

Extracción de características del negocio



**Tecnológico
de Monterrey**

A 28 de marzo de 2025

Extracción de características de “Wuupi”

Equipo 3:

Braulio González Esquivel || A01425274

Lucero Jannete López García || A01736938

Pilar Méndez Briones || A01736843

Rubén Uriel Flores Bello || A01737098

Analítica de datos y herramientas de inteligencia artificial II (Gpo 501)

Profesores: Juan Manuel Ahuactzin Larios,

Candy Yuridiana Alemán Muñoz,

Rigoberto Cerino Jiménez y

Alfredo García Suárez

CC: Sexto semestre profesional – Escuela de
Negocios – Licenciatura en Administración y
Transformación de Negocios.

Extracción de características de



Índice

Contenido

Índice	2
Introducción	3
Acondicionamiento de la información	4
Método numérico.....	5
Método numérico, corrección.....	6
Método “string”	7
Valore atípicos.....	9
Transformación de la información en modelos gráficos	11
Análisis gráfico de personas relacionadas.....	11
Análisis de juegos y sus características.....	13
Bibliografía.....	19

Introducción

Wuupi es un startup mexicano que desarrolla soluciones tecnológicas innovadoras para mejorar la calidad de vida de personas con discapacidad y sedentarismo. Su línea de productos combina videojuegos terapéuticos con controles adaptativos, permitiendo realizar terapias físicas y conductuales.

La empresa nació de las experiencias personales de sus fundadores con familiares que enfrentan dificultades para acceder a terapias efectivas y atractivas. Su tecnología se basa en algoritmos pedagógicos y herramientas de medición de progreso, lo que permite realizar un seguimiento detallado del progreso de los usuarios.

Principales actividades (Wuupi. (s.f.):

- Se especializa en el desarrollo de algoritmos pedagógicos para personalizar las experiencias de juego.
- Incorpora mediciones de avances de los usuarios en términos de tiempo de reacción, eficiencia motriz y otros indicadores clave.
- Su plataforma integra un tablero interactivo, que permite a terapeutas y administradores visualizar métricas y generar informes detallados.
- Utiliza métricas de uso en cada sesión de juego, analizando la interacción del usuario en términos de aciertos, errores y patrones de respuesta.
- Su enfoque es proporcionar herramientas que facilitan la interpretación de datos y que permiten adaptar las terapias a cada usuario.

Acondicionamiento de la información

Analizando el archivo en formato “.csv”, se pudo encontrar 5865 filas y 13 columnas de las cuales seis fueron de tipo “float64” (botón correcto, tiempo de interacción, número de interacción, autopush, tiempo de lección y tiempo de sesión), y siete columnas de tipo “object” (administrador, usuario, minijuego, color presionado, dificultad, fecha, juego).

Un punto importante para destacar es que la columna “fecha” se debe convertir a formato “datetime” para futuros análisis donde queramos recuperar información de ciertas fechas o periodos.

```
Data columns (total 13 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Administrador                          5865 non-null   object
1   Usuario                                5865 non-null   object
2   botón correcto                       5103 non-null   float64
3   tiempo de interacción                  5103 non-null   float64
4   mini juego                             5709 non-null   object
5   número de interacción                  5103 non-null   float64
6   color presionado                       5103 non-null   object
7   dificultad                             5865 non-null   object
8   fecha                                  5865 non-null   object
9   Juego                                  5865 non-null   object
10  auto push                              5103 non-null   float64
11  tiempo de lección                      5688 non-null   float64
12  tiempo de sesión                       5259 non-null   float64
dtypes: float64(6), object(7)
memory usage: 595.8+ KB
```

Elaboración propia, (con base en la base de datos principal de Wuupi): tabla de información de objetos por columna, (marzo de 2025).

A continuación, se muestra una tabla en donde se puede identificar que, ocho del total de columnas, no cuentan con información, es decir, no tienen valores nulos.

Administrador	0
Usuario	0
botón correcto	762
tiempo de interacción	762
mini juego	156
número de interacción	762
color presionado	762
dificultad	0
fecha	0
Juego	0
auto push	762
tiempo de lección	177
tiempo de sesión	606
dtype:	int64

Elaboración propia, (con base en la base de datos principal de Wuupi): tabla de identificación de valores nulos por columna, (marzo de 2025).

Una vez obtenida esta información, se procedió a procesar cada columna de acuerdo con su tipo de formato de información.

Método numérico

Para las columnas “botón correcto” y “autopush” se sustituyó el valor nulo por la cantidad específica de “2” con el fin de poder diferenciar los valores no específicos de los otros valores existentes dentro de esa columna, es decir, “1” y “0”; respectivamente significando que, la interacción fue correcta o, que la interacción no fue correcta. Mejora de las posibilidades de comprensión y procesamiento posterior de los datos.

```
binarias = ['botón correcto', 'auto push']  
df[binarias] = df[binarias].fillna(2).astype(int)
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por una cifra en concreto, (marzo de 2025).

Método numérico, corrección

Para el siguiente método, cuatro columnas fueron sustituidas con el número “0” para no crear confusiones.

Tiempo de lección:

Como podemos observar el tiempo de lección tiene un tiempo final, por lo que, si lo sustituimos con algún número diferente, los cálculos sobre determinar tiempo podrían cambiar significativamente, así que seguimos el mismo patrón de ceros.

```
df['tiempo de lección'] = df['tiempo de lección'].fillna(0)
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por una cifra en concreto¹, (marzo de 2025).

Tiempo de sesión:

Una vez que la lección acaba ya no hay más tiempo de sesión, por lo que, si sustituimos los valores con un número diferente de cero, podríamos alternar los cálculos.

```
df['tiempo de sesión'] = df['tiempo de sesión'].fillna(0)
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por una cifra en concreto², (marzo de 2025).

Tiempo interacción:

Al ya no tener más tiempo de lección y sesión, significa que las interacciones acabaron lo que contaría cómo cero, al no tener más interacciones por parte del paciente.

```
df['tiempo de interacción'] = df['tiempo de interacción'].fillna(0)
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por una cifra en concreto³, (marzo de 2025).

Número de interacción:

Mismo que tiempo de interacción, al no haber interacciones por parte del paciente no se puede contabilizar ningún valor más, por lo que, se deben sustituir con cero para conocer la diferencia.

```
df['número de interacción'] = df['número de interacción'].fillna(0)
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por una cifra en concreto⁴, (marzo de 2025).

A continuación, se presenta una imagen en la que podemos observar en dónde se tienen que hacer los cambios dentro de la base de datos.

tiempo de lección	tiempo de sesión	tiempo de interacción	número de interacción
0	0	5.399169	1
0	0	1.2834	2
0	0	2.700226	3
0	0	3.050262	4
0	0	4.750256	5
0	0	7.400463	6
0	0	13.16735	7
0	0	22.46778	8
0	0	14.51718	9
0	0	15.61735	10
0	0	16.80135	11
0	0	18.16748	12
63.80254			

Elaboración propia, (con base en la base de datos principal de Wuupi): valores iniciales antes de generar la sustitución de valores nulos en cada una de las columnas, (marzo de 2025).

Método “string”

Este método se usa generalmente para sustituir los valores nulos de un objeto contenga como tipo de información un texto, en este caso en particular, se identificaron dos columnas que manejan este tipo de información, por lo que para que

fueran más fáciles de interpretar y evitar confusiones, se les aplicó este método, sustituyendo caracteres por otros más comprensibles.

A continuación, se presenta una imagen en donde se puede observar las columnas (datos) de tipo texto que se buscan, además, se muestran los valores nulos, mismos que se sustituirán por un tipo de texto distinto para una mejor interpretación.

Administrador	Usuario	botón correcto	mini juego	color presionado	dificultad	fecha	Juego	auto push	tiempo de lección	tiempo de sesión
nicolas	nicolas	0	Restaurante	violet	Episodio 2	#####	Astro	0	0	0
nicolas	nicolas	0	Restaurante	violet	Episodio 2	#####	Astro	0	0	0
nicolas	nicolas	0	Restaurante	violet	Episodio 2	#####	Astro	0	0	0
nicolas	nicolas	0	Restaurante	violet	Episodio 2	#####	Astro	0	0	0
nicolas	nicolas	0	Restaurante	green	Episodio 2	#####	Astro	0	0	0
nicolas	nicolas		Restaurante		Episodio 2	#####	Astro		76.45556	
nicolas	nicolas				Episodio 2	#####	Astro			1330.133

Elaboración propia, (con base en la base de datos principal de Wuupi): tabla de información de columnas, (marzo de 2025).

Minijuego:

Para esta columna se decidió rellenarla con texto para continuar con la secuencia de palabras en una columna, al tener nulos, conviene llamarlos como “Sin interacción”, para evitar confusiones.

```
df['mini juego'] = df['mini juego'].fillna('Sin interacción')
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por un texto en concreto¹, (marzo de 2025).

Color presionado:

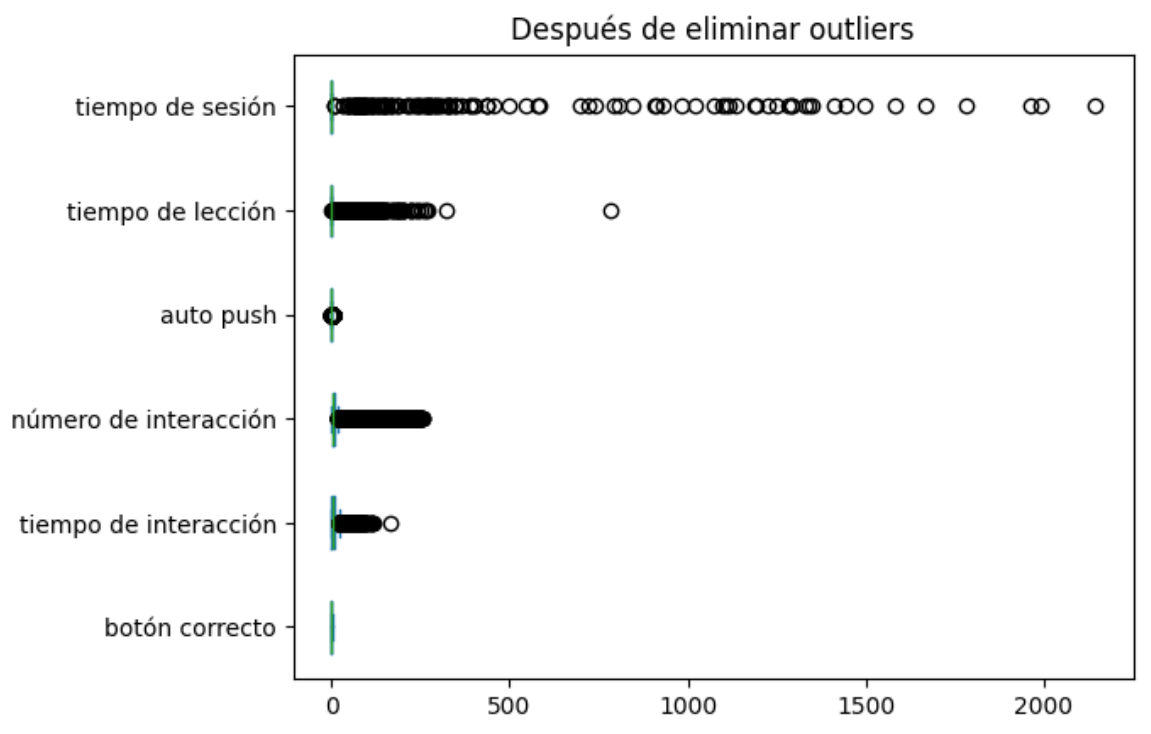
En esta columna hay valores nulos debido a dos factores porque acabó la lección o porque acabó la sesión, por lo que el valor nulo se sustituye con “sin color”, haciendo referencia a que no hay más interacciones.

```
df['color presionado'] = df['color presionado'].fillna('sin color')
```

Elaboración propia, (con base en la base de datos principal de Wuupi): sustitución de valores nulos por un texto en concreto², (marzo de 2025).

Valore atípicos

Una vez eliminando los valores nulos, se procede a eliminar los “valores sobresalientes” (outlayers) que se encuentran en las columnas “tiempo de sesión”, “tiempo de sección”, “autopush”, “número de interacción” y “tiempo de interacción”. Los valores nulos existentes dentro de estas columnas se les sustituyó por el método de promedio, para no generar alteraciones.



Elaboración propia, (con base en la base de datos principal de Wuupi): eliminación y vista de los valores normales sin valores sobresalientes, (marzo de 2025).

Una vez que se eliminan los valores nulos y los valores sobresalientes el resultado de la operación resulta en una base de datos limpia con posibilidad de ser interpretada de manera más ágil. Lo anterior permite conocer información útil y hacerla comprensible para las demás personas a través de métodos de generación de apoyos gráficos.

A continuación, se muestra una imagen con los nuevos tipos de datos (almacén de información) por columna, en donde observamos los cambios realizados.

Es preciso mencionar también, que ningún valor se está eliminando, solo se transforma o intercambia por otro valor que resulte más fácil de interpretar y tratar para posteriores adaptaciones de información.

```
RangeIndex: 5865 entries, 0 to 5864
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   Administrador          5865 non-null   object
1   Usuario                5865 non-null   object
2   mini juego             5865 non-null   object
3   color presionado       5865 non-null   object
4   dificultad             5865 non-null   object
5   fecha                  5865 non-null   datetime64[ns]
6   Juego                  5865 non-null   object
7   botón correcto         5865 non-null   int32
8   tiempo de interacción  5865 non-null   float64
9   número de interacción  5865 non-null   float64
10  auto push              5865 non-null   float64
11  tiempo de lección      5865 non-null   float64
12  tiempo de sesión       5865 non-null   float64
dtypes: datetime64[ns](1), float64(5), int32(1), object(6)
memory usage: 572.9+ KB
```

Elaboración propia, (con base en la base de datos principal de Wuupi): visualización de nuevos tipos de datos por columna, (marzo de 2025).

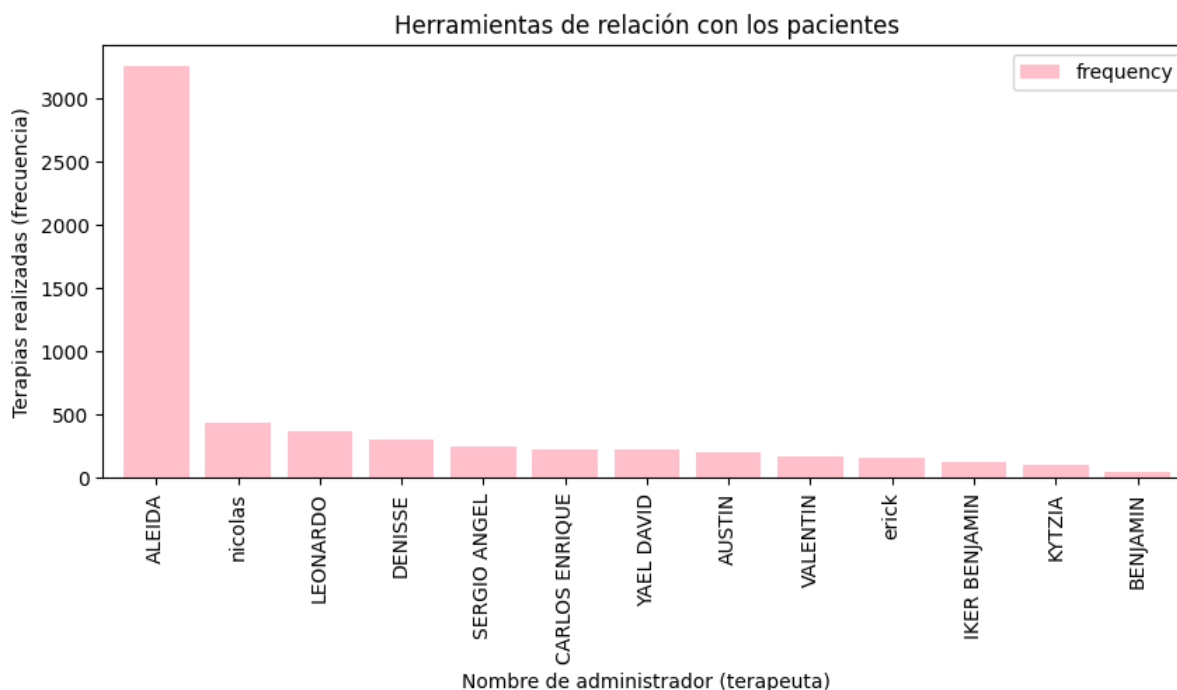
Transformación de la información en modelos gráficos

Análisis gráfico de personas relacionadas

Para el análisis de las columnas tipo categóricas, tomamos en cuenta su total general (siete), sin embargo, para este caso particular, se utilizará el gráfico de mayor conveniencia para la visualización más comprensible de los datos.

Interacciones con el administrador (terapeuta):

Se puede observar que quién tiene más registros es la terapeuta Aleida, por lo que nos podríamos acercar a ella para verificar que tipos de datos son más importante a la hora de analizar los resultados de un paciente y cómo podríamos mejorar el servicio.

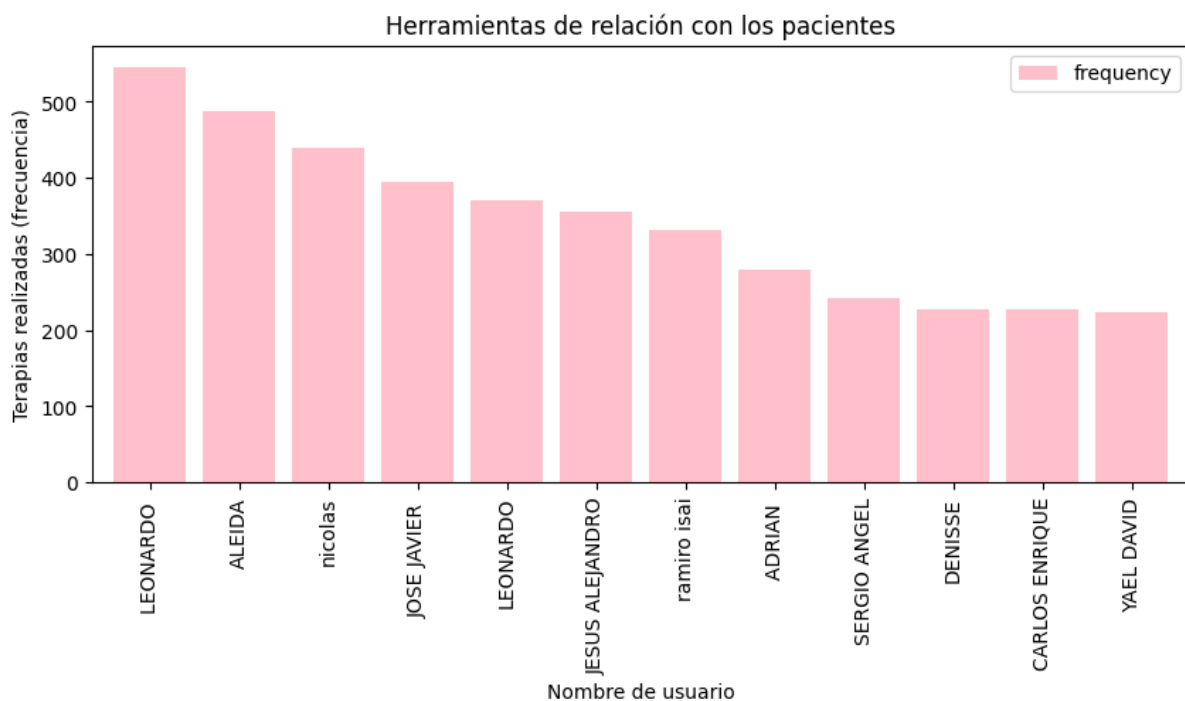


Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de barras que muestra la interacción de los terapeutas con las personas tratadas (sesiones), [marzo de 2025].

Se puede identificar que, “ALEIDA” es la institución o terapeuta que más ha interactuado con pacientes a través de las sesiones o terapias realizadas, además, hay un cambio abrupto después del mismo, esto puede significar que, es la empresa con más antigüedad o también, con la que más se han hecho pruebas antes de proceder con los pacientes reales.

Interacciones con los pacientes (personas tratadas):

A través de esta gráfica podemos encontrar patrones de comportamiento de acuerdo con el usuario y cuántas veces ha asistido a una terapia, y así mismo poder entablar qué tipo de discapacidad tiene, si bien no se puede generalizar, pero sí encontrar características específicas.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de barras que muestra la interacción de los clientes con los videojuegos (por sesiones), [marzo de 2025].

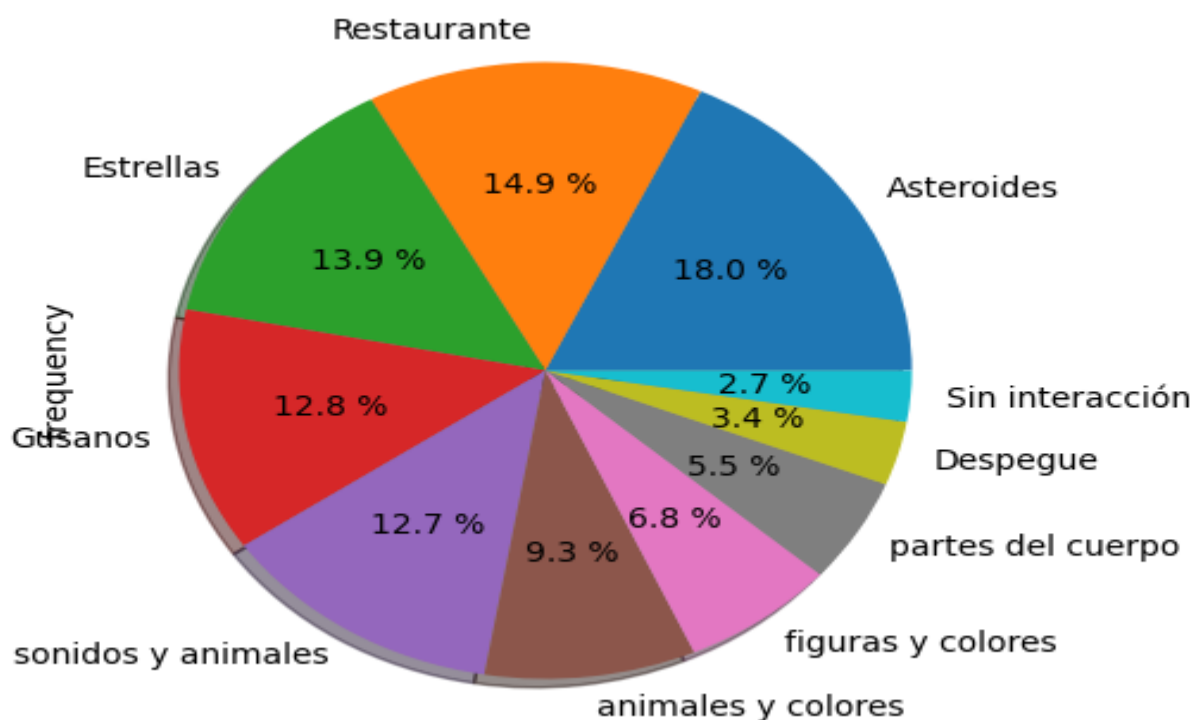
En esta nueva gráfica, se nota que la diferencia no es significativa entre pacientes interactuando con los videojuegos, pero podemos notar que, el cliente/paciente que más ha interactuado con los mismos, es “Leonardo”, con un total de casi 550 terapias realizadas (sesiones), seguido de “ALEIDA”, en donde se genera la siguiente pregunta: “¿ALEIDA es una institución, terapeuta o paciente?”, misma que sería interesante conocer recopilando información a través del Wuupi.

Análisis de juegos y sus características

Wuupi cuenta con bastantes minijuegos terapéuticos que permiten a los terapeutas medir la progresión de sus pacientes, muchos de los videojuegos se encuentran en fases preliminares, por lo que algunos son más usados que otros (preferidos por los terapeutas u organizaciones que adquieren este servicio).

A continuación, se puede observar en la siguiente gráfica creada, cómo es que existen tres principales juegos más utilizados, siendo estos últimos asteroides, restaurante, y estrellas respectivamente. Esto se puede deberse a un enfoque más terapéutico por parte del juego, preferencias del terapeuta o institución o bien, como ya se ha explicado antes, debido a que muchos juegos aún se encuentran en fases preliminares de desarrollo.

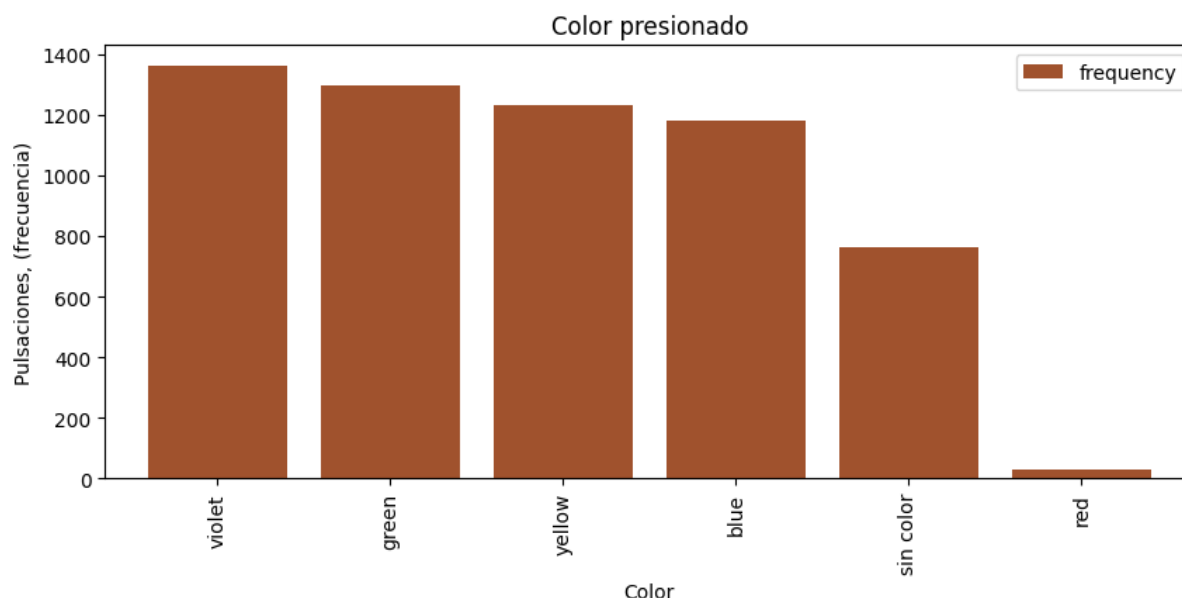
Se utiliza una gráfica de pastel para poder visualizar de mejor manera los datos presentados.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de pastel que muestra la distribución y porcentaje de uso de los videojuegos (en sesiones realizadas), [marzo de 2025].

Color presionado:

En esta gráfica, usamos la de barras para observar la comparación entre los colores. La mayoría tiene una gran participación dentro de la plataforma, pero el rojo es el que casi nunca lo aprieta porque si los pacientes lo hacen, se eliminan del juego.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de barras que muestra el número de pulsaciones realizadas en cada botón (en sesiones realizadas), [marzo de 2025].

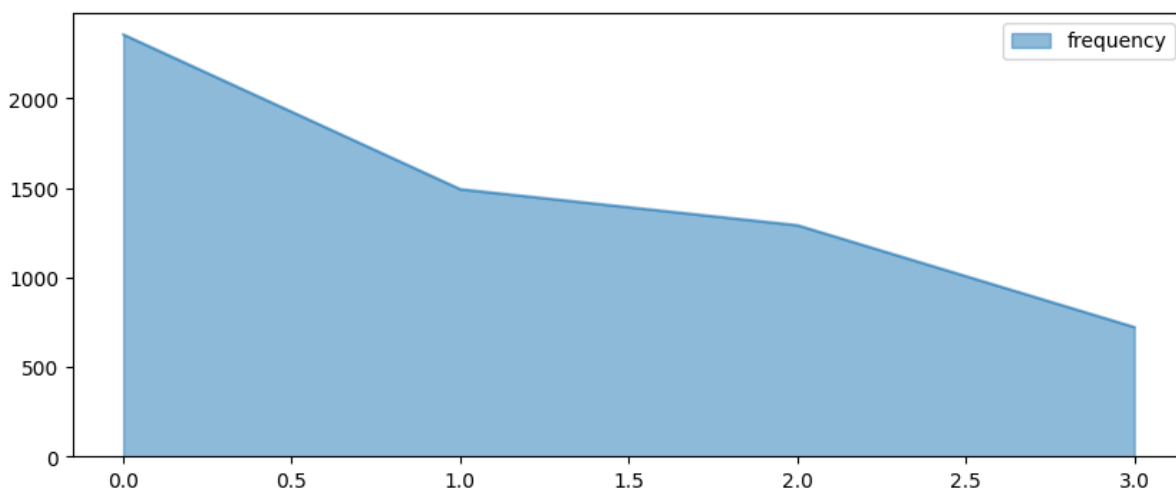
Podemos notar, que el botón rojo es el menos pulsado si se toman en cuenta todas las sesiones, esto debido a que el botón rojo, dentro de los videojuegos, solo se usa en situaciones específicas en donde, el paciente, no tiene que presionar, es decir, el pequeño porcentaje de pulsaciones que se muestran dentro de la gráfica, son errores de los pacientes dentro de sus participaciones en las sesiones creadas por los terapeutas.

También se puede notar que, el botón más pulsado por los pacientes es el morado o violeta, seguido del botón verde, a partir de ahí, todos los botones se pulsan de manera parecida, lo que representa que el algoritmo de los videojuegos se encuentra bien diseñado y balanceado para crear un ambiente aleatorio.

Dificultad:

Aquí usamos un gráfico de área porque podemos observar que tanto incrementa o disminuye la frecuencia de acuerdo con el nivel o episodio.

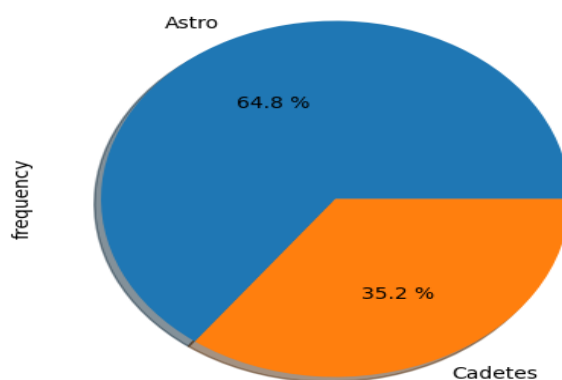
Observamos que conforme el episodio va incrementando, la frecuencia de su uso va disminuyendo, esto se puede interpretar de manera sencilla, pues a mayor dificultad, menos porcentaje de pacientes acceden a los nuevos capítulos.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica que muestra la disminución de actividad conforme al nivel de dificultad (total de sesiones), [marzo de 2025].

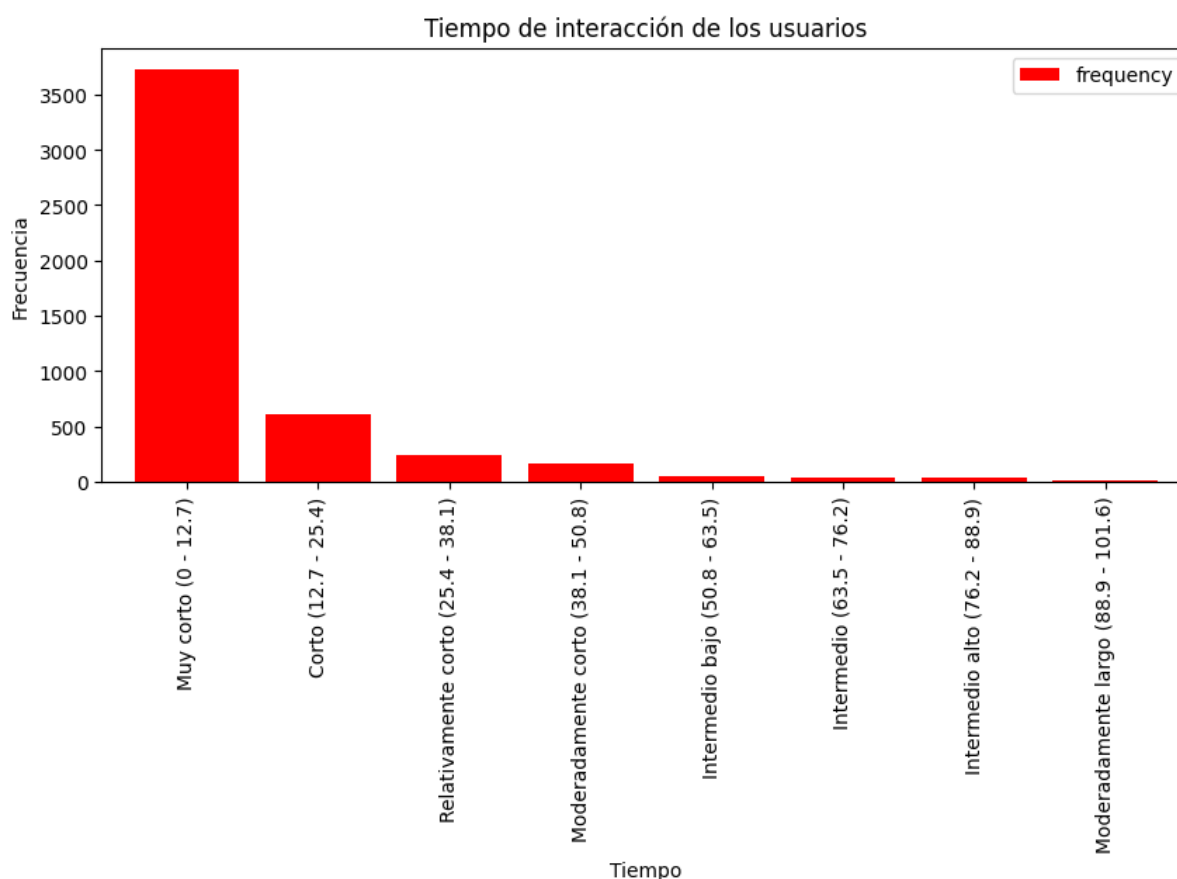
Juego:

Escogimos una de pastel para observar la comparación entre el juego de Astros y cadetes, y podemos observar una gran frecuencia en el de Astro, dando a entender que es el más preferido dentro de los jugadores de acuerdo con el tipo de paciente y sus necesidades de terapia.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de pastel que muestra la distribución de uso entre los dos videojuegos más usados, (marzo de 2025).

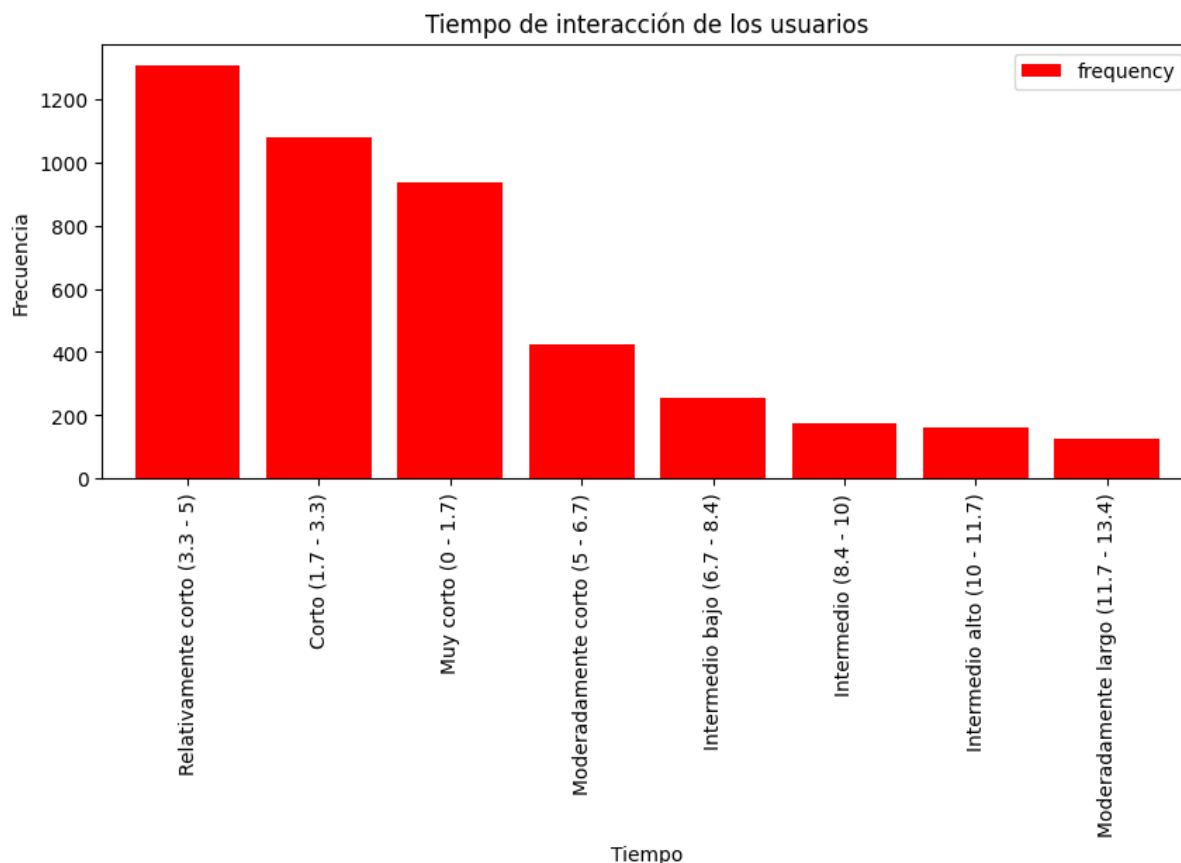
El tiempo de interacción se dividió en subcategorías con rangos continuos, con el fin de facilitar la interpretación en métodos visuales de la información, a continuación, se muestran las categorías realizadas para el tiempo de interacción, en donde, el mayor número de interacciones, las tienen los primeros lugares.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de barras que muestra la distribución en rangos, el tiempo de interacción (por cada sesión individual), [marzo de 2025].

Podemos observar, que generalmente, las sesiones son cortas, en un periodo de 0 a 12 minutos sin exceder ese límite, fuera de este las cifras caen considerablemente, pasando de casi 4000 a menos de mil las interacciones que se hacen en un tiempo superior al ya antes mencionado.

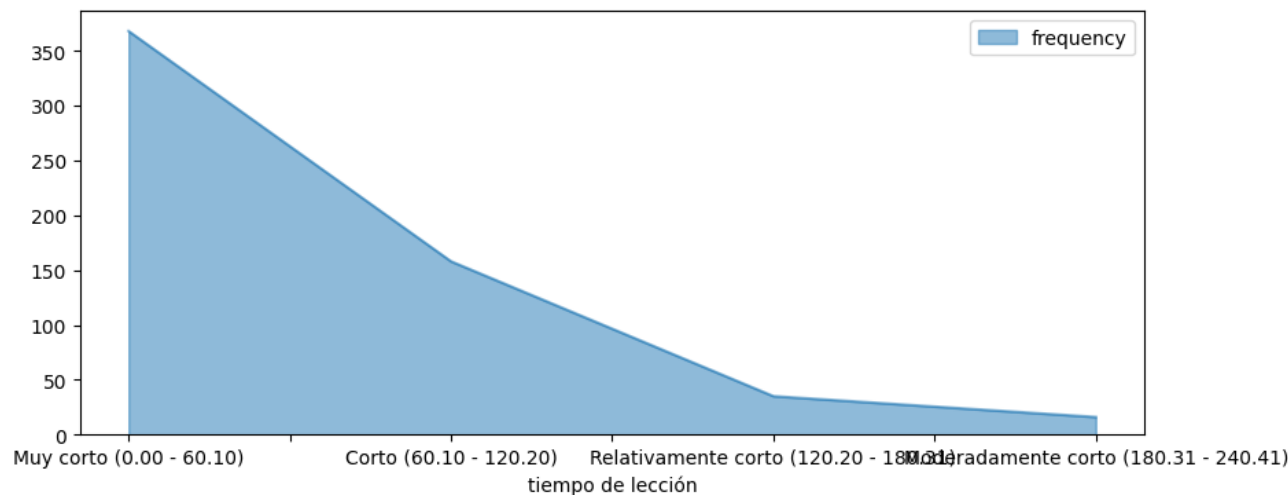
Con el fin de optimizar aún más los rangos realizados y en consecuente, poder visualizar de mejor manera los cambios que se tienen, se redujo el rango de interacción, por lo que ahora, se observa de mejor manera el tipo de cambio entre columnas dentro de la gráfica utilizada.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de barras que muestra la distribución en rangos, el tiempo de interacción (por cada sesión individual), [marzo de 2025].

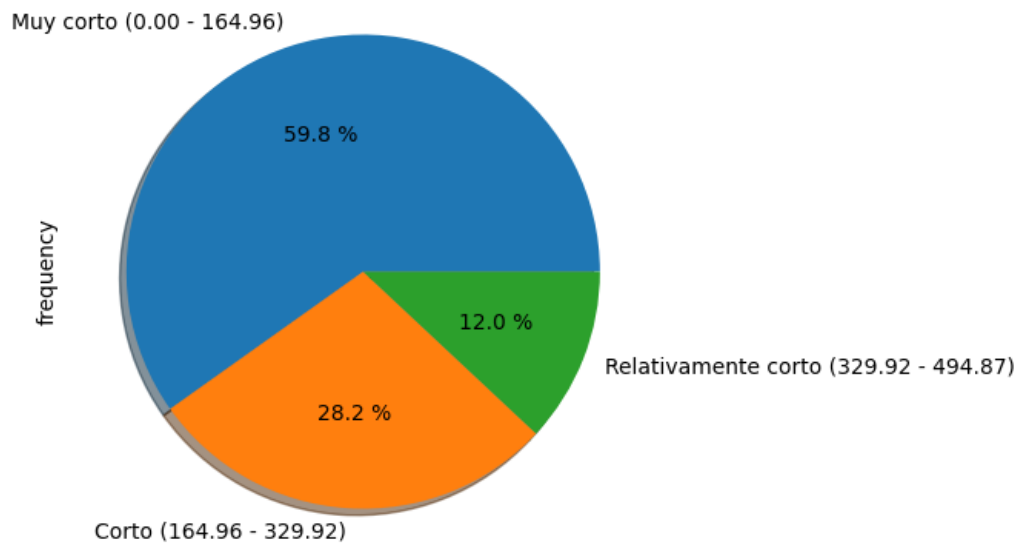
En esta nueva gráfica, podemos observar cómo, dentro del rango “de 0 a 13 minutos” (primero de los ocho rangos presentados en la anterior gráfica), se generan ocho rangos nuevos, para poder observar la distribución de la concentración total de casi 95% de la población o muestra, es decir, observamos la nueva distribución que toma el 95% de los datos concentrados en el primer rango creado en la primera gráfica presentada.

A continuación, se muestra el tiempo de lección por intervalos creados, es decir, el número de terapias a pacientes. Nuevamente observando que el tiempo aumenta de acuerdo con el intervalo categórico otorgado a cada posición.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica que muestra la distribución en rangos, del tiempo de consulta, [marzo de 2025].

Finalmente, se realiza una gráfica de pastel para poder observar el tiempo de cada sesión en comparación con las anteriores, pudiendo observar y predecir futuros incidentes.



Elaboración propia, (con base en la base de datos principal de Wuupi): gráfica de pastel que muestra la distribución de los tiempos de consulta más frecuentes, [marzo de 2025].

Bibliografía

Wuupi. (s.f.). Wuupi. Recuperado el 9 de marzo de 2025, de
<https://wuupi.com.mx/>

El Economista. (2023, 10 de octubre). Conoce los 4 proyectos ganadores de los UP Innovation Awards 2023. Recuperado el 9 de marzo de 2025, de
<https://www.eleconomista.com.mx/el-empresario/Conoce-los-4-proyectos-ganadores-de-los-UP-Innovation-Awards-2023-20231010-0113.html>

Valor Compartido. (2023, 4 de diciembre). Finalizan mentorías a emprendimientos sociales tecnológicos de T-Systems. Recuperado el 9 de marzo de 2025, de
<https://valor-compartido.com/finalizan-mentorias-a-emprendimientos-sociales-tecnologicos-de-t-systems/>