Subject Abdelrahman Adel Abdelfattah Date 17012296

Sheet 6

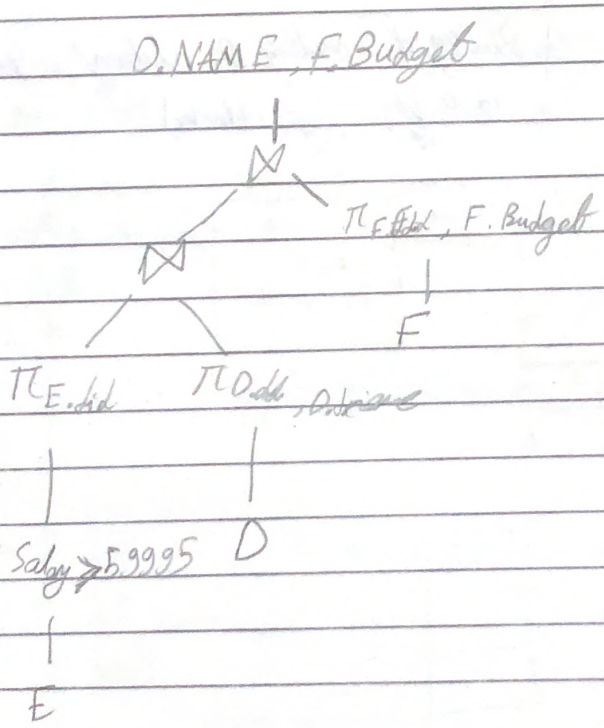
1) Assume Relation R is stored in bR blocks & the file resulting from the operation is stored in b RESULTS.

if <attribute list> includes key of R then the cost is 2*bR (1 b read & 1 b write in result) But if <attribute list> of length K doesn't include the key of R then we sort before eliminating duplicates for the result

Key ⟶ 2*bR

No key ⟶ $3*bR + K*bR*\log_2 bR$

2) a)



(E ⋈ D) ⋈ F

Subject _____ Date / /

b) Clustering on E.SALARY
Secondary on D.FLOOR
Primary on D.DID & Primary E.FDID

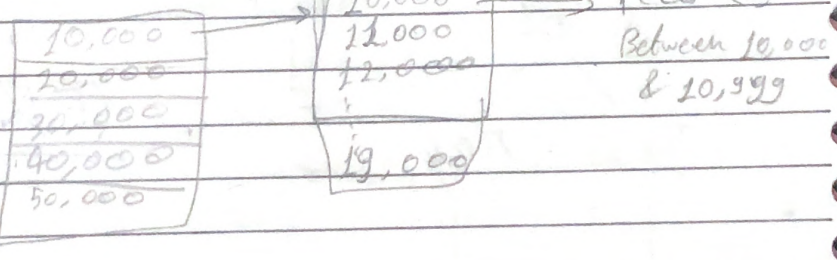
Each has 2 levels

Range Salary: 10,000 to 60,000
Distinct Hobby: 200 values
Distinct FLOOR: 2 values

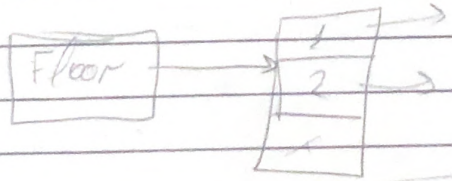
number of records:
EMP. ~~HOBBY~~ = 50,000
DEPT = 5,000
FIN = 5,000

factors
- Blocking of 3 relations is 20
- Blocking of 2 resulting 2 relations is 10
- 12 Buffer Pages (blocks)

Clustering of Salary



Secondary of Floor



Primary of DID

Subject _____ Date / /

i) Salary > 59,000 ⟶ 2,000 records in total

E.HOBBY = "Football" ⟶ 250 records in total

Salary > 59,000 AND Football & Uniform ⇒ 5 records
Sol

5 Football in every 1000

5 records satisfy the condition

ii) first we will read 1,000 records in buffer and get rid of all that doesn't have "Football"

$\frac{1000}{12} = \frac{250}{3} \simeq 84$ times access the disk

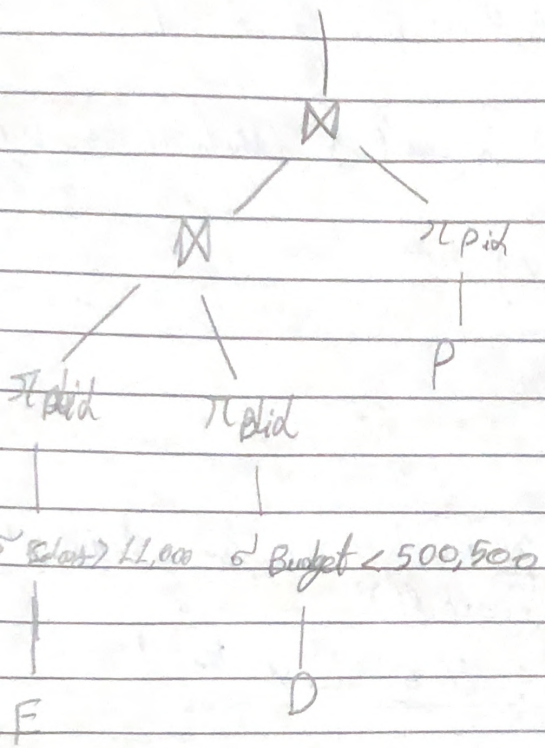
iii) E.Salary > 59,995 ⇒ records = 5

① $\sigma_{salary>59,995}$ E
② $\pi_{D.ID}$ E
③ $\pi_{D.ID, dname}$ D
④ E ⋈ D
⑤ $\pi_{Dname, FID}$ F
⑥ F ⋈ (E ⋈ D)

- with this it creates least records
- D & E have similar attributes, so Join first
- then Join at the end, before Joining with Project.

Subject _____ Date / /

3) E.eid, D.did, P.pid



E → 20,000 , 200 records / block , 100 blocks bfr DP = 1

D → 5,000 , 100 records / block , 50 blocks

P → 1,000 , 2 records / block , 500 blocks

Salary = 10,000 → 30,000 (Secondary)

E.did = 500 (Clustering)

Budget = 500,000 → 1,000,000 (Clustering)

Primary (P.pid) buffer = 12 blocks

E.Salary > 11,000 → 19,000 records → 95 blocks

D.budget < 500,500 → 500 records → 5 blocks

buffer Time ⇒ for Joining 9 Times

Subject _____ Date / /

Buffer Times for Join = 9 Times

(E ⋈ D) → 5 blocks Max

(E ⋈ D) ⋈ F Join => $\frac{505}{12}$ = 43 Times access

Total 51 Times

51 * 6 R