

TEPC01 : Certified Data Science Specialist (CDSS) (อบรมเชิงปฏิบัติการพร้อมสอบประกาศนียบัตรในระดับสากล)**Description :****Program Certified by iTrain Asia Pte Ltd**

โดยในหลักสูตรจะให้ความรู้เกี่ยวกับแนวทางและเทคนิคที่คุณจำเป็นต้องทราบในการทำ Data Science ตั้งแต่ขั้นตอนการเก็บข้อมูล, การจัดการข้อมูล, การวิเคราะห์ข้อมูล จนถึงขั้นตอนการนำข้อมูลมาช่วยตัดสินใจ เพื่อให้คุณสามารถเริ่มต้นและมีทักษะด้าน Data Science

เมื่อผู้เรียนได้ผ่านการเรียน และทดสอบความรู้และความสามารถตามกำหนดเกณฑ์ของหลักสูตร จะได้รับ E-Certificate และ Digital Badge ในระดับสากล

Instructor :**Dr. Tarun Sukhani**

Principal Trainer

Training Date : **13-17 May 2024**fee : **34000 ฿** (ราคายังไม่รวม Vat 7%)Days & Duration : **5 Day(s) | 30 Hour(s)**Time : **09:00:00 - 16:00:00**Language : **English**Venue : **Online by Zoom**Type : **Online**Category : **Professional Certification Program****Objectives :****Certified Data Science Specialist (CDSS)****Introduction:**

Our lives are flooded by large amount of information, but not all of them is useful data. Therefore it is essential for us to learn how to applying data science to every aspect of our daily life from personal finances, reading, lifestyle to making business decisions. Leveraging on this data to make our life easier, or unlock new economic value for a business, is what you are going to learn in this course.

This course is a hands-on guided course for you to learn the concepts, tools, and techniques that you need to begin learning data science. We will cover the key topics from data science to big data, and the processes of gathering, cleaning and handling data. This course is well balanced between theory and practical, and key concepts are taught using case studies references. Upon completion, participants will be able to perform the basic data handling tasks, collect and analyze data, and present them using industry standard tools.

Upon completion of this course, you will be able to:

- Identify appropriate model for different data types.
- Create your own data process and analysis workflow.
- Define and explain the key concepts and models relevant to data science.
- Differentiate key data ETL process, from cleaning, processing to visualization.
- Implement algorithms to extract information from dataset.
- Apply best practices in data science, and familiar with standard tools.

Target Group :**Who Should Attend:**

The course is vendor-neutral. It is suitable for Department Head, Business Analyst, Market Analyst, Data Analyst, BI practitioner, IT Managers, Software Developer, System Analyst, Startups who are working with Data, and any Business / IT Executives who are interested in learning data science, or who want to begin their career as data scientist.

Target audience:

This workshop is intended for individuals who are interested in learning data science, or who want to begin their career as a data scientist.

Prerequisites:

All participants should have basic understanding of data, relations, and basic knowledge of mathematics.

EXAMINATION:

- No. of Questions: 50 Questions
- Duration - 2 hours
- Exam Type - Multiple Choice Questions (MCQ)
- Compulsory Passing Rate - 70%

Benefits :

- Participants will be equipped and ready to perform the basic data handling tasks, collect and analyze data, and present them using industry standard tools.

Course Outline :

DAY 1**Introduction to Data Science**

- What is Data?
- Types of Data
- What is Data Science?
- Statistical thinking
- Knowledge Check
- Lab Activity

Data Processes

- Extract, Transform and Load (ETL)
- Data Cleansing
- Aggregation, Filtering, Sorting, Joining
- Data Workflow
- Knowledge Check
- Lab Activity

Data Quality

- Raw vs Tidy Data
- Key features of data quality
- Maintenance of data quality
- Data profiling
- Data completeness and consistency

Life of a data scientist

- Identify problem
- Define question
- Define ideal dataset
- Obtain data
- Analyze data
- Interpret results
- Distribute results
- Knowledge Check

DAY 2**Beginning Databases**

- Types of Databases
- Relational Databases
- NoSQL
- Hybrid database
- Knowledge check
- Lab activity

Structured Query Language (SQL)

- Performing CRUD (Create, Retrieve, Update, Delete)
- Designing a Real world database
- Normalizing a table
- Knowledge check
- Lab Activity

Introduction to Python

- Basics of Python language
- Functions and packages
- Python lists
- Functional programming in Python
- Numpy and Scipy
- iPython
- Knowledge check
- Lab Activity

Lab: Exploring data using Python

DAY 3**Data Gathering**

- Obtain data from online repositories
- Import data from local file formats (json, xml)
- Import data using Web API
- Scrape website for data
- Knowledge check
- Lab Activity
- Instructor-led case study

Exploratory Data Analysis

- What is EDA?
- Goals of EDA
- The role of graphics
- Handling outliers
- Dimension reduction Introduction to R
- Features of R
- Vectors
- Matrices and Arrays
- Data Frame
- Input / Output

Lab: Exploring data using R**DAY 4****Introduction Text Mining**

- What is Text Mining?
- Natural Language Processing
- Pre-processing text data
- Extracting features from documents
- Using BeautifulSoup
- Measuring document similarity
- Knowledge check
- Lab activity

Supervised Learning

- What is prediction?
- Sampling, training set, testing set.
- Constructing a decision tree.
- Knowledge check
- Lab Activity

DAY 5**Presenting Data**

- Choosing the right visualization
- Plotting data using Python libraries
- Plotting data using R
- Using Jupyter Notebook to validate scripts
- Knowledge check
- Lab activity

Data Analysis Presentation

- Using Markdown language
- Convert your data into slides
- Data presentation techniques
- The pitfall of data analysis
- Knowledge check
- Lab activity

Group presentation Lab: Mini Project**Big Data Landscape**

- What is small data?
- What is big data?
- Big data analytics vs Data Science
- Key elements in Big Data (3Vs)
- Extracting values from big data
- Challenges in Big data

Big data Tools and Applications

- Introducing Hadoop Ecosystem
- Cloudera vs Hortonworks
- Real world big data applications
- Knowledge check
- Group discussion

What's next?

- Preview of Data Science Specialist
- Showing advanced data analysis techniques
- Demo: Interactive visualizations

EXAMINATION:

- No. of Questions: 50 Questions
- Duration - 2 hours
- Exam Type - Multiple Choice Questions (MCQ)
- Compulsory Passing Rate - 70%

Payment Condition :

Payment can be made by:

1. Cash or Credit Card or Bank Cheque payable to "สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ" (a post-dated cheque is not accepted) on the first day of the service or within the last day of the service.
2. Account transfer and send the proof of the payment (the deposit slip) via email ttd@swpark.or.th
 - ธนาคารกรุงเทพ สาขาอุทยานวิทยาศาสตร์
Saving Account Number: 080-0-00001-0
Account Name: สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ
 - ธนาคารกรุงไทย สาขาตลาดไท
Saving Account Number: 152-1-32668-1
Account Name: สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

Notes:

- Withholding tax (3%) is exempt.
- Should you need to withdraw, you must send the notice of the withdrawal in writing no later than 7 working days before the commencement date. The cancellation less than 7 days will be subject to a fine of 40% of the fee.
- Software Park Thailand reserves the rights to cancel courses due to unforeseen circumstances.

Contact Person :

For more information, contact our course coordinator on:

เสกสรรค์ สังสุข (อิฐ)

Mr. Seksun Sungsook

☎ : +662 583 9992 Ext. 81421

☎ : +6681 913 1828

✉ : seksun.sun@nstda.or.th

SOFTWARE PARK
THAILAND

You are encouraged to use the course schedule as a guide to plan your training. The schedule is accessible at www.swpark.or.th for more information.

