Measures of Variability: Takeaways ⇪

Syntax

- Writing a function that returns the mean absolute deviation of a vector:

```
mean_absolute_deviation <- function(vector) {

    distances  <-  abs(vector - mean(vector))

    sum(distances) / length(distances)

}
```

- Writing a function that returns the variance of a vector:

```
variance <- function(vector) {

    distances  <-  (vector - mean(vector))**2

    sum(distances) / length(distances)

}
```

- Writing a function that returns the standard deviation of a vector:

```
standard_deviation <- function(vector) {

    distances  <-  (vector - mean(vector))**2

    sqrt(sum(distances) / length(distances) )

}
```

- Computing the variance of a vector with a R base function:

```
var(vector)
```

- Computing the standard deviation of a vector with a R base function:

```
sd(vector)
```

Concepts

- There are many ways we can measure the **variability** of a distribution. These are some of the measures we can use:
  - **The range**
  - **The mean absolute deviation**
  - **The variance**
  - **The standard deviation**

- Variance and standard deviation are the most used metrics to measure variability. To compute the standard deviation and the variance for a **population**, we can use the formulas:

- To compute the standard deviation and the variance for a **sample**, we need to add the **Bessel's correction** to the formulas above:

- **Sample variance** is the only unbiased estimator we learned about, and it's unbiased only when we sample with replacement.

Resources

- An intuitive introduction to variance and standard deviation.
- Useful documentation:
    - R base function `var()`
    - R base function `sd()`