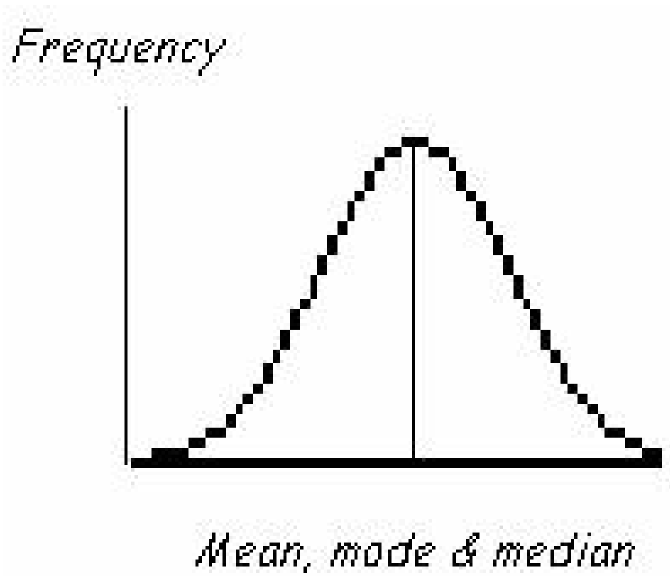
The Need for Sampling

- In order to carry out primary research it is necessary to use **sampling**
- This is because it would be impossible to ask every single customer, or prospective customer
- Sampling involves selecting a few people to interview, and so it is important that they are representative of the market being looked at
- **The Normal Distribution** is a statistical model which is often used to show why sampling is necessary

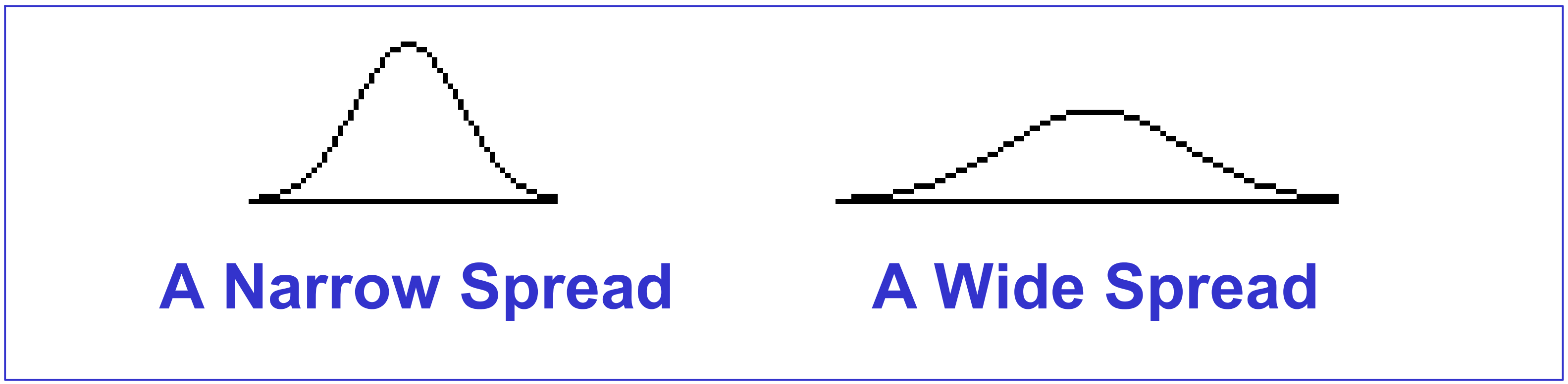
The Normal Distribution

- Tells a business what the expected range of outcomes from a particular population will be
- It is useful where businesses have used a large scale sample
- A Normal distribution is a symmetrical frequency distribution, which can be split in 2 equal halves, so that the mean, mode and median are equal
- Such a distribution can be shown diagrammatically:



The Normal Distribution (cont'd)

- It is the spread of the data that determines the steepness or shallowness of a curve:



- The spread is measured by the STANDARD DEVIATION
- This measures how far on average a figure is from the mean of the distribution.

The Normal Distribution (cont'd)

- As such in order to use the normal distribution we need to be able to calculate the mean and standard deviation
- For a normal distribution these are calculated as follows:

> - Mean = np
> - Standard Deviation = $\sqrt{npq}$
> - Where:
>   - n = the sample size
>   - p = the probability of an event occurring
>   - q = the probability of an event not occurring

Dividing The Normal Distribution Up

- Any normal distribution can be divided into 3 sections by adding or subtracting the standard deviation upto 3 times either side of the mean
- Mathematicians have discovered that:

68% of the sample will lie between ± 1 Standard Deviation

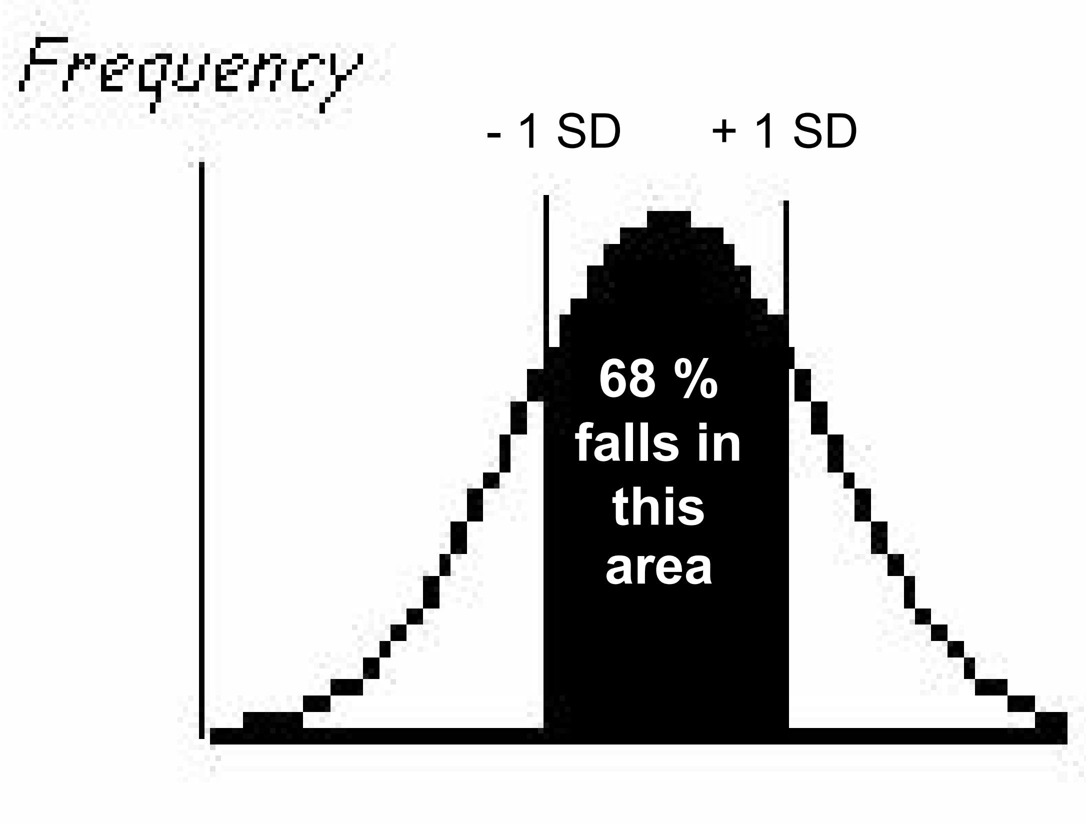
95% of the sample will lie between ± 2 Standard Deviations

99.9% of the sample will lie between ± 3 Standard Deviations

  - (don't worry about why!)
- These are known as "**confidence levels**"
- They can be shown diagrammatically on a normal distribution curve.

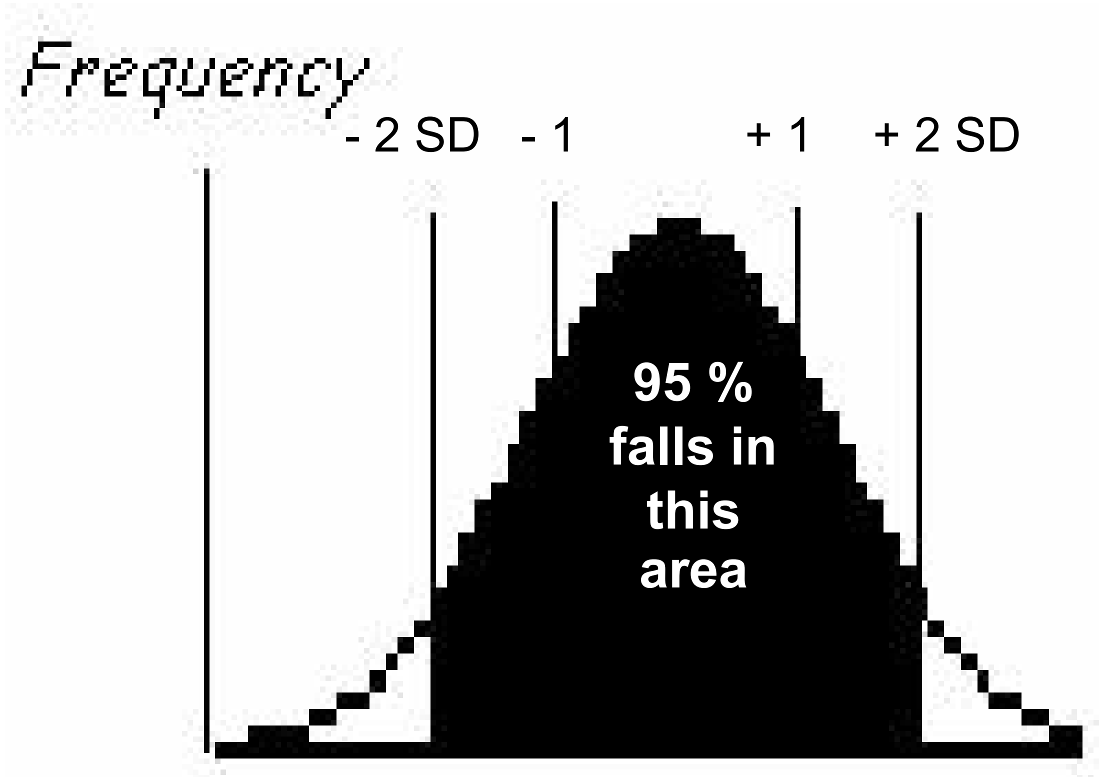
68 % Confidence Level

- The diagram shows that 68% of the population will fall within the area between - 1 SD and + 1 SD:



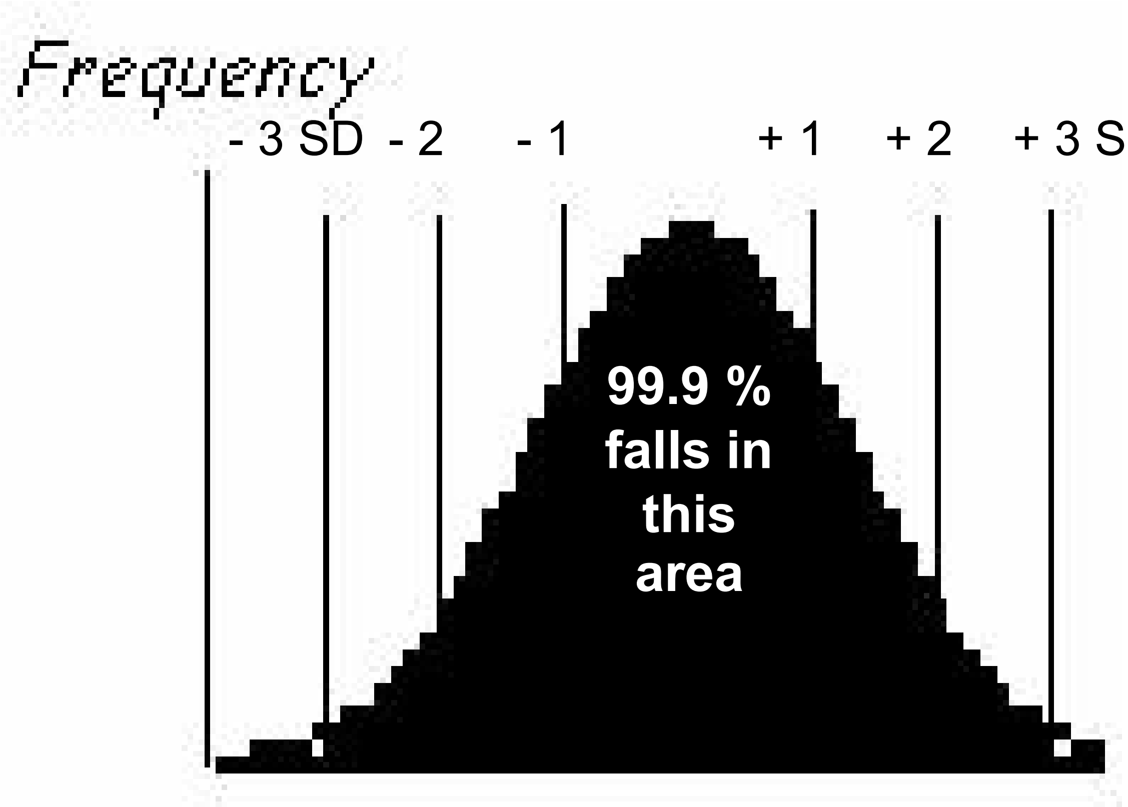
95 % Confidence Level

- The diagram shows that 95% of the population will fall within the area between - 2 SD and + 2 SD:



99.9 % Confidence Level

- The diagram shows that 99.9% of the population will fall within the area between - 3 SD and + 3 SD:



Using the Normal Distribution

- Cadbury’s can only afford to launch one new chocolate bar, but have 2 possible new products to choose from.
- They conduct research and ask 150 people whether they prefer product X to product Y, and assume that the sample is normally distributed.
- 60% of those asked preferred product X, but can they be 95% confident that the majority of their customers will prefer product X?

The Solution

- Using the normal distribution we know that:
  - n = 150
  - p = 0.6
  - q = (1-0.6) = 0.4
  - Therefore:
    - mean = 150 * 0.6 = **90**
    - SD = $\sqrt{150 * 0.6 * 0.4}$ = **6 or 4% (6/150 * 100)**
  - Since 95% confidence level is ± 2SD we know that the true result may lie between (60% - 8%) and (60% + 8%)
  - This means that we can be 95% confident that between 52% and 68% of people will prefer product X
  - Since 51% represents a majority then Cadbury's can confidently launch product X rather than product Y!