



Optimizer

Exercises

Exercise 1

- The following relations are given (primary keys are underlined):
 - COMPETITION(CompetitionId, CName, Place, Date, Discipline)
 - ATHLETE(AthleteId, AName, State, SAId, BirthDate)
 - SPORT_ASSOCIATION(SAId, SAName, Address, City)
 - PARTICIPATION(CompetitionId, AthleteId, Rank, Time)

- Assume the following cardinalities:
 - $\text{card}(\text{COMPETITION}) \approx 10^5$ tuples
 - COMPETITION.Discipline: 100 distinct values
 - $\text{MIN}(\text{COMPETITION.Date}) = 1/1/1998$
 - $\text{MAX}(\text{COMPETITION.Date}) = 31/12/2007$
 - $\text{card}(\text{ATHLETE}) \approx 10^4$ tuples
 - ATHLETE.State: 100 distinct values
 - $\text{MIN}(\text{ATHLETE.BirthDate}) = 1/1/1970$
 - $\text{MAX}(\text{ATHLETE.BirthDate}) = 31/12/1990$
 - $\text{card}(\text{SPORT_ASSOCIATION}) \approx 10^3$ tuples
 - $\text{card}(\text{PARTICIPATION}) \approx 4 * 10^6$ tuples
 - $\text{MIN}(\text{PARTICIPATION.Rank}) = 1$
 - $\text{MAX}(\text{PARTICIPATION.Rank}) = 40$
 - Selectivity $\text{HAVING COUNT}(\ast) \geq 5$ equal to $\frac{1}{10}$

Query 1

```
SELECT SAName, City, COUNT(*)
FROM SPORT_ASSOCIATION S, ATHLETE A
WHERE S.SAId = A.SAId
and State = 'Italy'
and BirthDate > 1979
and AthleteId IN
    (SELECT AthleteId
     FROM COMPETITION C, PARTICIPATION P
     WHERE C.CompetitionId = P.CompetitionId
     and Discipline= ' Breaststroke '
     and Date ≥ 2003
     and Rank ≥ 8
     GROUP BY AthleteId
     HAVING COUNT(*) ≥ 5 )
GROUP BY S.SAId, SAName, City;
```