KDnuggets

7 Steps to Mastering SQL for Data Science

Previous post
Next post

Tags: 7 Steps, Data Science, Database, Relational Databases, SQL

Follow these 7 steps to go from SQL data science newbie to seasoned practitioner quickly. No nonsense, just the necessities.

Pages: 1 2

By Matthew Mayo, KDnuggets.

Step 4: Creating, Dropping, Deleting

Our second set of commands include those used to CREATE and DROP tables, as well as to DELETE records. With an understanding of this growing collection of commands, suddenly much of what could be referred to as *regular data management and query* is attainable (with practice, of course).

Create

- Creating Tables (from SQL Course)
- CREATE Table (from SQL Zoo)

Drop

- Drop a Table (from SQL Course)
- DROP Table (from SQL Zoo)

Delete

- Deleting Records (from SQL Course)
- DELETE (from SQL Zoo)

Step 5: Views and Joins

On to some slightly more advanced SQL topics. First, we have a look at views, which can be thoughts of as virtual tables populated by the results of queries, useful for a number of different scenarios including application development, data security, and eased data sharing.

First, get the details on what views are:

- Using Views (from Tutorials Point)

For beginner SQL practitioners in the data science realm, I would call views a "nice to have." Since the focus is likely to be more on data exploration, I would say that the next topic, joins, is a "must have."

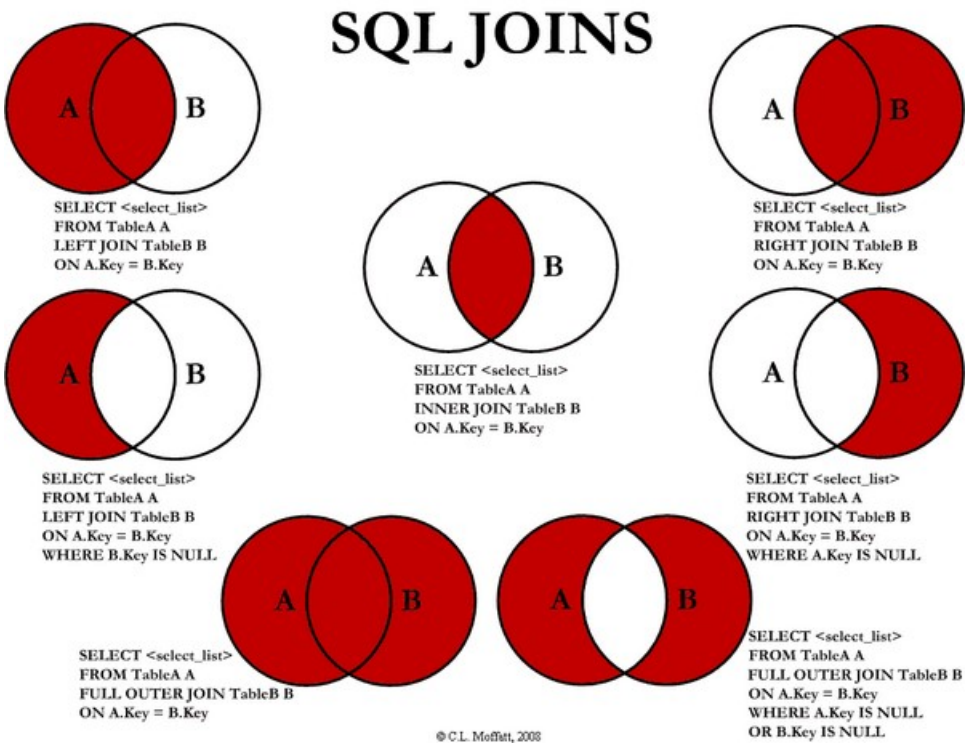
Read about what joins are, and why they are important (and get some examples):

- What Are Joins? (from sql-joins.com)
- SQL JOINS (from Mode Analytics)

Joins come in different flavors, and likely one of the more complex topics you will cover while learning SQL is getting them straight. That's really more of a testament to the ease of SQL than the actual difficulty of learning about joins.

Watch a video explaining inner joins, then check out a video on outer and cross joins:

- Inner Joins
- Outer & Cross Joins



A visual representation of SQL joins.

See this visual representation of SQL joins:

- Visual Representation of SQL Joins (from Code Project)

Finally, this tutorial reviews both joins and views:

- MySQL Tutorial 2: Views and Joins (from arachnoid.com)

Step 6: SQL for Data Science

OK, so you've made some headway with learning SQL. You can query some data, create and manage some tables, make a view if you had to, and even use joins in some more complex querying. But why are you learning all this again? For data science, right? Let's take a bit of a break from the technical to get an idea of this very topic.

Here are a couple of discussions regarding what SQL can be used for in data science:

- SQL For Data Science (from yhat)
- How do data scientists use SQL? (from Quora)

Step 7: SQL Integration with Python, R

Often we will find that SQL is embedded in, or called from within, software written in other programming languages as part of a larger system. For example, in web development you may find PHP or Ruby or some other language making calls to a database via SQL to input, modify, or retrieve application-related data. In data science, you may see SQL being called as part of some application written in, perhaps, Python or R. To that end, having an understanding of how these languages play with SQL is not a bad idea.

SQL queries with Python and SQLite.

Python

To gain an understanding of how Python and SQL can work together, read Sebastian Raschka's fantastic and detailed post on using SQLite in Python:

- SQLite in Python Tutorial

R

Here are a pair of resources for achieving R and SQL integration, which approach the topic from different directions:

- SQL and R (from Simple Talk)
- RSQLite

Further

If you feel like a relentless regimen of reading about an SQL topic and following up with exercises, I recommend going through the following (freely-available) book:

- Learn SQL the Hard Way, by Zed A. Shaw

Related:

- 7 Steps to Mastering Machine Learning With Python
- 7 Steps to Understanding Deep Learning
- R Learning Path: From beginner to expert in R in 7 steps

Pages: 1 2

Previous post

Next post

Top Stories Past 30 Days

Most Popular

1. Top 10 Machine Learning Algorithms for Beginners
2. Top 10 TED Talks for Data Scientists and Machine Learning Engineers
3. Quantum Machine Learning: An Overview
4. Comparing Machine Learning as a Service: Amazon, Microsoft Azure, Google Cloud AI
5. The Art of Learning Data Science
6. Want to Become a Data Scientist? Try Feynman Technique
7. Docker for Data Science

Most Shared

1. Top 10 TED Talks for Data Scientists and Machine Learning Engineers
2. Comparing Machine Learning as a Service: Amazon, Microsoft Azure, Google Cloud AI
3. The Art of Learning Data Science
4. How Docker Can Help You Become A More Effective Data Scientist
5. A Beginners Guide to Data Engineering – Part I
6. How To Grow As A Data Scientist
7. Web Scraping Tutorial with Python: Tips and Tricks

Latest News

- Expanding Self-Service BI with AI-Powered Analytics, Fe...
- Which Machine Learning Algorithm be used in year 2118?
- Introduction to Python Ensembles
- Top 15 Scala Libraries for Data Science in 2018
- Pray.com: Sr Data Engineer
- Join RE•WORK & AI experts in London, Boston and ...

Top Stories
Last Week

Most Popular

1. NEW Comparing Machine Learning as a Service: Amazon, Microsoft Azure, Google Cloud AI
2. ▲ Top 10 Machine Learning Algorithms for Beginners

3. NEW **Want to Become a Data Scientist? Try Feynman Technique**
4. NEW **Data Structures Related to Machine Learning Algorithms**

5. NEW **The 8 Neural Network Architectures Machine Learning Researchers Need to Learn**
6. NEW **How To Grow As A Data Scientist**
7. NEW **Top 10 TED Talks for Data Scientists and Machine Learning Engineers**

Most Shared

1. **Web Scraping Tutorial with Python: Tips and Tricks**

2. **Data Structures Related to Machine Learning Algorithms**
3. **The 8 Neural Network Architectures Machine Learning Researchers Need to Learn**
4. **Avoid Overfitting with Regularization**
5. **Understanding Learning Rates and How It Improves Performance in Deep Learning**
6. **My Journey into Deep Learning**
7. **Automated Text Classification Using Machine Learning**

More Recent Stories

- Join RE•WORK & AI experts in London, Boston and Hong Ko...
- Why Data Scientists Must Know About Change Management
- 5 Machine Learning Projects You Should Not Overlook
- Fast.ai Lesson 1 on Google Colab (Free GPU)
- Top tweets, Jan 31 – Feb 6: #DeepLearning for Natural...
- AI & Machine Learning: the key skills every software engi...
- Building a Daily Bitcoin Price Tracker with Coindeskr and Shin...
- Deep Feature Synthesis: How Automated Feature Engineering Works
- KDnuggets 18:n06, Feb 7: 5 Fantastic Practical Machine Lear...
- NYU Stern MS in Business Analytics
- Register for DataScience: Elevate Livestream, Feb 22
- Top January Stories: Docker for Data Science; Top 10 TED Talks...
- 2018 Predictions for the Analytics & Data Science Hiring ...
- 5 Fantastic Practical Machine Learning Resources
- The Doing Part of Learning Data Science
- The Voleon Group: Machine learning research, software developm...
- Future Trends in Biometrics
- Challenge Yourself to Think, Mar 19-22, Las Vegas
- Top Stories, Jan 29 – Feb 4: Web Scraping Tutorial with ...
- A Simple Starter Guide to Build a Neural Network

X