

Sia data la seguente base di dati relativa ad una compagnia aerea

AEREO	ID	Costruttore	Modello	NPosti	AnnoC	AnnoR
VOLO	(Sigla, Partenza, Arrivo, Orario)					
AEROPORTO	(Sigla, Città)					
VIAGGIO	(ID, SiglaVolo, Aereo, Data)					
PERSONALE	(CE, ID, Nome, Cognome, DataNascita)					
EQUIPAGGIO	(Viaggio, Pers)					

1a) Trovare i modelli e numero posti di aerei revisionati prima del 2019 con cui sono stati effettuati voli in partenza da Milano nel 2021

1b) Trovare i modelli di aerei che non hanno mai volato verso New York partendo Napoli nel 2018.

1a) $API9 = \sigma_{ANNO \leq 2019} (AEREO) \text{ VOLI} = (VIAGGIO \bowtie_{SIGLA \text{ VOLO} = VOLO.SIGLA} VOLO) \bowtie_{PARTENZA = AEREO.PARTENZA} AEREO$

$VM = \sigma_{AEROPORTO.CITTA = MILANO \wedge VIAGGIO.DATA \geq 01/01/2021 \wedge VIAGGIO.DATA \leq 31/12/2021} (VOL) \quad Q' = VM \bowtie_{VIAGGIO.AEREO = API9.ID} API9$

Query finale: $Q = \pi_{MODELLO, NPOSTI}(Q')$

1b) Trovo prima quelli che hanno volato verso New York partendo da Napoli nel 2018 ALMENO una volta.

$V18 = \sigma_{DATA \geq 01/01/2018 \wedge DATA \leq 31/12/2018} (VIAGGIO) \quad VPARTN = \sigma_{CITTA = NAPOLI} (VOLO \bowtie_{PARTENZA = AEREO.PARTENZA} AEREO)$

$VARRN = \sigma_{CITTA = NEWYORK} (VOLO \bowtie_{PARTENZA = AEREO.PARTENZA} AEREO) \quad VNN = VPARTN \cap VARRN$

$Q' = \pi_{MODELLO}(V18 \bowtie_{AEREO.ID = VIAGGIO.AEREO} VNN)$ Query finale: $Q = \pi_{MODELLO}(AEREO) - Q'$

2) Siano dati lo schema $R=ABCDE$ e l'insieme di dipendenze funzionali

$F = \{AE \rightarrow BD, BE \rightarrow AC, D \rightarrow AC, B \rightarrow CF\}$

2a) Determinare le tre chiavi dello schema

2b) Dire se lo schema è 3NF e giustificare l'affermazione

2c) Trovare una decomposizione dello schema tale che:

- ogni sottoschema è 3NF

- la decomposizione preserva le dipendenze

- la decomposizione ha un join senza perdita.

2a) Noto che E non appare mai come determinato, quindi sarà parte di una chiave.

Noto che CF non appare mai come determinante, quindi non sarà chiave.

Controllo: $\bullet ABE^+ : R \Rightarrow$ devo controllare AE e BE. $\bullet BDE^+ : R \Rightarrow$ controllo DE.

$\bullet ADE^+ : R$, i sottoinsiemi già li ho marcati come "da controllare".

$AE^+ : R \Rightarrow$ e' chiave. A ed E singolarmente non determinano nulla.

$BE^+ : R \Rightarrow$ e' chiave. $DE^+ : R$ e' chiave. **CHIAVI = AE, BE, DE.**

2b) No, D non e' superchiave, ed AC non e' primo, ma $D \rightarrow AC \in F$!

2c) Ricorro una copertura minimale di F, minimizzo i determinati, ed ottengo:

$F = \{AE \rightarrow B, AE \rightarrow D, BE \rightarrow A, BE \rightarrow C, D \rightarrow A, D \rightarrow C, B \rightarrow C, B \rightarrow F\}$, ora controllo i sottoinsiemi dei determinanti.

$\bullet A^+ = \{A\} \bullet AE^+ = \{E\}$, $AE \rightarrow B$ rimane, $AE \rightarrow D$ rimane. $\bullet B^+ = \{B, C, F\}$ $BE \rightarrow C$ viene sostituito da $B \rightarrow C$. Ottengo:

$F = \{AE \rightarrow B, AE \rightarrow D, BE \rightarrow A, D \rightarrow A, D \rightarrow C, B \rightarrow C, B \rightarrow F\}$, ora controllo le dipendenze ridondanti:

$\bullet AE^+_{F \setminus \{AE \rightarrow B\}} = \{A, E, D, C\} \Rightarrow$ non e' ridondante. $\bullet AE^+_{F \setminus \{AE \rightarrow D\}} \not\supseteq D \bullet A \notin BE^+_{F \setminus \{BE \rightarrow A\}} \bullet A \notin D^+_{F \setminus \{D \rightarrow A\}} \bullet C \notin D^+_{F \setminus \{D \rightarrow C\}} \bullet C \notin B^+_{F \setminus \{B \rightarrow C\}} \bullet F \notin B^+_{F \setminus \{B \rightarrow F\}}$.

F e' già minimale. Applico l'algoritmo! $\rho = \{AEB, AED, AD, CD, BC, BF\}$, AEB contiene una chiave \Rightarrow lossless Join.

3) E' dato un file di 145.500 record. Ogni record occupa 400 byte. Un blocco contiene 2048 byte. Un puntatore a blocco occupa 4 byte. Si utilizza una organizzazione hash con record distribuiti uniformemente tra 250 bucket.

3a) Calcolare l'occupazione della bucket directory e dei bucket

3b) Calcolare il costo medio di una ricerca considerando chiavi di hash univoche

3c) Quanti bucket occorrerebbero per avere un costo medio di ricerca minore o uguale a 20 accessi?

3a) Puntatori x Blocco : $2048/4 = 512$ blocchi x BDirectory = $\lceil 250/512 \rceil = 1$. Record x Bucket = $\frac{\lceil 145.500 \rceil}{250} = 582$

Record x Blocco = $\lfloor (2048 - 4)/400 \rfloor = \lfloor 2040/400 \rfloor = 5$ block x bucket = $\lceil 582/5 \rceil = 117$ 3b) $\lceil 117/2 \rceil = 59$ accessi

3c) $20 \geq \frac{x}{2} \Rightarrow x \leq 40 \Rightarrow \frac{x}{5} \leq 40 \Rightarrow x \leq 200 \Rightarrow \frac{145.500}{x} \leq 200 \Rightarrow x \leq 728$ bucket