



Business Intelligence per Big Data

Esercizi
Data Warehouse

DBG
M



Data Base and Data Mining Group of Politecnico di Torino

DBG
M

AA 2022-2023 - *Politecnico di Torino*



Esercizio 1: Progettazione DW

Si vogliono analizzare le informazioni relative agli accessi che sono stati effettuati dai pazienti presso il reparto di Pronto Soccorso degli ospedali italiani. Il reparto di Pronto Soccorso garantisce gli interventi diagnostico-terapeutici di emergenza/urgenza in risposta al bisogno manifestato dall'utenza.

I pazienti possono raggiungere il reparto di Pronto Soccorso in ambulanza o autonomamente con un mezzo proprio. Ai pazienti che accedono al pronto soccorso viene assegnato un codice colore in base ai sintomi denunciati dal paziente e ai segnali vitali riscontrati. Il codice colore può assumere un solo valore tra rosso, giallo, verde e bianco in base al livello di gravità della condizione del paziente e quindi il grado di urgenza nel procedere con la visita del paziente.

Una volta ricevuto il codice colore, il paziente viene visitato dai medici in servizio presso il reparto. Qualora ritenuto necessario, il paziente è anche sottoposto ad uno o più esami medici che permettano di approfondire e chiarire la sua condizione. Al termine, viene rilasciata una diagnosi. Nei casi più critici viene effettuato il ricovero del paziente presso l'ospedale; altrimenti il paziente viene dimesso.

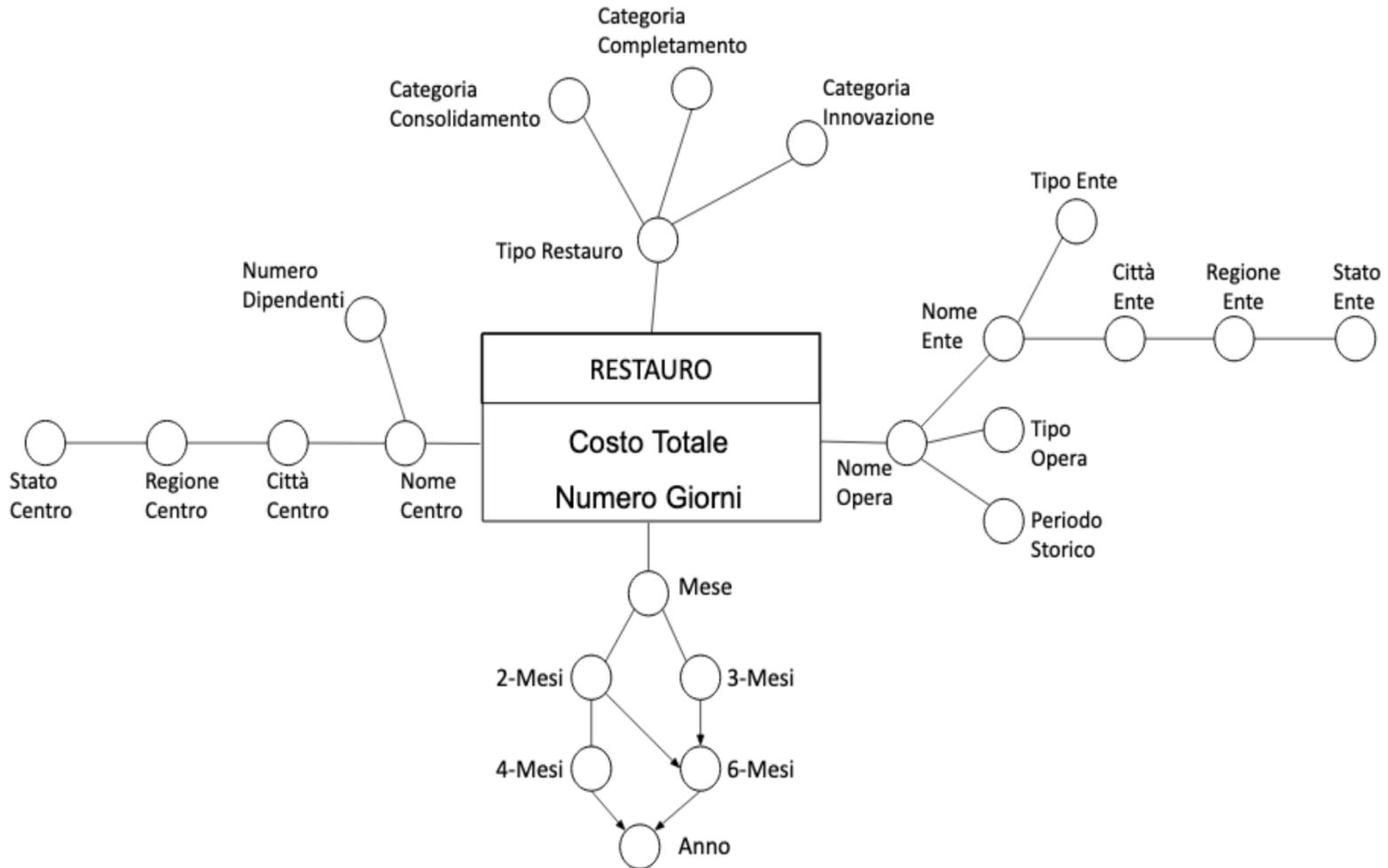
Si vuole analizzare (1) il numero medio di ore di permanenza in pronto soccorso per paziente e (2) il numero medio di visite per medico, in base a:

- modalità di arrivo al Pronto Soccorso (ambulanza oppure mezzo proprio)
- codice colore (bianco, verde, giallo e rosso)
- ospedale, caratterizzato dalla sua collocazione geografica (espressa in termini di città, provincia, regione), e dall'indicazione se siano presenti nell'ospedale uno o più tra i seguenti reparti: cardiologia, pediatria, ortopedia e geriatria. Per ogni ospedale è anche noto se sia una struttura pubblica o accreditata
- data, giorno del mese, giorno festivo, mese, mese dell'anno, quadrimestre, semestre, anno
- sesso e fascia di età del paziente (<13, 13-18, 19-40, 41-70, >70)
- diagnosi rilasciata, caratterizzata dall'elenco degli esami medici eseguiti. Ogni esame è caratterizzato dalla specialità medica a cui si riferisce. Inoltre, ogni esame può essere richiesto in più diagnosi diverse.



Esercizio 2: Progettazione logica di DW

Dato il seguente DFM di un DW, scrivere il corrispondente modello logico relazionale





Esercizio 3: SQL esteso

Dato il seguente schema relazione di un DW, scrivere l'interrogazione utilizzando il linguaggio SQL esteso

CENTRO_RESTAURO (IDCentro, NomeCentro, CittàCentro, RegioneCentro, StatoCentro, NumeroDipendenti)

OPERA (IDOpera, NomeOpera, TipoOpera, PeriodoStorico, NomeEnte, TipoEnte, CittàEnte, RegioneEnte, StatoEnte)

TEMPO (IDTempo, Mese, 2-Mesi, 3-Mesi, 4-Mesi, 6-Mesi, Anno)

TIPO_RESTAURO (IDTipoRestauro, TipoRestauro, CategoriaConsolidamento, CategoriaCompletamento, CategoriaInnovazione)

RESTAURO (IDOpera, IDCentro, IDTempo, IDTipoRestauro, CostoTotale, NumeroGiorni)

Separatamente per città dell'ente e trimestre (attributo 3-Mesi), considerando solo le opere del periodo storico 'Rinascimento', visualizzare:

- il numero di tipi diversi di opera restaurati
- la percentuale del numero giorni di restauro rispetto al totale annuale per città dell'ente
- il totale cumulativo del costo al trascorrere dei trimestri, separatamente per anno



Esercizio 4: SQL esteso

Dato il seguente schema relazione di un DW, scrivere l'interrogazione utilizzando il linguaggio SQL esteso

CENTRO_RESTAURO (IDCentro, NomeCentro, CittàCentro, RegioneCentro, StatoCentro, NumeroDipendenti)

OPERA (IDOpera, NomeOpera, TipoOpera, PeriodoStorico, NomeEnte, TipoEnte, CittàEnte, RegioneEnte, StatoEnte)

TEMPO (IDTempo, Mese, 2-Mesi, 3-Mesi, 4-Mesi, 6-Mesi, Anno)

TIPO_RESTAURO (IDTipoRestauro, TipoRestauro, CategoriaConsolidamento, CategoriaCompletamento, CategoriaInnovazione)

RESTAURO (IDOpera, IDCentro, IDTempo, IDTipoRestauro, CostoTotale, NumeroGiorni)

Separatamente per anno e regione dell'ente, considerando solo i tipi di restauro che appartengono ad entrambe le categorie consolidamento e completamento, visualizzare

- il costo medio mensile
- la percentuale del costo totale rispetto al costo totale complessivo per anno e stato dell'ente
- la posizione in una graduatoria (rank), separatamente per anno, in ordine decrescente di numero medio di giorni di restauro mensili



Esercizio 5: SQL Esteso

Dato il seguente schema relazione di un DW, scrivere l'interrogazione utilizzando il linguaggio SQL esteso

CENTRO_RESTAURO (IDCentro, NomeCentro, CittàCentro, RegioneCentro, StatoCentro, NumeroDipendenti)

OPERA (IDOpera, NomeOpera, TipoOpera, PeriodoStorico, NomeEnte, TipoEnte, CittàEnte, RegioneEnte, StatoEnte)

TEMPO (IDTempo, Mese, 2-Mesi, 3-Mesi, 4-Mesi, 6-Mesi, Anno)

TIPO_RESTAURO (IDTipoRestauro, TipoRestauro, CategoriaConsolidamento, CategoriaCompletamento, CategoricalInnovazione)

RESTAURO (IDOpera, IDCentro, IDTempo, IDTipoRestauro, CostoTotale, NumeroGiorni)

Considerando solo i centri di restauro con più di 20 dipendenti, separatamente per ogni tipo di opera, anno e nome del centro, visualizzare

- Il numero totale di giorni di restauro
- Il numero medio di giorni di restauro per trimestre (attributo 3-Mesi)
- La percentuale del costo totale rispetto al costo totale complessivo della regione in cui si trova il centro
- la posizione in una graduatoria (rank) in ordine crescente rispetto al numero totale di giorni di restauro, separatamente per ogni anno e tipo di opera



Esercizio 6: Viste Materializzate

CENTRO_RESTAURO (IDCentro, NomeCentro, CittàCentro, RegioneCentro, StatoCentro, NumeroDipendenti)

OPERA (IDOpera, NomeOpera, TipoOpera, PeriodoStorico, NomeEnte, TipoEnte, CittàEnte, RegioneEnte, StatoEnte)

TEMPO (IDTempo, Mese, 2-Mesi, 3-Mesi, 4-Mesi, 6-Mesi, Anno)

TIPO_RESTAURO (IDTipoRestauro, TipoRestauro, CategoriaConsolidamento, CategoriaCompletamento, CategoriaInnovazione)

RESTAURO (IDOpera, IDCentro, IDTempo, IDTipoRestauro, CostoTotale, NumeroGiorni)

Dato lo schema logico precedente, considerare le seguenti query di interesse:

- Considerando solo i centri restauri siti in Germania, visualizzare il numero complessivo di giorni di restauro e il relativo costo separatamente per tipo restauro e anno.
- Visualizzare il costo medio dei restauri per giorno di lavoro separatamente per semestre e regione del centro restauro.
- Considerando i tipi di restauro appartenenti alla categoria innovazione, visualizzare il costo totale dei restauri separatamente per stato in cui è sito un centro restauro.

Dato lo schema logico precedente, si svolgano le seguenti attività

- Definire una vista materializzata con CREATE MATERIALIZED VIEW, in modo da ridurre il tempo di risposta delle query di interesse da (a) a (c) sopra riportate. In particolare si specifichi la query SQL associata al Blocco A nella seguente istruzione. **Si ricorda che le istruzioni SQL per le query (a), (b) e (c) non sono richieste.**

```
CREATE MATERIALIZED VIEW ViewCentroRestauro  
BUILD IMMEDIATE  
REFRESH FAST ON COMMIT  
AS
```

Blocco A