SQL under pressure

SQL / Learn and Be Curious / Earn Trust

Full guide: 11h 50m | Core route: 4h | Read this map first, then use the final recall strip.

BEFORE STUDY Population and joins → ranking/windows → time/sequences → JSON/history/reconciliation → dialect and plan → learning/trust stories.

01 / Declare the query contract

Output grain, columns, population, time zone/cutoff, ties, nulls, duplicates and order. Identify join cardinality before typing.

Use small logical stages. Walk a tiny fixture, then explain scans, sorts, movement and output size. State the SQL dialect.

Logical order reminder: FROM/JOIN → WHERE → GROUP/HAVING → windows → final projection/order. Filter window results in an outer stage or supported QUALIFY.

02 / Joins, nulls and aggregation

Preserve unmatched parents deliberately. A right-side WHERE condition can undo a left join. COUNT(*) counts the placeholder; COUNT(child_key) counts matches.

Use NOT EXISTS for a null-safe anti-match. Preaggregate independent child facts before joining. SUM(DISTINCT amount) is not a fanout repair.

Filter rows before grouping and groups after aggregation. SUM(CASE...1 ELSE 0 END) counts flags; COUNT(...ELSE 0) also counts zeros.

03 / Ranking, versions and frames

ROW_NUMBER: exact row count with a total tie order. RANK: ties with gaps. DENSE_RANK: ties without gaps. Match the requested tie policy.

Select authoritative latest versions, then remove tombstones. Conflicting equal versions need a rule; arbitrary ordering is not resolution.

Declare PARTITION BY, ORDER BY and frame. ROWS counts rows; RANGE groups by ordering values/peers. LAST_VALUE needs a frame reaching the intended end.

Window order does not guarantee final output order. Include required history before filtering display dates.

04 / Time, sessions and product metrics

Use half-open time ranges when appropriate. A three-row window equals three days only with one complete row per day.

Sessions: new ID when gap from previous event exceeds the threshold. Streaks: deduplicate entity-day first, then group consecutive dates.

Funnels require ordered qualifying events within the stated window. Retention excludes unobservable future periods; missing future data is not churn.

Rolling distinct users cannot be obtained by summing daily distinct counts.

05 / History, reconciliation and shape

As-of joins: key plus start<=event_time<end, with nonoverlapping valid intervals. Separate missing row from stored NULL using a presence key.

Compare keys, values and multiplicities, not totals only. UNION deduplicates complete selected rows; UNION ALL preserves them.

Pivot via conditional aggregation at a declared grain. Median needs deliberate odd/even handling and null policy.

All-of requirements can use double NOT EXISTS; an empty requirement set usually matches all candidates unless business rules say otherwise.

06 / JSON, hierarchy and scale

Valid JSON does not prove field type. Distinguish absent, null and numeric text. Outer array expansion preserves parents with empty arrays when required.

Recursive traversal needs an anchor, step, cycle/visited rule and bounds. UNION may not stop cycles when depth/path changes each row.

Inspect actual plan evidence before tuning. Filter/project early when semantics permit; reduce fanout before sorting or shuffling.

SQLite practice syntax is not Redshift or Oracle syntax. Translate JSON, dates, frames, recursion and QUALIFY deliberately.

07 / Learn and Be Curious

Real knowledge gap → deliberate learning/experiment → evidence that changed your belief → applied decision → outcome and sharing.

Prepare a primary and backup. Explain what you initially misunderstood and what would disconfirm your current belief. Reading volume is not the result.

08 / Earn Trust and live correction

Name your own error or commitment gap, communicate verified versus unresolved scope, repair the consequence and show follow-through.

If the query is wrong, state the defect, walk the counterexample, correct the logic and update the conclusion. Do not quietly edit while defending the old claim.

Keep personal facts and numbers consistent. Credit partners; avoid blame and unsupported claims that trust was fully restored.

FINAL 5-MINUTE RECALL Write one join + window cold • test null/tie/zero/boundary • explain frame and grain • rehearse learning + trust incidents • correct a wrong answer aloud • ask about metric review practices.

ELI5: Say what the answer means, write the query, then try to break it.

Sources: SQLite window semantics | Redshift QUALIFY