

SQL

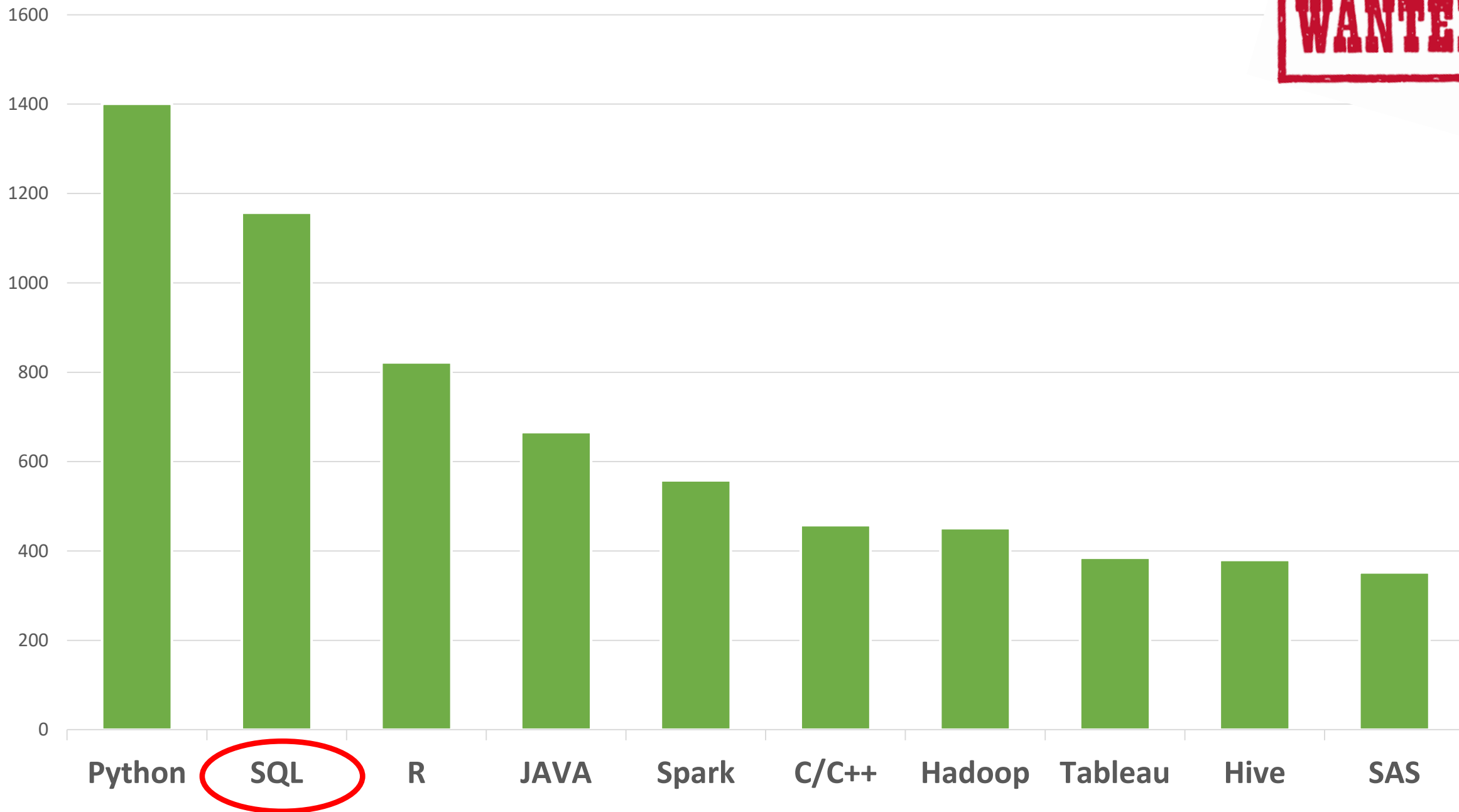
Dr. Villanes

Introduction

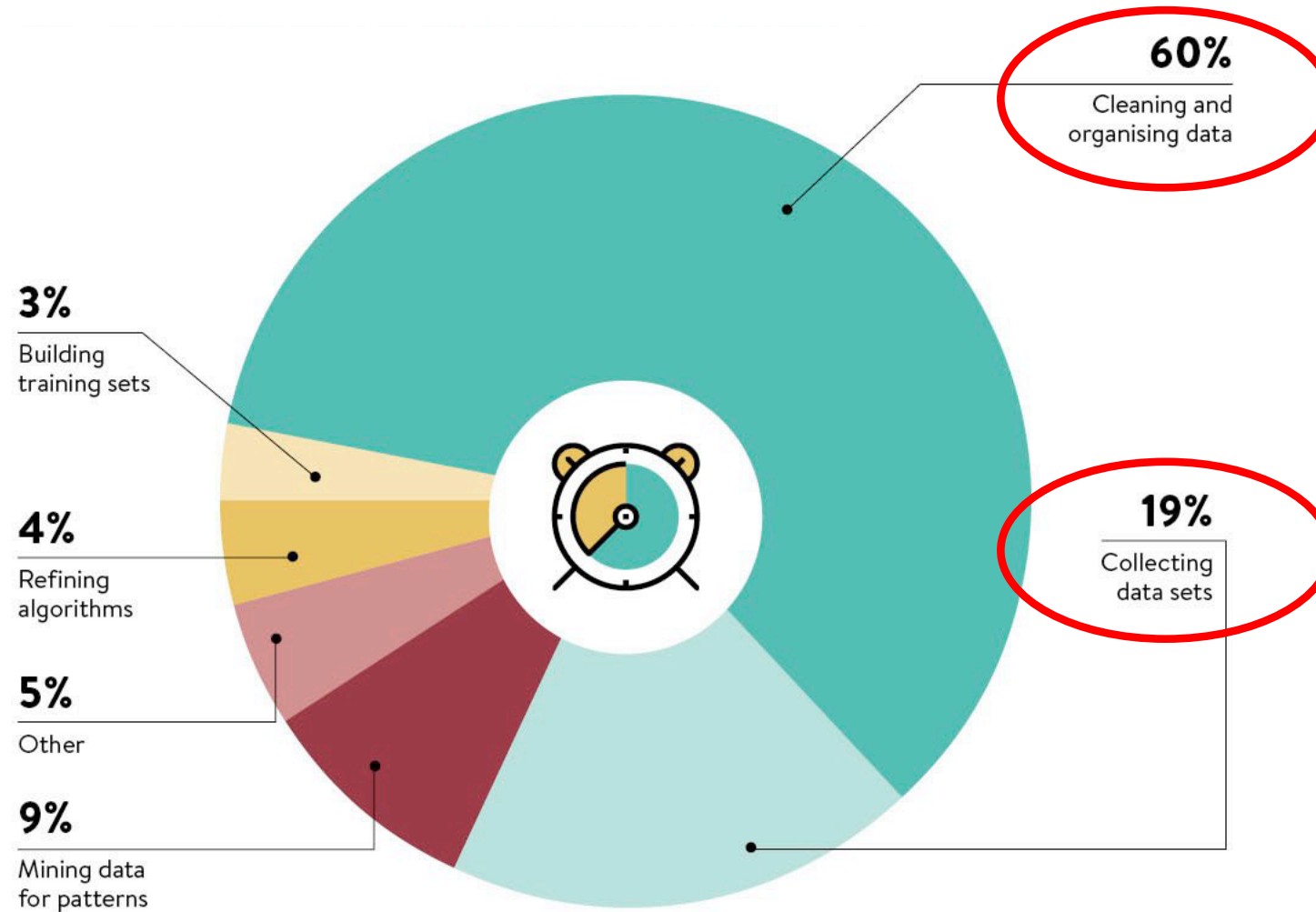


2,583 jobs

related to data science



What data scientists spend the most time doing



Source: CrowdFlower 2016

Programming tools available for data extraction/cleaning/organizing...



SQL



uses SQL-like syntax!

What is SQL?

- *Structured Query Language* (SQL) is a **standardized** language originally designed as a relational database query tool.
- SQL is currently used in **many** software products to retrieve and update data.



Why is SQL a **standardized** language?

- SQL became a **standard** of the American National Standards Institute (ANSI) in 1986, and of the International Organization for Standardization (ISO) in 1987.
- Since then, the standard has been **revised** to include a larger set of features.
- Despite the existence of such standards, most SQL code is **not completely portable** among different database systems without adjustments.



How is SQL currently being used in **many** software products?

- Although most database systems use SQL, **most of them also have their own additional proprietary extensions** that are usually only used on their system.
- However, the standard SQL commands such as "Select", "Insert", "Update", "Delete", "Create", and "Drop" can be used to accomplish almost everything that one needs to do with a database.

SQL is a standard... BUT...

Microsoft SQL Server

Oracle

Extension

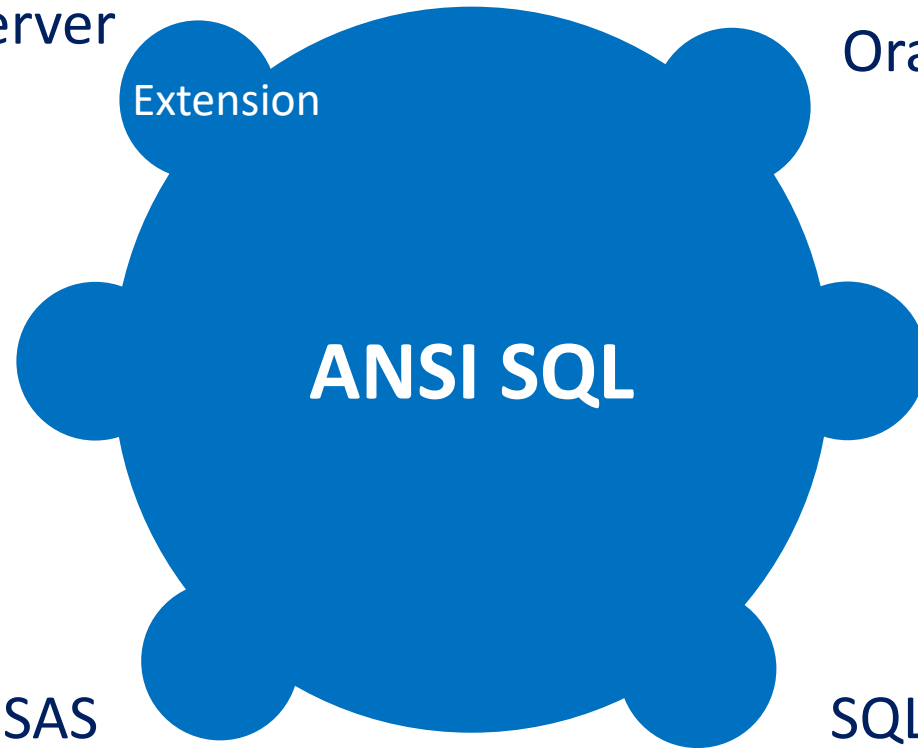
MySQL

ANSI SQL

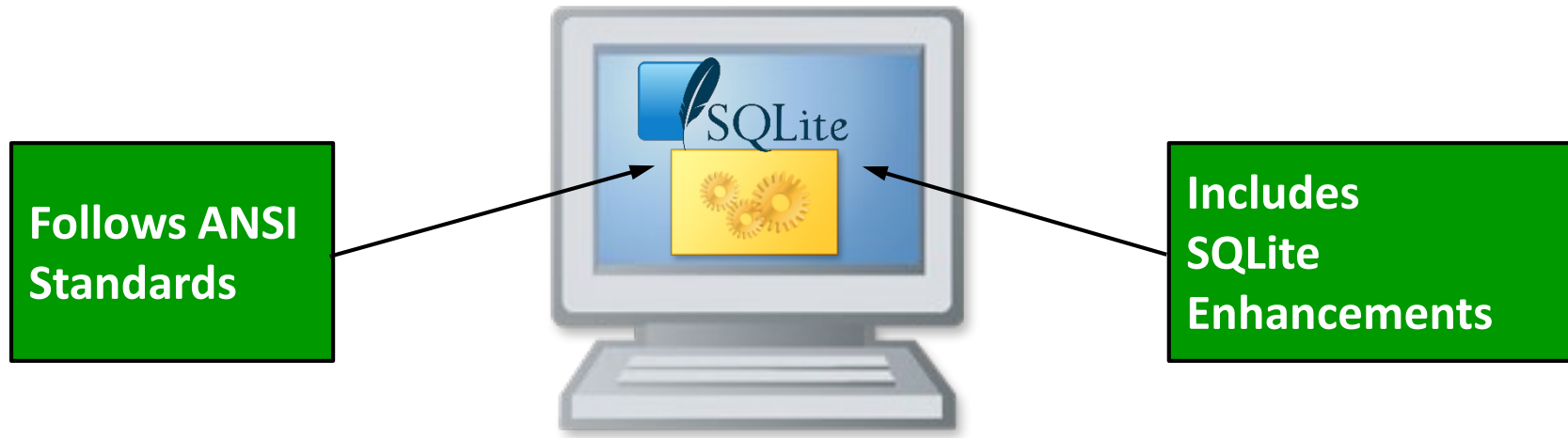
PostgreSQL

SAS

SQLite



For example...



*Small. Fast. Reliable.
Choose any three.*

[Home](#) [About](#) [Documentation](#) [Download](#) [License](#) [Support](#) [Purchase](#)

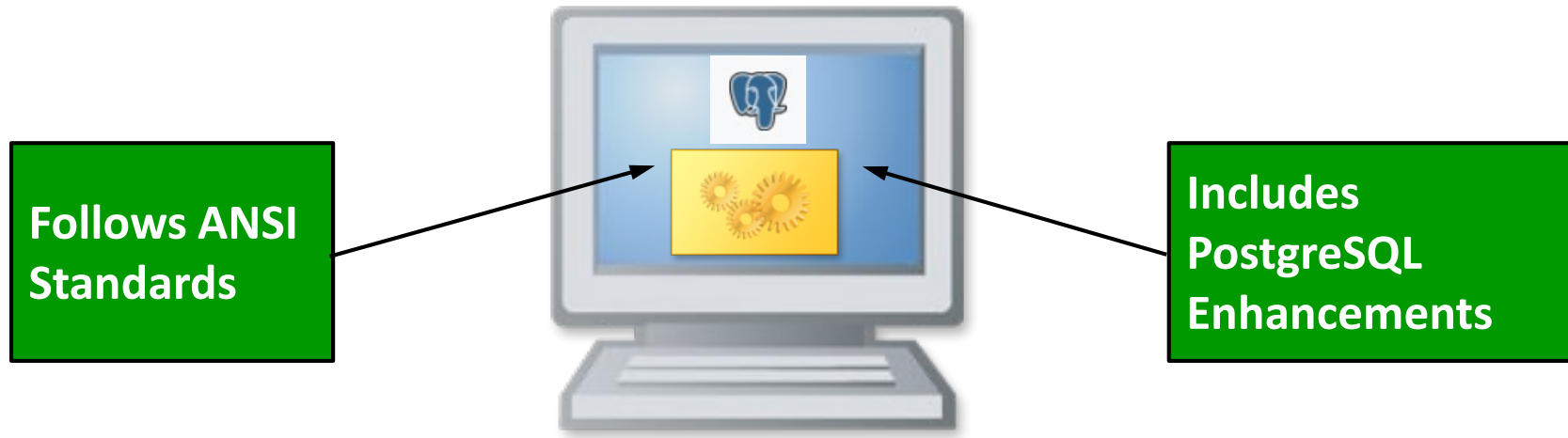
[Search](#)

SQL As Understood By SQLite

SQLite understands most of the standard SQL language. But it does omit some features while at the same time adding a few features of its own. This document attempts to describe precisely what parts of the SQL language SQLite does and does not support. A list of [SQL keywords](#) is also provided. The SQL language syntax is described by [syntax diagrams](#).

<https://www.sqlite.org/lang.html>

What about PostgreSQL?

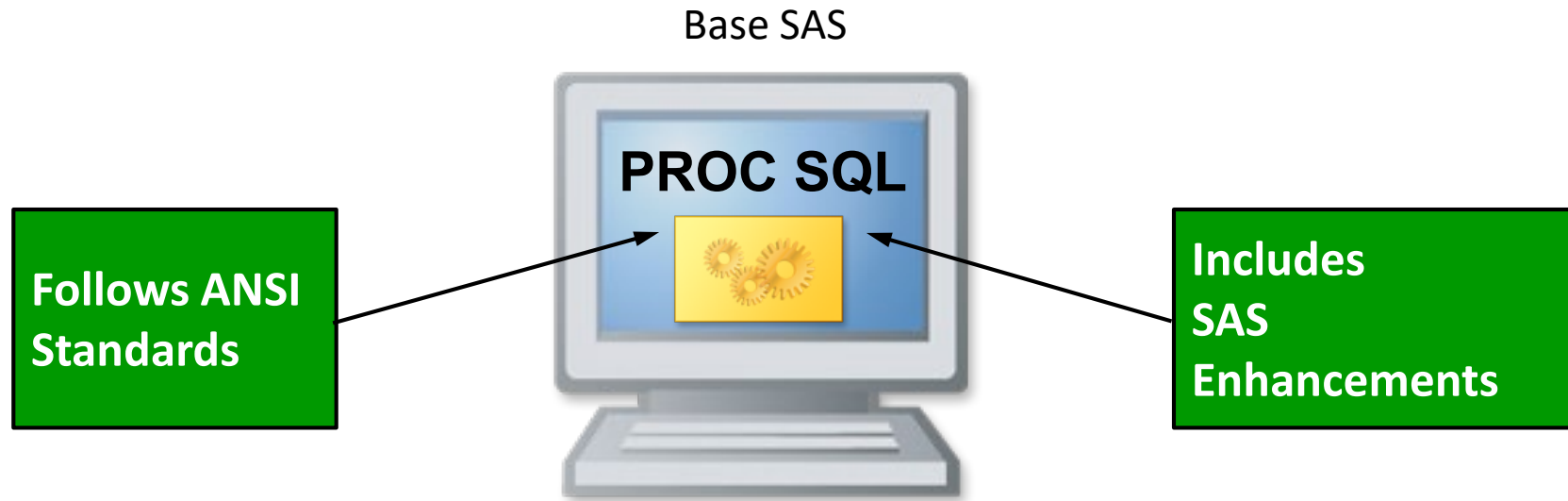


1. What Is PostgreSQL?

PostgreSQL is an object-relational database management system (ORDBMS) based on **POSTGRES, Version 4.2**, developed at the University of California at Berkeley Computer Science Department. POSTGRES pioneered many concepts that only became available in some commercial database systems much later.

PostgreSQL is an open-source descendant of this original Berkeley code. It supports a large part of the SQL standard and offers many modern features.

What about SAS?



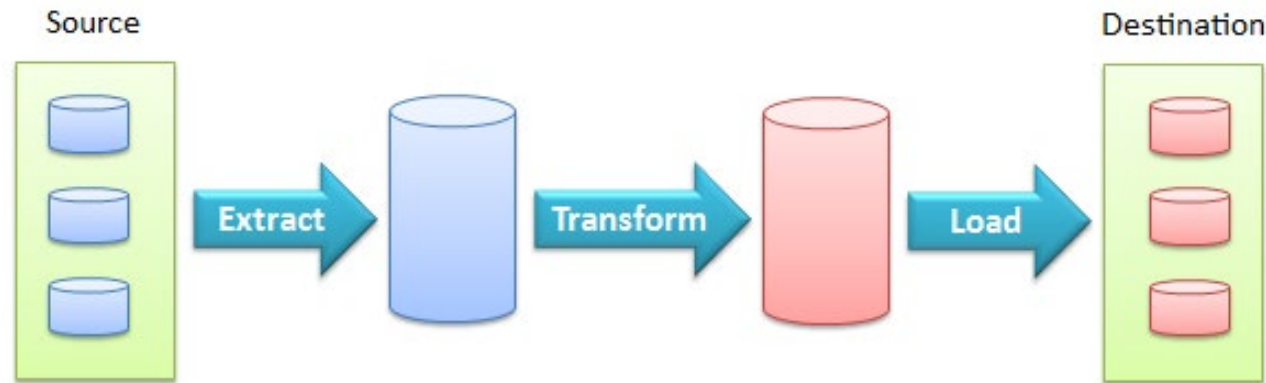
PROC SQL and the ANSI Standard:

<http://support.sas.com/documentation/cdl/en/sqlproc/63043/HTML/default/viewer.htm#p0corah7t43hy7n1r7efplde5c8z.htm>

Where is SQL mostly used?

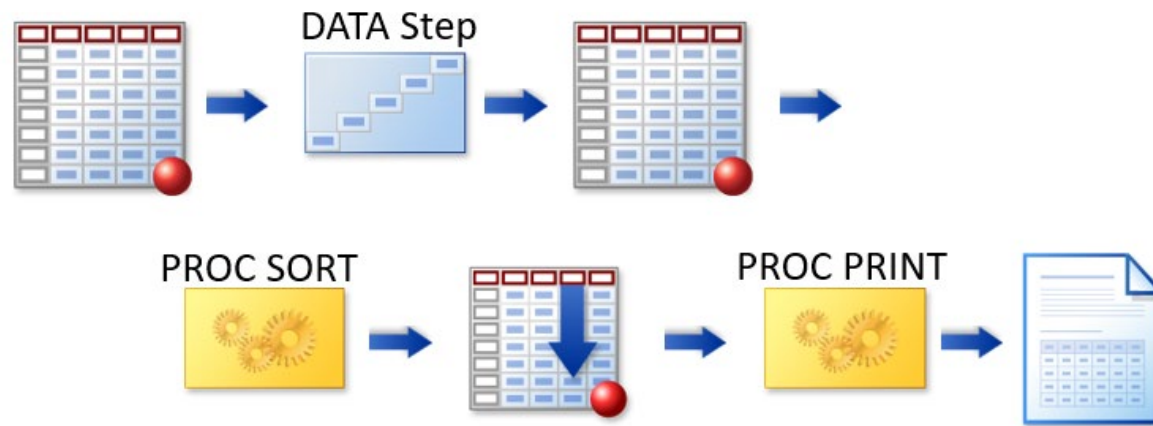
ETL: extract, transform, and load

Challenge: Data resides in multiple locations and in many formats.



SQL is very useful for the ETL process

Why is SQL so awesome?



Why is SQL so awesome?

... and because they will ask you SQL questions in interviews!