dimajix

df.select(columns...)

df.alias(table_alias)

df.join(other_table, join_condition)

df.where(condition)

df.groupBy(grouping)
.agg(columns)

df.where(condition)

df.orderBy(columns...)

df.limit(n)

SELECT
    columns...
FROM table AS table_alias
JOIN other_table AS other_table_alias
    ON join_condition
WHERE condition
GROUP BY grouping
HAVING condition
ORDER BY columns...
LIMIT n
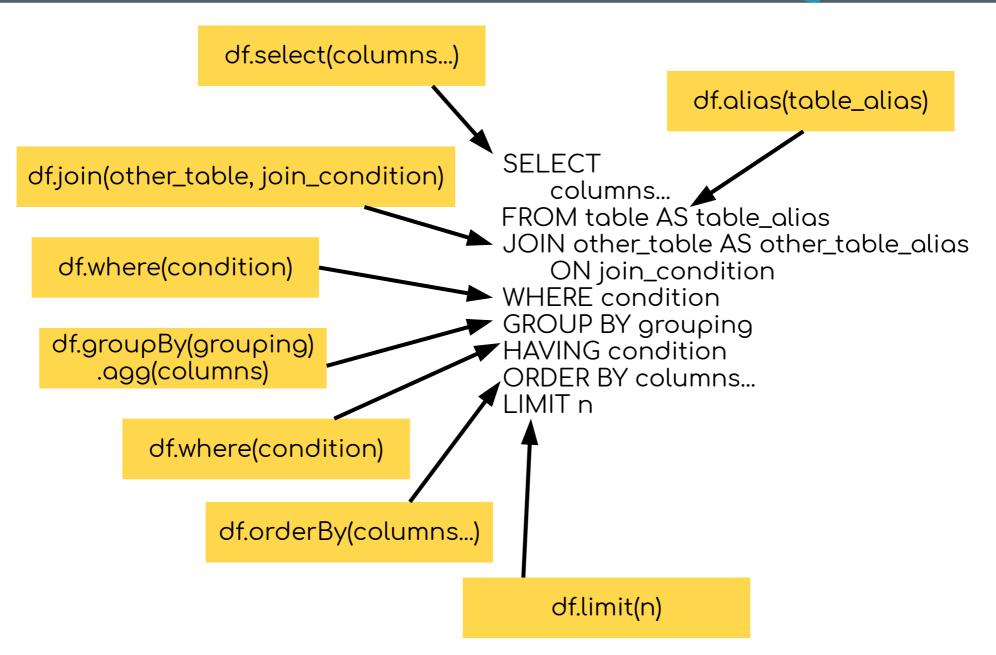
# SQL

```
SELECT
    columns...
FROM table AS table_alias
JOIN other_table AS other_table_alias
    ON join_condition
WHERE condition
GROUP BY grouping
HAVING condition
ORDER BY columns...
LIMIT n
```

# Spark

```
table.alias(„table_alias")
    .select(columns)
    .join(other_table.alias(„other_table_alias),
        join_condition)
    .where(condition)
    .groupBy(grouping)
    .agg(columns)
    .where(condition)
    .orderBy(columns)
    .limit(n)
```