

# Glossary

## Data Analytics

### Terms and Definitions

---



## A

**Action-oriented question:** A type of question that generates answers that can lead to action or change (Example: What design features will make our packaging easier to recycle?)

**Analytical skills:** Qualities and characteristics associated with solving problems using facts

**Analytical thinking:** Identifying and defining a problem and then solving it by using data in an organized, step-by-step manner

**Attribute:** A characteristic or quality of data used to label a column in a table

## B

**Business task:** The question or problem that data analysis answers for a business

## C

**Cloud:** A place to keep data online, rather than a computer hard drive

**Context:** The condition in which something exists or happens

# D

**Data:** A collection of facts

**Data analysis:** The collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making

**Data analysis process:** Carrying out the six phases of ask, prepare, process, analyze, share, and act in order to gain data-driven insights that lead to informed decisions

**Data analyst:** Someone who collects, transforms, and organizes data in order to help make informed decisions

**Data analytics:** The science of data

**Database:** A collection of data stored in a computer system

**Data design:** How you organize information

**Data-driven decision making:** Using facts to guide business strategy

**Data ecosystems:** The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data

**Data life cycle:** The sequence of stages that data experiences, which include plan, capture, manage, analyze, archive, and destroy

**Data science:** Creating new ways of modeling and understanding the unknown by using raw data

**Data strategy:** The management of the people, processes, and tools used in data analysis

## F

**Fairness:** Ensuring that your analysis doesn't create or reinforce bias

**Formula:** A set of instructions that performs a specific calculation using the data in a spreadsheet

**Function:** A preset command that automatically performs a specific process or task using the data in a spreadsheet

## G

**Gap analysis:** A method for examining and evaluating how a process works currently in order to get where you want to be in the future

## L

**Leading question:** A question that encourages people to respond in a certain way (Example: These are the best sandwiches ever, aren't they?)

## M

**Measurable question:** A question that generates answers that can be quantified and assessed (Example: How many times was our video shared on social media the first week it was posted?)

# O

**Observation:** All of the attributes for something contained in a row of a data table

# P

**Problem types:** The different types of tasks that data analysts encounter including: making predictions, categorizing things, spotting something unusual, identifying themes, discovering connections, and finding patterns

# Q

**Query:** The way we use SQL to communicate with the database

**Query language:** A computer programming language that allows you to retrieve and manipulate data from a database

# R

**Relevant question:** A question that has significance to a problem (Example: What environmental factors changed in Durham, North Carolina that could have caused Pine Barrens tree frogs to disappear from the Sandhills Region?)

**Root cause:** The reason why a problem occurs

# S

**SMART methodology:** A method for determining a question's effectiveness. An effective question is Specific, Measurable, Action-oriented, Relevant, and Time-bound (SMART)

**Specific question:** A question that is simple, significant, and focused on a single topic or a few closely related ideas (Example: How many kids achieve the recommended 60 minutes of physical activity at least five days a week?)

**Spreadsheet:** An electronic document in which data is arranged in the rows and columns of a grid and can be manipulated and used in calculations

**Stakeholders:** People that have invested time, interest, and resources into the projects you'll be working on as a data analyst

**Structured thinking:** The process of recognizing the current problem or situation, organizing available information, revealing gaps and opportunities, and identifying the options

## T

**Technical mindset:** The ability to break things down into smaller steps or pieces and work with them in an orderly and logical way

**Third-party logistics:** The partnership between a shipping company and a business to ship products when the business doesn't have its own trucks, planes, or ships

**Time-bound question:** A question that specifies the timeframe to be addressed in the answers (Example: What environmental factors changed in Durham, North Carolina, *between 1983 and 2004* that could cause Pine Barrens tree frogs to disappear from the Sandhills Region?)

## U

**Unfair question:** A question that makes assumptions or is difficult to answer honestly  
(Example: What do you love most about our exhibits?)

V

**Visualization:** The graphical representation of information