

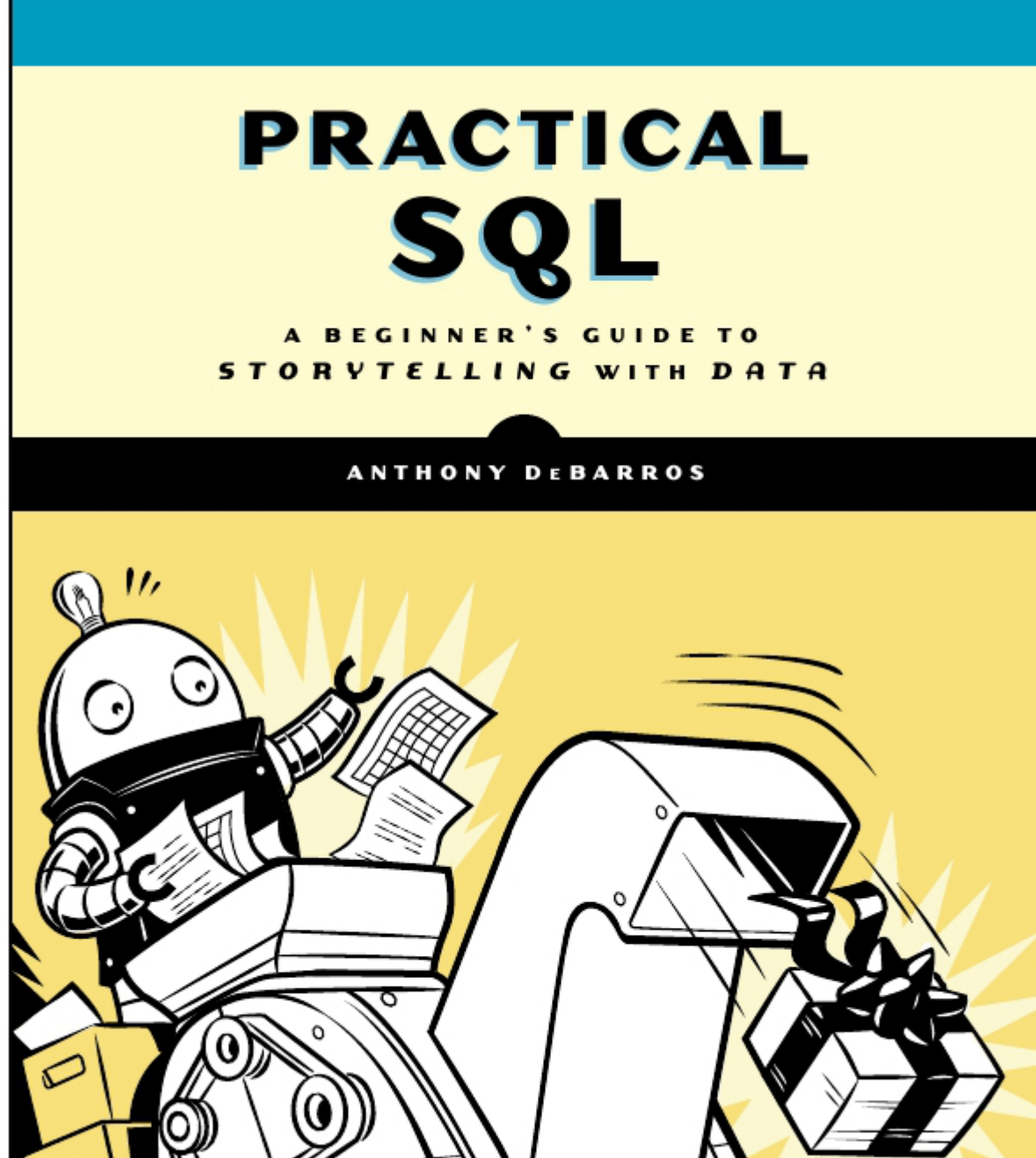
# Einführung in SQL

Selektieren / Ordnen / Filtern

Prof. Dr. Jan Kirenz

HdM Stuttgart

Dieses Tutorial  
basiert auf  
"A Beginner's  
Guide to  
Storytelling  
with Data"  
from Anthony  
DeBarros (2018)



# Setup in R

## Verbindung zur lokalen Datenbank herstellen

- In R kann das Paket `RPostgres` genutzt werden
- [RPostgres Übersicht](#)
- [RPostgres Dokumentation](#)
- Das Passwort `pw` muss zuvor als Objekt gespeichert werden:

```
pw = "hier Passwort eingeben"
```

```
library(DBI)  
library(RPostgres)
```

```
con <- dbConnect(RPostgres::Postgres(),  
  dbname = "postgres",  
  host = "localhost",  
  port = 5432,  hier den eigenen Port angeben (wurde bei der  
  user = "postgres",  
  password = pw)
```

# Tabelle erstellen

- Tabelle erstellen mit CREATE TABLE.
- Wir erstellen zunächst eine Variable mit dem Namen `id` und versehen diese mit dem Datentyp `bigserial`.

```
CREATE TABLE teachers (  
  id bigserial,  
  first_name varchar(25),  
  last_name  varchar(50),  
  school     varchar(50),  
  hire_date  date,  
  salary     numeric  
);
```

# Daten in Tabelle einfügen.

- Bei Zahlenwerten sind keine Anführungszeichen erforderlich.

```
INSERT INTO teachers (first_name, last_name, school, hire_date, salary)
VALUES ('Janet', 'Smith', 'F.D. Roosevelt HS', '2011-10-30', 36200),
      ('Lee', 'Reynolds', 'F.D. Roosevelt HS', '1993-05-22', 65000),
      ('Samuel', 'Cole', 'Myers Middle School', '2005-08-01', 43500),
      ('Samantha', 'Bush', 'Myers Middle School', '2011-10-30', 36200),
      ('Betty', 'Diaz', 'Myers Middle School', '2005-08-30', 43500),
      ('Kathleen', 'Roush', 'F.D. Roosevelt HS', '2010-10-22', 38500);
```

# Datenexploration mit SELECT

- Tabelle anzeigen

```
SELECT *  
FROM teachers;
```

id	first_name	last_name	school	hire_date	salary
1	Janet	Smith	F.D. Roosevelt HS	2011-10-30	36200
2	Lee	Reynolds	F.D. Roosevelt HS	1993-05-22	65000
3	Samuel	Cole	Myers Middle School	2005-08-01	43500
4	Samantha	Bush	Myers Middle School	2011-10-30	36200
5	Betty	Diaz	Myers Middle School	2005-08-30	43500
6	Kathleen	Roush	F.D. Roosevelt HS	2010-10-22	38500

# Datenexploration mit SELECT

- Abfrage an Datenbank stellen:

```
SELECT first_name  
FROM teachers;
```

---

<b>first_name</b>
-------------------

Janet
-------

Lee
-----

Samuel
--------

Samantha
----------

Betty
-------

Kathleen
----------

---

# Datenexploration mit SELECT

```
SELECT first_name, last_name, salary  
FROM teachers;
```

first_name	last_name	salary
Janet	Smith	36200
Lee	Reynolds	65000
Samuel	Cole	43500
Samantha	Bush	36200
Betty	Diaz	43500
Kathleen	Roush	38500



# Datenexploration mit SELECT

## DISTINCT

- Auswahl von einzigartigen Fällen

```
SELECT DISTINCT school  
FROM teachers;
```

<b>school</b>
Myers Middle School
F.D. Roosevelt HS

# Datenexploration mit SELECT

## DISTINCT

```
SELECT DISTINCT school, salary  
FROM teachers;
```

<b>school</b>	<b>salary</b>
Myers Middle School	36200
F.D. Roosevelt HS	65000
Myers Middle School	43500
F.D. Roosevelt HS	38500
F.D. Roosevelt HS	36200

# Datenexploration mit SELECT

## ORDER BY

- Der Befehl verändert nicht die Original-Tabelle.
- ORDER BY sortiert automatisch aufsteigend (ASC für ascending).
- Falls eine absteigende Reihenfolge gewünscht wird, muss der Befehl DESC für descending hinzugefügt werden.

# Datenexploration mit SELECT

## ORDER BY

```
SELECT first_name, last_name, salary  
FROM teachers  
ORDER BY salary DESC;
```

first_name	last_name	salary
Lee	Reynolds	65000
Samuel	Cole	43500
Betty	Diaz	43500
Kathleen	Roush	38500
Janet	Smith	36200
Samantha	Bush	36200

# Datenexploration mit SELECT

- Aufgabe: Erstellen Sie eine Tabelle mit den zuletzt eingestellten Lehrern pro Schule. Die Ergebnisse sollen nach Schule (aufsteigend) und Einstellungsdatum (absteigend) sortiert werden.

last_name	school	hire_date
Smith	F.D. Roosevelt HS	2011-10-30
Roush	F.D. Roosevelt HS	2010-10-22
Reynolds	F.D. Roosevelt HS	1993-05-22
Bush	Myers Middle School	2011-10-30
Diaz	Myers Middle School	2005-08-30
Cole	Myers Middle School	2005-08-01

# Datenexploration mit SELECT

## WHERE

```
SELECT last_name, school, hire_date  
FROM teachers  
WHERE school = 'Myers Middle School'
```

last_name	school	hire_date
Cole	Myers Middle School	2005-08-01
Bush	Myers Middle School	2011-10-30
Diaz	Myers Middle School	2005-08-30

<b>Zeichen</b>	<b>Bedeutung</b>	<b>Beispiel</b>
=	Equal to	WHERE school = 'Myers Middle School'
<>	or != Not equal to	WHERE school <> 'Myers Middle School'
>	Größer als	WHERE salary > 20000
<	Kleiner als	WHERE salary < 20000
>=	Größer als oder gleich	WHERE salary >=20000
<=	Kleiner als	WHERE salary <= 60500
BETWEEN	Zwischen	WHERE salary BETWEEN 20000 AND 40000
IN	Innerhalb	WHERE last_name IN ('Bush', 'Roush')
LIKE	Wie (case sensitive)	WHERE first_name LIKE 'Sam %'
ILIKE	Wie (case insensitive)	WHERE first_name ILIKE 'sam %'
NOT	Negiert eine der oberen Bedingungen	WHERE first_name NOT LIKE 'sam %'

# Datenexploration mit SELECT

## WHERE

- Aufgabe: In welcher Schule unterrichtet Janet?

first_name	school
Janet	F.D. Roosevelt HS



# Datenexploration mit SELECT

## WHERE !=

```
SELECT school  
FROM teachers  
WHERE school != 'F.D. Roosevelt HS';
```

<b>school</b>
F.D. Roosevelt HS
F.D. Roosevelt HS
Myers Middle School
Myers Middle School
Myers Middle School
F.D. Roosevelt HS

# Datenexploration mit SELECT

WHERE <

- Welche Lehrer\_innen wurden vor dem 01 Januar 2001 eingestellt?

first_name	last_name	hire_date
Lee	Reynolds	1993-05-22

# Datenexploration mit SELECT

## WHERE >=

```
SELECT first_name, last_name, salary  
FROM teachers  
WHERE salary >= 43500;
```

first_name	last_name	salary
Lee	Reynolds	65000
Samuel	Cole	43500
Betty	Diaz	43500

# Datenexploration mit SELECT

## WHERE BETWEEN

```
SELECT first_name, last_name, school, salary  
FROM teachers  
WHERE salary BETWEEN 40000 AND 65000
```

first_name	last_name	school	salary
Lee	Reynolds	F.D. Roosevelt HS	65000
Samuel	Cole	Myers Middle School	43500
Betty	Diaz	Myers Middle School	43500

# Datenexploration mit SELECT

## LIKE, ILIKE und WHERE

- Prozentzeichen: (%) Platzhalter für ein oder mehrere Zeichen.
- Unterstrich: (\_) Platzhalter für genau *ein* Zeichen
- Beispiel für die Suche nach dem Namen Baker mit LIKE (case sensitive):
  - LIKE 'b%'
  - LIKE '%ak%'
  - LIKE '\_aker%'
  - LIKE 'ba\_er'

# Datenexploration mit SELECT

## LIKE, ILIKE und WHERE

- Aufgabe: Welche Vornamen beginnen mit "sam" (mit Berücksichtigung der Groß- und Kleinschreibung)?

**first\_name**

- Diese Abfrage führt zu keinem Resultat, da LIKE case sensitive ist.

# Datenexploration mit SELECT

## WHERE ILIKE

- Aufgabe: Welche Vornamen beginnen mit "sam" (ohne Berücksichtigung der Groß- und Kleinschreibung)?

<b>first_name</b>
Samuel
Samantha

# Datenexploration mit SELECT

## LIKE & ILIKE

- Generell ist es empfehlenswert, **ILIKE** in Verbindung mit Platzhaltern zu nutzen, um nicht versehentlich Resultate aus der Suche auszuschließen.
- Dadurch werden auch Ergebnisse mit eingeschlossen, die versehentlich falsch in die Datenbank eingetragen worden sind.
- Da LIKE und ILIKE nach Mustern in Daten suchen, kann die Suche relativ lange dauern.



# Datenexploration mit SELECT

## WHERE, AND und OR

- Es ist oft sinnvoll, Operatoren in einer Abfrage zu verbinden.
- Dies kann mit AND und OR umgesetzt werden.

```
SELECT *  
FROM teachers  
WHERE school = 'Myers Middle School'  
AND salary < 40000;
```

id	first_name	last_name	school	hire_date	salary
4	Samantha	Bush	Myers Middle School	2011-10-30	36200

# Datenexploration mit SELECT

## WHERE & OR

- Aufgabe: Auflistung aller Personen mit Nachnamen Cole oder Bush.

id	first_name	last_name	school	hire_date	salary
3	Samuel	Cole	Myers Middle School	2005-08-01	43500
4	Samantha	Bush	Myers Middle School	2011-10-30	36200

# Datenexploration mit SELECT

## WHERE, AND und OR

- Aufgabe: Welche Lehrer sind an der F.D. Roosevelt HS beschäftigt und haben entweder einen Verdienst unter 38.000 oder über 40.000?

id	first_name	last_name	school	hire_date	salary
1	Janet	Smith	F.D. Roosevelt HS	2011-10-30	36200
2	Lee	Reynolds	F.D. Roosevelt HS	1993-05-22	65000

# Datenexploration mit SELECT

## Kombinierte Abfragen

Beispiel für das allgemeine Schema eine Abfrage:

```
SELECT column_names  
FROM table_names  
WHERE criteria  
ORDER BY column_names
```

```
SELECT first_name, last_name, school, hire_date, salary
FROM teachers
WHERE school LIKE '%Roos%'
ORDER BY hire_date DESC;
```

first_name	last_name	school	hire_date	salary
Janet	Smith	F.D. Roosevelt HS	2011-10-30	36200
Kathleen	Roush	F.D. Roosevelt HS	2010-10-22	38500
Lee	Reynolds	F.D. Roosevelt HS	1993-05-22	65000

# Datenexploration mit SELECT

## Übungen

- Welche Lehrer\_innen (Name und Vorname) sind an welchen Schulen angestellt? Das Ergebnis soll nach Schulen und Nachnamen geordnet werden (aufsteigend).
- Welche Lehrer\_innen, deren Vorname mit "S" beginnen, haben ein Einkommen größer als 40.000?
- Geben Sie das Einkommen der Lehrer\_innen an, die ab 2010-01-01 eingestellt worden sind. Das höchste Gehalt soll ganz oben in der Tabelle aufgelistet sein.

# Datenexploration mit SELECT

## Übungen

- Welche Lehrer\_innen (Name und Vorname) sind an welchen Schulen angestellt? Das Ergebnis soll nach Schulen und Nachnamen geordnet werden (aufsteigend).

```
SELECT school, last_name, first_name  
FROM teachers  
ORDER BY school, last_name ASC
```

<b>school</b>	<b>last_name</b>	<b>first_name</b>
F.D. Roosevelt HS	Reynolds	Lee
F.D. Roosevelt HS	Roush	Kathleen
F.D. Roosevelt HS	Smith	Janet
Myers Middle School	Bush	Samantha
Myers Middle School	Cole	Samuel
Myers Middle School	Diaz	Betty



# Datenexploration mit SELECT

## Übungen

- Welche Lehrer\_innen, deren Vorname mit "S" beginnen, haben ein Einkommen größer als 40.000?

salary	last_name	first_name
43500	Cole	Samuel

# Datenexploration mit SELECT

## Übungen

- Geben Sie das Einkommen der Lehrer\_innen an, die ab 2010-01-01 eingestellt worden sind. Das höchste Gehalt soll ganz oben in der Tabelle aufgelistet sein.

salary	last_name	first_name	hire_date
38500	Roush	Kathleen	2010-10-22
36200	Smith	Janet	2011-10-30
36200	Bush	Samantha	2011-10-30

# Vielen Dank!

Prof. Dr. Jan Kirenz

HdM Stuttgart  
Nobelstraße 10  
70569 Stuttgart

