

A friendly introduction to SQL

Jorge Hopkins (JHOP)



Agenda

- (Brief) Intro to relational databases and SQL
- Structured Query Language (SQL)
- Data Ingestion & Automation

(Brief) Intro to relational databases and SQL

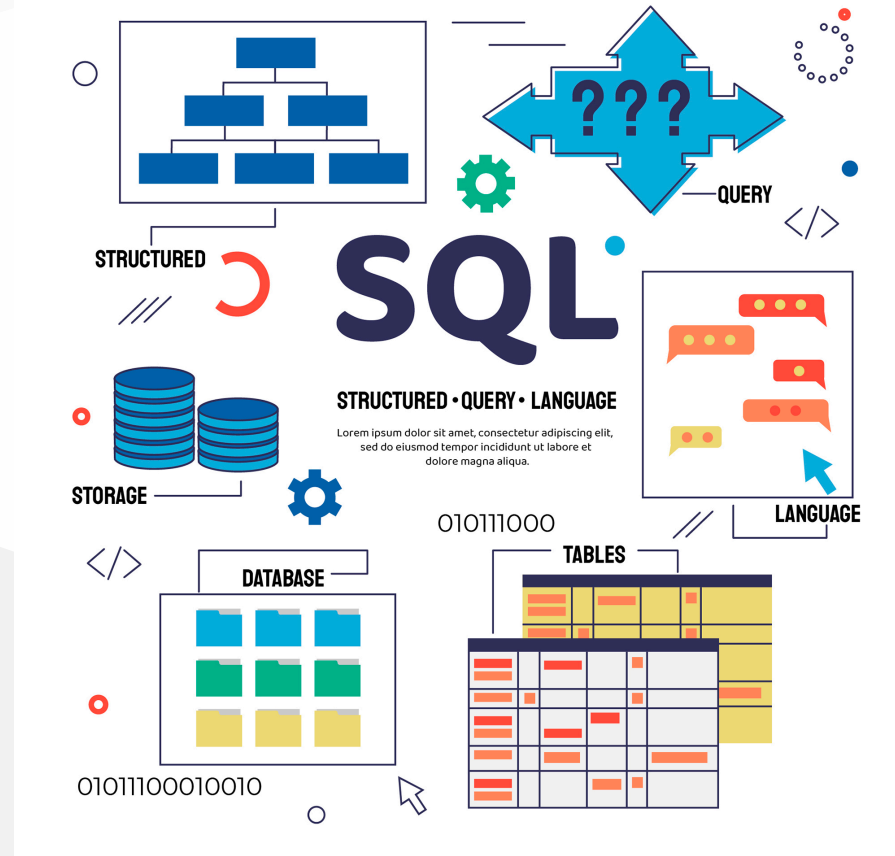
*"A relational database (RDB) is a way of structuring information in tables, rows, and columns. An RDB has the ability to establish links—or **relationships**—between information by **joining** tables, which makes it easy to understand and gain insights about the relationship between various data points" -- Google Cloud*

(Brief) Intro to relational databases and SQL

What do you need?

- DB/Server
- Schema
- Tables
- Data!

Structured Query Language (SQL)



Structured Query Language (SQL)

Querying and Filtering

```
SELECT *  
FROM <_SCHEMA_>.<_TABLE_>  
WHERE <_CONDITION_>
```

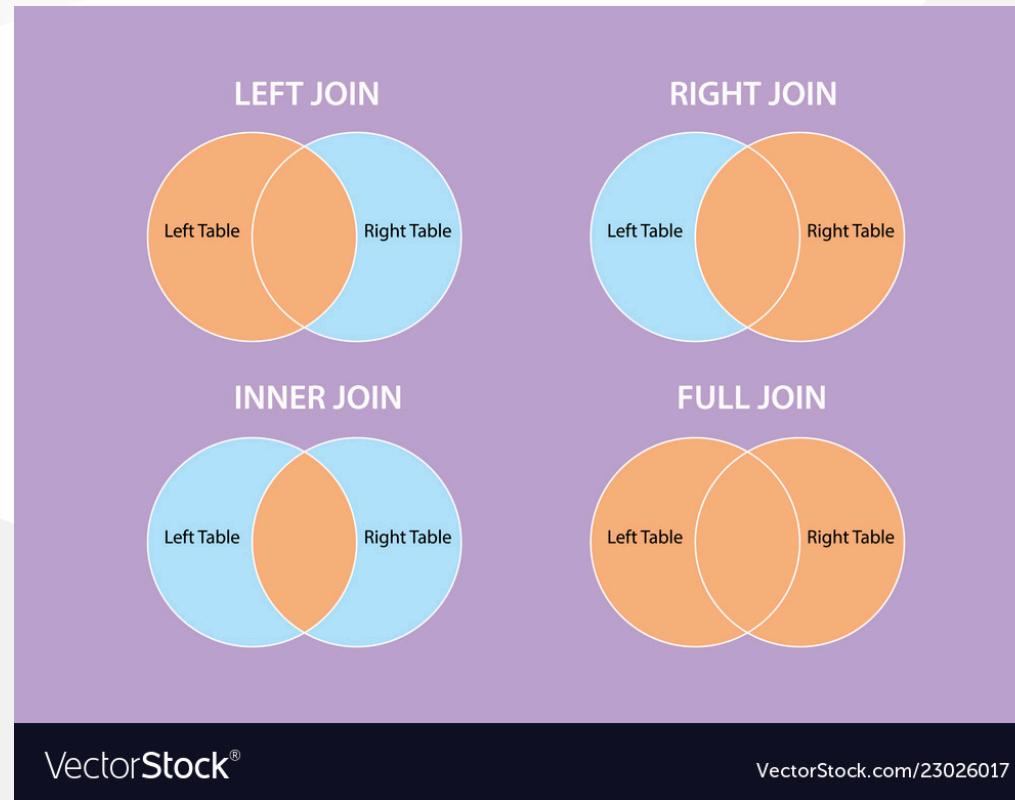
Structured Query Language (SQL)

Grouping

```
SELECT <GROUP_FIELD>, COUNT(<FIELD>), SUM(<FIELD>), AVG(<FIELD>), MAX(<FIELD>)  
FROM <_SCHEMA_>.<_TABLE_>  
WHERE <_CONDITION_>  
GROUP BY <GROUP_FIELD>  
HAVING <_CONDITION_>
```

Structured Query Language (SQL)

Joins



Structured Query Language (SQL)

Subqueries and CTE's

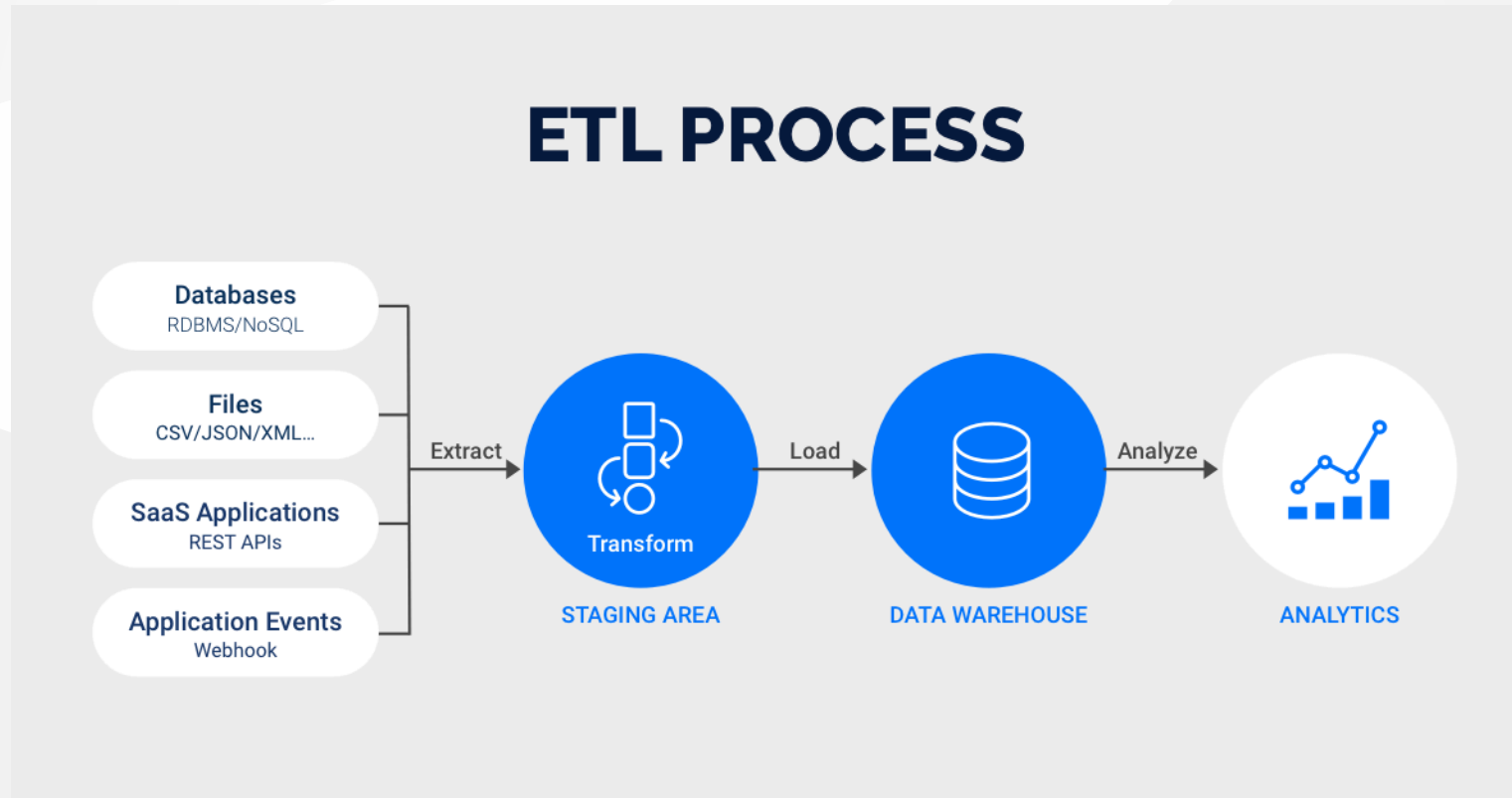
```
WITH CALCULATION_CTE AS (  
    SELECT <GROUP_FIELD>, COUNT(<FIELD>), SUM(<FIELD>), AVG(<FIELD>), MAX(<FIELD>)  
    FROM <_SCHEMA_>.<_TABLE_1>  
    WHERE <_CONDITION_>  
    GROUP BY  
    )  
  
SELECT *  
FROM <_SCHEMA_>.<_TABLE_2_> T  
JOIN CALCULATION_CTE C ON T.KEY = C.KEY
```

Structured Query Language (SQL)

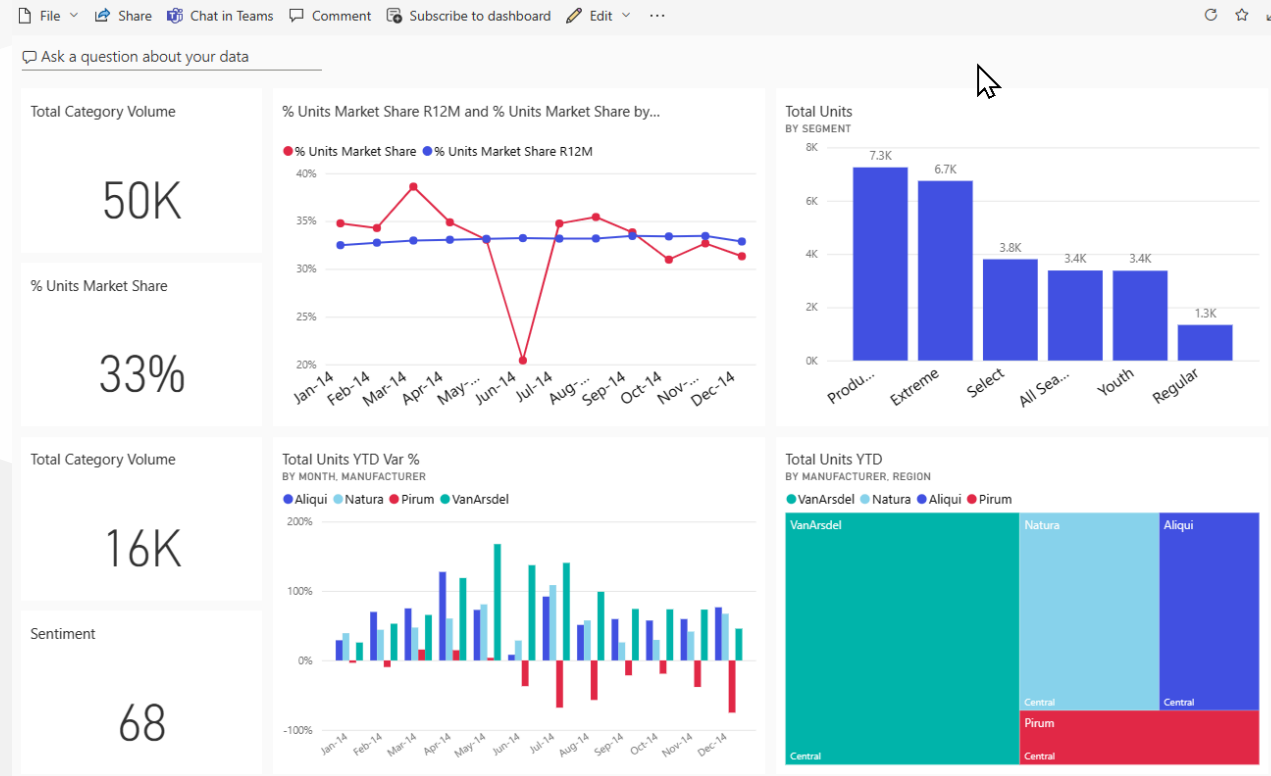
Window Functions

```
LAG(<FIELD_1>, 1) OVER (ORDER BY <FIELD_2>),  
LEAD(<FIELD_1>, 1) OVER (ORDER BY <FIELD_2>),  
RANK() OVER (PARTITION BY <FIELD_1> ORDER BY <FIELD_2>),  
NTILE(<NUMBER_OF_BUCKETS>) OVER  
    (PARTITION BY <FIELD_1> ORDER BY <FIELD_2>)  
    AS PERCENTILE,
```

Data Ingestion & Automation



Data Ingestion & Automation



Thank you!

Jorge Hopkins (JHOP)

