

Datenbanken

Grenzen der Tabellenkalkulation

Öffne die Datei buch.xlsx beantworte die folgenden Fragen, indem du die Datei durchsuchst:

1. Wie viele Seiten hat das dickste Buch?

 2. Aus welchem Jahr ist das älteste Buch? (0 = Erscheinungsdatum unbekannt)

 3. Wie viele Bücher sind vor dem Jahr 2000 erschienen?

 4. Was kosten alle Bücher zusammen? (Denke an die mehrfachen Exemplare!)

 5. Von wie vielen verschiedenen Verlagen gibt es Bücher?

 6. Bei wie vielen Büchern kommt im Titel das Wort / der Wortbestandteil „deutsch“ vor?

 7. Von welchem Verlag sind am meisten Bücher vorhanden und wie viele sind es?

 8. Aus welchen Fachbereichen gibt es Bücher?

 9. Wie viele Bücher gibt es aus dem Fachbereich Physik?

 10. Wie viele Bücher gibt es von jedem Autor?

- Beschreibe kurz, wie du bei den schwierigen Fragen vorgegangen bist?

Zusammenfassung:

Rechenblätter erreichen sehr schnell ihre Grenzen, wenn

- _____
- _____
- _____

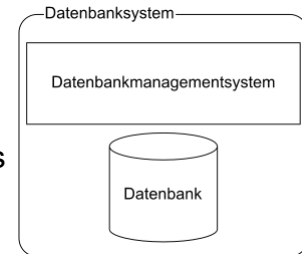
In solchen Anwendungsfällen ist eine _____ besser geeignet.

Datenbank – was ist das eigentlich?

In einem **Datenbanksystem (DBS)** werden die Daten derart gespeichert, dass ein schneller Zugriff auch auf komplexe Informationen möglich ist.

Ein DBS besteht aus:

- der eigentlichen **Datenbank (DB)**, in der die Daten in Form von Tabellen gespeichert werden.
- Dem **Datenbank-Management-System (DBMS)** ist das Verwaltungsprogramm, welches die Zugriffe (Einfügen, Löschen, Verändern, Abfragen, Benutzerrechte, ...) auf die Daten regelt.



Aufbau von Datenbanken

Eine relationale **Datenbank** besteht aus einer bzw. mehreren gegebenenfalls zusammenhängenden **Tabellen**. Die Tabellen beinhalten die zu verwaltenden Daten.

In der _____ wird in jeder _____ ein anderes Müsli beschrieben.

In jeder _____ wird eine _____ des jeweiligen Müslis beschrieben.

In jeder _____ steht der _____ der Eigenschaft.

ArtNr	Name	Marke	Typ	Hauptzutat	Preis	Bestand
1	Bircher MüsliDrink	MyMuesli	Drink	Bircher	1.65	54
2	Blaubeer MüsliDrink	MyMuesli	Drink	Blaubeeren	1.65	68
3	Apfel-Banane MüsliDrink	MyMuesli	Drink	Apfel-Banane	1.65	13
4	Bircher-Müsli	MyMuesli	Dose	Bircher	6.9	50
5	Himbircher-Müsli	MyMuesli	Dose	Bircher	6.9	1

In Fachsprache:

Eine Tabelle stellt einen _____ für gleichartige Objekte dar.

In jeder Zeile wird ein anderes _____ beschreiben.

Die Spalten sind die _____ der Objekte und in jeder Zelle steht der _____ für das jeweilige Attribut des Objekts.

Unterschiede zwischen einer Datenbanktabelle und einem Tabellenblatt in Excel

Datenbanktabelle	Tabellenblatt

Aufgaben des Datenbank-Management-Systems (DBMS)

- _____

- _____
- _____

- _____

Die Datenbank MySQL und das Webinterface phpMyAdmin

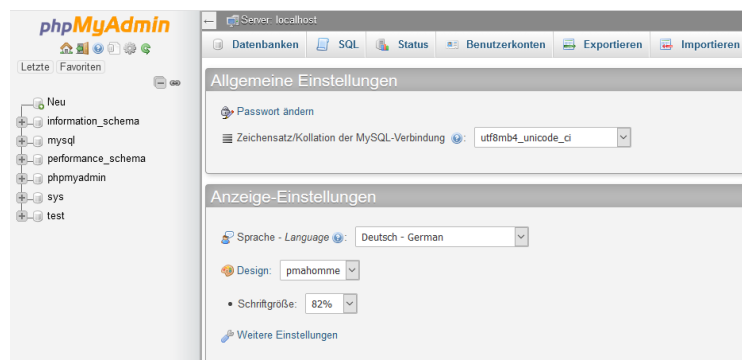
Auf einem Server (z. B. im Internet) ist das DBS **MySQL** installiert.

Der Zugriff vom Client-PC aus erfolgt über den Webbrowser. Dabei verbindet man sich mit dem Webinterface (grafische Administrationsoberfläche) **phpMyAdmin** des DBS.

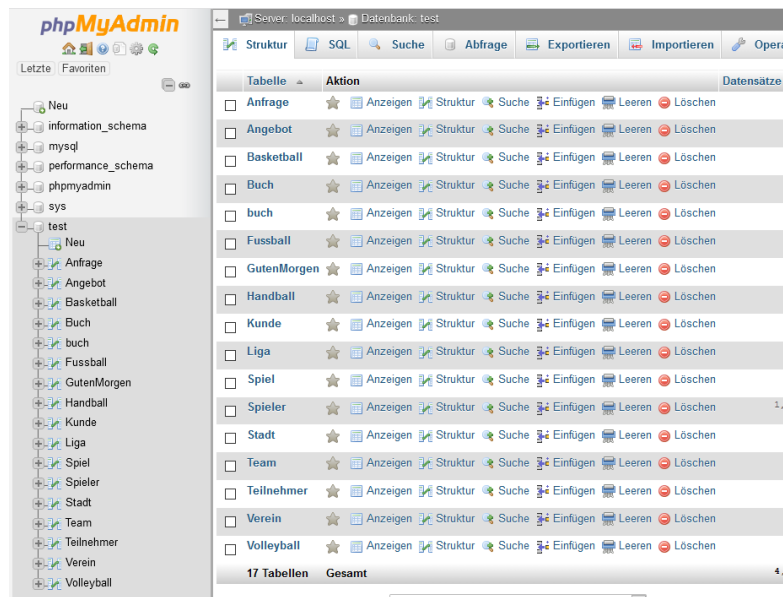


- Man erreicht die das DBS unter der URL:
<http://joachimhofmann.org/phpmyadmin>
- Der Benutzername besteht aus: **9x-Nachname** der Klasse, einem Bindestrich und dem Nachnamen (keine Umlaute: ae oe ue; kein ß: ss)
- Das Passwort lautet abc und kann geändert werden.

Nach der Anmeldung siehst du den Startbildschirm (das Home):



Mit einem Klick auf eine Datenbank gelangst du zu deren Tabellen, die du wiederum anklicken kannst, um sie anzusehen oder mit ihnen zu arbeiten.



Erste Abfragen in SQL

SQL steht für _____.

Mit Hilfe von SQL kann man _____ einer Abfrage an eine Datenbank _____.

Man formuliert in SQL also einfach, was man als Ergebnis haben möchte.

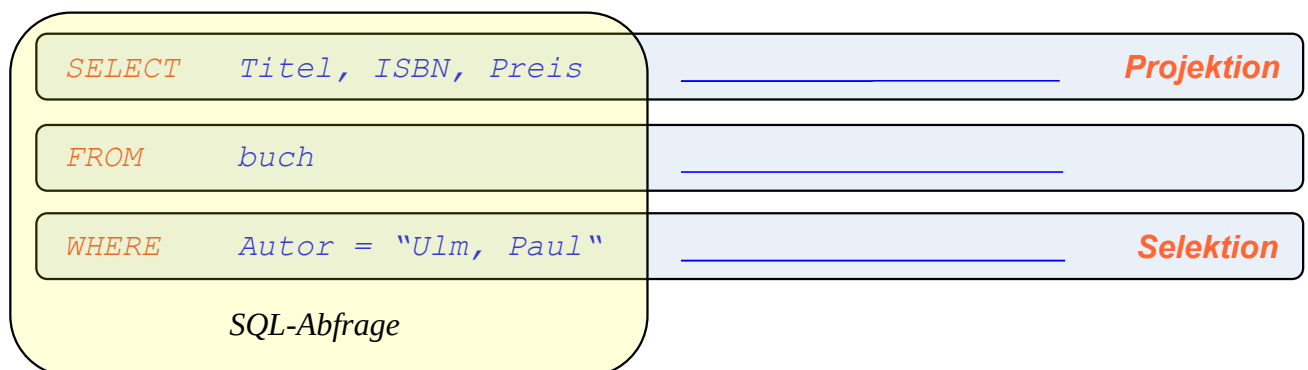
Beispiel:

Du befindest dich in der Datenbank und möchtest dort eine Abfrage an die Tabelle *buch* richten.

BuchNr	Titel	Fachbereich	Autor	Verlag	ISBN	Preis	Seitenzahl	Erscheinungsjahr	Exemplaranzahl
1	Abenteuer Bundesrepublik	G	Karlsfeld, Ulrich	Fasan	978-3-14345-456-5	39.00	93	1994	4
2	Absolutismus	G	Worishofen, Johannes	Reiher	978-3-99-078907-0	64.00	82	1991	3
3	Adenauer und die deutsche Frage	G	Regensburg, Josef	Reiher	978-3-99-758422-7	31.00	117	1993	3
4	Adolf Hitler	G	Solnhofen, Karl	Kohlrahe	978-3-23-324442-7	31.00	80	1976	4
5	Afrikas Weg in die Gegenwart	G	Illertissen, Herbert	Bachstelze	978-3-887-06890-5	34.00	82	1968	2

Du interessierst dich für **Titel**, **ISBN** und **Preis** aller Bücher der **Tabelle buch**, deren **Autor „Ulm, Paul“** heißt.

In SQL formuliert man das dann so:



Als Ergebnis erhältst du dann eine (kleinere) Tabelle, in der alle Datensätze stehen, welche die Bedingung deiner _____ erfüllen. Du siehst auch nur noch diejenigen Attribute und deren Werte, welche du durch die _____ angegeben hast.

Eine **SQL-Abfrage** besteht in der Regel aus folgenden drei Teilen:

1. **Projektion**: Hier werden die interessanten Attribute (**Spalten**) ausgewählt.
2. Angabe der **Tabelle(n)** die durchsucht werden soll(en).
3. **Selektion**: Hier werden durch eine (mehrere) Bedingung(en), die Datensätze (**Zeilen**) ausgewählt.

Praktische Übungsaufgaben

1. Liste alle Bücher der Tabelle buch auf, deren Erscheinungsjahr 1981 ist.

Zeige Titel und Fachbereich.

Anzahl Datensätze: _____

2. Liste alle Bücher der Tabelle buch auf, deren Erscheinungsjahr größer als 1981 ist.

Zeige Titel und Erscheinungsjahr.

Anzahl Datensätze: _____

3. Liste alle Bücher der Tabelle buch auf, die vom Verlag Kohlrahe sind.

Zeige alle Spalten.

Anzahl Datensätze: _____

Logische Funktionen in SQL

Hierfür gibt es die Schlüsselwörter _____ , _____ .

Beispiele:

Welche Bücher haben mehr als 100 Seiten und kosten 39 €?

SELECT *

FROM buch

WHERE _____

Welche Bücher werden für den Fachbereich Informatik oder Physik angeboten?

SELECT *

FROM buch

WHERE _____

Welche Bücher werden von den Verlagen Uhu und Zeisig veröffentlicht?

SELECT *

FROM buch

WHERE _____

Welche Bücher sind nicht vom Verlag Kohlrahe veröffentlicht und sind dem Fachbereich Informatik oder Mathematik zugeordnet? (Tipp: Nutze passende Klammerung, um die passende Logik zu erzeugen!)

SELECT *

FROM buch

WHERE _____

IN: Mehrere Werte zulassen

Hierfür gibt es das Schlüsselwort _____ .

Beispiel:

Gib alle Bücher aus, die in den Jahre 2000, 2003 und 2004 erschienen sind.

SELECT *

FROM buch

WHERE Erscheinungsjahr _____

BETWEEN: Durchgehende Wertebereiche

Hierfür gibt es das Schlüsselwort _____ .

Beispiel:

Gib alle Bücher aus, die zwischen 50 und 100 Seiten haben.

SELECT *

FROM buch

WHERE Seitenzahl _____

ORDER BY: Ergebnistabelle sortieren

Eine **vierte Zeile** wird ergänzt:

_____	_____
_____	_____

Beispiel:

Gib alle Büchertitel inklusive Seitenzahl aus, die weniger als 100 Seiten haben. Sortiere so, dass die Bücher mit den meisten Seiten oben stehen.

SELECT Titel, Seitenzahl

FROM buch

WHERE _____

Wertvolle Helfer

Mit dem Schlüsselwort **DISTINCT** kann man in SQL _____
_____.

Kommt ein Attributwert in der Ergebnistabelle mehrmals vor, so kann man dies mit DISTINCT _____.

Beispiel:

Von welchen verschiedenen Verlagen gibt es Bücher? (keine Mehrfachnennung)

SELECT _____ **Verlag**

FROM buch

Manchmal sucht man nicht nach einem Attributwert als Ganzes sondern nur nach einem **Teil des Attributwerts**.

In diesem Fall benutzt man den **%-Operator**, der als _____
_____ gilt.

Der Vergleich erfolgt dann nicht mit dem = sondern mit dem **LIKE-Operator**.

Die logische Funktion der Verneinung wird mit dem Schlüsselwort **NOT** wird z.B. bei Abfragen mit **LIKE** benötigt, da man hier nicht den Operator **!=** verwenden kann.

Beispiel:

Bei welchen Büchern kommt im Titel das Wort / der Wortbestandteil „deutsch“ vor?

SELECT *

FROM buch

WHERE Titel _____

Bei welchen Büchern kommt im Titel **nicht** der Wortbestandteil „deutsch“ vor?

SELECT *

FROM buch

WHERE Titel _____

Übungsaufgabe

Welche Verlage haben „fink“ im Namen? Jeder Verlag soll nur einmal genannt werden!

SELECT _____

FROM buch

WHERE _____

Aggregatfunktionen in SQL

Der Begriff **Aggregat** bedeutet, _____
_____.

In SQL gibt es sog. **Aggregat- oder Gruppenfunktionen**. Sie können aus mehreren Werten einer Spalte etwas berechnen. Außerdem können sie **nicht** bei in der WHERE-Klausel verwendet werden.

Es gibt folgende 5 Aggregatfunktionen:

- | | | |
|----|-------|---|
| 1. | _____ | Minimum aller Werte |
| 2. | _____ | Maximum aller Werte |
| 3. | _____ | Summe aller Werte |
| 4. | _____ | Mittelwert aller Werte (engl.: average) |
| 5. | _____ | zähle alle Datensätze |

Beispiele anhand der Tabelle ***buch***:

1. Wie viele Seiten hat das dickste Buch?

SELECT _____
FROM buch

2. Aus welchem Jahr ist das älteste Buch? (0 = Erscheinungsjahr unbekannt)

SELECT _____
FROM buch
WHERE _____

3. Wie viele Bücher sind nicht älter als 20 Jahre?

SELECT _____
FROM buch
WHERE _____

4. Was kosten alle Bücher zusammen?

SELECT _____
FROM buch

5. Wie viele Exemplare sind im Durchschnitt pro Buch vorhanden?

SELECT _____
FROM buch

AS: Umbenennung der Spaltenbeschriftungen

Nach der Aggregatfunktion gibt man mittels **AS** "neuer Spaltenname" die gewünschte Spaltenüberschrift an:

Beispiel:

Wie viele Bücher kosten mehr als 40 €. Benenne die Spalte in Anzahl Bücher > 40€ um.

SELECT _____
FROM buch
WHERE Preis > 40

Unterabfragen

Unterabfragen sind eigenständige SQL-Abfragen, die in einer anderen SQL-Abfrage (in der Regel) bei der **WHERE-Klausel integriert** werden. Die Unterabfragen müssen **geklammert** sein.

Die zwei häufigsten Fälle in denen Unterabfragen benötigt werden, sind:

- Wenn nach einem **Attribut in Kombination mit einem aggregierten Attribut** gefragt wird. Diese dürfen nie zusammen in einer Projektion stehen.
- Wenn die Selektion anhand von **Attributwerten aus anderen Tabellen** erfolgt.

Beispiele:

Welche Spieler haben die meisten Tore geschossen? Zeige Spielernamen und Tore.

```
SELECT _____  
FROM Spieler  
WHERE _____  
_____
```

Welche Spieler spielen beim Verein 1. FC Nürnberg? Zeige Spielernamen. (Die Vereins_ID wissen wir noch nicht bzw. wollen sie nicht händisch nachschauen!)

```
SELECT _____  
FROM Spieler  
WHERE _____  
_____  
_____
```

GROUP BY

Die **Group-By**-Anweisung gruppiert Zeilen (Datensätze), die **denselben Wert** für das übergebene Attribut aufweisen. Häufig wird die Group-By-Anweisung in **Kombination mit Aggregatfunktionen** verwendet, um Informationen über die Werte eines Attributs zu erhalten. Group-By-Anweisung reiht sich als vierter Bestandteil in die SQL-Abfrage ein.

Beispiele:

Wie viele Spieler aus den jeweiligen Ländern spielen in der Bundesliga? Benenne die Ergebnisspalte entsprechend um.

SELECT _____
FROM Spieler

Was kostet das teuerste Buch der jeweiligen Fachbereiche? Benenne die Ergebnisspalte entsprechend um und der Fachbereich mit dem teuersten Buch soll oben stehen.

SELECT _____
FROM buch

HAVING-Klausel

Die Having-Klausel ist der sechste und letzte Baustein für die SQL-Abfragen (in Reihenfolge: **SELECT, FROM, WHERE, GROUP BY, HAVING, ORDER BY**).

Das **HAVING** ist prinzipiell eine **zweite Bedingung/Selektion** (zweites WHERE), die nach der Gruppierung ausgeführt wird und sich somit auf die „Gruppen“ bezieht. Somit kann es eine HAVING-Klausel nur geben, wenn es auch ein GROUP BY gibt.

Im Gegensatz zum WHERE dürfen beim HAVING **Aggregatsfunktionen** (AVG(), COUNT(), MAX(), MIN(), SUM()) angewendet werden!

Beispiele:

- Es sollen die Durchschnittstore von den deutschen Spielern nach der jeweiligen Trikotnummer ausgegeben werden. Dabei sollen nur die Trikotnummern ausgegeben werden, bei denen der Durchschnitt größer als 1 ist. Die Trikotnummer mit den höchsten Durchschnittstoren soll oben stehen.
- Welcher Autor hat mit allen seinen Büchern mehr als 500 Seiten veröffentlicht? Hierbei sollen nur die Bücher betrachtet werden, die weniger als 50€ kosten. Der Autor mit den meisten veröffentlichten Seiten soll oben stehen.

Eine sehr komplexe SQL-Abfrage kann neben den sechs Bausteinen – die Aggregatsfunktionen enthalten können – die Informationen über mehrere Tabellen (Stoff der 10. Jahrgangsstufe) zusammenfügen und zusätzlich noch Unterabfragen enthalten, die ebenfalls komplex sein können.

Aufbau und Veränderung von Tabellen in Datenbanken

Eine **Tabelle** stellt einen _____ für gleichartige Objekte dar.

Bei Tabellen entspricht eine Zeile einem _____.

Die Spalten sind die _____ der Objekte.



B U C H
Titel Autor Seitenzahl Verlag

buch_1	buch_2	buch_3
Titel = Informatik 2 Autor = Hubwieser Seitenzahl = 179 Verlag = Klett	Titel = Informatik I Autor = Brichzin Seitenzahl = 159 Verlag = Oldenbourg	Titel = Debian Linux Autor = Amberg Seitenzahl = 893 Verlag = Mitpress

B U C H			
Titel	Autor	Seitenzahl	Verlag
Informatik 2	Hubwieser	179	Klett
Informatik I	Brichzin	159	Oldenbourg
Debian Linux	Amberg	893	Mitpress

Für jedes Attribut muss ein geeigneter _____ gewählt werden.

Gängige Datentypen sind:

int	Ganze Zahl	
decimal	Kommazahl	(Dezimalpunkt: 1.36 nicht 1,36)
varchar	Zeichenfolge	(Anführungszeichen: "Hallo")
date	Datum	(Format: JJJJ-MM-TT in Anführungszeichen)
time	Zeit	(Format: hh:mm oder hh:mm:ss in Anführungszeichen)

Somit ergibt sich das _____ :

BUCH (Titel : varchar ; Autor : varchar ; Seitenzahl : int ; Verlag : varchar)
--

B U C H

Titel : varchar Autor : varchar Seitenzahl : int Verlag : varchar
--

Mit dem Schema modelliert man die logische Struktur der Tabelle, bevor man sie im DBS anlegt.

Aufgabe 1:

Eine Autovermietung möchte ihre Autos in einer Datenbank speichern. Dazu braucht man die Marke (VW, Audi, ...), Modell (Golf, Passat, ...), Kennzeichen, Kilometerstand, Ausleihdatum, Ausleihuhrzeit, Ausleiher, PreisproStunde (z.B. 19,49€) .

Gib eine erweiterte Klassenkarte oder das Schema der nötigen Tabelle an.

Aufgabe 2:

Für einen Sportverein sollen die angebotenen Trainings in einer Datenbanktabelle gespeichert werden. Dazu braucht man die Sportart, den Wochentag1 des Trainings, die Uhrzeit1, den Wochentag2, die Uhrzeit2, den Jahresbeitrag für diese Sportart sowie die maximale Teilnehmerzahl und die aktuelle Teilnehmerzahl.

Gib eine erweiterte Klassenkarte oder das Schema der nötigen Tabelle an.

Aufgabe 3:

In der Schule möchten wir die Schüler in einer Datenbanktabelle speichern. Dazu brauchen wir deren Nachnamen, Vornamen, Konfession, Geschlecht, Geburtsdatum, Klasse.

Gib eine erweiterte Klassenkarte oder das Schema der nötigen Tabelle an.