Homework 3

Gabriele Lorenzo
s314913

December 2022

1: Relational Algebra

Standard Representation:

The corresponding algebric expression for the query is:

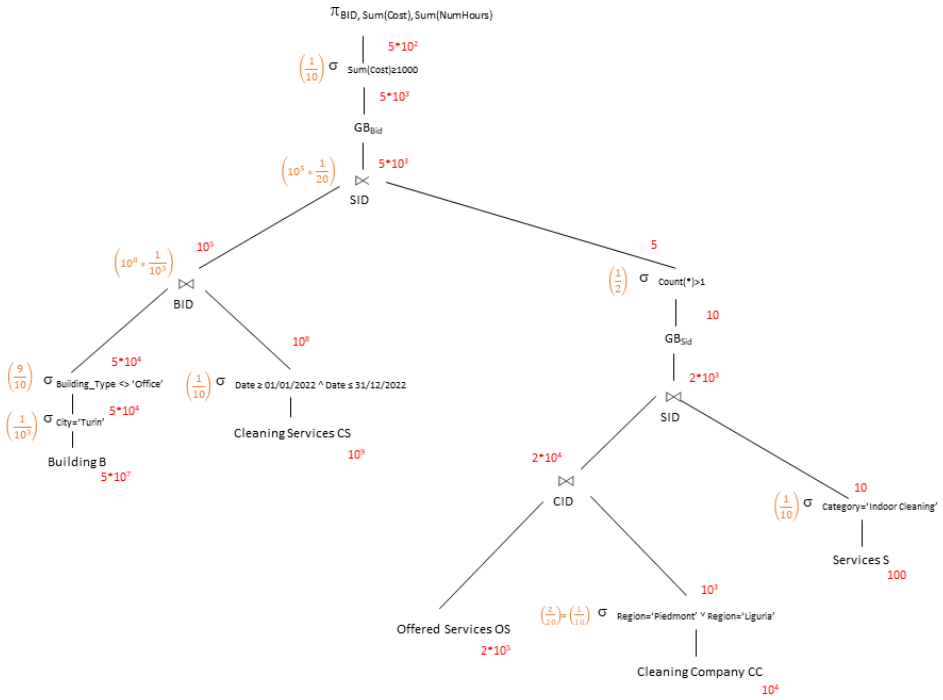

Figure 1: Relational Algebra expression without Group By anticipation

	B-CS	OS-CC	[OS-CC]-S	Anti-Semi Join
Join	Hash Join	Nested Loop (Inner table right)	Nested Loop (Inner table right)	Nested Loop (Inner table right)

	SID	BID
Group By	Group By Hash	Group By Hash

Group By anticipation:

We can perform a group by anticipation on the SID group by.

The resulting algebric expression is:

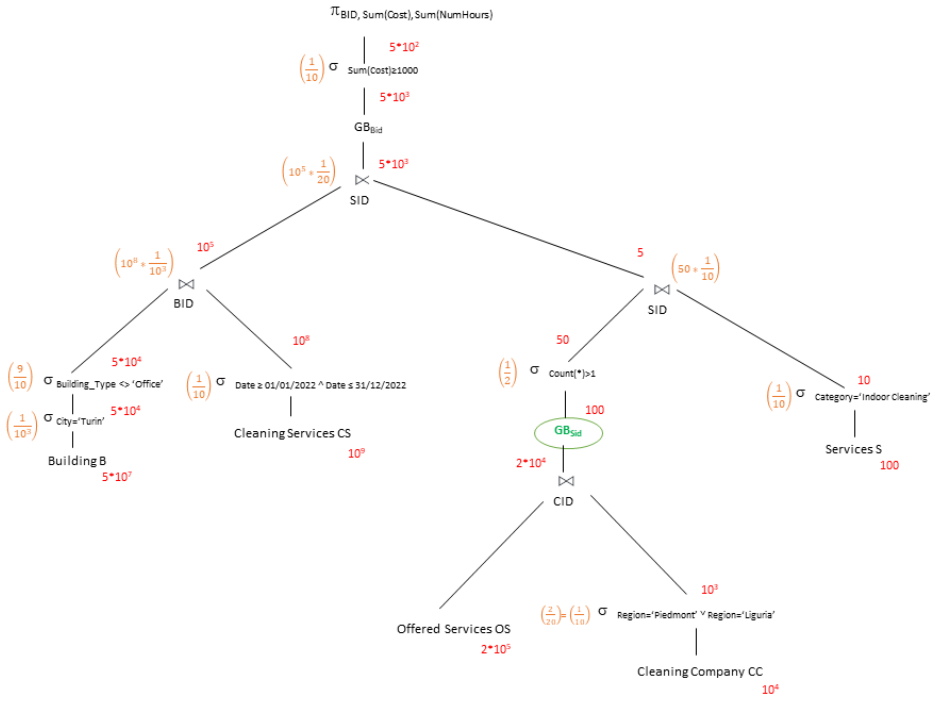

Figure 2: Relational Algebra expression with Group By anticipation

We can see that the cardinality before the join between the tables [OS-CC]-S, is now reduced to 50, from an original value of $2 * 10^4$.

	B-CS	OS-CC	[OS-CC]-S	Anti-Semi Join
Join	Hash Join	Nested Loop (Inner table right)	Nested Loop (Inner table right)	Nested Loop (Inner table right)

	SID	BID
Group By	Group By Hash	Group By Hash

2: Physical Structures Access

Without Indexes:

Table	Access Path
Building B	Table Access Full + Filter
Cleaning-Services CS	Table Access Full + Filter
Cleaning-Company CC	Table Access Full + Filter
Offered-Services OS	Table Access Full
Services S	Table Access Full + Filter

With Indexes:

Table	Index	Access Path
Building B	Secondary Hash Index on City	Index Range Scan + access by RowID
Cleaning-Services CS	Secondary $B^+ - Tree$ Index on Date	Index Range Scan + access by RowID
Cleaning-Company CC	Secondary Hash Index on Region	Index Range Scan + access by RowID
Offered-Services OS	/	Table Access Full
Services S	/	Table Access Full + Filter

- **Building B:**

  Secondary Hash Index on City: good selectivity $\frac{1}{1000}$ and equality predicate. I could have created a covering index on City and BID attributes, but the maintenance cost would be too high.

- **Cleaning-Services CS**

  Secondary $B^+ - Tree$ Index: average selectivity $\frac{1}{10}$ and interval predicate. I use an Index Range scan because I have to retrieve the BID attribure in order to perform the join.

- **Cleaning-Company CC**

  Secondary Hash Index: average selectivity $\frac{1}{10}$ and equality predicate. I use an Index Range scan because I have to retrieve the CID attribure in order to perform the join.

- **Offered-Services OS**

  There is no need for an index.

- **Services S**

  The table is small, there is no need to create and index.