Glossary

Data Analytics

Terms and definitions from Course 1

A

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

B

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

C

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 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 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

F

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

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

I

J

K

L

M

N

O

Oversampling: The process of increasing the sample size of nondominant groups in a population. This can help you better represent them and address imbalanced datasets

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

P

Q

Query: A request for data or information from a database

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

R

Root cause: The reason why a problem occurs

S

Self-reporting: A data collection technique where participants provide information about themselves

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

Spreadsheet: A digital worksheet

SQL: (Refer to Structured Query Language)

T

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

U

V

Visualization: (Refer to data visualization)

W

X

Y

Z