Peter Shen
BMI 701 Assignment 1

Question 1.

	Query	Result
A	`select count(*) from ` + "`patients`" + ` where` `` `patients`.`sex` = 'F'; ``	16751
B	`select count(*) from ` + "`diag3d`" + ` where diag3d.diag =` `'(250) Diabetes mellitus';`	1261
C	`select count(*) from `patients`` `join `pmh` on `patients`.`id` = `pmh`.`id`` `join `drugs` on `patients`.`id` = `drugs`.`id`` `where `pmh`.`arthritis` = 1` `AND `drugs`.`drug` = 'celecoxib';`	101

D	`select * from (`
	`    select count(`visitreason`) as 'count',`
	`    visitreason`
	`    from `patients` group by `visitreason`) as A`
	`order by `count` desc`
		1925
		General
		medical
		examination

E	`select count(*) from (` `    select count(*) as `count` from `patients`` `    join `tests` on `patients`.`id` = `tests`.`id`` `    where `tests`.`test` in ('glucose', 'hba1c')` `    group by `patients`.`id`) as `A`` `where `A`.`count` = 2;`	395

Question 2:

a)
```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
where `pmh`.diabetes = 1;
```

Result: 2955

```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
join `tests` on `tests`.id = `patients`.id
where `pmh`.diabetes = 1 and `tests`.test = 'hba1c'
```

Result: 441

441/2955 = 14.9%

b)
```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
where `pmh`.htn = 1
```

Result: 7026

```
select count(*) from `patients` where id in (
    select `patients`.id from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`.htn = 1)
and sbp is null or sbp = '';
```

Result: 1992

1992/7026 = 28.4%

c)
```
select count(`patients`.id) from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`. hyperlipid = 1
```
Result: 3975

```
select count(*) from (select id from `drugs` where `drugs`.id in
(
    select `patients`.id from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`.hyperlipid = 1) group by id) as A
```

Result: 3463
3975 - 3463 = 512 who are not taking drugs

512 / 3975 = 12.9%

Question 2 but looking at just males
a)
```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
where `pmh`.diabetes = 1 and `patients`.sex = 'M';
```

Result: 1336

```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
join `tests` on `tests`.id = `patients`.id
where `pmh`.diabetes = 1 and `tests`.test = 'hba1c' and
`patients`.sex = 'M';
```

Result: 207

207/1336 = 15.5%

b)
```
select count(*) from `patients`
join `pmh` on `patients`.id = `pmh`.id
where `pmh`.htn = 1 and `patients`.sex = 'M';
```

Result: 3072

```
select count(*) from `patients` where id in (
    select `patients`.id from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`.htn = 1 and `patients`.sex = 'M')
and sbp is null or sbp = '';
```

Result: 887

887/3072 = 28.9%

c)
```
select count(`patients`.id) from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`. hyperlipid = 1 and `patients`.sex = 'M'
```
Result: 1887

```
select count(*) from (select id from `drugs` where `drugs`.id in
(
    select `patients`.id from `patients`
    join `pmh` on `patients`.id = `pmh`.id
    where `pmh`.hyperlipid = 1 and `patients`.sex = 'M') group
by id) as A
```

Result: 1623
1887 - 1623 = 264 who are not taking drugs

264 / 1887 = 14.0%

Overall, recommended interventions are less followed now compared to 2003.

Question 3

```
select count(sbp) from patients where sbp like '%0';
```
Result: 5908
```
select count(sbp) from patients where sbp is not null or sbp !=
'';
```
Result: 15630

```
select count(dbp) from patients where dbp like '%0';
```
Result: 6208
```
select count(dbp) from patients where dbp is not null or dbp !=
'';
```
Result: 15618

Do another pass on sbp that end in other numbers:

Ends In	Records
0	5908
1	416
2	2255
3	392
4	1601
5	549
6	1400
7	445
8	2201
9	463

There does seem to be a preference for measurements to end in 0

Question 4

Question:
Our surroundings such the food we eat, the cultures we follow, and our lifestyles can attribute
to certain disease progressions.
Which region has the highest number of patients diagnosed with hyperlipid and what age group
do most of these patients belong to?

```
select `patients`.region as 'Region', count(*) as `Number of
Patients` from `patients` join `pmh` on `patients`.id = `pmh`.id
where `pmh`.hyperlipid = 1 group by `patients`.region;
```
Result:

Region	Number of Patients
Midwest	880
Northeast	935
South	1311
West	849

```
select `patients`.age, count(*) as `Number of Patients` from
`patients` join `pmh` on `patients`.id = `pmh`.id where
`pmh`.hyperlipid = 1 and `patients`.region = 'South' group by
age order by `Number of Patients` desc
```

Result: Most of these patients are from late 50s to early 70s.