# BIG DATA DATA SCIENCE FOUNDATION COURSE

IKOMPASS COURSE

## BIG DATA FOUNDATION

### FEATURES

- 3 Days classroom
- Certification (optional)
- Based on best practices
- Essential Knowledge

### EXAM

- A — **Paper based** — You will take the exam after class
- B — **40 Questions** — You will have 50 minutes to complete the exam
- C — **Certification** — Cloud Credential Council
- D — **Type of Questions** — Multiple choice and closed book

### TYPICAL DAY

- AM 9-10 — **Concept explanation** — Our trainers will explain the concept and show you the steps
- AM 10-12 — **Learn by Doing** — You will complete the steps that the trainer showed you
- PM 1-3 — **Challenge of the Day** — Each day you will work on a case study
- PM 3-5 — **Demo & Retrospective** — We will randomy chose one groups' work to review

## REGISTRATION

EMAIL **contact@ikompass.edu.sg**

CALL **+65 66381203**

ENROLL **www.ikompass.edu.sg**

# WHAT IS COVERED

## CHARACTERISTICS

History
Overview
Characteristics
Current best practices
challenges

## BIG DATA TECHNOLOGIES

foundational knowledge on hadoop, mongodb and other big data related technologies. we cover case studies in the social media and retail sector. you will also learn about the hdfc insfratstucture.

## DATA SOURCES

Enterprise Systems
Data Warehouses
Unstructured Data

## PANDAS

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

## SECOND ORDER

There are two main classes of machine learning algorithms: 1. supervised and 2. unsupervised learning. Exactly what does learning entail? At its most basic, learning involves specifying a model structure f that hopefully can extract regularities for the data or problem at hand.

## ANALYTICS

This chapter discusses how we can process a dataset and understand its basic characteristics. We will cover more complex methods like data mining, classification, and so on

## ALGORITHMS

frameworks are well suited for large-scale search and indexing applications. In fact, Google came up with the original MapReduce framework specifically to facilitate the various operations involved with web searching.

## CLOUD DEPLOYMENT

Computing clouds provide on-demand, horizontal, scalable computing resources with no upfront capital investment, making them an ideal environment to perform occasional large -scale Hadoop computations.

## OTHER STUFF.....

Cloud computing in terms of:
BIG DATA,
NoSql,
Social Media
MACHINE LEARNING

# CASE STUDY

Hypothetical Big Data project use case: Cybersecurity measures within a company in relation to insider threats. The company hosts thousands of applications for various business functions. The context will be User Behavior Analytics. Signals include, login meta data for each application, location data, network data, employee data, performance appraisal data, travel data, deaktop activity data. The analytics is focused on determining a risk score based for each user. One can think of Big Data as the raw data available in sufficient volume, variety and velocity. Volumes here refers to terabytes of data. Variety refers to the different dimensions of data. Velocity refers to the rate of change.

A bank can use credit card information to develop models that's more predictive about future credit behavior. This provides better financial access. What you purchased, frequency of purchase, how often do you pay back, where do you spend money are better predictors of payment credibility than a simple one dimensional credit score.

Big Data and analytics is made possible due to the digital breadcrumbs we leave. Digital breadcrumbs include things like location data, browsing habits, information from health apps, credit card transactions etc.,

The data lets us create mathematical models of how people interact, what motivates us, what influences our decision making process and how we learn from each other.

# EXPERT TRAINERS

## ROSHAN

Roshan is a Data Scientist and has over 14 years experience in building models and big data. He has worked with fortune 500 companies running large scale software projects in various roles including that of project manager, solution architect, sponsor and development lead. Roshan has been developing cloud applications for his enterprise clients in the area of finance and payment systems. His main strengths lies leveraging insights obtained from analyzing massive amounts of data.

## SANAT

Sanat has extensive expertise in Cloud computing and Virtualization technologies. He brings to the table hands-on knowledge in Virtual Desktop Infrastructure (VDI) – private cloud. His technical proficiencies cover Private/Public Cloud technologies- Amazon EC2, Windows Azure, Eucalyptus, OpenStack. Sanat is a Certified Trainer for Cloud and Virtualization Essentials courses from Cloud Credential Council.

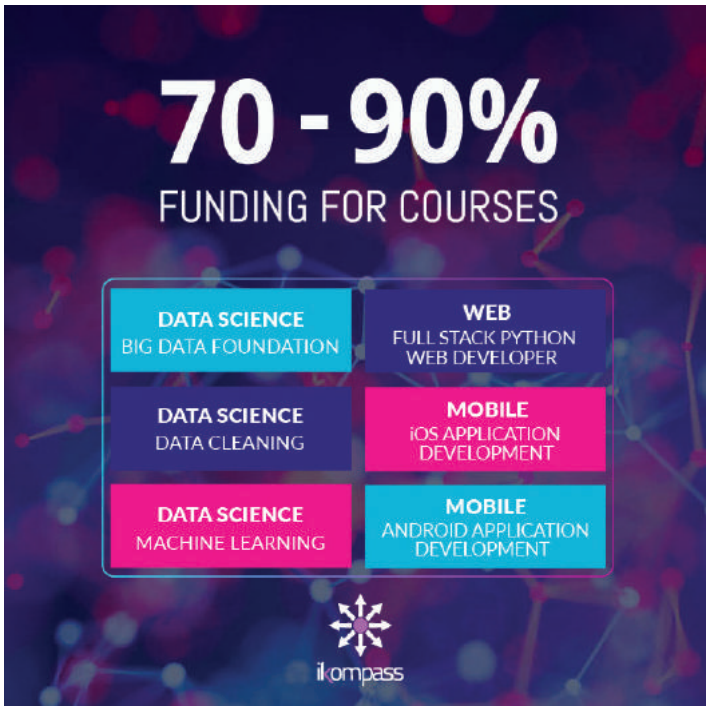
CERTIFIED TRAINING PARTNER
CLOUD CREDENTIAL COUNCIL

ITPACS

PMI Registered Education Provider
Project Management Institute

| DATA SCIENCE BIG DATA FOUNDATION | Singaporean | | | | | Singapore PR | Local Student |
|---|---|---|---|---|---|---|---|
| | SELF SPONSORED<br>Above 40 | SELF SPONSORED<br>Less than 40 | COMPANY SPONSORED (MNC)<br>Above 40 | COMPANY SPONSORED (MNC)<br>Less than 40 | COMPANY SPONSORED (SME)<br>Less than & Above 40 | COMPANY SPONSORED OR SELF SPONSORED<br>or less than 40 or above 40 | |
| COURSE FEES S$ | 2590 | 2590 | 2590 | 2590 | 2590 | 2590 | 2000 |
| EXAM FEES S$ | 227 | 227 | 227 | 227 | 227 | 227 | 227 |
| GST S$ | 197.19 | 197.19 | 197.19 | 197.19 | 197.19 | 197.19 | 155.89 |
| FUNDING | 90% | 70% | 90% | 70% | 90% | 70% | 100% |
| CITREP CLAIM S$ | 2535.3 | 1971.9 | 2535.3 | 1971.9 | 2535.3 | 1971.9 | 2227 |
| FEES AFTER FUNDING | 478.89 | 1042.29 | 478.89 | 1042.29 | 478.89 | 1042.29 | 155.89 |

Participant needs to make full payment to the training provider. Funding will be reimbursed by the relevant government agency after course completion.




## Registration

EMAIL **contact@ikompass.edu.sg**

CALL **+65 66381203**

ENROLL **www.ikompass.edu.sg**