

Relational algebra operators - selection

Relational algebra is the mathematical basis for performing queries against a relational database. Operations are performed against relations – resulting in relations. Because the result of relational algebra operation is a relation, operations can be stacked up against each other. More on this as we go forward.

For the following examples, we are going to be using a relation Employees with the following schema:

Employee (Emp_id:int, Emp_name:string, Emp_office:int)

Selection

Selection is represented by the symbol σ (sigma). With the projection operation, you can select sets of tuples (more easily thought of as rows) to be pulled out of the relation. Let's look at the following relation instance of Employee:

Employee		
Emp_id	Emp_name	Emp_office
1001	Bob	10
1002	Alice	11
1003	Sandy	10
1004	Larry	11
1005	Susan	11

If we apply the selection operation to this instance of Employee, we are saying we want a specific set of rows out of this instance. The statement:

$\sigma_{\text{Emp_id} > 1004}(\text{Employee})$ - which is also written as: $\sigma_{\text{Emp_id} > 1004} \text{Employee}$

is saying, select all of the rows whose Emp_id is greater than 1005 from this instance of the relation Employee. The result of this operation is a relation instance of Employee that looks like this:

Employee		
Emp_id	Emp_name	Emp_office
1005	Susan	11

The operation basically said "Build a new relation instance that consists of only the rows whose employee ids are greater than 1004 from the original relation instance of Employee. Note the schema of the resulting relation is the same as the original instance.

And & or operations are allowed in this as well. If the operation had looked like this:

$\sigma_{Emp_id > 1004 \vee Emp_name = Sandy}$ Employee

The resulting relation would look like:

Employee		
Emp_id	Emp_name	Emp_office
1003	Sandy	10
1005	Susan	11

Where the rows of the initial relation whose Emp_id > 1004 OR whose Emp_name = Sandy were selected.