

Exercici Queries SQL

En quin ordre s'hauran d'importar les dades?

Primer - Taula CARRIERS
Segon - Taula AIRPORTS
Tercer – Taula FLIGHTS

Durant la importació de les dades de la taula flights, els camps de l'arxiu csv d'on s'importen estan ben correlacionats amb els camps de la taula?

No

Per què creus que és això?

La primera línia conte el nom de les columnes, no les dades dels registres, i a més a més el fitxer csv te 5 columnes mes que la tabla de la base de dades.

Inclou totes les consultes en un arxiu .sql i afegeix-les al teu repositori de github.

Les consultes són les següents:

1. Quantitat de registres de la taula de vols:

Resultat:

total

4758

```
SELECT COUNT(*) AS total  
FROM flights
```

2. Retard promig de sortida i arribada segons l'aeroport origen.

Resultat:

```
SELECT Origin,  
AVG(ArrDelay) AS prom_arribades,  
AVG(DepDelay) AS prom_sortides  
FROM flights  
GROUP BY Origin
```

Origin	prom_arribades	prom_sortides
ABE	17.1250	15.7500
ABQ	3.8158	6.8684
ABY	4.5000	19.0000
AGS	19.0000	15.3333
ALB	-2.3000	1.4000
AMA	1.2500	0.0000
ANC	1.8571	3.5000
ASE	7.0000	10.0000
ATL	9.7248	10.0891
ATW	0.6667	1.6667
AUS	1.0000	1.1250
AVL	-7.6667	-3.3333
AZO	14.0000	19.5000
BDL	-5.5200	2.2400
BET	0.5000	1.5000
BFL	-6.0000	0.0000
BGR	45.0000	60.0000
BHM	5.7500	8.3750
BIL	14.0000	4.5000
BIS	17.0000	0.0000
BNA	6.8800	15.3400
BOI	0.0909	1.9091
BOS	5.2110	7.8899
BTR	7.6667	-0.8333
BTV	-11.4444	-4.0000

3. Retard promig d'arribada dels vols, per mesos, anys i segons l'aeroport origen. A més, volen que els resultat es mostrin de la següent forma (fixa't en l'ordre de les files):

Resultat:

origin	colYear	colMonth	prom_arribades
ABE	1988	5	73.0000
ABE	1988	11	26.0000
ABE	1990	1	-4.0000
ABE	1990	12	-5.0000
ABE	1991	7	10.0000
ABE	1992	6	-5.0000
ABE	2005	1	-18.0000
ABE	2006	1	60.0000
ABQ	1987	10	10.0000
ABQ	1989	1	7.0000
ABQ	1990	4	-3.0000
ABQ	1991	1	6.0000
ABQ	1993	11	3.0000
ABQ	1994	1	12.0000
ABQ	1995	1	-8.0000
ABQ	1995	7	-1.0000
ABQ	1995	8	7.0000
ABQ	1995	9	-2.0000
ABQ	1996	1	-1.0000
ABQ	1996	7	0.0000
ABQ	1997	9	-8.0000
ABQ	1997	10	12.0000
ABQ	1997	11	-11.0000
ABQ	1997	12	14.0000
ABQ	1998	1	-5.0000

```
SELECT Origin, colYear, colMonth, AVG (ArrDelay) AS prom_arribades
FROM flights
GROUP BY Origin, colYear, colMonth ORDER BY Origin
```



4. Retard promig d'arribada dels vols, per mesos, anys i segons l'aeroport origen (mateixa consulta que abans i amb el mateix ordre). Però a més, ara volen que en comptes del codi de l'aeroport es mostri el nom de la ciutat.

Resultat:

City ▲ 1	colYear ▲ 2	colMonth ▲ 3	prom_arribades
Albany	1988	3	-6.0000
Albany	1989	3	0.0000
Albany	1992	12	-18.0000
Albany	1993	10	0.0000
Albany	1994	10	-11.0000
Albany	2001	9	-5.0000
Albany	2003	3	-4.0000
Albany	2004	10	1.0000
Albany	2006	1	-8.0000
Albany	2006	3	17.0000
Albany	2007	6	5.0000
Albany	2008	12	15.0000
Albuquerque	1987	10	10.0000
Albuquerque	1989	1	7.0000
Albuquerque	1990	4	-3.0000
Albuquerque	1991	1	6.0000
Albuquerque	1993	11	3.0000
Albuquerque	1994	1	12.0000
Albuquerque	1995	1	-8.0000
Albuquerque	1995	7	-1.0000
Albuquerque	1995	8	7.0000
Albuquerque	1995	9	-2.0000
Albuquerque	1996	1	-1.0000
Albuquerque	1996	7	0.0000
Albuquerque	1997	9	-8.0000

```
SELECT City, colYear, colMonth,
AVG(ArrDelay) AS prom_arribades
FROM flights LEFT JOIN usairports
ON flights.Origin = usairports.IATA
GROUP BY City, colYear, colMonth
ORDER BY City, colYear, colMonth
```

5. Les companyies amb més vols cancel·lats, per mesos i any. A més, han d'estar ordenades de forma que les companyies amb més cancel·lacions apareguin les primeres.

Resultat:

```
SELECT UniqueCarrier, colYear, colMonth,
COUNT(cancelled) AS total_cancelled
FROM flights
WHERE Cancelled=1
GROUP BY UniqueCarrier
ORDER BY total_cancelled DESC
```

UniqueCarrier	colYear	colMonth	total_cancelled ▼ 1
DL	1988	1	3
CO	1988	2	1
PI	1988	7	1
AS	1989	8	1
TW	1987	11	1
UA	1990	2	1
UA	1988	4	1
US	1990	8	1
UA	1989	2	1
AS	1988	9	1
DL	1990	5	1
AA	1989	5	1
US	1987	12	1

6. L'identificador dels 10 avions que més distància han recorregut fent vols.

Resultat:

```
SELECT TailNum, SUM(Distance) AS  
totalDistance  
FROM flights  
WHERE TailNum IS NOT NULL  
AND TailNum != ""  
GROUP BY TailNum  
ORDER BY SUM(Distance) DESC LIMIT 0,10
```

TailNum	totalDistance ▾ 1
N543UA	7675
N12238	6829
N580JB	6277
N5DCAA	5300
N637DL	5223
N5DJAA	5046
N653UA	5005
N614UA	4950
N615DL	4948
N128DL	4584

7. Companyies amb el seu retard promig només d'aquelles les quals els seus vols arriben al seu destí amb un retràs promig major de 10 minuts.

Resultat:

```
SELECT UniqueCarrier, AVG(ArrDelay) AS avgDelay  
FROM flights  
GROUP BY UniqueCarrier  
HAVING avgDelay > 10  
ORDER BY avgDelay DESC
```

UniqueCarrier	avgDelay
TZ	38.5714
F9	32.8571
B6	20.3611
PI	17.7273
9E	16.4000
YV	13.2727
OH	12.9310
EV	12.8857
FL	11.7872