Database Homework 2

0416037 李家安

OS: Fedora 23 Server
RAM: DDR4 8G * 2
Storage 480G SSD

1.
SQL:

```
SELECT COUNT(*)
FROM (
    SELECT DISTINCT Origin, Dest
    FROM records
) T;
```

TIME:
27.64 sec

ANS:
6901

```
MariaDB [test]> SELECT COUNT(*) FROM (SELECT DISTINCT Origin, Dest FROM records) T;
+----------+
| COUNT(*) |
+----------+
|     6901 |
+----------+
1 row in set (27.64 sec)
```

2.
SQL:

```
SELECT COUNT(*) AS flaitAmount
FROM records
WHERE Dest='JFK' AND ActualElapsedTime BETWEEN 60 AND 180;
```

TIME:
6.11 sec

ANS:
303868

```
MariaDB [test]> SELECT COUNT(*) AS flaitAmount FROM records WHERE Dest='JFK' AND ActualElapsedTime BETWEEN 60 AND 180;
+------------+
| flaitAmount |
+------------+
|      303868 |
+------------+
1 row in set (6.11 sec)
```

3.
SQL:

```
SELECT model, COUNT(*) AS modelAmount
FROM plane_data
WHERE model LIKE '737%'
GROUP BY model;
```

TIME:
0.01 sec

ANS:

```
MariaDB [test]> SELECT model, COUNT(*) FROM plane_data WHERE model LIKE '737%' GROUP BY model;
ERROR 2006 (HY000): MySQL server has gone away
No connection. Trying to reconnect...
Connection id:    787
Current database: test

+-----------+----------+
| model     | COUNT(*) |
+-----------+----------+
| 737-230   |        1 |
| 737-236   |        4 |
| 737-282   |        1 |
| 737-282C  |        1 |
| 737-2P6   |        1 |
| 737-2X6C  |        1 |
| 737-2Y5   |        1 |
| 737-301   |        7 |
| 737-317   |        4 |
| 737-322   |       64 |
| 737-33A   |        4 |
| 737-3A4   |        9 |
| 737-3B7   |       16 |
| 737-3G7   |       20 |
| 737-3H4   |      147 |
| 737-3K2   |        2 |
| 737-3L9   |        2 |
| 737-3Q8   |        9 |
| 737-3S3   |        4 |
| 737-3T5   |        3 |
| 737-3TO   |       50 |
| 737-3Y0   |        6 |
| 737-401   |       13 |
| 737-490   |       16 |
| 737-4B7   |       27 |
| 737-4Q8   |       22 |
| 737-4S3   |        1 |
| 737-522   |       30 |
| 737-524   |       56 |
| 737-5H4   |       25 |
| 737-705   |        1 |
| 737-724   |       36 |
| 737-73A   |        2 |
| 737-76N   |       27 |
| 737-76Q   |        2 |
| 737-790   |       19 |
| 737-7AD   |        1 |
| 737-7BD   |       32 |
| 737-7H4   |      308 |
| 737-7Q8   |        2 |
| 737-824   |      100 |
| 737-832   |       68 |
| 737-890   |       33 |
| 737-8FH   |        1 |
| 737-924   |       12 |
| 737-924ER |        8 |
| 737-990   |       12 |
+-----------+----------+
47 rows in set (0.01 sec)
```

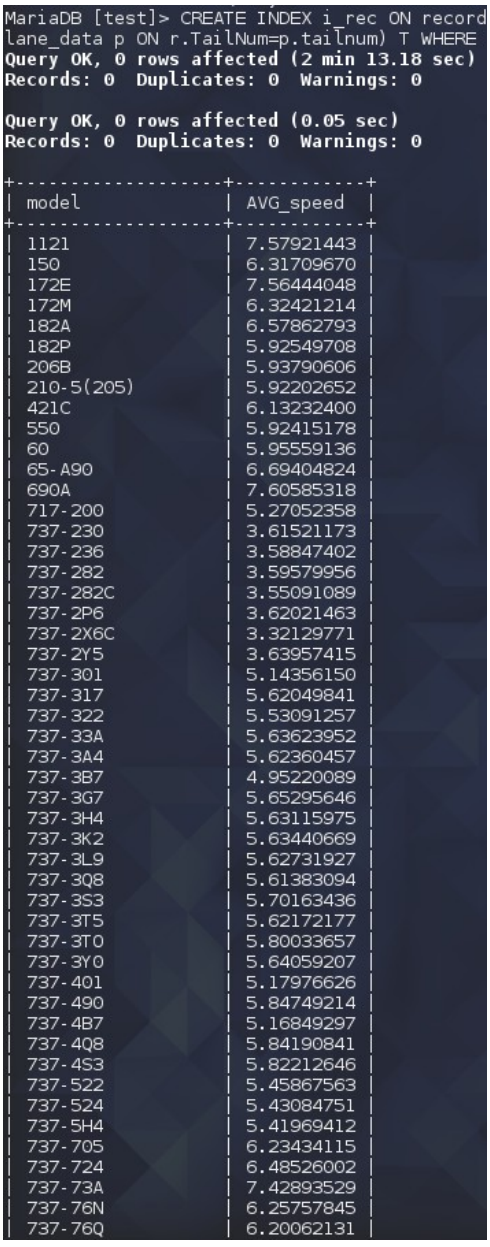
4.
SQL:

```
CREATE INDEX i_r4
ON records(TailNum);
CREATE INDEX i_p4
ON plane_data(tailnum);
```

```
SELECT model, AVG(Distance)/AVG(ActualElapsedTime) AS AVG_speed
FROM (
  SELECT model, ActualElapsedTime, Distance
  FROM records r RIGHT JOIN plane_data p ON r.TailNum=p.tailnum
) T
WHERE model IS NOT NULL AND Distance IS NOT NULL AND ActualElapsedTime IS NOT
NULL AND ActualElapsedTime!=0
GROUP BY model;

DROP INDEX i_r4
ON records;
DROP INDEX i_p4
ON plane_data;
```

TIME:
2 min 13.18 sec + 0.05 sec + 1 min37.78 sec = 3 min 56.01 sec

ANS:



```
MariaDB [test]> CREATE INDEX i_rec ON record
lane_data p ON r.TailNum=p.tailnum) T WHERE
Query OK, 0 rows affected (2 min 13.18 sec)
Records: 0  Duplicates: 0  Warnings: 0

Query OK, 0 rows affected (0.05 sec)
Records: 0  Duplicates: 0  Warnings: 0

+------------------+-----------+
| model            | AVG_speed |
+------------------+-----------+
| 1121             | 7.57921443 |
| 150              | 6.31709670 |
| 172E             | 7.56444048 |
| 172M             | 6.32421214 |
| 182A             | 6.57862793 |
| 182P             | 5.92549708 |
| 206B             | 5.93790606 |
| 210-5(205)       | 5.92202652 |
| 421C             | 6.13232400 |
| 550              | 5.92415178 |
| 60               | 5.95559136 |
| 65-A90           | 6.69404824 |
| 690A             | 7.60585318 |
| 717-200          | 5.27052358 |
| 737-230          | 3.61521173 |
| 737-236          | 3.58847402 |
| 737-282          | 3.59579956 |
| 737-282C         | 3.55091089 |
| 737-2P6          | 3.62021463 |
| 737-2X6C         | 3.32129771 |
| 737-2Y5          | 3.63957415 |
| 737-301          | 5.14356150 |
| 737-317          | 5.62049841 |
| 737-322          | 5.53091257 |
| 737-33A          | 5.63623952 |
| 737-3A4          | 5.62360457 |
| 737-3B7          | 4.95220089 |
| 737-3G7          | 5.65295646 |
| 737-3H4          | 5.63115975 |
| 737-3K2          | 5.63440669 |
| 737-3L9          | 5.62731927 |
| 737-3Q8          | 5.61383094 |
| 737-3S3          | 5.70163436 |
| 737-3T5          | 5.62172177 |
| 737-3T0          | 5.80033657 |
| 737-3Y0          | 5.64059207 |
| 737-401          | 5.17976626 |
| 737-490          | 5.84749214 |
| 737-4B7          | 5.16849297 |
| 737-4Q8          | 5.84190841 |
| 737-4S3          | 5.82212646 |
| 737-522          | 5.45867563 |
| 737-524          | 5.43084751 |
| 737-5H4          | 5.41969412 |
| 737-705          | 6.23434115 |
| 737-724          | 6.48526002 |
| 737-73A          | 7.42893529 |
| 737-76N          | 6.25757845 |
| 737-76Q          | 6.20062131 |
| 737-76Q          | 6.20062131 |
| 737-790          | 6.40997187 |
| 737-7AD          | 6.14800444 |
| 737-7BD          | 6.10818417 |
| 737-7H4          | 6.18263788 |
| 737-7Q8          | 6.76318653 |
| 737-824          | 6.59269711 |
| 737-832          | 6.40819303 |
| 737-890          | 6.86663541 |
| 737-8FH          | 6.90343341 |
| 737-924          | 6.36654034 |
| 737-924ER        | 6.47393227 |
| 737-990          | 6.57038558 |
| 747-2B5F         | 6.27681370 |
| 747-422          | 7.18762836 |
| 747-451          | 7.29874891 |
| 757-212          | 6.23553789 |
| 757-222          | 6.82843053 |
| 757-223          | 6.81594520 |
| 757-224          | 6.40956774 |
| 757-225          | 6.32397594 |
| 757-231          | 6.43156732 |
| 757-232          | 6.45465432 |
| 757-23N          | 5.86610514 |
| 757-251          | 6.39125387 |
| 757-26D          | 6.33942076 |
| 757-2B7          | 6.29981612 |
| 757-2G7          | 7.11922366 |
| 757-2Q8          | 6.45164933 |
| 757-2S7          | 7.02637507 |
| 757-324          | 6.57105390 |
| 757-33N          | 6.57667358 |
| 757-351          | 6.92979338 |
| 767-201          | 5.92969210 |
| 767-223          | 7.22926227 |
| 767-224          | 6.50409810 |
| 767-2B7          | 5.85262697 |
| 767-322          | 7.32504361 |
| 767-323          | 7.62607051 |
| 767-324          | 6.78811666 |
| 767-332          | 6.51246180 |
| 767-33A          | 7.89083233 |
| 767-3CB          | 7.89741696 |
| 767-3G5          | 7.86719683 |
| 767-3P6          | 6.68247046 |
| 767-424ER        | 8.00700211 |
| 767-432ER        | 7.27992262 |
| 777-222          | 7.44301694 |
| 777-224          | 6.34772310 |
| 777-232          | 6.13505564 |
| 777-232LR        | 5.76615570 |
| A-1B             | 7.60256708 |
| A109E            | 7.59517853 |
| A318-111         | 6.09970628 |
| A319-111         | 6.20712747 |
| A319-112         | 5.41568641 |
| A319-114         | 5.90231410 |
| A319-131         | 6.52662734 |
| A319-132         | 6.47297979 |
| A320-211         | 5.99630477 |
| A320-212         | 6.11491495 |
| A320-214         | 6.43510028 |
| A320-231         | 6.37077731 |
| A320-231         | 6.37077731 |
| A320-232         | 6.48876264 |
| A321-211         | 6.71483051 |
| A330-223         | 7.22528426 |
| A330-323         | 7.67633678 |
| AS 355F1         | 7.66224954 |
| ATR 72-212       | 2.74553212 |
| ATR-72-212       | 2.73771299 |
| C90              | 5.95289395 |
| CL-600-2B19      | 4.85590261 |
| CL-600-2C10      | 5.45294912 |
| CL600-2D24       | 5.06319043 |
| DA 20-A1         | 6.62522792 |
| DC-7BF           | 7.64956717 |
| DC-9-31          | 4.58660499 |
| DC-9-32          | 4.84488484 |
| DC-9-41          | 4.35939609 |
| DC-9-51          | 4.34947282 |
| DC-9-82(MD-82)   | 5.93632326 |
| DC-9-83(MD-83)   | 6.26075946 |
| DHC-8-102        | 2.27010924 |
| DHC-8-202        | 2.81330731 |
| E-90             | 7.63165932 |
| EMB-120          | 3.08263509 |
| EMB-120ER        | 3.01349673 |
| EMB-135ER        | 3.97316605 |
| EMB-135KL        | 4.84945909 |
| EMB-135LR        | 4.54655144 |
| EMB-145          | 4.64566446 |
| EMB-145EP        | 3.81440759 |
| EMB-145LR        | 4.70883683 |
| EMB-145XR        | 5.65934396 |
| ERJ 190-100 IGW  | 4.89868378 |
| EXEC 162F        | 6.04066074 |
| F85P-1           | 6.79776371 |
| FALCON XP        | 6.28940718 |
| FALCON-XP        | 6.23671130 |
| G-IV             | 7.63317140 |
| HST-550          | 5.92618167 |
| KITFOX IV        | 5.93953479 |
| MD 83            | 6.12734214 |
| MD-88            | 5.09086429 |
| MD-90-30         | 5.81234880 |
| OTTER DHC-3      | 6.31159660 |
| PA-28-180        | 5.29787766 |
| PA-31-350        | 6.43698629 |
| PA-32R-300       | 5.92648959 |
| PA-32RT-300      | 7.55686650 |
| S-50A            | 5.95046499 |
| S-76A            | 7.64682381 |
| S55A             | 5.93225288 |
| SAAB 340B        | 2.46667663 |
| T210N            | 5.90888471 |
| T337G            | 6.10547217 |
| VANS AIRCRAFT RV6 | 5.95543667 |
+------------------+-----------+
165 rows in set (1 min 37.87 sec)

Query OK, 0 rows affected (0.05 sec)
Records: 0  Duplicates: 0  Warnings: 0

Query OK, 0 rows affected (0.03 sec)
Records: 0  Duplicates: 0  Warnings: 0
```

5.
SQL:
TIME:
ANS:

6.
SQL:

```
SELECT DISTINCT model
FROM plane_data
WHERE model IS NOT NULL AND tailnum NOT IN (
  SELECT DISTINCT TailNum
  FROM records
  WHERE year='2008'
);
```

TIME:
14 min 20.93 sec

ANS:
11AC
M-5-235C

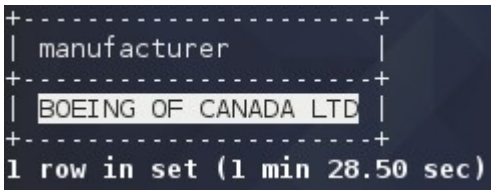
7.
SQL:

```
CREATE INDEX i_r7
ON records(TailNum);
CREATE INDEX i_p7
ON plane_data(tailnum);

SELECT DISTINCT manufacturer
FROM (
  SELECT manufacturer, ArrDelay
  FROM records r, plane_data p
  WHERE r.TailNum=p.tailnum AND ArrDelay IS NOT NULL AND ArrDelay!=0
) T
GROUP BY manufacturer
HAVING AVG(ArrDelay) > 30;

DROP INDEX i_r7
ON records;
```

```
DROP INDEX i_p7
ON plane_data;
```

TIME:
1 min 56.13 sec + 0.05 sec + 1 min 28.50 sec = 3 min 25.07 sec

ANS:
BOEING OF CANADA LTD



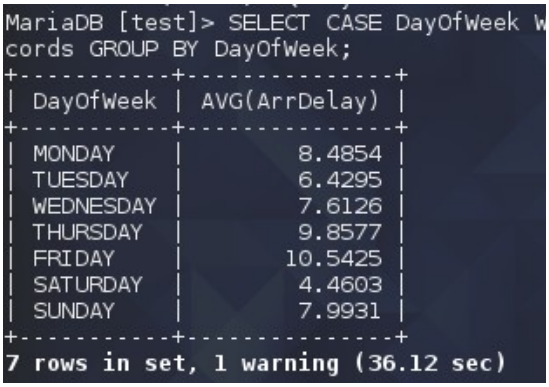
8.
(a) week

SQL:

```
SELECT CASE DayOfWeek
  WHEN 1 THEN 'MONDAY'
  WHEN 2 THEN 'TUESDAY'
  WHEN 3 THEN 'WEDNESDAY'
  WHEN 4 THEN 'THURSDAY'
  WHEN 5 THEN 'FRIDAY'
  WHEN 6 THEN 'SATURDAY'
  WHEN 7 THEN 'SUNDAY'
END AS DayOfWeek , AVG(DepDelay)
FROM records
GROUP BY DayOfWeek;
```

TIME:
36.12 sec

ANS:
Saturday 遇到的 Delay 會最少



(b) time

SQL:

```
SELECT  CONCAT(ROUND(CRSDepTime/100, 0), 'hour') AS hour, AVG(ArrDelay)
```

FROM records
GROUP BY ROUND(CRSDepTime/100, 0);

TIME:
52.60 sec

ANS:
5:00~6:00 遇到的 Delay 會最少

```
MariaDB [test]> SELECT  CONCAT(ROUND(CRSDepTime/100, 0), 'hour') AS hour, AVG(ArrDelay)
    -> FROM records
    -> GROUP BY ROUND(CRSDepTime/100, 0);
+--------+--------------+
| hour   | AVG(ArrDelay) |
+--------+--------------+
| 0hour  |       2.6105 |
| 1hour  |       1.9386 |
| 2hour  |       1.2364 |
| 3hour  |       8.4130 |
| 4hour  |       4.6221 |
| 5hour  |      -0.6995 |
| 6hour  |       0.1291 |
| 7hour  |       1.1390 |
| 8hour  |       2.2587 |
| 9hour  |       3.3129 |
| 10hour |       4.3412 |
| 11hour |       5.1508 |
| 12hour |       6.3103 |
| 13hour |       8.0736 |
| 14hour |       9.7145 |
| 15hour |      11.0143 |
| 16hour |      12.0590 |
| 17hour |      13.9883 |
| 18hour |      14.3408 |
| 19hour |      14.5807 |
| 20hour |      14.6129 |
| 21hour |      13.0642 |
| 22hour |       9.3605 |
| 23hour |       7.1158 |
| 24hour |       7.1268 |
+--------+--------------+
25 rows in set (52.60 sec)
```

9.
SQL:

```
CREATE INDEX i_r9
ON records(TailNum);
CREATE INDEX i_p9
ON plane_data(tailnum);

SELECT Year-issY, AVG(ArrDelay)
FROM (
  SELECT CAST(RIGHT(issue_date,4) AS int) AS issY, r.Year, ArrDelay
  FROM plane_data p RIGHT JOIN records r ON p.tailnum=r.TailNum
) T
WHERE issY IS NOT NULL AND Year IS NOT NULL AND Year >= issY AND issY!=0
GROUP BY Year-issY
ORDER BY Year-issY;

DROP INDEX i_r9
ON records;
DROP INDEX i_p9
ON plane_data;
```

TIME:

1 min 54.19 sec + 0.05 sec + 2 min 23.07 sec = 4 min 17.31 sec

ANS:
我認為會，利用 SQL 找出飛行年與飛機發行年(也就是當時飛機年齡)與平均 delay 時間，發現越老的飛機，delay 時間也會越久。

Year-issY	AVG(ArrDelay)
0	8.1484
1	7.7235
2	8.5124
3	8.6219
4	8.6038
5	8.4808
6	7.8200
7	7.8488
8	8.3449
9	8.2232
10	8.1060
11	8.3685
12	8.6385
13	9.0975
14	8.7778
15	8.8641
16	9.0206
17	9.3694
18	9.0141
19	9.6461
20	9.4080
21	11.0360
22	9.8619
23	9.8201
24	8.7323
26	5.9467
27	7.5969
28	6.2576
29	10.4055
30	9.7152
31	11.3856
32	12.8707

32 rows in set, 4254 warnings (2 min 23.07 sec)

10.
(a) season

SQL:

```
SELECT  season, SUM(ArrDelay) AS sum_delay
FROM (SELECT ArrDelay,
  CASE
    WHEN Month IN (3,4,5) THEN 'Spring'
    WHEN Month IN (6,7,8) THEN 'Summer'
    WHEN Month IN (9,10,11) THEN 'FALL'
    WHEN Month IN (12,1,2) THEN 'Winter'
  END AS season
  FROM records
) T
GROUP BY season
ORDER BY FIELD(season, 'Spring', 'Summer', 'Fall', 'Winter');
```

TIME:
53.57 sec

ANS:

季節感覺會，夏天的 delay 就明顯高很多。

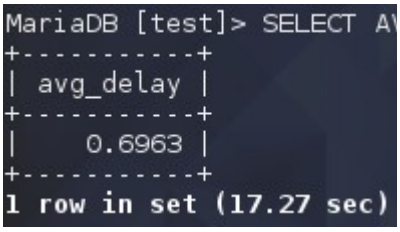
(b) weather

SQL:

```
SELECT AVG(WeatherDelay) AS avg_delay
FROM records
WHERE WeatherDelay IS NOT NULL;
```

TIME:
17.27 sec

ANS:
天氣感覺不大會影響 delay，我試著找出平均 WeatherDelay，發現它並不大。



11.
SQL:
TIME:
ANS:

12.
SQL:

```
CREATE INDEX i_r12
ON records(TailNum);
CREATE INDEX i_p12
ON plane_data(tailnum);

SELECT `type`, AVG(Taxiln) AS avgTaxiln
FROM plane_data p, records r
WHERE p.tailnum=r.TailNum AND `type` IS NOT NULL
GROUP BY type
ORDER BY AVG(r.Taxiln);

DROP INDEX i_r12
```

ON records;
DROP INDEX i_p12
ON plane_data;

TIME:
2 min 0.48 sec + 0.04 sec + 4 min 22.52 sec = 6 min 23.44 sec

ANS:
我試圖尋找飛機 type 與起飛時間的關係，結果發現似乎沒有太大的關係。

```
MariaDB [test]> SELECT `type`, AVG(Taxiln) AS avgTaxiln
    -> FROM plane_data p, records r
    -> WHERE p.tailnum=r.TailNum AND `type` IS NOT NULL
    -> GROUP BY type
    -> ORDER BY AVG(r.Taxiln);
 +---------------------+-----------+
| type                | avgTaxiln |
+---------------------+-----------+
| Corporation         |    7.4559 |
| Foreign Corporation |    7.8078 |
| Partnership         |    8.2712 |
| Co-Owner            |    8.3172 |
| Individual          |    8.3262 |
+---------------------+-----------+
5 rows in set (4 min 22.52 sec)
```