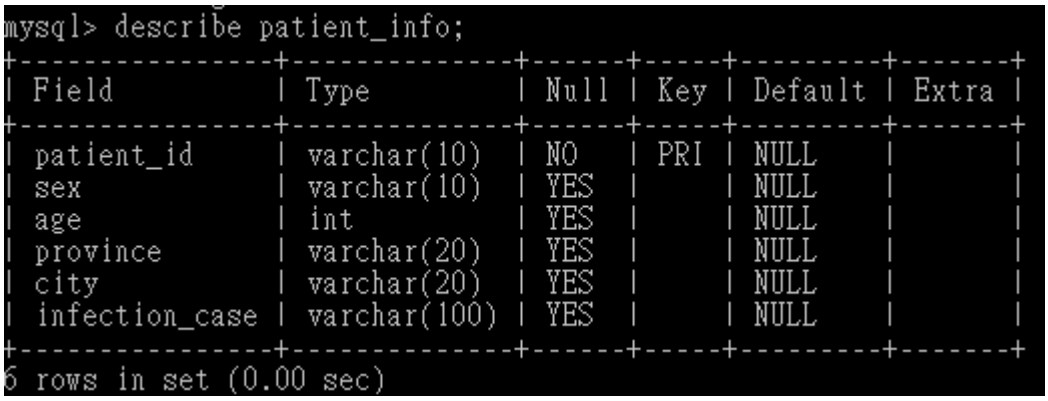
Part 1 - Data Science for COVID-19 in South Korea:

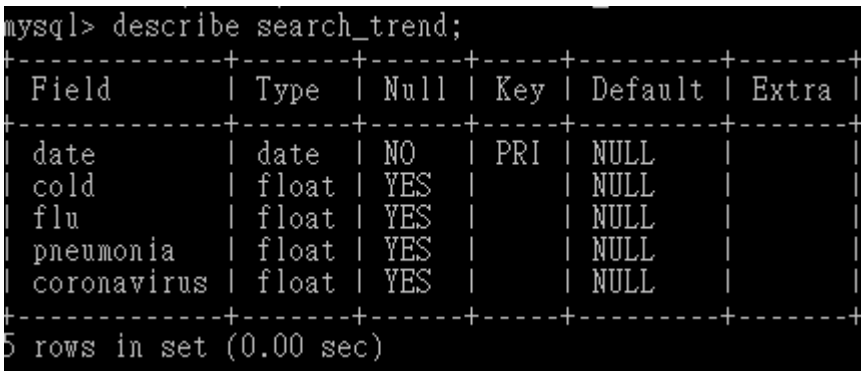
- A. Create Tables
    - What is the difference between type “char” and type “varchar”?
        - Varchar is variable length and char is fixed length.
    - How many bytes it should take for “tinyint”, “smallint”, “mediumint”, “int”? And what’s the range they can express?
        - Tinyint 1byte range (-128 to 127) ,smallint 2byte range(-32,768 to 32767), mediumint 3byte range(-8388608 to 8388607), int 4byte range(-2147483648 to 2147483647).
    - What do you think about this DB schema? If you can change this table architecture, how would you modify it and why?
        - Overall I think this database is good, and you can get a lot of information out of it without writing too much query ◦ If I can change it, it

will be to add a column that represent the amount of increment per day in some of the tables.

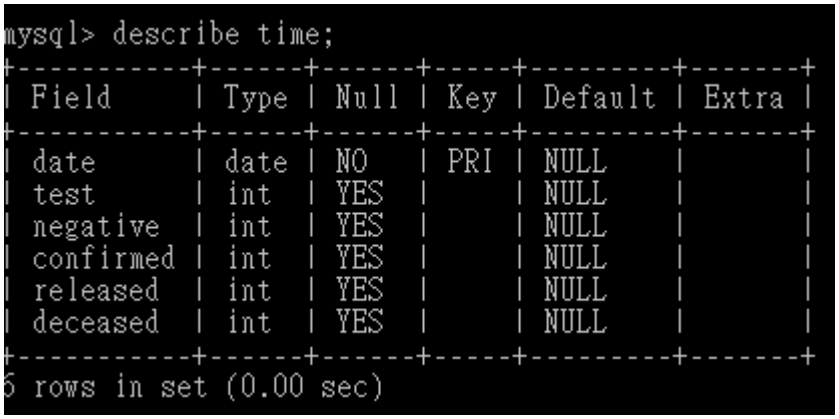
- Screen shots



```
mysql> describe patient_info;
+----------------+-------------+------+-----+---------+-------+
| Field          | Type        | Null | Key | Default | Extra |
+----------------+-------------+------+-----+---------+-------+
| patient_id     | varchar(10) | NO   | PRI | NULL    |       |
| sex            | varchar(10) | YES  |     | NULL    |       |
| age            | int         | YES  |     | NULL    |       |
| province       | varchar(20) | YES  |     | NULL    |       |
| city           | varchar(20) | YES  |     | NULL    |       |
| infection_case | varchar(100)| YES  |     | NULL    |       |
+----------------+-------------+------+-----+---------+-------+
6 rows in set (0.00 sec)
```



```
mysql> describe search_trend;
+-------------+-------+------+-----+---------+-------+
| Field       | Type  | Null | Key | Default | Extra |
+-------------+-------+------+-----+---------+-------+
| date        | date  | NO   | PRI | NULL    |       |
| cold        | float | YES  |     | NULL    |       |
| flu         | float | YES  |     | NULL    |       |
| pneumonia   | float | YES  |     | NULL    |       |
| coronavirus | float | YES  |     | NULL    |       |
+-------------+-------+------+-----+---------+-------+
5 rows in set (0.00 sec)
```



```
mysql> describe time;
+-----------+------+------+-----+---------+-------+
| Field     | Type | Null | Key | Default | Extra |
+-----------+------+------+-----+---------+-------+
| date      | date | NO   | PRI | NULL    |       |
| test      | int  | YES  |     | NULL    |       |
| negative  | int  | YES  |     | NULL    |       |
| confirmed | int  | YES  |     | NULL    |       |
| released  | int  | YES  |     | NULL    |       |
| deceased  | int  | YES  |     | NULL    |       |
+-----------+------+------+-----+---------+-------+
6 rows in set (0.00 sec)
```

```
mysql> describe time_age;
+-----------+------+------+-----+---------+-------+
| Field     | Type | Null | Key | Default | Extra |
+-----------+------+------+-----+---------+-------+
| date      | date | NO   | PRI | NULL    |       |
| age       | int  | NO   | PRI | NULL    |       |
| confirmed | int  | YES  |     | NULL    |       |
| deceased  | int  | YES  |     | NULL    |       |
+-----------+------+------+-----+---------+-------+
4 rows in set (0.00 sec)
```

```
mysql> describe time_gender;
+-----------+-------------+------+-----+---------+-------+
| Field     | Type        | Null | Key | Default | Extra |
+-----------+-------------+------+-----+---------+-------+
| date      | date        | NO   | PRI | NULL    |       |
| sex       | varchar(10) | NO   | PRI | NULL    |       |
| confirmed | int         | YES  |     | NULL    |       |
| deceased  | int         | YES  |     | NULL    |       |
+-----------+-------------+------+-----+---------+-------+
4 rows in set (0.00 sec)
```

```
mysql> describe time_province;
+-----------+-------------+------+-----+---------+-------+
| Field     | Type        | Null | Key | Default | Extra |
+-----------+-------------+------+-----+---------+-------+
| date      | date        | NO   | PRI | NULL    |       |
| province  | varchar(20) | NO   | PRI | NULL    |       |
| confirmed | int         | YES  |     | NULL    |       |
| released  | int         | YES  |     | NULL    |       |
| deceased  | int         | YES  |     | NULL    |       |
+-----------+-------------+------+-----+---------+-------+
5 rows in set (0.01 sec)
```

```
mysql> describe region;
+--------------------------+-------------+------+-----+---------+-------+
| Field                    | Type        | Null | Key | Default | Extra |
+--------------------------+-------------+------+-----+---------+-------+
| code                     | int         | NO   | PRI | NULL    |       |
| province                 | varchar(20) | YES  |     | NULL    |       |
| city                     | varchar(20) | YES  |     | NULL    |       |
| elementary_school_count  | int         | YES  |     | NULL    |       |
| kindergarten_count       | int         | YES  |     | NULL    |       |
| university_count         | int         | YES  |     | NULL    |       |
| elderly_population_ratio | float       | YES  |     | NULL    |       |
| elderly_alone_ratio      | float       | YES  |     | NULL    |       |
| nursing_home_count       | int         | YES  |     | NULL    |       |
+--------------------------+-------------+------+-----+---------+-------+
9 rows in set (0.00 sec)
```

```
mysql> describe weather;
+----------------------+-------+------+-----+---------+-------+
| Field                | Type  | Null | Key | Default | Extra |
+----------------------+-------+------+-----+---------+-------+
| code                 | int   | NO   | PRI | NULL    |       |
| date                 | date  | NO   | PRI | NULL    |       |
| avg_temp             | float | YES  |     | NULL    |       |
| most_wind_direction  | int   | YES  |     | NULL    |       |
| avg_relative_humidity| float | YES  |     | NULL    |       |
+----------------------+-------+------+-----+---------+-------+
5 rows in set (0.00 sec)
```

- C. Query tasks

12. 我用 match_info 內所有比賽的歷史資料算出每一個隊伍過去的平均勝率、敗率、平手率，再以每場比賽三家賭場所開的賠率去算下注每個隊伍勝、敗、平手，所賺到的錢的期望值各是多少。所以當今天有一場比賽，而我知道雙方的隊伍，就可以藉由這個表來判斷投資哪一方會賺最多。若是不管哪一方期望值都是負的那麼就可以直接放棄這場比賽。

B365_win_exp 是根據 B365 這個賭場的賠率算出投資勝利的期望值，B365_lose_exp 是輸、B365_even_exp 是平手。另外還有 WH_win_exp、SJ_win_exp 等等。