

Applied Neuroscience for Data Visualization

Rafael Camacho-Aguilar, Máster en Investigación Empresarial¹, Javier Rojas-Segura, Máster en Investigación Empresarial², Marco Martínez-Mora, Máster en Dirección de Empresas³, Jose Martínez-Villavicencio, Doctor en Dirección de Empresas⁴, ^{1, 2,3,4} *Tecnológico de Costa Rica, Costa Rica, fitacam@gmail.com, jarojas@tec.ac.cr, marco.martinez@tec.ac.cr, jomartinez@tec.ac.cr*

Abstract – The exponential increase in complexity and volume of data in the contemporary world has driven the adoption of Visual Analytics (VA) tools to accelerate knowledge extraction through effective visual representation of metrics and trends. This study aims to evaluate their effectiveness in practice, particularly during information analysis, and their influence on decision-making. To conduct this evaluation, physiological response measures from neuroscience, particularly eye-tracking using webcams and Real Eye® software, will be employed during an experimental design intended to compare the effectiveness of visual data analysis with traditional tabular information analysis. The results of this experiment are expected to provide valuable insights for both the academics and practitioners in the field, as well as data users in general.

Keywords—Neuroscience, visual analytics, data analysis, experimental design, innovation.

Neurociencia Aplicada a la Visualización de los Datos

Rafael Camacho-Aguilar, Máster en Investigación Empresarial¹, Javier Rojas-Segura, Máster en Investigación Empresarial², Marco Martínez-Mora, Máster en Dirección de Empresas³, Jose Martínez-Villavicencio, Doctor en Dirección de Empresas⁴,^{1, 2,3,4} Tecnológico de Costa Rica, Costa Rica, fitacam@gmail.com, jarojas@tec.ac.cr, marco.martinez@tec.ac.cr, jomartinez@tec.ac.cr

Resumen— El aumento exponencial de la complejidad y volumen de datos en el mundo contemporáneo ha impulsado la adopción de herramientas de Visual Analytics (VA) como medio para acelerar la extracción de conocimiento a través de una representación visual efectiva de métricas y tendencias. Este estudio se propone evaluar su efectividad en la práctica, particularmente durante el análisis de información, y su influencia en la toma de decisiones. Para llevar a cabo esta evaluación, se emplearán medidas de respuestas fisiológicas de neurociencia, particularmente el seguimiento ocular utilizando cámaras web y el software Real Eye® durante un diseño experimental destinado a comparar la eficacia del análisis visual de datos con el análisis tradicional tabular de información. Se espera que los resultados de este experimento proporcionen conocimientos valiosos tanto para la comunidad académica como para los profesionales en el campo y usuarios de datos en general.

Palabras clave— Neurociencia, visual analytics, análisis de datos, diseño experimental.

I. INTRODUCCIÓN

Existe una creciente demanda de soluciones digitales para la gestión del rendimiento, y las herramientas visuales han experimentado altas tasas de adopción en todas las industrias durante los últimos años a medida que las empresas han buscado mejorar la precisión y la velocidad de sus decisiones a través de una mejor visualización de las métricas y tendencias del rendimiento [1].

Este rápido crecimiento de la demanda de herramientas de visualización se produce principalmente porque la tecnología ha mejorado rápidamente la capacidad de registro de información más no así la capacidad para su procesamiento y análisis. Varios investigadores han descrito este vacío previamente [2], [3], [4], señalando que se debe principalmente al rápido desarrollo de tecnologías de gestión y almacenamiento de datos, y nuestra limitada capacidad humana para generar conocimiento a partir de los datos recopilados. Esta brecha, además, ha llevado a un aumento en la actividad de investigación en el campo del VA como medida para reducir esta brecha [5]. En otras palabras, tenemos mucha más información de la que nos es posible analizar utilizando las herramientas tradicionales.

Es por esto por lo que VA se está convirtiendo cada vez más en una tecnología destacada para las organizaciones, que

buscan obtener conocimientos a partir de conjuntos de datos complejos y de gran tamaño [6].

El término VA fue acuñado por el Departamento de Seguridad Nacional de Estados Unidos, que enfrentó este problema de capacidad y velocidad de análisis durante su lucha contra el terrorismo, razón que los llevo a encomendar al Centro Nacional de Visualización y Análisis (NVACTM por sus siglas en inglés) la tarea de facilitar conocimiento analítico avanzado. Es ahí donde se define al VA como la ciencia del razonamiento facilitado por interfaces visuales interactivas [2]. Definición que aún hoy, continúa evolucionando y ensanchando su alcance. Por ejemplo, si comparamos esta temprana definición con una más contemporánea en donde se define al VA como un enfoque integrado y combinando de: visualización, algorítmicos, y la interacción humana y su razonamiento analítico [5]. Como vemos se han incorporado una amplia gama de dominios y disciplinas en tan solo 20 años de historia.

Esta dinámica expansiva del campo del VA se mantiene, pues algunos autores mencionan la importancia y necesidad de mejorar las herramientas de VA, incluir el apoyo narrativo, es decir que se mejoren no solamente la comprensión de datos, sino también que se brinde un apoyo en la elaboración de una historia, es decir, todo análisis de información además de ser visual, debe contar una historia que lo haga más atractivo a los seres humanos, lo que nos da una idea muy clara del devenir del campo en el futuro cercano, la incorporación del *Story Telling* [7].

De esta manera se pretende evaluar la intersección de dos campos emergentes, el VA, y el análisis de datos, centrándose dentro del entorno empresarial. Este estudio busca experimentar el análisis de datos en tareas sencillas, de una manera más gráfica comparándola con las formas tradicionales de análisis tabular. Siendo el objetivo evaluar técnicas de VA en la práctica, durante el análisis de información, así como sus efectos en la toma de decisiones. A través de la medición de respuestas oculares a diferentes estímulos de análisis de datos. Para lo que será necesario:

- a. Diseñar y realizar experimentos para medir las respuestas fisiológicas de los usuarios al interactuar con técnicas de VA versus tablas tradicionales al analizar información.

b. Determinar si existen superioridad en la toma de decisiones en aquellos grupos que analizan información de manera grafica versus datos tabulares.

c. Evaluar los beneficios y desafíos asociados con la implementación de VA en los sistemas de *Business Intelligence*.

d. Proporcionar recomendaciones para las organizaciones que buscan integrar VA en sus sistemas de *Business Intelligence* para mejorar la toma de decisiones y el rendimiento general, basándose en los resultados de las mediciones fisiológicas.

II. REVISIÓN DE LITERATURA

El surgimiento del VA se puede atribuir en gran medida a la fuerte necesidad de la seguridad nacional de los Estados Unidos de analizar datos complejos, como información incompleta, inconsistente o potencialmente engañosa, desde los ataques terroristas del 11 de septiembre de 2001 [2].

Podemos analizar ambos términos del VA. En primer lugar, la intervención humana, el análisis, emerge como un componente esencial en la evaluación de datos y la generación de conocimiento y es importante mencionar que se espera que dicho conocimiento sea imparcial, es decir, independiente del individuo que lo lleve a cabo, y que pueda ser replicado por otros con resultados consistentes. En segundo lugar, el término visualización se presenta como un concepto ambiguo, pues puede abarcar tanto una disciplina de investigación específica como tecnologías concretas, así como los resultados visuales en sí mismos [3].

Sin embargo, fueron [2] quienes presentan la primera definición ampliamente aceptada para VA, publicada en su libro considerado como seminal dentro del campo, y lo concretaron como "La ciencia del razonamiento analítico asistido por interfaces visuales interactivas" sembrando las bases para toda una nueva etapa de investigación en el campo.

De manera casi paralela al surgimiento del VA, [8] proponía su "Mantra de búsqueda de información visual" introduciendo las bases para para diseños orientados a facilitar la exploración visual de datos, es decir el nacimiento de los tableros de información. Este mantra se describe a continuación: a) descripción general, b) acercamiento y filtros, c) detalles según demanda.

Si bien es cierto, a este punto en la historia el VA ya estaba fundamentado, fue hasta el año 2010 que se adiciona un elemento clave para su desarrollo dentro el campo empresarial, pues se amplía el concepto y se le dota de un modelo que lo define como una combinación de visualizaciones automatizadas e interactivas para una comprensión efectiva, razonamiento y toma de decisiones sobre la base de cuestiones muy grandes y complejas [9], [10]. Esta nueva definición no solamente vincula al VA con la formación de conocimiento, y la toma de decisiones fundamentadas en datos, si no que de alguna manera le otorga el grado de proceso de análisis, por ende, permite a los profesionales explotar nuevo conocimiento esencial para que

las empresas se adapten e innoven en la nueva realidad digital [11].

A. El Proceso de Visual Analytics (PVA)

En este proceso se combinan métodos automatizados de VA, estrechamente ligados con la interacción humana para obtener conocimiento a partir de datos. Es decir, no se trata únicamente de un vaciado de datos en forma gráfica, es en realidad un proceso interactivo e iterativo, de prueba y error, en donde el usuario retroalimenta al modelo de análisis basándose en los datos visualizados. En la Fig. 1 podemos analizar dicho proceso, el más aceptado actualmente y que es la base conceptual a la que nos apegaremos en este proyecto.

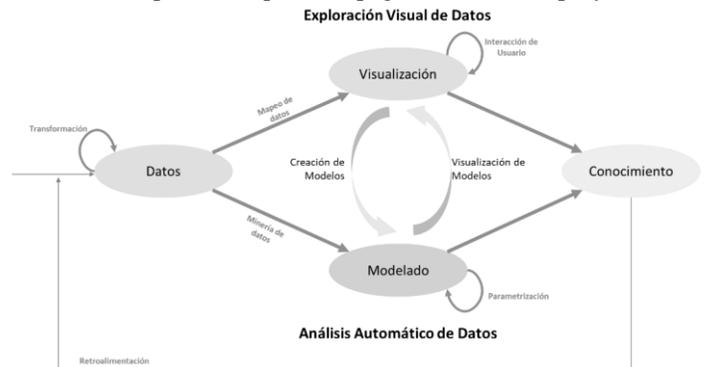


Fig. 1 PROCESO DE VISUAL ANALYTICS (PVA)

Este proceso de sinergia y colaboración entre métodos de exploración visual y análisis automático constituye la esencia distintiva del VA, marcando su diferenciación versus otros métodos de análisis de datos. Dicha integración facilita un refinamiento continuo y una evaluación detallada de los resultados analíticos. Por ende, la combinación de la exploración visual de datos con análisis automáticos modelados puede, con frecuencia, resultar en una mejora sustancial de los resultados analíticos [12].

Podríamos afirmar, que fue posterior a la introducción del PVA, que la mayoría de la producción académica se orienta a su estudio, y la aplicación de este en campos específicos, y al surgimiento de nuevas técnicas para el manejo de conjuntos de datos cada vez más complejos y su análisis automatizado. Por ejemplo, una de las aplicaciones más destacadas de la visualización son los resultados de la modelización climática, simulaciones sofisticadas de fenómenos climáticos complejos con datos de sensores remotos mediante visualizaciones de estructuras que de otro modo serían invisibles [4].

Otro ejemplo de aplicación surge debido al alta tasa de adopción de las redes sociales, que ha creado un dominio difícil de analizar debido al volumen de datos en constante expansión. La visualización de información es clave para resaltar características de interés en las redes sociales y es objeto de estudio del VA [13].

Adicionalmente, la visualización de gráficos dinámicos se centra en el desafío de representar la evolución de las

relaciones entre entidades, representadas en diagramas legibles, escalables y efectivos [14].

Existen también usos prácticos, aplicados a soluciones cotidianas, por ejemplo, [15] propuso un método de VA para la planificación dinámica de la evacuación de edificios, donde el usuario puede visualizar el progreso de evacuación y tomar mejores decisiones en la planificación de la evacuación.

Incluso de manera casi irónica, ha sido descrito como una herramienta útil para el diseño de productos relacionados con VA. Pues se ha concluido que la aplicación del VA para respaldar la toma de decisiones requiere un enfoque mucho más avanzado de comprensión de la situación, así como una proyección precisa de los eventos futuros de las tareas analíticas del usuario [16]. En otras palabras, hay una necesidad inmensa de exploración, incluso en el contexto específico de su aplicación y en el caso de toma de decisiones, la predicción futura es relevante, justamente algo que VA facilita.

Más recientemente, se ha argumentado una necesidad urgente de establecer una estructura común para acelerar la investigación y el desarrollo de nuevas técnicas en VA. Es decir, aún no existe una propuesta que organice y sistematice la producción académica relacionada al VA. Esta necesidad no ha sido completamente valorada ni discutida en investigaciones recientes [5].

Hasta ahora, se ha señalado como la mayoría de las investigaciones anteriores sobre análisis visual se han centrado en comprender las necesidades y los desafíos de los analistas de datos de diversos campos, sin embargo, se sabe menos sobre las tareas y los desafíos de los tomadores de decisiones organizacionales y cómo las herramientas de VA podrían ayudar [17], incluso se ha señalado que el objetivo final del VA es obtener información sobre un problema en cuestión [9], lo que no necesariamente nos remite a estudiar la forma de hacerlo, sino más bien nos remite a un resultado esperado, el conocimiento generado.

Es posible presentar una comprobación práctica a esta teoría expuesta y generar precedentes que impulsen más y mejores investigaciones orientadas a la optimización del VA para la toma de decisiones, lo que constituye nuestro reto actualmente.

Es por ello por lo que se ha definido la comprobación de diferencias significativas entre métodos gráficos utilizados para el análisis de datos versus el proceso tabular tradicional, mediante el planteamiento de las siguientes hipótesis:

H_0 = No hay diferencia significativa en el análisis de datos entre grupos expuestos a estímulos visuales y grupos expuestos a estímulos tabulares.

H_1 = Existe una diferencia significativa en el análisis de datos entre grupos expuestos a estímulos visuales y los grupos expuestos estímulos tabulares.

III. METODOLOGÍA

El enfoque de investigación es cuantitativo, buscando medir el impacto de la integración de técnicas de análisis visual durante el análisis de datos y la toma de decisiones en comparación con técnicas no gráficas. Se propone un diseño de experimentos controlados (DOE) de Psicología Experimental, para medir las respuestas fisiológicas de los usuarios al interactuar con técnicas visuales de análisis versus técnicas no gráficas, además de determinar la tasa de respuestas correctas que tiene cada grupo ante los distintos estímulos. Específicamente presentaron estímulos apareados, uno tabular y uno gráfico de manera aleatoria a cada participante, para medir sus respuestas fisiológicas utilizando equipos de neurociencia y su capacidad de responder correctamente a partir de cada estímulo.

El método de recopilación de datos fue a través de la plataforma en línea RealEye®. Que tiene metodologías para neuromarketing de seguimiento ocular, específicamente el *Face Reading* y el *Eye Tracking*.

A. Población y Muestra

Población (N): Todos los individuos que realizan análisis de datos y que podrían estar expuestos a herramientas visuales y tabulares de información. Esto podría incluir a analistas de datos, científicos de datos, estudiantes de estadística, investigadores, agentes de seguros, oficinistas en general, y otros.

Muestra (n): La muestra debe ser representativa de la población para que permita la generalización de los resultados. Para esto se utilizaron estudiantes de grado de la carrera de Administración de Empresas, así como también egresados practicantes analistas de información de cualquier ámbito. Un total de 30 personas en Costa Rica fueron incluidos en el estudio. No se aplicó cuota de edad ni género.

Tamaño de muestra mínimo deseado:

$$n = \frac{Z^2 * p * (1-p)}{E^2} = \frac{1,645^2 * 0,5 * (1-0,5)}{0,10^2} = 67,651 \cong 68 \text{ impresiones}$$

Donde:

- n es el tamaño de la muestra.
- Z es el valor Z para el nivel de confianza deseado ($Z=1,645$ para 90% de confianza).
- p es la proporción estimada de la población que posee la característica de interés, asumimos que $p=0,5$ (para maximizar el tamaño de la muestra)
- E es el margen de error (10% en este caso)

B. Diseño de Investigación Experimental y Variables Medidas

Diseño con posprueba de valoración de conocimientos para la toma de decisiones (Fig. 2) Los participantes serán expuestos aleatoriamente al tratamiento experimental o de control de manera aleatoria, y posteriormente deberán responder preguntas relacionadas con el estímulo presentado. Cada participante podrá ver un estímulo de forma aleatoria, ya sea visual o tabular. Para garantizar la diversificación del sesgo existirán cuatro exposiciones en total a cada persona, de manera que un mismo participante generara un total de cuatro impresiones, donde cada impresión es producto de un estímulo gráfico o tabular de manera aleatoria.

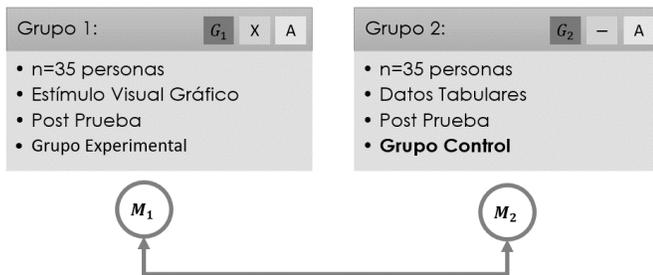


Fig. 2 DISEÑO EXPERIMENTAL

El tratamiento experimental se constituye de un análisis de datos a partir de estímulos gráficos, o bien tableros de visualización. En la Tabla I se muestran las variables del estudio.

TABLA I
VARIABLES DEL ESTUDIO

Tipo	Descripción de la variable
Variable Independiente	Tipo de Estímulo: Método de análisis utilizado Y1 Estímulo Visual Gráfico Y0 Estímulo Tabular
Variables dependientes	Tasa de Aciertos: Número de respuestas correctas al interpretar los datos $TA = \frac{\text{Total Aciertos}}{\text{Total Preguntas}}$ Tasa de Desaciertos: Número de respuestas incorrectas al interpretar los datos $TD = \frac{\text{Total Desaciertos}}{\text{Total Preguntas}}$ X1 = fixation_total_count : representa el número total de fijaciones registradas durante una sesión de rastreo ocular. Una "fijación" ocurre cuando la mirada se detiene en un punto específico. X2 = fixation_average_duration_ms : duración promedio de las fijaciones, medida en milisegundos. X3 = fixation_tfff_ms (Time to First Fixation): Este es el tiempo que toma para que ocurra la primera fijación en un objeto o área específica. X4 = fixation_total_time_spent_ms : Representa el tiempo total gastado en fijaciones durante la sesión de rastreo ocular, medido en milisegundos. X5 = fixation_first_fixation_average_duration_ms : Se refiere a la duración promedio de la primera fijación en cada elemento o área visualizada. X6 = visit_total_count : Es el número total de visitas a un área de interés específica. X7 = revisit_total_count : Cuenta cuántas veces un área de interés ha sido visitada más de una vez. X8 = revisit_average_count : Es el número promedio de veces que las áreas de interés han sido revisadas.

C. Detalle de los Estímulos a Utilizar

Para la elaboración de los estímulos se utilizó información demográfica de países y su nivel de ingresos tomada de [18], sitio *open access* y de libre reproducción. Los gráficos son tomados directamente de su herramienta de visualización, y de manera alternativa se reproducen de manera tabular a partir de la misma información para los estímulos tabulares.

Como un ejemplo de los materiales utilizados se puede observar la Fig. 3, que corresponde a una de las preguntas y sus respectivos elementos de estímulo. Recordemos que en total existen cuatro preguntas cada una con un elemento gráfico y otro tabular que llevan a la misma respuesta, el resto de los estímulos se pueden observar en los anexos del proyecto.



Fig. 3 ESTÍMULOS APAREADOS

IV. RESULTADOS

En total el experimento se aplicó a 30 personas que en total generaron 120 impresiones, o medidas de respuesta a los estímulos que les fueron presentados, tengamos en cuenta que nuestra base de análisis por lo tanto será el número de impresiones para las variables de respuesta fisiológica, y para efectos descriptivos de la muestra y preguntas de toma de

decisiones lo serán el número de personas. En la Fig. 4 se muestran la composición etaria y de genero declaradas.

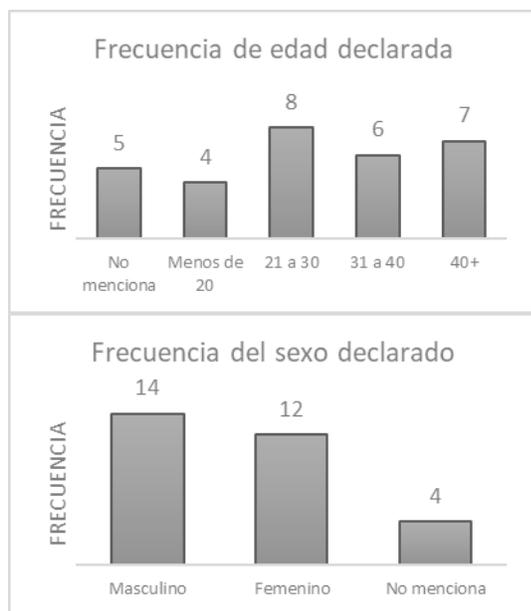


Fig. 4 DEMOGRAFÍA

Ahora bien, en la Tabla II se muestran los resultados de las impresiones totales obtenidas, así como su tipo, como impresión se tiene a la suma de veces que un participante fue expuesto a un estímulo y la recolección de métricas respectiva. Cada participante fue expuesto aleatoriamente a cuatro estímulos, tanto gráficos como tabulares de manera que el resultado obtenido nos proporciona un número suficiente de muestra para realizar conclusiones significativas, inclusive con un mayor margen de error que el propuesto inicialmente.

TABLA II
IMPRESIONES CAPTURADAS

Total Impresiones	Impresiones de Estimulo Gráfico	Impresiones de Estimulo Tabular
120	59	61

A. Mapas de Calor

Uno de los resultados más reveladores se desprende del análisis de los mapas de calor, para esto utilizaremos los estímulos apareados mostrados en la metodología, correspondientes a la pregunta de simetría, ya que todos los comportamientos fueron similares. En los anexos aportaremos todas las preguntas y mapas de calor obtenidos.

El mapa de calor muestra la distribución del comportamiento visual de cada estímulo (ver Fig. 5). En ambos casos fue extenso, es decir los panelistas pusieron su atención en la totalidad del elemento, es decir ambas

herramientas de análisis parecen requerir un esfuerzo visual a este punto. Resalta incluso como los espacios en blanco en ambos casos captaron atención, de manera similar a lo que hizo el enunciado o el espacio de respuesta.

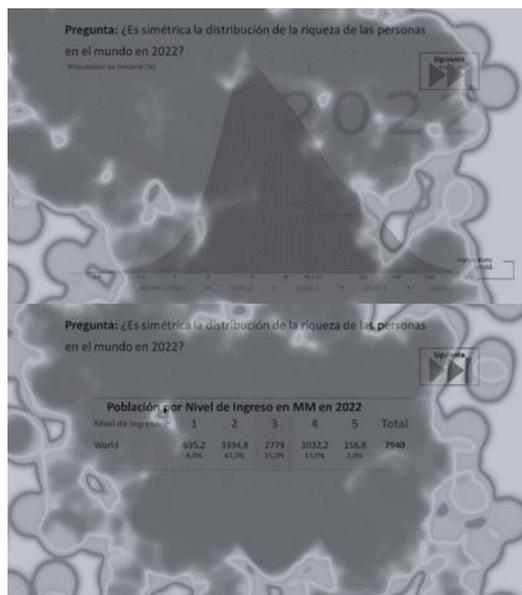


Fig. 5 MAPA DE CALOR: PREGUNTA DE SIMETRÍA

Como área de particular interés, para este pareado de estímulos (así como para todos los demás) se determinó el espacio donde se revela la respuesta a la pregunta del enunciado, de manera tal que se espera una mayor atención ahí, sin embargo, no fue este el resultado obtenido al comparar el mapa de calor con el área de interés definida. En la Figura 6 mostramos el área para cada elemento, misma que se encuentra señalada por el cuadro rojo, nótese como si bien es cierto que hay atención importante en esta área, es similar a otras.

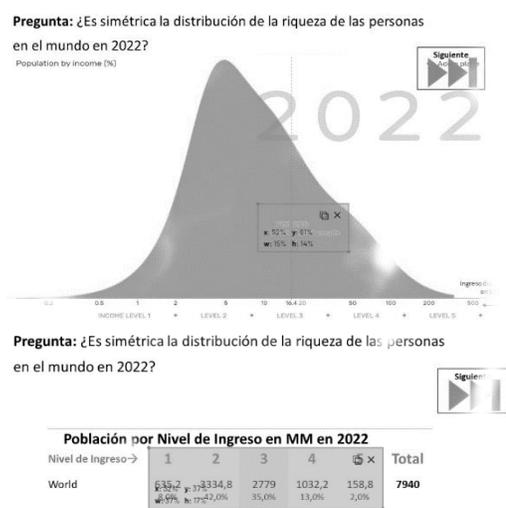


Fig. 6 MAPA DE CALOR: PREGUNTA DE SIMETRÍA AOI

Ahora bien, en donde si es posible determinar diferencias en el comportamiento ocular lo encontramos en los patrones de movimiento detectados. Estos patrones nos revelan que sí se cumple la premisa inicial trazada en torno al área de interés. Como se muestra en la Fig. 7, el patrón de movimiento efectivamente gira en torno al área de respuesta dentro del estímulo, sin embargo, si existen diferencias principalmente en el elemento tabular donde se revela una concentración de las miradas hacia área de interés que no es tan evidente en el elemento gráfico. Estos resultados son consistentes para todos los estímulos, aportamos los relacionados con la pregunta de línea de pobreza.

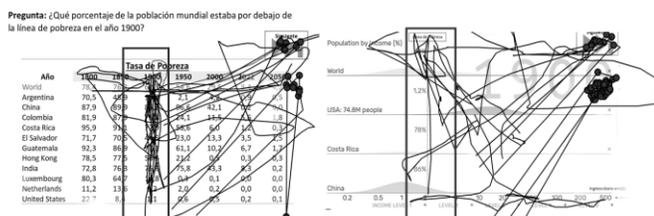


Fig. 7 MAPA DE CALOR: PREGUNTA DE POBREZA AOI

B. Tasa de Respuestas Obtenidas

Otro de los elementos analizados fue la capacidad de responder de los panelistas a preguntas relacionadas con la información analizada, ya sea en forma visual o tabular, con el objetivo de validar la capacidad de accionar una vez realizado el ejercicio.

En la Tabla III se muestran las tasas de aciertos y error obtenidas en cada pregunta realizada, en donde se evidencia que las respuestas tanto correctas e incorrectas pueden ser asociadas con ambos tipos de estímulo, pues no se evidencian patrones en ningún caso.

TABLA III
TASA DE ERRORES Y ACIERTOS POR TIPO DE ESTÍMULO

Pregunta	¿Qué porcentaje de la población mundial estaba por debajo de la línea de pobreza en el año 1900?		¿Es simétrica la distribución de la riqueza de las personas en el mundo en 2022?		¿Cuál es el rango de ingreso por día en US\$ en el nivel 3 de ingresos?		¿Cuál país tiene más personas del total de su población nacional dentro del nivel 4 de ingresos en 2022?	
	Gráfico	Tabular	Gráfico	Tabular	Gráfico	Tabular	Gráfico	Tabular
Respuesta correcta	62,30%		No		De 8 a 32 US\$ por día		USA	
Tipo Estímulo Mostrado	Gráfico	Tabular	Gráfico	Tabular	Gráfico	Tabular	Gráfico	Tabular
Frecuencia obtenida respuesta								
Correctas	11	12	16	9	7	17	13	13
Incorrectas	2	2	1	3	2	1	3	1
No lo sabía	1	2	0	1	3	0	0	0
Total	14	16	17	13	12	18	16	14
Tasa Aciertos	78,6%	75,0%	94,1%	69,2%	58,3%	94,4%	81,3%	92,9%
Tasa Errores	21,4%	25,0%	5,9%	30,8%	41,7%	5,6%	18,8%	7,1%

En la Tabla IV se muestran las tasas de aciertos y error obtenidas para todo el estudio, en donde se concluye mediante una prueba de significancia que, dado que el valor p es 0.747, que es mayor que 0.1, no se puede rechazar la hipótesis nula, es decir una vez mas no hay evidencia estadística suficiente para afirmar que existe una asociación significativa entre los estímulos presentados y las respuestas obtenidas sean aciertos

o no. Los resultados sugieren que cualquier diferencia observada en las tasas de respuestas no se pueden atribuir a los estímulos.

TABLA IV
TASA DE ERRORES Y ACIERTOS POR TIPO DE ESTÍMULO TOTAL

	Estímulos Gráficos	Estímulos Tabular
Respuesta correcta	47	51
Incorrecta o desconocida	12	10
Total	59	61
Tasa Aciertos	80%	84%
Tasa Errores	20%	16%

Estadística de prueba chi-cuadrado: 0.104, Valor p: 0.747, Grados de libertad: 1

En las Fig. 8 y 9 se muestran los resultados obtenidos para cada una de las variables de interés, las hemos dividido por temas de conveniencia por sus unidades de medición ya sean las de tiempo (milisegundos) y las de conteo de ocasiones (veces). Es posible observar que no existen grandes divergencias en cuanto la media de cada tipo de estímulo para todas las variables, además de la gran cantidad de valores que se alejan de las medias (*outliers*) esto nos indica la gran variabilidad observada en los resultados de algunas variables, como *aoi_fixation:tfff_ms* por citar una.

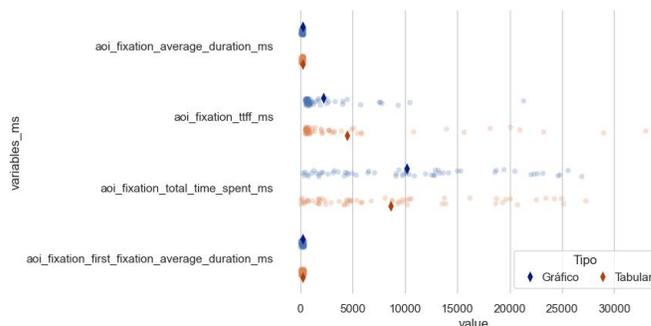


Fig. 8 VARIABLES MEDIDAS EN MILISEGUNDOS POR TIPO DE ESTÍMULO

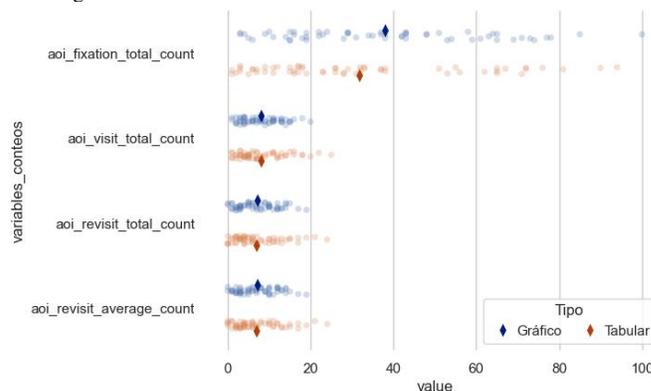


Fig. 9 VARIABLES OBSERVABLES POR TIPO DE ESTÍMULO

En la Tabla V adjuntamos los estadísticos de todas las variables del estudio, en donde es posible constatar la dispersión encontrada durante el experimento, así como el

amplio espectro de posibilidades de análisis de los datos que se desprende de estos resultados.

TABLA V
ESTADÍSTICOS DE LAS VARIABLES ESTUDIADAS

	fixation_tot al_count	fixation_av erage_dura tion_ms	fixation_ttf f_ms	fixation_tot al_time_sp ent_ms	fixation_fir st_fixation average_du ration_ms	visit_total_ count	revisit_tota l_count	revisit_aver age_count
n	119	119	119	119	119	119	119	119
media	34,9	237,7	3362,9	9403,3	240,2	8,0	7,0	7,0
std	24,6	83,0	5892,3	7802,1	116,2	5,1	5,1	5,1
min	1	84	501	84	84	1	0	0
25%	16	146	637	2423	129	4	3	3
50%	29	267	898	8754	215	7	6	6
75%	54	309	2783	13'718	380	11	10	10
max	100	360	33'008	27'296	400	25	24	24

C. Relación entre Variables

Profundizando un poco más en la relación de las variables observadas, podemos determinar cómo existe una distribución muy similar en los resultados observados para ambos estímulos y como ninguna de ellas parece ajustarse a una distribución normal, por lo que se hace necesario un tipo de análisis no paramétrico, para determinar la relación entre variables de una manera más adecuada.

En la Fig. 10 hemos realizado un ploteo de las distribuciones de las variables medibles en Milisegundos para analizar sus distribuciones y su varianza de manera gráfica. Hemos mascarado las variables para facilitar el grafico, remitirse a la metodología si lo requiere. En este enfoque de múltiple para visualizar la distribución de todas las variables junto con todas sus relaciones por pares. Importante resaltar que solo existen diferencias en donde se registraron “outliers”

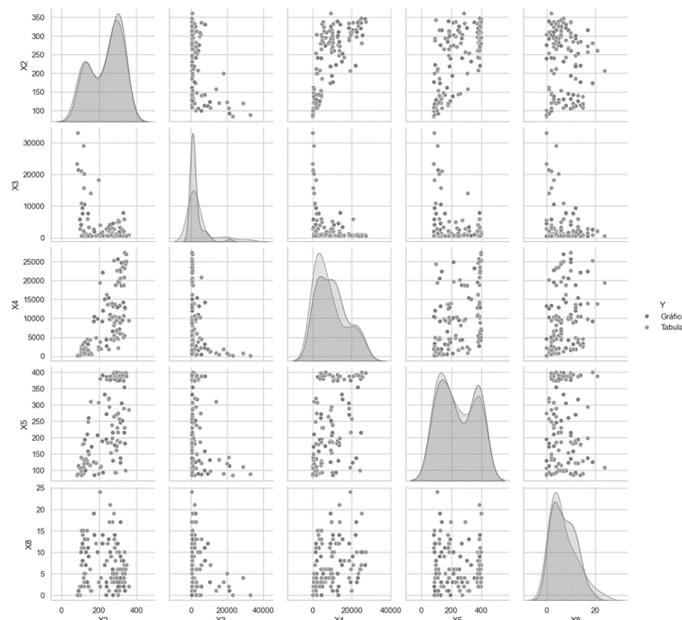


Fig. 10 DISTRIBUCIÓN DE LAS VARIABLES DE TIEMPO POR TIPO DE ESTÍMULO APLICADO

En la Fig. 11, de manera muy similar y siguiendo el mismo análisis anterior hemos realizado un ploteo de las distribuciones de las variables medibles conteos para analizar sus distribuciones y su varianza de manera gráfica. En donde puede notarse como las variables de conteos de visitas tienen una distribución casi idéntica entre ellas lo que denota problemas de colinealidad.

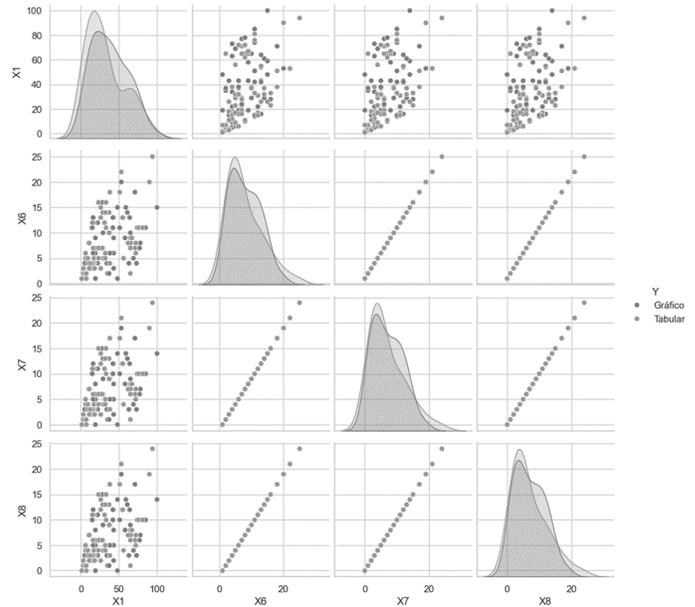


Fig. 11 DISTRIBUCIÓN DE LAS VARIABLES DE CONTEO POR TIPO DE ESTÍMULO APLICADO

D. Regresión Logística

Como último paso de análisis hemos realizado un modelo de ajuste que permita determinar con mayor precisión el grado de predicción que tienen los estímulos presentados durante el estudio sobre las variables medidas. Los resultados obtenidos con este modelo en general no tuvieron un buen ajuste a los datos, dado el pseudo R-cuadrado bajo y el p-valor del test de razón de verosimilitudes que no indica mejoras significativas sobre el modelo nulo. La mayoría de los coeficientes no son significativos, excepto X3 que muestra un pequeño efecto. Esto puede indicar que otros predictores no capturados en el modelo podrían ser necesarios para explicar la variable dependiente, por lo que no es posible rechazar la hipótesis nula. En la Tabla VI proporcionamos la salida obtenida con este análisis.

- [8] B. Shneiderman, «The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations», en *The Craft of Information Visualization*, B. B. Bederson y B. Shneiderman, Eds., en Interactive Technologies. , San Francisco: Morgan Kaufmann, 2003, pp. 364-371. doi: 10.1016/B978-155860915-0/50046-9.
- [9] D. A. Keim, F. Mansmann, J. Schneidewind, y H. Ziegler, «Challenges in Visual Data Analysis», en *Tenth International Conference on Information Visualisation (IV'06)*, jul. 2006, pp. 9-16. doi: 10.1109/IV.2006.31.
- [10] D. Keim, J. Kohlhammer, G. Ellis, y F. Mansmann, *Mastering the Information Age Solving Problems with Visual Analytics*. Eurographics Association, 2010. Accedido: 31 de julio de 2024. [En línea]. Disponible en: <https://diglib.org/handle/10.2312/14803>
- [11] J. Rojas-Segura, M. Faith-Vargas, y J. Martínez-Villavicencio, «Conceptualizing digital transformation using semantic decomposition», *TEC Empres.*, vol. 17, n.o 3, pp. 63-75, dic. 2023, doi: 10.18845/te.v17i3.6850
- [12] G.-D. Sun, Y.-C. Wu, R.-H. Liang, y S.-X. Liu, «A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges», *J. Comput. Sci. Technol.*, vol. 28, n.º 5, pp. 852-867, sep. 2013, doi: 10.1007/s11390-013-1383-8.
- [13] D. Best, J. Bruce, S. Dowson, O. Love, y L. McGrath, «Web-Based Visual Analytics for Social Media», *Proc. Int. AAAI Conf. Web Soc. Media*, vol. 6, n.º 4, Art. n.º 4, 2012, doi: 10.1609/icwsm.v6i4.14363.
- [14] F. Beck, M. Burch, S. Diehl, y D. Weiskopf, «The State of the Art in Visualizing Dynamic Graphs.», *EuroVis STARS*, 2014, Accedido: 31 de julio de 2024. [En línea]. Disponible en: https://www.visus.uni-stuttgart.de/documentcenter/forschung/visualisierung_und_visual_analytics/eurovis14-star.pdf
- [15] R. Reddy, M. Höferlin, M. Dambier, y D. Weiskopf, «Visual Analytics for Dynamic Evacuation Planning.», en *EuroVA@ EuroVis*, 2012. Accedido: 31 de julio de 2024. [En línea]. Disponible en: <http://diglib.cgv.tugraz.at/bitstream/handle/10.2312/PE.EuroVAST.EuroVA12.013-017/013-017.pdf?sequence=1&isAllowed=y>
- [16] O. Adagha, R. M. Levy, y S. Carpendale, «Towards a product design assessment of visual analytics in decision support applications: a systematic review», *J. Intell. Manuf.*, vol. 28, n.º 7, pp. 1623-1633, oct. 2017, doi: 10.1007/s10845-015-1118-5.
- [17] E. Dimara, H. Zhang, M. Tory, y S. Franconeri, «The Unmet Data Visualization Needs of Decision Makers Within Organizations», *IEEE Trans. Vis. Comput. Graph.*, vol. 28, n.º 12, pp. 4101-4112, dic. 2022, doi: 10.1109/TVCG.2021.3074023.
- [18] «Gapminder». Accedido: 29 de mayo de 2024. [En línea]. Disponible en: <https://www.gapminder.org/>
- [19] J. Sweller, «Cognitive Load Theory and E-Learning», en *Artificial Intelligence in Education*, G. Biswas, S. Bull, J. Kay, y A. Mitrovic, Eds., Berlin, Heidelberg: Springer, 2011, pp. 5-6. doi: 10.1007/978-3-642-21869-9_3
- [20] P. W. M. Van Gerven, F. Paas, J. J. G. Van Merriënboer, y H. G. Schmidt, «Memory load and the cognitive pupillary response in aging», *Psychophysiology*, vol. 41, n.o 2, pp. 167-174, 2004, doi: 10.1111/j.1469-8986.2003.00148.x