

# Zusammenfassung Datenbankdesign (PDF)

## Der Designprozess

Durch ein gutes Datenbankdesign

- Werden Redundanzen in den Daten vermieden
- Wird die Korrektheit und Vollständigkeit der Daten sichergestellt

Der Designprozess einer Datenbank ist ein iterativer Prozess. Folgende Schritte führen dich zum Ziel:

- Zweck der Datenbank bestimmen
- Die relevanten Informationen festlegen
- Die Informationen in Tabellen und Spalten organisieren
- Primärschlüssel und Fremdschlüssel bestimmen
- Das Design verfeinern und Normalisierungsregeln anwenden

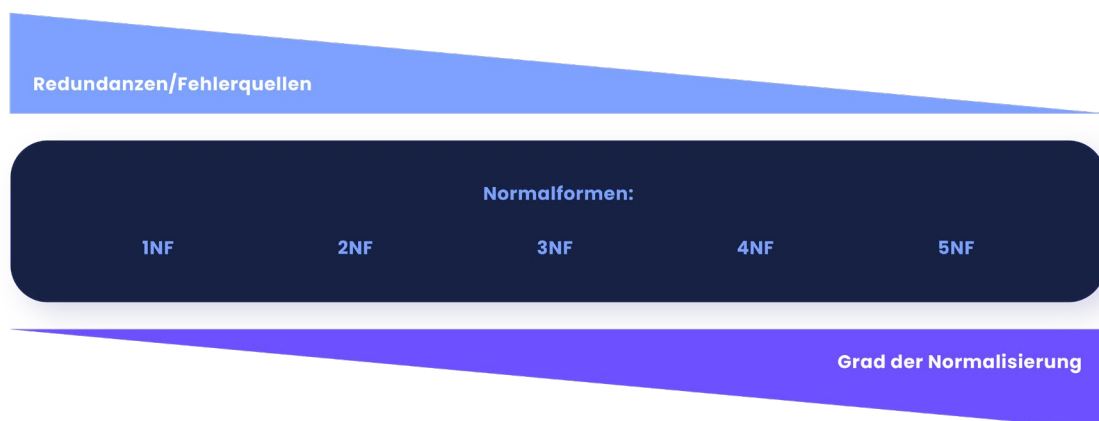
## Normalisierung

Normalisierung ist ein stufenweiser Prozess, bei dem durch Ändern und Hinzufügen von Tabellen sichergestellt wird, dass es keine Redundanzen in den Daten gibt. So werden Änderungs-, Löscher- und Einfügeanomalien vermieden. Die Stufen der Normalisierung sind die Normalformen. Es gibt viele Normalformen, doch in der Praxis relevant sind die Normalformen 1NF, 2NF und 3NF. Oft wird eine Datenbank in der 3NF als normalisiert bezeichnet.

- Erste Normalform (1NF): alle Attribute sind nicht weiter teilbar.
- Zweite Normalform (2NF): Attribute hängen immer nur vom ganzen Primärschlüssel ab.
- Dritte Normalform (3NF): Attribute hängen ausschließlich vom Primärschlüssel ab.

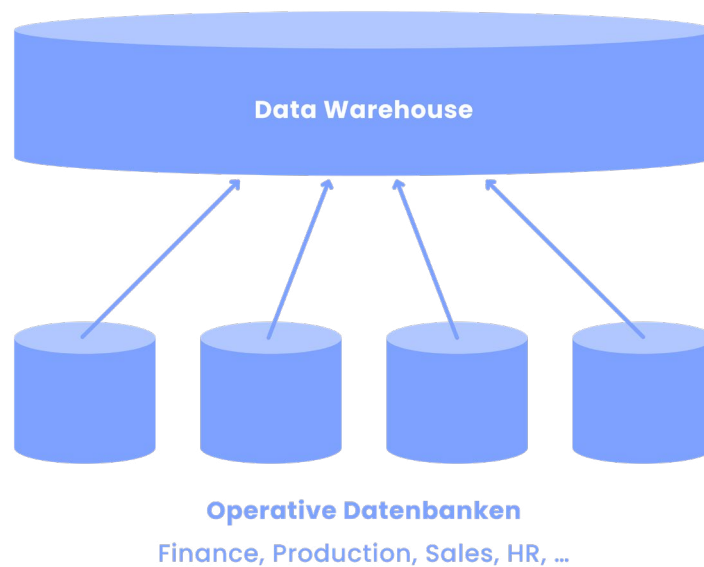
Merke dir den folgenden Satz für eine Datenbank in der 3NF:

**Jedes Attribut der Tabelle hängt vom Primärschlüssel ab, vom ganzen Primärschlüssel, und von nichts anderem als dem Primärschlüssel.**



## **Operative versus entscheidungsorientierte Daten**

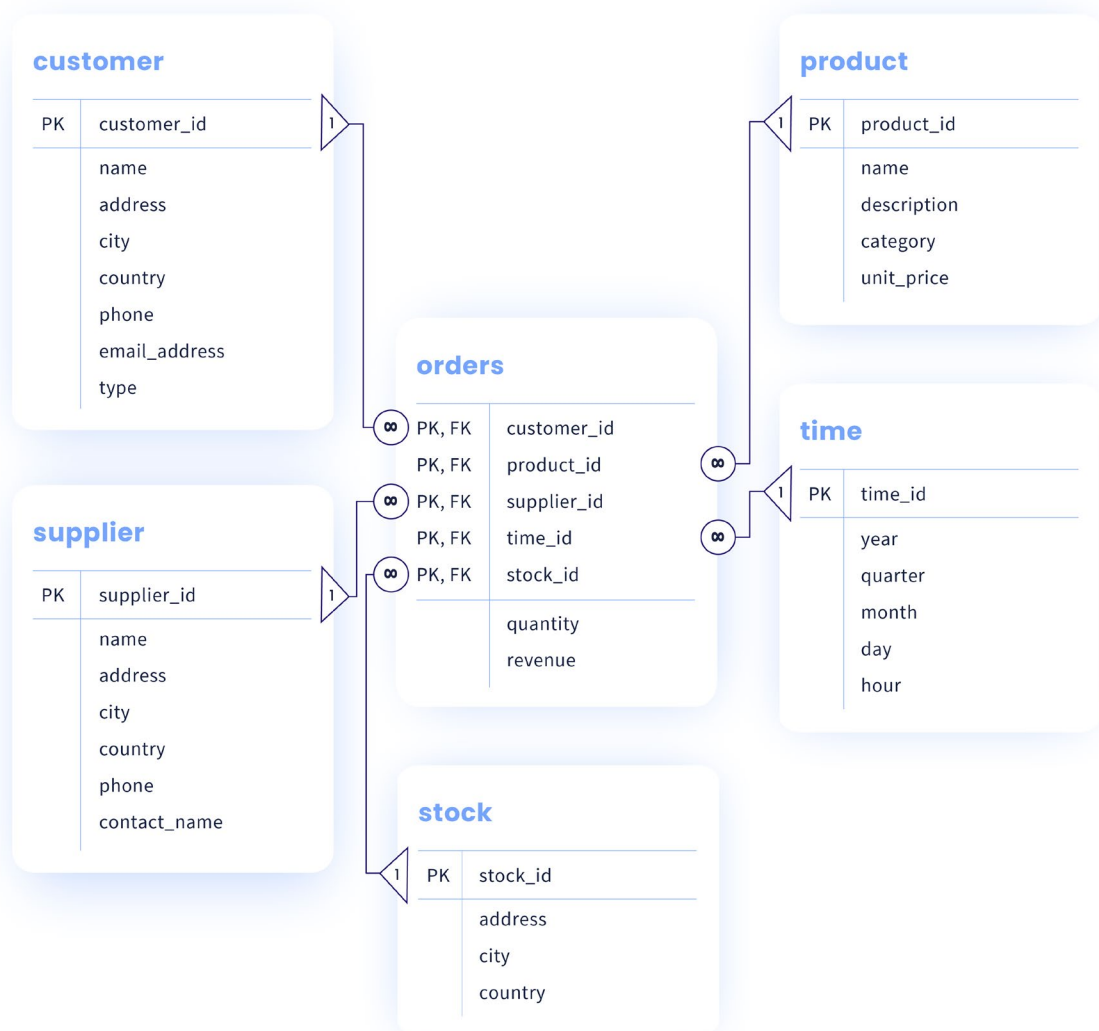
Eine wichtige Unterscheidung im Zweck einer Datenbank ist die Frage, ob es sich um operative Daten oder entscheidungsorientierte Daten handelt. Entscheidungsorientierte Datenbanken (oft Data Warehouse genannt) dienen der Analyse, zum Beispiel von Geschäfts- oder Umsatzzahlen. Operative Datenbanken ermöglichen die Geschäftsprozesse, zum Beispiel werden darin Transaktionen, Buchungen oder Bestellungen abgebildet.



## Star-Schema

Ein Design, dass sich als besonders effektiv für entscheidungsorientierte Datenbanken herausgebildet hat, ist das Star-Schema. Dabei sind mehrere Dimensionstabellen sternförmig um eine Faktentabelle organisiert. Die Faktentabelle enthält Kenn- oder Ergebniszahlen, die den laufenden Geschäftsbetrieb abbilden. Es können Kosten, Umsätze, Gewinne oder Ähnliches sein.

Die Dimensionstabellen enthalten deskriptive Daten. Durch das Design können analytische Fragenstellungen durch leicht verständliche Abfragen mit wenigen Joins beantwortet werden. Die Dimensionstabellen sind allerdings nicht normalisiert, was man aber in Kauf nimmt, weil die Nutzenden ausschließlich lesend auf die Datenbank zugreifen.



## Snowflake-Schema

Das Snowflake-Schema ist ein Design, das auf das Star-Schema aufbaut. Dabei werden die Dimensionstabellen zum Teil oder vollständig normalisiert. Das macht bei den Abfragen zusätzliche Joins nötig, kann aber in bestimmten Fällen die Performance verbessern. Ob und welche Normalisierungen sinnvoll sind, ist von der konkreten Anwendung abhängig.

