

INFSCI 2710

Database Management

Data Mart Assignment

By

Sai Supreetha Varanasi (sav52@pitt.edu)

Vaishnavi Deshpande (vad28@pitt.edu)



Fall 2017

School of Computing and Information
University of Pittsburgh

Table of Contents

Deliverable 1:	3
1. ER Diagram of star schema:	3
2. DDL scripts for the fact table and dimension tables:	3
Deliverable 2:	6
1. What are the top 5 player birth countries grouped by Descending salary?	6
2. Who are the top 5 overall highest paid player of all time grouped by Descending salary?	7
3. Create a view that sums salaries and groups records by if players are inducted into the hall of fame.	7

Deliverable 1:

1. ER Diagram of star schema:

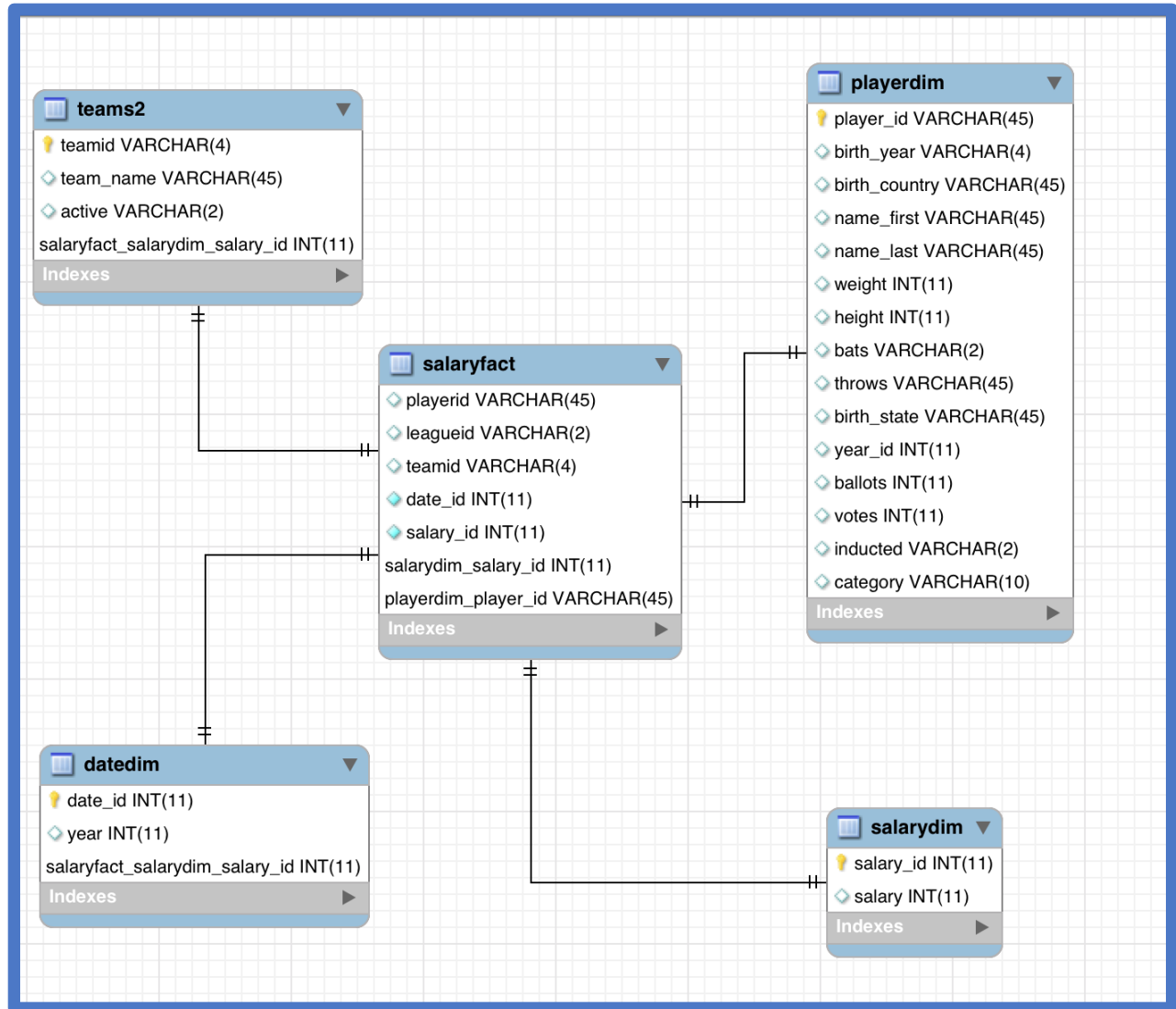


Fig 1: ER Diagram of star schema

2. DDL scripts for the fact table and dimension tables:

```
CREATE SCHEMA `salary` ;
```

```
CREATE TABLE `salary`.`hof2` (
```

```
`playerid` VARCHAR(45) NOT NULL,  
`year_id` INT NULL,  
`ballots` INT NULL,  
`votes` INT NULL,  
`inducted` VARCHAR(2) NULL,  
`category` VARCHAR(10) NULL,  
PRIMARY KEY (`playerid`));
```

```
CREATE TABLE `salary`.`player2` (  
  `player_id` VARCHAR(45) NOT NULL,  
  `birth_year` VARCHAR(4) NULL,  
  `birth_country` VARCHAR(45) NULL,  
  `birth_city` VARCHAR(45) NULL,  
  `name_first` VARCHAR(45) NULL,  
  `name_last` VARCHAR(45) NULL,  
  `weight` INT NULL,  
  `height` INT NULL,  
  `bats` VARCHAR(2) NULL,  
  `throws` VARCHAR(45) NULL,  
  `birth_state` VARCHAR(45) NULL,  
  PRIMARY KEY (`player_id`));
```

```
CREATE TABLE `salary`.`salary2` (  
  `playerid` VARCHAR(45) NOT NULL,  
  `leagueid` VARCHAR(2) NULL,  
  `teamid` VARCHAR(4) NULL,  
  `year` INT NULL,  
  `salary` INT NULL,  
);
```

```
CREATE TABLE `salary`.`teams2` (  
  `teamid` VARCHAR(4) NOT NULL,  
  `team_name` VARCHAR(45) NULL,
```

```
`active` VARCHAR(2) NULL,  
PRIMARY KEY (`teamid`));
```

```
CREATE TABLE playerdim as  
(select player_id, birth_year, birth_country,  
name_first, name_last,  
weight, height, bats, throws, birth_state, year_id,  
ballots, votes, inducted, category  
from player2 LEFT outer join hof2  
on player2.player_id = hof2.playerid);
```

```
CREATE TABLE `salary`.`datedim` (  
`date_id` INT NOT NULL AUTO_INCREMENT,  
`year` INT(11) NULL,  
PRIMARY KEY (`date_id`));
```

```
CREATE TABLE `salary`.`salarydim` (  
`salary_id` INT(11) NOT NULL AUTO_INCREMENT,  
`salary` INT(11) NULL,  
PRIMARY KEY (`salary_id`));
```

```
ALTER TABLE datedim AUTO_INCREMENT=1;  
insert into datedim(year)  
select distinct year from salary2;
```

```
#select * from datedim;
```

```
ALTER TABLE salarydim AUTO_INCREMENT=1;  
insert into salarydim(salary)  
select distinct salary from salary2;
```

```
#select * from salarydim;
```

```
CREATE TABLE salaryfact as
(select salary2.playerid, leagueid, teamid, date_id,
salary_id
from salary2, salarydim, datedim
where
salary2.year=datedim.year
and
salary2.salary = salarydim.salary);
```

Deliverable 2:

1. What are the top 5 player birth countries grouped by Descending salary?

```
create view player_salary as
(select playerid, salary
from salaryfact, salarydim
where salaryfact.salary_id=salarydim.salary_id
);
```

```
select playerid, birth_state, birth_country, salary
from playerdim, player_salary
where player_id=playerid
order by salary desc
limit 5;
```

playerid	birth_state	birth_country	salary
verlaju01	VA	USA	28000000
howarry01	MO	USA	25000000
leec102	AR	USA	25000000
howarry01	MO	USA	25000000
leec102	AR	USA	25000000

2. Who are the top 5 overall highest paid player of all time grouped by Descending salary?

```
select playerid, name_first, name_last, salary
from playerdim, player_salary
where player_id=playerid
order by salary desc
limit 5;
```

playerid	name_first	name_last	Salary
verlaju01	Justin	Verlander	28000000
howarry01	Ryan	Howard	25000000
leec102	Cliff	Lee	25000000
howarry01	Ryan	Howard	25000000
leec102	Cliff	Lee	25000000

3. Create a view that sums salaries and groups records by if players are inducted into the hall of fame.

```
create view sum_salary as
(
    select sum(salary) as total_salary, inducted
    from hof2, player_salary
    where hof2.playerid=player_salary.playerid
    group by inducted
);
select * from sum_salary;
```

total_salary	inducted
98329370	N
43190001	Y