

Data: 13/06/2018

Autor: JRVG

Contingut:

1	EXERCICI	2
1.1	Enunciat.....	2
1.1.1	Info.txt	2
1.1.2	Model de la BBDD.....	3
2	Consultes	4
2.1	Quantitat de registres de la taula de vols:	4
2.2	Retard promig de sortida i arribada segons l'aeroport origen.....	4
2.3	Retard promig d'arribada dels vols, per mesos i segons l'aeroport origen.	5
2.4	Retard promig d'arribada dels vols, per mesos i segons l'aeroport origen	6
2.5	Les companyies amb més vols cancelats.	7
2.6	Companyies amb el seu promig només d'aquelles les quals els seus vols arriben al seu destí amb un retràs major de 10 minuts.....	8
2.7	L'identificador dels 10 avions que més kilòmetres han recorregut fent vols comercials.....	9

1 EXERCICI

Exercicis workbench

<https://drive.google.com/file/d/1rh-PD3piYDgLnwhm-ZOKHw-yLmHnZxD/view?usp=sharing>

1.1 Enunciat

07/06/2018 [2.1 Enunciat.docx](#)

13/06/2018 [Enunciat.pdf](#)

1.1.1 Info.txt

Les dades d'aquest dataset consisteixen en tots els vols comercials realitzats a EEUU l'any XXX.

El dataset original conté tots els vols des de l'octubre de 1987 fins l'Abril de 2008 i ocupa 12GB, per això hem decidit limitar-ho a un sol any.

En el següent enllaç podeu trobar la informació d'aquest dataset, on indica el significat de cada columna i possibles valors que pot prendre.

Vols: <http://stat-computing.org/dataexpo/2009/the-data.html>

Dades addicionals (Aeroports (Airports), companies (Carriers) i avions (Planes)): <http://stat-computing.org/dataexpo/2009/supplemental-data.html>

On canviar la variable --secure-file-priv:

On apunta la variable --secure-file-priv: <C:\ProgramData\MySQL\MySQL Server 8.0\Uploads>

Punts a tenir en compte:

· Sembla que si importes dades des de la interfície (botó dret al icono de la BD --> table data import wizard) no dóna problemes, però si es vol fer des d'un script de forma (LOAD DATA INFILE), pot donar problemes per motius de seguretat. Per a resoldre això s'han d'importar les dades des d'un directory especificat per MySQL, que es pot consultar mitjançant la comanda: SHOW VARIABLES LIKE "secure_file_priv";

· Options File és a Server --> Options File

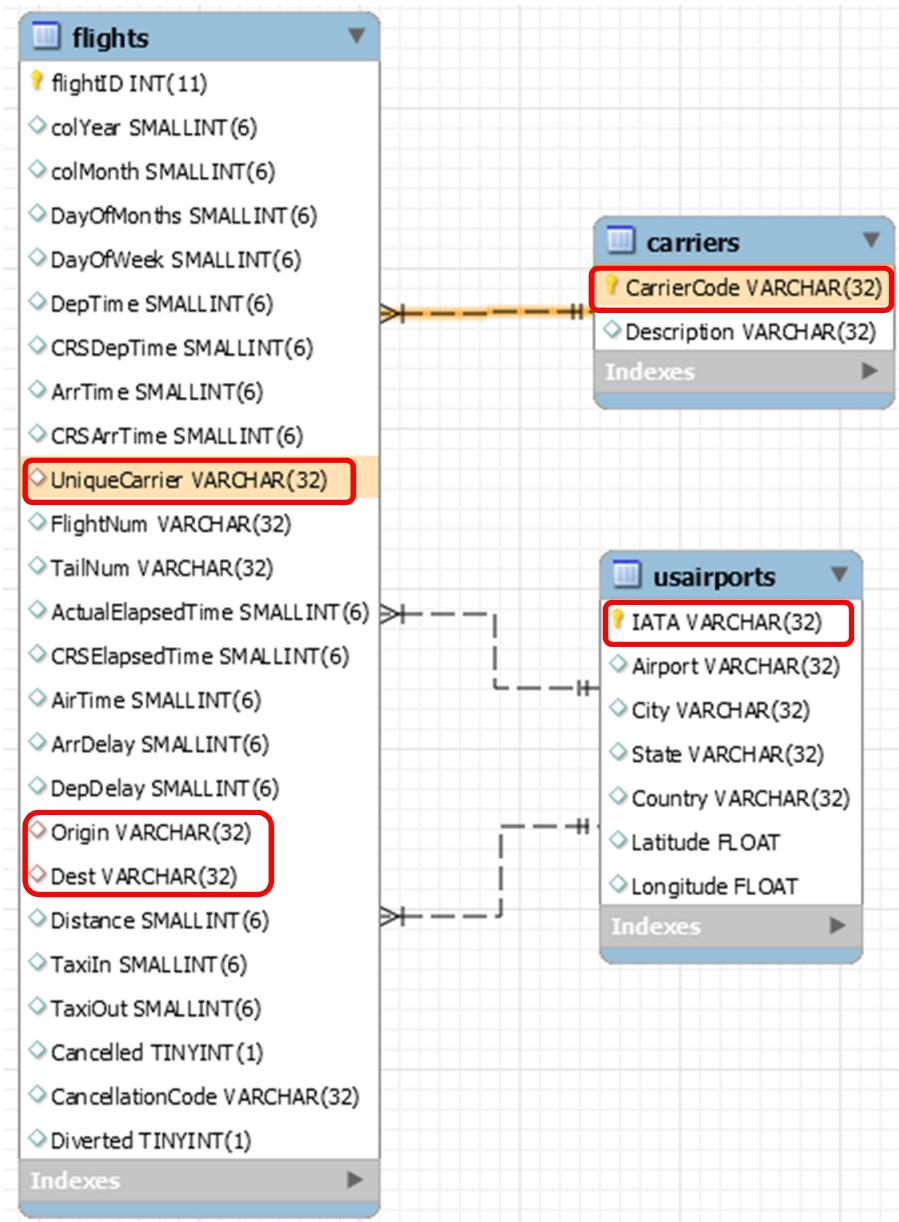
SHOW VARIABLES LIKE "secure_file_priv";

Variable_name	Value
secure_file_priv	C:\ProgramData\MySQL\MySQL Server 8.0\Uploads\

show variables like 'sql_mode';

Variable_name	Value
sql_mode	STRICT_TRANS_TABLES,NO_ENGINE_SUBSTITUTION

1.1.2 Model de la BBDD



2 Consultes

2.1 Quantitat de registres de la taula de vols:

The screenshot shows a MySQL Workbench interface with three tabs: '02_Airport_test', 'airports_test', and '03_usairlineflights_Qualities'. The '03_usairlineflights_Qualities' tab is active. The query window displays the following code:

```
56 /* Consulta 1 • Quantitat de registres de la taula de vols:
58 • SELECT COUNT(*) FROM usairlineflights.flights; # 1777
--
```

The results grid shows one row with the value 1777.

2.2 Mitjana del retard de sortida i arribada segons l'aeroport origen.

The screenshot shows a MySQL Workbench interface with three tabs: '02_Airport_test', 'airports_test', and '03_usairlineflights_Qualities'. The '03_usairlineflights_Qualities' tab is active. The query window displays the following code:

```
60 /* Consulta 2 • Retard promig (Mitjana del retard) de sortida i arribada segons l'aeroport origen.
83 # Interpretació 1: L'aeroport d'origen queda qualificat pel seu codi IATA.
84 • USE `usairlineflights`;
85 • SELECT
86     AVG(f.DepDelay) AS 'Departure average delay',
87     AVG(f.ArrDelay) AS 'Arrival average delay',
88     f.Origin
89 FROM
90     flights As f
91 GROUP BY f.Origin;
```

The results grid shows the following data:

Departure average delay	Arrival average delay	Origin
9.6800	7.8800	AMA
18.2378	19.5778	ATL
7.8292	7.7542	AUS
9.4954	6.9450	BDL
12.7317	14.7317	BHM
8.6343	11.4595	BNA
15.5128	12.7692	BOS

The screenshot shows a MySQL Workbench interface with three tabs: '02_Airport_test', 'airports_test', and '03_usairlineflights_Qualities'. The '03_usairlineflights_Qualities' tab is active. The query window displays the following code:

```
93 # Interpretació 2: L'aeroport d'origen s'explica amb el seu nom, no amb el seu codi IATA.
94 • SELECT
95     AVG(f.DepDelay) AS 'Departure average delay',
96     AVG(f.ArrDelay) AS 'Arrival average delay',
97     ua.Airport AS 'Airport'
98 FROM
99     flights As f
100    INNER JOIN
101        airports AS ua ON f.Origin = ua.IATA
102 GROUP BY f.Origin
103 ORDER BY ua.Airport ASC;
```

The results grid shows the following data:

Departure average delay	Arrival average delay	Airport
9.6800	7.8800	Amarillo International
7.8292	7.7542	Austin-Berstrom International
12.7317	14.7317	Birmingham International
9.4954	6.9450	Bradley International
15.5128	12.7692	Gen Edw L Logan Intl
8.6343	11.4595	Nashville International
18.2378	19.5778	William B Hartsfield-Atlanta Intl

2.3 Mitjana del retard d'arribada dels vols, per mesos i segons l'aeroport origen.

A més, volen que els resultat es mostrin de la següent forma:

LAX, 2000, 01, retard

LAX, 2000, 02, retard

...

LAX, 2000, 12, retard

LAX, 2001, 01, retard

...

ONT, 2000, 01, retard

ONT, 2000, 02, retard

etc.

The screenshot shows a MySQL Workbench interface with three tabs at the top: '02_Airport_test', 'airports_test', and '03_usairlineflights_Queries*'. The '03_usairlineflights_Queries*' tab is active and contains the following SQL code:

```
105 /* Consulta 3 • Mitjana del retard d'arribada dels vols, per mesos i segons l'aeroport origen.
125 • SELECT
126     Origin,
127     colYear AS 'Year',
128     colMonth AS 'Month',
129     AVG(ArrDelay) AS 'Average delay'
130
131     flights
132 GROUP BY Origin
133 ORDER BY colYear ASC , colMonth DESC;
```

Below the code is a 'Result Grid' table with the following data:

Origin	Year	Month	Average delay
BNA	2000	12	11.4595
BOS	2000	12	12.7692
AMA	2000	12	7.8800
ATL	2000	12	19.5778
AUS	2000	12	7.7542
BDL	2000	12	6.9450
BHM	2000	12	14.7317

2.4 Mitjana del retard d'arribada dels vols, per mesos 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.

Los Angeles, 2000, 01, retard

Los Angeles, 2000, 02, retard

The screenshot shows the MySQL Workbench interface with three tabs: '02_Airport_test', 'airports_test', and '03_usairlineflights_Queries*'. The '03_usairlineflights_Queries*' tab contains the following SQL code:

```
136 /* Consulta 4 • Mitjana del retard d'arribada dels vols, per mesos i segons l'aeroport origen.
155 • USE `usairlineflights`;
156 • SELECT
157     ua.City AS 'City',
158     f.colYear AS 'Year',
159     f.colMonth AS 'Month',
160     AVG(f.ArrDelay) AS 'Arrival delay'
161 FROM
162     usairports AS ua
163     INNER JOIN
164     flights AS f ON ua.IATA = f.Origin
165 GROUP BY ua.City
166 ORDER BY ua.City;
```

The 'Result Grid' shows the following data:

City	Year	Month	Arrival delay
Amarillo	2000	12	7.8800
Atlanta	2000	12	19.5778
Austin	2000	12	7.7542
Birmingham	2000	12	14.7317
Boston	2000	12	12.7692
Nashville	2000	12	11.4595
Windsor Locks	2000	12	6.9450

The screenshot shows the MySQL Workbench interface with the same three tabs. The '03_usairlineflights_Queries*' tab now includes a comment '#La consulta feta a la inversa' before the SELECT statement, which is highlighted with a red rectangle.

```
168 #La consulta feta a la inversa
169 • SELECT
170     ua.City AS 'City',
171     f.colYear AS 'Year',
172     f.colMonth AS 'Month',
173     AVG(f.ArrDelay) AS 'Arrival delay'
174 FROM
175     flights AS f
176     INNER JOIN
177     usairports AS ua ON f.Origin = ua.IATA
178 GROUP BY ua.City
179 ORDER BY ua.City;
```

The 'Result Grid' shows the same data as the previous screenshot:

City	Year	Month	Arrival delay
Amarillo	2000	12	7.8800
Atlanta	2000	12	19.5778
Austin	2000	12	7.7542
Birmingham	2000	12	14.7317
Boston	2000	12	12.7692
Nashville	2000	12	11.4595
Windsor Locks	2000	12	6.9450

2.5 Les companyies amb més vols cancel·lats.

A més, han d'estar ordenades de forma que les companyies amb més cancelacions apareguin les primeres.

The screenshot shows a MySQL query editor interface. The top bar has tabs for '02_Airport_test', 'airports_test', and '03_usairlineflights_Queries*' (which is active). Below the tabs is a toolbar with various icons. The main area contains the following SQL code:

```
181
182  /* Consulta 5 • Les companyies amb més vols cancel·lats.
197  • USE `usairlineflights`;
198  • SELECT
199      c.CarrierCode AS 'Carrier code',
200      c.Description AS 'Carrier description',
201      SUM(f.Cancelled) AS 'Cancellations'
202  FROM
203      carriers AS c
204          INNER JOIN
205      flights AS f ON c.CarrierCode = f.UniqueCarrier
206  WHERE
207      f.Cancelled = '1'
208  GROUP BY f.UniqueCarrier
209  ORDER BY SUM(f.Cancelled) DESC;
```

Below the code is a result grid with the following data:

Carrier code	Carrier description	Cancellations
AA	American Airlines Inc.	7

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

The screenshot shows a MySQL Workbench interface with three tabs at the top: '02_Airport_test', 'airports_test', and '03_usairlineflights_Queries*'. The '03_usairlineflights_Queries*' tab is active. Below the tabs is a toolbar with various icons. The main area contains a code editor with the following SQL query:

```

209 ORDER BY SUM(f.Cancelled) DESC;
210
211 /* Consulta 6 • Companies amb la seva mitjana només d'aquelles les quals els seus
212 • USE `usairlineflights`;
213 # Interpretro que només cal considerar els retards que superen els 10 minuts i most
214 • SELECT
215     c.Description AS 'Carrier description',
216     AVG(f.ArrDelay) AS 'Arrival delay'
217 FROM flights AS f INNER JOIN carriers as c ON f.UniqueCarrier = c.CarrierCode
218 WHERE f.ArrDelay > 10
219 GROUP BY f.UniqueCarrier

```

Below the code editor is a 'Result Grid' section with the following data:

	Carrier description	Arrival delay
American Airlines Inc.	45.7068	
Delta Air Lines Inc.	44.5258	

Incluc una consulta addicional que no es demana en què es té en compte tot l'espai mmostral.

En aquest cas en tenir en compte tots els retards, que inclouen les mostres sense retards, s'emmascara el resultat.

Ja que les companyies que s'anticipen en l'arrivada poden compensar els retards.

CONCLUSIÓ:

És un exemple clar de que quan es vol mesurar un paràmetre convé teneir clar que és important definir quin conjunt de dades cal considerar per poder obtenir-lo.

The screenshot shows a MySQL Workbench interface with three tabs at the top: '02_Airport_test', 'airports_test', and '03_usairlineflights_Queries*'. The '03_usairlineflights_Queries*' tab is active. Below the tabs is a toolbar with various icons. The main area contains a code editor with the following SQL query:

```

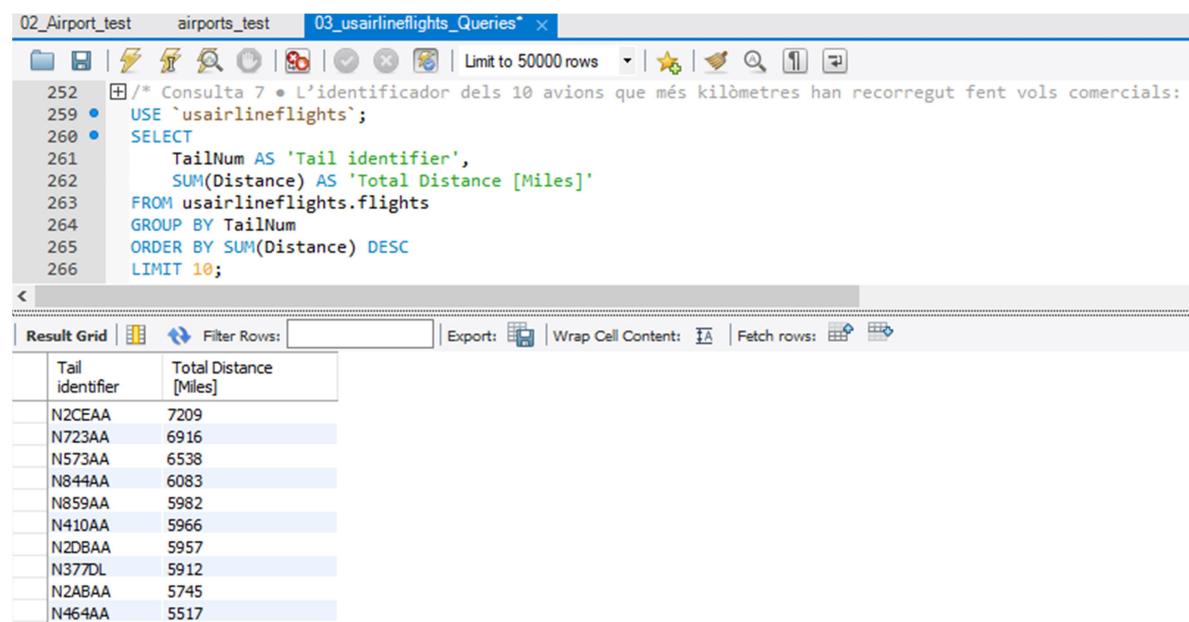
237 WHERE f.ArrDelay > 10
238 GROUP BY f.UniqueCarrier
239 ORDER BY c.Description ASC;
240
241 # Interpretació diferent. Només la incluc com a ampliació de l'exercici.
242 # Considerar tots els retards i mostrar només les companyies amb una mitjana de retard que superi els 10 minuts.
243 • USE `usairlineflights`;
244 • SELECT
245     c.Description AS 'Carrier description',
246     AVG(f.ArrDelay) AS 'Arrival delay'
247 FROM flights AS f INNER JOIN carriers as c ON f.UniqueCarrier = c.CarrierCode

```

Below the code editor is a 'Result Grid' section with the following data:

	Carrier description	Arrival delay
Delta Air Lines Inc.	22.9644	

2.7 L'identificador dels 10 avions que més kilòmetres han recorregut fent vols comercials



The screenshot shows a MySQL Workbench interface with a query editor and a results grid. The query editor contains the following SQL code:

```
252 /* Consulta 7 • L'identificador dels 10 avions que més kilòmetres han recorregut fent vols comercials:
253 • USE `usairlineflights`;
254 • SELECT
255     TailNum AS 'Tail identifier',
256     SUM(Distance) AS 'Total Distance [Miles]'
257     FROM usairlineflights.flights
258     GROUP BY TailNum
259     ORDER BY SUM(Distance) DESC
260     LIMIT 10;
```

The results grid displays the top 10 entries from the query:

Tail identifier	Total Distance [Miles]
N2CEAA	7209
N723AA	6916
N573AA	6538
N844AA	6083
N859AA	5982
N410AA	5966
N2DBAA	5957
N377DL	5912
N2ABAA	5745
N464AA	5517