

What is the variation in statistics?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the variation in statistics?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=158984>

The variation in statistics refers to the degree of dispersion or spread of data points in a given dataset. It measures the differences between individual data points and the overall average or central tendency of the data. This variation is an important aspect in understanding the distribution and behavior of a dataset, as it can provide insights into the level of diversity and consistency within the data. Variations can be calculated using various statistical measures such as variance, standard deviation, and range, and can help in making informed decisions and drawing accurate conclusions from data analysis.

Statistics - Variation

Variation is a measure of how spread out the data is around the center of the data.

The Variation of the Data

Measures of variation are statistics of how far away the values in the observations (data points) are from each other.

There are different measures of variation. The most commonly used are:

Range, Quartiles and Percentiles, Interquartile Range, Standard Deviation

Measures of variation combined with an average (measure of center) gives a good picture of the distribution of the data.

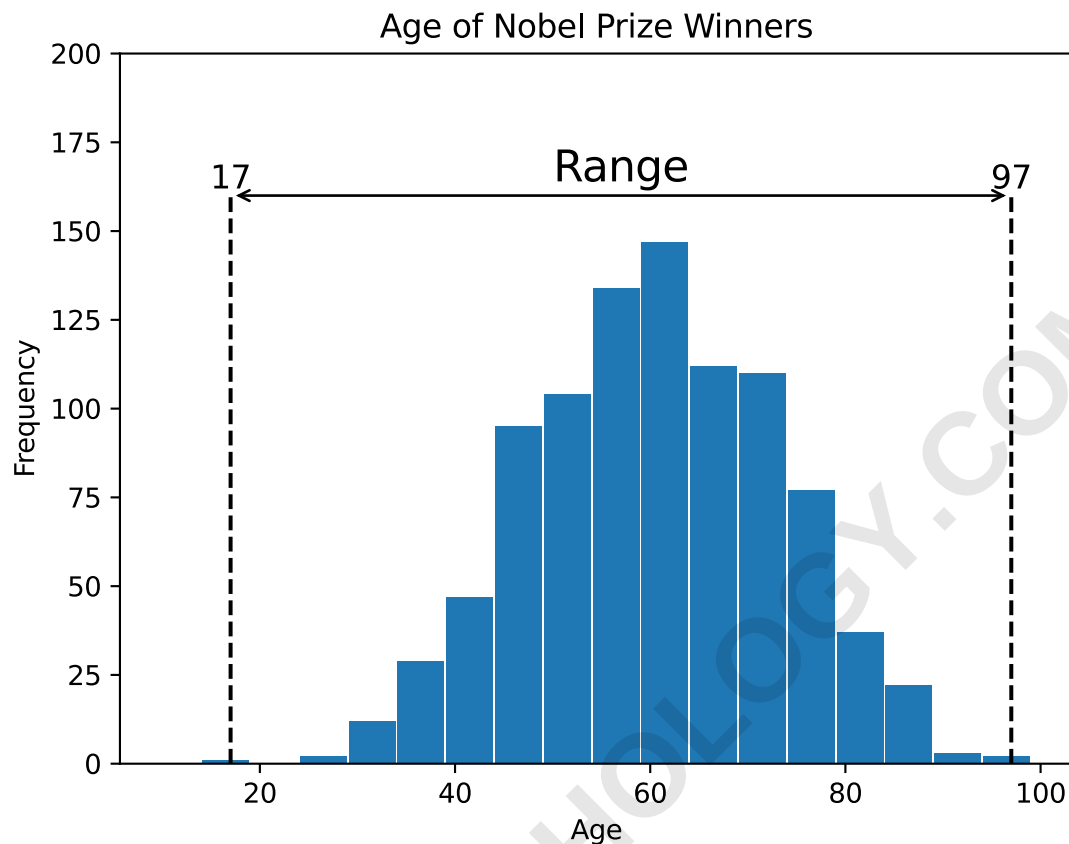
Note: These measures of variation can only be calculated for numerical data.

Range

The range is the difference between the smallest and the largest value of the data.

Range is the simplest measure of variation.

Here is a histogram of the age of all 934 Nobel Prize winners up to the year 2020, showing the **range**:



The youngest winner was 17 years and the oldest was 97 years. The range of ages for Nobel Prize winners is then 80 years.

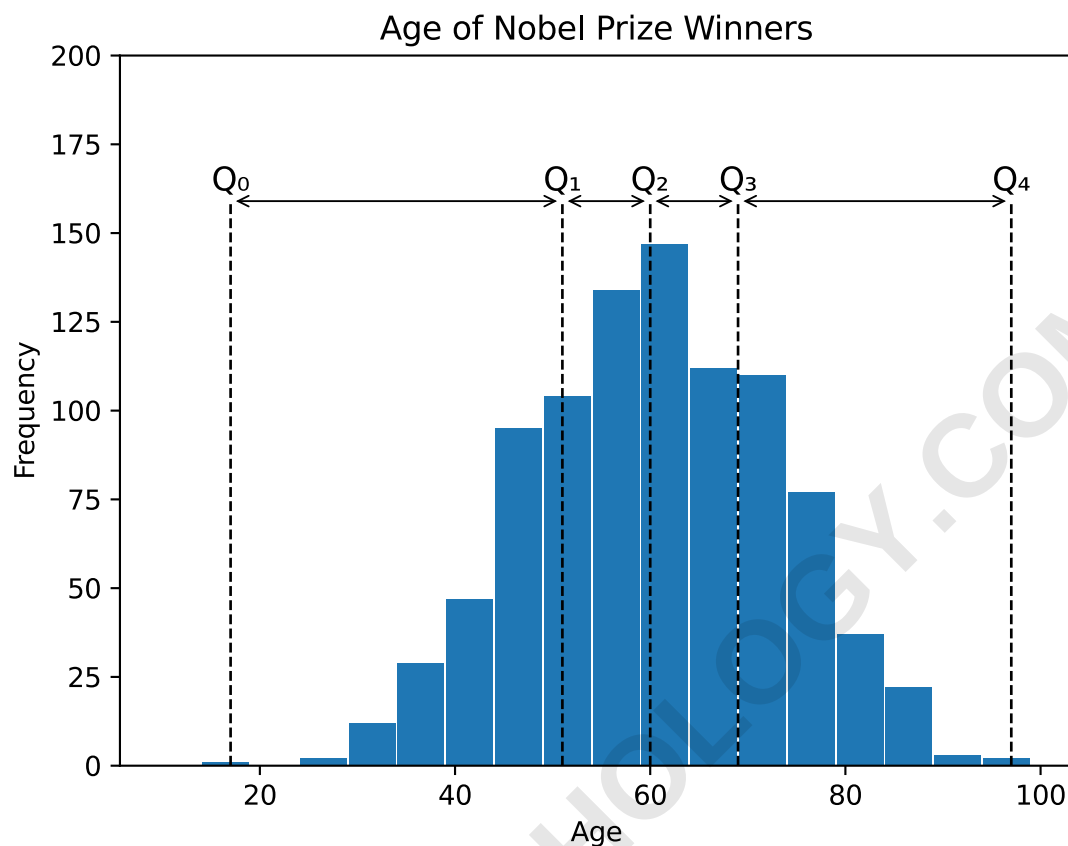
Quartiles and Percentiles

Quartiles and percentiles are ways of separating equal numbers of values in the data into parts.

Quartiles are values that separate the data into four equal parts.

Percentiles are values that separate the data into 100 equal parts.

Here is a histogram of the age of all 934 Nobel Prize winners up to the year 2020, showing the **quartiles**:



The quartiles (Q_0, Q_1, Q_2, Q_3, Q_4) are the values that separate each quarter.

Between Q_0 and Q_1 are the 25% lowest values in the data. Between Q_1 and Q_2 are the next 25%. And so on.

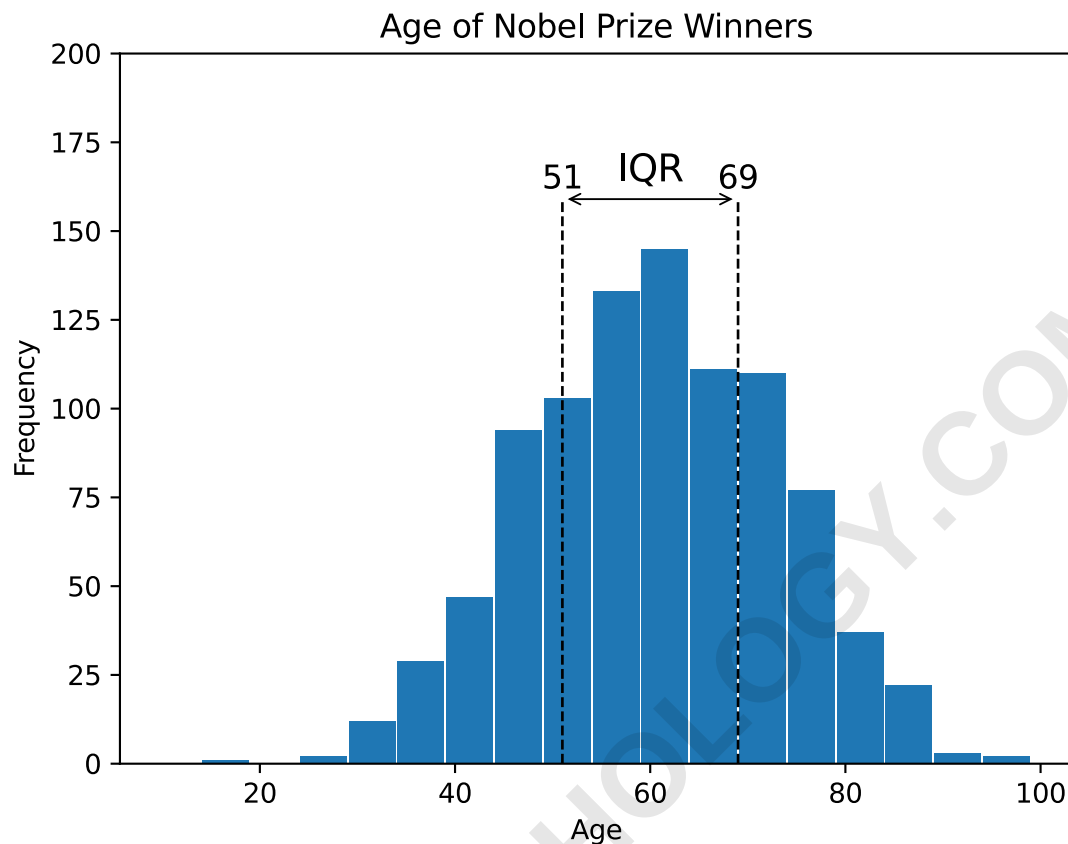
Q_0 is the smallest value in the data. Q_2 is the middle value (median). Q_4 is the largest value in the data.

Interquartile Range

Interquartile range is the difference between the first and third quartiles (Q_1 and Q_3).

The 'middle half' of the data is between the first and third quartile.

Here is a histogram of the age of all 934 Nobel Prize winners up to the year 2020, showing the **interquartile range (IQR)**:



Here, the middle half of is between 51 and 69 years. The interquartile range for Nobel Prize winners is then 18 years.

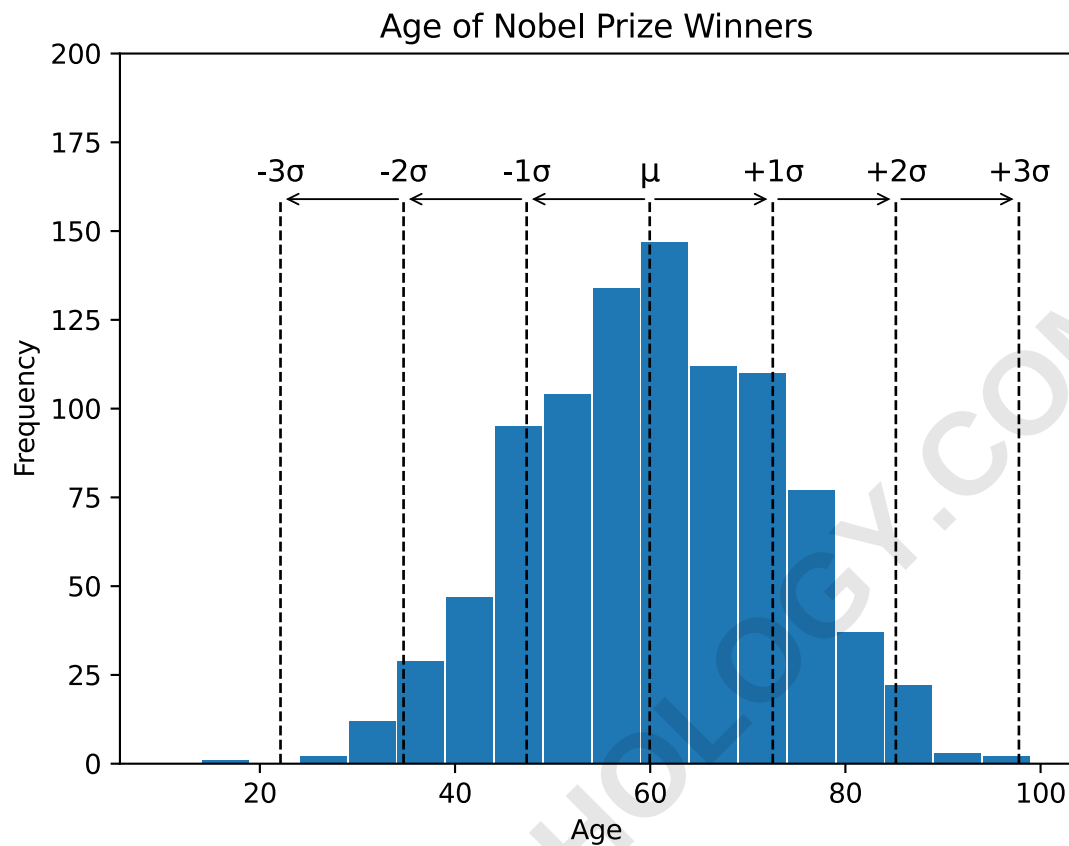
Standard Deviation

Standard deviation is the most used measure of variation.

Standard deviation (σ) measures how far a 'typical' observation is from the average of the data (μ).

Standard deviation is important for many statistical methods.

Here is a histogram of the age of all 934 Nobel Prize winners up to the year 2020, showing **standard deviations**:



Note: Values within one standard deviation (σ) are considered to be typical.

Values outside three standard deviations are considered to be **outliers**.

★+1 W3schools PathfinderTrack your progress - it's free!

Log in

Sign Up