

Моргунов Е. П.

Полная инструкция для изучения языка SQL

(МАИ, 2023 г.)

ПОДГОТОВКА РАБОЧЕЙ СРЕДЫ

Для подготовки рабочей среды на своем компьютере можно воспользоваться материалами, размещенными на моем Яндекс-диске <https://disk.yandex.ru/d/PaYZnpoQsR-JaA>. Там находится виртуальная машина, полностью подготовленная к работе, и файл-инструкция по ее установке – Prepare_for_SQL.pdf. В файле readme.pdf также есть полезная информация, начать подготовку к работе можно с него.

УЧЕБНЫЕ МАТЕРИАЛЫ

1. Моргунов, Е. П. PostgreSQL. Основы языка SQL : учеб. пособие / Е. П. Моргунов ; под ред. Е. В. Рогова, П. В. Лузанова. – СПб. : БХВ-Петербург, 2018. – 336 с. : ил. – ISBN 978-5-9775-4022-3. https://edu.postgrespro.ru/sql_primer.pdf

2. Моргунов, Е. П. Администрирование информационных систем [Текст] : учеб. пособие / Е. П. Моргунов, О. Н. Моргунова ; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2015. – 90 с. http://www.morgunov.org/docs/inf_sys_admin.pdf

Здесь находятся все исходные тексты, включенные в это пособие:

<http://www.morgunov.org/programming.html>

Кроме того, эти пособия можно найти в каталоге /home/WORK/Databases в Debian'ах.

3. Презентации и видеозаписи лекций:

<https://postgrespro.ru/education/university/sqlprimer>

Презентация 0 «Вводная лекция» посвящена описанию основ теории баз данных. Она построена на основе учебников по базам данных, перечисленных в конце презентации.

Презентации 1–8 построены на основе учебника по SQL (см. п. 1, главы 3–10).

Презентация 9 «Программирование на стороне сервера» построена на основе учебного пособия (см. п. 2, глава 4). Исходные тексты, приведенные в презентации, находятся здесь <https://edu.postgrespro.ru/sqlprimer/sqlprimer-2019-msu-09.tgz>

Презентация 10 «Полнотекстовый поиск» построена на основе главы 12 документации на Постгрес <https://postgrespro.ru/docs/postgresql/12/textsearch>

Исходные тексты, приведенные в презентации, находятся здесь

<https://edu.postgrespro.ru/sqlprimer/sqlprimer-2019-msu-10.tgz>

ЗАДАНИЯ

Всего в нашем курсе 12 заданий: 11 текущих и одно – финальное.

ДЗ – это домашнее задание.

ДЗ 1. Элементы теории баз данных

Это вводная лекция <https://edu.postgrespro.ru/sqlprimer/sqlprimer-2019-msu-00.pdf> Мы ее не будем рассматривать. Ее могут посмотреть самостоятельно те из вас, у кого не было курса по базам данных, а потом мы можем провести для таких студентов дополнительное занятие, если будет у вас такое желание. Если же специальное занятие по этой лекции не потребуется, но у вас все же будут вопросы по этой лекции, вы можете написать мне письма с вопросами, на которые я постараюсь ответить.

Задание:

1. Подумать над выбором предметной области для выполнения финальной (зачетной) работы. Выбирайте предметную область, которая вам интересна и в которой вы разбираетесь или хотите разобраться.

2. Сделать краткое описание выбранной предметной области (1–2 страницы). Если описание получится более объемным, не беда. Ведь это описание затем войдет в финальный отчет.

3. Попытаться сформулировать требования к будущей базе данных (можно опираться на вводную лекцию). Все презентации лекций находятся здесь:

<https://postgrespro.ru/education/university/sqlprimer>

Фактически, в первой работе нужно выполнить следующие стадии проектирования базы данных:

- Планирование БД
- Определение системы
- Сбор и анализ требований

См. раздел 0.6 «Жизненный цикл разработки системы с базой данных» во вводной лекции.

Планирование БД – это краткое описание тех преимуществ, которые даст компании реализация базы данных, какие проблемы позволит решить эта база данных, что она позволит улучшить в плане обработки данных и т.д.

Определение системы и Сбор и анализ требований можно объединить в один этап.

Конечно, прикладывать формы документов, как это сделано в лекции, не нужно (все-таки это небольшая работа, а не курсовой проект). Напишите требования к данным, требования к транзакциям, т. е. наиболее типичные (вероятные) запросы, которые может пожелать выполнить пользователь. Пусть этих запросов будет не 20 штук, а буквально 4-5, самых важных. Эти запросы на данном этапе пишутся не на языке SQL, а на русском языке. Примеры того, как сформулировать требования к данным и к транзакциям, есть во вводной лекции (см. слайды 97–106).

В результате у вас должно получиться:

– краткое описание выбранной предметной области в виде связного рассказа;

– краткое описание тех преимуществ, которые даст компании или пользователю реализация базы данных, какие проблемы позволит решить эта база данных (это буквально один абзац);

– требования к данным и транзакциям, наиболее типичные запросы пользователя к базе данных, которые должны обеспечить пользователя полезной информацией.

Если вы смешаете в один блок требования к транзакциям и типичные запросы пользователя к БД, это вполне допустимо.

ДЗ 2 ... ДЗ 9 выполняются на основе учебника «PostgreSQL. Основы языка SQL»

https://edu.postgrespro.ru/sql_primer.pdf

Далее указаны главы этого учебника и номера упражнений, которые нужно проделать (в скобках). Упражнения нужно брать из раздела «Контрольные вопросы и задания», который есть в каждой главе учебника.

ДЗ 2. Глава 3 (упражнения 1–4).

ДЗ 3. Глава 4 (упражнения 2, 4, 8, 12, 15, 21, 30, 33, 35).

ДЗ 4. Глава 5 (упражнения 2, 9, 17, 18).

ДЗ 5. Глава 6 (упражнения 2, 7, 9, 13, 19, 21, 23).

ДЗ 6. Глава 7 (упражнения 1, 2, 4).

ДЗ 7. Глава 8 (упражнения 1, 3).

ДЗ 8. Глава 9 (упражнения 2, 3).

ДЗ 9. Глава 10 (упражнения 3, 6, 8).

ДЗ 10. Программирование на стороне сервера.

Выполняется на основе презентации <https://edu.postgrespro.ru/sqlprimer/sqlprimer-2019-msu-09.pdf> и главы 4 учебного пособия «Администрирование информационных систем» (см. выше). Учебную базу данных можно наполнить информацией, функциями и триггерами с помощью команд, выполняемых в командной строке Debian:

```
createdb ais -U postgres
psql -d ais -f adj_list.sql -U postgres
```

Файл adj_list.sql находится в составе архива исходных текстов, прилагаемых к учебному пособию. Они находятся здесь: http://www.morgunov.org/docs/inf_sys_admin_prg.tgz

Эти исходные тексты также есть в виртуальной машине ОС Debian в каталоге /home/WORK/Databases/Admin_DB/UTF-8. В составе этих примеров есть и те, которые не приведены в тексте пособия, но они могут быть полезными при выполнении финального задания.

Задание. Нужно проделать упражнения 12–18 на стр. 86–88 учебного пособия «Администрирование информационных систем». Почти все эти упражнения являются простыми, ознакомительными, не требующими программирования.

ДЗ 11. Полнотекстовый поиск.

Задание выполняется на основе презентации 10 «Полнотекстовый поиск» и главы 12 документации на Постгрес <https://postgrespro.ru/docs/postgresql/12/textsearch>

Задание. Придумать и реализовать пример использования полнотекстового поиска, аналогичный (можно более простой или более сложный) тому примеру с библиотечным каталогом, который был приведен в презентации. Можно использовать исходные тексты, приведенные в презентации: <https://edu.postgrespro.ru/sqlprimer/sqlprimer-2019-msu-10.tgz>

ФИНАЛЬНОЕ ЗАДАНИЕ

1. Спроектировать базу данных для выбранной предметной области.

Самые первые этапы разработки вы выполните в первом задании (см. выше). А в финальном задании выполняете остальные этапы. При этом нужно следовать общей методологии проектирования баз данных: сначала необходимо создать концептуальную модель данных с использованием ER-диаграмм, затем построить логическую модель, выполнив отображение сущностей и связей в отношения, в завершение нужно выполнить физическое проектирование, создав реляционные таблицы в среде целевой СУБД PostgreSQL. Можно использовать ту нотацию, которая вам больше нравится: П. Чена, «вороньи лапки» или UML-нотацию, как в лекции и в учебнике Т. Коннолли. Для рисования диаграмм можно использовать любой бесплатный редактор, умеющий рисовать ER-диаграммы. Можно нарисовать их в Word'e или аккуратно от руки (и сфотографировать).

Число таблиц должно быть равно 8–10 – для группы «Проектирование высоконагруженных Интернет-сервисов», 4–5 – для остальных групп.

Почему дается интервал (8–10 и 4–5)? Потому что число таблиц может быть 8, а не 10, но при этом могут использоваться более разнообразные типы данных и может быть больше столбцов в таблицах.

Обязательно должны быть созданы триггеры (и триггерные функции к ним) и хранимые функции (процедуры) на языке PL/pgSQL. Число триггеров должно быть не менее 2. Число хранимых функций должно быть не менее 2. Эти функции и триггеры не обязательно должны быть сложными. Цель – научиться их применять с пользой для дела.

2. Ввести небольшое количество записей в таблицы базы данных, чтобы можно было продемонстрировать типичные запросы к базе данных. В запросах должны быть использованы:

- подзапросы;
- общие табличные выражения (CTE) (хотя бы одно) и
- оконные функции (Window functions) (хотя бы одна).

3. Создавать приложение (интерфейс пользователя) не требуется.

Конечно, создать его не запрещается, но это не означает, что при отсутствии интерфейса пользователя оценка будет снижена, а при его наличии повышена. Мотивом для его создания может быть стремление к совершенству, желание превратить эту учебную разработку в полноценный продукт и т. д.

В том случае, если приложение не разрабатывалось, нужно заранее подготовить несколько типичных запросов к базе данных и сохранить их в отдельных текстовых файлах. Для демонстрации этих запросов их можно вызывать как извне утилиты psql, так и изнутри нее. Например, для выполнения запроса, содержащегося в файле file_with_request.sql, нужно сделать так:

```
psql -d your_database -f file_with_request.sql -U postgres
```

Для выполнения запроса, сохраненного в файле, изнутри утилиты psql, можно сделать так (конечно, нужно учитывать, в каком каталоге находится этот файл, возможно, будет удобнее и проще написать полный путь к файлу):

```
\i file_with_request.sql
```

4. Подготовить отчет.

Нужно поместить в него:

- краткое описание предметной области и требования к базе данных (т. е. результат выполнения самого первого домашнего задания);
- концептуальную, логическую и физическую схемы базы данных;
- типичные запросы к базе данных, сохраненные в виде отдельных текстовых файлов;
- резервную копию базы данных, созданную с помощью утилиты pg_dump, входящей в состав PostgreSQL. С помощью этой копии можно будет быстро восстановить вашу базу данных и те данные, которые вы в нее ввели.

Краткое описание предметной области, концептуальную и логическую схемы базы данных нужно поместить в документ в формате doc или pdf (как вам удобнее).

Физическую схему базы данных нужно представить в отчете в виде текстового файла, содержащего SQL-команды для создания таблиц, представлений (если они используются), триггеров и триггерных функций. В этом файле должны быть сделаны комментарии. Имея такой файл, создать все объекты базы данных можно будет таким образом:

```
psql -d your_database -f file_with_SQL_commands.sql -U postgres
```

В качестве примера такого файла можно посмотреть файл adj_list.sql в каталоге /home/WORK/Databases/Admin_DB/UTF-8.

Для каждого отношения (таблицы) необходимо указать номер нормальной формы, в которой это отношение находится, и кратко (буквально, в два слова) обосновать, из чего это следует. Если какое-либо отношение не находится хотя бы в 3НФ, необходимо обосновать, почему принято такое проектное решение.

ЗАЧЕТ

1. Краткий рассказ о предметной области.

2. Демонстрация работы с вашей базой данных путем выполнения SQL-запросов, заранее сохраненных в текстовых файлах, сопровождаемая устными пояснениями о порядке их выполнения, смысле выводимых результатов, особенностях конкретных SQL-команд.

3. Ответы на теоретические вопросы, связанные с темой вашей работы.

4. Предоставление отчета.