COURSE GLOSSARY

Introduction to Data Visualization with ggplot2

Aesthetic mapping (Aesthetic): The assignment of data variables to visual properties (such as x position, y position, color, size, shape, or alpha) so values are encoded on visual scales

Alpha (alpha-blending): The transparency aesthetic that controls the opacity of plot elements to help reveal overlapping data and convey density

Attribute: A fixed visual setting provided directly in a geom (for example color = "red" or size = 2) that sets appearance rather than mapping a data column to a visual property

Data (as a grammatical element): The dataset or data frame that supplies the variables and observations to be plotted, and which can be inherited or overridden by individual layers

Data visualization: The practice of representing data graphically to reveal patterns, trends, and relationships so that viewers can analyze and communicate findings efficiently and accurately

element_blank: A theme setting that removes a specific non-data element from a plot entirely so it is not drawn (commonly used to simplify or declutter visuals)

Explanatory visualization: A carefully designed, often labor-intensive graphic created to communicate specific insights or stories to a broader or non-specialist audience

JSON (JavaScript Object Notation): A lightweight, text-based, schema-less format for representing structured data as nested key-value pairs, commonly used by APIs and configuration files

Exception (Error): A runtime event indicating that something went wrong in code execution (often called an error), which, if unhandled, will terminate the program

Geom (Geometry layer): The layer that defines the geometric object used to represent data (such as points, bars, lines, or ribbons) and is implemented in ggplot2 via geom_ functions

geom_bar: A geom that, by default, counts observations in each category mapped to the x-axis (uses stat = "count"), whereas geom_col plots bar heights directly from a provided y value without recounting

geom_histogram: A geom that bins a single continuous x variable and draws bars to show the frequency or count per bin, implemented with an underlying stat (stat_bin)

geom_point: The ggplot2 geometry that draws individual points for x–y data and commonly accepts aesthetics like color, shape, size, and alpha

Grammar of Graphics: A conceptual framework that treats plots as compositions of layered grammatical elements (e.g., data, aesthetics, geoms, scales, themes) which together specify how data is mapped to visual forms

Jitter: A specific position adjustment that adds small random noise to point coordinates to reduce overplotting and reveal point density in discrete or rounded data

Log transformation: A data transformation that replaces values with their logarithm to reduce skew, stabilize variance, or linearize multiplicative relationships for clearer plotting and modeling

Outlier: An observation with an unusually large or small value that can strongly influence summary statistics or fitted models and may require inspection, transformation, or exclusion

Overplotting: The visual problem that occurs when many points overlap in a plot, obscuring data structure and density and requiring remedies such as jittering, alpha transparency, or alternative geoms

Position adjustments: Methods that change how overlapping or categorical elements are placed (for example identity, dodge, stack, fill, jitter, or nudge) to improve readability and avoid visual ambiguity

Scale (scale functions): Functions (scale_*) that control how data values are translated to visual values, including options like limits, breaks, labels, expand, and custom palettes to adjust axes and aesthetic ranges

Stat (statistic) layer: The component that performs statistical transformations (for example binning, counting, or smoothing) before plotting, often invoked implicitly by specific geoms like geom_histogram (stat_bin).

Theme (Themes layer): The layer that controls all non-data ink (text, lines, and rectangles) such as fonts, grid lines, axis appearance, and legend positioning to produce consistent, publication-quality styling