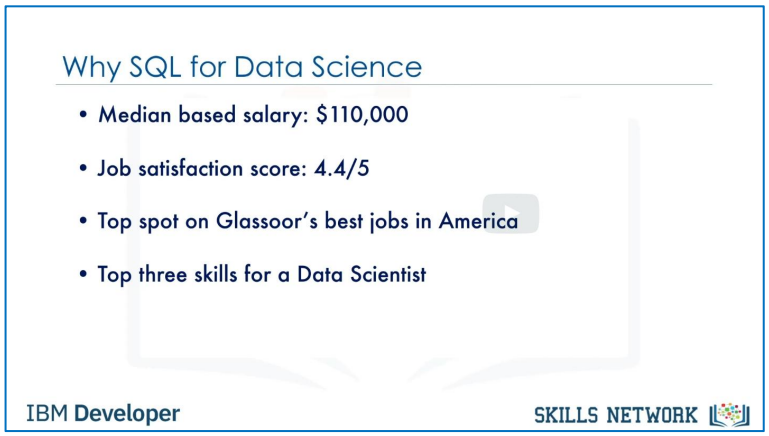
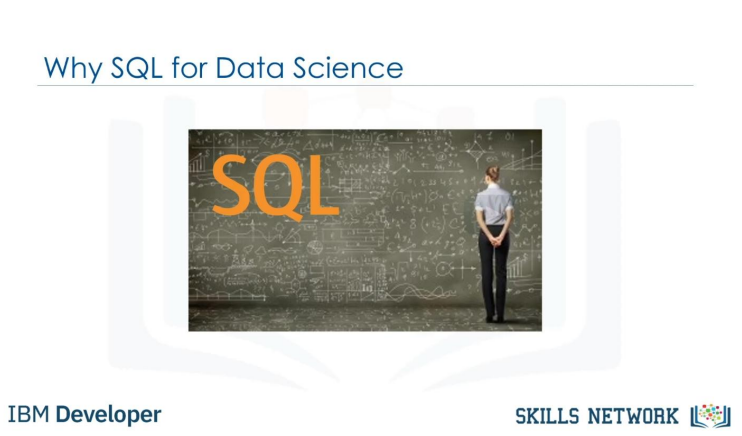
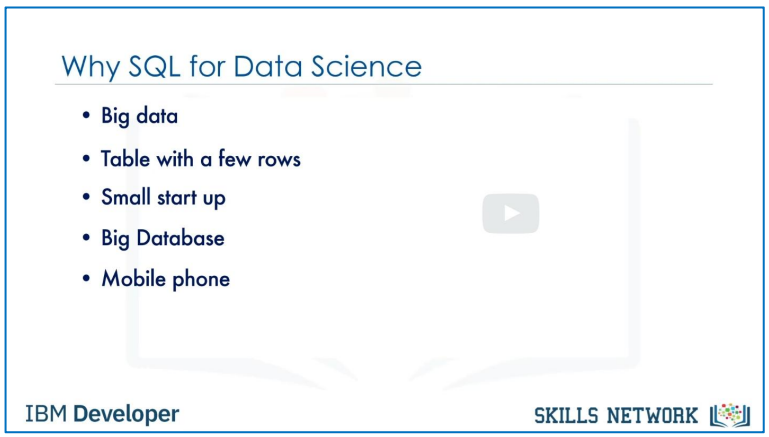
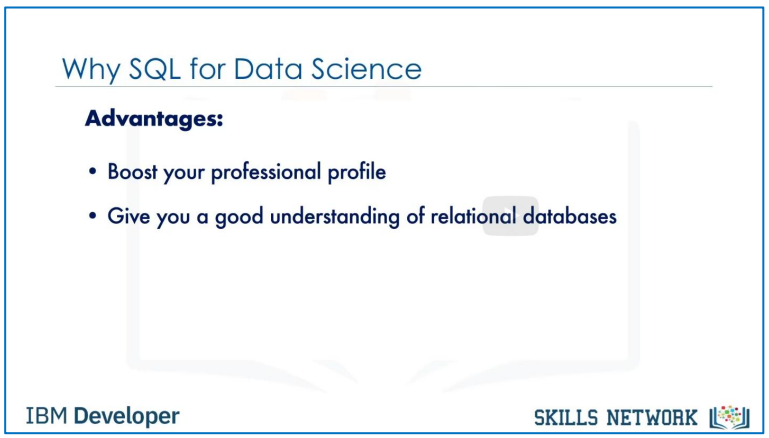
About this Course

Welcome to SQL for Data Science







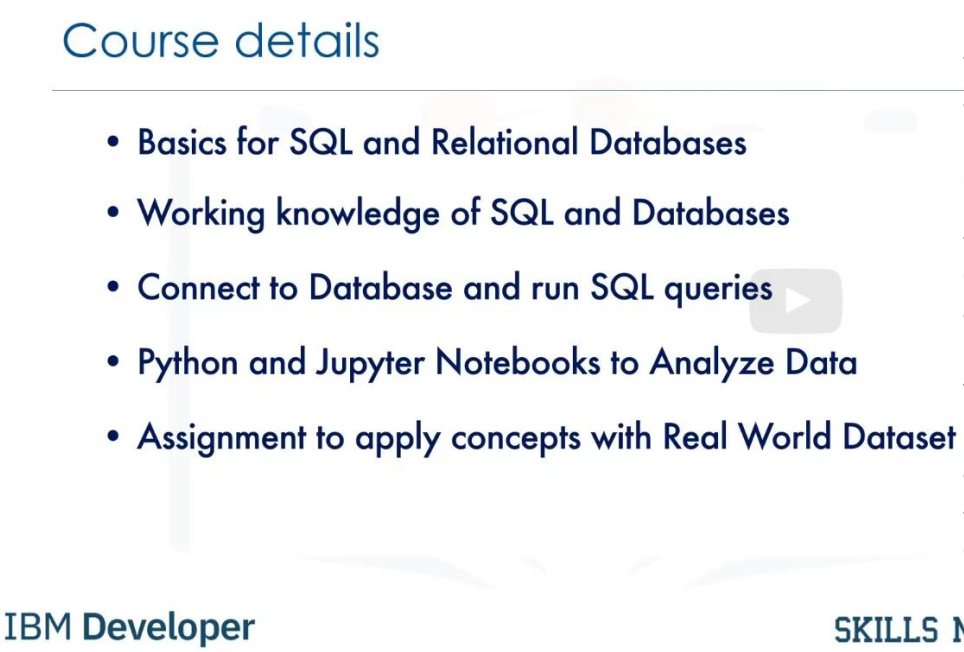


Before you step into the field of data science, it is vitally important that you set yourself apart by mastering the foundations of this field. One of the foundational skills that you will require is SQL.

SQL is a powerful language that's used for communicating with databases. Every application that manipulates any kind of data needs to store that data somewhere; whether it's big data, or just a table with a few simple rows for government, or a small startup, or a big database that spans over multiple servers or a mobile phone that runs its own small database.

Here are some of the advantages of learning SQL for someone interested in data science.
- SQL will boost your professional profile as a data scientist, as it is one of the most sought after skills by hiring employers.
- Learning SQL will give you a good understanding of relational databases. Tapping into all this information requires being able to communicate with the databases that store the data.

Even if you work with reporting tools that generate SQL queries for you, it may be useful to write your own SQL statements so that you need not wait for other team members to create SQL statements for you.



In this course, you will learn the basics of both the SQL language and relational databases.

The course includes interesting quizzes and hands on lab assignments, where you can get experience working with databases.
In the first few modules, you work directly with the database and develop a working knowledge of SQL.

Then, you will connect to a database and run SQL queries like a data scientist typically would, where you will use Python and Jupyter notebooks to connect to relational databases to access and analyze data.
There is also an assignment included towards the end of the course, where you will get an opportunity to apply the concepts that you learned.

总结： So, let's get started with SQL for data science.

Course Overview

Much of the world's data resides in databases. SQL (or Structured Query Language) is a powerful language which is used for communicating with and extracting data from databases. A working knowledge of databases and SQL is a must if you want to become a data scientist.

The purpose of this course is to introduce relational database concepts and help you learn and apply foundational knowledge of the SQL language. It is also intended to get you started with performing SQL access in a data science environment.

The emphasis in this course is on hands-on and practical learning. As such, you will work with real databases, real data science tools, and real-world datasets. You will create a database instance in the cloud. Through a series of hands-on labs, you will practice building and running SQL queries. You will also learn how to access databases from Jupyter notebooks using SQL and Python.

Who should take this course?

This course is intended for existing or aspiring Data Scientists who want practical knowledge of SQL to query databases and know-how to execute SQL from Jupyter Notebooks using Python. It is also useful for Data Analysts, Application Developers, and Data Engineers.

Pre-requisites

There are no prerequisites for this course other than general familiarity with using computers and a desire to learn. No prior knowledge of databases, SQL, or programming is required. Some experience with Python will be an asset, it is not a must, as we will share the Python working knowledge required for completing this course.

Learning Objectives

In this course you will learn about:

- The fundamentals of relational databases and basic SQL commands that you can use to create, manage and query them.
- More advanced SQL commands that enable you to group and sort the results of queries, use built-in functions, and include results from multiple tables.
- Accessing databases programmatically with Python using Jupyter Notebooks.

总结:

Syllabus

Module 1 - Getting Started with SQL

- Introduction to Databases
- SELECT Statement
- SELECT Statement Examples
- Hands-on Lab: Simple SELECT Statements
- COUNT, DISTINCT, LIMIT
- Hands-on Lab: COUNT, DISTINCT, LIMIT
- INSERT Statement
- UPDATE and DELETE Statements
- Summary & Highlights
- Practice Quiz
- Graded Quiz

Module 2 - Introduction to Relational Databases and Tables

- Relational Database Concepts
- How to Create a Database Instance on Cloud
- Hands-on Lab: Sign up for IBM Cloud, Create Db2 Service Instance and Get started with the Db2 Console
- Types of SQL Statements (DDL vs. DML)
- CREATE TABLE Statement
- ALTER, DROP and Truncate Tables
- Examples to CREATE and DROP Tables
- Hands-on Lab: CREATE, ALTER, TRUNCATE, DROP
- Hands-on Lab: Create and Load Tables using SQL Scripts
- Summary & Highlights
- Practice Quiz
- Graded Quiz

Module 3 - Intermediate SQL

- Using String Patterns and Ranges
- Sorting Result Sets
- Grouping Result Sets
- Hands-on Lab: String Patterns, Sorting & Grouping
- Summary & Highlights
- Practice Quiz
- Built-in Database Functions
- Date and Time Built-in Functions
- Hands-on Lab: Built-in Functions
- Sub-Queries and Nested Selects
- Hands-on Lab: Sub-Queries and Nested SELECTs
- Working with Multiple Tables
- Hands-on Lab: Working with Multiple Tables
- Summary & Highlights
- Practice Quiz
- Graded Quiz

Module 4 - Accessing Databases using Python

- How to Access Databases Using Python
- Writing Code Using DB-API
- Connecting to a Database Using ibm_db API
- Hands-on Lab: Create Database Credentials
- Hands-on Lab: Connecting to a Database Instance
- Creating Tables, Loading Data, and Querying Data
- Hands-on Lab: Creating Tables, Inserting and Querying Data
- Introducing SQL Magic
- Hands-on Lab: Analyzing a Real World Data Set
- Summary & Highlights
- Practice Quiz
- Graded Quiz

Module 5 - Final Exam

- Instructions
- Final Exam

Module 6 - Assignment Preparation: Working with real-world data sets and built-in SQL functions

- Working with Real World Datasets
- Getting Table and Column Details
- LOADing Data
- Hands-on Lab: Practice Querying Real World Datasets

Module 7 - Course Assignment

- Instructions for Peer-Graded Assignment
- Jupyter Notebook with Problems for Peer-Reviewed Assignment
- Peer Review: Submit Your Work and Review Your Peers

总结: