# Calculating Cohen’s d Effect Size with Excel: A Step-by-Step Guide

Authored by

**Mohammed looti**

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Calculating Cohen’s d Effect Size with Excel: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from https://statistics.arabpsychology.com/?p=13482

In the field of statistics, researchers frequently seek to determine if a meaningful difference exists between two distinct groups. This determination is traditionally initiated through a formal hypothesis test, the primary output of which is the calculation of a p-value. The p-value serves as the gatekeeper for drawing conclusions about population differences, guiding the rejection or retention of the null hypothesis based on a predetermined significance level.

When the p-value falls below the chosen significance threshold (such as $\alpha$ = 0.05), we conclude that the observed difference is **statistically significant**. This finding confirms that the difference is unlikely to be due to mere random chance. However, relying solely on significance can be misleading; a statistically significant result might represent a minuscule, practically irrelevant distinction, especially in studies involving very large sample sizes. Modern statistical reporting demands a metric that quantifies the magnitude of the difference, moving beyond the simple binary question of "does a difference exist?"

To address this critical need for measuring practical importance, researchers utilize measures of effect size. Effect size provides a standardized, interpretable scale for quantifying the strength of the relationship between variables or the magnitude of the difference between group means. This standardization ensures that findings are comparable across diverse research contexts and measurement instruments.

One of the most robust and widely adopted metrics for quantifying the difference between two means is **Cohen's d**. This standardized mean difference expresses the raw difference between groups in units of standard deviation, making it universally interpretable. Understanding its calculation is essential for fully grasping the impact of research findings, and this comprehensive tutorial details exactly how to calculate this measure efficiently using Microsoft Excel.

## Defining Cohen's d: Formula and Components

Cohen's d provides a standardized value that facilitates the comparison of mean differences across studies that may use wildly different scales (e.g., comparing a study measuring weight in kilograms to one measuring anxiety on a 100-point scale). It transforms the raw difference into a standard deviation metric. The fundamental formula for calculating Cohen's d for two independent groups is defined as:

Cohen's d = ($\overline{x}_{1}$ - $\overline{x}_{2}$) / pooled $SD$

To properly execute this calculation, it is necessary to clearly define each of the components involved:

$\overline{x}_{1}$ = The **arithmetic mean** of the first group of observations.
$\overline{x}_{2}$ = The **arithmetic mean** of the second group of observations.

pooled $SD$ = The pooled standard deviation, which represents a weighted average of the variability within both groups. In the simplified calculation often used for Cohen's d (assuming equal sample sizes), this is calculated using the formula: $sqrt{(s_{1}^{2} + s_{2}^{2}) / 2}$.

The pooled standard deviation in the denominator is critical because it standardizes the mean difference, effectively removing the influence of the original measurement scale. This procedure allows the final result, $d$, to be interpreted as the number of standard deviations by which the two group means differ.

## Theoretical Background: The Importance of Standardized Mean Difference

While traditional inferential statistics often focus narrowly on rejecting the null hypothesis, contemporary research standards emphasize the vital role of **effect size reporting**. As established, a highly significant p-value can be achieved with trivial real-world differences if the study employs a massive sample size. Conversely, a small study might fail to reach statistical significance despite exhibiting a large, practically important effect. Cohen's d effectively bridges this gap by providing a measure of magnitude that is statistically independent of the sample size ($n$).

Cohen's d is categorized as a metric of standardized mean difference, meaning the calculation scales the raw difference between the means by dividing it by the groups' shared variability (the pooled standard deviation). This crucial standardization allows researchers to easily interpret the effect, regardless of the original units of measurement. The resulting standardized score is the key feature that makes **Cohen's d** invaluable for conducting meta-analysis, where results from multiple studies using differing instruments must be integrated and compared on a common scale.

For the calculation of Cohen's d to be accurate, we generally assume that the variability (specifically the variance and standard deviation) across the two groups is relatively equivalent. This is known as the assumption of homogeneity of variances. This assumption justifies the use of the pooled standard deviation in the denominator, which provides a stable estimate of the population standard deviation. Should the variances be substantially different, researchers may opt for an alternative effect size measure, such as Hedges' $g$, which applies a correction for bias, although Cohen's $d$ remains the most frequently reported effect size for two-group comparisons.

## Practical Example: Structuring Data in Excel

To illustrate the calculation of Cohen's d, we will utilize a hypothetical dataset comparing the effectiveness of two distinct training methods, Group 1 and Group 2, based on their performance scores. The calculation process requires three core summary statistics for each group: the mean ($overline{x}$), the standard deviation ($s$), and the sample size ($n$).

The initial and most critical step in Microsoft Excel is establishing a clear and organized layout for these summary statistics. Proper organization minimizes the risk of calculation errors by ensuring that all subsequent formulas reference the correct cells. We designate specific rows to store the mean, standard deviation, and sample size for both groups, placing Group 1's data in column B and Group 2's data in column C.

**Step 1: Enter the Required Data.** Ensure that the necessary summary information is readily available. If you begin with raw data (individual scores), you must first use Excel functions like `AVERAGE()` and `STDEV.S()` to derive the mean and standard deviation, respectively. For the purpose of this tutorial, we assume the summary statistics are already calculated and are entered into cells B2 through C4, as displayed in the image below.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | **Group 1** | **Group 2** | | |
| 2 | **Mean** | 15.2 | 14 | | |
| 3 | **Std. Dev.** | 4.4 | 3.6 | | |
| 4 | **n** | 39 | 34 | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |

## Calculating the Numerator: Difference in Means

The numerator of the Cohen's d equation is the simplest element to calculate: the raw difference between the two group means ($overline{x}_{1}$ - $overline{x}_{2}$). This result quantifies the distance separating the two groups in the original units of measurement before any standardization occurs. This raw difference sets the stage for the standardization process that follows.

**Step 2: Calculate the difference in means.** In Excel, this is achieved through a straightforward subtraction operation. Assuming the mean of Group 1 is located in cell B2 and the mean of Group 2 is in cell C2, the appropriate formula to enter into a new, designated cell (e.g., E2) would be: `=B2-C2`. This result is the direct measure of the effect we are standardizing.

The sign of the result indicates direction: a positive result means Group 1 has the higher mean, while a negative result means Group 2 is higher. While the sign is important for directional claims, the interpretation of the magnitude of **Cohen's d** typically relies on its absolute value, showing only how far apart the groups are, irrespective of which group is superior.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | **Group 1** | **Group 2** | | | | |
| 2 | **Mean** | 15.2 | 14 | | **Difference in means** | 1.2 | =B2-C2 |
| 3 | **Std. Dev.** | 4.4 | 3.6 | | | | |
| 4 | **n** | 39 | 34 | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |
| 9 | | | | | | | |
| 10 | | | | | | | |
| 11 | | | | | | | |
| 12 | | | | | | | |
| 13 | | | | | | | |
| 14 | | | | | | | |
| 15 | | | | | | | |

## Calculating the Denominator: Pooled Standard Deviation

The denominator, the **pooled standard deviation ($SD_{pooled}$)**, is typically the most intricate part of the calculation. It serves as a combined, weighted estimate of the variability common to both groups, under the assumption that they share an underlying population standard deviation. For the simplified version of Cohen's $d$, where sample sizes are assumed equal, the pooled $SD$ is calculated as the square root of the average of the squared standard deviations (variances).

**Step 3: Calculate the pooled standard deviation.** This step requires carefully managing the sequence of operations within Excel using nested functions. We must first square the standard deviations of both groups, sum those squared values, divide the sum by two, and then take the square root of the entire result to return to the standard deviation metric. If the standard deviation for $s_1$ is in cell B3 and $s_2$ is in cell C3, the complete formula to calculate the pooled standard deviation (using the `SQRT` function) would be entered into a new cell (e.g., E5):

```
=SQRT((B3^2 + C3^2)/2)
```

It is essential to verify the correct use of parentheses to force Excel to perform the operations in the correct statistical order: squaring first, then averaging the variances, and finally taking the square root. This resulting value, the pooled $SD$, is the crucial standardized unit against which the mean difference will be measured.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | **Group 1** | **Group 2** | | | | | |
| 2 | **Mean** | 15.2 | 14 | | **Difference in means** | 1.2 | | |
| 3 | **Std. Dev.** | 4.4 | 3.6 | | **Pooled standard deviation** | 4.01995 | =SQRT((B3^2+C3^2)/2) | |
| 4 | **n** | 39 | 34 | | | | | |
| 5 | | | | | | | | |
| 6 | | | | | | | | |
| 7 | | | | | | | | |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

## Finalizing the Calculation and Interpreting the Result

With the calculation of the mean difference (numerator, calculated in Step 2, cell E2) and the pooled standard deviation (denominator, calculated in Step 3, cell E5) complete, the final step to obtain Cohen's d is a simple division operation.

**Step 4: Calculate Cohen's d.** In a final cell (e.g., E6), divide the mean difference by the pooled standard deviation. The precise Excel formula is simply: `=E2/E5`. This delivers the final, standardized effect size measure.

For our specific example, this calculation yields the result of **0.29851**. This signifies that the mean performance score of Group 1 is approximately 0.3 standard deviations higher than the mean performance score of Group 2. This standardized value is now ready for contextual interpretation.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | **Group 1** | **Group 2** | | | | |
| 2 | **Mean** | 15.2 | 14 | | **Difference in means** | 1.2 | |
| 3 | **Std. Dev.** | 4.4 | 3.6 | | **Pooled standard deviation** | 4.01995 | |
| 4 | **n** | 39 | 34 | | **Cohen's D** | 0.29851 | =F2/F3 |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |
| 9 | | | | | | | |
| 10 | | | | | | | |
| 11 | | | | | | | |
| 12 | | | | | | | |
| 13 | | | | | | | |
| 14 | | | | | | | |
| 15 | | | | | | | |
| 16 | | | | | | | |
| 17 | | | | | | | |
| 18 | | | | | | | |
| 19 | | | | | | | |
| 20 | | | | | | | |

## Interpreting the Magnitude: Cohen's Benchmarks

Once the numerical value for **Cohen's d** is successfully obtained, the crucial next step is to interpret its practical magnitude. Jacob Cohen established widely accepted guidelines, often termed "rules of thumb," to help categorize the size of the effect. While these benchmarks are advisory and not absolute rules, they offer a common framework for communicating the real-world relevance of findings to both peers and non-specialist audiences.

The traditional framework for interpreting the absolute value of Cohen's d uses the following categories:

**$d$ = 0.2** = Categorized as a **Small effect size**. This indicates a minor or subtle difference between the two group means, often difficult to observe without statistical analysis.
**$d$ = 0.5** = Categorized as a **Medium effect size**. This difference is noticeable and represents a moderate level of practical impact.
**$d$ = 0.8** = Categorized as a **Large effect size**. This signifies a substantial and clearly important difference between the group means, often visible to the naked eye.

In the context of our calculated result of **0.29851**, the effect size falls between the small (0.2) and medium (0.5) thresholds, but is generally interpreted as closer to a **small effect size**. This nuanced outcome has profound implications for how the study's results should be discussed. It reveals that although a traditional p-value test might indicate a statistically significant difference between the two group means, the actual difference in practical terms (the effect size) is relatively minor or weak.

Ultimately, the interpretation of Cohen's d must always be contextualized within the specific field of study. While Cohen's benchmarks provide helpful starting points, what constitutes a "large" or "small" effect can vary significantly depending on the costs, risks, and domain of the research. For example, in public health, an effect size considered "small" might be deemed highly important if the intervention is inexpensive and scalable to millions of people. Therefore, reporting both the p-value and the effect size (Cohen's d) is crucial for providing the most complete and honest picture of the research findings.