

# 1

## Прежде чем начать

---

Практически каждое наше действие оставляет цифровой след. Каждая покупка, каждая пройденная миля и каждый переход по ссылке в интернете заносятся в колоссальный, все разрастающийся массив данных, превратившийся для многих компаний в их самый ценный актив. Эти данные, как правило, хранятся в реляционной базе данных (БД), обеспечивающей их безопасность, масштабируемость и доступность для непрерывного чтения и правки бесчисленным количеством пользователей.

Однако как именно пользователи взаимодействуют с данными в реляционной БД? И, что еще важнее, каким образом осуществляется чтение критически важной для вашей компании информации? Ответ на этот вопрос, а также предмет рассмотрения в книге, — *язык структурированных запросов* (Structured Query Language), широко известный под аббревиатурой *SQL*.

### **1.1. Актуальность SQL**

Возможно, вы задаетесь вопросом: а стоит ли SQL того, чтобы потратить на его изучение целых 24 часа? Можете не сомневаться — владение этим языком по праву считается одним из ценнейших навыков для любого специалиста, работающего с информацией. Огромное количество данных в современном мире хранится в системах управления базами данных (СУБД) от разных производителей, таких как Oracle, Microsoft и IBM, и почти все они применяют SQL, причем на протяжении достаточно долгого времени.

В то время как многие языки программирования живут всего несколько лет, SQL остается стандартом для работы с реляционными базами данных на протяжении

уже нескольких десятилетий — и в ближайшем будущем его позиции ничто не угрожает. Таким образом, мастерство, которое вы приобретете и отточите, изучая это руководство и выполняя рекомендуемые упражнения, — это долгосрочная инвестиция в ваше профессиональное будущее.

Возможно, одно из ключевых преимуществ SQL — исключительная простота его освоения, обусловленная тем обстоятельством, что он максимально приближен к естественному английскому языку. Если вы знакомы с английским, команды и синтаксис, применяемые в SQL, покажутся вам интуитивно понятными. К примеру, для выборки имен и фамилий всех покупателей из Канады достаточно составить такой SQL-запрос:

```
SELECT FirstName, LastName FROM Customers WHERE Country = 'Canada';
```

Видите, как все просто? Пусть вам пока и не ясна каждая часть этого выражения, уверяю вас, что по прочтении первых нескольких глав этой книги вы научитесь свободно составлять такие конструкции.

## **1.2. Подходит ли вам эта книга?**

Руководств, учебных пособий, видеокурсов и сайтов, предлагающих обучить вас SQL, существует великое множество. Однако большинство из них рассчитаны на аудиторию с опытом разработки программ. Зачастую они начинаются с изложения истории языка, затем переходят к обсуждению его широкого понятийного аппарата, после чего следуют главы, сгруппированные по принципу демонстрации работы различных команд. Возможно, это и неплохой подход, но он не особо учитывает интересы нетехнических специалистов, таких как я сам, которым необходимо знать SQL.

Хотя я пользуюсь SQL уже свыше 20 лет, начинал я отнюдь не программистом. Впервые я столкнулся с базами данных на должности администратора данных, где я отвечал за импорт данных из различных источников в реляционную БД. Мне нужно было считывать поступившие данные, чтобы проверять, все ли прошло успешно, — и единственным инструментом для этого был SQL, который мне и пришлось освоить.

И хотя тогда я почти ничего не смыслил в программировании, я довольно быстро разобрался с SQL. Если вы умеете писать по-английски, будьте уверены — у вас все получится. Вы убедитесь в этом сами, читая книгу: большинство команд и операторов SQL интуитивно понятны и звучат точь-в-точь как их английские эквиваленты.

### **1.2.1. Спектр применения SQL**

Разумеется, данные нужны не только ИТ-отделу. Взять, к примеру, бизнес-аналитиков: при помощи SQL они могут оперативно извлекать и анализировать

информацию об операционных тенденциях для принятия обоснованных бизнес-решений. Маркетологи могут воспользоваться SQL для выявления полезных инсайтов о недавних рекламных кампаниях и поиска новых точек роста для бизнеса, а финансисты — для получения критически важной бизнес-отчетности, необходимой для обеспечения соответствия нормативным требованиям.

Все эти данные являются основой существования и развития любой современной компании, а успех напрямую зависит от того, способны ли сотрудники самых разных отделов работать с реляционными БД и на основе извлекаемых данных принимать судьбоносные для бизнеса решения. Цель этой книги — помочь вам, как одному из таких специалистов, овладеть SQL и развить соответствующие навыки. Если пока ваш главный инструмент — электронные таблицы, это уже неплохо: вы идеально подготовлены к старту.

Если же вы программист, администратор БД или специалист по анализу данных, эта книга предназначена и для вас — просто метод подачи материала здесь несколько иной. В отличие от большинства учебников, которые с первых страниц обрушивают на вас поток терминологии, вы сразу приступите к решению реальных задач на SQL, постигая концепции и значения терминов постепенно, в ходе работы.

В то же время моя цель — не просто научить вас набору SQL-команд. Продвигаясь шаг за шагом, я покажу вам, как применять элементы языка SQL для решения рабочих задач, вне зависимости от вашего уровня подготовки в области программирования.

### **1.2.2. Диалекты SQL**

Несмотря на то что изучать SQL мы будем на примере MySQL, практически все рассматриваемые нами концепции и приемы работы с языком универсальны и подойдут для любой реляционной базы данных. Таким образом, обретенные вами знания будут применимы к любой из следующих СУБД:

- IBM DB2;
- MariaDB;
- Microsoft SQL Server;
- MySQL;
- Oracle;
- PostgreSQL.

Сформировав прочный фундамент знаний в области применения SQL, вы сможете работать с данными в любой из перечисленных систем. Правда, стоит иметь в виду: в каждой из них встречаются свои особенности. Я буду отмечать их по ходу книги, чтобы вы чувствовали себя уверенно с любым диалектом SQL.

### **1.2.3. ИИ и SQL**

С появлением генеративного искусственного интеллекта (ИИ) резонно задать вопрос: а стоит ли вообще изучать SQL, когда можно просто попросить ИИ-ассистента, скажем, ChatGPT, написать любой необходимый SQL-код? Безусловно, ИИ выглядит соблазнительной альтернативой изучению языка. Но подвох вот в чем — без прочного знания SQL вы даже не сможете проверить, будет ли код, предоставленный таким инструментом, давать надлежащие результаты. К тому же, чтобы нейросеть сгенерировала конкретный SQL-запрос, вам нужно предоставить сведения о вашей базе данных, что может быть запрещено корпоративными правилами.

Впрочем, у генеративного ИИ есть и своя ниша — он может стать отличным помощником для тех, кто уже уверенно владеет SQL. Когда вы достигнете уровня, на котором будете точно понимать, как работает ваш код, эти инструменты можно будет задействовать для быстрого поиска узких мест в производительности или для анализа логики выполнения запросов. Однако для интерпретации предоставляемых рекомендаций и объяснений все равно необходимы знания SQL, сформировать которые и призвано это руководство.

## **1.3. Как построена книга**

Концепция работы с книгой проста: по одной главе в день. Разумеется, читать можно не только в обеденный перерыв, хоть это и идеальное время: на главу уйдет минут 40, и у вас еще останется минут 20 на то, чтобы закрепить материал на практике, пока вы допиваете кофе.

### **1.3.1. Основные главы**

Главы 1 и 2 помогут вам быстро войти в курс дела. Они не только объяснят, что такое таблица и как правильно формулировать запросы, но и познакомят с инструментарием, который пригодится вам в дальнейшем. В некотором смысле это самые важные главы во всей книге.

Главы с 3-й по 22-ю содержат основной учебный материал, на освоение которого уйдет около месяца, даже если месяц такой короткий, как февраль. Не на каждую главу потребуется целый час, но очень важно идти по порядку: каждая следующая глава опирается на сведения и команды из предыдущих. Можно, конечно, читать и по нескольку глав в день, но я настоятельно рекомендую не спешить: осваивайте по одной главе в день и не жалейте времени на практику. Умение сосредоточиться на небольшом наборе понятий и примеров обеспечивает оптимальный режим для быстрого закрепления знаний. Как говаривал великий баскетбольный тренер Джон Вуден: «Поспешай не торопясь».

### **1.3.2. Практические занятия**

Почти в каждой главе вас ждет небольшой практикум — специальное задание, позволяющее незамедлительно опробовать новые команды и концепции в деле. Относитесь к этим упражнениям не как к контрольной проверке ваших знаний, а как к возможности применить и закрепить ваши новые навыки работы с SQL. Разумеется, ответы на практические задачи приведены в конце каждой главы, но поверьте: именно их самостоятельное прохождение — ключ к эффективному усвоению пройденного материала.

### **1.3.3. Дальнейшее изучение**

Книга призвана стать отправной точкой для тех, кто делает первые шаги в SQL и лишь приоткрывает дверь в мир возможностей работы с реляционными базами данных. Поэтому в конце некоторых глав вы найдете рекомендации по дальнейшему изучению рассмотренных аспектов языка. Если у вас есть время и желание, ознакомьтесь с этими ресурсами, чтобы обогатить ваш растущий набор профессиональных навыков в области SQL.

## **1.4. Подготовка рабочего места**

Не теряя времени, приступим к настройке всего, что необходимо для практики. Вся процедура займет буквально несколько минут и не потребует мощного компьютера. Нам нужно установить две бесплатные программы, а затем выполнить готовый SQL-сценарий, чтобы сгенерировать данные для практикума.

### **1.4.1. Установка MySQL и MySQL Workbench**

Первый шаг — загрузка MySQL и ее установка на выбранный вами компьютер. MySQL не только совершенно бесплатна, но и входит в число самых популярных реляционных СУБД в мире.

Мы также установим MySQL Workbench — именно в этой среде мы и будем выполнять все запросы из книги. Она потребляет минимальный объем ресурсов, так что можете смело ставить MySQL Workbench даже на ноутбук.

Инструкции по скачиванию и установке этих двух приложений доступны в моем GitHub-репозитории по адресу <https://mng.bz/PN18>. Поскольку программный пакет MySQL часто обновляется, номера версий, которые вы увидите, могут быть новее, чем указанные в документации. Не беспокойтесь об этом — все примеры из книги будут работать и на свежих сборках.

### **1.4.2. Выполнение практических работ**

На протяжении всей книги мы будем работать с одним и тем же набором данных, размещенным в базе данных `sqlnovel`. Это заказы вымышленного издательства, которое специализируется на книгах про SQL. Мы рассмотрим эти данные немного позднее, а пока создадим базу данных и заполним ее тестовым содержанием, запустив подготовленный SQL-сценарий.

Инструкции по настройке базы данных `sqlnovel` также размещены по адресу <https://mng.bz/PN18>, и они еще проще, чем процедура установки MySQL и MySQL Workbench. Сейчас вы просто запустите готовый сценарий, а ближе к концу книги мы вернемся к нему и разберем, как он устроен. К тому моменту вы уже сможете создавать собственные наборы данных!

## **1.5. Полезные онлайн-ресурсы**

В книге вас ждет множество примеров и заданий для самостоятельной работы. Я призываю вас набирать все сценарии самостоятельно и при желании оформлять SQL-код в стиле, отличном от представленного в этом руководстве. Так вы лучше почувствуете язык. При наборе SQL-инструкций вы можете столкнуться с непонятными вам ошибками. В связи с этим в онлайн-репозитории размещены все SQL-сценарии, фигурирующие в издании. Заглядывайте туда лишь для того, чтобы свериться или найти причину ошибки. Поверьте, когда вы пишете код сами, а не копируете его, вы учитесь в разы быстрее.

## **1.6. SQL в действии — с первого же дня**

Подобно другим изданиям серии «За 24 часа», первостепенная задача книги — прямо в процессе освоения материала научить вас эффективно применять его на практике. Так, каждая глава посвящена конкретному компоненту языка SQL и дает его краткий обзор, при этом основное внимание уделяется сценариям из реальной жизни. Дополнительно в конце каждой главы предусмотрено закрепление обретенных навыков: выполнение упражнений в форме практической работы.

Как я отмечал ранее, если ваша цель — углубленное изучение теории и истории реляционных баз данных, существует масса других изданий, способных стать хорошим помощником на этом пути. Несмотря на то что во многих частях руководства обсуждаются технические детали и тонкости, каждая глава подчинена одной цели — обеспечить вашу оперативную результативность при решении прикладных задач.

Итак, достаточно о книге. Приступаем к работе с SQL!

# Ваш первый SQL-запрос

---

Первую главу мы закончили словами об оперативной результативности — и сейчас безотлагательно перейдем к делу! Начнем с рассмотрения того, как данные хранятся в реляционной базе, и разберемся с их структурой. Это поможет вам увереннее оперировать терминами, которые встретятся в книге. Не волнуйтесь: терминов будет совсем немного, да и слова эти хорошо вам известны и применяются в обычной жизни. Просто теперь мы уточним их в контексте хранения данных в реляционных БД.

Кроме того, вы построите ваш первый запрос. На всякий случай уточним: под *запросом* (query) подразумевается выполнение SQL-команды с целью извлечения данных из базы. Чем дальше вы углубитесь в книгу, тем больше запросов напишете — и тем увереннее будете себя чувствовать в SQL. Если вы еще не установили MySQL и MySQL Workbench (см. главу 1) и не запустили сценарий `Create_SQLNovel_database.sql` для создания учебной БД, — самое время это сделать, чтобы можно было сразу приступить к работе. Однако, прежде чем начать выполнять запросы, давайте сначала взглянем на сами данные.

## **2.1. Кто работал с Excel — уже знаком с таблицами БД**

Хотя для изучения SQL это и не обязательно, опыт работы с Excel или другими табличными редакторами очень пригодится. Быть может, вы об этом и не задумывались, но структура электронных таблиц весьма схожа с устройством ключевых объектов БД. В этой главе мы познакомимся с рядом терминов, которые помогут разобраться, как хранятся данные в реляционной базе и как с ними работает система управления базами данных (СУБД).

Мы не просто собираем данные и загружаем их в реляционную БД как попало; мы организуем и сохраняем их в объектах, соответствующих характеру самих данных. Эти объекты называются *таблицами*. Как правило, мы группируем данные в таблицы, соотнося их с элементами предметной области, такими как заказы, клиенты или платежи. Таблицы являются составными компонентами любой реляционной базы данных и по своей структуре во многом похожи на электронные таблицы, поэтому сравнение с таблицами Excel поможет понять термины, применяемые в этой главе и далее в книге.

Если вы не знакомы с основными терминами, служащими для описания электронной таблицы, взгляните на типичную электронную таблицу на рис. 2.1, содержащую сведения о нескольких книгах про SQL (разумеется, вымышленных)<sup>1</sup>. Считайте такого рода информацию массивом данных, или, как часто говорят, *набором данных* (dataset). Все довольно просто, не так ли? Этот набор представлен в электронной таблице, однако если бы он хранился в таблице базы данных, его структура осталась бы практически идентичной.

	A	B	C	D	E
1	<b>Название</b>	<b>Цена</b>	<b>Аванс</b>	<b>Гонорар</b>	<b>Дата публикации</b>
2	Pride and Predicates	\$ 9.95	\$ 5,000.00	15%	4/30/2015
3	The Join Luck Club	\$ 9.95	\$ 6,000.00	12%	2/6/2016
4	The Catcher in the Try	\$ 8.95	\$ 5,000.00	10%	4/3/2017
5	Anne of Fact Table	\$ 12.95	\$ 10,000.00	15%	1/12/2018
6	The DateTime Machine	\$ 7.95	\$ 5,500.00	15%	2/4/2019
7	The Great GroupBy	\$ 10.95	\$ -	20%	12/23/2019
8	The Call of the While	\$ 8.95	\$ 2,500.00	15%	3/14/2020
9	The Sum Also Rises	\$ 7.95	\$ 5,000.00	12%	11/12/2021

**Рис. 2.1.** Пример электронной таблицы, содержащей пять столбцов (A–E) с информацией о книгах. Ее структура очень похожа на таблицу базы данных

Я обещал не сыпать сложными терминами, но три основных понятия, связанных с таблицами, следует усвоить обязательно, прежде чем вы начнете применять

<sup>1</sup> В заглавиях вымышленных «книг про SQL» обыгрываются названия следующих классических литературных произведений: «Гордость и предубеждение» Джейн Остин (*Pride and Prejudice*), «Клуб радости и удачи» Эми Тан (*The Joy Luck Club*), «Над пропастью во ржи» Джерома Сэлинджера (*The Catcher in the Rye*), «Энн из Зеленых Мезонинов» Люси Мод Монтгомери (*Anne of Green Gables*), «Машина времени» Герберта Джорджа Уэллса (*The Time Machine*), «Великий Гэтсби» Фрэнсиса Скотта Фицджеральда (*The Great Gatsby*), «Зов предков» Джека Лондона (*The Call of the Wild*), «Фиеста» Эрнеста Хемингуэя (*The Sun Also Rises*). — *Примеч. пер.*

SQL для чтения и обработки данных. В начале главы я уже говорил, что вы наверняка уже слышали эти слова раньше:

- столбец;
- строка;
- значение.

Следуя самому простому определению, таблица — это структура из одного или нескольких *столбцов* (column) данных. Столбцы располагаются вертикально, подобно архитектурным колоннам. На рис. 2.2 мы видим столбцы «Название», «Цена», «Аванс», «Гонорар» и «Дата публикации», причем столбец «Название» выделен. Иногда столбец называют *полем* (field), однако в формальной спецификации языка SQL такой термин не применяется.

Следующий ключевой термин — это *строка* (row). Строками являются горизонтальные записи данных в таблице. Каждая такая запись представляет собой отдельный элемент из множества тех объектов, которым посвящена приводимая таблица, — в нашем примере одну конкретную книгу. Посмотрите на рис. 2.3: все строки имеют одинаковую структуру и повторяют один и тот же порядок столбцов. Это необходимо, поскольку в каждой строке должны быть представлены все столбцы таблицы.

	A	B	C	D	E
1	<u>Название</u>	<u>Цена</u>	<u>Аванс</u>	<u>Гонорар</u>	<u>Дата публикации</u>
2	Pride and Predicates	\$ 9.95	\$ 5,000.00	15%	4/30/2015
3	The Join Luck Club	\$ 9.95	\$ 6,000.00	12%	2/6/2016
4	The Catcher in the Try	\$ 8.95	\$ 5,000.00	10%	4/3/2017
5	Anne of Fact Table	\$ 12.95	\$ 10,000.00	15%	1/12/2018
6	The DateTime Machine	\$ 7.95	\$ 5,500.00	15%	2/4/2019
7	The Great GroupBy	\$ 10.95	\$ -	20%	12/23/2019
8	The Call of the While	\$ 8.95	\$ 2,500.00	15%	3/14/2020
9	The Sum Also Rises	\$ 7.95	\$ 5,000.00	12%	11/12/2021

**Рис. 2.2.** Электронная таблица с названиями книг. Выделенный столбец «Название» подчеркивает вертикальное расположение столбцов

Эти строки также пронумерованы в левой боковой панели, причем, в таблицах не всегда присутствует явная нумерация. Следует отметить, что слова «*строка*» (row) и «*запись*» (record) часто означают одно и то же — в некоторых программах строки называются записями. Тем не менее в большинстве реляционных СУБД корректным термином для обозначения горизонтальных записей в таблицах является именно «*строка*».

	A	B	C	D	E
1	<b>Название</b>	<b>Цена</b>	<b>Аванс</b>	<b>Гонорар</b>	<b>Дата публикации</b>
2	Pride and Predicates	\$ 9.95	\$ 5,000.00	15%	4/30/2015
3	The Join Luck Club	\$ 9.95	\$ 6,000.00	12%	2/6/2016
4	The Catcher in the Try	\$ 8.95	\$ 5,000.00	10%	4/3/2017
5	Anne of Fact Table	\$ 12.95	\$ 10,000.00	15%	1/12/2018
6	The DateTime Machine	\$ 7.95	\$ 5,500.00	15%	2/4/2019
7	The Great GroupBy	\$ 10.95	\$ -	20%	12/23/2019
8	The Call of the While	\$ 8.95	\$ 2,500.00	15%	3/14/2020
9	The Sum Also Rises	\$ 7.95	\$ 5,000.00	12%	11/12/2021

**Рис. 2.3.** Электронная таблица с названиями книг. Выделенная первая строка подчеркивает горизонтальное расположение строк

И наконец, последний (по крайней мере, на текущий момент) термин — это *значение* (value), которое представляет собой отдельный фрагмент данных на пересечении строки и столбца таблицы. Каждая строка содержит по одному значению для каждого столбца. Так, на рис. 2.4 значение столбца «Название» в последней строке нашего набора данных — «The Sum Also Rises», а значение столбца «Цена» в этой строке — «\$7.95». Стоит отметить, что хотя в каждой строке должны быть представлены все столбцы, сами значения могут оставаться пустыми. К примеру, в строке с названием «The Great GroupBy» значение «Аванс» не заполнено.

	A	B	C	D	E
1	<b>Название</b>	<b>Цена</b>	<b>Аванс</b>	<b>Гонорар</b>	<b>Дата публикации</b>
2	Pride and Predicates	\$ 9.95	\$ 5,000.00	15%	4/30/2015
3	The Join Luck Club	\$ 9.95	\$ 6,000.00	12%	2/6/2016
4	The Catcher in the Try	\$ 8.95	\$ 5,000.00	10%	4/3/2017
5	Anne of Fact Table	\$ 12.95	\$ 10,000.00	15%	1/12/2018
6	The DateTime Machine	\$ 7.95	\$ 5,500.00	15%	2/4/2019
7	The Great GroupBy	\$ 10.95	\$ -	20%	12/23/2019
8	The Call of the While	\$ 8.95	\$ 2,500.00	15%	3/14/2020
9	The Sum Also Rises	\$ 7.95	\$ 5,000.00	12%	11/12/2021

**Рис. 2.4.** Выделено значение столбца «Цена» в строке с названием «The Sum Also Rises» — одно из множества других значений в таблице

Ну что ж, с табличной терминологией пока закончим. Теперь самое интересное — давайте узнаем, как с помощью всего этого делать запросы к таблицам!