**Synopses of Modules**

Each PDC is attained after the completion of two modules. The structure of the PDC and the individual module synopsis is given below:

**1) PDC1 in Fundamentals of Data Science**

**Module 1: Introduction to Statistics for Data Science**

This module provides participants with an introduction to elementary probability theory and statistical concepts and principles that lay the foundation to understand and learn the statistical procedures and methods in the subsequent modules. The topics covered include descriptive statistics, rules of probability, probability distributions of discrete and continuous random variables, sampling distributions, statistical estimation and hypothesis testing.

**Module 2: Introduction to Programming for Data Science**

This module provides participants with the fundamental skills to code applications to retrieve, manipulate, process and visualize data using the Python programming language. Participants learn key concepts such as what structured and unstructured data are, and how they can create and manipulate relational and NoSQL databases to explore data and to create visualizations that can help them gain useful insights from it.

**2) PDC2 in Visual and Textual Data Analytics**

**Module 3: Practical Computer Vision**

Participants will learn various image processing techniques through openCV and python programming practice. Participants will learn how to apply machine learning algorithms and deep artificial neural network to carry out vision analytics tasks such as image classification, segmentation, object detection, pose estimation, action recognition, OCR and generative CV. Participants will also learn mainstream computer vision models, pipelines, frameworks and industry trends. After this module, participants are able to develop and evaluate a computer vision solution to address a real-life problem through project work, for example, in the area of healthcare image analysis or product quality inspection.

**Module 4: Practical Natural Language Processing**

This module will immerse the participants in the rapidly evolving field of natural language processing (NLP), which focuses on enabling computers to learn and understand human languages. In this module, participants will learn to leverage text-based data to perform various analyses such as sentiment analysis, document classification, and text summarization. They will also be introduced to Generative AI and Large Language Models (LLMs), learning to fine-tune state-of-the-art pretrained models. Additionally, participants will develop a question-and-answer chatbot and learn to generate captions/text from images. Throughout the course, participants will have the opportunity to work on datasets based on real-world scenarios and apply what they have learned to solve practical business problems.