**Logotipo

Descripción generada automáticamente con confianza media**

**TRABAJO FINAL DE ESTADÍSTICA INFERENCIAL**

**INTEGRANTES:  
ZHARICK MOLINA**

**RICARDO WATS**

**HENRY SAENZ**

**NICOLL FONTALVO**

**DOCENTE:**

**NATALIA HERNANDEZ-VARGAS**

**UNIVERSIDAD DEL NORTE**

**BARRANQUILLA-ATLÁNTICO**

**28 DE MAYO DE 2023**

**Contexto disciplinar**

En la era contemporánea, la estadística se destaca como uno de los pilares de enorme aporte para la investigación, más valiosos en el mundo de la ciencia, ya que permite relacionar supuestos con el mundo real y comprobar la relevancia y certeza de algunos fenómenos de estudio para el campo de aplicación objetivo, bien sea de índole social, económico, político, médico, etc. Si bien, también es sabido que dentro del campo de acción de la ciencia de datos se destaca a la estadística como una de las ramas de aplicación de mayor influencia, se debe generar una compaginación más ecuánime, como lo expresa Weihs (2018) “únicamente por medio del balance entre las ciencias computacionales y la estadística, es que se puede llegar a resultados exitosos”, obteniendo como puentes de acción, la validación de herramientas computacionales modernas y herramientas que involucren la participación de la ciencia de datos en la práctica de la estadística.

Desde una óptica más cercana, la estadística toma elementos propios de sí misma para validar y cuestionar fenómenos de estudio generales y específicos, partiendo del planteamiento de hipótesis, los cuales son adaptados para su agilización en modelos computacionales más eficientes dentro de los que participa la ciencia de datos para la optimización de estos, del mismo modo, para la estadística no es diferencial el campo de acción de la misma, pues es un elemento transversal de acción multidisciplinaria, por lo que es válido, incluso indagar fenómenos que vayan desde enfermedades endémicas y sus causas hasta la plasticidad neural del cerebro.

Un estudio titulado “*Training Effects on Dual-Task Performance: Are There Age-Related Differences in Plasticity of Attentional Control?”* indagaba acerca de la concentración y plasticidad cerebral de un grupo etario específico de adultos y jóvenes divididos entre hombres y mujeres, para prescribir los patrones y métodos de enseñanza que mayormente promueven la plasticidad neuronal y la memoria. Durante la realización del estudio se implementó diversas técnicas de obtención de datos de manera cuantitativa para el muestreo de los experimentos y de las pruebas que se aplicarían a los individuos, dentro de las cuales se incluía una sesión de preentrenamiento, entrenamiento y post entrenamiento. Estas, inducían al usuario a preparar su cerebro de manera eficiente y así obtener y evaluar los resultados deseados. Cada uno de los participantes completó 1 hora de pruebas neuropsicológicas en las cuales respondían preguntas acerca de su salud mental y de sus datos demográficos, para introducir a los investigadores en un perfil más específico. El segundo día de experimento los participantes completaron una sesión de preentrenamiento que tardó alrededor de 1 hora y el protocolo de este fue clasificar los grupos según el rango etario; cuya variable se dividía en VP y FP (Young adults and Old adults ) previo al inicio de las pruebas de entrenamiento para tener una clasificación y una muestra antes y después de medias independientes para los resultados.

Durante la fase de entrenamientos se realizaron 5 sesiones de testeo y medición de las diferentes pruebas sobre la concentración, la retención, la memoria y la división de las tareas que se realizó en un rango de tiempo aproximado de una semana cuyo tiempo de realización de cada tarea tenía en promedio 1 hora. En la última fase del estudio correspondiente al post entrenamiento pasadas dos semanas se analizaron y contrastaron los resultados de los participantes con un test adicional que permitiría comparar las medias de los resultados del entrenamiento con los resultados de este nuevo post entrenamiento (Bherer & Hramer, 2005).

Según el estudio de Bherer y Kramer (2005), “El método de recolección y obtención de los resultados para el muestreo realizado implementó bloques de trabajo, es decir, división de los grupos de edad para aplicar pruebas de preguntas que combinaban 20 preguntas de lógica y 20 preguntas de observación e intuición, asimismo otros tipos de pruebas como la retención visual y auditoria y la combinación de estímulos”. Cada uno de los resultados de esta prueba arrojan un puntaje de acuerdo con un rango previamente estipulado por los investigadores que permitirá clasificar la calidad de las respuestas y la apropiada distribución de las tareas de los participantes de la prueba. Estos resultados numéricos son pasados por un filtro con el modelo de clasificación ‘ANOVAs’ que comparaba los puntajes con una función matemática para el éxito del análisis neuropsicológico.

Para la realización del estudio, se tomó como referencia el anterior artículo que trata sobre la atención, la atención residual y la plasticidad neuronal. Se desarrolló un experimento titulado ‘SNEK’ el cual involucra la reacción y la memoria de los individuos por medio de un videojuego con lo que cada participante llevo a cabo un *perform* de acuerdo con 1 de los 4 grupos en los que se dividió la población de muestra con el fin de contrastar condiciones como la cercanía de la pantalla y los colores del juego con el resultado y así evaluar el tiempo de reacción y la cantidad de obstáculos evadidos. El experimento toma como eje de evaluación dos variables del experimento; el color de fondo (Blanco o negro), además de tener en cuenta la edad de los encuestados y los antecedentes que tienen con un juego similar. En la *tabla 1* se muestra la clasificación de las variables implementadas junto con las variables de testeo de la prueba.

**Tabla 1**

*Definición y nombre de variables*

|  |  |
| --- | --- |
| Nombre | Definición |
| Edad | Años de edad. |
| Sexo | Sexo asignado al nacer. |
| AntC | Antecedentes jugando videojuegos. |
| ColorF | Color de fondo en el juego. |
| Score | Puntaje obtenido en el juego. |
| Treac | Cantidad de veces de uso de teclas. |
| CantObs | Cantidad de obstáculos presentados. |

**Planteamiento de hipótesis.**

**Hipótesis 1.**

* Hipótesis estadística: Existe correlación entre las variables del puntaje obtenido en el juego y del tiempo de reacción.
* Hipótesis nula: No existe correlación entre las variables del puntaje obtenido en el juego y del tiempo de reacción.

**Hipótesis 2.**

* Hipótesis estadística: Existe correlación entre la cantidad de obstáculos y el tiempo de reacción.
* Hipótesis nula: No existe correlación entre las variables de la cantidad de obstáculos y del tiempo de reacción.

**Hipótesis 3.**

* Hipótesis estadística: Existen diferencia entre las medias del puntaje obtenido de los hombres y las mujeres.
* Hipótesis nula: No existen diferencia entre las medias del puntaje obtenido de los hombres y las mujeres.

**Hipótesis 4.**

* Hipótesis estadística: Existe diferencia entre las medias de los antecedentes del jugador y la edad de este.
* Hipótesis nula: No existe diferencia entre las medias de los antecedentes del jugador y la edad de este.

**Hipótesis 5.**

* Hipótesis estadística: Existe asociación entre el color del fondo del juego y el sexo del jugador.
* Hipótesis nula: Existe independencia entre el color de fondo del juego y el sexo del jugador.

**Métodos y Resultados.**

**Método de Recolección de Datos.**

Para la recolección de los datos necesarios para la realización del estudio se implementó un formulario de uso único por los encuestadores con base a las respuestas de los individuos del experimento con el fin de evitar errores al completar la información. A cada uno de los participantes se les asignó una serie de características gráficas específicas de interés dentro del juego de manera aleatoria. A continuación, en la tabla 2 se muestra la ficha técnica de la encuesta realizada.

**Tabla 2**

*Ficha técnica de la encuesta*

|  |  |
| --- | --- |
| Realizada por: | Zharick Molina, Henry Sáenz, Nicoll Fontalvo y Ricardo Watts. |
| Nombre de la encuesta: | Plasticidad Neuronal y Agilidad al jugar “Snek”. |
| Población: | Hombres y mujeres pertenecientes a cualquier rango de edad con la habilidad de aprender a jugar “Snek”. |
| Muestra: | Ciudadanos colombianos con la habilidad de aprender a jugar “Snek”. |
| Tamaño de la muestra: | Ochenta (80) personas. |
| Fecha de recolección de datos: | 8 de mayo de 2023 a 17 de mayo de 2023. |
| Técnica de recolección de datos: | Encuesta cara a cara. |
| Objetivo de la encuesta: | Recoger datos suficientes para afirmar o descartar hipótesis relacionadas a la plasticidad neuronal y agilidad de los encuestados en el juego “SNEK”. |
| Número de preguntas formuladas: | Siete (7). |
| Tipo de preguntas aplicadas: | Abiertas con respuestas cortas. |

El estudio engloba una totalidad de 7 variables de interés, cuya información fue completada a partir de la encuesta realizada, estas fueron el centro de la comprobación de 5 pruebas de hipótesis, donde cada una de las variables está relacionada a una pregunta de la encuesta con el fin de determinar si existía alguna relación entre las mismas. En la tabla 3 se muestra la definición de cada una de las variables.

**Tabla 3**

*Definición de variables implicadas en el estudio.*

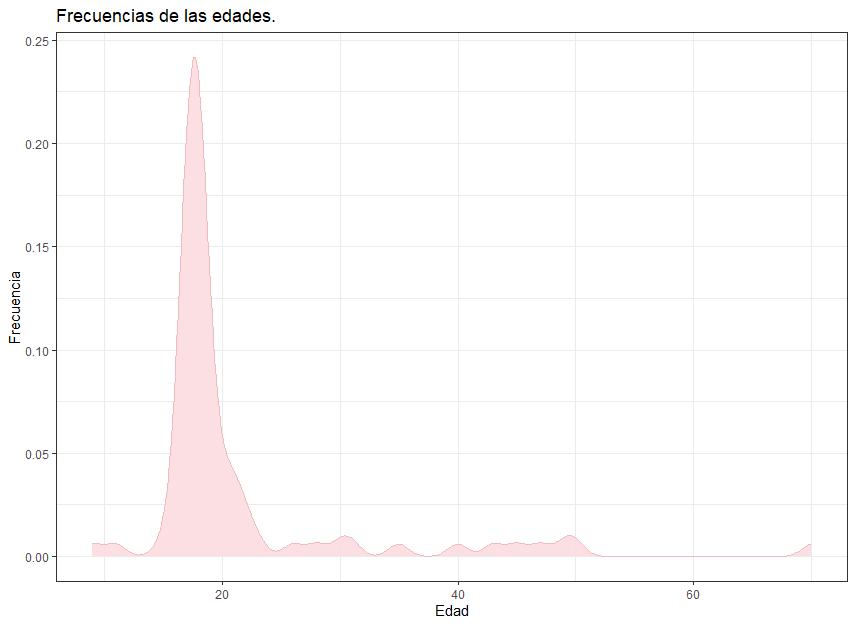
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nombre | Definición | Pregunta | Tipo | Categorías |
| Edad | Años de edad. | Edad | Numérica | - |
| Sexo | Sexo asignado al nacer. | ¿Cuál es el sexo del jugador? | Categórica | Hombre  Mujer |
| AntC | Antecedentes jugando videojuegos. | ¿Ha jugado anteriormente? | Categórica | Sí  No |
| ColorF | Color de fondo en el juego. | ¿Qué color de fondo se utilizó? | Categórica | Blanco  Negro |
| Score | Puntaje obtenido en el juego. | ¿Cuál fue el score obtenido? | Numérica | - |
| Treac | Cantidad de veces de uso de teclas. | ¿Cuál fue el tiempo promedio de reacción? | Numérica | - |
| CantObs | Cantidad de obstáculos presentados. | ¿Cuántos obstáculos se presentaron? | Numérica | - |

**Descripción de la Muestra.**

A continuación, en las figuras 1, figura 2, figura 3, figura 4, figura 5, figura 6 y figura 7 se describe gráficamente la muestra con un gráfico respectivo para cada variable involucrada.

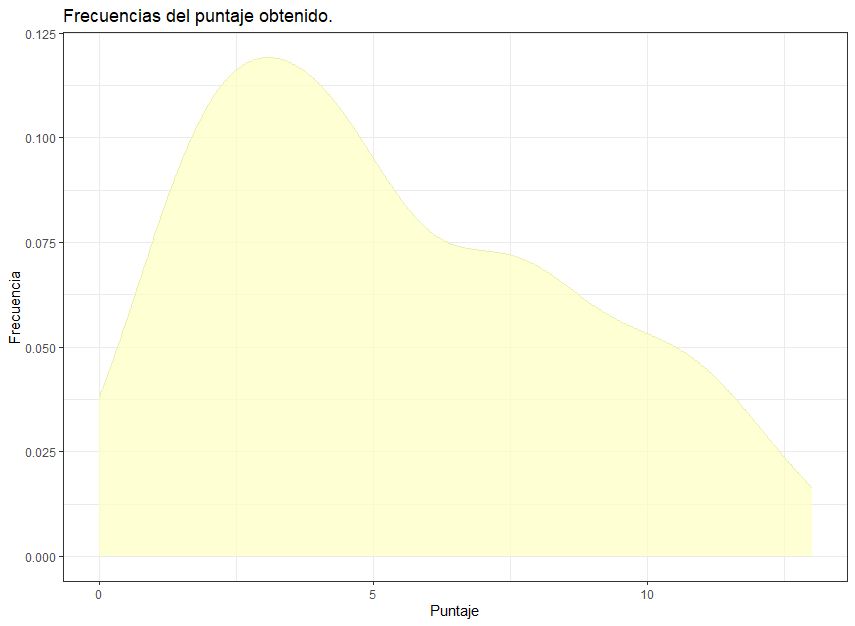
**Figura 1**

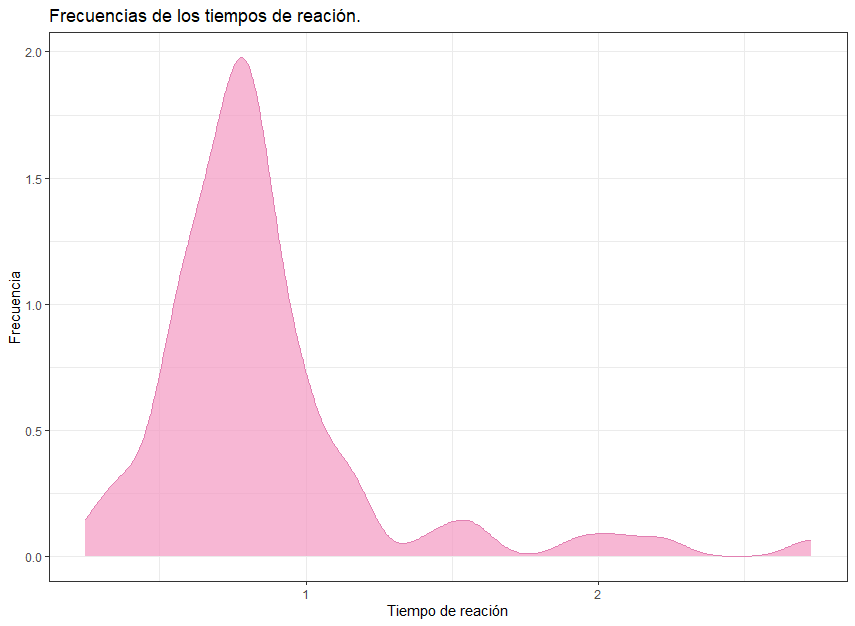
*Frecuencia de edades*



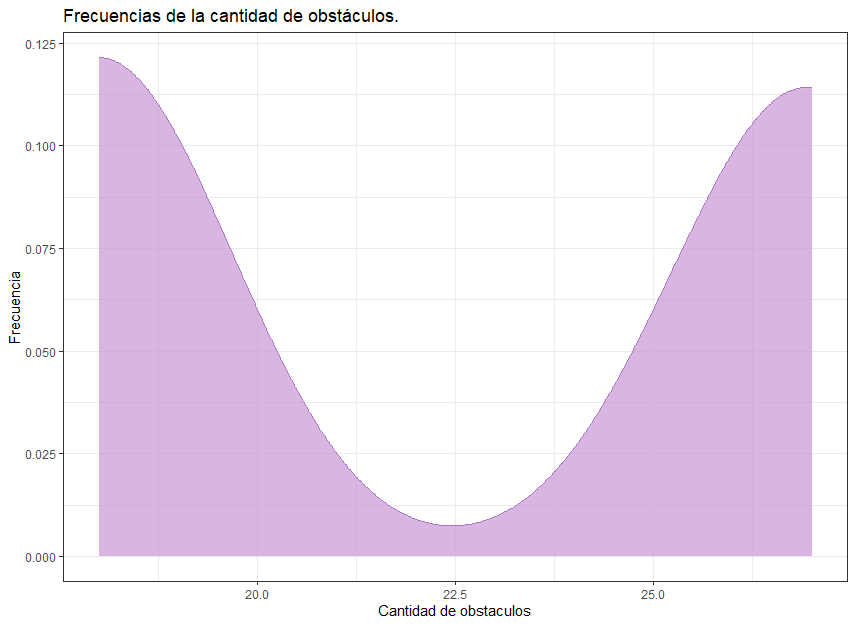
**Figura 2**

*Frecuencia del puntaje obtenido*



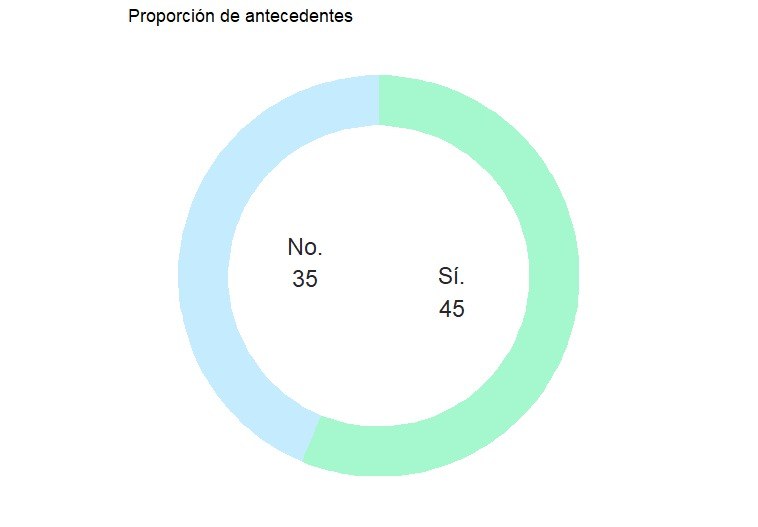
**Figura 3**

*Frecuencia de los tiempos de reacción*

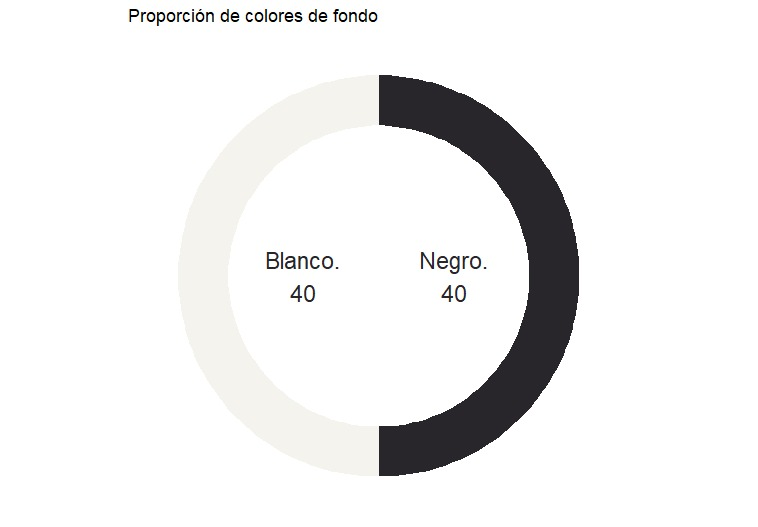
**Figura 4**

*Frecuencia de la cantidad de obstáculos*

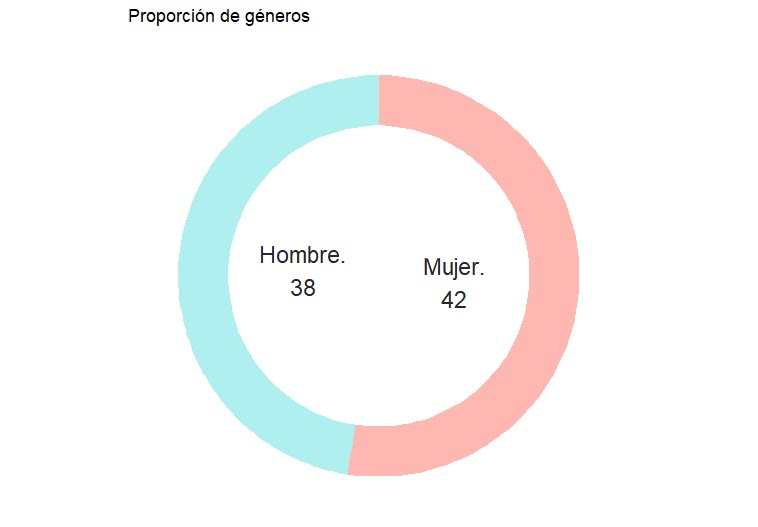
**Figura 5**

*Proporción de antecedentes*

**Figura 6**

*Proporción de colores de fondo*

**Figura 7**

*Proporción de géneros*

**Prueba de Normalidad.**

Con el fin de analizar si los datos recolectados provienen de una distribución normal (y cumplir con las condiciones de algunas pruebas como Correlación de Pearson y Comparación de medias para muestras independientes), se realizó la prueba de Kolmogorov-Smirnov para las variables numéricas. En la tabla 4 se muestran los resultados obtenidos.

**Tabla 4**

*Prueba de normalidad de las variables implicadas en el estudio*

|  |  |  |
| --- | --- | --- |
| Variable | D | P-valor |
| Edad | 0.32888 | 6.094e-08 |
| Score | 0.14812 | 0.05977 |
| Treac | 0.14449 | 0.07084 |
| CantObs | 0.34606 | 9.532e-09 |

Debido a los resultados mostrados anteriormente, se llega a las siguientes conclusiones:

1. La variable Edad no proviene de una distribución normal.
2. La variable CantObs (cantidad de obstáculos presentados) no proviene de una distribución normal.
3. La variable Score (puntaje obtenido en el juego) no proviene de una distribución normal con un nivel de significancia del 10%.
4. La variable Treac no proviene de una distribución normal con un nivel de significancia del 10%.

Por fines del estudio, se tomaron todas las variables como provenientes de distribuciones normales cumpliendo así las condiciones de las pruebas necesarias.

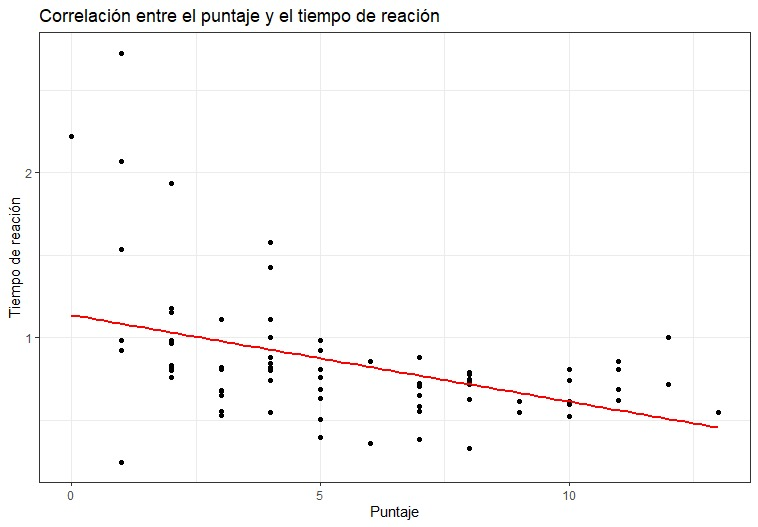
**Plan de Análisis de Resultados.**

Se realizaron distintas pruebas estadísticas teniendo en cuenta la naturaleza de las variables implicadas en las hipótesis anteriormente definidas y sus respectivos supuestos. En cada una de las pruebas se hizo uso del software estadístico R.

Se realizó una prueba de Correlación de Pearson para comprobar la hipótesis 1: *existe correlación entre las variables del puntaje obtenido en el juego y del tiempo de reacción*, donde se halló que existe correlación entre el puntaje obtenido y el tiempo de reacción (t = -4.2305, df = 78, p-value = 6.297e-05) como se muestra en la figura 8.

**Figura 8.**

*Correlación entre las variables puntaje obtenido y tiempo de reacción*



Posteriormente, se realizó una prueba de Correlación de Pearson para comprobar la hipótesis 2; *existe correlación entre la cantidad de obstáculos y el tiempo de reacción*. Donde se concluyó que la correlación entre las variables es nula para un nivel de significancia del 10% (t = 1.3355, df = 78, p-value = 0.1856) esto se observa en la figura 9.

**Figura 9**

*Correlación entre las variables cantidad de obstáculos y tiempo de reacción.*

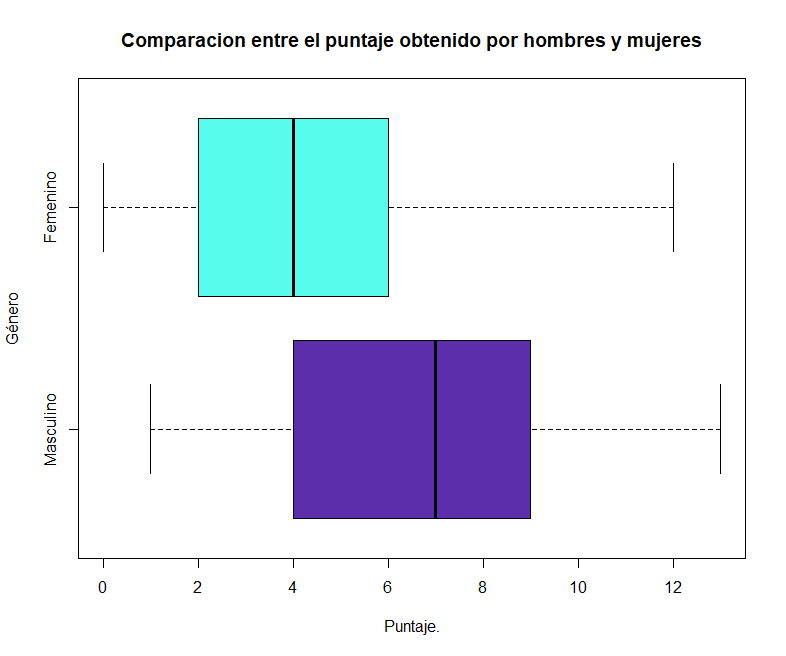
Gráfico, Gráfico de líneas

Descripción generada automáticamente

También se hizo uso de la prueba de Welch para comprobar la hipótesis 3; *Existe diferencia entre las medias del puntaje obtenido de los hombres y las mujeres*. Con lo que se concluyó que sí existe diferencia entre las medias de los resultados ya que el intervalo de confianza es mayor a cero (0.5631783, 3.421784) como se evidencia en la figura 10.

**Figura 10.**

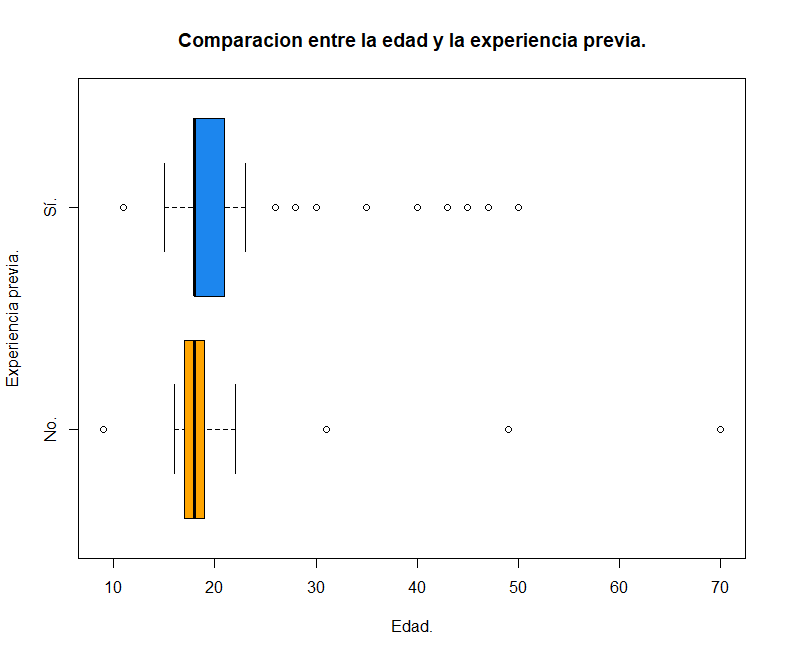
*Diferencia entre las medias del puntaje obtenido de hombres y mujeres.*



Para la hipótesis 4; *Existe diferencia entre las medias de los antecedentes del jugador y la edad de este*, se implementó nuevamente la prueba de Welch con el que se analiza el intervalo de confianza de la media, con lo que se obtuvo que dentro del intervalo se incluye al 0 (-2.936265, 2.24029), lo que es criterio suficiente para concluir que no se puede determinar si hay o no una relación, así mismo se evidencia la distribución en la figura 11.

**Figura 11.**

*Comparación de las edades y si existen antecedentes con un juego parecido.*



Por último, para la hipótesis 5; *existe asociación entre el color del fondo del juego y el sexo del jugador*, se llevó a cabo la prueba Chi Cuadrado de Pearson que arrojó como resultado que existe independencia entre el color de fondo del juego y el sexo del jugador con un p-valor mayor a todos los niveles de significancia (X-squared = 0.050125, df = 1, p-value = 0.8228) como se evidencia en la gráfica 12.

**Figura 12**

*Asociación entre el color de fondo del juego y el sexo del jugador.*

Gráfico, Gráfico de barras

Descripción generada automáticamente

**Discusión, Conclusiones y Recomendaciones.**

**Resumen de la discusión.**

La plasticidad neuronal ha sido temática central de estudio de múltiples artículos e investigación partiendo del uso de técnicas de análisis con enfoques diversos, desde estudios tradicionales de la flexibilidad creativa hasta pruebas y test psicológicos. Los video juegos, por su parte, también han representado un papel participativo y relativamente novedoso en el entorno de estudio de esta temática. Investigaciones como *Training Effects on Dual-Task Performance: Are There Age-Related Differences in Plasticity of Attentional Control?*, implementan los resultados tomando como herramienta los video juegos para la aplicación de estos, dentro de dicho estudio tomamos como ejemplo el modelo de video juego como herramienta de muestreo.

**Evaluación de la muestra.**

Para realizar este tipo de estudios donde se quiere comprobar hipótesis que se tienen de una población a partir de una pequeña parte de ella a la que se tiene acceso y a la que le llamamos muestra, para que esto se pueda hacerse con un nivel de veracidad se necesita minimizar cada uno de los sesgos que podrían presentarse al seleccionar una muestra. En este caso la muestra seleccionada fueron personas del común, escogidas al azar por los encuestadores, sin embargo, como se mencionó al comienzo de este informe, para ejecutar la encuesta debía de haber alguien del equipo para digitar los datos y sabemos que el contexto en el que se desarrollan los encuestadores y de donde escogieron los participantes, no representa a toda la población a la que se quería llegar. Así mismo si se tiene en cuenta las edades de los participantes se puede ver que la mayoría de los encuestados son personas jóvenes y hay muy pocas personas que representen a otras edades.

**Recomendaciones para próximos estudios relacionados.**

La realización de este trabajo de investigación nos permitió ver más de cerca como se trabaja de manera real y desde lo más inicial como lo es establecer una población objetivo, siguiendo con la elección de la muestra y la recolección de los datos a partir de esta experiencia podemos resaltar dos recomendaciones importantes al momento de trabajar en un estudio con una temática similar. Tomando como referencia la experiencia vivida con este estudio es importante recalcar la rigurosidad que hay que tener al momento de dividir la muestra si esto es un aspecto importante para el estudio, en este caso se dividió por el sexo y también por las características visuales del juego, es importante intentar hacer que la muestra sea pareja en todos los sentidos, para así poder tener mayor confianza en las afirmaciones que se hagan como conclusiones.

Por otro lado, usar una herramienta diseñada con características propias necesarias para recolectar los datos del estudio facilita mucho el trabajo de encuestar y tabular los datos, por ejemplo, el desarrollo del juego “Snek” fue hecho para arrojar datos en un formato que permitiera trabajar con los datos de manera que fuera fácil la tabulación de los mismos, si esto no hubiese sido de esta manera el trabajo se hubiera extendido mucho más ya que habría que dedicarle mucho más tiempo a la limpieza y estandarización de los datos. De igual forma usar una herramienta de fácil manejo de datos, preferiblemente que tabule los datos según la encuesta, como Google Forms.

Finalmente, el formato del juego a usar es de suma importancia, debido a la facilidad de este para enviárselo a posibles encuestados y así obtener mayor cantidad de datos.

**Referencias**

Bherer, Louis., Kramer, Arthur F., Peterson, Matthew S., Colcombe, Stanley., Erickson, Kirk., Becic, Ensar., *Training Effects on Dual-Task Performance: Are There Age-Related Differences in Plasticity of Attentional Control?*, Psychology and Aging, Vol 20(4), Dec, 2005 pp. 695-709. Publisher: American Psychological Association.

Weihs, C., Ickstadt, K. *Data Science: the impact of statistics. Int J Data Science* Anal **6**, 189–194 (2018). <https://doi.org/10.1007/s41060-018-0102-5>