

- birth_year
 - there are tuples where birth_year is not 4-digit number(including space, which is recognized from sql as 1 digit)
 - So, update those birth_year to default value

```
update PatientInfo
set birth_year = 9999
where birth_year not like '____'
```

- infected_by
 - there are tuples where infected_by is form of global_num, not patient_id
 - So, update those value to patient_id whose global_num is same from previous tuple's value

```
update PatientInfo
set infected_by = (
  select patient_id
  from PatientInfo as PI
  where PI.global_num = CAST(substr(PatientInfo.infected_by , 2, 5) as
  INTEGER)
)
WHERE infected_by not like '_____'
```

- contact_number
 - there are values that are too high(ex. 1000000772, 1000000796)(It's too big compared to the population in Korea) or non-digit values(ex. '-')
 - So, update those non-ideal values to default value

```
update PatientInfo
set contact_number = 0
where contact_number > 10000 OR contact_number = '-'
```

- age
 - age value is like '%s' or 'None'(which is default value). But, there are values that are not like '%s'(ex. 30)
 - So, update those non-ideal values to '%s' type
- symptoms_onset_date
 - there are values that are not date-type 4digit-2digit-2digit(ex. space value)
 - So, update those values to default value
 - symptoms_onset_date may be later than confirmed_date, but due to the nature of the disease, it was not modified.

```
update PatientInfo
set symptom_onset_date = '9999-12-31'
```

```
where symptom_onset_date not like '____-__-__'
```