

```

----- DATA ANALYST ACADEMY -----
----- PROJECT WORK FACEBOOK -----

```

```

/*

```

Giorno 2: Estrazione dati, reportistica e presentazione

3) Preparazione queries di estrazione dati (DQ/BI) (2/3 ore)

a. Definizione degli obbiettivi dell'estrazione dati (testuale)

b. Preparazione di queries SQL per il raggiungimento di tali obbiettivi

c. L'output sarà un foglio di testo con almeno 5 coppie di obbiettivo analitico/query SQL per rispondere a tale requisito analitico

```

*/

```

```

/*

```

CONTESTO DI BUSINESS

La società cinematografica NetFliz desidera analizzare la possibilità di creare nuovi film, e analizzare la possibilità di espandersi in altre aree di investimento, come il business della musica digitale.

Per questo ha deciso di contrattare una società di analisi dei dati, chiedendo di analizzare i profili Facebook dei suoi utenti.

L'obiettivo di questa analisi è vedere il profilo comportamentale degli utenti, analizzando il genere di film che piacciono e se trascorrono il loro tempo in altre aree ricreative.

```

*/

```

```

USE FACEBOOK2;

```

```

SELECT * FROM UTENTI;

```

```

SELECT * FROM COMPORTAMENTALE;

```

```

SELECT * FROM INTERESSI;

```

```

SELECT * FROM UTENTE_COMPORTAMENTALE;

```

```

SELECT * FROM UTENTE_INTERESSI;

```

-- 1) Valuta su cosa gli utenti trascorrono il loro tempo libero.

```

SELECT CATEGORIA, COUNT(I.CATEGORIA) AS TOP_CATEGORIA

```

```

FROM UTENTI U, UTENTE_INTERESSI UI, INTERESSI I

```

```

WHERE U.UTENTE_ID = UI.UTENTE_FK_INT

```

```

      AND UI.INTERESSE_FK = I.INTERESSE_ID

```

```

GROUP BY CATEGORIA

```

```

ORDER BY TOP_CATEGORIA DESC;

```

/* Si può vedere che i clienti trascorrono il loro tempo

guardando film e ascoltando musica,

la differenza tra loro è praticamente nulla.

```

*/

```

-- 2)

```

SELECT COUNT(CATEGORIA_TONO_POST) AS TOP_TONO_POST, 'ALEGRO' AS TONO

```

```

FROM UTENTI U, UTENTE_COMPORTAMENTALE UE, COMPORTAMENTALE C

```

```

WHERE U.UTENTE_ID = UE.UTENTE_FK_COMP

```

```

      AND UE.COMPORTAMENTO_FK = C.COMPORTAMENTO_ID

```

```

      AND CATEGORIA_TONO_POST LIKE '%_ALLEGRO%'

```

UNION

```
SELECT COUNT(CATEGORIA_TONO_POST) AS TOP_TONO_POST, 'POLEMICO' AS TONO
FROM UTENTI U, UTENTE_COMPORAMENTALE UE, COMPORAMENTALE C
WHERE U.UTENTE_ID = UE.UTENTE_FK_COMP
      AND UE.COMPORTAMENTO_FK = C.COMPORTAMENTO_ID
      AND CATEGORIA_TONO_POST LIKE '%_POLEMICO%'
```

UNION

```
SELECT COUNT(CATEGORIA_TONO_POST) AS TOP_TONO_POST, 'RANDOM' AS TONO
FROM UTENTI U, UTENTE_COMPORAMENTALE UE, COMPORAMENTALE C
WHERE U.UTENTE_ID = UE.UTENTE_FK_COMP
      AND UE.COMPORTAMENTO_FK = C.COMPORTAMENTO_ID
      AND CATEGORIA_TONO_POST LIKE '%_RANDOM%'
```

UNION

```
SELECT COUNT(CATEGORIA_TONO_POST) AS TOP_TONO_POST, 'SHITPOST' AS TONO
FROM UTENTI U, UTENTE_COMPORAMENTALE UE, COMPORAMENTALE C
WHERE U.UTENTE_ID = UE.UTENTE_FK_COMP
      AND UE.COMPORTAMENTO_FK = C.COMPORTAMENTO_ID
      AND CATEGORIA_TONO_POST LIKE '%_SHITPOST%'
```

UNION

```
SELECT COUNT(CATEGORIA_TONO_POST) AS TOP_TONO_POST, 'AUGURI' AS TONO
FROM UTENTI U, UTENTE_COMPORAMENTALE UE, COMPORAMENTALE C
WHERE U.UTENTE_ID = UE.UTENTE_FK_COMP
      AND UE.COMPORTAMENTO_FK = C.COMPORTAMENTO_ID
      AND CATEGORIA_TONO_POST LIKE '%_AUGURI%'
```

ORDER BY TOP_TONO_POST DESC;

```
/*
----- PROJECT WORK FACEBOOK -----
Supponendo che le persone consumino contenuti multimediali
dello stesso "Tono" dei loro post,
l'Azienda NetFliz dovrebbe concentrarsi sulla produzione di
contenuti multimediali "Alegro".
*/
-- 3) AGERAGE AGE

SELECT AVG(DATEDIFF(YEAR,DT_NASCITA, GETDATE())) AS AVG_AGE
FROM UTENTI
WHERE DT_NASCITA >= '1950';

-- 4) RANGE OF AGE
SELECT MIN(DATEDIFF(YEAR,DT_NASCITA, GETDATE())) AS AGE,
       'MIN_AGE' AS 'RANGE'
FROM UTENTI
```

UNION

```
SELECT MAX(DATEDIFF(YEAR,DT_NASCITA, GETDATE())) AS AGE,  
       'MAX_AGE' AS 'RANGE'  
FROM UTENTI  
;
```

--5) GENDER

```
SELECT GENERE, COUNT(GENERE) AS QUANTITA  
FROM UTENTI  
GROUP BY GENERE  
ORDER BY QUANTITA DESC;
```

-- GENERATE CALCULATED COLUMN AGE:

--ALTER TABLE UTENTI ADD AGE AS DATEDIFF(YEAR,DT_NASCITA, GETDATE())

```
SELECT * FROM UTENTI;
```