Types of Data

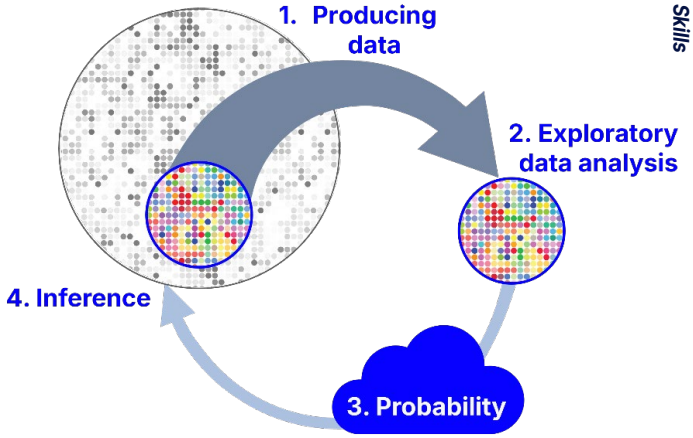
Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is the process of organising and summarising data to discover patterns and relationships.

EDA helps us:

- Summarise data
- Identify patterns and trends
- Compare groups
- Detect unusual values (outliers)
- Investigate relationships between variables
- Choose appropriate graphs and statistics

EDA is often called Descriptive Statistics.



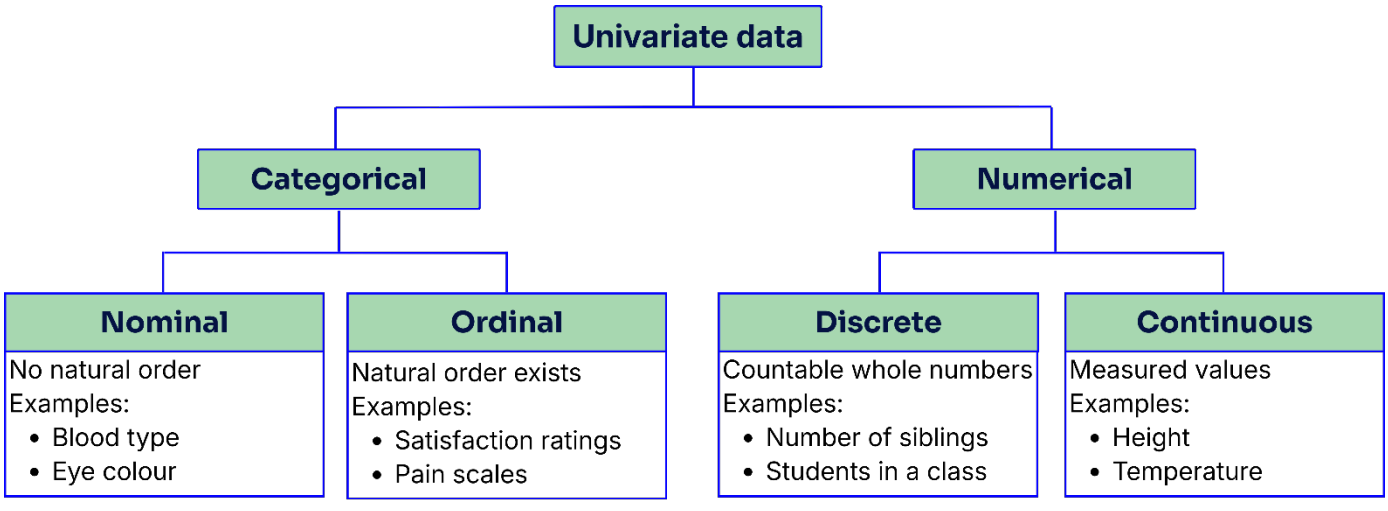
What is a variable?

A variable is a measurable or describable characteristic that can differ between observations. When analysing data, we may investigate:

Data type	No. of variables	Purpose
Univariate	One	Describe and summarise a variable
Bivariate	Two	Investigate relationships between variables

The number of variables being studied influences the choice of graphs, summary statistics, and analysis methods.

Types of variables



CRICOS 00103D | RTO 4909 | TEQSA PRV12151 (Aust)

Choosing displays for EDA

The type of univariable determines the most appropriate display.

Variable type	Numerical measures:	Graphical displays:
Nominal	Frequency Relative frequency	Bar chart, pie chart
Ordinal		Ordered bar chart
Discrete	Measures of centre Measures of spread	Dot plot, stem-and-leaf plot
Continuous		Histogram, box plot

Bivariate analysis

The purpose of bivariate analysis is to investigate relationships, associations, or differences between variables. When examining relationships, we consider two variables:

- Explanatory (X): The variable that may influence or predict another variable
- Response (Y): The outcome or result being measured

Relationship:

Explanatory Variable → Response Variable

		Response	
		Categorical	Quantitative
Explanatory	Categorical	C → C	C → Q
	Quantitative	Q → C	Q → Q

Variable types and relationships

The type of variables being investigated determines the graphs and statistical methods used.

Relationship type	Common displays:
Categorical → Categorical	Two-way table
Categorical → Quantitative	Side-by-side box plots
Quantitative → Quantitative	Scatter plot

Related helpsheets

- The Statistical Process
- Types of Studies
- Producing Data

CRICOS 00103D | RTO 4909 | TEQSA PRV12151 (Aust)