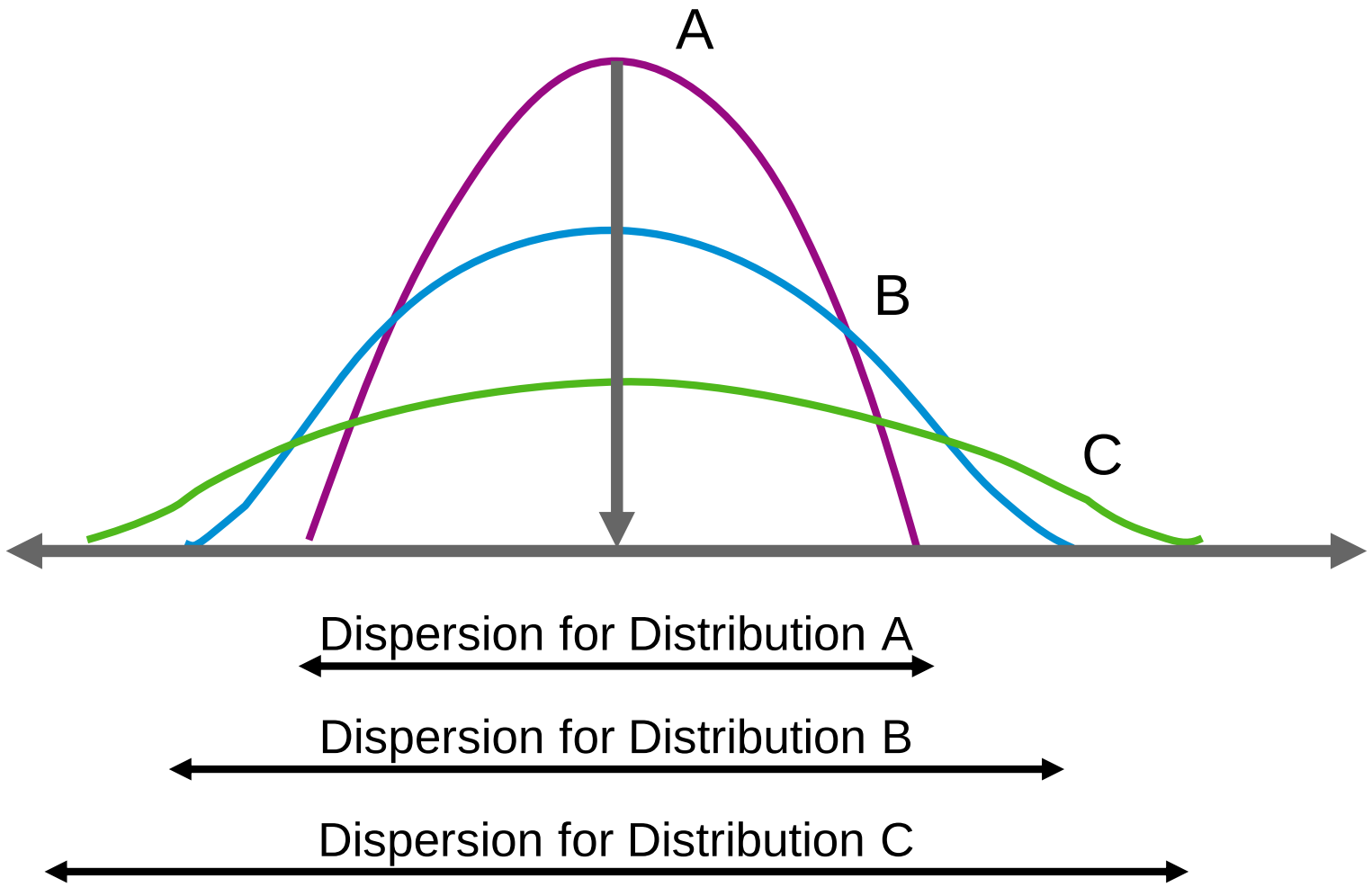
Week 2: Descriptive Statistics

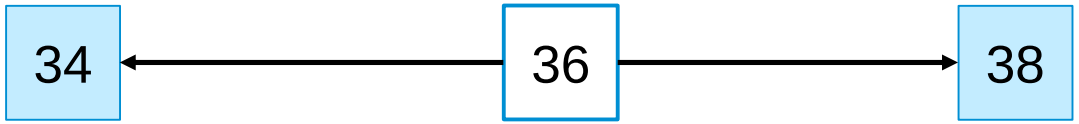
Unit 5: Measures of Dispersion

Introduction

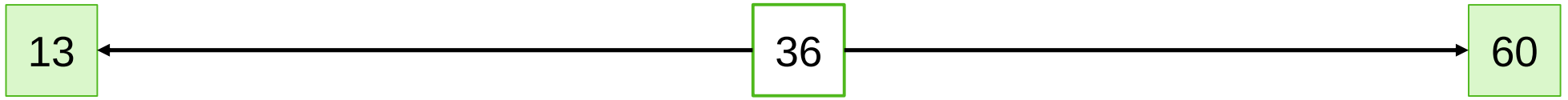


Measure of variation

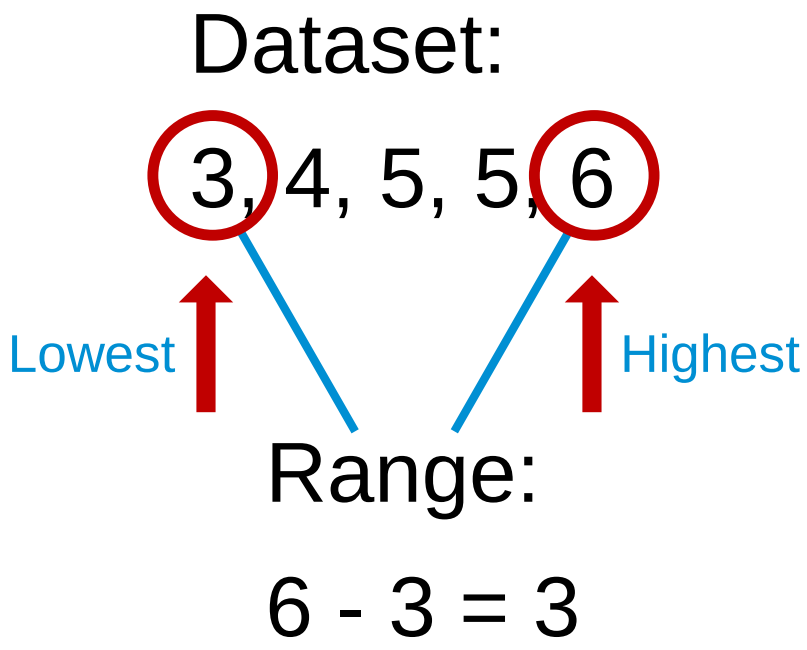
Ages of Group A: 34, 35, 35, 37, 37, 38 → Average is 36 years



Ages of Group B: 13, 15, 17, 55, 56, 60 → Average is 36 years



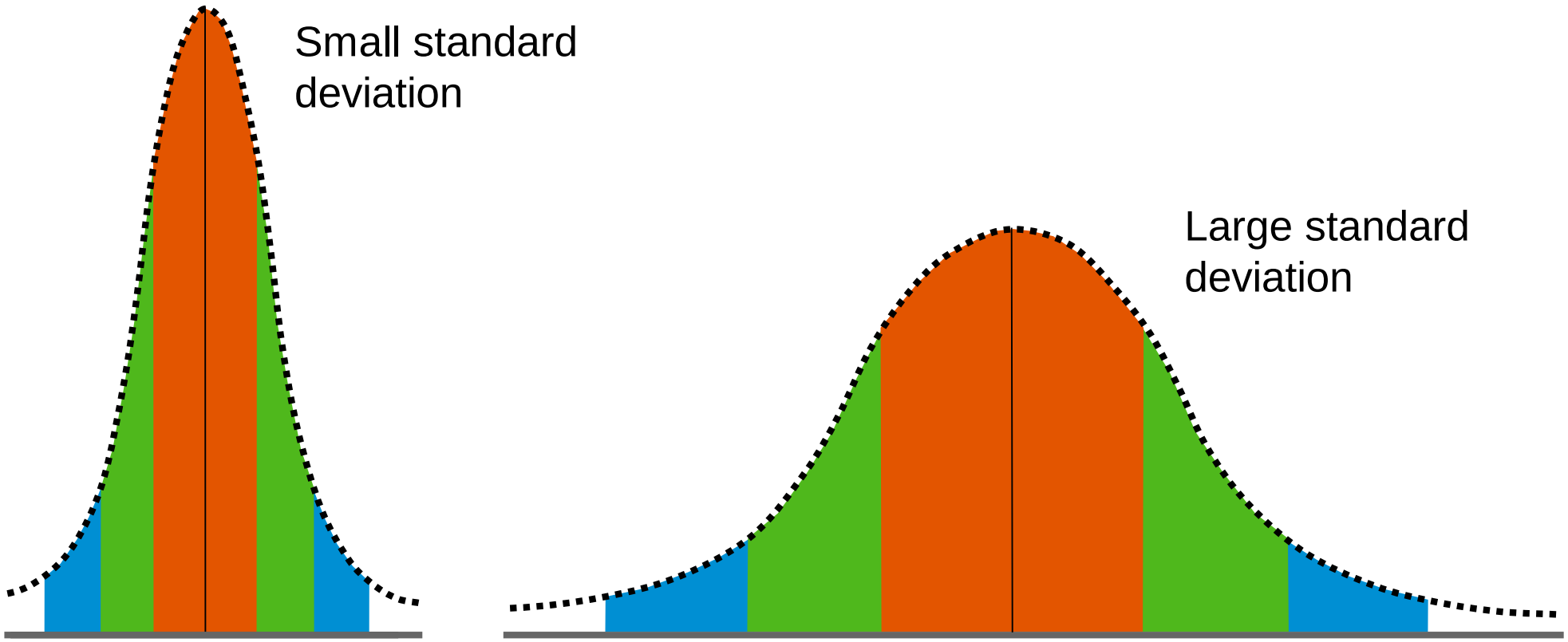
Range



The three data sets below may seem to have the same variability but in fact the variation is not the same:

Data Set A: 1, 2, 3, 3, 4, 5 Range: 5 –1 = 4
Data Set B: 1, 1, 3, 3, 5, 5 Range: 5 –1 = 4
Data Set C: 1, 1, 1, 5, 5, 5 Range: 5 –1 = 4

Standard deviation



Standard deviation for sample or population

- Standard deviation can be applied to a sample and a population
- The formulas to calculate standard deviation are:

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

The population
Number = N
Mean = μ
Standard deviation = σ

Sample standard deviation:

$$s_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

The sample
Sample size = **n**
Sample mean = $\bar{x}$
Sample standard deviation = **s**

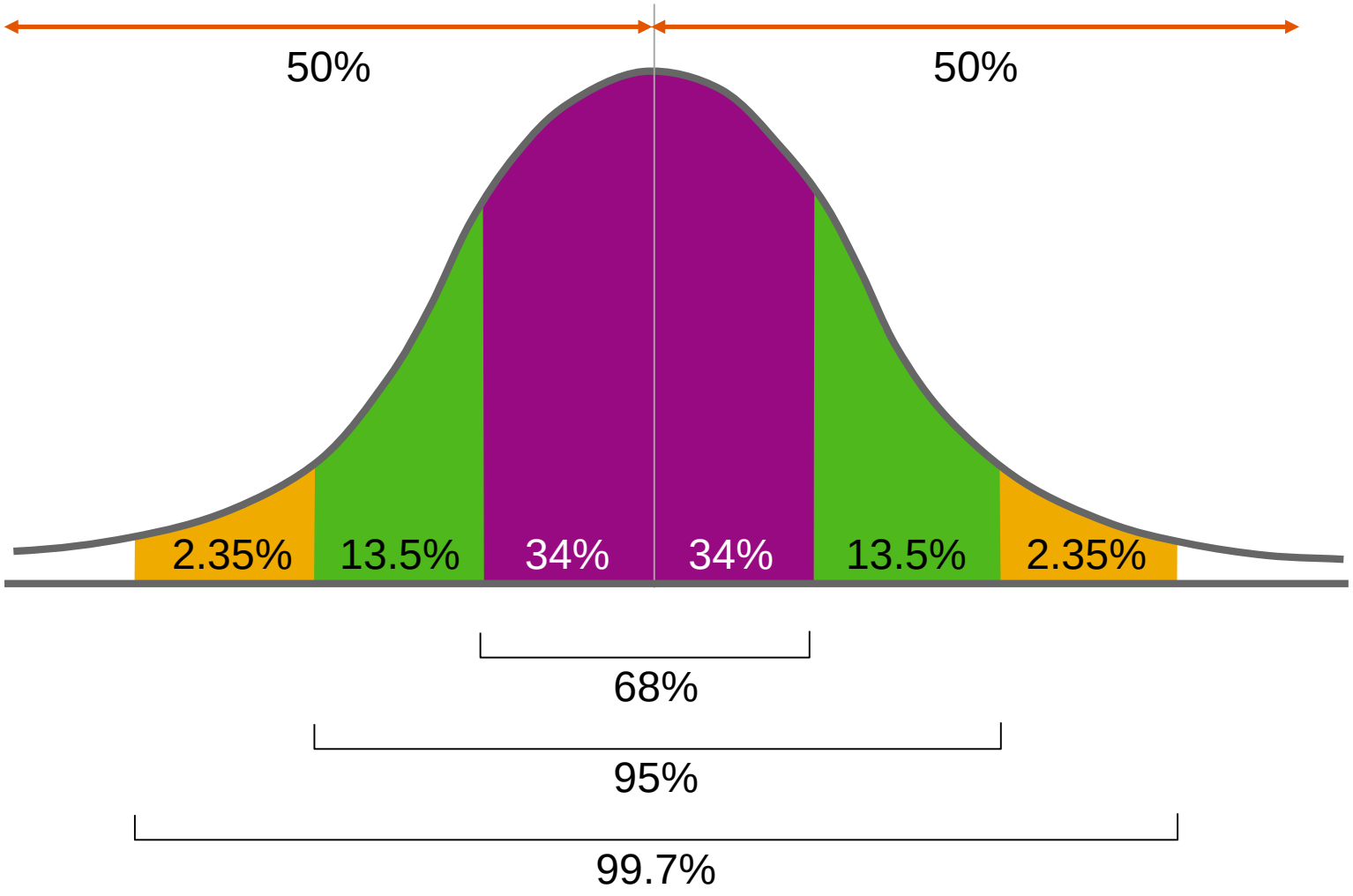
For more information see:

https://www.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/variance-standard-deviation-sample/a/population-and-sample-standard-deviation-review

and

https://www.mathsisfun.com/data/standard-deviation-formulas.html

Using the mean and standard deviation to describe data – The empirical rule



Variance

Sample Variance

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

Sample Standard Deviation

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Symbols for variance and standard deviation:

$$s^2 = Sample\ variance$$
$$s = Sample\ standard\ deviation$$

$$\sigma^2 = Population\ variance$$
$$\sigma = Population\ standard\ deviation$$

Summary

- You have seen why looking at the range or average values of a distribution without considering the dispersion will not give you a full understanding of the data.
- Common measures include variance and standard deviation.
- You have also been introduced to the Empirical Rule that helps explain the spread in a distribution.

Thank you.

Contact information:

open@sap.com

Follow all of SAP

www.sap.com/contactsap

THE BEST RUN SAP®