week 3 assignment

Submitted by Maryam Heidari

ALY 6030-71430 – Data Warehousing and SQL

Northeastern University

Date of Submission: 21 November 2019

Introduction:

For this assignment, we are going to make a model by using free on-line tool ERDPluse
And answer the question of the story by SQL code.

Analysis:

1. Identify Dimensions (first three only) from the two-dimension tables.

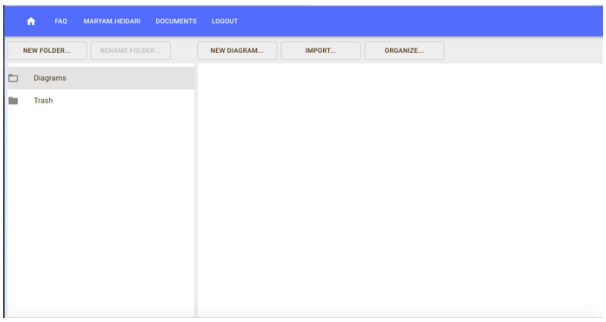
For the business table, the first three dimensions are ims_org_id (the ID of hospital or business), business name (name of the hostpital or hospital) and license beds (number of beds they allow to have).

For the bed type table, all dimensions are bed Id (the ID of bed), bed code (each bed, base n their usage, has a code) and bed desc (where each bed type is going to use, which department).

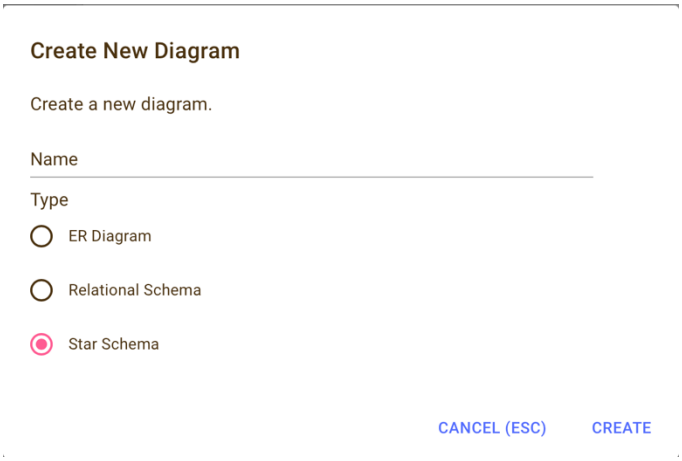
2. Identify the Facts (first three only) from the single fact table.

For the fact table, the first three facts are number of bed, number of license beds that they have and number of staffed beds.

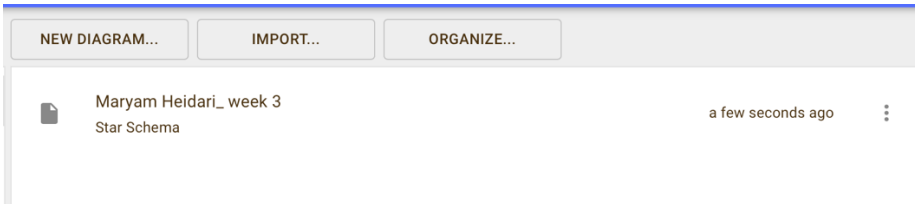
3. In order to skate out a schema star, I used free online tool ERDPlus. After singing up, I have empty table that I can make my model.



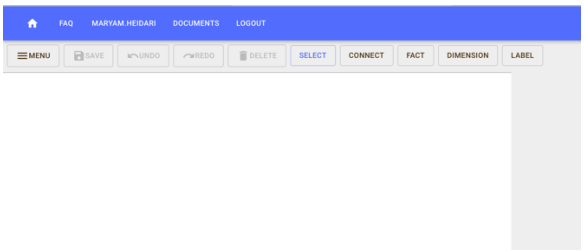
Click on new diagram, and we have below window



I chose the star schema and give it the name and create it. The result is:



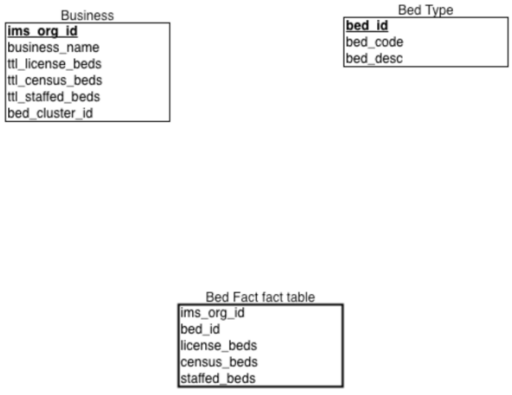
Now I am open it.



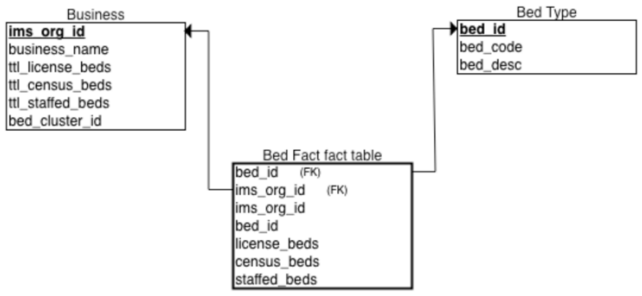
Now I have empty schema that I can make my model on it.

I clicked on Dimension, give it name to may table. In column part, I clicked on add to add my columns. After putting the name and data type of each columns, I chose primary key. The result is:

After that, I clicked on Fact option and made a fact table. And the result is:

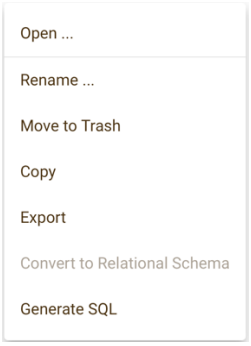


Then I clicked on connect option and made relationships between tables.

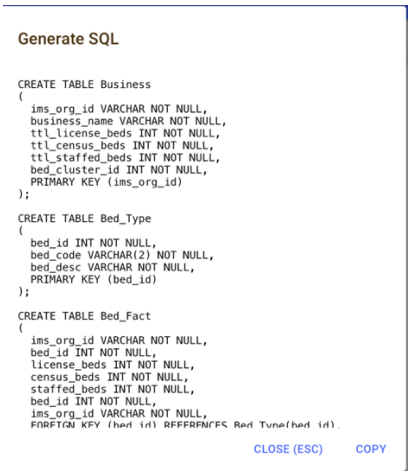


Now, I used this schema as a database in mySQL. In order to do that, I clicked on the dots in below picture

and chose the Generate SQL option



And now, I have a SQL code to use in mySQL.



4 to 7. In mySQL, I made a new schema and add the name of my schema before name of tables in the above code. The code is

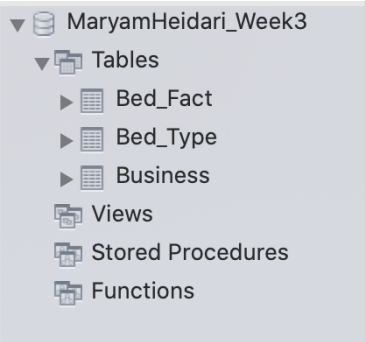
```sql
#create schema
CREATE SCHEMA `MaryamHeidari_Week3`;

#paste the code of star schema
CREATE TABLE `MaryamHeidari_Week3`.`Business`
(
  ims_org_id VARCHAR(45) NOT NULL,
  business_name VARCHAR(45) NOT NULL,
  ttl_license_beds INT NOT NULL,
  ttl_census_beds INT NOT NULL,
  ttl_staffed_beds INT NOT NULL,
  bed_cluster_id INT NOT NULL,
  PRIMARY KEY (ims_org_id)
);

CREATE TABLE `MaryamHeidari_Week3`.`Bed_Type`
(
  bed_id INT,
  bed_code VARCHAR(2),
  bed_desc VARCHAR(45),
  PRIMARY KEY (bed_id)
);
```

```sql
CREATE TABLE `MaryamHeidari_Week3`.`Bed_Fact`
(
  ims_org_id VARCHAR(45) NOT NULL,
  bed_id INT NOT NULL,
  license_beds INT NOT NULL,
  census_beds INT NOT NULL,
  staffed_beds INT NOT NULL,
  FOREIGN KEY (bed_id) REFERENCES Bed_Type(bed_id),
  FOREIGN KEY (ims_org_id) REFERENCES business(ims_org_id)
);
```

And the result is:



I filled each of tables by right click on them, chose the import data option and chose their table, except bed_type table, I imported its data by code which is:

```
#fill the table
INSERT INTO MaryamHeidari_Week3.Bed_Type (bed_id, bed_code, bed_desc)
VALUES (1,'BU','Burn'),
(2,'CC','CCU'),
(3,'DE','Detox ICU'),
(4,'IC','ICU'),
(5,'MS','Med/Surg'),
(6,'NE','NeoNatal ICU'),
(7,'NU','Nursery'),
(8,'NF','Nursing Facility'),
(9,'OT','Other'),
(10,'PD','Pediatric ICU'),
(11,'PR','Premature ICU'),
(12,'PS','Psychiatric'),
(13,'PI','Psych ICU'),
(14,'RH','Rehabilitation'),
(15,'SI','SICU'),
(16,'SN','Skilled Nursing'),
(17,'SP','Special Care'),
(18,'TO','Total'),
(19,'TR','Trauma ICU'),
(20,'DD','Developmental Disability');
```

Then, in order to make a table that mentioned in pert 7.

I select the variable in mentioned in that table, and I used the CASE… THEN function to make a column for ICU and SICU, I have to say if there is no ICU or SCIU put the zero in that cell

because if I filled it with null or NA I would have trouble for sorting table. Moreover, I used the left join to join all of my tables to fact beds table, then group by business ID, and finally order by ICU result and limit the final table to 10 rows. I repeat these steps for license, census and staffed.

For License:

```
#License
SELECT a.ims_org_id, c.business_name,
(CASE
WHEN b.bed_desc like '%ICU%'  THEN (a.license_beds)
ELSE 0
END) as License_ICU,
(CASE
WHEN b.bed_desc like '%SICU%'  THEN (a.license_beds)
ELSE 0
END) as License_SICU
FROM maryamheidari_week3.bed_fact a
left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id
group by a.ims_org_id
order by License_ICU desc
LIMIT 10;
```

ims_org_id	business_name	License_ICU	License_SICU
INS00005214	Saint Marys Hospital	264	0
INS00001172	Saints Mary and Elizabeth Hospital	174	0
INS00001639	Emory University Hospital	169	0
INS00011224	G Werber Bryan Psychiatric Hospital	139	0
INS00003552	Childrens Hospital of Wisconsin, Inc	138	0
INS00005612	Brigham and Womens Hospital	138	138
INS00005078	Jewish Hospital	136	0
INS00004254	Heartland Regional Medical Center	135	0
INS00011312	The Devereux Texas Treatment Network	132	0
INS00006375	Texas Health Presbyterian Hospital Dallas	131	0

In above table, we can see you many beds license each top ten hospital has for ICU and SICU. The only hospital how has a license for SICU is Brigham and women's hospital which is 138.

For Census

```
19    #census
20 •  SELECT a.ims_org_id, c.business_name,
21    (CASE
22    WHEN b.bed_desc like '%ICU%'  THEN (a.census_beds)
23    ELSE 0
24    END) as census_ICU,
25    (CASE
26    WHEN b.bed_desc like '%SICU%'  THEN (a.census_beds)
27    ELSE 0
28    END) as census_SICU
29    FROM maryamheidari_week3.bed_fact a
30    left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
31    left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id
32    group by a.ims_org_id
33    order by census_ICU desc
34    LIMIT 10;
```

ims_org_id	business_name	census_ICU	census_SICU
INS00005214	Saint Marys Hospital	147	0
INS00005612	Brigham and Womens Hospital	124	124
INS00001639	Emory University Hospital	120	0
INS00001172	Saints Mary and Elizabeth Hospital	105	0
INS00005078	Jewish Hospital	98	0
INS00011224	G Werber Bryan Psychiatric Hospital	97	0
INS00014916	Bayhealth Medical Center, Inc	95	0
INS00004254	Heartland Regional Medical Center	92	0
INS00003552	Childrens Hospital of Wisconsin, Inc	91	0
INS00011312	The Devereux Texas Treatment Network	90	0

In above table, we can see you many beds Census each top ten hospital has for ICU and SICU. The only hospital how has a census for SICU is Brigham and women's hospital which is 124, which make sense because it is only hospital who has license for SICU.

For staffed

```
36    # staffed
37 •  SELECT a.ims_org_id, c.business_name,
38    (CASE
39    WHEN b.bed_desc like '%ICU%'  THEN (a.staffed_beds)
40    ELSE 0
41    END) as staffed_ICU,
42    (CASE
43    WHEN b.bed_desc like '%SICU%'  THEN (a.staffed_beds)
44    ELSE 0
45    END) as staffed_SICU
46    FROM maryamheidari_week3.bed_fact a
47    left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
48    left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id
49    group by a.ims_org_id
50    order by staffed_ICU desc
51    LIMIT 10;
```

ims_org_id	business_name	staffed_ICU	staffed_SICU
INS00005214	Saint Marys Hospital	219	0
INS00001172	Saints Mary and Elizabeth Hospital	174	0
INS00001639	Emory University Hospital	169	0
INS00011224	G Werber Bryan Psychiatric Hospital	139	0
INS00005612	Brigham and Womens Hospital	138	138
INS00003552	Childrens Hospital of Wisconsin, Inc	138	0
INS00004254	Heartland Regional Medical Center	135	0
INS00011312	The Devereux Texas Treatment Network	132	0
INS00006375	Texas Health Presbyterian Hospital Dallas	131	0
INS00000886	Trinity Medical Center	130	0

In above table, we can see you many beds staffed each top ten hospital has for ICU and SICU. The only hospital how has a census for SICU is Brigham and women's hospital which is 138, which make sense because it is only hospital who has license for SICU.

Now, it is time to find the answer for all of three question in the story part. The first one is Total licensed beds (total beds allowed by state license) for ICU and SICU

```sql
#Q1
SELECT
sum(CASE
WHEN b.bed_desc like '%ICU%'  THEN (a.license_beds)
ELSE 0
END) as license_ICU,
sum(CASE
WHEN b.bed_desc like '%SICU%'  THEN (a.license_beds)
ELSE 0
END) as license_SICU
FROM maryamheidari_week3.bed_fact a
left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id;
```

And the answer is:

	license_ICU	license_SICU
▶	107123	5222

So for top ten hospital, total ICU bed license is 107123and the total SICU bed license is 5222.

The second one is Total census beds (total beds at the hospital)

```sql
#Q2
SELECT
sum(CASE
WHEN b.bed_desc like '%ICU%'  THEN (a.census_beds)
ELSE 0
END) as census_ICU,
sum(CASE
WHEN b.bed_desc like '%SICU%'  THEN (a.census_beds)
ELSE 0
END) as census_SICU
FROM maryamheidari_week3.bed_fact a
left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id;
```

The answer is

	census_ICU	census_SICU
▶	69143	3753

So for top ten hospital, total ICU bed census is 69143 and the total SICU bed census is 3753.

And the third one is Total staffed beds (total beds for which staffing, e.g., physicians and nurses, exists)

```
#Q3
SELECT
sum(CASE
WHEN b.bed_desc like '%ICU%'  THEN (a.staffed_beds)
ELSE 0
END) as staffed_ICU,
sum(CASE
WHEN b.bed_desc like '%SICU%'  THEN (a.staffed_beds)
ELSE 0
END) as staffed_SICU
FROM maryamheidari_week3.bed_fact a
left join maryamheidari_week3.bed_type b on a.bed_id = b.bed_id
left join maryamheidari_week3.business c on a.ims_org_id = c.ims_org_id;
group by a.ims_org_id
order by census_ICU desc
LIMIT 10;
```

And the answer is

	staffed_ICU	staffed_SICU
▶	96026	4802

So, for top ten hospital, total ICU bed staffed is 96026 and the total SICU bed staffed is 4802.

Conclusion:

In this assignment, I had a lot of challenge for import data into the tables. Most of them caused by McOS, but overall, I learned to how to define the fringe keys, the relationship between child and parent tables. Moreover, I learnt to use the CASE (WHEN …THEN) function, different kind of JOIN.

I also worked with star schema to make a table and define their relations.