Glossary

Data Analytics
Terms and Definitions

A

Action - oriented question: A question whose answers lead to change

Algorithm: A process or set of rules followed for a specific task

Analytical skills: Qualities and characteristics associated with using facts to solve problems

Analytical thinking: The process of identifying and defining a problem, 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

AVERAGE: A spreadsheet function that returns an average of the values from a selected range

B

Big data: Large, complex datasets typically involving long periods of time, which enable data analysts to address far-reaching business problems

Borders : Lines that can be added around two or more cells on a spreadsheet

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

C

Cell reference: A cell or a range of cells in a worksheet typically used in formulas and functions

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

Context: The condition in which something exists or happens

COUNT: A spreadsheet function that counts the number of cells in a range that meet a specific criteria

D

Dashboard: A tool that monitors live, incoming data

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: The six phases of ask, prepare, process, analyze, share, and act whose purpose is to gain insights that drive informed decision-making

Data analyst: Someone who collects, transforms, and organizes data in order to draw conclusions, make predictions, and drive informed decision-making

Data analytics: The science of data

Data design: How information is organized

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

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

Data-inspired decision-making: Exploring different data sources to find out what they have in common

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

Data science: A field of study that uses raw data to create new ways of modeling and understanding the unknown

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

Data visualization: The graphical representation of data

Database: A collection of data stored in a computer system

Dataset: A collection of data that can be manipulated or analyzed as one unit

E

Equation: A calculation that involves addition, subtraction, multiplication, or division (also called a math expression)

F

Fairness: A quality of data analysis that does not c reate or reinforce bias

Fill handle: A box in the lower - right -hand corner of a selected spreadsheet cell that can be dragged through neighboring cells in order to continue an instruction

Filtering: The process of showing only the data that meets a specifi ed criteria while hiding the rest

Formula: A set of instructions used to perform a calculation using the data in a spreadsheet

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

G

Gap analysis: A method for examining and evaluating the current state of a process in order to identify opportunities for improvement in the future

H

Header: The first row in a spreadsheet that labels the type of data in each column

I

J

K

L

Leading questio n: A question that steers people toward a certain response

M

Math expression: A calculation that involves addition, subtraction, multiplication, or division (also called an equation)

Math function   : A function that is used as part of a mathematical formula

MAX: A spreadsheet function that returns   the largest numeric value from a range of cells

Measurable question: A question whose answers can be quantified and assessed

Metric: A single, quantifiable type of data that is used for measurement

Metric goal: A measurable goal set by a company and evaluated using metrics

MIN: A spreadsheet function that returns   the smallest numeric value from a range of cells

N

O

Observation: The attributes that describe a piece of data contained in a row of a table

Open data: Data that is available to the public

Operator: A symbol that names the operation or calculation to be performed

Order of operations: Using parentheses to group together s preadsheet values in order to clarify the order in which operations should be performed

P

Pivot chart: A chart created from the fields in a pivot table

Pivot table: A data summarization tool used to sort, reorganize, group, count, total, or

average data

Problem domain: The area of analysis that encompasses every activity affecting or affected by a problem

Problem types: The various problems that data analysts encounter, including categorizing things, discovering connections, finding patterns, identifying t   hemes, making predictions, and spotting something unusual

Q

Qualitative data: A subjective and explanatory measure of a quality or characteristic

Quantitative data: A specific and objective measure, such as a number, quantity, or range

Query: A request for data or information from a database

Query language: A computer programming language used to communicate with a database

R

Range: A collection of two or more cells in a spreadsheet

Reframing: Restating a problem or challenge, then redirecting it toward a p otential resolution

Relevant question: A question that has significance to the problem to be solved

Report: A static collection of data periodically given to stakeholders

Return on investment (ROI):   A formula that uses the metrics of investment and profi   t to evaluate the success of an investment

Revenue: The total amount of income generated by the sale of goods or services

Root cause: The reason why a problem occurs

S

Scope of work (SOW):   An agreed-upon outline of the tasks to be performed during a project

Small data: Small, specific data points typically involving a short period of time, which are useful for making day -to-day decisions

SMART methodology: A tool for determining a question's effectiveness based on whether it is specific, measurable, action-oriented, relevant, and time-bound

Sorting: The process of arranging data into a meaningful order to make it easier to understand, analyze, and visualize

Specific question: A question that is simple, significant, and focused on a single topic or a few closely related ideas

Spreadsheet: A digital worksheet

SQL: Refer to Structured Query Language

Stakeholders: People who invest time and resources into a project and are interested in its outcome

Structured Query Language: A computer programming language used to communicate with a database

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

SUM: A spreadsheet function that adds the values of a selected range of cells

T

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

Time-bound question: A question that specifies a timeframe to be studied

Turnover rate: The rate at which employees voluntarily leave a company

U

Unfair question: A question that makes assumptions or is difficult to answer honestly

V

Visualization: (Refer to data visualization)

W

X

Y

Z