

华东师范大学软件学院实验报告

实验课程：数据库系统实践

实验报告编号：02

成员：徐玮鸿（10195101536），何启凡（10184304123），周子涵（10195101522）

一、实验环境

datagrip、python、POSTGRE、VUE、VUETIFY、VUE.CLI、Springboot2.5.0、JDK11、mybatis

二、实验内容与实验步骤

6. Choose an appropriate platform for your application;
7. Acquire the large “production” dataset;
8. Test the SQL statements you developed for Task 5 in the large database;

三、实验过程与分析

6. Choose an appropriate platform for your application:

目前我们的代码都放在了 github 上。

Sample database driven web application source code can be found in the GitHub repository: <https://github.com/AidenHe2000/Covid>

前端将通过 VUE, VUETIFY 作为主要框架。

举 medicine 作为 demo:

初始界面如下:

Medicine | Test

生产厂商: 白云 药物种类: 中药

药物:

药物介绍:

显示页数: 1

显示记录数: 84

SEARCH

id	药物名称	生产厂商	药物介绍	药物类型
1	白云退烧胶囊	白云	用于退烧。孕妇慎用。	中药
2	白云退烧颗粒	白云	用于退烧。孕妇慎用。	中药
3	白云退烧冲剂	白云	用于退烧。孕妇慎用。	中药
4	白云退烧片	白云	用于退烧。孕妇慎用。	中药
5	白云清血胶囊	白云	用于退烧。孕妇慎用。	中药
6	白云清血颗粒	白云	用于退烧。孕妇慎用。	中药
7	白云清血冲剂	白云	用于退烧。孕妇慎用。	中药
8	白云清血片	白云	用于退烧。孕妇慎用。	中药
9	白云布洛芬胶囊	白云	用于退烧。孕妇慎用。	中药
10	白云布洛芬颗粒	白云	用于退烧。孕妇慎用。	中药

Rows per page: 10 1-10 of 84

搜索功能展示:

Medicine | Test

生产厂家 康泰克 药物种类 退烧药

药名 退烧

药物介绍

显示页数 1

显示记录数 84

SEARCH

id	药物名称	生产厂家	药物介绍	药物类型
61	康泰克退烧胶囊	康泰克	用于退烧。孕妇慎用。	西药
62	康泰克退烧颗粒	康泰克	用于退烧。孕妇慎用。	西药
63	康泰克退烧冲剂	康泰克	用于退烧。孕妇慎用。	西药
64	康泰克退烧片	康泰克	用于退烧。孕妇慎用。	西药

Rows per page 10 1-4 of 4

翻页功能展示：

Medicine | Test

生产厂家 康泰克 药物种类 退烧药

药名

药物介绍

显示页数 1

显示记录数 84

SEARCH

id	药物名称	生产厂家	药物介绍	药物类型
71	康泰克退烧冲剂	康泰克	用于退烧。孕妇慎用。	西药
72	康泰克退烧片	康泰克	用于退烧。孕妇慎用。	西药

Rows per page 10 1-12 of 12

7.Acquire the large “production” dataset:

本项目数据由Python的Faker库生成。与疫情相关的所有数据被分为八张关系表: hospital, doctor, patient, diagnosis, image, prescription, medicine, user 是用来维护登录信息的, 与其他关系表无关。大多数关系表之间都有外键依赖, 具体情况如下表。

关系表名	外键依赖于
doctor	hospital
patient	doctor, hospital
diagnosis	doctor, patient
prescription	patient, medicine, doctor
image	patient, doctor

每张关系表通过代码生成数据的规模如下所示：

file name	description	scale now
doctor.py	generate doctors	100,000
hospital.py	generate hospitals	10,000
user.py	generate users	200,000
diagnosis.py	generate diagnosis	4,000,000
image.py	generate images	2,000,000
patient.py	generate patients	600,000
medicine.py	generate medicine	84
prescription.py	generate prescription	400,000

8. Test the SQL statements you developed for Task 5 in the large database:
 为了使查询速度在单表查询和多表连接查询时都能够更快，许多关系表在除主键之外的其他属性上（特别是有关外键依赖的属性）建立索引，具体如下：

doctor

索引名	属性
doctor_hospital	hospital_id
main	doctor_id, hospital_id

patient

索引名	属性
index_hospital	patient_id, hospital_id
patient_hospital	hospital_id
patient_doctor	doctor_id
index_doctor	patient_id, doctor_id

diagnosis

索引名	属性
diagnosis_patient	patient_id
diagnosis_doctor	doctor_id

prescription

索引名	属性
prescription_doctor	doctor_id
prescription_medicine	medicine_id
prescription_patient	patient_id

单表查询和多表查询所用时间(最坏情况):

关系表	时间 (ms)
Diagnosis	19
Doctor	37
Doctor, hospital	30
Hospital	16
Image	417
Medicine	32

Patient	15
Patient, hospital, doctor	444
Prescription	15
Prescription, doctor, medicine	32

四、实验结果总结

在本次实验中我们通过 python 完成了大规模随机数据的制造，制作了一个关于 medicine 的 demo，通过前后端分离的方式各自推进自己的工作，并在最后对接成功。目前 medicine 已经能够访问后端数据。不仅如此，我们完成了索引的建立，并通过大量测试得到其查询时间的相关数据。按时完成了任务。

五、附录

我们的仓库 <https://github.com/AidenHe2000/Covid>