Lesson summary

Module 1 Lesson 1: From Problem to Approach

Congratulations! You have completed this lesson. At this point in the course, you know:

- Foundational methodology, a cyclical, iterative data science methodology developed by John Rollins, consists of 10 stages, starting with Business Understanding and ending with Feedback.
- CRISP-DM, an open source data methodology, combines several of the data-related methodology stages into one stage and omits the Feedback stage resulting in a six-stage data methodology.
- The primary goal of the Business Understanding stage is to understand the business problem and determine the data needed to answer the core business question.

Business Understanding

Understanding the Question

Data science methodology begins with seeking clarification and attaining a business understanding

1. Clearly defining the question is vital as it guides the analytic approach
2. Understanding the goal of the person asking the question is crucial in establishing a clearly defined question
3. Objectives that support the goal should be identified to prioritize and plan for problem-solving
4. Engagement of different stakeholders is vital to determine requirements and clarify questions

IBM DATA SCIENCE METHODOLOGY Skills Network

- During the Analytic Approach stage, you can choose from descriptive, diagnostic, predictive, and prescriptive analytic approaches, whether to use machine learning with clustering associations.

Analytic Approach

Determine Appropriate Approach

The second stage of the data science methodology involves selecting the analytic approach in the context of business requirements.

1. Identifying what type of patterns will be needed to address the question most effectively.
2. Different types of patterns and relationships in data require different approaches to effectively address the questions being asked
3. Descriptive:
   Current Status
   Diagnostic:
   Statistical Analysis
   Predictive:
   Forecasting
   Prescriptive:
   Recommendations
4. Machine Learning enables computers to learn without explicit programming and identify patterns in data that might otherwise not be accessible

IBM DATA SCIENCE METHODOLOGY Skills Network

- Decision tree classification is a predictive analytics approach that's easy for non-data scientists to implement

Author(s)

Dr. Pooja
Patsy R. Kravitz

Changelog

Date	Version	Changed by	Change Description
2023-08-09	0.1	Patsy R. Kravitz	Initial version created
2023-08-13	0.2	Dr. Pooja	Inforgraphic included