Tutorial 3 Database Design

Designed by ZHU Yueming. A small part of descriptions of basic concepts in this tutorial are borrowed from the Stephane Faroult's Slide and Wikipedia. Thanks for Chen Shijie help us check the tutorial and do simple modification

Experimental Objective

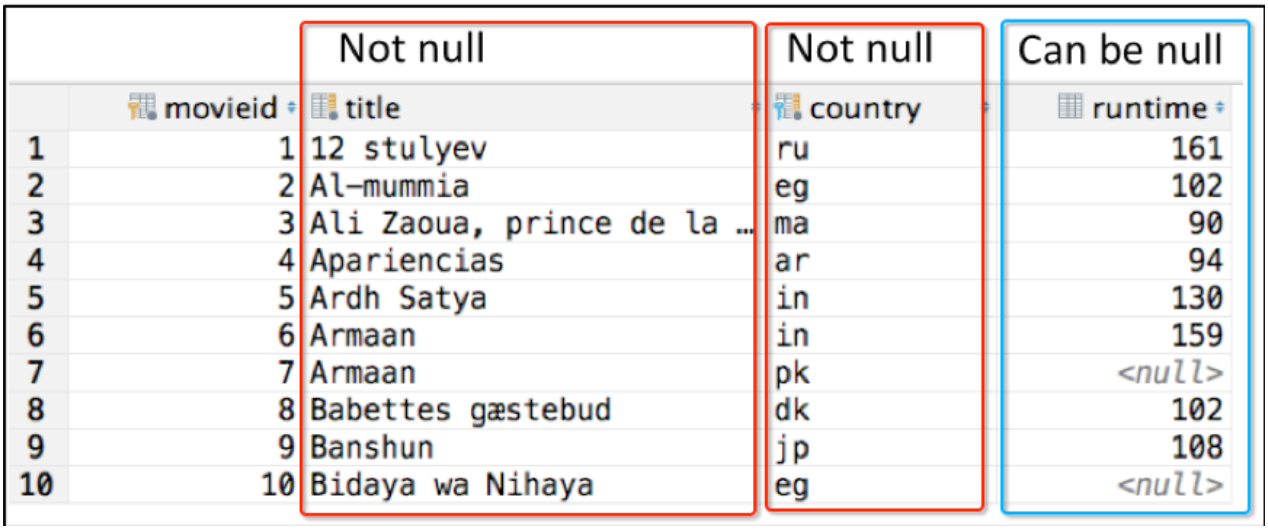
- Understand basic constraints in database design
- Understand 3 normal form of database design
- Learn how to build ER diagram using MySQL workbench
- Learn how to design a database according to the requirement document.

Constraints

Constraints are declarative rules that the DBMS apply to ensure the integrity of data. DBMS will check constraints every time new data is added, changed or deleted, to prevent any inconsistency. Any operation that violates a constraint fails and returns an error.

1. NOT NULL.

This field cannot be empty.



	movieid	title	country	runtime
1	1	12 stulyev	ru	161
2	2	Al-mummia	eg	102
3	3	Ali Zaoua, prince de la …	ma	90
4	4	Apariencias	ar	94
5	5	Ardh Satya	in	130
6	6	Armaan	in	159
7	7	Armaan	pk	<null>
8	8	Babettes gæstebud	dk	102
9	9	Banshun	jp	108
10	10	Bidaya wa Nihaya	eg	<null>

2. UNIQUE

Each value in the specified column(s) is unique.

- Unique for one column

	Primary Key	Unique Column	Can't be unique
	country_code	country_name	continent
1	dz	Algeria	AFRICA
2	ao	Angola	AFRICA
3	bj	Benin	AFRICA
4	bw	Botswana	AFRICA
5	bf	Burkina Faso	AFRICA
6	bi	Burundi	AFRICA
7	cm	Cameroon	AFRICA
8	cf	Central African Republic	AFRICA
9	td	Chad	AFRICA
10	km	Comoros	AFRICA

- Unique for multiple columns

		unique (surname, first_name)		
	peopleid	first_name	surname	gender
7	206	Lexi	Alexander	F
8	207	Ross	Alexander	M
9	208	Sarah	Alexander	F
10	209	Terry	Alexander	M
11	213	James	Algar	M
12	496	James	Arness	M
13	965	James	Baskett	M
14	1119	James	Belushi	M
15	1267	James	Best	M
16	1489	James	Bobin	M

3. PRIMARY KEY

Primary key specifies the main key for the table, which is:

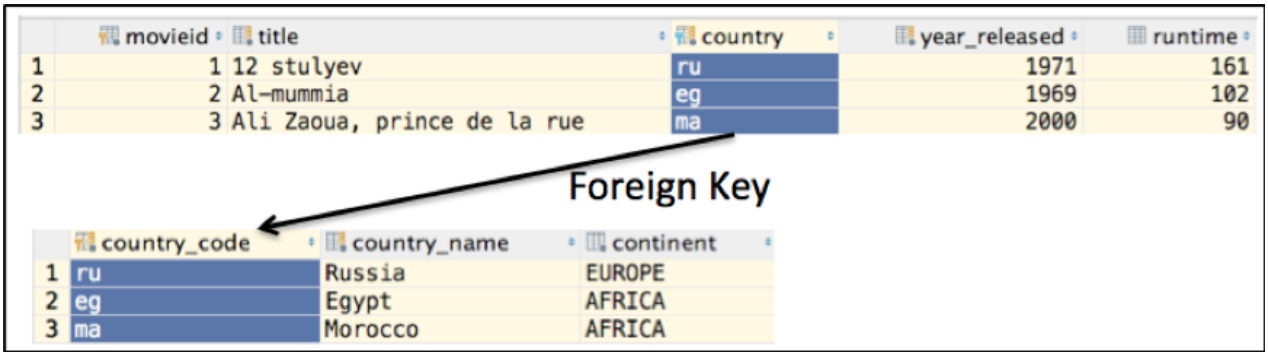
- Mandatory (the additional NOT NULL doesn't hurt but is redundant)
- Unique (no duplicates allowed in the column)

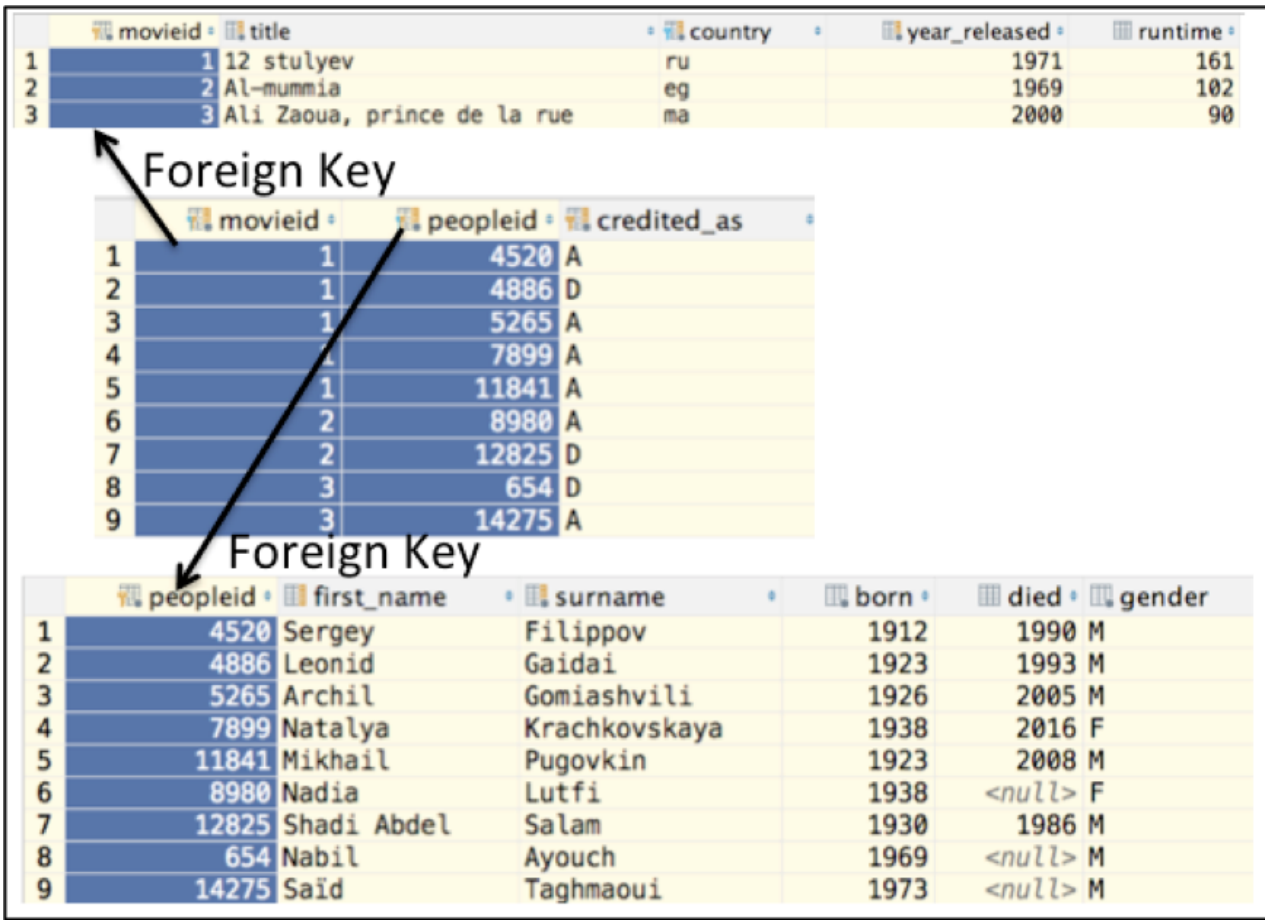
	movieid	title	country	year_released	runtime
1	1	12 stulyev	ru	1971	161
2	2	Al-mummia	eg	1969	102
3	3	Ali Zaoua, prince…	ma	2000	90
4	4	Apariencias	ar	2000	94
5	5	Ardh Satya	in	1983	130
6	6	Armaan	in	2003	159
7	7	Armaan	pk	1966	<null>
8	8	Babettes gæstebud	dk	1987	102
9	9	Banshun	jp	1949	108
10	10	Bidaya wa Nihaya	eg	1960	<null>

4. FOREIGN KEY

Foreign key indicates that the column must reference a key (Only primary keys and columns declared as UNIQUE) of another table.

- Constraints are used to prevent actions that break the connection between tables.
- Constraints also prevent illegal data from being inserted into the column.





Normal Form

1NF

A relation is in first normal form if and only if the domain of each attribute contains only atomic (indivisible) values, and the value of each attribute contains only a single value from that domain.

Comparison 1:

The first one doesn't satisfy 1NF, because the full name can be split to first name and surname, lifetime can be split to born time and death time. After that we can fetch data of first name, surname, born time and death time directly.

wrong

	peopleid	fullname	lifetime
1	1	O'Shea Jackson (Jr.),	1991, still alive
2	2	Oscar Beregi (Sr.),	1876, 1965
3	3	null, 50 Cent	1975, still alive
4	4	null, Aaliyah	1979, 2001
5	5	null, Aamani	1973, still alive
6	6	Willie, Aames	1960, still alive
7	7	Caroline, Aaron	1952, still alive
8	8	Quinton, Aaron	1984, still alive
9	9	Dodo, Abashidze	1924, 1990
10	10	Diego, Abatantuono	1955, still alive

correct

peopleid	first_name	surname	born	died
4	<null>	Aaliyah	1979	2001
5	<null>	Aamani	1973	<null>
6	Willie	Aames	1960	<null>
7	Caroline	Aaron	1952	<null>
8	Quinton	Aaron	1984	<null>
9	Dodo	Abashidze	1924	1990
10	Diego	Abatantuono	1955	<null>

Comparison 2:

The first one doesn't satisfy 1NF, because the column actor1 and actor2 are similar, so that we should merge them into one column.

First design: wrong

	title	year_released	runtime	actor1	actor2
1	Zhēngfā Tàipíng Yáng	2016	90	Brandon Routh	Yuqi Zhang
2	Zeoi¹ Lung⁴	2017	112	Andy Lau	Donnie Yen

Also a wrong design, but it is better than the first one

	id	title	year_released	runtime	actor
1	5197	Zhēngfā Tàipíng Yáng	2016	90	Brandon Routh
2	5197	Zhēngfā Tàipíng Yáng	2016	90	Yuqi Zhang
3	9141	Zeoi¹ Lung⁴	2017	112	Andy Lau
4	9141	Zeoi¹ Lung⁴	2017	112	Donnie Yen

2NF

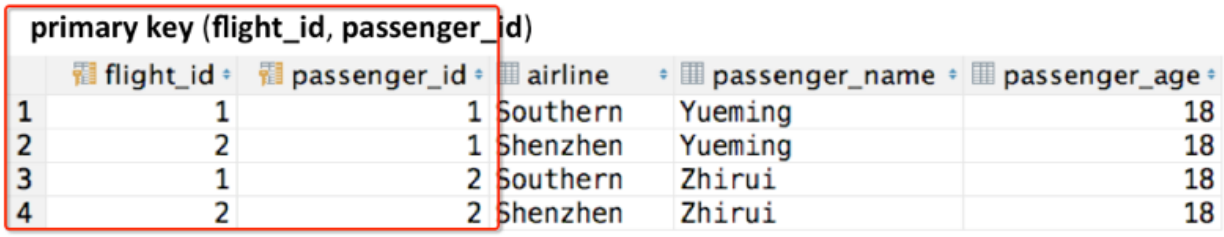
A relation satisfying 2NF must:

- be in 1NF
- not have any non-prime attribute that is dependent on any proper subset of any candidate key of the relation. A non-prime attribute of a relation is an attribute that is not a part of any candidate key of the relation.

The primary keys in the first table are flightid and passengerid, but the first one is a wrong design bause the columns passengerName and passengerAge are not related to fightid (primary key).

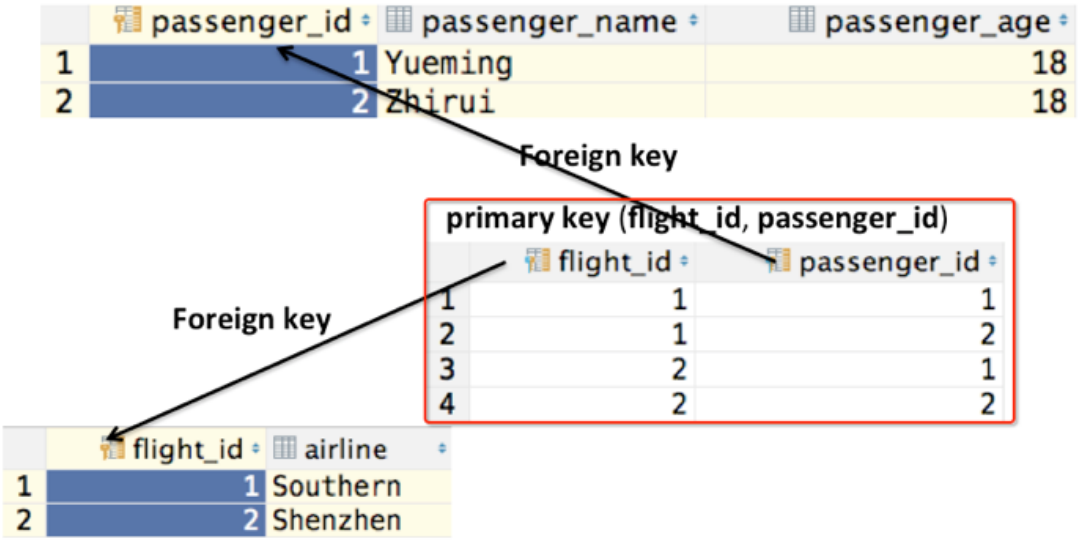
If we want to describe **many-to-many relationship** between two entities, we need to separate those two entities into two tables and then add a **relation table** to describe the relationship itself (as in the second case).

Wrong design



	flight_id	passenger_id	airline	passenger_name	passenger_age
1	1	1	Southern	Yueming	18
2	2	1	Shenzhen	Yueming	18
3	1	2	Southern	Zhirui	18
4	2	2	Shenzhen	Zhirui	18

Correct design



	passenger_id	passenger_name	passenger_age
1	1	Yueming	18
2	2	Zhirui	18

primary key (flight_id, passenger_id)

	flight_id	passenger_id
1	1	1
2	1	2
3	2	1
4	2	2

	flight_id	airline
1	1	Southern
2	2	Shenzhen

3NF

A relation satisfying 2NF must:

- be in 2NF
- all the attributes in a table are determined only by the candidate keys of that relation and not by any non-prime attributes.

We can see from the figure that the column city and province are not related to the primary key (flightid). Those columns are more suitable for describing cities. In this case, we only use the id of cities, which serves as a foreign key referencing the city table, in the flight table,

If we want to describe **one-to-many relationship**, e.g. A only has one B, but B could have more than one A, we can design a foreign key in A table referencing the primary key in B table.

Wrong design

primary key

	airline_id	airline_name	city	province
1	1	Southern	GuangZhou	GuangDong
2	2	Shenzhen	ShenZhen	GuangDong
3	3	Capital	BeiJing	BeiJing
4	4	HK Express	HK	HK
5	5	Cathay Pacific	HK	HK

Correct design

primary key

	airline_id	airline_name	city_id
1	1	Southern	1
2	2	Shenzhen	2
3	3	Capital	3
4	4	HK Express	4
5	5	Cathay Pacific	4

Foreign key

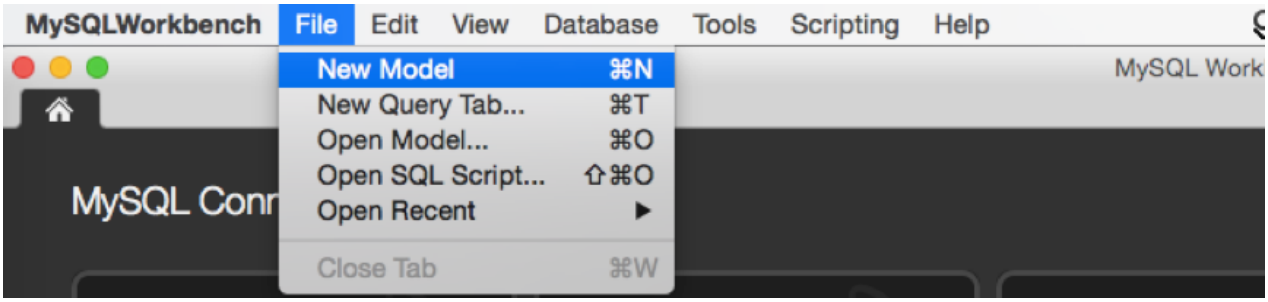
primary key

	city_id	city_name	province
1	1	GuangZhou	GuangDong
2	2	ShenZhen	GuangDong
3	3	BeiJing	BeiJing
4	4	HK	HK

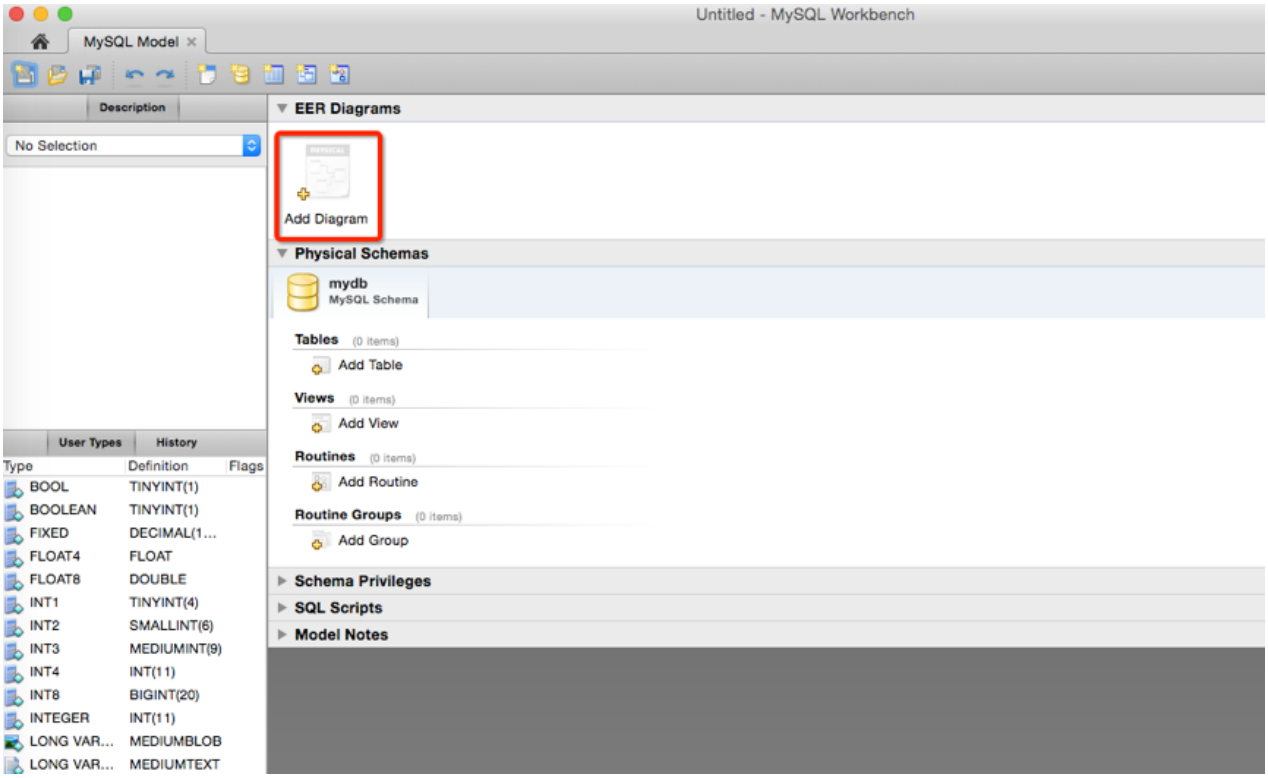
Build ER diagram Using Mysql Workbench

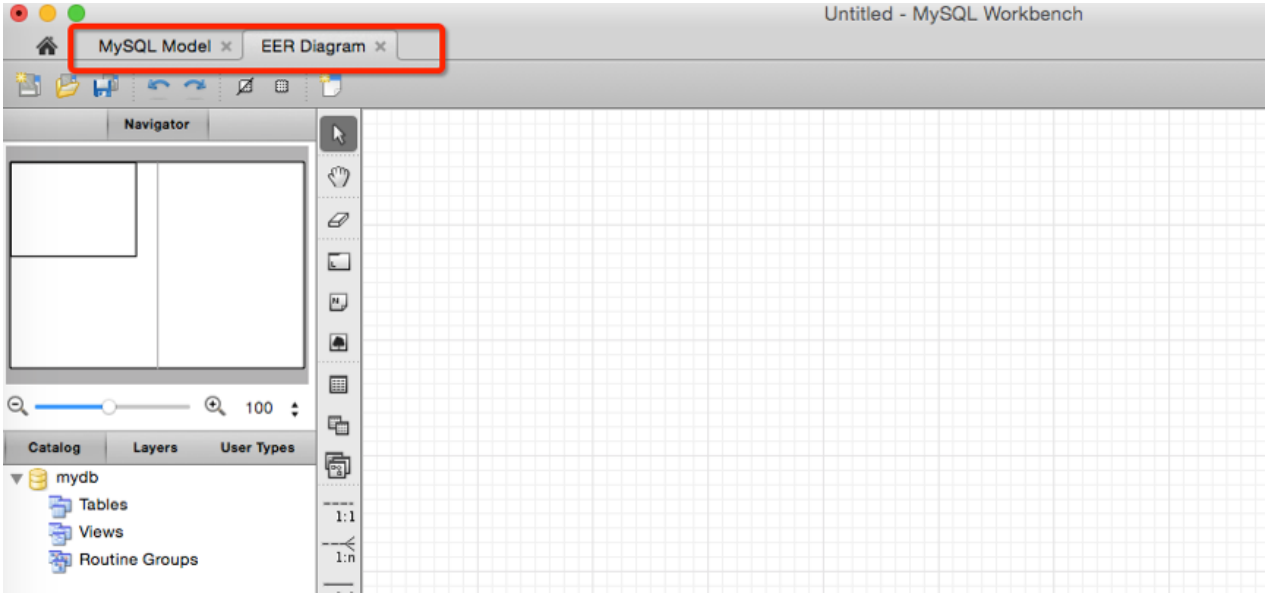
How to create ER diagrams using MySQL workbench without connection?

Step 1: File—New Model

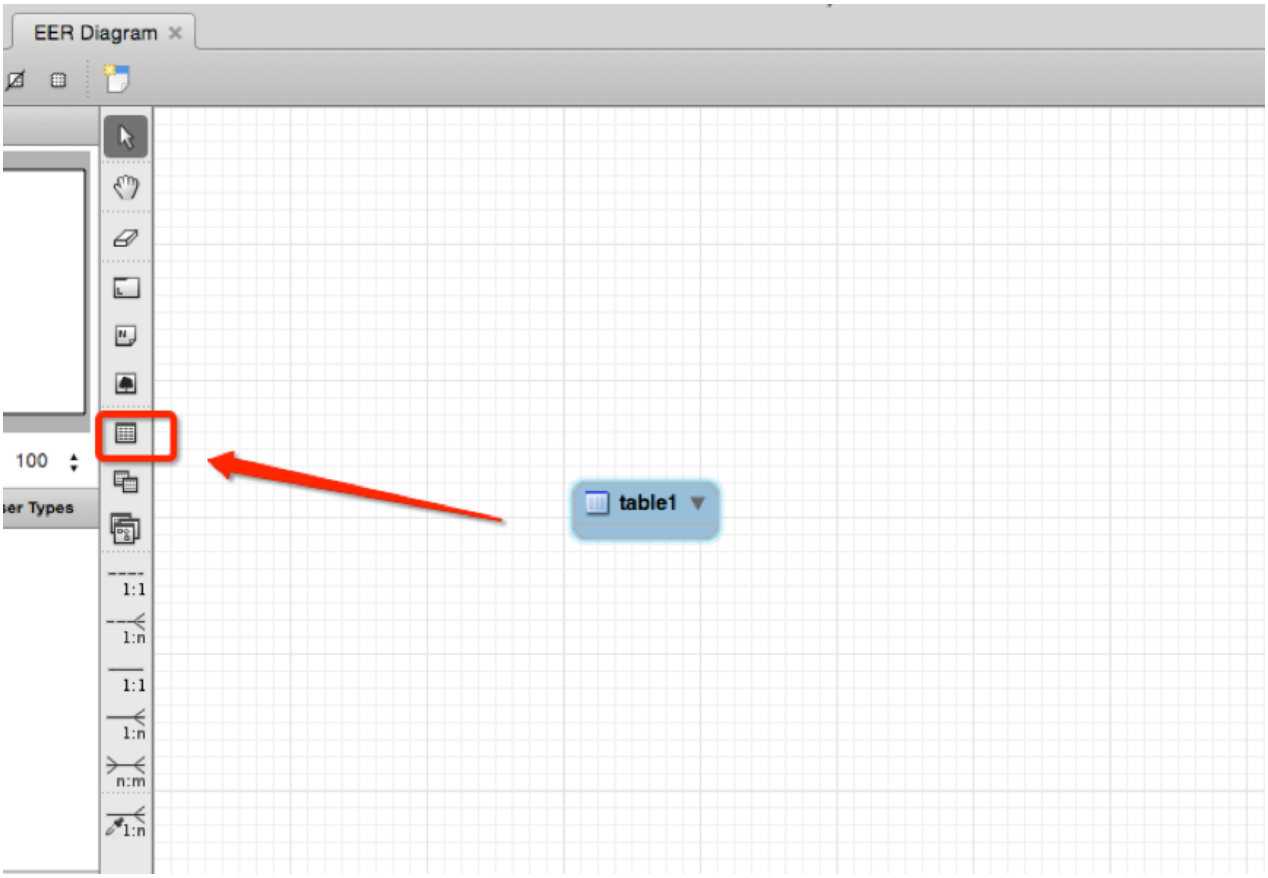


Step2: Double click Add diagram

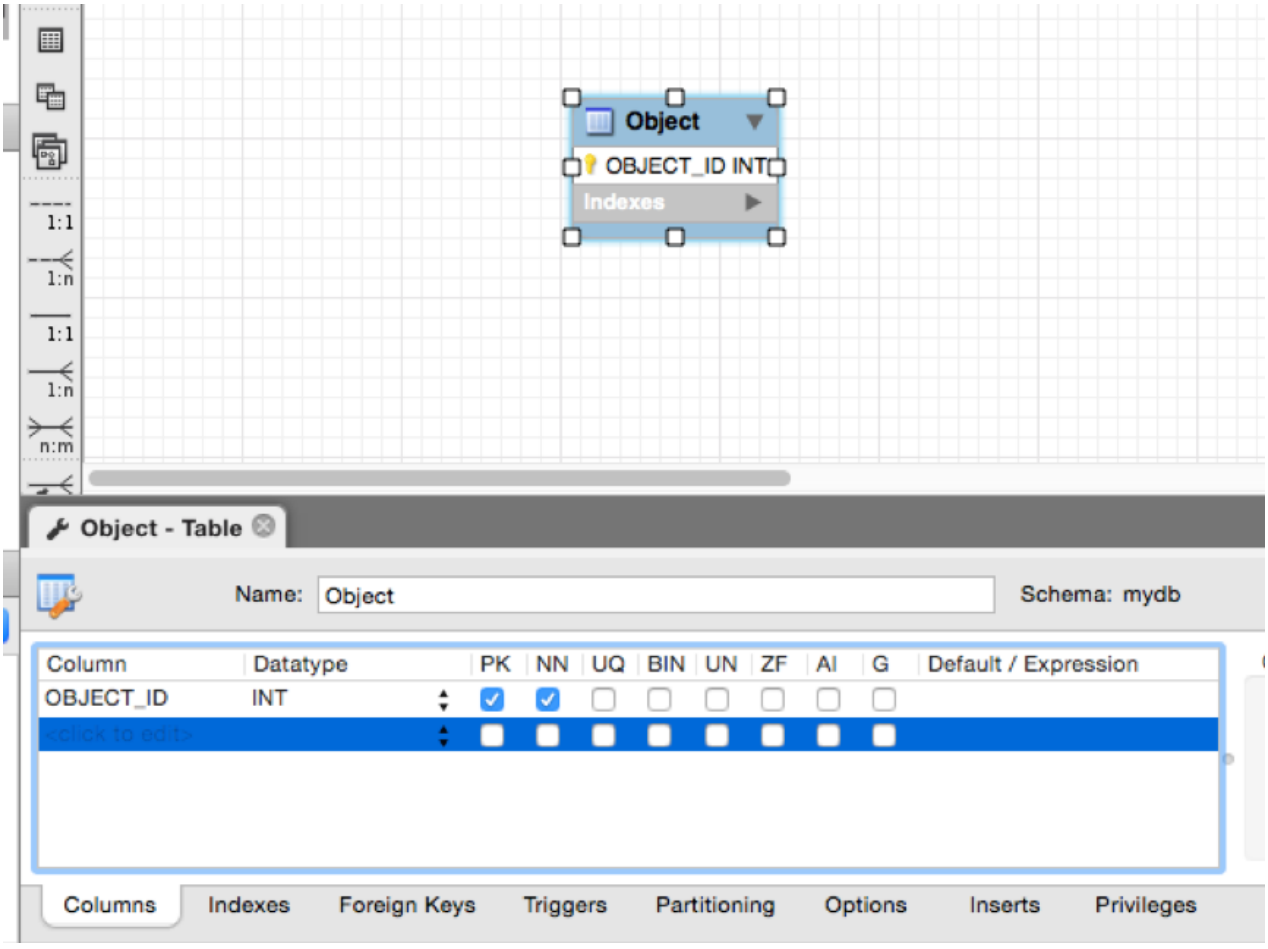




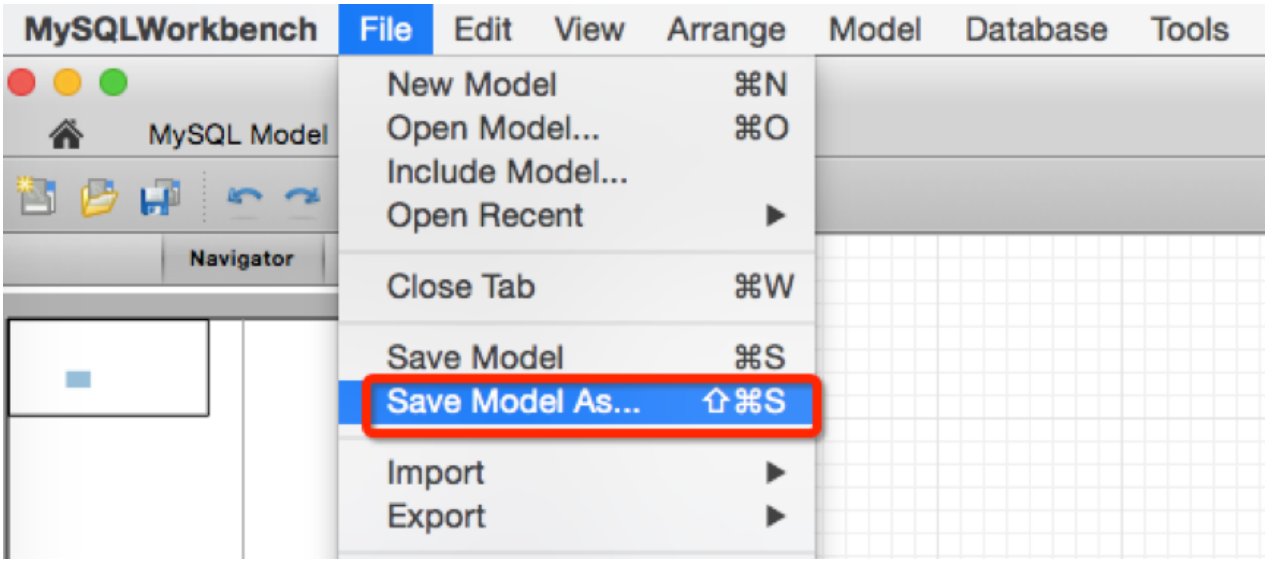
Step3: Add table to ER diagram by clicking this button



Step 4: Double click table in ER diagram and edit it. Give it a name and add other columns.



Step 5: When you finish editing your ER diagram, do not forget to save it



WHAT TO SUBMIT

Design a simple database which contains three tables as follows:

- Student (name, student_id, department, gender)
- Department (name, location, website)
- Course (name, course_number, department, credit)

Other requirements describe as follows:

- The relationship between Student and Course is many-to-many.
- The relationship between Course and Department is many-to-one.
- The relationship between Student and Department is many-to-one.

Please use mysql workbech to design a simple database that can match all requirements above, and then submit a pdf file into sakai website as week3 submission.