Glossary

A

Activity (CPA) [p. 473] A task to be completed as part of a project. Activities are represented by the edges in the project diagram.

Activity network [p. 473] An activity network is a weighted directed graph that shows the required order of completion of the activities that make up a project. The weights indicate the durations of the activities they represent.

Adjacency matrix [p. 399] A square matrix showing the number of edges joining each pair of vertices in a graph.

Algorithm [p. 441] A step-by-step procedure for solving a particular problem that involves applying the same process repeatedly. Examples include Prim's algorithm and the Hungarian algorithm.

Allocation [p. 451] Allocation is the process of assigning a series of tasks to different members of a group in a way that enables the tasks to be completed for the minimum time or cost.

Annuity [p. 351] An annuity is a compound interest investment from which regular payments are made.

Arc [p. 219] The part of a circle between two given points on the circle. The length of the arc of a circle is given by $s = r\left(\frac{\tilde{}}{180}\right)°$, where r is the radius of the circle and $\tilde{}$ is the angle in degrees subtended by the arc at the centre of the circle.

Arithmetic sequences [p. 166] A sequence is arithmetic if it satisfies the recurrence relation: $t_{n+1} = t_n + d$ and a starting point usually t_1. Arithmetic sequences are used to model linear growth and linear decay situations. The rule for the nth term of an arithmetic sequence is: $t_n = t_1 + (n - 1)d = a + (n - 1)d$, where $a = t_1$ is the starting value.

B

Backward scanning [p. 485] Backward scanning is the process of determining the LST for each activity in a project activity network.

Balance [p. 324] The balance of a loan or investment is the amount owed or accrued after a period of time.

Bipartite graph [p. 446] A graph whose set of vertices can be split into two subsets, X and Y, in such a way that each edge of the graph joins a vertex in X and a vertex in Y.

Bivariate data [p. 2] Data in which each observation involves recording information about two variables for the same person or thing. An example would be data recording the height and weight of the children in a preschool.

Bridge [p. 394] An edge in a **connected graph** that, if removed, would leave the graph no longer connected.

C

Capacities (flow network) [p. 455] The weights of the directed edges in a flow network are called **capacities**. They give the maximum amount that can move between the two points in the flow network represented by these vertices in a particular time interval. This could be, for example,

the maximum amount of water in litres per minute or the maximum number of cars per hour.

Categorical variable [p. 2] A variable used to represent characteristics of individuals, for example place of birth, house number. Categorical variables come in types, nominal and ordinal.

Causal relationship [p. 84] When a change in the **explanatory variable** leads to a change in the **response variable**, this is known as a causal relationship.

Centring [p. 124] If smoothing takes place over an even number of data values, the smoothed values do not align with an original data value. A second stage of smoothing is carried out to centre the smoothed values at an original data value.

Coefficient of determination (R^2) [p. 45] A coefficient which gives a measure of the predictive power of a regression line. It gives the percentage of variation in the RV that can be explained by the variation in the EV.

Complete graph [p. 394] A graph with edges connecting all pairs of vertices.

Compound interest [p. 199] Where the interest paid on a loan or investment is added to the principal and subsequent interest is calculated on the total.

Compounding period [p. 290] The compounding period is the time period for the calculation of interest for an investment or loan. Typical compounding periods are yearly, quarterly, monthly or daily.

Connected graph [p. 394] A **connected graph** is a graph that has no isolated vertices and no separate parts.

Continuous variable [p. 3] A variable representing a quantity that is measured rather than counted, for example the weights of people in kilograms.

Coordinated Universal Time (UTC) [p. 241] A measure of time used to regulate time across the world. Equivalent to **GMT**.

Correlation coefficient r [p. 36] A statistical measure of the strength of the linear association between two numerical variables.

Cost matrix [p. 448] A cost matrix is a table that contains the cost of allocating objects from one group, such as people, to objects from another group, such as tasks. The cost can be money, or other factors such as the time taken to complete the project.

Critical path [p. 483] The project path that has the longest completion time.

Critical path analysis [p. 483] A project planning method in which activity durations are known with certainty.

Cut [p. 457] A line dividing a directed (flow) graph into two parts in a way that separates all 'sinks' from their 'sources'.

Cut capacity [p. 457] The capacity of a cut is the sum of the capacities of the cuts passing through the cut that represents flow from the source to the sink. Edges that represent flow from the sink to the source do not contribute to the capacity of the cut.

Cycle (graphs) [p. 413] A **walk** with no repeated vertices that starts and ends at the same vertex. *See also* **circuit**.

Cycle (time series) [p. 115] Periodic movement in a time series but over a period greater than a year.

D

Degree of a vertex (deg(A)) [p. 391] The number of edges attached to the vertex. The degree of vertex A is written as deg(A).

Deseasonalise [p. 130] The process of removing seasonality in time series data.

Directed graph (digraph) [p. 395] A graph or network in which directions are associated with each of the edges.

Discrete variable [p. 3] A variable representing a quantity that is determined by counting, for example, the number of people waiting in a queue.

Dummy activity [p. 477] An artificial activity of zero time duration added to a project diagram to ensure that all predecessor activities are properly accounted for.

E

Earliest starting time (EST) [p. 484] The earliest time an activity in a project can be started.

Edge [p. 391] A line joining one vertex in a graph or network to another vertex or itself (a loop).

Effective annual rate of interest [p. 303] Used to compare the interest paid on loans (or investments) with the same annual nominal interest rate r but with different compounding periods (daily, monthly, quarterly, annually, other).

Elements [p. 450] The numbers or symbols displayed in a matrix.

Equivalent graph [p. 405] *see* **isomorphic graphs**.

Eulerian trail [p. 416] A walk in a graph or network that includes every edge just once (but does not start and finish at the same vertex). To have an eulerian walk (but not an eulerian circuit), a network must be connected and have exactly two vertices of odd degree, with the remaining vertices having even degree.

Euler's formula [p. 407] The formula $v - e + f = 2$, which relates the number of vertices, edges and faces in a connected graph.

Explanatory variable [p. 4] When investigating associations in **bivariate data**, the explanatory variable (*EV*) is the variable used to explain or predict the value of the **response variable** (*RV*).

Extrapolation [p. 75] Using a mathematical model to make a prediction *outside* the range of data used to construct the model.

F

Face [p. 406] An area in a graph or network that can only be reached by crossing an edge. One such area is always the area surrounding a graph.

Flat-rate depreciation [p. 180] Depreciation where the value of an item is reduced by the same amount each year. Flat-rate depreciation is equivalent, but opposite, to simple interest.

Float (slack) time [p. 483] The amount of time available to complete a particular activity that does not increase the total time taken to complete the project.

Flow [p. 455] Flow is the transfer of material, such as water, gas or traffic through a directed network.

Flow network [p. 455] A **flow network** occurs where the directed edges of the graph represent the flow of material from one vertex to another. The weight of an edge of a flow network is called the **capacity** of that edge.

Forward scanning [p. 484] Forward scanning is the process of determining the EST for each activity in a project activity network.

G

Geometric decay [p. 188] When a recurrence rule involves multiplying by a factor less than one, the terms in the resulting sequence are said to decay geometrically.

Geometric growth [p. 188] When a recurrence rule involves multiplying by a factor greater than one, the terms in the resulting sequence are said to grow geometrically.

Geometric sequences [p. 187] A sequence is geometric if it satisfies the recurrence relation: $t_{n+1} = r \times t_n$ and a starting point usually t_1. Geometric sequences are used to model geometric growth and linear decay situations. The rule for the nth term of a geometric sequence is $t_n = r^{n-1}t_1 = r^{n-1}a$, where $a = t_1$ is the starting value.

Graph or network [pp. 391, 500] A collection of points called vertices and a set of connecting lines called edges.

Great circle [p. 223] A circle on a sphere whose plane passes through the centre of the sphere. The shortest distance between two points on a sphere is along an arc of the great circle passing through the two points. *See also* **small circle**.

Greenwich Mean Time (GMT) [p. 241] Equivalent to **UTC**, this is a measure of time centred around Greenwich, England and is used across the world.

H

Hamiltonian cycle [p. 420] A hamiltonian path that starts and finishes at the same vertex.

Hamiltonian graph [p. 420] A Hamiltonian graph is a graph that contains a **Hamiltonian cycle**.

Hamiltonian path [p. 420] A path through a graph or network that passes through each vertex exactly once. It may or may not start and finish at the same vertex.

Hungarian algorithm [p. 448] An algorithm for solving allocation (assignment) problems.

I

Immediate predecessor [p. 473] An activity that must be completed immediately before another one can start.

Intercept (of a straight line) [p. 60] Where the regression line cuts across the y-axis.

Interest [p. 178] The amount of money paid (earned) for borrowing (lending) money over a period of time.

Interest rate [p. 115] The rate at which interest is charged or paid. Usually expressed as a percentage of the money owed or lent.

International Date Line [p. 242] An imaginary line through the Pacific Ocean that corresponds to 180° **longitude**.

Interpolation [p. 75] Using a regression line to make a prediction *within* the range of values of the explanatory variable.

Irregular (random) fluctuations [p. 117] Unpredictable fluctuations in a time series. Always present in any real world time series plot.

Isolated vertex [p. 393] A vertex that is not connected to any other vertex. Its degree is zero.

Isomorphic graphs [p. 405] Equivalent graphs. Graphs that have the same number of edges and vertices that are identically connected.

L

Latest start time (LST) [p. 487] The latest time an activity in a project can begin, without affecting the overall completion time for the project.

Latitude [p. 224] The angle or angular distance north or south of the equator.

Least squares method [p. 56] One way of finding the equation of a regression line. It minimises the sum of the squares of the residuals. It works best when there are no outliers.

Linear decay [p. 169] When a recurrence rule involves subtracting a fixed amount, the terms in the resulting sequence are said to decay linearly.

Linear growth [p. 169] When a recurrence rule involves adding a fixed amount, the terms in the resulting sequence are said to grow linearly.

Linear regression [p. 58] The process of fitting a straight line to bivariate data.

Longitude [p. 224] The angle or angular distance east or west of the **prime meridian**.

Loop [p. 391] An edge in a graph or network that joins a vertex to itself.

M

Matrix [p. 399] A rectangular array of numbers or symbols set out in rows and columns within square brackets (*pl*: matrices).

Maximum flow (graph) [p. 455] The capacity of the 'minimum' cut.

Mean ($\bar{x}$) [p. 60] The balance point of a data distribution. The mean is given by $\bar{x} = \frac{\sim x}{n}$, where $\sim x$ is the sum of the data values and n is the number of data values. Best used for symmetric distributions.

Median [p. 89] The median (M) is the middle value in a data distribution. It is the midpoint of a distribution dividing an ordered data set into two equal parts. Can be used for skewed or symmetric distributions.

Meridian [p. 224] Semi-great circles that pass through north and south poles.

Meridians of longitude [p. 224] Semi-great circles which pass through the north and south poles.

Minimum cut (graph) [p. 458] The cut through a graph or network with the minimum capacity.

Minimum spanning tree [p. 441] The spanning tree of minimum length. For a given connected graph, there may be more than one minimum spanning tree.

Modelling [pp. 198, 285] Mathematical modelling is the use of a mathematical rule or formula to represent real-life situations.

Moving mean smoothing [p. 121] In three-moving mean smoothing, each original data value is replaced by the mean of itself and the value on either side. In five-moving mean smoothing, each original data value is replaced by the mean of itself and the two values on either side.

Multiple edge [p. 393] Where more than one edge connects the same two vertices in a graph.

N

Network [pp. 424, 526] A set of points called vertices and connecting lines called edges, enclosing and surrounded by areas called faces.

Nominal interest rate [p. 290] The annual interest rate for a loan or investment that assumes the compounding period is 1 year. If the compounding period is less than a year, for example monthly, the actual or **effective interest rate** will be greater than r.

Numerical variable [p. 2] A variable used to represent quantities that are counted or measured. For example, the number of people in a queue, the heights of these people in cm. Numerical variables come in types: discrete and continuous.

O

Outliers [pp. 26, 117] Data values that appear to stand out from the main body of a data set.

P

Parallels of latitude [p. 224] Small circles whose planes are parallel to that of the equator.

Path [p. 412] A **walk** with no repeated vertices. *See also* **trail**.

Percentage frequency [p. 11] Frequency expressed as a percentage.

Perpetuity [p. 376] An investment where an equal amount is paid out on a regular basis forever.

Planar graph [p. 405] A graph that can be drawn in such a way that no two edges intersect, except at the vertices.

Precedence table [p. 474] A table that records the activities of a project, their immediate predecessors and often the duration of each activity.

Prim's algorithm [p. 441] An algorithm for determining a minimum spanning tree in a connected graph.

Prime meridian [p. 225] The meridian located at 0° which passes through Greenwich, England.

Principal (P) [p. 178] The initial amount borrowed, lent or invested.

R

Radius [p. 219] The distance from the centre to any point on the circle (sphere). Half the diameter.

Recurrence relation [pp. 167, 281] A relation that enables the value of the next term in a sequence to be obtained by one or more current terms. Examples include 'to find the next term, add two to the current term' and 'to find the next term, multiply the current term by three and subtract five'.

Reducing-balance depreciation [p. 201] When the value of an item is reduced by the same percentage each year. Reducing-balance depreciation is equivalent to, but opposite to, compound interest.

Reducing-balance loan [p. 324] A loan that attracts compound interest, but where regular repayments are also made. In most instances the repayments are calculated so that the amount of the loan and the interest are eventually repaid in full.

Reseasonalise [p. 130] The process of converting seasonal data back into its original form.

Residual [p. 59] The vertical distance from a data point to a straight line fitted to a scatterplot is called a residual:

residual = actual value − predicted value

Residuals are sometimes called *errors of prediction*.

Residual plot [p. 68] A plot of the residuals against the explanatory variable. Residual plots can be used to investigate the linearity assumption.

Response variable [p. 4] The variable of primary interest in a statistical investigation.

S

Scatterplot [p. 20] A statistical graph used for displaying bivariate data. Data pairs are represented by points on a coordinate plane, the EV is plotted on the horizontal axis and the RV is plotted on the vertical axis.

Seasonal indices [p. 130] Indices calculated when the data shows seasonal variation. Seasonal indices quantify seasonal variation. A seasonal index is defined by the formula:

$$\text{seasonal index} = \frac{\text{value for season}}{\text{seasonal average}}$$

For seasonal indices, the average is 1 (or 100%).

Seasonality [p. 115] The tendency for values in the time series to follow a seasonal pattern, increasing or decreasing predictably according to time periods such as time of day, day of the week, month, or quarter.

Sequence [p. 163] A list of numbers or symbols written down in succession, for example 5, 15, 25, . . .

Shortest path [p. 424] The path through a graph or network with minimum length.

Simple graph [p. 394] A graph with no loops or multiple edges.

Simple interest [pp. 178, 285] Interest that is calculated for an agreed period and paid only on the original amount invested or borrowed.

Sink [p. 455] *See* **sink and source**.

Sink and source [p. 455] In a flow network, a **source** generates flow while a **sink** absorbs the flow.

Slope (of a straight line) [p. 60] The slope of a straight line is defined to be: slope $= \frac{rise}{run}$. The slope is also known as the gradient.

Small circle [p. 223] Any circle on a sphere whose plane does not pass through the centre of the sphere. *See also* **great circle**.

Smoothing [p. 121] A technique used to eliminate some of the variation in a time series plot so that features such as seasonality or trend are more easily identified.

Source [p. 455] *See* **sink and source**.

Spanning tree [p. 439] A subgraph of a connected graph that contains all the vertices of the original graph, but without any multiple edges, circuits or loops.

Standard deviation (*s*) [p. 60] A summary statistic that measures the spread of the data values around the mean. The standard deviation is given by $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$

Strength of a linear relationship [p. 36] Classified as weak, moderate or strong. Determined by observing the degree of scatter in a scatterplot or calculating a correlation coefficient.

Structural change (time series) [p. 116] A sudden change in the established pattern of a time series plot.

Subgraph [p. 395] Part of a graph that is also a graph in its own right.

T

Time series data [p. 107] A collection of data values along with the times (in order) at which they were recorded.

Time series plot [p. 107] A line graph where the values of the response variable are plotted in time order.

Time zone [p. 241] A region of the Earth that has a uniform standard time or local time. There are 24 time zones in total.

Trail [p. 412] A walk with no repeated edges. *See also* **path**.

Tree [p. 439] A **connected graph** with no circuits, multiple edges or loops.

Trend [p. 113] The tendency for values in the time series to generally increase or decrease over a significant period of time.

Trend line forecasting [p. 143] Using a line fitted to an increasing or decreasing time series to predict future values.

Two-way frequency table [p. 8] A frequency table in which subjects are classified according to two categorical variables. Two-way frequency tables are commonly used to investigate the associations between two categorical variables.

U

Unit-cost depreciation [p. 180] Depreciation based on how many units have been produced or consumed by the object being depreciated. For example, a machine filling bottles of drink may be depreciated by 0.001 cents per bottle it fills.

V

Variable [p. 2] A symbol used to represent a number or group of numbers.

Vertex (graph) [p. 391] The points in a graph or network (*pl* vertices).

W

Walk [p. 411] Any continuous sequence of edges, linking successive vertices, that connects two different vertices in a graph. *See also* **trail** and **path**.

Weighted graph [p. 424] A graph in which a number representing the size of some quantity is associated with each edge. These numbers are called weights.