

# Terms and definitions from Course 1,

## Module 4

**PACE workflow:** A framework that provides an initial structure to guide the process of data analytics; PACE stands for plan, analyze, construct, and execute

**Plan stage:** Stage of the PACE workflow where the scope of a project is defined and the informational needs of the organization are identified

**Analyze stage:** Stage of the PACE workflow where the necessary data is acquired from primary and secondary sources and then cleaned, reorganized, and analyzed

**Construct stage:** Stage of the PACE workflow where data models and machine learning algorithms are built, interpreted, and revised to uncover relationships within the data and help unlock insights from those relationships

**Execute stage:** Stage of the PACE workflow where a data professional will present findings with internal and external stakeholders, answer questions, consider different viewpoints, and make recommendations

## Terms and definitions from previous modules

### A

**Active listening:** Refers to allowing team members, bosses, and other collaborative stakeholders to share their own points of view before offering responses

**Aggregate information:** Data from a significant number of users that has eliminated personal information

**Analytics Team Manager:** A data professional who supervises analytical strategy for an organization, often managing multiple groups

**Artificial intelligence (AI):** Refers to computer systems able to perform tasks that normally require human intelligence

## B

**Business Intelligence Analyst:** (Refer to **Business Intelligence Engineer**)

**Business Intelligence Engineer:** A data professional who uses their knowledge of business trends and databases to organize information and make it accessible; also referred to as a Business Intelligence Analyst

## C

**Chief Data Officer:** An executive-level data professional who is responsible for the consistency, accuracy, relevancy, interpretability, and reliability of the data a team provides

## D

**Data anonymization:** The process of protecting people's private or sensitive data by eliminating PII

**Data cleaning:** The process of formatting data and removing unwanted material

**Data Engineer:** A data professional who makes data accessible, ensures data ecosystems offer reliable results, and manages infrastructure for data across enterprises

**Data professional:** Any individual who works with data and/or has data skills

**Data science:** The discipline of making data useful

**Data Scientist:** A data professional who works closely with analytics to provide meaningful insights that help improve current business operations

**Data stewardship:** The practices of an organization that ensure that data is accessible, usable, and safe

## E

**Edge computing:** A way of distributing computational tasks over a bunch of nearby processors (i.e., computers) that is good for speed and resiliency and does not depend on a single source of computational power

## H

**Hackathon:** An event where programmers and data professionals come together and work on a project

## I

**Interpersonal skills:** Traits that focus on communicating and building relationships

## J

**Jupyter Notebook:** An open-source web application used to create and share documents that contain live code, equations, visualizations, and narrative text

## M

**Machine learning:** The use and development of algorithms and statistical models to teach computer systems to analyze patterns in data

**Mentor:** Someone who shares knowledge, skills, and experience to help another grow both professionally and personally

**Metrics:** Methods and criteria used to evaluate data

## N

**Nonprofit:** A group organized for purposes other than generating profit; often aims to further a social cause or provide a benefit to the public

## O

**Open data:** Data that is available to the public and free to use, with guidance on how to navigate the datasets and acknowledge the source

## P

**Personally identifiable information (PII):** Information that permits the identity of an individual to be inferred by either direct or indirect means

**Python:** A general-purpose programming language

## R

**RACI chart:** A visual that helps to define roles and responsibilities for individuals or teams to ensure work gets done efficiently; lists who is responsible, accountable, consulted, and informed for project tasks

## S

**Sample:** A segment of a population that is representative of the entire population

# T

**Tableau:** A business intelligence and analytics platform that helps people visualize, understand, and make decisions with data