Describing Data

Describing data distributions



Whenever you are asked to describe or compare a data set, comment on:

1. Shape
2. Centre
3. Spread
4. Outliers

Shape

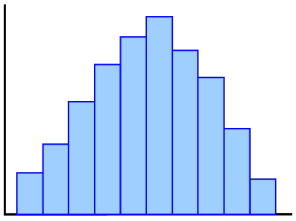
The shape describes how the data is distributed.

Symmetric distribution

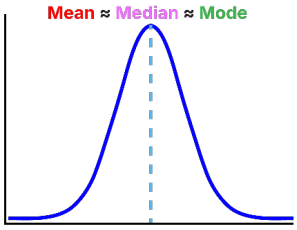
A symmetric distribution is balanced around the centre.

Features

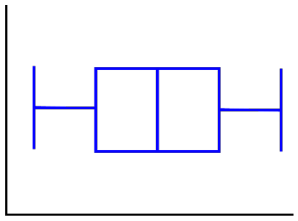
- Both sides are similar



- Mean ≈ Median ≈ Mode



- No obvious skew

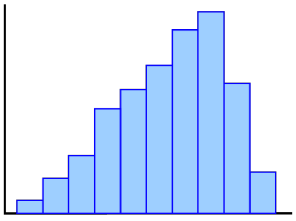


Negatively skewed

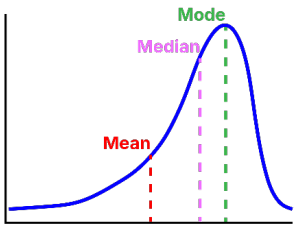
A negatively skewed distribution has a long tail to the left.

Features

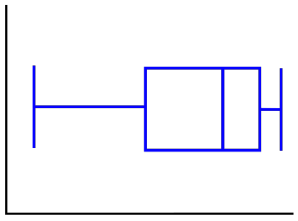
- Most values are on the right



- Mean < Median



- Low values stretch the distribution to the left

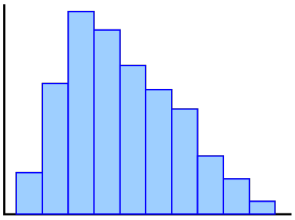


Positively skewed

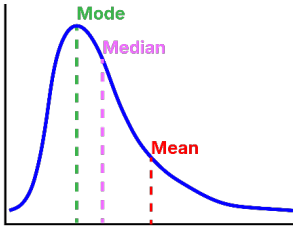
A positively skewed distribution has a long tail to the right.

Features

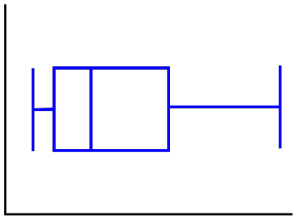
- Most values are on the left



- Mean > Median



- High values stretch the distribution to the right

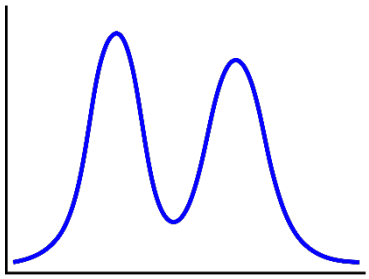


CRICOS 00103D | RTO 4909 | TEQSA PRV12151 (Aust)

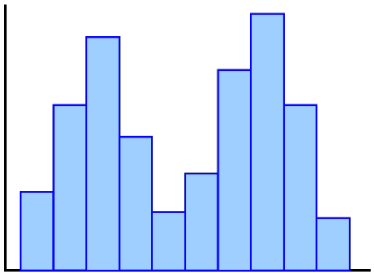
Bimodal
A bimodal distribution has two distinct peaks or modes.

Features

- Two common groups of values



- Two modes



Useful sentence starters

Shape

- The distribution is symmetric
- The distribution is positively skewed
- The distribution is negatively skewed
- The distribution is bimodal

Comparing shape

- Both datasets are approximately symmetric
- Dataset A is more positively skewed than Dataset B
- Dataset B is more negatively skewed than Dataset A
- Dataset A is bimodal, while Dataset B is unimodal
- The distributions have similar/different shapes

Centre

- The median is ...
- The mean is ...
- The typical value is ...

Comparing centre

- Dataset A has a higher median than Dataset B
- Dataset B has a higher mean than Dataset A
- The centres of the two datasets are similar
- Dataset A has a greater typical value than Dataset B

Related helpsheets

- Measures of Centre
- Measures of Spread
- Standard Deviation
- Descriptive Statistics in Jamovi

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