

DataBase HW2

0416213 林彥傑

- 所有sql, result的csv格式可至[此github連結](#)

Environment

MySQL 5.7.17 with MySQL Workbench 6.3.9 CE

前置作業(方便以下的Query使用，避免sql太長)

- 建 3 個 View : MaxSalaries, TeamSalaries, WSWinnerTeam
- MaxSalaries
 - 儲存每年每隊的最高薪球員的playerID
 - columns : (yearID, teamID, lgID, playerID, salary)

```
1 create view `MaxSalaries` as (  
2 select  
3     S.yearID, S.teamID, S.lgID, S.playerID, S.salary  
4 from  
5     Salaries S, (  
6         select S2.yearID, S2.teamID, max(S2.salary) as max_sal  
7         from Salaries S2  
8         group by S2.yearID, S2.teamID  
9     ) as temp  
10 where  
11     S.yearID = temp.yearID and  
12     S.teamID = temp.teamID and  
13     S.salary = temp.max_sal  
14 );
```

- TeamSalaries
 - 儲存大聯盟歷年各隊的團隊薪資總和
 - columns : (yearID, lgID, teamID, franchID, total_sal)

```
1 CREATE VIEW `TeamSalaries` AS (  
2     SELECT  
3         T.yearID, T.lgID, T.teamID, T.franchID, S.total_sal  
4     FROM  
5         Teams T, (  
6             SELECT t.yearID, t.teamID, sum(t.salary) AS total_sal  
7             FROM Salaries t  
8             GROUP BY t.yearID, t.teamID  
9         ) AS S  
10    WHERE  
11        T.yearID = S.yearID AND  
12        T.teamID = S.teamID  
13 );
```

- WSWinnerTeam
 - 儲存每年的世界大賽冠軍球隊
 - columns : (yearID, lgID, teamID, franchID)

```
1 create view `WSWinnerTeam` as (  
2 select  
3     T.yearID, T.lgID, T.teamID, T.franchID  
4 from  
5     Teams T
```

```
6     where
7         T.WSWin = 'Y'
8     );
```

- [view sql檔案](#)

Query

Q1: How many players are recorded in the database ?

```
1 SELECT COUNT(*)
2 FROM Master;
```

- result : 19105

Q2: How many countries do the players come from ?

```
1 SELECT count(distinct M.birthCountry)
2 FROM Master M;
```

- result : 54

Q3: List yearID and playerID who hit the home run between 50 and 60.

```
1 SELECT B.yearID, B.playerID
2 FROM Batting B
3 WHERE B.HR BETWEEN 50 AND 60;
```

- result: <https://github.com/jay16213/Database-practice/blob/master/q3.csv>

Q4: How many parks whose name contains 'Stadium' are recorded?

```
1 SELECT COUNT(*)
2 FROM Parks P
3 WHERE P.`park.name` LIKE '%Stadium%';
```

- result : 29

Q5: How long is the player's career ?

```
1 SELECT AVG(M.finalGame - M.debut)
2 FROM Master M;
```

- result : 4.668254383669197

Q6: Find the players' schools who made their debut at May 2000.

```
1 SELECT DISTINCT C.schoolID
2 FROM CollegePlaying C, Master M
3 WHERE
4     M.playerID = C.playerID AND
5     M.debut BETWEEN '2000-05-01' AND '2000-05-31';
```

- result : <https://github.com/jay16213/Database-practice/blob/master/q6.csv>

Q7: How much is the highest and average salary in each year ?

- 平均薪水

```
1 SELECT
2     S.yearID, AVG(S.salary)
3 FROM
4     Salaries S
5 GROUP BY
6     S.yearID;
```

- result: <https://github.com/jay16213/Database-practice/blob/master/q7-avg.csv>
- 最高薪水

```
1 SELECT
2     S.yearID, MAX(S.salary)
3 FROM
4     Salaries S
5 GROUP BY
6     S.yearID;
```

- result: <https://github.com/jay16213/Database-practice/blob/master/q7-max.csv>

Q8: In the previous question, who are those highest paid players in each year ?

```
1 SELECT
2     S.yearID, S.salary, M.nameFirst, M.nameLast
3 FROM
4     Master M, (
5         SELECT T.yearID , T.salary , T.playerID
6         FROM Salaries T, (
7             SELECT T2.yearID, MAX(T2.salary) as max_sal
8             FROM Salaries T2
9             GROUP BY T2.yearID
10        ) as T2
11        WHERE
12            T.yearID = T2.yearID AND
13            T.salary = T2.max_sal
14    ) as S
15 WHERE
16     M.playerID = S.playerID
17 ORDER BY
18     S.yearID ASC;
```

- result: <https://github.com/jay16213/Database-practice/blob/master/q8.csv>

Q9: List the salary and the player that is the highest paid player whose team win the World Series and the highest paid player in Boston Red Sox duaring the Curse of the Bambino.

- 貝比魯斯魔咒: 西元1920 - 2004年
- 但此資料庫沒有西元1984年(含)之前的球員薪資資料
- 因此查詢期間為西元1985 - 2004年

```
1 SELECT
2     S.yearID, S.teamID AS WS_winner,
3     CONCAT(M.nameFirst, ' ', M.nameLast) AS Name, S.salary,
4     CONCAT(M2.nameFirst, ' ', M2.nameLast) AS BOS_Player_Name, S2.salary
5 FROM
6     WSWinnerTeam W, MaxSalaries S, MaxSalaries S2, Master M, Master M2
```

```

7 WHERE
8     W.yearID = S.yearID AND S.yearID = S2.yearID AND
9     M.playerID = S.playerID AND
10    M2.playerID = S2.playerID AND
11    W.teamID = S.teamID AND
12    S2.teamID = 'BOS' AND
13    W.yearID BETWEEN 1985 AND 2004
14 ORDER BY
15     S.yearID;

```

- result: <https://github.com/jay16213/Database-practice/blob/master/q9.csv>

Q10: Why Ichiro Suzuki is a good player? Answer by your own view.

- 為求公平, 本題在大聯盟出賽平均只計當年度出賽超過81場之選手出賽數, 打擊數據平均只計當季打數超過486個的打者的數據(162場 x 每場平均3打數)
- 出賽穩定無傷痛
 - 鈴木一郎每季出賽數皆超過MLB每季平均20-30場, 生涯16年平均出賽數約為147.1場

```

1 SELECT
2     B.yearID, AVG(B.G) AS Suz_G, AVG(B2.G) AS Avg_G
3 FROM
4     Batting B, Batting B2
5 WHERE
6     B.playerID = 'suzukic01' AND
7     B.yearID = B2.yearID AND
8     B2.G >= 81
9 GROUP BY
10    B.yearID;

```

- result: https://github.com/jay16213/Database-practice/blob/master/q10-game_played.csv
- 打擊成績出色
 - 2001 - 2010連續10年打擊率超過3成
 - 2001 - 2010連續10年單季安打數超過200支
 - 大聯盟單季最多安打數紀錄保持人(2004年單季262支), 至今無人能破
 - 歷年單季總三振數為平均的1/2 - 2/3之間, 顯示其選球能力及與投手纏鬥之能力

```

1 SELECT
2     B.yearID,
3     AVG(B.H) AS Avg_H, Suz.H AS Suz_H,
4     AVG(B.SO) AS Avg_SO, Suz.SO AS Suz_SO,
5     Suz.A AS Suz_Avg
6 FROM
7     Batting B, (
8         SELECT
9             B2.yearID,
10            SUM(B2.H) AS H,
11            SUM(B2.H)/SUM(B2.AB) AS A,
12            SUM(B2.SO) AS SO
13        FROM
14            Batting B2
15        WHERE
16            B2.playerID = 'suzukic01'
17        GROUP BY
18            B2.yearID
19    ) AS Suz
20 WHERE
21     B.yearID = Suz.yearID AND
22     B.ab >= 486
23 GROUP BY
24     B.yearID;

```

- result: <https://github.com/jay16213/Database-practice/blob/master/q10-batting.csv>

Bonus: 找出歷年來團隊薪資最便宜 & 最貴的世界大賽冠軍球隊

```
1  (  
2  SELECT  
3      Ws.yearID, Ws.teamID, Ws.franchID, Ws.total_sal  
4  FROM  
5      TeamSalaries Ws, (  
6          SELECT  
7              MIN(T.total_sal) as MIN_team_sal  
8          FROM  
9              WSWinnerTeam W, TeamSalaries T  
10         WHERE  
11             W.yearID = T.yearID AND  
12             W.teamID = T.teamID  
13     ) AS Temp  
14 WHERE  
15     Ws.total_sal = Temp.MIN_team_sal  
16 )  
17 UNION  
18 (  
19 SELECT  
20     Ws.yearID, Ws.teamID, Ws.franchID, Ws.total_sal  
21 FROM  
22     TeamSalaries Ws, (  
23         SELECT  
24             MAX(T.total_sal) as MAX_team_sal  
25         FROM  
26             WSWinnerTeam W, TeamSalaries T  
27         WHERE  
28             W.yearID = T.yearID AND  
29             W.teamID = T.teamID  
30     ) AS Temp  
31 WHERE  
32     Ws.total_sal = Temp.MAX_team_sal  
33 );
```

- result: <https://github.com/jay16213/Database-practice/blob/master/bonus.csv>

yearID	teamID	franchID	total_sal
1987	MIN	MIN	6397500
2009	NYA	NYN	201449189

- 若不考慮通貨膨脹等問題，則由上述結果可知，1987年的馬林魚僅花費約640萬美金就贏得當年的世界大賽冠軍，2009年的洋基隊則花了2億多美金才贏得冠軍，是有薪資紀錄以來最豪華的奪冠球隊
- 查詢MaxSalaries，可知2009年洋基隊年薪最高的A-ROD當年年薪為3300萬美金，是640萬美金的5倍有餘