

Database homework 1 report

Task A:

1. Task_A.sql transcript:

```
mysql> source task_A.sql
Query OK, 0 rows affected (0.02 sec)

Query OK, 0 rows affected (0.02 sec)

Query OK, 0 rows affected (0.04 sec)

Query OK, 0 rows affected (0.02 sec)

Query OK, 0 rows affected (0.03 sec)

Query OK, 138 rows affected (0.01 sec)
Records: 138  Deleted: 0  Skipped: 0  Warnings: 0

Query OK, 182527 rows affected (1.71 sec)
Records: 182527  Deleted: 0  Skipped: 0  Warnings: 0

Query OK, 1825270 rows affected (29.75 sec)
Records: 1825270  Deleted: 0  Skipped: 0  Warnings: 0

Query OK, 1089969 rows affected (8.97 sec)
Records: 1089969  Deleted: 0  Skipped: 0  Warnings: 0

Query OK, 1825270 rows affected (30.47 sec)
Records: 1825270  Deleted: 0  Skipped: 0  Warnings: 0
```

2. 'Describe' command for each table:

```
mysql> describe champ;
+-----+-----+-----+-----+-----+-----+
| Field          | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| champion_name  | varchar(15)   | NO   |     | NULL    |       |
| champion_id    | int(11)       | NO   | PRI | NULL    |       |
+-----+-----+-----+-----+-----+-----+
2 rows in set (0.00 sec)
```

```
mysql> describe match_info
-> ;
+-----+-----+-----+-----+-----+-----+
| Field        | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| match_id     | int(11)       | NO   | PRI | NULL    |       |
| duration     | int(11)       | YES  |     | NULL    |       |
| version      | varchar(15)   | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

```
mysql> describe participant;
```

Field	Type	Null	Key	Default	Extra
player_id	int(11)	NO	PRI	NULL	
match_id	int(11)	NO	MUL	NULL	
player	tinyint(4)	YES		NULL	
champion_id	int(11)	NO		NULL	
ss1	varchar(15)	YES		NULL	
ss2	varchar(15)	YES		NULL	
position	varchar(13)	NO		NULL	

```
7 rows in set (0.00 sec)
```

```
mysql> describe stat;
```

Field	Type	Null	Key	Default	Extra
player_id	int(11)	NO	PRI	NULL	
win	tinyint(1)	YES		NULL	
item1	smallint(6)	YES		NULL	
item2	smallint(6)	YES		NULL	
item3	smallint(6)	YES		NULL	
item4	smallint(6)	YES		NULL	
item5	smallint(6)	YES		NULL	
item6	smallint(6)	YES		NULL	
kills	tinyint(4)	YES		NULL	
deaths	tinyint(4)	YES		NULL	
assists	tinyint(4)	YES		NULL	
longesttimespentliving	smallint(6)	YES		NULL	
doublekills	tinyint(4)	YES		NULL	
triplekills	tinyint(4)	YES		NULL	
quadrakills	tinyint(4)	YES		NULL	
pentakills	tinyint(4)	YES		NULL	
legendarykills	tinyint(4)	YES		NULL	
goldearned	mediumint(9)	YES		NULL	
firstblood	tinyint(1)	YES		NULL	

```
19 rows in set (0.00 sec)
```

```
mysql> describe teamban;
```

Field	Type	Null	Key	Default	Extra
match_id	int(11)	NO	PRI	NULL	
team	char(1)	NO		NULL	
champion_id	int(11)	NO		NULL	
banturn	tinyint(4)	NO	PRI	NULL	

```
4 rows in set (0.01 sec)
```

3. Questions:

(1) What the difference between type “char” and type “varchar”?

char: 固定大小的字符，預設大小是一，可以透過後方加上括號來指定大小。可以存比輸入預設的字符長度還小的字串，但使用的空間

仍會是原本預設，會浪費空間。

varchar: 動態方式儲存自符，跟 char 幾乎相同，但在儲存較原本預設串長度還小的字串時，只會使用本身字串所需的記憶體，較省空間。

(2) Type “boolean” would be stored as which type in MySQL?

Tinyint(1) (從 stat 裡的 firstblood 可以看到)

(3) How many bytes it should take for “tinyint”, “smallint”, “mediumint”, “int”? (e.g. 8 bytes for “bigint”) And what’s the range they can express? (e.g. from -1000 to 1000)

Type	# of byte	Range
tinyint	1	0~255
smallint	2	-8,388,608~ 8,388,607
mediumint	3	-32,768~ 32,767
int	4	-2,147,483,648~ 2,147,483,647

(4) What do you think about this table schema? If you can change this table architecture, how would you modify it and why?

- 基本上在做下面的題目時就會發現有做到多次的 champion_name 跟 champion_id 互相配對的地方，這樣會造成很大的麻煩，所以應該把 participant 裡面直接加上 champion_name，會方便很多。

Task C:

1. SQL Screenshot:

```
SELECT COUNT(champion_id) AS cnt
FROM champ;
```

Query result:

```
mysql> source 1.sql
+-----+
| cnt |
+-----+
| 138 |
+-----+
1 row in set (0.01 sec)
```

2. SQL Screenshot:

```
SELECT COUNT(S.A) AS cnt
FROM(
    SELECT DISTINCT SUBSTRING_INDEX(version, '.',2) AS A
    FROM match_info
) S;
```

Query result:

3. SQL Screenshot:

```
SELECT S.champion_name, S.cnt
FROM
(
  SELECT C.champion_name, P.cnt
  FROM champ C,
  (
    SELECT champion_id, COUNT(*) AS cnt
    FROM participant
    WHERE position LIKE '%JUNGLE%'
    GROUP BY champion_id
  ) P
  WHERE P.champion_id = C.champion_id
) S
ORDER BY S.cnt DESC
LIMIT 3;
```

Query result:

champion_name	cnt
Lee Sin	56598
Master Yi	23385
Graves	19767

3 rows in set (1.01 sec)

4. SQL Screenshot:

```
SELECT match_id, CONCAT(FLOOR(duration/3600), ':', FLOOR((duration-3600* FLOOR(duration/3600))/60),
  | ':', duration-3600* FLOOR(duration/3600)-60*FLOOR((duration-3600* FLOOR(duration/3600))/60)) AS time
FROM match_info
ORDER BY duration DESC
LIMIT 5;
```

Query result:

match_id	time
146486	1:23:11
69303	1:20:14
581	1:16:59
70361	1:15:6
176628	1:13:34

5 rows in set (0.18 sec)

5. SQL Screenshot:

```

SELECT CASE
    WHEN F.win = 1 then 'win'
    WHEN F.win = 0 then 'lose'
    END
    AS win_lose,
    F.cnt
FROM
(
    SELECT COUNT(A.ma) AS cnt, A.win
    FROM
    (
        SELECT P.match_id AS ma, AVG(S.longesttimespentliving) AS av, S.win
        FROM participant P,
        (
            SELECT player_id, win, longesttimespentliving
            FROM stat
        ) S
        WHERE P.player_id = S.player_id
        GROUP BY P.match_id, S.win
    ) A
    WHERE A.av > 1200
    GROUP BY A.win
) F;

```

Query result:

```

mysql> source hw1-6.sql
+-----+-----+
| win_lose | cnt |
+-----+-----+
| lose     | 338 |
| win      | 806 |
+-----+-----+
2 rows in set (22.62 sec)

```

6. SQL Screenshot:

```

SELECT MA.position, C.champion_name
FROM champ C,
(
    SELECT D.position, MAX(D.cnt) AS cnt
    FROM
    (
        SELECT P.position, P.champion_id, COUNT(P.player_id) AS cnt
        FROM match_info M,
        (
            SELECT position, champion_id, player_id, match_id
            FROM participant
            WHERE position = 'DUO_CARRY' OR position = 'DUO_SUPPORT' OR position = 'JUNGLE' OR position = 'MID' OR position = 'TOP'
        ) P
        WHERE M.duration BETWEEN 2399 AND 3001 AND M.match_id = P.match_id
        GROUP BY P.position, P.champion_id
    ) D
    GROUP BY D.position
) MA,
(
    SELECT P.position, P.champion_id, COUNT(P.player_id) AS cnt
    FROM match_info M,
    (
        SELECT position, champion_id, player_id, match_id
        FROM participant
        WHERE position = 'DUO_CARRY' OR position = 'DUO_SUPPORT' OR position = 'JUNGLE' OR position = 'MID' OR position = 'TOP'
    ) P
    WHERE M.duration BETWEEN 2399 AND 3001 AND M.match_id = P.match_id
    GROUP BY P.position, P.champion_id
) AL
WHERE MA.position = AL.position AND MA.cnt = AL.cnt AND AL.champion_id = C.champion_id
ORDER BY FIELD(MA.position, 'DUO_CARRY', 'DUO_SUPPORT', 'JUNGLE', 'MID', 'TOP')
;

```

Query result:

```
mysql> source hw1-7.sql
+-----+-----+
| position | champion_name |
+-----+-----+
| DUO_CARRY | Caitlyn       |
| DUO_SUPPORT | Thresh       |
| JUNGLE    | Lee Sin      |
| MID       | Ahri         |
| TOP       | Riven        |
+-----+-----+
5 rows in set (21.00 sec)
```

7. SQL Screenshot:

```
SELECT MA.position, C.champion_name, MA.kda
FROM champ C,
(
  SELECT P.position, S.kda, P.champion_id
  FROM
  (
    SELECT position, champion_id, player_id
    FROM participant
    WHERE position = 'DUO_CARRY' OR position = 'DUO_SUPPORT' OR position = 'JUNGLE' OR position = 'MID' OR position = 'TOP'
  ) P,
  (
    SELECT player_id, (kills+assists)/deaths AS kda
    FROM stat
    WHERE deaths != 0
  ) S
  WHERE P.player_id = S.player_id
) AL,
(
  SELECT P.position, MAX(S.kda) AS kda
  FROM
  (
    SELECT position, champion_id, player_id
    FROM participant
    WHERE position = 'DUO_CARRY' OR position = 'DUO_SUPPORT' OR position = 'JUNGLE' OR position = 'MID' OR position = 'TOP'
  ) P,
  (
    SELECT player_id, (kills+assists)/deaths AS kda
    FROM stat
    WHERE deaths != 0
  ) S
  WHERE P.player_id = S.player_id
) MA
GROUP BY P.position
WHERE AL.position= MA.position AND AL.kda = MA.kda AND C.champion_id = AL.champion_id
ORDER BY FIELD(MA.position, 'DUO_CARRY', 'DUO_SUPPORT', 'JUNGLE', 'MID', 'TOP')
;
```

Query result:

```
mysql> source hw1-8.sql
+-----+-----+-----+
| position | champion_name | kda    |
+-----+-----+-----+
| DUO_CARRY | Jhin          | 46.0000 |
| DUO_SUPPORT | Janna        | 48.0000 |
| JUNGLE    | Evelynnn     | 44.0000 |
| MID       | Brand         | 45.0000 |
| TOP       | Pantheon      | 40.0000 |
+-----+-----+-----+
5 rows in set (38.01 sec)
```

8. SQL Screenshot:

```
SELECT DISTINCT champion_name
FROM champ
WHERE champion_id NOT IN (
    SELECT T.champion_id
    FROM teamban T, match_info M
    WHERE M.version LIKE '7.7%' AND M.match_id = T.match_id
)
ORDER BY champion_name;
```

Query result:

```
mysql> mysql> source hw1-9.sql
+-----+
| champion_name |
+-----+
| Kayn          |
| Ornn          |
| Rakan         |
| RekSai        |
| Sion          |
| Xayah         |
+-----+
6 rows in set (0.42 sec)
```

9. SQL Screenshot:

```

SELECT SUBSTRING_INDEX(M.version, '.',2) AS `version`
    , SUM(CASE
            WHEN S.win = 1 THEN 1
            WHEN S.win = 0 THEN 0
            END) AS win_cnt
    , SUM(CASE
            WHEN S.win = 0 THEN 1
            WHEN S.win = 1 THEN 0
            END) AS lose_cnt
    , SUM(CASE
            WHEN S.win = 1 THEN 1
            WHEN S.win = 0 THEN 0
            END) / COUNT(*) AS win_ratio
FROM match_info M, (
    SELECT player_id, win
    FROM stat
) S,
(
    SELECT L.player_id, L.match_id
    FROM
    (
        SELECT P.player_id, P.match_id, P.player
        FROM
        (
            SELECT champion_id
            FROM champ
            WHERE champion_name = 'Lee Sin'
        ) C, participant P
        WHERE C.champion_id = P.champion_id
    ) L,
    (
        SELECT P.player_id, P.match_id, P.player
        FROM
        (
            SELECT champion_id
            FROM champ
            WHERE champion_name = 'Teemo'
        ) C, participant P
        WHERE C.champion_id = P.champion_id
    ) T
    WHERE L.match_id = T.match_id AND (
        1 = CASE
            WHEN L.player BETWEEN 1 AND 5 AND T.player BETWEEN 1 AND 5 THEN 1
            WHEN L.player BETWEEN 6 AND 10 AND T.player BETWEEN 6 AND 10 THEN 1
            ELSE 0
        END
    )
) PID
WHERE PID.player_id = S.player_id AND PID.match_id = M.match_id
GROUP BY SUBSTRING_INDEX(M.version, '.',2);

```

Query result:


```
mysql> source hw1-18.sql
```

version	win_cnt	lose_cnt	win_ratio
4.10	2	1	0.6667
4.12	0	1	0.0000
4.15	1	1	0.5000
4.17	0	1	0.0000
4.18	0	1	0.0000
4.19	0	1	0.0000
4.21	1	1	0.5000
4.9	1	0	1.0000
5.1	1	2	0.3333
5.12	1	0	1.0000
5.13	0	1	0.0000
5.15	0	1	0.0000
5.19	1	0	1.0000
5.20	2	0	1.0000
5.21	0	2	0.0000
5.24	1	1	0.5000
5.5	1	0	1.0000
5.6	0	1	0.0000
5.7	1	0	1.0000
6.1	0	1	0.0000
6.13	1	0	1.0000
6.14	1	0	1.0000
6.18	1	1	0.5000
6.19	1	0	1.0000
6.2	1	1	0.5000
6.20	3	2	0.6000
6.21	0	2	0.0000
6.22	2	1	0.6667
6.23	3	2	0.6000
6.24	4	3	0.5714
6.5	1	0	1.0000
6.6	0	1	0.0000
6.8	1	0	1.0000
6.9	1	1	0.5000
7.10	282	304	0.4812
7.2	2	1	0.6667
7.3	0	1	0.0000
7.4	1	1	0.5000
7.5	2	2	0.5000
7.6	2	5	0.2857
7.7	32	29	0.5240
7.8	210	237	0.4098
7.9	527	464	0.5318

```
43 rows in set (22.95 sec)
```

10. SQL Screenshot:

```

SELECT PL.champion_name, SUM(CASE
    WHEN S1.win = 1 THEN 1
    WHEN S1.win = 0 THEN 0
    END) / COUNT(S1.win) AS win_ratio, AVG((S1.kills + S1.assists) / S1.deaths) AS self_kda,
    AVG(S1.gold_earned) AS self_avg_gold, PL.enemy_champ_name, AVG((S2.kills + S2.assists) / S2.deaths) AS enemy_kda,
    AVG(S2.gold_earned) AS enemy_avg_gold, COUNT(*) AS battle_record
FROM stat S1, stat S2, (
    SELECT P.champion_id, P.player_id, C.champion_name, R.enemy_id, R.enemy_champ_name
    FROM participant P, champ C,
    (
        SELECT P.player, P.match_id, P.player_id AS enemy_id, C.enemy_champ_name
        FROM
        (
            SELECT champion_id, champion_name AS enemy_champ_name
            FROM champ
            WHERE champion_name = 'Renekton'
        ) C, participant P
        WHERE C.champion_id = P.champion_id AND P.position = 'TOP'
    ) R
    WHERE P.match_id = R.match_id AND C.champion_id = P.champion_id AND (
        1 = CASE
            WHEN P.player BETWEEN 1 AND 5 AND R.player BETWEEN 6 AND 10 THEN 1
            WHEN P.player BETWEEN 6 AND 10 AND R.player BETWEEN 1 AND 5 THEN 1
            ELSE 0
        END
    )
) PL
WHERE S1.player_id = PL.player_id AND S2.player_id = PL.enemy_id
GROUP BY PL.champion_name, PL.enemy_champ_name
HAVING COUNT(*) > 100
ORDER BY SUM(CASE
    WHEN S1.win = 1 THEN 1
    WHEN S1.win = 0 THEN 0
    END) / COUNT(S1.win) DESC
LIMIT 5;

```

Query result:

champion_name	win_ratio	self_kda	self_avg_gold	enemy_champ_name	enemy_kda	enemy_avg_gold	battle_record
Karthus	0.5693	3.29208346	12300.2993	Renekton	2.18649302	11918.0511	137
Fiddlesticks	0.5690	3.35047286	10848.6465	Renekton	2.74150964	11824.2121	297
Ivern	0.5667	5.55251289	10095.7881	Renekton	2.65322662	11586.7952	420
Teemo	0.5556	2.64719751	12100.0920	Renekton	2.46523929	11821.7240	576
KogMaw	0.5534	2.97679645	12185.1529	Renekton	2.56089202	11668.3689	412

5 rows in set, 7577 warnings (20.87 sec)

11. SQL Screenshot:

```

SELECT (CASE
    WHEN P.ss1 = 'Flash' THEN P.ss2
    WHEN P.ss2 = 'Flash' THEN P.ss1
    END) AS s,
    SUM(CASE
        WHEN S.win = 1 THEN 1
        WHEN S.win = 0 THEN 0
        END) / COUNT(S.win) AS win_ratio, COUNT(*) AS cnt
FROM stat S, participant P
WHERE ((P.ss1 = 'Flash' and P.ss2 = 'Ignite') OR (P.ss2 = 'Flash' and P.ss1 = 'Ignite')) AND S.player_id = P.player_id
GROUP BY (CASE
    WHEN P.ss1 = 'Flash' THEN P.ss2
    WHEN P.ss2 = 'Flash' THEN P.ss1
    END);

SELECT (CASE
    WHEN P.ss1 = 'Flash' THEN P.ss2
    WHEN P.ss2 = 'Flash' THEN P.ss1
    END) AS s,
    SUM(CASE
        WHEN S.win = 1 THEN 1
        WHEN S.win = 0 THEN 0
        END) / COUNT(S.win) AS win_ratio, COUNT(*) AS cnt
FROM stat S, participant P
WHERE ((P.ss1 = 'Flash' and P.ss2 = 'Teleport') OR (P.ss2 = 'Flash' and P.ss1 = 'Teleport')) AND S.player_id = P.player_id
GROUP BY (CASE
    WHEN P.ss1 = 'Flash' THEN P.ss2
    WHEN P.ss2 = 'Flash' THEN P.ss1
    END);

```

Query result:

```
mysql> source 11.sql
+-----+-----+-----+
| s      | win_ratio | cnt    |
+-----+-----+-----+
| Ignite | 0.5073    | 324125 |
+-----+-----+-----+
1 row in set (14.91 sec)

+-----+-----+-----+
| s      | win_ratio | cnt    |
+-----+-----+-----+
| Teleport | 0.4978    | 346424 |
+-----+-----+-----+
1 row in set (14.96 sec)
```

Description:

因為可能是 ss1 是 flash 或者是 ss2 是 flash，所以必須分成兩個討論，把兩個可能 or 起來之後就可以得到兩個不同狀態的表格，可以知道在”flash+ ignite” 的時候結果勝率比較大。

12. Description:

在 stat 裡面有很多的 kills 紀錄，是跟連殺率有關的，所以我整理了連殺率來看哪些角色的連殺技能比較強。

對一個角色而言，連殺的計算方式如下:

連殺率

$$= \frac{\sum_{\text{玩家}} (\text{doublekills} \times 2 + \text{triplekills} \times 3 + \text{quadrakills} \times 4 + \text{pentakills} \times 5)}{\sum_{\text{玩家}} \text{kills}}$$

SQL Screenshot

```
SELECT C.champion_name, KK.position, KK.fast_kill_rate
FROM champ C,
(
  SELECT SUM(S.doublekills* 2+ S.triplekills* 3+ S.quadrakills* 4+ S.pentakills* 5)/ SUM(S.kills) AS fast_kill_rate, P.position, P.champion_id
  FROM stat S, participant P
  WHERE P.player_id = S.player_id AND S.kills != 0
  GROUP BY P.position, P.champion_id
  ORDER BY SUM(S.doublekills* 2+ S.triplekills* 3+ S.quadrakills* 4+ S.pentakills* 5)/ SUM(S.kills)
) KK
WHERE C.champion_id = KK.champion_id AND KK.fast_kill_rate> 0.4;
```

Query result:

```
mysql> source 12.sql
```

champion_name	position	fast_kill_rate
Jinx	DUO_CARRY	0.4001
Vayne	DUO_CARRY	0.4043
Graves	DUO_SUPPORT	0.4074
Ahri	SOLO	0.4167
Katarina	JUNGLE	0.4180
Vayne	NONE	0.4227
Kassadin	DUO_SUPPORT	0.4333
Tryndamere	DUO_CARRY	0.4348
Orianna	SOLO	0.4444
Nidalee	NONE	0.4545
Lissandra	DUO_CARRY	0.4583
Maokai	DUO_CARRY	0.4634
KhaZix	DUO_CARRY	0.4667
Yasuo	NONE	0.4667
Master Yi	JUNGLE	0.4778
Katarina	SOLO	0.5000
Kalista	NONE	0.5000
Master Yi	SOLO	0.5000
Karthus	SOLO	0.5000
Hecarim	DUO_CARRY	0.5000
Master Yi	MID	0.5008
Master Yi	TOP	0.5241
Nidalee	DUO_CARRY	0.5417
Katarina	DUO_CARRY	0.5504
Swain	SOLO	0.6250
Master Yi	DUO_CARRY	0.6291
Evelynn	SOLO	0.6667
Syndra	NONE	0.6667
Tahm Kench	DUO_CARRY	0.6667
Udyr	DUO	0.8333
Renekton	DUO_CARRY	0.9000
Irelia	SOLO	1.2500

32 rows in set (28.40 sec)