



TOOLBOX

# SUMMARISING DATA



MACQUARIE  
University  
SYDNEY · AUSTRALIA

## Frequency Tables

To summarise one categorical variable we need to:

**a. Construct a table showing the:**

- Variable name
- Name of each category
- The count, proportion and/or
- Percentage of observations in each category

**b. The table should also show the sample size (n).**

## Excel Commands

### Mean

=AVERAGE(array of numbers)

### Median

=Median(array of numbers)

### Variance

=VAR.S(array of numbers)

### Standard Deviation

=STDEV.S(array of numbers)

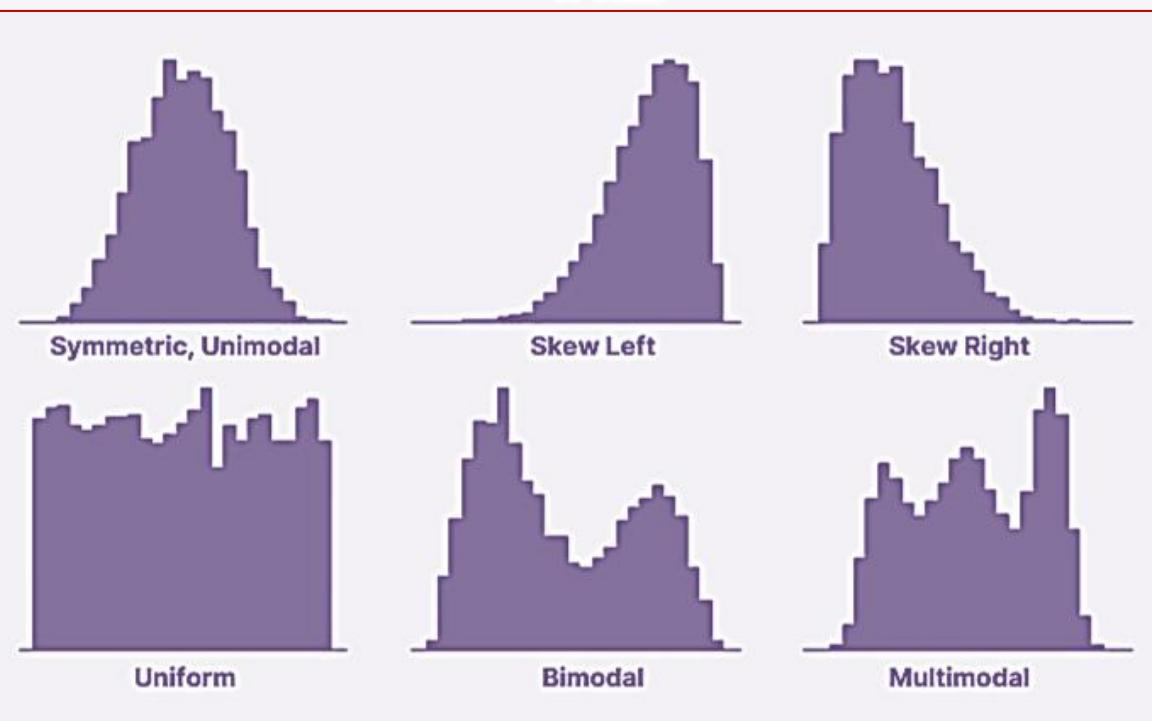
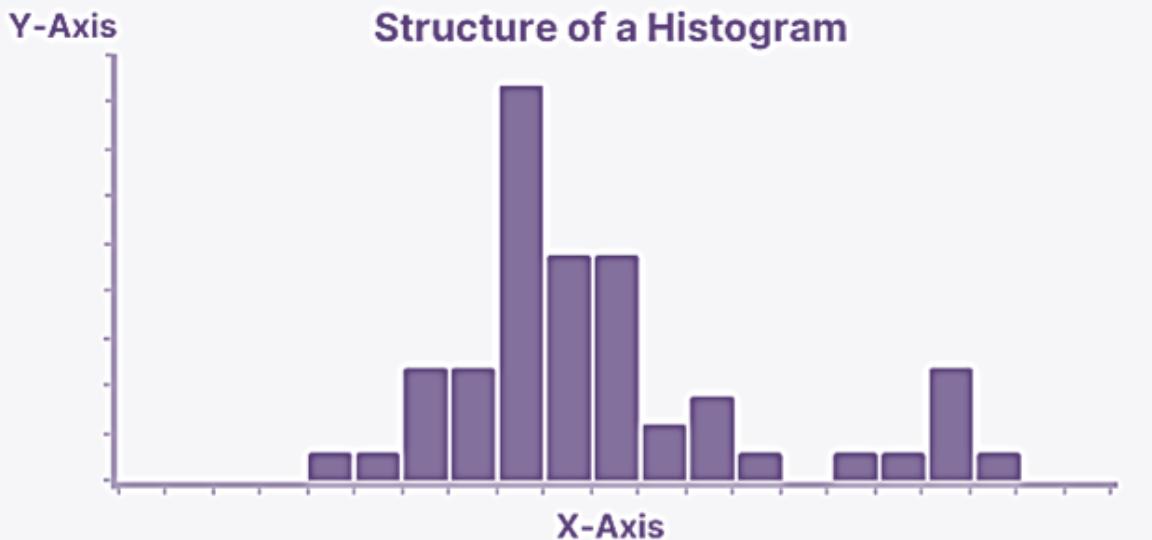
### Range of Values

=MAX(array of numbers)-MIN(array of numbers)

### Interquartile Range

=QUARTILE.EXC(array of numbers, 3)-QUARTILE.EXC(array of numbers, 1)

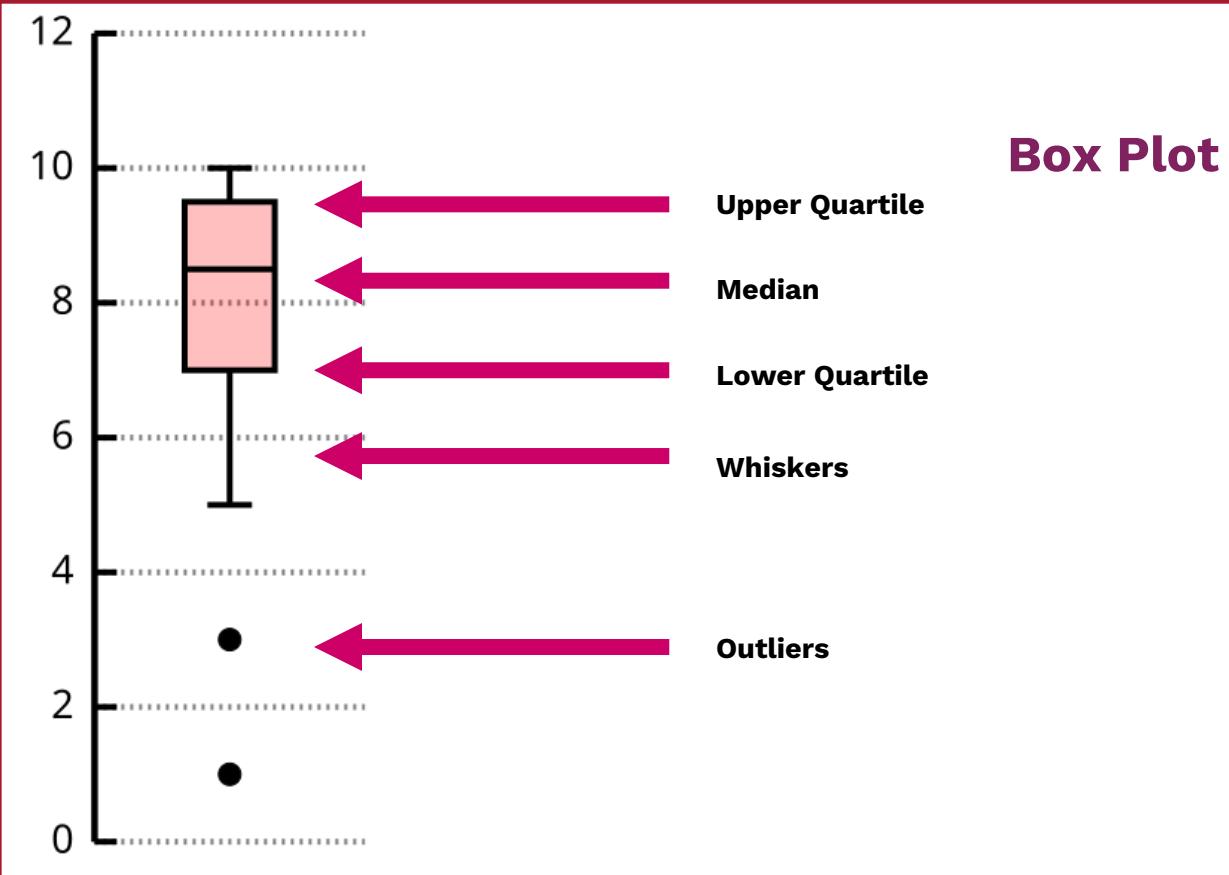
## Structure of a Histogram



## Notation

For variable  $X$  we label observations using a subscript i.e.

- The first observation:  $X_1$
- The second observation:  $X_2$
- The n-th observation:  $X_n$



Sample Statistic	Population Parameter
Mean	$\bar{x}$
Standard Deviation	$\sigma$
Variance	$\sigma^2$