

Small Sample Size

Beschreibung

Die Nutzung von zu kleinen Stichprobengrößen.

Kleine Stichprobengrößen führen zu schlechter Präzision, da sie die Vielfalt einer größeren Population oft nicht richtig abbilden. Dadurch entstehen fehlende Repräsentativität, verzerrte oder verfälschte Ergebnisse und reduzierte statistische Aussagekraft

<https://smart.dhgat.com/small-sample-size-why-its-bad-for-research-outcomes/>

Darstellung

Angabe zur Stichprobengröße leicht versteckt

Stichprobengröße im Kontexttext verallgemeinert (z.B. „Das deutsche Volk wurde befragt...“ aber auf dem Diagramm wird sichtbar, dass nur 20 Menschen befragt wurden)

Nutzung von zu kleinen Stichprobengrößen

Dynamische Umsetzung

Wert für Stichprobengröße zufällig verkleinern

Schlechtes/kein narratives Design

Beschreibung

Ohne narratives Design fehlen Bezugspunkte, wodurch Zahlen/Werte zufällig oder bedeutungslos wirken

Darstellung

Korrekte Diagramme, aber wichtige Achsenbeschriftungen fehlen

Irreführende, reduzierte, generalisierte Kontextbeschreibung

Titel extrem vage

Dynamische Umsetzung

Zufälliges Weglassen von Werten/Beschriftungen

Kontexttext und Titel absichtlich generisch und oberflächlich halten

Intend-Based Data/ Intentional Framing

Beschreibung

Absichtsbasierte Auswahl und Betonung bestimmter Aspekte einer Information, um die Wahrnehmung und Interpretation zu beeinflussen und eine bestimmte Reaktion oder Schlussfolgerung zu fördern

<https://designbuddy.substack.com/p/intent-frame>

Darstellung

Diagramm/Statistik bleibt normal

Kontextbeschreibung ist wertend/emotional (positiv, negativ oder dramatisch)

Dynamische Umsetzung

Zufälliger Einsatz von bewertenden Wörtern/Phrasen in Kontextbeschreibung oder Titel

Nichtssagende Statistiken

Beschreibung

Eine Statistik, die keine richtige Aussage hat oder Kategorien kombiniert, die nichts miteinander zutun haben. Erzeugt künstliche Zusammenhänge, die inhaltlich nicht begründbar sind

<https://pmc.ncbi.nlm.nih.gov/articles/PMC9645302/>

Darstellung

Diagramm/Statistik mit zufälligen oder irreführend benannten Kategorien (z.B. Balkendiagramm zur durchschnittlichen Schlafdauer: balken 1: Testgruppe A, balken 2: Männer Ü40, Balken 3: Hundebesitzer)

Pauschalisierte Kontextbeschreibung, die keinen relevanten Bezug zur Statistik hat

Dynamische Umsetzung

Zufällige Auswahl der Kategorien sowie Einsetzung von unsinnigen Kausalitäten in Kontextbeschreibung

Auswahl günstiger Zeitfenster

Beschreibung

Je nach Auswahl der Zeitfenster können Trends vollständig anders aussehen, es wird also auf einen Zeitraum gezoomt der für die Argumentation besonders stützend agiert, auch wenn das das Gesamtbild etwas völlig anderes vermitteln würde

<https://www.clearerthinking.org/post/are-you-being-misled-by-statistics>

Darstellung

Kontextbeschreibung über etwas, was für gewöhnlich über einen längeren Zeitraum gemessen wird. Diagramm ist dann jedoch nur auf einen sehr kleinen Zeitraum gezoomt

Dynamische Umsetzung

Zufälliger Zoom auf kleinen Zeitraum mit großem Anstieg/Fall (am besten gegenteilig zur gesamtstatistik, also wenn chart z.B. im gesamtzeitraum einen Anstieg erlebt, wird auf ein Zeitfenster in dem der chart stark fällt gezoomt)

Einheitswechsel

Beschreibung

Durch den Wechsel von relativen und absoluten Einheiten werden Trends und Unterschiede entweder dramatisiert oder abgeschwächt. Informationsgehalt wird nicht verändert, subjektive Wahrnehmung wird jedoch massiv beeinflusst.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7419489/>

Darstellung

Wechsel zwischen Prozentwerten und Absoluten zahlen in Diagramm und kontexttext

Dynamische Umsetzung

???

Korrealtion != Kausalität

Beschreibung

Zwei Variablen können gemeinsam steigen oder fallen, ohne dass die eine die andere verursacht.
Durch manipulierten Kontexttext kann jedoch eine scheinbare Ursache-Wirkung-Beziehung suggeriert werden. („Wenn X steigt, steigt auch Y, also verursacht X Y“)
<https://journals.sagepub.com/doi/full/10.1177/2515245917745629>

Darstellung

Diagramm zeigt zwei Kurven mit sehr ähnlichem Verlauf
Kontexttext macht daraus eine direkte Ursache
Mögliche Drittvariablen werden ignoriert

Dynamische Umsetzung

System generiert zwei zufällig positiv oder negativ korrelierende Datensätze,
Kontexttext verstärkt die Schein-Ursache („Wenn X steigt, zieht Y direkt nach.“)

Verwechslung von Median und Mittelwert:

Beschreibung:

Viele Menschen verwechseln den Median und den Durchschnitt. Der Median ist oft gemeint, da er aus allen Werten nach Sortierung den Wert am nächsten zur Mitte beschreibt, der Durchschnitt hingegen beschreibt den Wert aller Werte geteilt durch die Anzahl welcher sich signifikant ändern kann wenn Werte weit von der Mitte weg verändert werden, ansteigen oder fallen.

Darstellung:

Es kann ein zusätzlicher Wert für den Median und für den Mittelwert gezeigt werden. Beide Werte können über eine extra Linie visualisiert werden und sich gegebenenfalls auf den anderen Wert innerhalb der Statistik verschieben.

Dynamische Umsetzung:

Jede Statistik besitzt einen Median und einen Mittelwert, algorithmisch kann dann gefiltert werden in welchen Datensätzen diese Werte wie stark auseinander gehen und wo Vermittlungspotenzial besteht.

"Cherry Picking" der Inhalte

Beschreibung:

Inhalte können nach Bedarf zugeschnitten oder selektiv gefiltert werden (z.B. nach Zeitperiode, Periode, Verhaltenskurven), ohne den kompletten Datensatz oder alle Werte zu zeigen oder zu benennen. Dies fördert unvollständige Meinungsbildung und ermöglicht es leicht, Aussagen fälschlicherweise zu belegen.

Darstellung:

Unvollständige Datensätze können anschließend beim auflösen der Aufgabe um fehlende Werte ergänzt werden

Dynamische Umsetzung:

Leicht auf im Prinzip jede Statistik umsetzbar durch vorheriges einschränken der Daten mit hilfe algorithmischer Logik zum stellen der Aufgabe und anschließendem zeigen der fehlenden Inhalte

Ableitungsstatistiken:

Beschreibung:

Ableitungsstatistiken beschreiben die Steigung und nicht die absoluten Werte eines Datensatzes, Ableitungen beschreiben somit das allgemeine Verhalten der Daten und nicht den derzeitigen Stand.

Darstellung:

Ableitungsstatistiken sind genau so darstellbar wie normale Statistiken, wahlweise könnten auch Plot-Diagramme genutzt werden gegebenenfalls mit Verbindungslien

Dynamische Umsetzung:

Es können separate Ableitungsstatistiken aus absoluten Statistiken gebildet werden und beide können verglichen oder gegenübergestellt werden. Dieser Prozess funktioniert nur in eine Richtung es müssen daher bereits absolute Werte vorliegen um eine solche Aufgabe stellen zu können

Absolute/Relative Increases/Decreases von X:**Beschreibung:****Darstellung:**

Dynamische Umsetzung:

Fehlende Streuung/Varianz (Nur Mittelwerte):**Beschreibung:**

Die Streuung einer Statistik bzw. die Varianz einzelner Werte kann durch zusätzliche Boxen innerhalb von Datensätzen visualisiert werden, welche Tendenzen oder die Aufteilung der Werte beschreiben. Die Verteilung von Werten kann genau so zur Meinungsbildung und zur Vergewisserung von Aussagen beitragen, obwohl diese sehr oft weggelassen werden.

Darstellung:

Es können extra Balken oder "Plots" hinzugefügt werden, um fehlende Streuung zu visualisieren. Außerdem können extra Linien innerhalb ergänzt werden, die Abgrenzungen oder Perzentile zeigen und mit diesen beschriftet sind.

Dynamische Umsetzung:

Jede Statistik besitzt Streuung und eine spezifische Verteilung der Werte. Diese kann immer durch genannte Darstellungsoptionen ergänzt werden. Algorithmisch kann auch hier wieder erfasst werden, in welche Statistik die Streuung wie stark ausgeprägt ist und einen relevanten Teil der Daten ausmacht.