

Databases

2. Single-table queries

Arni Magnusson

United Nations University
Fisheries Training Programme

6–9 Nov 2017

Outline

What is a database

purpose, design, data types

Create database

software, import data

Query

get data, join tables, SQL language

Interface

connect to database from other program

Goals

After this database course, you should:

1. **Understand** what a database is, and how it works
2. Be able to **create** a simple database
3. Be able to **get data** from any database

Database

How does a database work?

Database

Tables

- A database is a collection of **tables**
- Related tables are joined using **key** columns
- Each table column has one **data type**

Database

Tables

- A database is a collection of **tables**
- Related tables are joined using **key** columns
- Each table column has one **data type**

Queries

- Users ask for specific data with a **query**
- Queries are written in a standard language called **SQL**

Database



Visual query builder

- Create – Query design
- Add table
- Drag columns
- Click **View** to toggle between query and result
- View SQL

SQL

The SQL language

SELECT, FROM

Select table columns

```
SELECT column FROM table
```

```
SELECT col1, col2 FROM table
```

SELECT, FROM

Select table columns

```
SELECT column FROM table
```

```
SELECT col1, col2 FROM table
```

All columns

```
SELECT * FROM table
```

SELECT, FROM

Alias

SELECT column AS *newname* FROM table

SELECT, FROM

Alias

```
SELECT column AS newname FROM table
```

Arithmetic

```
SELECT col1 + col2 AS newname FROM table
```

```
SELECT col1 - col2 AS newname FROM table
```

```
SELECT col1 * col2 AS newname FROM table
```

```
SELECT col1 / col2 AS newname FROM table
```

WHERE

Condition (number)

```
SELECT * FROM table WHERE col = 0
```

```
SELECT * FROM table WHERE col > 0
```

```
SELECT * FROM table WHERE col BETWEEN 1 AND 3
```

```
SELECT * FROM table WHERE col IN (1,3)
```

WHERE

Condition (number)

```
SELECT * FROM table WHERE col = 0
```

```
SELECT * FROM table WHERE col > 0
```

```
SELECT * FROM table WHERE col BETWEEN 1 AND 3
```

```
SELECT * FROM table WHERE col IN (1,3)
```

Many conditions

```
SELECT * FROM table WHERE col1 = 0 AND col2 > 0
```

```
SELECT * FROM table WHERE col1 = 0 OR col2 > 0
```

WHERE, ORDER BY

Condition (string)

```
SELECT * FROM table WHERE col = 'foo'
```

```
SELECT * FROM table WHERE col IN ('foo', 'bar')
```

```
SELECT * FROM table WHERE col LIKE '%fish%'
```

WHERE, ORDER BY

Condition (string)

```
SELECT * FROM table WHERE col = 'foo'
```

```
SELECT * FROM table WHERE col IN ('foo', 'bar')
```

```
SELECT * FROM table WHERE col LIKE '%fish%'
```

Condition (NA)

```
SELECT * FROM table WHERE col IS NOT NULL
```

Sort

```
SELECT * FROM table ORDER BY col
```

GROUP BY

Function

```
SELECT avg(col) FROM table
```

```
SELECT sum(col) FROM table
```

```
SELECT min(col) FROM table
```

```
SELECT max(col) FROM table
```

```
SELECT count(col) FROM table
```

GROUP BY

Function

```
SELECT avg(col) FROM table
```

```
SELECT sum(col) FROM table
```

```
SELECT min(col) FROM table
```

```
SELECT max(col) FROM table
```

```
SELECT count(col) FROM table
```

Aggregate

```
SELECT col1, avg(col2) FROM table GROUP BY col1
```

Overview of SQL commands

Required

SELECT

FROM

Overview of SQL commands

Required

SELECT
FROM

Condition

WHERE
BETWEEN ... AND
IN
LIKE
IS NULL
AND
OR
NOT

Overview of SQL commands

Required

SELECT
FROM

Sort

ORDER BY

Condition

WHERE
BETWEEN ... AND
IN
LIKE
IS NULL
AND
OR
NOT

Overview of SQL commands

Required

SELECT
FROM

Condition

WHERE
BETWEEN ... AND
IN
LIKE
IS NULL
AND
OR
NOT

Sort

ORDER BY

Function

avg
sum
min
max
count

Overview of SQL commands

Required

SELECT
FROM

Condition

WHERE
BETWEEN ... AND
IN
LIKE
IS NULL
AND
OR
NOT

Sort

ORDER BY

Function

avg
sum
min
max
count

Aggregate

GROUP BY

R interface

Querying from R

The RODB package

Connection between R and Access

- RODB can also connect R to Microsoft SQL Server

The RODBC package

Connection between R and Access

- RODBC can also connect R to Microsoft SQL Server
- Other R packages can connect to LibreOffice Base, MariaDB, MySQL, Oracle, PostgreSQL, SQLite, etc.

The RODBC package

Connection between R and Access

- RODBC can also connect R to Microsoft SQL Server
- Other R packages can connect to LibreOffice Base, MariaDB, MySQL, Oracle, PostgreSQL, SQLite, etc.

32 vs. 64 bits

If you have 32-bit Access, then you should connect from 32-bit R

The RODBC package

R functions

`odbcConnectAccess2007` Establish connection

`sqlTables` List all tables

`sqlQuery` Send query

`odbcClose` Close connection

The RODBC package

```
# Start R (32-bit)
# install.packages("RODBC")

library(RODBC)

con <- odbcConnectAccess2007("c:/database/onetable.accdb")

sqlTables(con)

sqlQuery(con, "SELECT * FROM catch")

odbcClose(con)
```

We create a “connection” object and call it `con`