

Mentir con exactitud: las mañas con datos de los políticos en Twitter

Por: Pablo Manrique. 24/01/2025

La guerra de los datos se ha convertido en parte del paisaje de las discusiones en Twitter, (ahora X). No es raro ver dos políticos mostrando con los mismos datos conclusiones diferentes. A pesar de que la estadística es un área de las matemáticas, los datos se pueden “torturar” o “novelar”, retorcerse hasta que digan lo que cada uno quiere escuchar, como dice, el exdirector del DNP e investigador económico, Jorge Iván González.

A continuación las 5 principales estrategias “estadísticas” que usan los políticos para amañar los datos a su gusto.

1. *Cherry picking*: escoger solo los datos que me sirven

Una de las técnicas más comunes es escoger datos atípicos (*outliers*) para argumentar una posición, ignorando la tendencia general. Esto se llama en la academia “*cherry picking*”, que traduce escoger la cereza.

Uno de los casos más recientes fue en el marco de la discusión por la Unidad de Pago por Capitación (UPC). Para calcular la UPC con la inflación, el Ministerio de Salud dijo que los datos de las EPS no eran fiables. Ellos pedían un aumento del 17%, y el Ministerio solo subió la UPC 5,4%. Mostró, por ejemplo, que en la base de datos se encontraban casos de gente fallecida con gastos reportados. Estos datos los uso para cuestionar la base con la que se calcula la UPC. “Aparecieron reportes de 23 410 personas fallecidas”, pero no propuso eliminar estas personas en lugar de toda la base de datos.

Frente a esto, Alejandro Gaviria, exministro de Salud y crítico del gobierno, explica que estas bases se llenan a mano, y por eso es usual que tengan errores. La práctica normal es sustraer las anomalías del análisis. Pero, en este caso, lo que hizo el Ministerio fue solo tenerlas en cuenta.

Aunque estas anomalías son reales, y pueden ser investigadas, no representan el comportamiento global de los sistemas. Sobre todo, usarlas para construir una narrativa es, como dice Gaviria, una forma de “mentir con exactitud”. Se enfoca en un error específico que respalda la narrativa deseada, mientras se ignora todo lo demás. Cuando esto sucede, la solución es tener acceso a todas las bases de datos para confirmar el grado de las anomalías, cosa que en este caso no ha hecho el gobierno, aunque informo que si lo hará.

2. Números absolutos versus relativos

Otra estrategia de manipulación consiste en presentar cifras absolutas o porcentuales, según lo que más convenga a la narrativa. Es decir, dar el número total versus el porcentaje

Por ejemplo, en la discusión sobre la deuda pública, el economista y exministro Jose Manuel Restrepo [publicó en Twitter](#) que la deuda en términos absolutos aumentó. “Bomba de tiempo la que deja el gobierno actual... en solo dos años, ¡no en cuatro!!, está aumentando la deuda en 38,600 millones de dólares!! ... o sea en solo dos años, tres veces la de todo el gobierno anterior!”. Es cierto, pasó de 759 a 932 billones, como lo mostró en el tweet, pero decir que la deuda se aumentó en tres veces más, no tiene en cuenta el crecimiento de la economía. Si se ve el crecimiento como porcentaje del PIB, en ese periodo fue solo de 1%.

Sin embargo, también es cierto que porcentualmente disminuyó con respecto al PIB si se mira anual y no trimestral, que es lo importante a la hora de discutir su sostenibilidad. Esa es la medida que usa el Comité Autónomo de la Regla Fiscal (Carf). Este tipo de interpretación depende de cuál dato se quiera enfatizar: mientras que los números absolutos suelen aumentar por la inflación, los datos porcentuales muestran una perspectiva más ajustada al contexto económico, como el crecimiento del PIB o la capacidad del país para pagar su deuda.

Otra versión de este truco también la usó el presidente Petro en 2024, en lugar de dar absolutos, dio cambios relativos pero sin tener en cuenta la inflación. Para decir que su subida del salario mínimo era mayor a la de Duque, afirmó que el aumento de Duque fue de 10% y el suyo del 12%. Pero si se analiza el aumento porcentual restando la inflación, es decir, teniendo en cuenta que los precios absolutos suben, resulta que el salario con Petro subió un 2,7% comparados con 4,4% de Duque. Como [mostró La Silla](#)

3. Eliminar la posibilidad de comparación

Otra estrategia usada por los gobiernos es alterar las categorías o metodologías de los datos publicados para dificultar las comparaciones con períodos anteriores. Esto no solo le quita contexto a la opinión pública, también complica el trabajo de los analistas que intentan entender si las políticas públicas han sido efectivas en el tiempo.

Por ejemplo, en este gobierno la Dian dejó de publicar cifras de recaudo en las mismas categorías históricas, lo que impide saber si realmente subió o no el recaudo del impuesto a la renta. Además, dejó de publicar las tablas comparativas que tenía de recaudo. En un año donde el bajo recaudo ha sido importante, haciendo más difícil calcular cuanto cae y en qué áreas cae el recaudo. Esto sucedió por primera vez con la llegada de Jario Villabonua a la Dian.

Informe de recaudo DIAN

Diciembre
2024



Recaudo acumulado a diciembre de 2024



\$267.20
billones



Recaudo bruto acumulado a diciembre
discriminado por tipo de impuesto

Valor recaudado y participación porcentual de los tributos con los que más se logró recaudo de enero a diciembre de 2024:

Concepto	Recaudo	Participación
Retención Renta	97.461.497	36,5%
Ventas	57.533.283	21,5%
Aduanas	41.127.160	15,4%
Otros	71.079.059	26,6%
Total	267.201.000	100%

Valores en millones de pesos.

Más que un truco estadístico, este tipo de estrategia es una táctica de comunicación. Al ocultar la información o cambiarla de formato, se dificulta la posibilidad de realizar análisis críticos o de mantener un debate público informado.

Después de publicar el informe, una o dos semanas después la Dian sí publica los metadatos para calcular las variaciones. Pero eso sucede luego de que se publican las noticias sobre el recaudo, que han sido negativas para el gobierno.

4. Escoger la línea base e ignorar la estacionalidad

Otra forma de manipular los datos es comenzar una serie temporal en un punto que favorezca la narrativa deseada. Es decir, comenzar o terminar la gráfica en el punto en el que el dato sostiene la tesis

Por ejemplo, Petro presentó datos sobre las quejas a las EPS intervenidas, mostrando que entre noviembre y diciembre estas disminuyeron. Sin embargo, si se analiza el año completo, las quejas en general estaban aumentando, como mostró [Colombia Factcheck](#).

Estos es engañoso en datos que tienen comportamientos estacionales. Por ejemplo, las quejas de las EPS, son estacionales, es decir, siempre tienen una variación parecida para cada mes. En diciembre siempre bajan las quejas porque la de gente se va de vacaciones y no va al médico o no tiene tiempo de tramitar las quejas.

5. Usar datos de dudosa procedencia

Finalmente, algunos políticos recurren a índices o datos de fuentes poco confiables para respaldar sus argumentos. Después de la publicación de un [artículo](#) en el diario de La República, Alejandro Gaviria retrinó un índice dudoso que afirmaba que el sistema de salud de Colombia había caído en comparación con otros países. El mismo Gaviria aceptó que era un índice dudoso y aceptó que cometió un error “No había mucha transparencia sobre la información”

Usar datos de dudosa procedencia no solo genera confusión, sino que también alimenta narrativas falsas que pueden tener consecuencias graves en la toma de decisiones.

Pablo Manrique Periodista/Desarrollador IA

Economista y Científico de datos, he trabajado en startups y consultoras como

analista de datos y desarrollador de inteligencia artificial. Escribo usando inteligencia artificial y minería de datos, y cubro economía

Fecha de creación

2025/01/24