



## Reading: Glossary terms from module 2

---

### Glossary terms from Course 1, Module 2

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

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

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

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

**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

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

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

**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

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

**Sample:** A segment of a population, often used to infer parameters of the whole population

### Terms and definitions from the previous module

#### D

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

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

**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

## 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

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

## P

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

## T

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

---