SQL for astro

Colin Navin

what is SQL?

- Structured Query Language – the standard language used to manipulate relational databases
- the acronym is not really accurate - it is not only used to query databases but can also create and edit them as well as create queries, forms, and reports

available software

- Commercial
  - Oracle
  - SAP Sybase
  - Microsoft SQL Server
- Free
  - Microsoft SQL Server Express
  - mySQL
  - PostgreSQL
  - SQLite (our weapon of choice for today)

overview

- relational database basics
- basic SQL syntax and usage
- online (astro) databases
- embedded SQL (in Python)
- one hour exam

relational database basics (1)

- database is a computerised record keeping system – a collection of data files
- users need to be able to add and delete files, insert data into files, retrieve data from files, change data in files, delete data from files

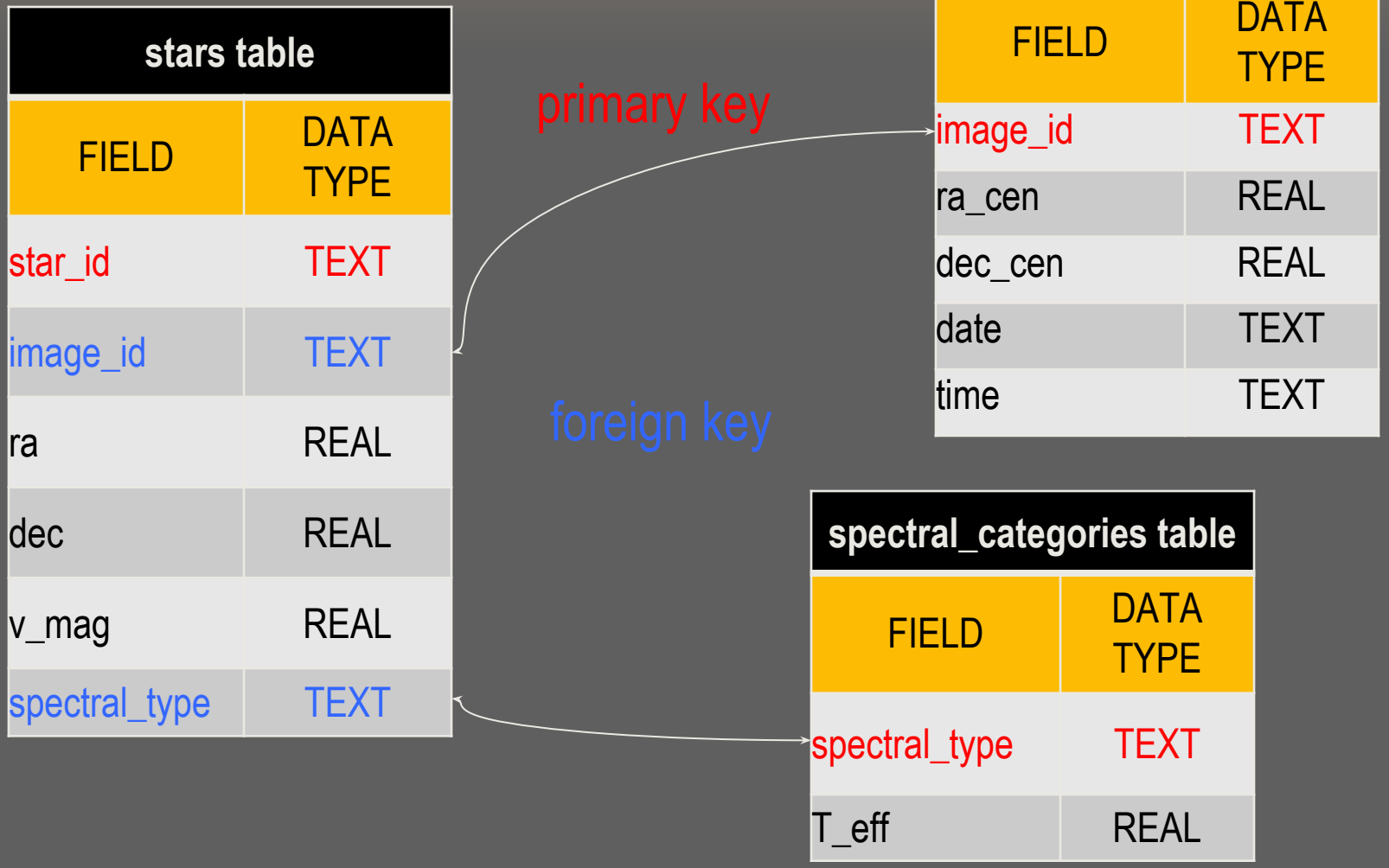
relational database basics (2)

- in a **relational** database, data is organised in one or more tables
- tables have columns (fields) and rows (records)
- columns have names and data types
- rows contain records or data for the columns
- Note: the physical files are (almost certainly) not stored as tables but the database system is designed to make it appear that way to the user

relational database basics (3)

- each table must have a primary key – a field that uniquely identifies each record in the table
- tables may have relationships with other tables
- relationships are managed by foreign keys – a field that contains the primary key of another table

test.db


stars table
FIELD
DATA TYPE
star_id
TEXT
image_id
TEXT
ra
REAL
dec
REAL
v_mag
REAL
spectral_type
TEXT
primary key
foreign key
images table
FIELD
DATA TYPE
image_id
TEXT
ra_cen
REAL
dec_cen
REAL
date
TEXT
time
TEXT
spectral_categories table
FIELD
DATA TYPE
spectral_type
TEXT
T_eff
REAL

test.db - stars table

star_id	image_id	ra	dec	v_mag	spectral_type
HIP29807	146-34-01	146.12	-34.32	6.78	O
HIP30973	146-33-01	146.33	-32.91	7.98	G
HD35294	145-33-01	144.82	-33.21	11.56	O
HD269324	145-33-01	145.22	-33.21	12.67	M
HD269334	145-33-01	145.28	-33.17	12.68	M
HD49091	146-34-01	145.82	-34.34	9.99	M
HD49126	146-34-01	145.82	-34.41	13.01	M
HD169334	146-34-01	146.5	-33.73	12.81	B
1252BT14	146-33-01	145.9	-33.24	10.02	M
1252BT12	146-34-01	146.26	-34.45	12.98	M
1252BT18	147-34-01	146.66	-33.84	8.42	F
HD48765	147-34-01	146.76	-33.56	8.45	M
HD24670	147-34-01	146.62	-34.34	9.07	M

test.db - images and spectral_categories tables

image_id	ra_cen	dec_cen	date	time
145-33-01	145.2	-33.1	20130321	1117
146-33-01	146.2	-33.1	20130321	1152
147-33-01	147.2	-33.1	20130321	1229
145-34-01	145.2	-34.1	20130322	1312
146-34-01	146.2	-34.1	20130322	1117
147-34-01	147.2	-34.1	20130322	1117

spectral_type	T_eff
O	50000
B	28000
A	10000
F	7500
G	6000
K	5000
M	3500

basic SQL syntax and usage

- [SQLite metacommands]
- SELECT statements to query the database
- INSERT, DELETE and UPDATE statements to change the data

SQLite metacommands

- .help
- .tables
- .schema
- .show
- .mode list|columns|csv
- .header on|off

SELECT statements

- basic syntax:

```
SELECT “column_a”, ..., “column_n”
FROM “table_a”
WHERE “condition”;
```

INSERT statements

- basic syntax:

```
INSERT INTO “table_a”
(“column_a”, ..., “column_n”)
VALUES (“value_a”, ...,“value_n”);
```

UPDATE statements

- basic syntax:

```
UPDATE ”table_a”
SET “column_a” = “value_a”, ..., “column_n”=
    “value_n”
WHERE “condition”;
```

DELETE statements

- basic syntax:

```
DELETE FROM “table_a”
WHERE “condition”;
```

online (astro) databases

- NASA/IPAC Infrared Science Archive (IRSA)
  - http://irsa.ipac.caltech.edu/index.html
- Sloan Digital Sky Survey
  - http://cas.sdss.org/astro/en/

embedded SQL (python)

- you can include SQL statements in many programming languages
- in particular for today you can embed SQLite3 statements in Python
- have a look at and run SQLtest.py