

QSTEP SQL Masterclass

James Tripp, Senior Academic Technologist, CIM (Warwick)

1st December 2020/ Microsoft Teams



James Tripp



WARWICK

- Background in Psychology (BSc, PhD)
- [Senior Academic Technologist](#) at [CIM](#) (Centre for Interdisciplinary Methodologies)
- SQL?
 - Large datasets
 - Analysis
 - System Administration

SQL?

- Structure Query Language. A domain specific language – designed for a specific purpose
- [Relational databases](#)
- Give the database the query. The database then gets the data in an optimal way.

See:

<https://en.wikipedia.org/wiki/SQL>

https://www.w3schools.com/sql/sql_intro.asp

Our Data

Table name: world_indicators

Row	country	country code	electricity	area
	United Kingdom	GBR	100	241930

Column

world_borders

iso3	population	geography
GBR	60244834	01060000 ...
...

E.g.,

```
SELECT column FROM table;
```

```
SELECT country FROM world_indicators;
```

```
SELECT population FROM world_borders;
```

Why use a database?

- Data storage
- Efficient data querying via SQL
- A little complicated...

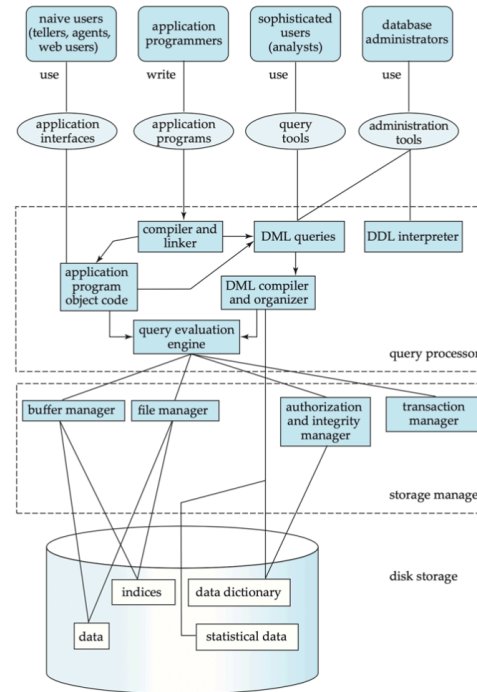


Figure 1.5 System structure.

Why use a database?

- Data storage
- Efficient data querying via SQL
- A little complicated...

WARWICK

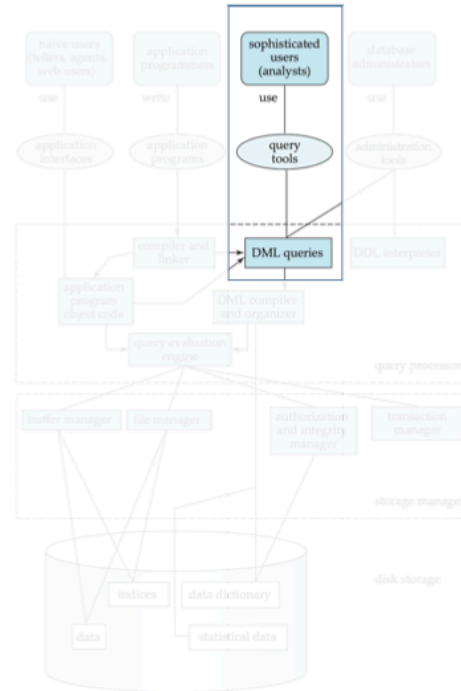


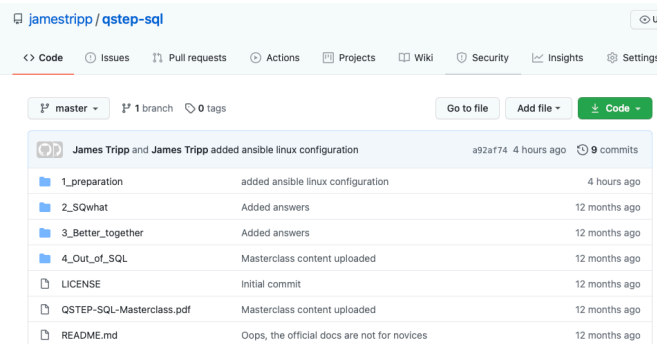
Figure 1.5 System structure.

Our brief

- We are data analysts/data scientists
- Question: What is the relationship between two world indicators across countries?
- Our data is in a table called world_indicators in the qstep database
- An additional table called world_borders contains the geospatial borders of countries. This may help with visualization work

Materials

- Located in Github
- <https://github.com/jamestripp/qstep-sql>



jamestripp / qstep-sql

< Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 1 branch 0 tags Go to file Add file Code

James Tripp and James Tripp added ansible linux configuration a92af74 4 hours ago 9 commits

1_preparation	added ansible linux configuration	4 hours ago
2_SQLwhat	Added answers	12 months ago
3_Better_together	Added answers	12 months ago
4_Out_of_SQL	Masterclass content uploaded	12 months ago
LICENSE	Initial commit	12 months ago
QSTEP-SQL-Masterclass.pdf	Masterclass content uploaded	12 months ago
README.md	Oops, the official docs are not for novices	12 months ago

Today

- Introduction (this presentation)
- Local installation of database and data (optional)
- SQWhat?
 - **Basic SQL introduction**
- Better together
 - **Aggregate function for data and joining tables**
- Out of SQL
 - **Taking data from the database**
 - **Using R for data analysis and visualisation**