**Question 2**

**### dim\_store**

create table db\_nsaraf.dim\_store as

select address, district, postal\_code, country, manager\_staff\_id, store\_id as store\_key

from sakila.address as a , sakila.country as c , sakila.store as s, city as ct

where s.address\_id = a.address\_id and a.city\_id = ct.city\_id and ct.country\_id = c.country\_id;

**#### dim\_staff**

create table db\_nsaraf.dim\_staff as

select first\_name, last\_name, staff\_id as staff\_key, staff.store\_id, active

from sakila.staff, sakila.store

where staff.store\_id = store.store\_id;

**#### dim\_film**

create table db\_nsaraf.dim\_film as

select film\_id as film\_key, film.\* from film;

####Alter all tables to declare film\_key, staff\_key and store\_key as PK or add an extra PK called film\_key, staff\_key, store\_key to each dimension table, AUTO INCREMENT, PK (without renaming store\_id, staff\_id and film\_id).

#technically you need to create separate dimensional keys instead of simply renaming store\_id as store\_key. But for now I will leave that upto you to know and follow for staff and film similar to the code below.

# alter table dim\_store add `store\_key` int AUTO\_INCREMENT PRIMARY KEY

###############Time Dimension

Use db\_nsaraf;

drop table T;

create table T as select film\_id from sakila.film;

# Use the below line of INSERT code repeatedly if you want to create a table with more rows.

#insert into T select film\_id from T;

# set the timespan based on the data in your transactional database

# select max(return\_date), min(return\_date), max(rental\_date), min(rental\_date) from sakila.rental;

# time span

SET @d0 = "2005-05-02";

SET @d1 = "2006-02-28";

SET @date = date\_sub(@d0, interval 1 day);

# set up the time dimension table

DROP TABLE IF EXISTS time\_dimension;

CREATE TABLE `time\_dimension` (

`date` date DEFAULT NULL,

`id` int NOT NULL,

`y` smallint DEFAULT NULL,

`m` smallint DEFAULT NULL,

`d` smallint DEFAULT NULL,

`yw` smallint DEFAULT NULL,

`w` smallint DEFAULT NULL,

`q` smallint DEFAULT NULL,

`wd` smallint DEFAULT NULL,

`m\_name` char(10) DEFAULT NULL,

`wd\_name` char(10) DEFAULT NULL,

PRIMARY KEY (`id`)

);

# populate the table with dates

INSERT INTO time\_dimension

SELECT @date := date\_add(@date, interval 1 day) as date,

# integer ID that allows immediate understanding

date\_format(@date, "%Y%m%d") as id,

year(@date) as y,

month(@date) as m,

day(@date) as d,

date\_format(@date, "%x") as yw,

week(@date, 3) as w,

quarter(@date) as q,

weekday(@date)+1 as wd,

monthname(@date) as m\_name,

dayname(@date) as wd\_name

FROM T

WHERE date\_add(@date, interval 1 day) <= @d1

ORDER BY date

;

####Truncating the time dimension to the weekly level from the above daily level

create table td as

select min(date) as wkbegin, y, m, yw, w from time\_dimension

group by y, m, w order by wkbegin;

**#### the fact table -- calculating the Number of Rental and Average Rental Duration for each combination of Store, Film and Staff**

create table fct\_rental as

select count(\*) as NumRentals,

avg((return\_date - rental\_date)/60/60/24) as AvgDuration, #or any other way you choose to do it

film\_key, store\_key, staff\_key, y as year, w as week

### here keep the Sakila ERD in view so that the joins are correct

from sakila.rental as r, sakila.inventory as i, sakila.staff as st, sakila.store as so, sakila.film as f, td

where

f.film\_id = i.film\_id

and i.inventory\_id = r.inventory\_id

and r.staff\_id = st.staff\_id

and so.store\_id = st.store\_id

### and now link with time\_dimension

and td.wkbegin <= rental\_date and date\_add(wkbegin, interval 7 day)> rental\_date

group by film\_key, store\_key, staff\_key, y, w ### here you need to simply state the dimension keys and time-based aggregation -- aggregating all sales for each week for each year for each film, for each store, for each staff member

order by y, w;

**Question 3**

**### dim\_store**

create table db\_nsaraf.dim\_store as

select address, district, postal\_code, country, manager\_staff\_id, store\_id as store\_key

from sakila.address as a , sakila.country as c , sakila.store as s, city as ct

where s.address\_id = a.address\_id and a.city\_id = ct.city\_id and ct.country\_id = c.country\_id;

**### dim\_actor**

Create table db\_nsaraf.dim\_actor as

Select actor\_id as actor\_key, \*

From sakila.actor;

**#### the fact table -- calculating the Number of Rental and Average Rental Duration for each combination of Store and Actor**

select count(\*) as NumRentals,

avg((return\_date - rental\_date)/60/60/24) as AvgDuration,

so.store\_id as store\_key, a.actor\_id as actor\_key ### dimension keys

### here keep the Sakila ERD in view to identify the tables you need and the joins

from film as f, film\_actor as fa, actor as a, inventory as i, rental as r, staff as st, store as so

where f.film\_id = fa.film\_id and fa.actor\_id = a.actor\_id

and i.film\_id = f.film\_id and i.rental\_id = r.rental\_id

and r.staff\_id = st.staff\_id and st.store\_id = so.store\_id

and r.rental\_date = time\_dimension.date

## Here I chose to use the date-level time dimension unlike in the Q2 solution

group by store\_key, actor\_key; ### here you need to simply state the dimension keys