

Il nucleo dell'esercizio consisteva nel capire se un cliente avesse fatto almeno un acquisto in un anno. Per fare questo occorreva estrapolare dalla data di vendita l'anno `to_char(data, 'yyyy')` e poi utilizzare la seguente query SQL

```
select count(distinct to_char(data, 'yyyy'))
  from acquisti
  where cliente = ??? and
        to_char(data, 'yyyy') > to_char(SYSDATE, 'yyyy')-5;
```

che conta in quanti valori distinti di anno l'utente ha acquistato.

La query poteva essere utilizzata dentro un ciclo definito da un cursore che scorreva i clienti

```
...
cursor c_clienti is
  select id from clienti;
anni integer;
begin
  for v_row in c_clienti
  loop
    select count(distinct to_char(data, 'yyyy')) into anni
      from acquisti
      where cliente = v_row.id and
            to_char(data, 'yyyy') > to_char(SYSDATE, 'yyyy')-5;
  end loop;
...

```

o, più efficientemente

```
cursor c_clienti is
select cliente, count(distinct to_char(data, 'yyyy')) as contaanni
  from acquisti
  where to_char(data, 'yyyy') > to_char(SYSDATE, 'yyyy')-5
group by cliente;
```

La soluzione precedente non considera i clienti che non hanno fatto acquisti negli ultimi 5 anni. Questi possono essere tralasciati o recuperati a posteriori mediante un outer-join.

```

explain plan for select PS_PARTKEY,SUM(PS_AVAILQTY)
  from PARTSUPP,SUPPLIER,NATION
  where PS_SUPPKEY=S_SUPPKEY and S_NATIONKEY=N_NATIONKEY
  GROUP BY PS_PARTKEY,N_NAME
  HAVING COUNT(*)>10;
@?/RDBMS/ADMIN/UTLXPLS;

```

Plan Table

Operation	Name	Rows	Bytes	Cost	Pstart	Pstop
SELECT STATEMENT		40K	1M	79226		
FILTER						
SORT GROUP BY		40K	1M	79226		
HASH JOIN		800K	37M	14258		
TABLE ACCESS FULL	NATION	25	725	1		
HASH JOIN		800K	15M	13906		
TABLE ACCESS FULL	SUPPLIER	10K	68K	65		
TABLE ACCESS FULL	PARTSUPP	800K	9M	5074		

$$NP_N = \lceil 25 \times 105 / (4096 \times 0,69) \rceil = 1$$

$$NP_S = \lceil 10.000 \times 143 / (4096 \times 0,69) \rceil = 506$$

$$NP_{PS} = \lceil 800.000 \times 142 / (4.096 \times 0,69) \rceil = 40.195$$

$$\text{Costo Hybrid Hash Join} = 3 \times (40.195 + 506) = 122.103$$

$$NP_{S+PS} = \lceil 800.000 \times (142 + 143) / (4.096 \times 0,69) \rceil = 80.673$$

$$\text{Costo Hybrid Hash Join} = 3 \times (80.673 + 1) = 242.022$$

$$NP_{N+S+PS} = \lceil 800.000 \times (105 + 142 + 143) / (4.096 \times 0,69) \rceil = 110.395$$

$$\text{Costo } (GB_{N+S+PS}) = 2 \times NP_{N+S+PS} \times (\lceil \log_{100}(NP_{N+S+PS}) \rceil + 1) = 2 \times 110.395 \times (3 + 1) = 883.160$$

Assumendo che tutte le nazioni abbiano nomi diversi

$$ET_{GB(PS_PARTKEY+N_NAME)} = \lceil \Phi(800.000, 200.000 \times 25) \rceil = \lceil 5.000.000 \times (1 - (1 - 1/5.000.000)^{800.000}) \rceil = 739.282$$

$$NP_{GB(PS_PARTKEY+N_NAME)} = \lceil (4 + 4) \times 739.282 / (4.096 \times 0,69) \rceil = 2.093$$

Il costo della selezione è pari al costo della lettura di $NP_{GB(PS_PARTKEY+N_NAME)}$

$$\text{Costo totale} = 122.103 + 242.022 + 883.160 + 2.093 = 1.249.378$$