses.



Géneros de juegos más utilizados por grupos globales de jugadores de Estados Unidos, Francia y Japón

Figura 6.25: Géneros de juegos utilizados últimamente por jugadores estadounidenses.

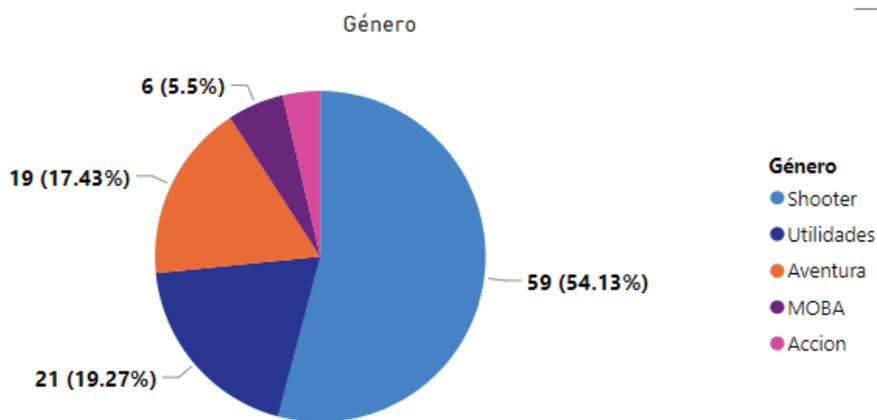


Figura 6.26: Géneros de juegos utilizados últimamente por jugadores franceses.

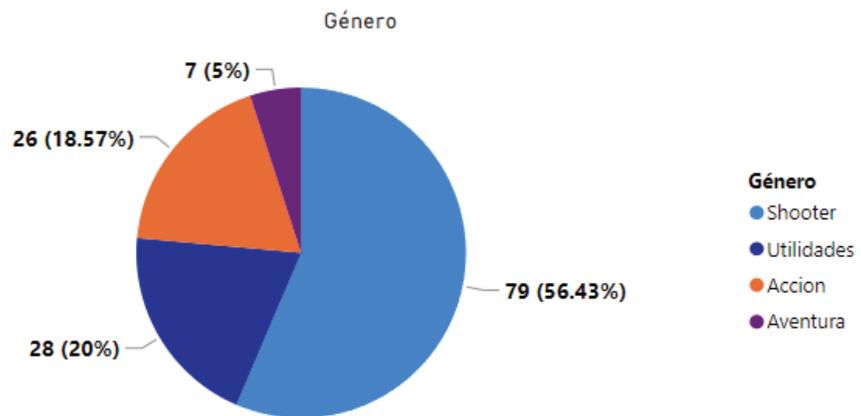
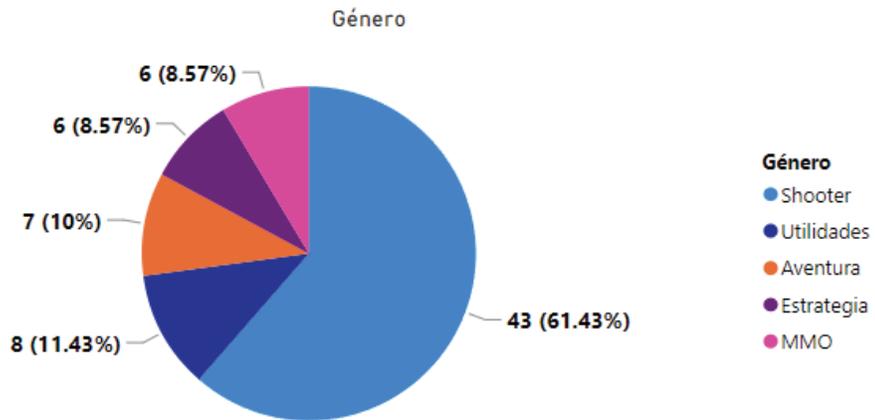


Figura 6.27: Géneros de juegos utilizados últimamente por jugadores japoneses.



7.1. Delphi expertos

Durante la segunda ronda del análisis de Delphi, se les preguntó cual consideraban el género más popular entre la población de estudio. El 83.3% argumentó que es acción, mientras que el 16.7% argumentó que es disparos. El estudio reveló que los géneros más populares son en orden: Aventura, Estrategia y Disparos. 6.13

Otro rubro donde se inquirió sobre el uso de dispositivos móviles reveló que los expertos consideran que todos los jugadores de la región utilizan dispositivos móviles más que cualquier otro tipo de dispositivo. El resultado que se obtuvo está de acuerdo con esa opinión, los videojuegos móviles, es decir Nintendo Switch y Teléfonos son utilizados por 61 de los 72 individuos, seguidos por 60 personas que usan la computadora de manera paralela. 6.11

Respecto al consumo monetario los expertos argumentaron que consideraban que un bajo porcentaje de jugadores paga por videojuegos, solamente entre 10 y 25% de los jugadores de la región. Los resultados muestran que solamente un 18% no paga por videojuegos y que el 82% gasta al menos 5\$ de manera mensual en compras directas, cosméticos o suscripciones. 6.3

7.2. Limitaciones del área rural

Se inquirió sobre las diferencias entre el área rural y el área urbana obteniendo múltiples opiniones sobre el tema. Algunos expertos argumentaron que debido a que no tienen la misma facilidad para realizar los pagos para compras de videojuegos el gasto mensual debería de ser menor. Otros expertos argumentaron que debido a la popularidad de los teléfonos inteligentes la cantidad de jugadores móviles debería ser más que el resto de dispositivos disponibles. Una tercera perspectiva argumentó que el acceso a la red es limitada por lo que los jugadores del área rural estarían drásticamente reducidos en el ámbito de juegos multijugador.

7.3. Generalización de la muestra

Como se mostró en la figura donde se dividió utilizando el género 6.1, debido a que hay una división muy pronunciada entre los individuos que indicaron género masculino y los que indicaron género femenino, se decidió no utilizar esta dimensión para los análisis de la muestra. Se incluyó una figura porcentual para determinar si existía alguna discrepancia entre las plataformas de adquisición de nuevos títulos sin ningún hallazgo evidente. 6.8

7.4. Resultados globales, comparación con resultados locales

En el caso de los resultados de carácter global, solamente se puede hacer un análisis de la plataforma Steam, segmentada para jugadores de computadora. Entre los resultados se encontró que la proporción de jugadores activos de las regiones de Japón y Francia con porcentajes de población inactiva de 40.5% y 44.6% se asemejan a el porcentaje de población inactiva de la región guatemalteca, con un 43.4%. Las razones de esto no se investigan a profundidad pero se considera que para futuros estudios, se examine el acceso a juegos multiplataforma y su relación con delimitadores geográficos e ingreso prescindible.

En el ámbito de los géneros de juegos más jugados en el ámbito global son los videojuegos de tipo Shooter, en todos los casos fue el género más jugado por más de 25%. Las utilidades en juegos de computadora compiten con los demás géneros por el segundo lugar debido a que ya que el dispositivo es utilizado para más que solo jugar videojuegos, la plataforma ofrece software para estas necesidades.

Otra característica segmentadora entre los resultados es la cantidad de videojuegos utilizados recientemente. En las regiones de Guatemala y Francia, con 9 y 10 juegos utilizados recientemente. Hay un decremento considerable comparado con las regiones de Estados Unidos y Japón, que se encuentran entre los 14 y 15 juegos respectivamente. Una nueva perspectiva para una investigación a profundidad que surge de esta comparación podría basarse en precios de títulos y la capacidad de adquisición de los mismos por la población.

- Este trabajo ayudó a clasificar los comportamientos de los individuos que pertenecen a la población de estudio. Se determinaron los géneros de videojuegos más rentables: Acción, Disparos y Estrategia. Además se determinó que el ámbito de gasto más efectivo son las compras directas.
- Se descubrió mediante a la información generada por el cruce entre plataformas utilizadas y gasto mensual en videojuegos, que los individuos que poseen un dispositivo específico para jugar videojuegos, son más propensos a gastar más dinero en videojuegos.
- Utilizando el cruce de información mencionado en la conclusión anterior, se determinó que los que utilizan el medio de descarga directa tienden a no gastar en nuevos juegos.
- En el área de dispositivos utilizados, lo más significativo es la ausencia de jugadores de Xbox, se determinó que menos del 1% de los encuestados juega utilizando ese dispositivo.
- Entre los hallazgos de nivel educativo se detectó una disminución en el gasto en las plataformas móviles Playstore y Appstore entre más nivel educativo tuviese el individuo. Esta disminución no afectó el consumo de Nintendo Shop.
- Debido a la disparidad entre individuos de género con el que se identifican, no se consideró este ámbito para más comparaciones de otros indicadores.
- La investigación de otras regiones geográficas reveló que los juegos más populares en las regiones de investigación son del género Disparos, específicamente el título Counter Strike Global offensive, debido a su gratuidad y compatibilidad con distintas capacidades de hardware de computadora.

Recomendaciones

Se recomienda para futuras investigaciones sobre las tendencias de la población de videojuegos, extender el alcance para abarcar los individuos ubicados bajo los 18 años y sobre 29 años y que según los expertos representan un porcentaje significativo de la población.

Una investigación más exhaustiva debe de tomar en cuenta la división entre los jugadores citadinos y los ubicados en áreas rurales, con el objetivo de otorgar perfiles mas acertados basados en esta demografía.

Se recomienda subdividir ciertos géneros de videojuegos, como acción y estrategia, con el objetivo de obtener información más específica en ámbitos como consumo monetario y dispositivos utilizados.

Ayudaría en la perfilación de los jugadores saber qué porcentaje de su ingreso representa el gasto de adquisición de nuevos títulos, microtransacciones y o cosméticos.

-
- [1] *Steam Web API Documentation*. <https://steamcommunity.com/dev>.
- [2] *Facet Analysis of Video Game Genres*. 2014. <https://doi.org/10.97762F14057>.
- [3] Chambers, John: *Software for Data Analysis*. Springer New York, 2008. <https://doi.org/10.1007%2F978-0-387-75936-4>.
- [4] Glez-Peña, Daniel, Anália Lourenço, Hugo López-Fernández, Miguel Reboiro-Jato y Florentino Fdez-Riverola: *Web scraping technologies in an API world*. Briefings in Bioinformatics, 15(5):788–797, Abril 2013, ISSN 1467-5463. <https://doi.org/10.1093/bib/bbt026>.
- [5] Ilyas, Ihab F. y Xu Chu: *Data Cleaning*. Association for Computing Machinery, New York, NY, USA, 2019, ISBN 9781450371520.
- [6] Jun, Seung Pyo, Hyoung Sun Yoo y San Choi: *Ten years of research change using Google Trends: From the perspective of big data utilizations and applications*. Technological Forecasting and Social Change, 130:69–87, may 2018. <https://doi.org/10.1016%2Fj.techfore.2017.11.009>.
- [7] McKinney, Wes: *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly, Beijing, first edición, 2013, ISBN 9781449319793 1449319793.
- [8] Milevska-Kostova, Neda y William Dunn: *Delphi Analysis*. páginas 423–436, Enero 2010.
- [9] Noticias, Agencia Guatemalteca de: *Conozca la población por edad y sexo que habita en Guatemala*, 2021. <https://agn.gt/conozca-la-poblacion-por-edad-y-sexo-que-habita-en-guatemala/>.
- [10] Osang, Udoimuk, Etta, Ushie y Offiong: *Methods of Gathering Data for Research Purpose and Applications Using IJSER Acceptance Rate of Monthly Paper Publication*. IOSR Journal of Computer Engineering, 15(2):59–65, 2013.
- [11] Paradis, Elise, Bridget O' Brien, Laura Nimmon, Glen Bandiera y Maria Martimianakis: *Design: Selection of Data Collection Methods*. Journal of graduate medical education, 8(2):263–4, 2016.
- [12] RI Co., Ltd gameage: *Segmentation Method*, 2021. <https://www.gameage.jp/en/methods/>.
- [13] Sohan, S.M., Craig Anslow y Frank Maurer: *A Case Study of Web API Evolution*. jun 2015. <https://doi.org/10.1109%2Fservices.2015.43>.
- [14] Suresh, KP y S Chandrashekara: *Sample size estimation and power analysis for clinical research studies*. Journal of Human Reproductive Sciences, 5(1):7, 2012. <https://doi.org/10.4103%2F0974-1208.97779>.

- [15] Veeder, Jane, Mark Stephen Pierce, Eugene Jarvis, John N. Latta, Heidi Therese Dangelmaier y Jez San: *Videogame industry overview (panel session)*. En *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH '95*. ACM Press, 1995. <https://doi.org/10.1145%2F218380.218521>.
- [16] Yamaguchi, Shinichi, Kotaro Iyanaga, Hirohide Sakaguchi y Tatsuo Tanaka: *The Substitution Effect of Mobile Games on Console Games: An Empirical Analysis of the Japanese Video Game Industry*. *The Review of Socionetwork Strategies*, 11(2):95–110, nov 2017. <https://doi.org/10.1007%2Fs12626-017-0014-1>.
- [17] Zubaydi, Maen y Ergun Gide: *Usage and Spending on Mobile Games: A Comparison Between Australian Single-Player Gamers and Multi-Player Gamers*. En *2019 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. IEEE, dec 2019. <https://doi.org/10.1109%2Fcde48274.2019.9162377>.

Arreglo: Conjunto de datos o una estructura de datos homogéneos que se encuentran ubicados en forma consecutiva en la memoria. 8, 9

Biblioteca: Conjunto de implementaciones funcionales codificadas en un lenguaje de programación. 8, 9

Funcion: Sección de un programa que calcula un valor de manera independiente al resto del programa. 7–9

Python: Lenguaje multiplataforma de alto nivel de código abierto gratuito. 8, 9, 20

Relacional: Tipo de base de datos que almacena y proporciona acceso a puntos de datos relacionados entre sí. 9, 10