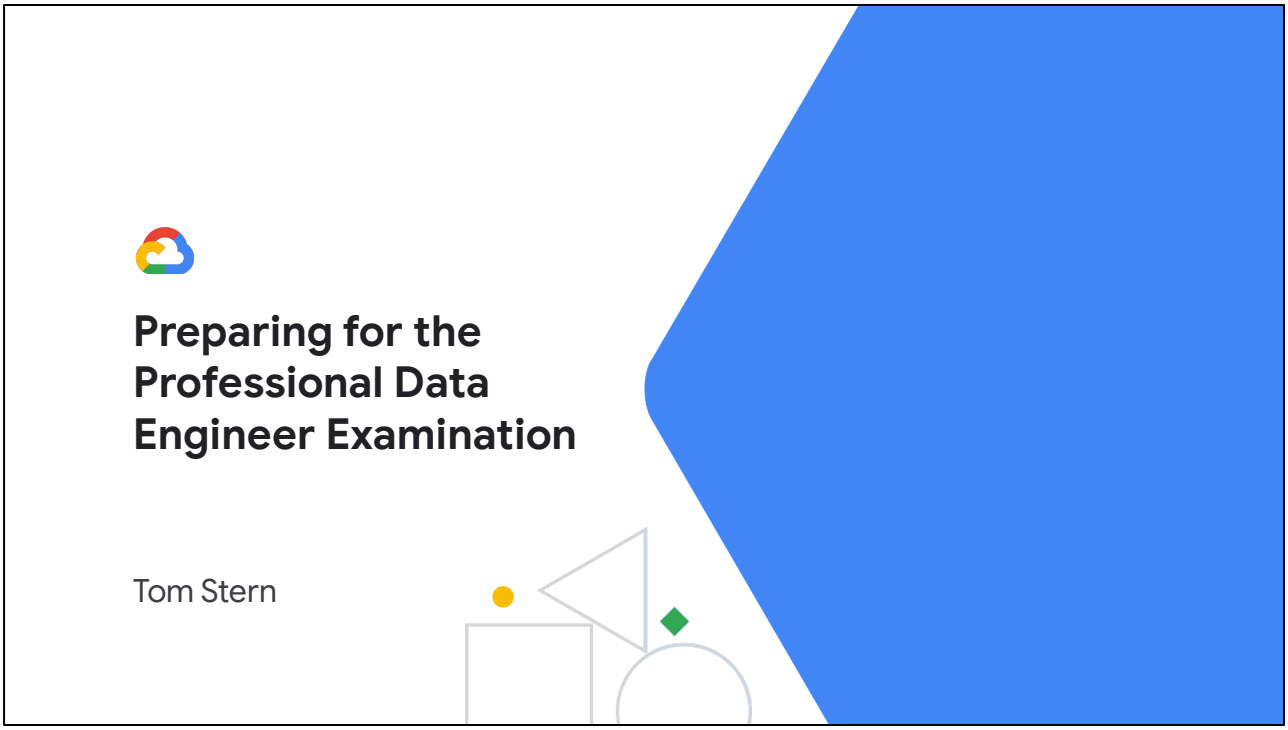
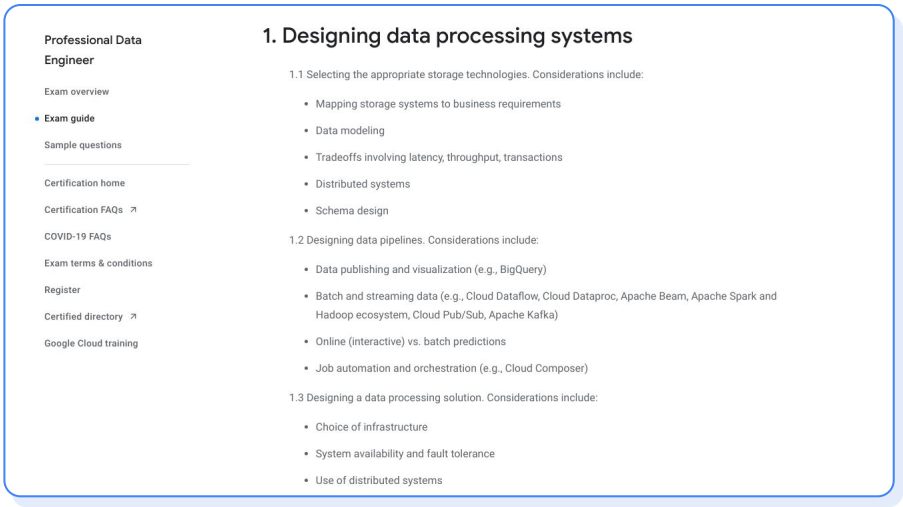


Hi, welcome to this course, Preparing for the Professional Data Engineer Exam.

I'm a technical curriculum developer with Google. My name is Tom Stern.

Data Engineer Exam Guide Outline

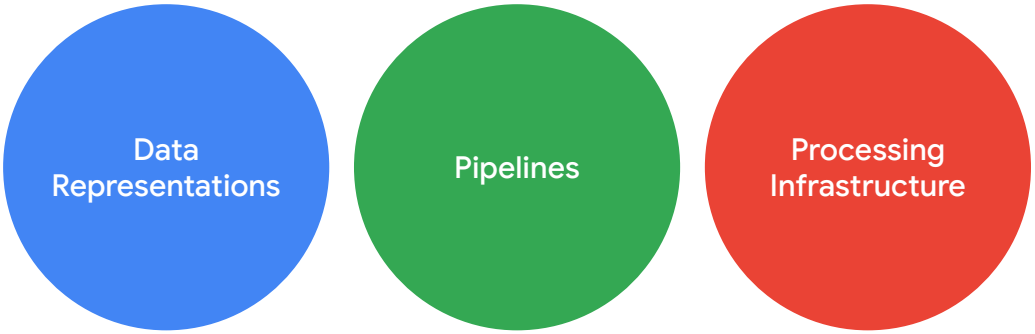


The tips sections of this course are organized around the data engineer exam guide outline. The outline divides the exam into sections that focus on specific priorities and information about the job role. So this course follows the outline, explaining each section of the course, providing tips, and highlighting information that would be useful to know.

Some items in the outline are important to the job of a professional data engineer, but they're not technical, or maybe they're not specific to Google. For example, a data engineer needs to know how to present solution proposals to customers and to communicate with executive staff. That isn't something we currently teach, but you should know it from experience or learn it on your own. I'll identify these items when they appear in the outline.

This course isn't about following some pattern exactly -- it is about being useful to help you decide the best ways for you to prepare for the exam.

Organizing information



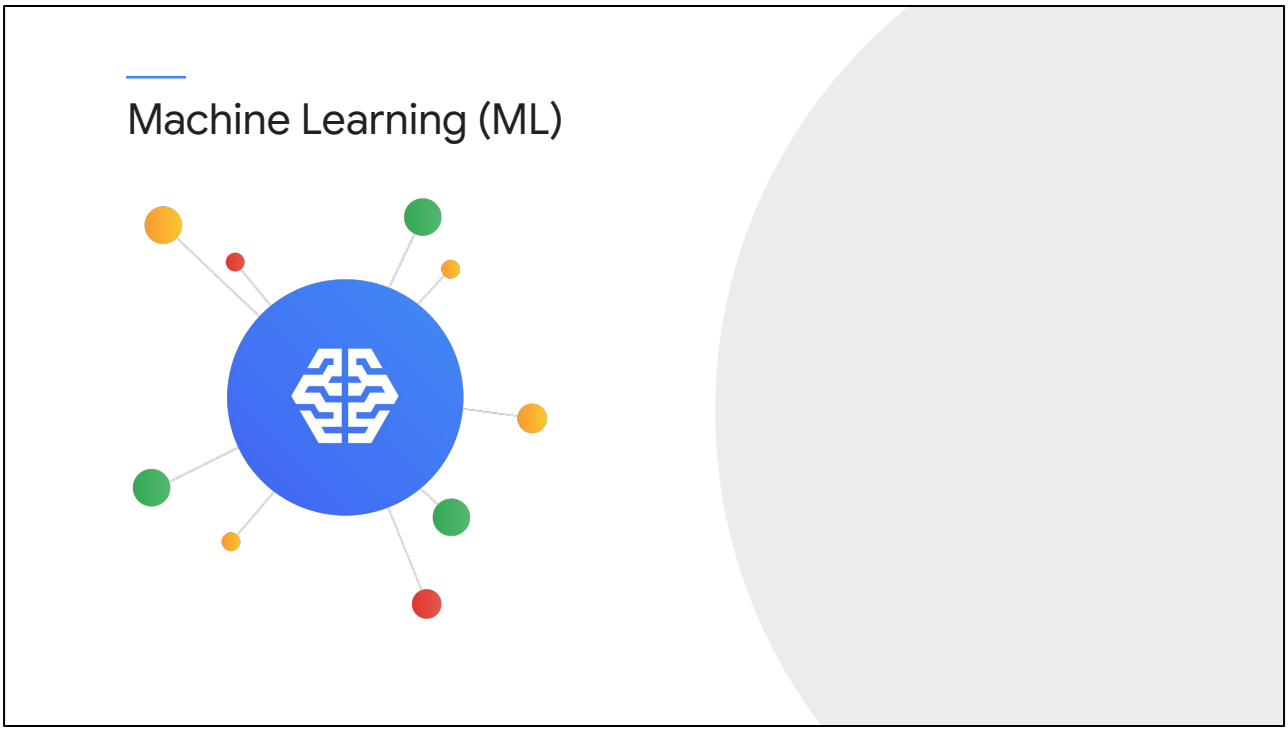
I think these three categories, of Data Representations, Pipelines, and Processing Infrastructure are a good way of organizing the information. And it helps you prepare for the exam by surfacing what is important. It is not just a list of details or trivial facts that are being tested.

It's the ability to perform the job, which means thinking through data engineering problems in the abstract. And using these categories of abstraction, it's a good way to organize your thinking about preparing for the exam.

Content organization

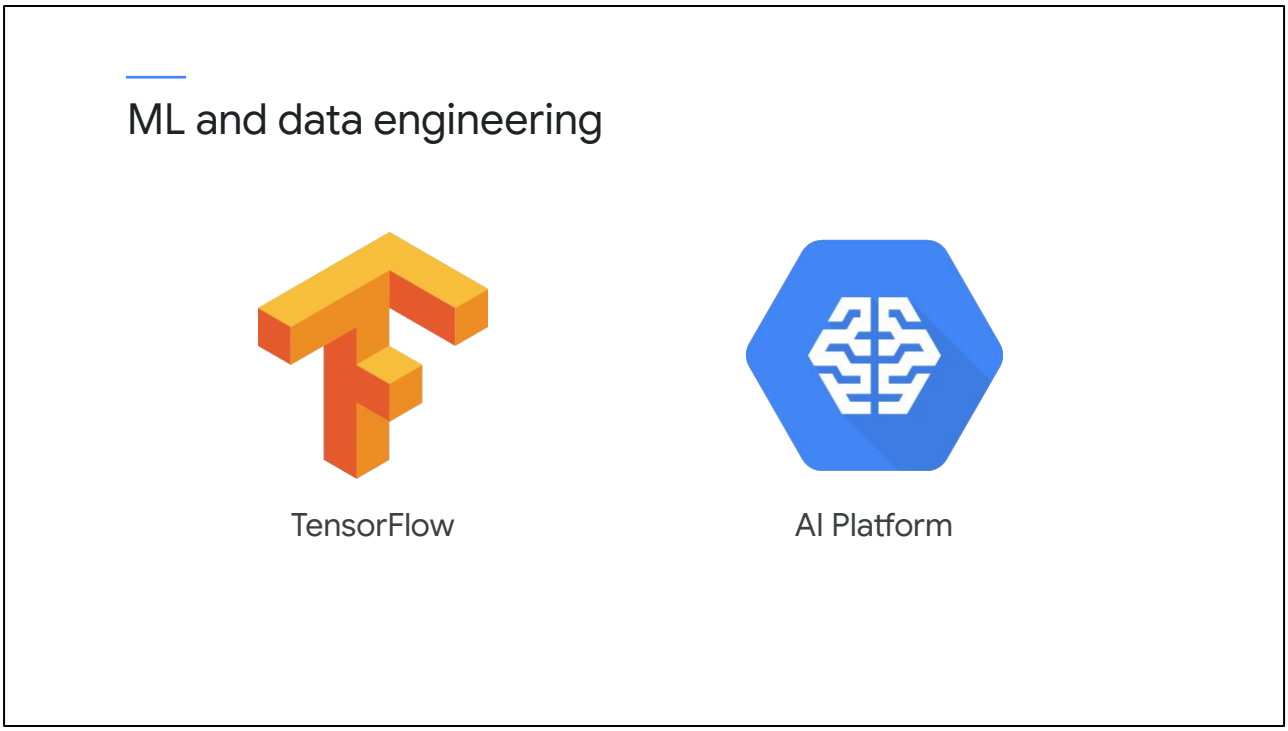
So wrapping up this point, the general organization of the class is that it follows the exam guide, but more specifically, we cover topics where they make the most sense for learning.

Finally, I want to highlight to you that we're all different. You have unique experiences, knowledge, and skills. So there isn't one right way to prepare for the exam. What's most important in this process is that you're exposed to many different ways to prepare and many different kinds of resources. And through this exposure, you can define your own unique and individualized approach to preparation.



A word about ML, which stands for machine learning.

In recent years, the industry evolved from database technologies to big data and data processing technologies. And now the industry is continuing to evolve from big data to machine learning. Big data didn't replace database technology. Hadoop isn't a replacement for MySQL. And machine learning and TensorFlow does not replace or eliminate big data or Hadoop in any way.



But what it does is it brings an entirely new perspective to data engineering. For the first time, we can take data that was complicated, and maybe even collected without any particular purpose in mind, and we can extract business intelligence from it.

Machine learning enables product innovation, making products better, and process innovation, making processes more efficient, and it brings a new kind of analysis to business decision-making. So ML isn't a subject that's tacked on or just included along with data engineering. It's a major part of data engineering, and you'll see it mentioned extensively in this course.

Agenda

- Understanding the Professional Data Engineer Certification
- Designing Data Processing Systems
- Building and Operationalizing Data Processing Systems
- Operationalizing Machine Learning Models
- Security, Policy, and Reliability
- Reviewing Resources and Next Steps
- Practicing for the Exam

Here's an agenda for this course.

Agenda

Understanding the Professional Data Engineer Certification

Designing Data Processing Systems

Building and Operationalizing Data Processing Systems

Operationalizing Machine Learning Models

Security, Policy, and Reliability

Reviewing Resources and Next Steps

Practicing for the Exam

The course begins by discussing what the certification is about, how it's positioned relative to other certifications, and more specifically, how it's designed relative to your job role, experience, and career aspirations.

Agenda

Understanding the Professional Data Engineer Certification

Designing Data Processing Systems

Building and Operationalizing Data Processing Systems

Operationalizing Machine Learning Models

Security, Policy, and Reliability

Reviewing Resources and Next Steps

Practicing for the Exam

The next three modules in the course cover specific preparation tips and technologies. Some of the information in these parts are what you might expect. Information on how to choose among related technologies, for example, under what conditions would you choose Bigtable over BigQuery? But there's other information.

One of the things we do in this course is highlight some sophisticated element of data engineering on Google Cloud. This is a way for us to convey many topics very fast. For example, one slide describes that subnets extend across zones within a region. This characteristic is unique to Google Cloud and different from most other cloud vendors. The reason for this design is it makes designed for reliability easier since adjacent instances can be in the same subnet but on different zones, giving them fault isolation. Now, that's a pretty sophisticated concept, but it's important for a data engineer to know when designing processing infrastructure on the Google Cloud.

So in this course, we don't teach you about regions or zones or subnets. That information is covered in our other courses. But what we do is highlight the skill you need to know for the job, and then you'll either know the dependent concepts on which it's based or you can go study them to fill in the gaps.

Agenda

- Understanding the Professional Data Engineer Certification
- Designing Data Processing Systems
- Building and Operationalizing Data Processing Systems
- Operationalizing Machine Learning Models
- Security, Policy, and Reliability
- Reviewing Resources and Next Steps
- Practicing for the Exam

It seems pretty obvious, but you should know that you can't expect to pass the exam from taking a single course. You can't learn everything you need to know to be a competent professional data engineer in a single course or in a single day. But what you will get from this class is a high-level overview of subject areas and tips, and practice with exam-taking skills, and that'll help you prepare.

If you've already been studying and preparing for the exam, this course will help you develop a good sense of what else you need to study or whether you're ready to attempt the exam.