



In unserem ersten Teil ([Datenbewusstsein Teil 1: Daten können dich auch verarschen!](#)) habe wir aufgezeigt, wie schnell man in eine Analyse-Falle auf Grund von falschen Daten tappen kann. Jetzt zeigen wir auf, wie man selbst ein Bewusstsein dafür entwickeln kann.

## Teil 1: Die Datenquellen

Es gibt jede Menge Daten, die von und über verschiedene Quellen gesammelt werden. Man unterscheidet in erster Linie Stammdaten, am Beispiel Mobilfunk sind das z.B. der Smartphone Typ, das Betriebssystem, Geschlecht oder Alter des Nutzers und hunderte weiterer Profilinformatoren. Dann gibt es bewegliche Daten wie z.B. Nutzungsinformationen der Apps oder das Konsumverhalten.

Mittlerweile werden in einem normalen Haushalt stündlich Millionen von Daten erzeugt: Smartphones, Smart TV, Smart Home, Smart Metering und mittlerweile auch

Haushaltsgeräte wie der Thermomix senden in verschiedene Cloudsysteme Datensätze en masse. Unser Business generiert noch ein Vielfaches davon und kombiniert diese Daten auf clevere Weise. Datenbewusstsein bedeutet, diese Quellen zu kennen und auf Ihre Validität hin zu verstehen.

## Teil 2: Daten brauchen Kontexte

Daten allein bieten keine Transparenz, genauso wenig bringt das reine Sammeln von Daten keinen Mehrwert. Daten müssen in einen Kontext gesetzt werden, um in Wissen umgewandelt zu werden. Beispiel: Nehmen wir uns die Abverkaufs- bzw. Umsatzdaten vor, die wir so gern mit den Vorjahresdaten vergleichen. Das kann man so einfach machen, ist aber gerade im Jahr 2022 ohne Kontext-Betrachtung wertlos. Wie war der geänderte Umsatzsteuersatz durch Corona? Wann war ein Teil-Lockdown, konnten die Menschen reisen und gab es Seiteneffekte durch den Cocooning Effekt? Kontexte aufzuzeigen die Einfluss auf die Daten haben, ist ein wichtiger Faktor der Berücksichtigt werden muss.

Hier mal ein paar Beispiele für Kontexte:

Korrelierende Daten

Ortsbezogene Daten

Daten mit Zeitbezug

Metadaten (Daten über Daten)

Referenzdaten als Basis und Vergleichsbetrachtung

Strukturdaten (Alter, Geschlecht, Fahrzeugtyp etc.)

Transaktionsdaten (Bewegungsdaten)

Auditdaten

Das Fachwissen des Auswertenden (elementar wichtig)

## Teil 3: Daten brauchen Qualität

Über Datenqualität wird seit Jahrzehnten geredet und speziell im Zeitalter der künstlichen Intelligenz ist sie besonders wichtig geworden. Getreu dem Motto „Shit in – Shit out“ kann mangelnde Datenqualität zu wertlosen Ergebnissen führen.

Folgende Fragen sollte man sich bei der Beurteilung von Ergebnissen bezüglich der verwendeten Datenbasis stellen:

- Wie vollständig sind die Daten?
- Sind die Daten eindeutig oder interpretierbar?
- Sind die Daten korrekt?
- Sind die Daten aktuell?
- Wie hoch ist der Genauigkeitsgrad?
- Sind die Daten konsistent?
- Besteht Redundanzfreiheit?
- Sind die Daten für die Auswertung relevant?

## Teil 4: Was noch viel wichtiger ist als die Daten

Ein ganz besonderer Grundsatz bei der Erstellung einer Analyse oder der Generierung von Informationen ist das Bewusstsein für den folgenden Sachverhalt:

**„Nur wer die richtigen Fragen stellt, kriegt auch eine richtige Antwort“**

Zuerst muss geklärt sein, was man denn überhaupt erreichen will. Beispiel: Wie viele Personen im Angestelltenverhältnis mit mittlerem Einkommen sind Mittwochs in der Stadt? Daraus ergibt sich der Datenbedarf, dessen Qualität nach den o.g. Kriterien festgelegt werden muss.

Dann stellt man fest, welche anderen Kontexte die Datengrundlage beeinflussen? Kann ein Festival, ein Bahnhof, ein Feiertag im Nachbarland ggf. diese Fragestellung beeinflussen?

Wie genau muss das Ergebnis sein? Reicht eine Tendenz, oder muss das Ergebnis exakt stimmen? Welche Daten machen zur Beantwortung der Frage überhaupt keinen Sinn? Im Extremfall können sie ein Ergebnis massiv verfälschen.

## Ein Fazit

Mit den Rezepten kann man eine Schnellanalyse durchführen und Fragen stellen, die

eventuell für den einen oder anderen unbequem sein könnten. Aber bitte nicht falsch verstehen: Es geht hier nicht darum, Projekte oder Lösungen damit zu verhindern, sondern diese vielmehr in der Qualität zu verbessern.

Beitragsbild von [Photo Mix](#) auf [Pixabay](#)