An ecci Venture

apex
GLOBAL
Learning for Excellence

Big Data Foundation
2 Days | Classroom Training

Content

Big Data Foundation Course

Introduction

Did you know that today, more than 2.5 Exabytes (2.5 billion gigabytes) of data is generated every single day? And 90% of all the data in the world was produced within the last 2 years?

According to a research conducted by McKinsey, Big Data will become a key basis of competition, underpinning new waves of productivity growth, innovation, and consumer surplus—as long as the right policies and enablers are in place. No wonder organizations nowadays are facing a shortage of Big Data professionals in the market.

Considering the increasing demand of Big Data and Analytics professionals, we at APEX Global, bring you an incredibly strong foundation course that will help your employees understand the basic concepts of Big Data. If your organization is ready to boost its analytical capabilities or embrace big data strategies, this course can act as your Launchpad.

With this course, we can help you build a strong foundation to let you explore large and complex datasets and communicate with data scientists by asking the right questions. We are market ready and internationally renowned for our course content, delivery and hand-on exercises.

APEX Global is the learning solutions arm of ECCI — the leading process improvement solutions provider in Southeast Asia.

Our sole aim is to promote performance excellence among professionals. We give our customers the advantage to achieve greater success through effective, experiential and results-oriented training delivery.

We have the experience of training over 300,000 professionals in the last decade, a strong pool of expert trainers and facilitators with expertise in a niche array of domains and a strong regional presence. We provide an extensive portfolio of high-quality, industry-specific and functional programs coupled with high quality comprehensive training materials to deliver our ultimate "promise"—the R.E.A.L learning experience.

Our clients continue to grow

With the ever-expanding landscape of Big Data, our clients trust us to deliver industry relevant courseware for their employees. We help our clients to upskill their internal talent with Big Data technologies, hands-on data analyzing experience and specialty in handling large chunks of data.

Course Overview

Who is this course for?

Our Big Data Foundation course is highly recommended for the following professionals (but not limited to) who have a keen interest in Big Data technologies and are looking to bring value to their organizations:

- **Software engineers**
- **System administrators**
- **IT architects**
- **Application developers**
- **Data analysts and scientists**

What can you expect?

We have designed this two-day program for those who are new to Big Data. It offers fundamental coverage of potential data sources that can be used to solve real-life business problems. The course also covers the practical aspects of approaching data science problems and getting started with Hadoop, the most popular Big Data framework.

Throughout the course, we will ask you to shape your data analytics skills by writing relevant Big Data questions, analyzing real world data issues and performing hand-on MapReduce tasks in virtual machines. By the end, you will be able to undertake a custom-made assignment, hands-on training and a final exam.

What will you learn?

Upon completion of our training program, you will be able to:

- **Explain the origin of Big Data which include three key sources: people, sensors and organizations.**
- **Exemplify real-world problems that establish the importance of Big Data.**
- **Discuss the impact and importance of 6 V's of Big Data: Volume, Velocity, Variety, Veracity, Valence and Value.**
- **Describe the 5-step process of approaching data science problems.**
- **Explain the programming models and architectural components that are used to analyze Big Data.**
- **Install and run a program with the use of Hadoop.**

"

We chose it because we deal with huge amounts of data. Besides, it sounds really cool.""
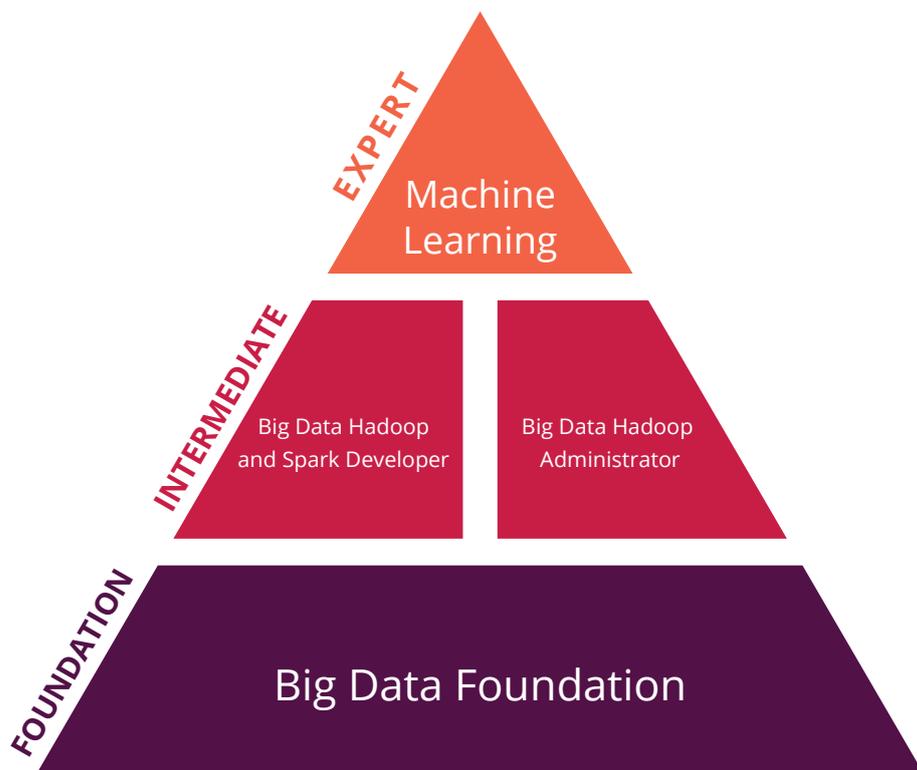
"

Larry Page, Co-founder of Google

Career Path

Get ready for an extraordinary data exploration journey with our bespoke Big Data career path! We understand that each learner has a distinct learning need. And, we cater to this need of IT Professionals in the Philippines and ASEAN region by making them dive deep into concepts of Big Data technologies, problems and systems.

Foundation

You can achieve the Foundation level by signing up for our Big Data Foundation course.

Big Data Foundation is a two-day course aimed to make you learn the basic concepts and terminologies working behind Big Data. With a detailed explanation of grounding concepts, it will introduce and describe the most commonly used framework, Hadoop.

Intermediate

After acquiring the Foundation level, the next step is to opt for either of the courses below based on your interest:

Big Data Hadoop and Spark Developer is a four-day course that will let help you master the skills of handling the Hadoop and Spark ecosystem. With hand-on practice sessions, you will become an expert in dealing with various ecosystem components like Hadoop, MapReduce, Impala, Sqoop, Apache Spark, Pig, Hive, Yarn, Flume and HBase.

Big Data Hadoop Administrator is a four-day course designed to provide detailed knowledge of the Hadoop framework, Hadoop cluste and HDFS. This course makes you proficient in concepts of core Hadoop distribution, cluster management solutions and Cloudera manager.

Expert

Join the elite league by enrolling in our Machine Learning course.

Machine Learning is a four-day course aimed to deliver expert knowledge of Machine Learning application and algorithms such as prediction, classification, clustering and regression. This course will also make you familiar with theories of Spark, Machine learning and deep learning.

Course Content

The Big Data Foundation course includes 5 modules followed by an assignment, a hands-on lab exercises and a final exam:

1.   Big Data: Why and Where?
2.   Characteristics of Big Data and Dimensions of Scalability
3.   Data Science: Getting Value out of Big Data
4.   Foundations for Big Data Systems and Programming
5.   Systems: Getting Started with Hadoop

Assignment

Hands-on lab exercises

Final exam

> **"**
> There were 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days."
>
> **"** Eric Schmidt, Google

MODULE 1

Big Data: Why and Where?

Big Data has been around us since ages. If you are wondering why data is called "big" and where did it come from, this module will clear all your doubts.

It intends to introduce you to the world of Big Data by explaining:

- The launch of Big Data
- Importance of Big Data
- Real-world problems of Big Data
- Three sources of Big Data: sensors, people and organization

With its relatable examples and easy explanations, you will be all set to continue with the rest of the learning after going through this module.

MODULE 2

Characteristics of Big Data and Dimensions of Scalability

If you have ever learned about Big Data earlier, you might have also heard of the "Big Vs". Generally, data analysts talk about just 5 Vs. But in this module, you will learn about the 6th V: Value. The Big Vs are:

- Volume
- Variety
- Velocity
- Veracity
- Valence
- Value

This module will explain the characteristics of all V's. It will also make you practice writing valid Big Data questions that you can put forward to Data Scientists.

Data Science: Getting Value out of Big Data

Big Data can benefit your organizations to limits you can ever imagine. This is all possible using a 5-step process that you can use to approach real-life data science problems.

In this module, we will talk about the following 5 steps:

1. Acquiring Data
2. Exploring and Pre-Processing Data
3. Analyzing Data
4. Communicating Results
5. Turning Insights into Action

Aside from these steps, this module will also explain how to build a Big Data strategy and the five P's of data science. These 5 Ps are:

1. Purpose
2. People
3. Process
4. Platforms
5. Programmability

MODULE 4

Foundations for Big Data Systems and Programming

Big Data exploration needs some programming systems and frameworks. For the purpose, this module has been designed to explain the relevant key programming concepts. It will describe:

• The Importance of programming frameworks
• The Meaning of Distributed File System
• Scalable Computing over the Internet
• Programming Models for Big Data

MODULE 5

Systems: Getting Started with Hadoop

Hadoop is a software framework that is used for processing large Big Data sets using the MapReduce programming model. This module will provide an in-depth understanding of Hadoop and MapReduce, while enabling you to perform a simple MapReduce task in the virtual machine. In addition, this module will cover:

• The Hadoop Distributed File System
• YARN
• Cloud Computing and Service Models

Course Assessment

Here are some quick facts about the exam:

The exam follows the proctored closed-book format.

The assessment is comprised of 40 multiple choice questions that should be answered within 1 hour. In order to pass the assessment, you will need a score of 70% or more.

Each multiple-choice item has one correct answer and three incorrect answers.

The mode of taking the exam is online or paper-based.

Once you pass the final assessment, you receive an APEX's certificate of completion.

Faculty Profile

Delivered by our established in-house experts, global network of trainers, speakers and facilitators through established processes, this certification course will introduce you to the most relevant and practical aspects of project management.

APEX Global Expert Council (AGEC) works closely with leading industry experts to consistently create, review and update the certification learning objectives. This ensures that what our clients learn is instantly applicable to any role and aligned to the needs of respective industries.

We collaborate with reputable partners on the design and development of course materials, so your learning is shaped by practical experience, expert insights and valuable case studies.

An ecci Venture

apex
GLOBAL
Learning for Excellence

For more information

Philippines	Malaysia	Vietnam	Singapore	India
T: (+632) 403 8668	T: (+603) 2788 4921	T: (+848) 3948 3142 to 43	T: (+865) 3158 6052	T: (+91) 44 4554 8438
F: (+632) 403 8358		F: (+848) 3811 4701		F: (+91) 44 4207 2357

 / apex-global   f / APEXGlobalLearning   in / company/apex-global-learning

For inquiries, email us at training@apexgloballearning.com

Disclaimer:
The trademarks/logos used in this document are owned by the respective brands.