

Analoge Daten - Fotos, LP, Kassette, Film, Bücher, Karteikarten

Attribute - Spalten einer Tabelle, weisen Namen und Wertebereich auf. Die einzelnen Daten heißen Attributwerte.

Begriffstabelle

Tabelle	Relationale Datenbank	Entity-Relationship-Modell
Tabelle	Relationstyp, Relationschema	Entitätstyp
Spalte	Attribut	Attribut
Zeile	Tupel, Datensatz	Entität
Inhalt	Relation	Entitätsmenge
Zelle	Attributwert	Attributwert
Beziehung	Beziehung, Relation	Beziehung, Relation
Begriff	Synonym	
Tabelle	Relation, Entitätstyp, Entitätsmenge	
Tabellezeile	Tupel, Datensatz, Entität	
Tabelle-spalte	Attribut, Merkmal	
Wertebereich	Domäne	
Relation	Tabelle, Beziehung	
Beziehung	Relation, Verknüpfung (von Tabellen)	

Beziehung - Verbindung von zwei Datenbanken, benutzt Primärschlüssel und Fremdschlüssel.

Beziehung - verschiedene Arten, wie sich Daten aufeinander beziehen können.

Einfache Beziehung (1: / \rightarrow) - Jede Entität bezieht sich auf genau **eine** andere Entität einer zweiten Entitätsmenge.

Konditionelle Beziehung (c: / \rightarrow^+) - Jede Entität bezieht sich auf **maximal eine** andere Entität einer zweiten Entitätsmenge. Findet sich nicht in einer relationalen Datenbank.

Mehrfache Beziehung (m: / \rightarrow^+) - Jede Entität bezieht sich auf **mindestens eine** andere Entität einer zweiten Entitätsmenge.

Mehrfachkonditionelle Beziehung (mc: / \rightarrow^+) - Jede Entität bezieht sich auf beliebig viele andere Entitäten einer zweiten Entitätsmenge. Findet sich nicht in einer relationalen Datenbank.

Datei - Zusammenfassung von Daten auf dem Computer.

Ausführbare Datei - Programm in Binär, Machine-language, Skript.

Nichtausführbare Datei - Programm als Quelltext, Textdatei, Audiodatei, Bilddatei, Datenbankdatei.

Datentypen - Kategorisierung einer Eingabe nach Form

CHAR(n) - Zeichenkette mit Länge n.

VARCHAR(n) - Zeichenkette mit maximaler Länge n.

BLOB - Binary Large Object, eg. Photo.

BOOLEAN - true o. false.

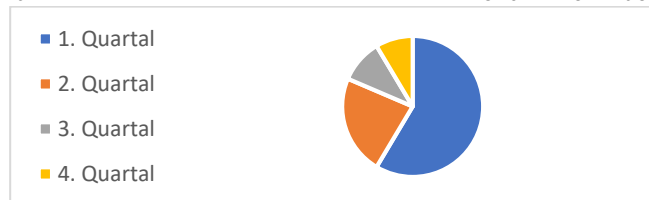
INTEGER - \mathbb{Z}

FLOAT - \mathbb{Q}

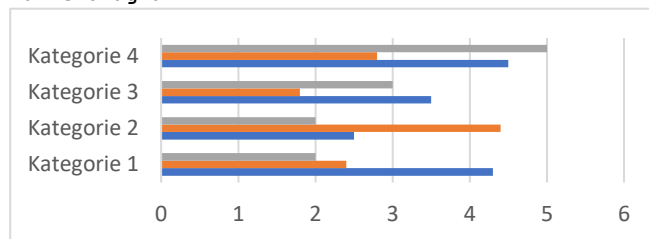
DATETIME - Datum und Zeit

Diagramme - Diagramme zur Datenvisualisierung.

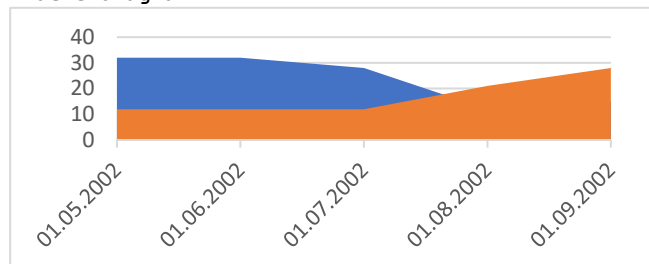
Kreisdiagramm



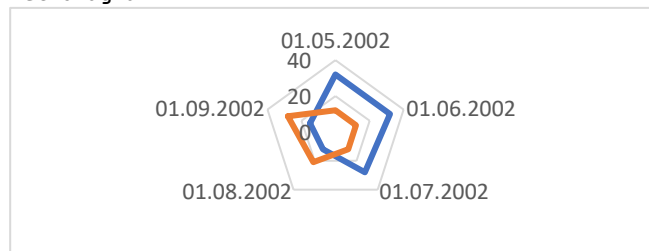
Balkendiagramm



Flächendiagramm



Netzdiagramm



Digitale Daten - Daten in elektronischer Form, können elektronisch bearbeitet werden.

Entity-Relationship-Modell - Konzeptuelle Darstellung von Datenbanken. Verwenden Tabellen, verbunden mit Pfeilen, bezeichnet mit Beziehung.

Konzeptionelles Datenmodell - Abstrahierte Darstellung ohne Zwischentabellen.

Logisches Datenmodell - Weiterentwicklung der konzeptionellen Darstellung, mit Zwischentabelle.

Physikalisches Datenmodell - Zeigt wie die Daten genau verarbeitet und abgespeichert werden.

Kardinalität - Beschreibt den Beziehungstyp zweier Entitäten in Fragerichtung. (siehe **Beziehung**)

Konsistenz - Abwesenheit von **Redundanzen**.

Namensgebung

- Tabellen haben bezeichnenden Namen.
- Tabellennamen sind Singular.
- Zwischentabellen werden alternativ auch **Tabelle1 Tabelle2** bezeichnet.
- Primärschlüssel ist Initialie des Tabellennamens und ID. eg. **XID**.
- Fremdschlüssel ist **FK_XID**.
- Die anderen Attribute haben einen bezeichnenden Namen.

Redundanz - Information, welche doppelt in Datenbank vorliegt. Kann mithilfe von Auslagerung einiger Attribute verhindert werden (Wohnorte/PLZ)

Schlüssel

Primärschlüssel - Datensatz-Identifizieren des Attribut in einer Tabelle. Bezieht sich auf einen einzigen Datensatz.

Fremdschlüssel - Verweist auf Datensatz in einer anderen Tabelle. Kann nur Werte, des jeweiligen Primärschlüssels annehmen. Muss nicht vergeben werden.

Skalentypen - Typen, nach welche Skalen kategorisiert werden könne, dies nach möglichen Operationen mithilfe Skala.

Nominalskala - Niedrigster Skalentyp

Definition:

Nominalskalen entsprechen dem niedrigsten Skalenniveau und dienen zur Darstellung und Klassifizierung qualitativer Eigenschaftsausprägungen ohne natürliche Ordnung.

Beispiele:

Staatsbürgerschaft, Geschlecht, Antworten vom Typ ja bzw. nein.

Mathematische Operationen:

=, ≠

Messbare Eigenschaften:

Häufigkeit: Bei Nominalskalen können lediglich sich unterscheidende Merkmalsausprägungen gezählt und die Häufigkeit einer Ausprägung dargestellt werden. Beispiel: In einer Klasse mit 24 Personen sind 13 Personen weiblich und 11 Personen männlich.

Ordinalskala - Zweitniedrigster Skalentyp

Definition:

Ordinalskalen erlauben die Aufstellung einer Rangordnung (besser/schlechter, größer/kleiner, häufiger/seltener usw.) mit Hilfe von Rangwerten und stehen bei den Skalenniveaus eine Stufe über der Nominalskala.

Beispiele:

Qualitätseinschätzung (sehr gut, gut, mittel, schlecht, sehr schlecht), Ränge beim Militär (der General steht über dem Leutnant, der Leutnant über dem Unteroffizier usw.).

Mathematische Operationen:

=, ≠; <, >

Neben dem Modus als Lageparameter kann auch ein Median identifiziert werden. Es kann jedoch kein Durchschnittswert errechnet werden. Beispiel: Ein Leutnant ist (rangmässig) > als ein Unteroffizier. Ein Soldat ist < als ein Unteroffizier.

Messbare Eigenschaften:

Häufigkeit, Rangfolge

Bei Ordinalskalen kann neben der Häufigkeit auch eine Rangfolge der Ausprägungen erstellt werden. Beispiel: In einer Befragung wird das Konsumverhalten in Bezug auf verschiedene Eissorten erfasst. Dabei müssen die Befragten für die Sorten Erdbeere, Vanille und Schoko angeben, ob sie diese (1) gar nicht, (2) kaum, (3) ganz gern oder (4) sehr gern essen. Aus dem Ergebnis lässt sich errechnen, ob Erdbeereis mehr Liebhaber findet als Vanilleeis. Nicht berechnen lässt sich aber, wie viel größer die Zuneigung zu Erdbeereis ist.

Kardinalskala (Intervallskala) - Zweithöchster Skalentyp

Definition:

Die Intervallskala zählt zu den metrischen Skalen, die quantitative Werte wiedergeben. Sie liegt beim Skalenniveau über der Nominal- und der Ordinalskala. Eine Intervallskala unterteilt sich immer in gleichgrosse Skalenabschnitte, hat aber keinen natürlichen Nullpunkt.

Beispiele:

Temperatur in Celsius, Jahreszahlen

Merkmal der Intervallskala ist, dass es zwar einen Nullpunkt geben kann, dieser aber "zufällig" gewählt worden ist. Es handelt sich nicht um den absoluten Nullpunkt oder um das "Nichts". Sowohl bei Fehler vorbehalten Keine Weitergabe oder Veränderung ohne ausdrückliche Gestattung durch Urheber.

der Temperatur in Celsius als auch bei den Jahreszahlen gibt es Werte, die unter 0 liegen. 0° Celsius = 32° Fahrenheit, 0° Fahrenheit = -18° Celsius.

Mathematische Operationen:

=, ≠; <, >, -, +

Bei Intervallskalen können die Lageparameter Modus, Median und arithmetisches Mittel berechnet (Durchschnittswert) werden.

Beispiel: 20° > 10°, 20° + 10° = 30°. Durchschnittstemperatur im Juli = 17.5°.

Messbare Eigenschaften:

Häufigkeit, Rangfolge, Abstand

Bei Intervallskalen kann die Häufigkeit ermittelt, die Rangfolge bestimmt und der Abstand zwischen zwei Ausprägungen berechnet werden.

Beispiel: Der Abstand von 15° zu 25° ist 10°.

Kardinalskala (Verhältnisskala) - Höchster Skalentyp

Definition:

Die Verhältnisskala (auch Rational-, Ratio- oder Proportionalskala genannt) ist das höchste Skalenniveau in der Statistik. Bei ihr handelt es sich um eine metrische Skala, um Unterschied zur Intervallskala existiert jedoch ein absoluter Nullpunkt.

Beispiele:

Temperatur in Kelvin, Masse, Grösse, Preis, Geschwindigkeit

Mathematische Operationen:

=, ≠; <, >, -, +, *, /

Es sind alle Operationen der Intervallskala möglich. Einzig bei diesem Skalenniveau sind zusätzlich Multiplikation und Division sinnvoll und erlaubt. Verhältnisse von Merkmalswerten dürfen also gebildet werden.

Beispiel: 10kg + 5kg = 15kg, 10kg * 2 = 20kg, die Hälfte von 100 CHF sind 50 CHF.

Messbare Eigenschaften:

Häufigkeit, Rangfolge, Abstand, absoluter Nullpunkt

Struktur

Unstrukturierte Daten - Dateien, welche schlecht durch den Computer verarbeitet werden können. Meist nur durch spezifische Programme. Photos, Grafiken, Videos, Audiodateien, unformatierter Text.

Semistrukturierte Daten - Haben keine generelle Struktur, tragen jedoch strukturinformation. XML, .docx, .xlsx.

Strukturierte Daten - Daten in vorgegebenem Format. in diesem Modul nur relationelle Datenbanken. Dazu werden Tabellen verwendet.

Def:

1. Entitätsmenge (Tabelle) hat einen eindeutigen Namen.
2. In einer Name ist der Attributname eindeutig für eine einzige Spalte/Attribut.
3. Jedes Attribut hat einen Datentyp.
4. Eindeutiger Primärschlüssel für ein Attribut.
5. Zeile (Einträge) einer Tabelle werden folgend benannt: **Entitäten, Datensätze, Tupel**.