



Capability Building

Data DRIVEN

Tratar la data con precaución

MÓDULO 6

Módulo 6:

Tratar la data con precaución

Si nos remontamos al inicio de nuestro proceso de aprendizaje, en el Módulo Uno el primer desafío al que nos enfrentamos fue la inmensa cantidad de datos que existen. Ahora que ya tenemos algunas pautas sobre cómo obtenerlos, el mayor desafío será analizarlos correctamente antes de un proceso de toma de decisiones.

Cuando analizamos datos, se nos presentan una serie de posibles errores sistemáticos que toda persona podría cometer (y que podrían afectar gravemente las decisiones que tome con las interpretaciones esos datos). ¿A qué nos referimos con errores sistemáticos? Por ejemplo, existe una tendencia natural a prestarle mayor atención a aquellas informaciones que confirman nuestras creencias e hipótesis, e ignorar aquellas que la refutan. Esta tendencia casi automática se conoce como 'sesgo de confirmación', y claramente atenta contra nuestra tan deseada objetividad al momento de analizar la data.

Sin embargo, **este no es el único sesgo que debemos prevenir**. Por desgracia, en nuestra mente existen cientos de sesgos.

Aunque los datos sean objetivos, las interpretaciones que hagamos de ellos siempre serán (por naturaleza) subjetivas. Por eso es que estos sesgos no sólo pueden conducirnos a errores de análisis, sino que inclusive han conducido a enormes desastres: desde la explosión del transbordador aeroespacial Challenger en 1986 por el sesgo de pensamiento grupal e ilusión de invulnerabilidad, hasta la tragedia de Chernobyl por sesgos retrospectivos, la historia está llena de ejemplos que suplican precaución.



Sesgos en la toma de decisiones

Sesgos cognitivos

Para darle una definición concreta, **un sesgo cognitivo es una interpretación sistemática errónea de la información disponible** que ejerce influencia en la manera en que procesamos los pensamientos, emitimos juicios y tomamos decisiones, y su concepto fue introducido por los psicólogos israelíes Kahneman y Tversky en 1972.

Esto se explica en el hecho de que cada segundo tu cerebro ejecuta millones de procesos mentales, por lo que busca los mecanismos para realizarlo rápidamente y ahorrar la mayor energía posible. Para ello, **se basa en preconcepciones y estructuras mentales existentes** que le permitan acomodar, categorizar y procesar la información de forma sistemática para tomar una decisión rápida ¿o crees que cuando nuestros ancestros veían un depredador podían detenerse a analizarlo? Por esta razón incorporamos sesgos en nuestra evolución: para actuar rápido, ahorrar recursos y evitar riesgos.

Para introducirnos en la gran cantidad de sesgos que configuran nuestras decisiones, vamos a profundizar en los cuatro sesgos cognitivos más presentes:

Los **sesgos cognitivos** son aquellos errores sistemáticos de juicio que todos cometemos inconscientemente y que **influyen en nuestras decisiones**. No podemos eliminar estos sesgos, pero podemos agudizar su pensamiento.



El efecto
ancla



El error de
confirmación



El error de
disponibilidad



El error
rápido/lento

	El efecto ancla	El error de confirmación	El error de disponibilidad	El error rápido/lento
Falsa Suposición	Consideramos todos los factores antes de tomar una decisión	Evaluamos objetivamente la situación antes de tomar una decisión	Tenemos buenos argumentos para tomar una decisión en particular	Creemos que intuitivamente hemos tomado la decisión correcta
Realidad	<p>El sesgo de anclaje describe la tendencia de las personas a confiar demasiado en la primera información que reciben sobre un tema. Independientemente de la precisión de esa información, las personas la usan como punto de referencia o ancla para hacer juicios posteriores.</p>	<p>Interpretamos la información de tal manera que afirma nuestra suposición o convicción existente. O viceversa: bloqueamos la información que contradice nuestras opiniones. Nadie busca contraargumentos en Google.</p>	<p>En psicología, el sesgo de disponibilidad es la tendencia humana a confiar en la información que nos viene fácilmente a la mente al evaluar situaciones o tomar decisiones. Debido a este sesgo, la gente cree que la información fácilmente disponible es más representativa de lo que es en realidad.</p>	<p>Existen dos maneras en las que el cerebro procesa información y piensa: Sistema 1 (rápido e intuitivo) y Sistema 2 (lento y prejuicioso). El sistema 1 predomina, ya que implica un abordaje más simple y de menor esfuerzo mental. Esta sobre simplificación de análisis nos hace equivocarnos al tomar una decisión.</p>
Solución	No confies en tus primeras impresiones	Asume que estás equivocado	No confies en la evidencia anecdótica	Lea los correos electrónicos de nuevo antes de enviarlos

Pero los sesgos cognitivos no son los únicos elementos que nos pueden llevar a decisiones equivocadas o a interpretar en forma errónea la información: existen otras **fuentes de error al momento de tomar decisiones** que tienen que ver con **el manejo de la data y el diseño de nuestros estudios y experimentos**.

Eso significa que existen dos tipos de errores a los que podemos enfrentarnos al momento de diseñar nuestros estudios y análisis:

- Error aleatorio: nace del hecho de que se trabaja con muestras de individuos y no con toda la población. Procede de la variabilidad inherente al muestreo y no puede eliminarse, pero sí controlarse mediante el aumento del tamaño muestral. Este tipo de error está muy relacionado con el concepto de precisión o fiabilidad.
- Error sistemático o sesgo: aparece cuando se introduce un error en el diseño del trabajo, ya sea en la selección de los individuos, en la información recogida, en el diseño de las preguntas de una encuesta o en su análisis. A diferencia de lo que ocurre con el error aleatorio no se atenúa al aumentar el tamaño de la muestra, y una vez introducido, es prácticamente imposible de remediar en la fase de análisis. Este tipo de error va muy ligado al concepto de validez., ya que los sesgos cognitivos pueden influir en la manera en que diseñamos nuestros estudios y generar sesgos sistemáticos a lo largo de todo el trabajo.

Mintiendo con la estadística

Otra precaución importante que debemos tener en el manejo de data ocurre al momento de realizar representaciones gráficas. A continuación, vamos a resumir cuatro errores que (malintencionados) pueden ser utilizados también para engañar y manipular informaciones y conclusiones.

Error #1: Romper las escalas muestra un drama donde no lo existe



En el gráfico de la izquierda da la impresión que el Producto Bruto Interno de EEUU está cayendo de manera impresionante. Sin embargo, es una tergiversación a partir de la manipulación gráfica de no mostrar la escala completa. Cuando en el gráfico de la derecha exhibimos la escala completa, vemos que en realidad el drama es mucho menor de lo quieren que creamos.

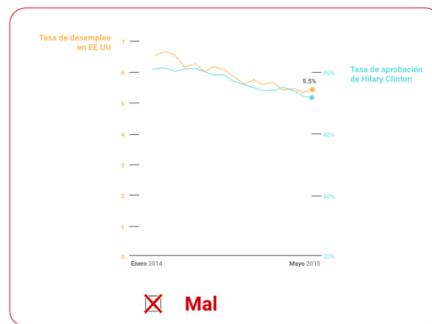
Error #2: Cambiar de eje



En los siguientes gráficos, a pesar de mostrar exactamente la misma información, el de la izquierda induce a pensar que el porcentaje de partidos ganados sobre total de partidos jugados (porcentaje de victorias) es significativamente más alto en el caso de Djokovic en comparación con Federer. Ocurre que en el primer gráfico, el porcentaje de un jugador está sobre el eje izquierdo, y el del otro jugador está sobre el eje derecho, y en distintas escalas.

Es válido y en muchos casos aconsejable utilizar el segundo para sumar claridad en situaciones específicas. Lo que definitivamente **no es válido hacerlo cuando se trata de la misma información** y resulta en un cambio de escalas.

Error #3: Mostrar correlación para implicar causalidad



Otro error común es mostrar una correlación e implicar una causalidad. Por ejemplo, la “correlación” de que la caída de los índices de aprobación de Hillary Clinton producen una caída en la tasa de empleo.

En realidad se trata de dos variables completamente independientes, sin relación alguna, que casualmente han descendido de manera similar en el último año.

Error #4: No contextualizar



Por último, otro error común es **no contextualizar la data**. En el gráfico de la izquierda, vemos que Chicago y New York son más peligrosas que Detroit ya que tienen entre 10% y 30% más asesinatos que Detroit.

Si bien el dato es correcto, la comparación no es válida porque no considera el tamaño de habitantes en cada ciudad.

En el marco o contexto adecuado, la información puede ser muy distinta: por eso hay que preguntarse **cuál es la variable que realmente importa**. Cuando vemos la cantidad de asesinatos por cada 100.000 habitantes, vemos que Detroit es la más peligrosa, y que ni Chicago ni New York están entre las primeras 5.



Capability Building

✉ hello@capabilitybuild.com

🌐 www.capabilitybuild.com

📍 175 SW 7th Street,
Suite 1501 Miami, FL 33130

SEGUINOS EN
NUESTRAS REDES!

