Certified Data Science Practitioner™ (CDSP) (Exam DSP-210)

Course Number: CNX0020

Course Length: 5 days

Overview:

For a business to thrive in our data-driven world, it must treat data as one of its most important assets. Data is crucial for understanding where the business is and where it's headed. Not only can data reveal insights, but it can also inform—by guiding decisions and influencing day-to-day operations. This calls for a robust workforce of professionals who can analyze, understand, manipulate, and present data within an effective and repeatable process framework. In other words, the business world needs data science practitioners. This course will enable you to bring value to the business by putting data science concepts into practice.

Course Objectives:

In this course, you will implement data science techniques in order to achieve organizational goals.

You will:

- Use data science principles to address business issues.
- Apply the extract, transform, and load (ETL) process to prepare datasets.
- Use multiple techniques to analyze data and extract valuable insights.
- Design a machine learning approach to address business issues.
- Train, tune, and evaluate classification models.
- Train, tune, and evaluate regression and forecasting models.
- Train, tune, and evaluate clustering models.
- Finalize a data science project by presenting models to an audience, putting models into production, and monitoring model performance.

Target Student:

This course is designed for business professionals who leverage data to address business issues. The typical student in this course will have several years of experience with computing technology, including some aptitude in computer programming.

However, there is not necessarily a single organizational role that this course targets. A prospective student might be a programmer looking to expand their knowledge of how to guide business decisions by collecting, wrangling, analyzing, and manipulating data through code; or a data analyst with a background in applied math and statistics who wants to take their skills to the next level; or any number of other data-driven situations.

Ultimately, the target student is someone who wants to learn how to more effectively extract insights from their work and leverage that insight in addressing business issues, thereby bringing greater value to the business.

This course is also designed to assist students in preparing for the CertNexus® Certified Data Science Practitioner™ (CDSP) (Exam DSP-210) certification.

Prerequisites:

To ensure your success in this course, you should have at least a high-level understanding of fundamental data science concepts, including, but not limited to: types of data, data science roles, the overall data science lifecycle, and the benefits and challenges of data science. You can obtain this level of knowledge by taking the CertNexus *DSBIZ™ (Exam DSZ-210)* course.

You should also have experience with high-level programming languages like Python. Being comfortable using fundamental Python data science libraries like NumPy and pandas is highly recommended. You can obtain this level of skills and knowledge by taking the Logical Operations course *Using Data Science Tools in Python®*.

In addition to programming, you should also have experience working with databases, including querying languages like SQL. Several Logical Operations courses can help you attain this experience:

- *Database Design: A Modern Approach*
- *SQL Querying: Fundamentals (Second Edition)*
- *SQL Querying: Advanced (Second Edition)*

Course Content

Lesson 1: Addressing Business Issues with Data Science

    Topic A: Initiate a Data Science Project

    Topic B: Democratize Data

    Topic C: Formulate a Data Science Problem

Lesson 2: Extracting, Transforming, and Loading Data

    Topic A: Extract Data

    Topic B: Transform Data

    Topic C: Load Data

Lesson 3: Analyzing Data

    Topic A: Examine Data

    Topic B: Explore the Underlying Distribution of Data

    Topic C: Use Visualizations to Analyze Data

    Topic D: Preprocess Data

Lesson 4: Designing a Machine Learning Approach

    Topic A: Identify Machine Learning Concepts

    Topic B: Identify Transformer-Based Deep Learning Concepts

    Topic C: Test a Hypothesis

Lesson 5: Developing Classification Models

    Topic A: Train and Tune Classification Models

    Topic B: Evaluate Classification Models