

GaussDB(for openGauss)数据库实验报告

1 GaussDB(for openGauss)数据库创建与维护实验

1.1 实验目的

1. 通过对GAUSSDB(FOR OPENGauss)数据库创建与访问：
 - (1) 了解华为云分布式数据库GAUSSDB(FOR OPENGauss)的软件环境和创建方法；
 - (2) 掌握并熟悉GAUSSDB(FOR OPENGauss)数据库软件的使用方法；
 - (3) 掌握并熟悉GAUSSDB(FOR OPENGauss)数据库软件的构成和相关工具；
 - (4) 通过GAUSSDB(FOR OPENGauss)数据库软件的使用，深入理解数据库系统的基本概念。
2. 通过创建GAUSSDB(FOR OPENGauss)数据库及进行相应的维护，了解并掌握GaussDB(for openGauss)数据库的创建与维护的不同方法和途径，进而通过这一具体的数据库理解实际数据库所包含的各要素。

1.2 实验平台及环境

- 华为云: GaussDB 2.7.2
- 数据库: PostgreSQL

1.3 实验内容

1. GAUSSDB(FOR OPENGauss)数据库软件的使用：
 - (1) 登录并运行GAUSSDB(FOR OPENGauss)数据库；
 - (2) 了解华为云数据库GAUSSDB(FOR OPENGauss)的机制；
 - (3) 熟悉GAUSSDB(FOR OPENGauss)数据库的各项功能。
2. 数据库创建与维护：
 - (1) 创建“疫情数据”数据库；
 - (2) 对数据库属性和参数进行相应的修改和维护；
 - (3) 练习数据库的删除等维护操作。

1.4 实验步骤

1. 熟悉GaussDB(for openGauss)数据库的创建过程；
2. 创建一个名为“疫情数据”的数据库；
3. 删除“疫情数据”数据库。

1.5 实验结果及分析

1.5.1 数据库的创建

1. 登录华为云平台，选择IAM用户登录，原华为云账号为 luxqbupt123，IAM用户名为 bupt+本人学号，密码为 bupt2024。
2. 进入控制台后点击数据库，进入数据库管理服务DAS，进入开发工具，环境切换到华南-广州-友好用户环境。
3. 新增数据库实例连接，数据库引擎选择GaussDB，数据库实例选择 gauss-3c93_root_0，角色为 master。

新建数据库



* 数据库名称

Epidemic_data

只能创建用户数据库

字符集

UTF8



Template ②

template0



Collation ②

Ctype ②

DBCMPATIBILITY ②

PostgreSQL



确定

取消

对应的SQL语句为：

```
1 CREATE DATABASE 疫情数据
```

1.5.2 数据库的删除

删除操作直接点击右侧删除即可。

删除数据库



确定执行 drop database Epidemic_data 吗？

该动作将彻底删除该库及里面的所有数据，且不可恢复，请谨慎操作！

库名	表数量	表大小	索引大小
Epidemic_data	--	--	--

是

否

对应SQL语句为：

```
1 DROP DATABASE 疫情数据
```

删除成功后，数据库 疫情数据 被删除。

1.6 实验小结

本次实验我通过华为云完成了数据库的创建和删除，实现了华为云数据库的使用方法，能通过华为云对数据库进行基本操作，目前没有遇到其他问题。

2 数据库表的创建与维护实验

2.1 实验目的

1. 通过进行数据库表的建立操作，熟悉并掌握GAUSSDB(FOR OPENGAUSS)数据库表的建立方法，理解关系型数据库表的结构，巩固PostgreSQL中关于数据库表的建立语句；
2. 通过进行数据库表数据的增加、删除和插入等维护操作，熟悉并掌握GAUSSDB(FOR OPENGAUSS)数据库数据的操作方法，巩固PostgreSQL中关于数据维护的语句。

2.2 实验平台及环境

- 华为云: GaussDB 2.7.2
- 数据库: PostgreSQL

2.3 实验内容

建立相应的表并熟悉基本操作，例如建表、对表进行增、删、改、查。

2.4 实验步骤

1. 熟悉课程实验背景知识；
2. 使用GAUSSDB(FOR OPENGAUSS)数据库软件创建相应的表；
3. 将提供的数据导入各表，掌握GAUSSDB(FOR OPENGAUSS)数据库数据导入的方法；
注意：
 - (1)表中空列的处理；
 - (2)表结构与数据类型的匹配。
4. 修改“病例基本信息”表数据，增加名为“备注”的列,数据类型为varchar()型；
5. 修改“病例基本信息”表数据，将“备注”列的数据类型改为int；
6. 修改“病例基本信息”表数据，删除“备注”列；
7. 删除“病例基本信息”数据表。

2.5 实验过程及分析

数据表的创建和导入

由于各表的创建步骤类似，因此此处只以病例基本信息表的创建和数据导入为例说明，首先执行如下SQL语句建表：

```
1 CREATE TABLE 病例基本信息 (
2     病例号 INT PRIMARY KEY,
3     省 VARCHAR(50) NULL,
4     市 VARCHAR(50) NULL,
5     区 VARCHAR(50) NULL,
6     日期 DATE NULL,
7     性别 VARCHAR(10) NULL,
8     年龄 INT NULL,
9     患者信息 VARCHAR(500) NULL,
10    其它信息 VARCHAR(500) NULL,
11    信息来源 VARCHAR(500) NULL
12 );
```

建表完成后，点击上侧的导入·导出中的导入，新建任务

新建任务

导入类型

sql

CSV

文件来源

上传文件

从OBS中选择

附件存放位置 ?

obs0630bupt

没有OBS桶? [创建OBS桶](#)

创建OBS桶免费，但保存文件将产生一定的费用。

选择附件



点击或将文件拖动到此处后上传文件 (.csv)

📎 病例基本信息表_1729764961361.csv

最大不能超过1GB，且只能上传一个附件

数据库

bupt2022211119

表

"bupt2022211119"."病例基本信息"

数据位置

第1行为属性

第1行为数据

字符集

UTF8

GBK

自动检测

选项

☐ 忽略报错,即SQL执行失败时跳过

☒ 导入完成后删除上传的文件

☐ 清空表,执行导入前先对相应的表执行Truncate操作

备注

创建导入任务

取消

注意，对存在空列的表需要使用excel将空列删除，创建导入任务后等待数据上传完成则数据导入成功。

修改“病例基本信息”表数据，增加名为“备注”的列,数据类型为vachar()型

执行SQL语句为:

```
1 | ALTER TABLE 病例基本信息 ADD COLUMN 备注 VARCHAR(100)
```

结果如下:

病例基本信息	
▼	columns
	病例号 (int4(32,0))
	省 (varchar(50))
	市 (varchar(50))
	区 (varchar(50))
	日期 (date)
	性别 (varchar(10))
	年龄 (int4(32,0))
	患者信息 (varchar(500))
	其它信息 (varchar(500))
	信息来源 (varchar(500))
	备注 (varchar(100))

修改“病例基本信息”表数据，将“备注”列的数据类型改为int

执行SQL语句为:

```
1 | ALTER TABLE 病例基本信息 ALTER COLUMN 备注 TYPE INT
```

结果如下:



修改“病例基本信息”表数据，删除“备注”列

执行SQL语句为:

```
1 | ALTER TABLE 病例基本信息 DROP COLUMN 备注;
```

删除“病例基本信息”数据表

执行SQL语句为:

```
1 | DROP TABLE 病例基本信息
```

即可删除成功。

2.6 实验总结

本次实验中第一个遇到的问题是空列的处理，对数据中的空列如果不进行处理的话导入数据时文件格式与表的格式就会不一致导致导入异常，经自己摸索和询问老师，对这种空列直接在excel中删除整列即可，第二个遇到的问题是 `varchar` 刚开始设置的长度有时会太小导致无法存下所有数据，导致导入异常，增大长度后即可正常导入。

在实验中还遇到设置主键和外键导致异常的问题，数据文件中的数据很多，同时不同数据的格式和可靠性也不同，可能有的数据的部分表项为空，所以不能够盲目的去设置主键和外键，需要谨慎考虑。

实验让我对基本的SQL建表以及修改表的操作更加熟悉，除上述问题以外没有遇到其他问题，同时希望能够更新一下第一次实验的PPT，PPT中的部分账号和密码已经更换，有一定的误导性。

3 数据查询实验

3.1 实验目的

通过对实验二建立的数据库关系表的各种查询的操作，加深对SQL语言和PostgreSQL查询语言的了解，掌握相关查询语句的语法及使用方法。

3.2 实验内容

1. 单表查询

- 1-1 查询国内确诊病例基本信息的所有信息来源。
- 1-2 给出河南省、西藏自治区、台湾省的英文名称和人口数。
- 1-3 查询2021年1月20日各省现有确诊病例数据，按现有确诊病例数降序排列输出。
- 1-4 顺义区中风险地区的数量。
- 1-5 计算截至2021年1月20日全国累计确诊病例数。
- 1-6 查询1005号病例确诊后，其所在市新增的所有确诊病例。
- 1-7 在“病例基本信息表”中查询石家庄市在2021年1月11日当天以及之前的所有60岁以上的患者信息。
- 1-8 统计截止到2020年12月30日美国累计确诊病例数最多的10个州。
- 1-9 统计截至2021年1月20号中国发病率最高的人群（人群按照年龄划分，儿童<18，18<=青壮年<60，老年>=60）。

2. 多表查询：

- 2-1 借助病例行程信息粗略查询曾去过“源升品质生活坊”的所有患者的基本信息。
- 2-2 根据病例行程信息表和病例基本信息表，查询行程信息中存在“家庭聚餐”的病例被确诊的日期。
- 2-3 对比中美两国累计确诊病例数，输出格式为(日期，中国累计确诊，美国累计确诊)。
- 2-4 计算截止到2021年1月20日，美国有些县的累计确诊是同一个州的其他县的2倍或以上，列出这些县，以及他们所在的州和他们的累计确诊。
- 2-5 计算世界上人口数排名前10位的国家地区。
- 2-6 列出美国人口超千万的大州中，截至2021年1月20日新冠肺炎疫情死亡率超过2%的州。
- 2-7 截至2021年1月20日，河北省哪些区出现了新冠确诊病例但不属于中高风险地区。
- 2-8 在病例行程信息表的基础上根据病例基本信息表，查询河北省病例的全部信息。

3. 嵌套查询：

- 3-1 查询披露的确诊患者信息中年龄最大的患者，输出其基本信息。(未注明年龄的患者不进行比较)。
- 3-2 查询2020年12月份新增确诊患者最多的城市。
- 3-3 结合“全国各省参考信息表”和“病例基本信息表”给出没有新增确诊病例或未披露病例信息的省份。
- 3-4 2021年1月20日全国中高风险地区所在省中，哪些省在1月20日没有新增确诊信息披露。
- 3-5 根据病例基本信息表查询一月份国内新增患者病例最多的城市。
- 3-6 查询除中美两国以外的其余国家中，进入2021年以来单日新增确诊病例始终不低于一万例的国家。

3.3 实验要求

1. 要求学生独立完成以上内容；
2. 按照实验步骤完成实验后，撰写报告内容。

3.4 实验结果及分析

1-1 查询国内确诊病例基本信息的所有信息来源

执行SQL语句为:

```
1 SELECT DISTINCT 信息来源
2 FROM 病例基本信息
```

查询结果如下:

以下是SELECT DISTINCT 信息来源 FROM 病例基本信息的执行结果集 ⓘ 该表不可编辑。

	信息来源
1	新京报
2	澎湃新闻
3	青冈县疾病预防控制中心
4	大众网
5	今晚报
6	中国新闻网
7	广西新闻网
8	辽宁日报
9	黑龙江卫健委网站
10	河北日报

当前第 1 页 上一页 下一页 50 条/页 跳转到 1 GO 共 44 条 获取总行数 刷新 单行详情

1-2 给出河南省、西藏自治区、台湾省的英文名称和人口数

执行SQL语句为:

```
1 SELECT 英文名称, 人口数
2 FROM 全国各省参考信息表
3 WHERE 中文名称 IN ('河南省', '西藏自治区', '台湾省')
```

查询结果如下:

	英文名称	人口数
1	Henan	96050000
2	Taiwan	23816775
3	Tibet	3440000

1-3 查询2021年1月20日各省现有确诊病例数据, 按现有确诊病例数降序排列输出

执行SQL语句为:

```
1 SELECT 省, 累计确诊-累计死亡-累计治愈 AS 现有确诊
2 from 全国各省累计数据统计表
3 WHERE 日期 = '2021-01-20'
4 ORDER BY 现有确诊 DESC
```


查询结果如下：

1	河北省	829
2	香港特别行政区	766
3	黑龙江省	259
4	吉林省	162
5	上海市	95
6	北京市	44
7	广东省	39
8	陕西省	26
9	辽宁省	25
10	天津市	24

当前第 1 页 上一页 下一页 50 条/页 跳转到 1 GO 共 34 条 获取总行数 刷新 单行详情

1-4 查询顺义区中风险地区的数量

执行SQL语句为：

```
1 SELECT COUNT(*)
2 FROM 全国城市风险等级表
3 WHERE 区 = '顺义区' AND 风险等级 = '中风险地区'
```

查询结果如下：

	count
1	4

1-5 计算截至2021年1月20日全国累计确诊病例数

执行SQL语句为：

```
1 SELECT SUM(累计确诊) AS 全国累计确诊
2 FROM 全国各省累计数据统计表
3 WHERE 日期 = '2021-01-20'
```

查询结果如下：

	全国累计确诊
1	98545

1-6 查询1005号病例确诊后，其所在市新增的所有确诊病例

执行SQL语句为：

```
1 SELECT *
2 FROM 病例基本信息 AS T1
3 WHERE EXISTS (
4     SELECT 市 ,日期
5     FROM 病例基本信息 AS T2
6     WHERE T2.病例号 = 1005 AND T1.日期 >= T2.日期 AND T1.市 = T2.市
7 )
```

通过简单的查询操作找到1005号病例所在市和确诊日期，然后再进行查找，查询结果如下：

	病例号	省	市	区	日期	性别	年龄
1	982	黑龙江省	黑河市	爱辉区	2021-01-07	女	54
2	1004	黑龙江省	黑河市	爱辉区	2021-01-06	女	41
3	1005	黑龙江省	黑河市	爱辉区	2021-01-06	女	18

1-7 在“病例基本信息表”中查询石家庄市在2021年1月11日当天以及之前的所有60岁以上的患者信息

执行SQL语句为：

```
1 SELECT *
2 FROM 病例基本信息
3 WHERE 市 = '石家庄市' AND 日期 <= '2021-01-11' AND 年龄 > 60
```

查询结果如下：

	病例号	省	市	区	日期	性别	年龄
1	925	河北省	石家庄市	井陉县	2021-01-08	女	65
2	741	河北省	石家庄市	藁城区	2021-01-11	女	78
3	743	河北省	石家庄市	藁城区	2021-01-11	女	75
4	744	河北省	石家庄市	藁城区	2021-01-11	女	68
5	746	河北省	石家庄市	藁城区	2021-01-11	女	77
6	747	河北省	石家庄市	藁城区	2021-01-11	男	70
7	751	河北省	石家庄市	藁城区	2021-01-11	女	67
8	755	河北省	石家庄市	藁城区	2021-01-11	女	68
9	757	河北省	石家庄市	藁城区	2021-01-11	女	68
10	758	河北省	石家庄市	藁城区	2021-01-11	女	70

当前第 1 页 上一页 下一页 50 条/页 跳转到 1 GO 共 84 条 获取总行数 刷新 单行详情

1-8 统计截止到2020年12月30日美国累计确诊病例数最多的10个州

执行SQL语句为：

```
1 SELECT 州,SUM(累计确诊) AS 州累计确诊
2 FROM 美国各州县确诊与死亡数统计表
3 GROUP BY 州
4 ORDER BY 州累计确诊 DESC
5 LIMIT 10
```

查询结果如下：

	州	州累计确诊
1	California	120240058
2	Texas	98530861
3	Florida	74199665
4	New York	53766209
5	Illinois	53577656
6	Ohio	36952029
7	Georgia	36808656
8	Pennsylvania	33587581
9	Tennessee	31115899
10	North Carolina	29604852

1-9 统计截至2021年1月20号中国发病率最高的人群（人群按照年龄划分，儿童<18，18<=青壮年<60，老年>=60）

执行SQL语句为：

```
1 SELECT
2     CASE
3         WHEN 年龄 < 18 THEN '儿童'
4         WHEN 年龄 >= 18 AND 年龄 < 60 THEN '青壮年'
5         ELSE '老年'
6     END AS 年龄段,
7     COUNT(*) AS 发病人数
8 FROM 病例基本信息
9 WHERE 日期 <= '2021-01-20'
10 GROUP BY
11     CASE
12         WHEN 年龄 < 18 THEN '儿童'
13         WHEN 年龄 >= 18 AND 年龄 < 60 THEN '青壮年'
14         ELSE '老年'
15     END
16 ORDER BY 发病人数 DESC
17 LIMIT 1;
```

首先将人群按年龄分类并计数，然后利用倒序输出最多发病人数即可，查询结果如下：

	年龄段	发病人数
1	青壮年	696

2-1 借助病例行程信息粗略查询曾去过“源升品质生活坊”的所有患者的基本信息

执行SQL语句为：

```
1 SELECT 病例基本信息.*
2 FROM 病例基本信息 NATURAL JOIN 病例行程信息表
3 WHERE 行程信息 LIKE '%源升品质生活坊%'
```

查询结果如下：

	病例号	省	市	区	日期
1	157	吉林省	通化市	东昌区	2021-01-18
2	167	吉林省	通化市	东昌区	2021-01-18
3	171	吉林省	通化市	东昌区	2021-01-18
4	568	吉林省	通化市	东昌区	2021-01-13

2-2 根据病例行程信息表和病例基本信息表，查询行程信息中存在“家庭聚餐”的病例被确诊的日期

执行SQL语句为：

```
1 SELECT 病例基本信息.日期
2 FROM 病例基本信息 NATURAL JOIN 病例行程信息表
3 WHERE 行程信息 LIKE '%家庭聚餐%'
```

查询结果如下：

	日期
1	2021-01-05
2	2021-01-04
3	2021-01-13
4	2021-01-13

2-3 对比中美两国累计确诊病例数，输出格式为(日期，中国累计确诊，美国累计确诊)

执行SQL语句为：

```
1 SELECT 日期,中国累计确诊,美国累计确诊
2 FROM (
3     SELECT 日期,SUM(累计确诊) AS 中国累计确诊
4     FROM 全国各省累计数据统计表
5     GROUP BY 日期
6 ) AS 中国确诊
7 NATURAL JOIN
8 (
9     SELECT 日期,SUM(累计确诊) AS 美国累计确诊
10    FROM 美国各州县确诊与死亡数统计表
11    GROUP BY 日期
12 ) AS 美国确诊
```

查询结果如下：

	日期	中国累计确诊	美国累计确诊
1	2021-01-09	96691	22225720
2	2020-12-13	94329	16408428
3	2020-12-26	95542	19067574
4	2021-01-05	96311	21140217
5	2020-11-22	92117	12347905
6	2020-11-29	92815	13495104

当前第 1 页 上一页 下一页 50 条/页 跳转到 1 GO 共 60 条 获取总行数 刷新 单行详情

2-4 计算截止到2021年1月20日，美国有些县的累计确诊是同一个州的其他县的2倍或以上，列出这些县，以及他们所在的州和他们的累计确诊

执行SQL语句为：

```
1 SELECT A.县, A.州, A.累计确诊
2 FROM 美国各州县确诊与死亡数统计表 AS A
3 WHERE A.日期 = '2021-01-20'
4     AND A.累计确诊 >= 2 * (
5         SELECT MAX(B.累计确诊)
6         FROM 美国各州县确诊与死亡数统计表 AS B
7         WHERE B.州 = A.州
8             AND B.县 != A.县
9             AND B.日期 = '2021-01-20'
10    )
```

查询结果如下：

	县	州	累计确诊
1	Hennepin	Minnesota	93303
2	Jefferson	Alabama	63040
3	Maricopa	Arizona	428624
4	Douglas	Nebraska	59230
5	Miami-Dade	Florida	350252
6	District of Columbia	District of Columbia	34403

2-5 计算世界上人口数排名前10位的国家地区

执行SQL语句为：

```
1 SELECT *
2 FROM (
3     SELECT 国家, 人口数
4     FROM 参考信息表
5     WHERE 组合码 = 国家 AND 人口数 IS NOT NULL
6 )
7 UNION (
8     SELECT 'china' AS 国家, SUM(人口数) AS 人口数
9     FROM 全国各省参考信息表
10    WHERE 组合码 LIKE '%, china'
11 )
12 ORDER BY 人口数 DESC
13 LIMIT 10
```

查询结果如下：

1	China	1428493105
2	India	1380004385
3	US	329466283
4	Indonesia	273523621
5	Pakistan	220892331
6	Brazil	212559409
7	Nigeria	206139587
8	Bangladesh	164689383
9	Russia	145934460
10	Mexico	127792286

2-6 列出美国人口超千万的大州中，截至2021年1月20日新冠肺炎疫情死亡率超过2%的州

执行SQL语句为：

```

1 SELECT 州, 州累计死亡/州累计确诊 AS 死亡率
2 FROM (
3     SELECT 州, SUM(累计死亡) AS 州累计死亡, SUM(累计确诊) AS 州累计确诊
4     FROM 美国各州县确诊与死亡数统计表 AS T
5     WHERE 日期 = '2021-01-20'
6     GROUP BY 州
7     HAVING 10000000 < (
8         SELECT 人口数
9         FROM 参考信息表 AS S
10        WHERE T.州 = S.省州 AND S.国家 = 'US' AND S.市县 IS NULL
11    )
12 )
13 WHERE 州累计确诊 > 0 AND 死亡率 > 0.02

```

查询结果如下：

	州	死亡率
1	Pennsylvania	.025154142288315
2	New York	.032494641786951

2-7 截至2021年1月20日，河北省哪些区出现了新冠确诊病例但不属于中高风险地区

执行SQL语句为：

```

1 SELECT 区
2 FROM 病例基本信息
3 WHERE 日期 <= '2021-01-20' AND 省 = '河北省' AND 区 IS NOT NULL
4 EXCEPT
5 SELECT DISTINCT 区
6 FROM 全国城市风险等级表
7 WHERE 省 = '河北省'

```

查询结果如下：

1	栾城
2	井陘县
3	裕华市
4	正定

2-8 在病例行程信息表的基础上根据病例基本信息表，查询河北省病例的全部信息

执行SQL语句为：

```

1 SELECT 行程号, 病例号, 年龄, 性别, 患者信息, 日期信息, 行程信息, 其它信息
2 FROM 病例行程信息表 NATURAL JOIN 病例基本信息
3 WHERE 省 = '河北省'

```

查询结果如下：

	行程号	病例号	年龄	性别	患者信息	日期信息	行程信息	其它信息
1	1	1	54	女	确诊病例1：女，54岁，蓟城区增村镇南桥寨村人，为1月17日第13号确诊病例的妻子。	2021年1月1日至10日	除1月3日12时步行到南桥寨惠兰粮油店买菜外，均居家无外出，期间3日、5日、7日3次将	暂无
2	2	1	54	女	确诊病例1：女，54岁，蓟城区增村镇南桥寨村人，为1月17日第13号确诊病例的妻子。	1月11日	转运至蓟城区指定隔离点进行集中隔离医学观察，期间12日、14日、16日、17日4次将	暂无
3	3	1	54	女	确诊病例1：女，54岁，蓟城区增村镇南桥寨村人，为1月17日第13号确诊病例的妻子。	1月18日	转运至石家庄市第三医院发热门诊就诊，核酸检测呈阴性，当日由120负压救护车转运至	暂无
4	4	1	54	女	确诊病例1：女，54岁，蓟城区增村镇南桥寨村人，为1月17日第13号确诊病例的妻子。	1月19日	诊群为确诊病例。	暂无
5	5	10	49	男	确诊病例10：男，49岁，蓟城区增村镇南桥寨村人，为1月13日第22号确诊病例的父兄。	2021年1月4日至12日	居家无外出，期间5日、7日、9日、11日4次核酸检测均为阴性	暂无
6	6	10	49	男	确诊病例10：男，49岁，蓟城区增村镇南桥寨村人，为1月13日第22号确诊病例的父兄。	1月13日	转运至蓟城区指定隔离点进行集中隔离医学观察，期间13日、15日、17日3次核酸检测	暂无
7	7	10	49	男	确诊病例10：男，49岁，蓟城区增村镇南桥寨村人，为1月13日第22号确诊病例的父兄。	1月19日	核酸检测呈阳性，当日由120负压救护车转运至石家庄市人民医院建华院区，诊断为确诊	暂无
8	8	100	85	男	确诊病例16：男，85岁，蓟城区增村镇南桥寨村人。	2020年12月30日至2021年1月10日	居家无外出，期间1月3日、5日、7日3次核酸检测均为阴性	暂无
9	9	100	85	男	确诊病例16：男，85岁，蓟城区增村镇南桥寨村人。	1月11日	转运至蓟城区指定隔离点进行集中隔离医学观察，期间12	暂无

3-1 查询披露的确诊患者信息中年龄最大的患者，输出其基本信息。（未注明年龄的患者不进行比较）

执行SQL语句为：

```
1 SELECT 性别，年龄，患者信息，其它信息
2 FROM 病例基本信息
3 WHERE 年龄 >= ALL (
4     SELECT 年龄
5     FROM 病例基本信息
6     WHERE 年龄 IS NOT NULL
7 )
```

查询结果如下：

	性别	年龄	患者信息	其它信息
1	女	95	张某某，女，95岁	退休人员，家庭住址为沈阳市皇姑区昆山西路塔湾新城小区。

3-2 查询2020年12月份新增确诊患者最多的城市

执行SQL语句为：

```
1 WITH 十二月各市新增确诊 AS (
2     SELECT 市，COUNT(*) AS 病例数
3     FROM 病例基本信息
4     WHERE 日期 >= '2020-12-01' AND 日期 <= '2020-12-31'
5     GROUP BY 市
6 )
7 SELECT 市，病例数
8 FROM 十二月各市新增确诊
9 WHERE 病例数 >= ALL (
10     SELECT 病例数
11     FROM 十二月各市新增确诊
12 )
```

查询结果如下：

	市	病例数
1	大连市	38

3-3 结合“全国各省参考信息表”和“病例基本信息表”给出没有新增确诊病例或未披露病例信息的省份

执行SQL语句为：

```
1 SELECT SUBSTRING(中文名称, 1, 2) AS 省份
2 FROM 全国各省参考信息表
3 EXCEPT (
4     SELECT DISTINCT SUBSTRING(省, 1, 2)
5     FROM 病例基本信息
6 )
```

由于两张表中有如 广西壮族自治区 和 广西省, 新疆维吾尔自治区 和 新疆维吾尔自治区 这样的差别, 所以在筛选的时候只选择前两个字进行匹配

查询结果如下：

	省份
1	福建
2	湖北
3	西藏
4	云南
5	海南
6	甘肃

当前第 1 页 上一页 下一页 50 条/页 跳转到 1 GO 共 21 条 获取总行数 刷新 单行详情

3-4 2021年1月20日全国中高风险地区所在省中, 哪些省在1月20日没有新增确诊病例披露

执行SQL语句为：

```
1 SELECT DISTINCT 省 AS 省份
2 FROM 全国城市风险等级表
3 EXCEPT (
4     SELECT DISTINCT 省
5     FROM 病例基本信息
6     WHERE 日期 = '2021-01-20'
7 )
```

查询结果如下：

	省份
1	吉林省
2	辽宁省

3-5 根据病例基本信息表查询一月份国内新增患者病例最多的城市

执行SQL语句为：


```

1 WITH 一月各市确诊 AS (
2     SELECT 市, COUNT(*) AS 病例数
3     FROM 病例基本信息
4     WHERE 日期 >= '2021-01-01' AND 日期 <= '2021-01-31'
5     GROUP BY 市
6 )
7 SELECT 市, 病例数
8 FROM 一月各市确诊
9 WHERE 病例数 >= ALL(
10     SELECT 病例数
11     FROM 一月各市确诊
12 )

```

查询结果如下：

	市	病例数
1	石家庄市	788

3-6 查询除中美两国以外的其余国家中，进入2021年以来单日新增确诊病例始终不低于一万例的国家

执行SQL语句为：

```

1 WITH 其余国家 AS (
2     SELECT 日期, 国家, SUM(累计确诊) AS 国家累计确诊
3     FROM 各国疫情数据统计表 AS S
4     WHERE 日期 >= '2021-01-01' AND 国家 NOT IN ('US','China') AND NOT EXISTS (
5         SELECT 国家
6         FROM 各国疫情数据统计表 AS T
7         WHERE S.国家 = T.国家 AND T.省州 IS NULL
8     )
9     GROUP BY 日期, 国家
10    UNION
11    SELECT 日期, 国家, 累计确诊
12    FROM 各国疫情数据统计表
13    WHERE 日期 >= '2021-01-01' AND 国家 NOT IN ('US','China') AND 省州 IS NULL
14 ),
15 每日新增确诊 AS (
16     SELECT A.日期, A.国家, A.国家累计确诊 - COALESCE(B.国家累计确诊, 0) AS 新增确诊
17     FROM 其余国家 AS A
18     JOIN 其余国家 AS B
19     ON A.国家 = B.国家 AND A.日期 = B.日期 + INTERVAL '1 day'
20 )
21 SELECT 国家
22 FROM 每日新增确诊
23 GROUP BY 国家
24 HAVING MIN(新增确诊) >= 10000;

```

本题较为复杂，首先各国疫情数据表中有的直接给出了国家每日累计确诊数，有的给的是各州的，需要分别统计并结合为其余国家，然后再通过后一天的累计确诊数减去前一天的累计确诊数计算为每日新增确诊数，再查找大于一万的

查询结果如下：

1	Russia
2	United Kingdom
3	Brazil

3.5 实验总结

在实验中，主要会遇到一些如使用单引号括起字符串，子查询连接时未申明别名的问题，在对多表查询以及嵌套查询中，需要进行一定的规划和思考来找到需要的数据，这些步骤使上课理论与实验实践相结合，加深了我对课程内容的理解和学习，也使我对SQL语法更加熟悉，平时上课没有注意到的细节也在实验中暴露出问题，并对我的错误观念进行了纠正，在这次实验中我受益匪浅。