

Materialization Strategies

Prof. Hasso Plattner

Tuple Reconstruction

Strategies:

- ☐ **Early** materialization
 - Decompress and decode data early and operate on strings
 - Create a row-wise data representation early on
- ☐ **Late** materialization
 - Operate on integers as long as possible
 - Operate on columns as long as possible

Example

Query:
List the number of Male
inhabitants per city in
Germany

```
SELECT city, COUNT(*)
FROM world_population
WHERE gender = "m"
AND country= "GER"
GROUP BY city
```

Table "world_population"

fname	lname	gender	country	city	birthday
Martin	Albrecht	m	GER	Berlin	08-05-1955
Michael	Berg	m	GER	Berlin	03-05-1970
Hanna	Schulze	f	GER	Bonn	04-04-1968
Ulrich	Schulze	m	GER	Bonn	10-20-1992
...	...	...	...	...	...

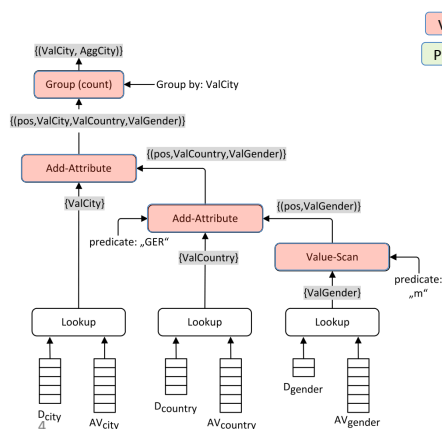
Dictionary encoded attribute vectors

53946	10435	0	68	357	15556
54368	25063	0	68	357	20882
30145	99645	1	68	443	20182
99312	99645	0	68	443	29147
fname	lname	gender	country	city	birthday

Comparison

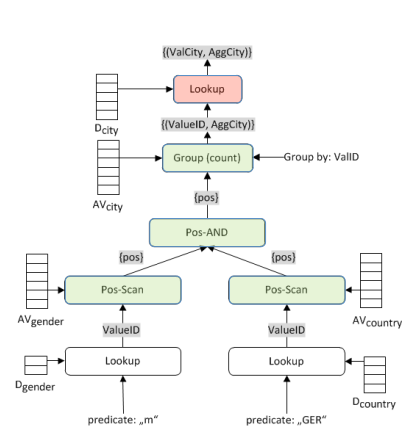
Early Materialization

Early dictionary lookup → working on
materialized values



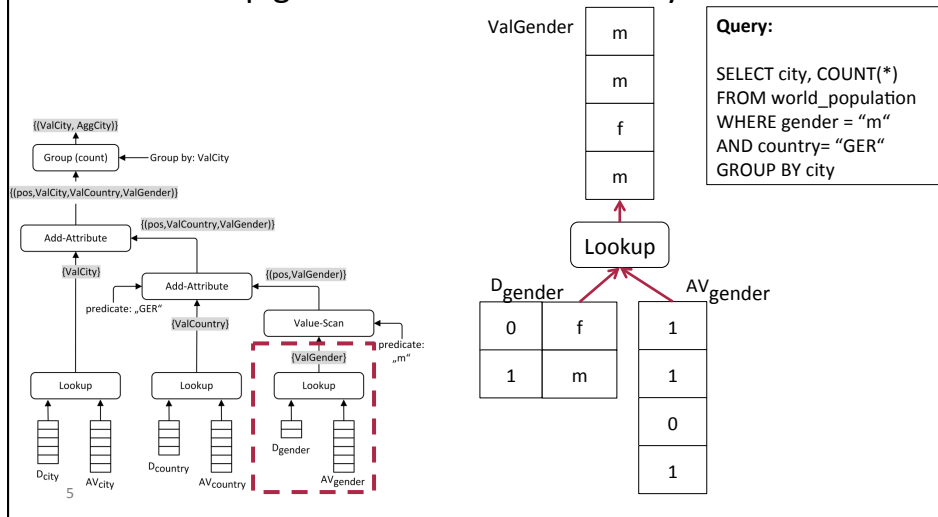
Late Materialization

Late dictionary lookup → working with
positional information



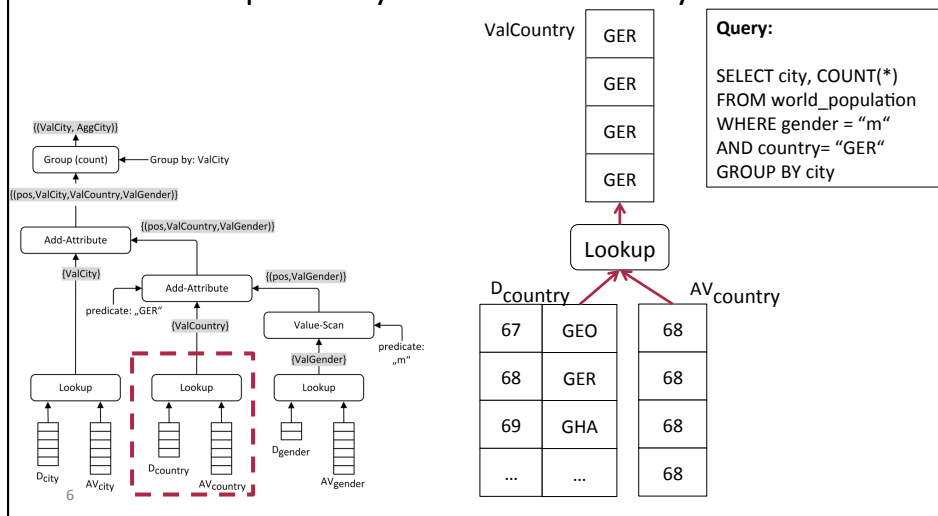
Early Materialization (Ia)

Look up gender values in dictionary



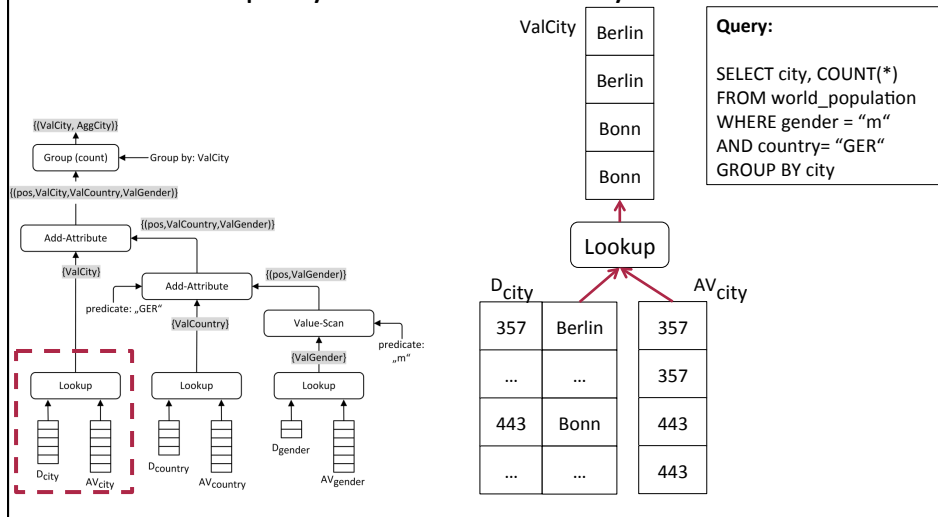
Early Materialization (Ib)

Look up country values in dictionary



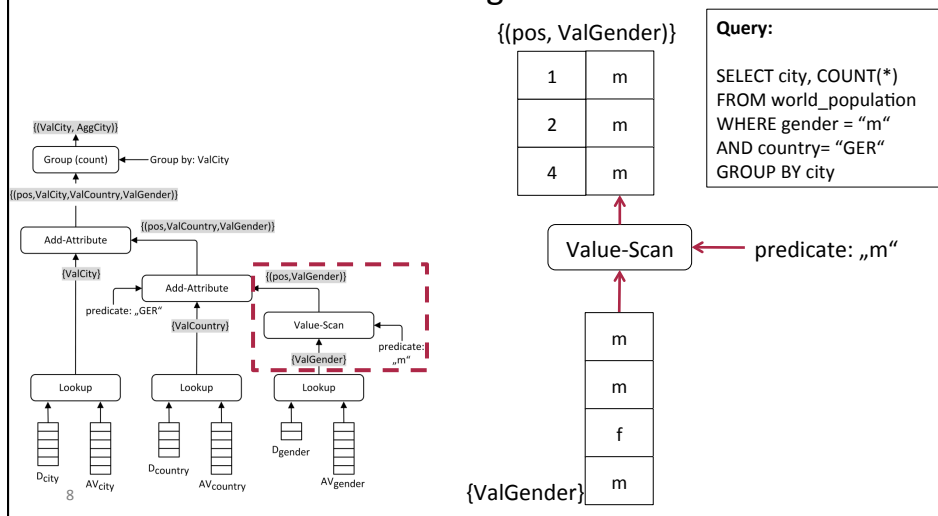
Early Materialization (1c)

Look up city values in dictionary



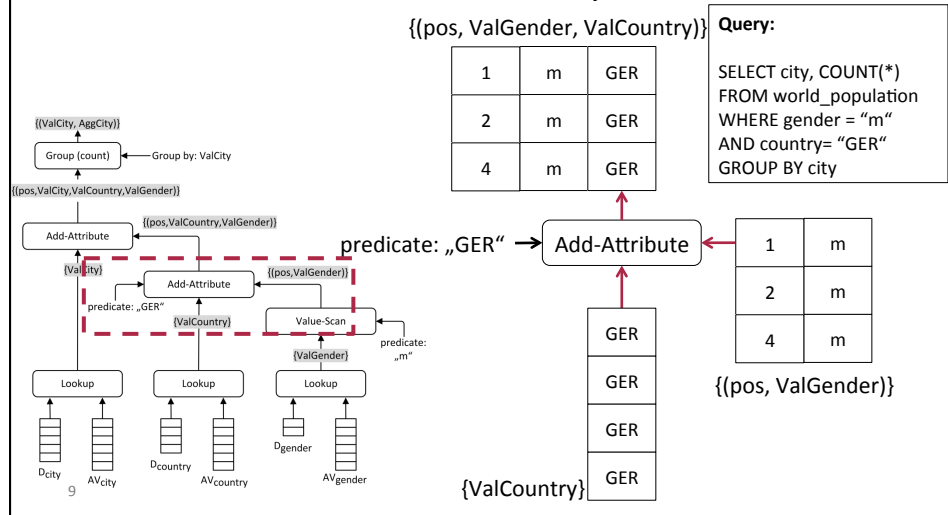
Early Materialization (2)

Scan for constraint on gender



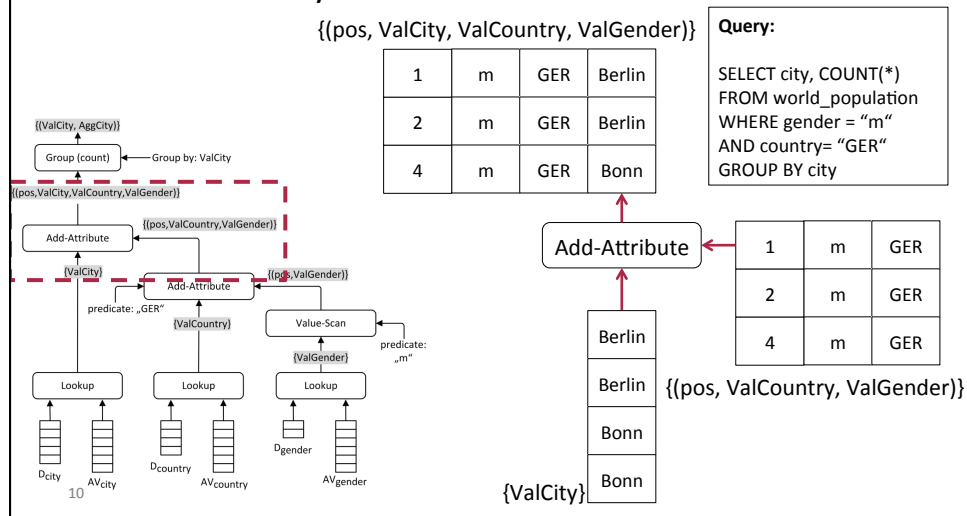
Early Materialization (3)

Scan for constraint on country and add to result

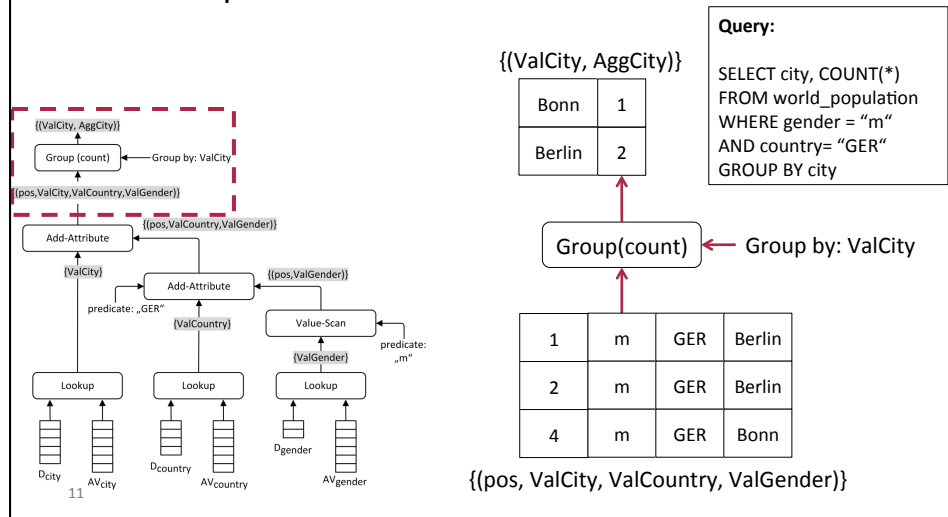


Early Materialization (4)

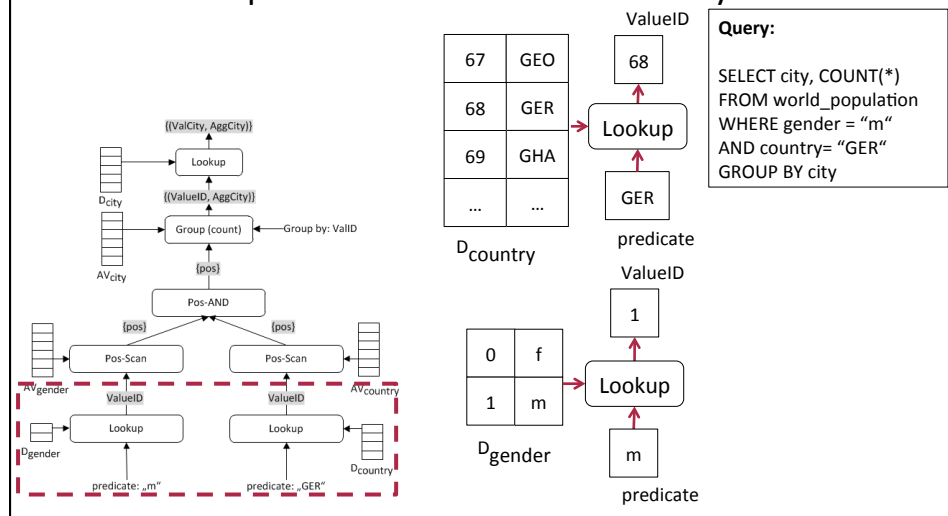
Add the city attribute vector to the result



Group and count

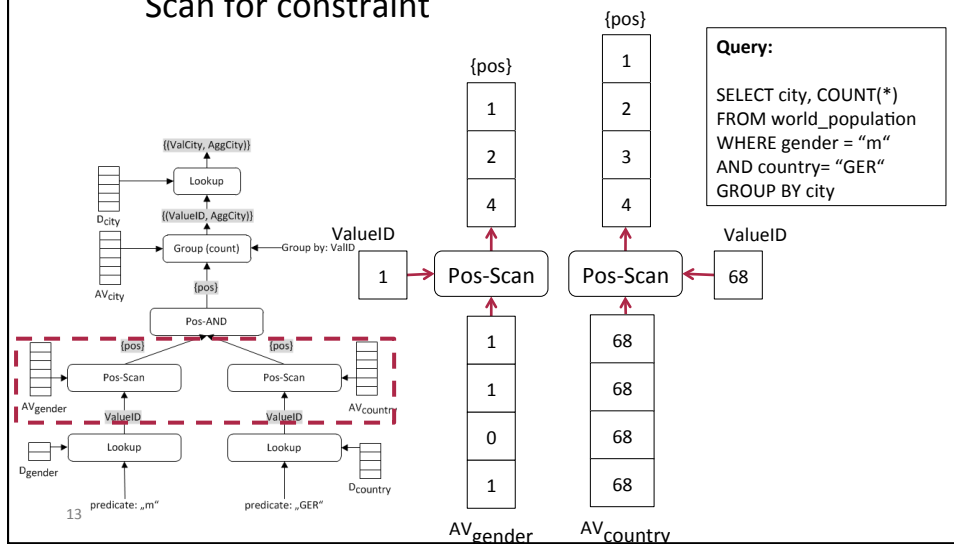


Look up constraint values in dictionary



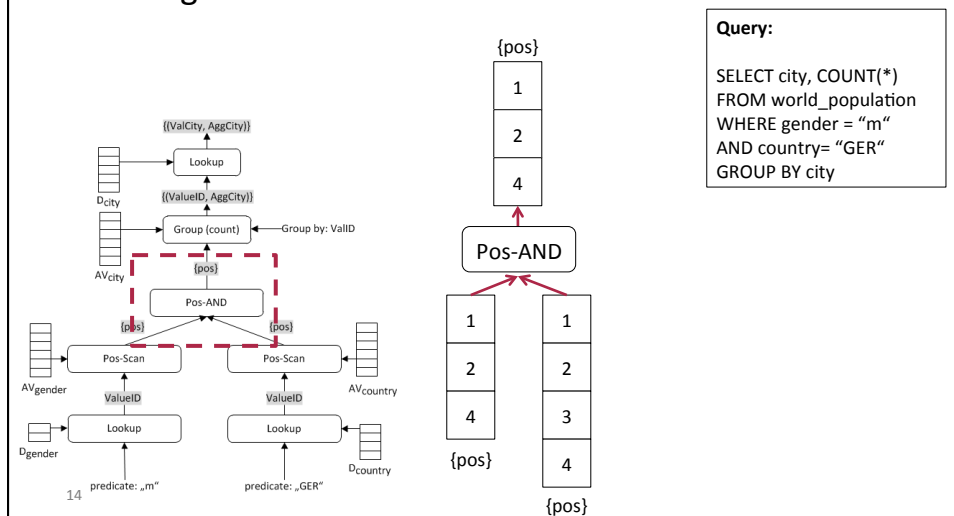
Late Materialization (2)

Scan for constraint



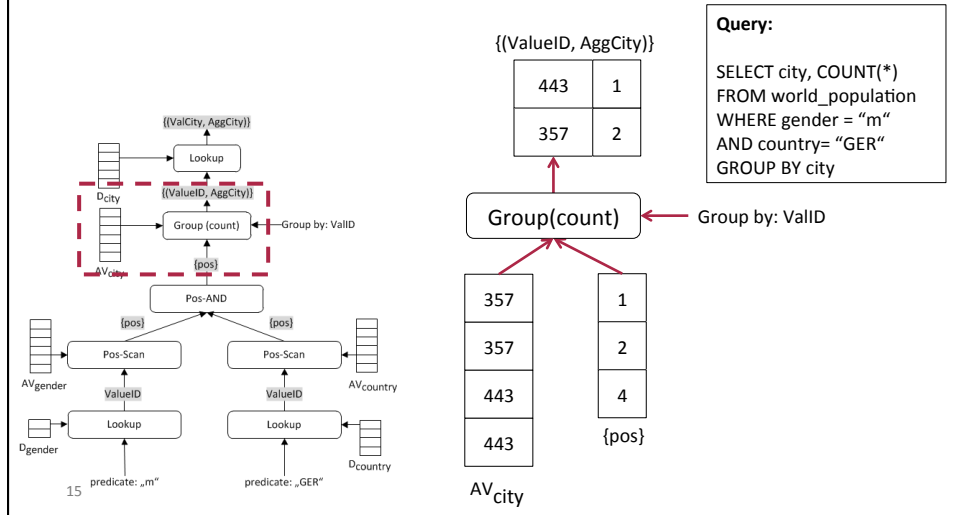
Late Materialization (3)

Logical AND



Late Materialization (4)

Filter attribute vector



Late Materialization (5)

Lookup values from the dictionary

