

Module 4: Fundamentals of Data Analysis

Glossary

Bimodal Distribution

A probability distribution with two different modes that appear as distinct peaks in the probability density function

Box Plot

A standardized way of displaying a dataset based on a five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles

Heat Map

A data visualization technique that shows the magnitude of a phenomenon as color in two dimensions

Inner Join

An operation that returns records that have matching values in both tables

Joint Plot

A plot comprising three plots, where one is placed horizontally at the top showing the distribution of the variable X, the second is placed on the right margin of the graph with the orientation set to vertical and shows the distribution of the variable Y, and the third is a scatterplot in the center that shows the relationship between X and Y.

Kernel Density Estimation (KDE)

A nonparametric way to estimate the probability density function of a random variable

Left Join

An operation that returns all records from the left table and the matched records from the right table

Normal Distribution

A type of continuous probability distribution for a real-valued random variable

Outer Join (or Full Join)

An operation that returns all records when there is a match in either the left or the right table

Right Join

An operation that returns all records from the right table and the matched records from the left table

Scatterplot

A type of plot that uses coordinates to display values for two variables in a dataset, typically

Skewed Distribution

A distribution with one of its tails longer than the other

Violin Plot

A type of plot that includes a marker for the median of the data, a box or marker indicating the interquartile range, and possibly all sample points if the number of the samples is not too high. Similar to a box plot, except that they also show the probability density of the data at different values, usually smoothed by a kernel density estimator.