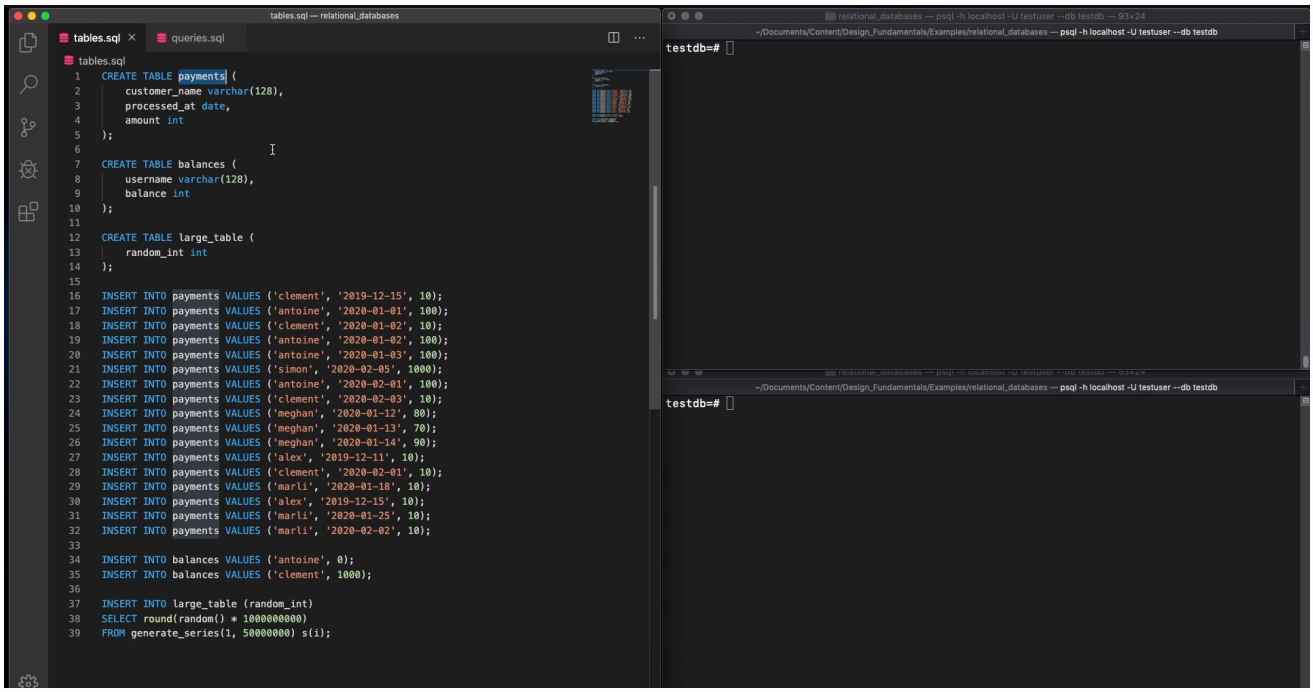
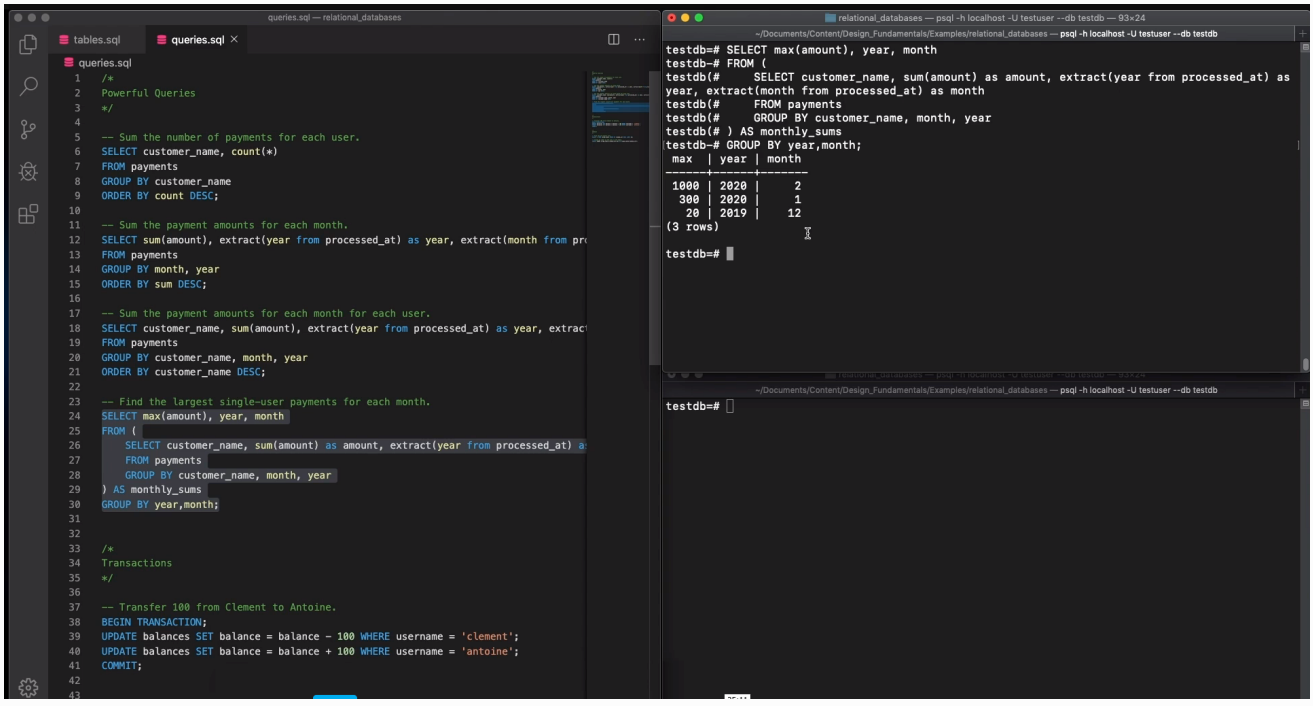
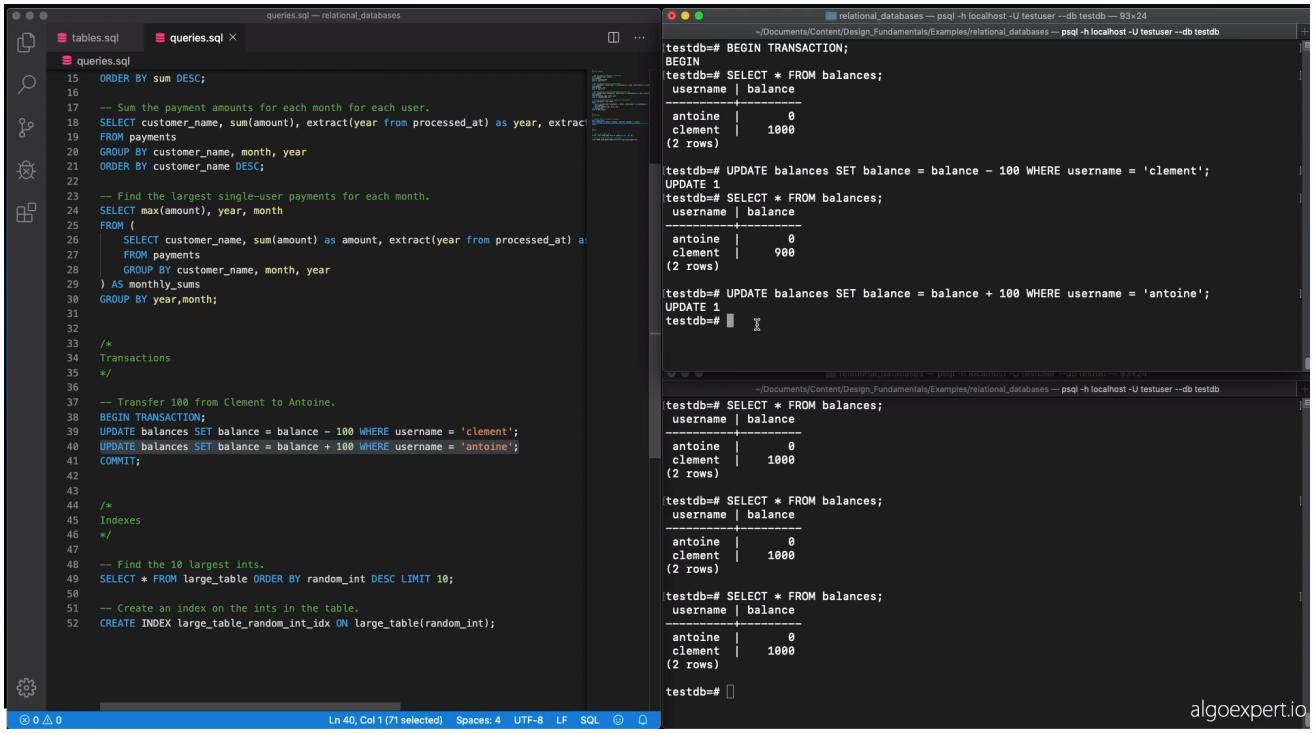
5/2/2021 - Relational Databases

ACID-

Database Index -







3 Prerequisites

Databases

Databases are programs that either use disk or memory to do 2 core things: **record** data and **query** data. In general, they are themselves servers that are long lived and interact with the rest of your application through network calls, with protocols on top of TCP or even HTTP.

Some databases only keep records in memory, and the users of such databases are aware of the fact that those records may be lost forever if the machine or process dies.

For the most part though, databases need persistence of those records, and thus cannot use memory. This means that you have to write your data to disk. Anything written to disk will remain through power loss or network partitions, so that's what is used to keep permanent records.

Since machines die often in a large scale system, special disk partitions or volumes are used by the database processes, and those volumes can get recovered even if the machine were to go down permanently.

Disk

Usually refers to either **HDD (hard-disk drive)** or **SSD (solid-state drive)**. Data written to disk will persist through power failures and general machine crashes. Disk is also referred to as **non-volatile storage**.

SSD is far faster than HDD (see latencies of accessing data from SSD and HDD) but also far more expensive from a financial point of view. Because of that, HDD will typically be used for data that's rarely accessed or updated, but that's stored for a long time, and SSD will be used for data that's frequently accessed and updated.

Memory

Short for **Random Access Memory (RAM)**. Data stored in memory will be lost when the process that has written that data dies.

Relational Database

A type of structured database in which data is stored following a tabular format; often supports powerful querying using SQL.

Non-Relational Database

In contrast with relational database (SQL databases), a type of database that is free of imposed, tabular-like structure. Non-relational databases are often referred to as NoSQL databases.

SQL

Structured Query Language. Relational databases can be used using a derivative of SQL such as PostgreSQL in the case of Postgres.

SQL Database

Any database that supports SQL. This term is often used synonymously with "Relational Database", though in practice, not *every* relational database supports SQL.

NoSQL Database

Any database that is not SQL-compatible is called NoSQL.

ACID Transaction

A type of database transaction that has four important properties:

- **Atomicity**: The operations that constitute the transaction will either all succeed or all fail. There is no in-between state.
- **Consistency**: The transaction cannot bring the database to an invalid state. After the transaction is committed or rolled back, the rules for each record will still apply, and all future transactions will see the effect of the transaction. Also named **Strong Consistency**.
- **Isolation**: The execution of multiple transactions concurrently will have the same effect as if they had been executed sequentially.
- **Durability**: Any committed transaction is written to non-volatile storage. It will not be undone by a crash, power loss, or network partition.

Database Index

A special auxiliary data structure that allows your database to perform certain queries much faster. Indexes can typically only exist to reference structured data, like data stored in relational databases. In practice, you create an index on one or multiple columns in your database to greatly speed up **read** queries that you run very often, with the downside of slightly longer **writes** to your database, since writes have to also take place in the relevant index.

Strong Consistency

Strong Consistency usually refers to the consistency of ACID transactions, as opposed to **Eventual Consistency**.

Eventual Consistency

A consistency model which is unlike **Strong Consistency**. In this model, reads might return a view of the system that is stale. An eventually consistent datastore will give guarantees that the state of the database will eventually reflect writes within a time period (could be 10 seconds, or minutes).

Postgres 

A relational database that uses a dialect of SQL called PostgreSQL. Provides ACID transactions.