

Homework 3

Data Science And Database Technology

The following relations are given (primary keys are underlined):

ACCOMODATION(CodA, NumberOfGuests, Address, City, Region)
SERVICE(CodS, ServiceName, ServiceType)
ACCOMODATION-HAS-SERVICE(CodA, CodS)
USER(CodU, FirstName, Surname, BusinessAccount, BirthDate, Address, City, Region)
BOOKING(CodA, StartDate, CodU, EndDate)

Assume the following cardinalities:

- $\text{card}(\text{ACCOMODATION}) = 10^5$ tuples,
distinct values of Region = 20
- $\text{card}(\text{SERVICES}) = 10^2$ tuples,
distinct values of ServiceType = 20
- $\text{card}(\text{ACCOMODATION-HAS-SERVICE}) = 10^6$ tuples,
- $\text{card}(\text{USER}) = 10^4$ tuples,
 $\text{MIN}(\text{DATE}(\text{BirthDate})) = 1/1/1930$,
 $\text{MAX}(\text{DATE}(\text{BirthDate})) = 31/12/2009$,
distinct values of Region = 20,
distinct values of BusinessAccount = 2 ("True", "False")
- $\text{card}(\text{BOOKING}) = 10^7$ tuples,
 $\text{MIN}(\text{Date}) = 1/9/2017$, $\text{MAX}(\text{Date}) = 31/08/2020$

Furthermore, assume the following reduction factor for the group by condition:

- Having $\text{COUNT}(\text{Distinct StartDate}) > 1 \approx 1/10$

Consider the following SQL query:

```
select A.CodA, count(Distinct StartDate)
from SERVICE S, ACCOMODATION-HAS-SERVICE AHS,
ACCOMODATION A, BOOKING B, USER U
where S.CodS=AHS.CodS and A.CodA=AHS.CodA and
U.CodU=B.CodU and B.CodA=A.CodA
and (S.ServiceType="Parking" or S.ServiceType="domestic
appliances")
and A.Region="Liguria" and B.StartDate>=1/5/20 and
B.StartDate<=31/8/20
and U.Region<>"Piemonte"
group by CodA
Having COUNT(Distinct StartDate)>1
```

Homework tasks

For the SQL query:

1. Report the corresponding algebraic expression and specify the cardinality of each node (representing an intermediate result or a leaf). If necessary, assume a data distribution. Also analyze the GROUP BY anticipation.
2. Select one or more secondary physical structures to increase query performance. Justify your choice and report the corresponding execution plan (join orders, access methods, etc.).