# How to Make a Q-Q Plot in Excel: A Step-by-Step Guide

Authored by

**stats writer**

December 28, 2025

The process of generating a Q-Q plot in Microsoft Excel is a critical skill for performing preliminary statistical analysis, particularly when testing for normal distribution. This visualization technique compares the distribution of your sample data against a theoretical distribution, typically the standard normal distribution.

While some specialized statistical software can generate these plots instantly, mastering the manual calculation steps within Excel provides a deeper understanding of how quantiles are derived. This comprehensive guide details the precise, multi-step process required to correctly calculate the necessary values and plot the empirical quantiles against the theoretical quantiles, culminating in a valid Q-Q plot.

## Introduction to Q-Q Plots and Their Purpose

A **Q-Q plot**, which is an abbreviation for "quantile-quantile" plot, serves as an essential graphical tool in statistics. Its primary function is to determine if a given set of data plausibly originated from a specific theoretical probability distribution. In academic research and applied statistics, the most frequent application of the Q-Q plot is to visually evaluate whether or not a data set follows a normal distribution.

Understanding the underlying distribution of data is paramount because many inferential statistical tests (such as t-tests or ANOVA) assume that the data are normally distributed. If this assumption is violated, the results of those tests may be unreliable. The Q-Q plot offers a highly intuitive, non-formal assessment of this distributional assumption.

This detailed tutorial provides a step-by-step walkthrough on how to generate a robust Q-Q plot for any data set using the calculation capabilities available in Excel.

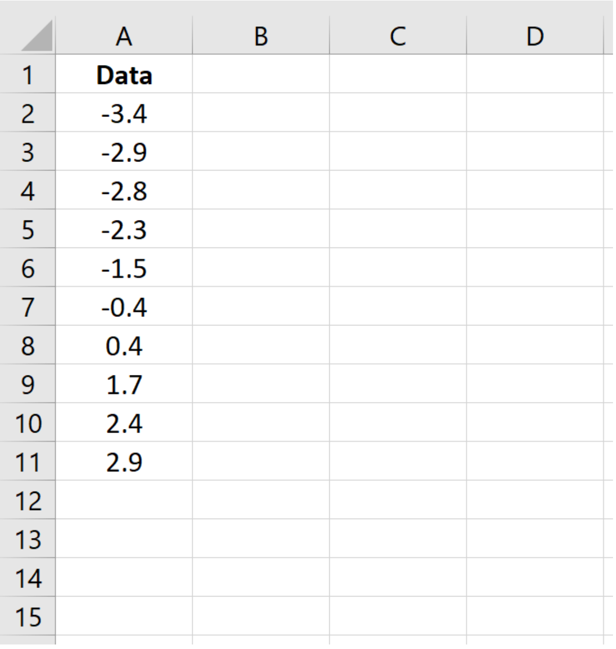
## Example: Calculating the Theoretical Quantiles in Excel

To construct a statistically accurate Q-Q plot, we must first calculate three derived variables for each observation in the data set: the rank, the empirical percentile, and the theoretical z-score (quantile). Follow these precise steps to prepare your data for plotting.

## Step 1: Data Preparation and Sorting the Observations

The initial requirement for creating a Q-Q plot is to organize your raw data. Each data point must be processed in ascending order to correctly calculate its corresponding rank and percentile.

Begin by entering your data into a single column in Excel. For this example, assume we have the following 10 data points entered into Column A:



| | A | B | C | D |
|---|---|---|---|---|
| 1 | **Data** | | | |
| 2 | -3.4 | | | |
| 3 | -2.9 | | | |
| 4 | -2.8 | | | |
| 5 | -2.3 | | | |
| 6 | -1.5 | | | |
| 7 | -0.4 | | | |
| 8 | 0.4 | | | |
| 9 | 1.7 | | | |
| 10 | 2.4 | | | |
| 11 | 2.9 | | | |
| 12 | | | | |
| 13 | | | | |
| 14 | | | | |
| 15 | | | | |

It is absolutely crucial that this data is sorted from the smallest value to the largest value. If your raw data is not yet sorted, navigate to the **Data** tab located on the top ribbon in Excel. Then, within the **Sort & Filter** group, click the **Sort A to Z** icon to arrange the values in ascending order. This ensures that the subsequent rank calculations are accurate.

## Step 2: Determining the Rank of Each Data Value

Once the data is sorted, the next step is to calculate the rank of each individual data point. The rank simply identifies the ordinal position of the observation within the sorted list.

In the adjacent column (Column B), use the following Excel formula to calculate the rank of the first data value (assuming the first value is in cell A2):

**=RANK(A2, $A$2:$A$11, 1)**

In this formula, `A2` is the value being ranked, `$A$2:$A$11` is the reference range (fixed using dollar signs to prevent shifting when copied), and the `1` specifies that the ranking should be performed in ascending order (from smallest to largest).

B2 =RANK(A2, $A$2:$A$11, 1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Data | Rank | | | | |
| 2 | -3.4 | 1 | | | | |
| 3 | -2.9 | | | | | |
| 4 | -2.8 | | | | | |
| 5 | -2.3 | | | | | |
| 6 | -1.5 | | | | | |
| 7 | -0.4 | | | | | |
| 8 | 0.4 | | | | | |
| 9 | 1.7 | | | | | |
| 10 | 2.4 | | | | | |
| 11 | 2.9 | | | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |

Apply this formula by copying it down to all the remaining cells in Column B. You should see ranks ranging from 1 to the total number of observations (N).

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Data | Rank | | | |
| 2 | -3.4 | 1 | | | |
| 3 | -2.9 | 2 | | | |
| 4 | -2.8 | 3 | | | |
| 5 | -2.3 | 4 | | | |
| 6 | -1.5 | 5 | | | |
| 7 | -0.4 | 6 | | | |
| 8 | 0.4 | 7 | | | |
| 9 | 1.7 | 8 | | | |
| 10 | 2.4 | 9 | | | |
| 11 | 2.9 | 10 | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |



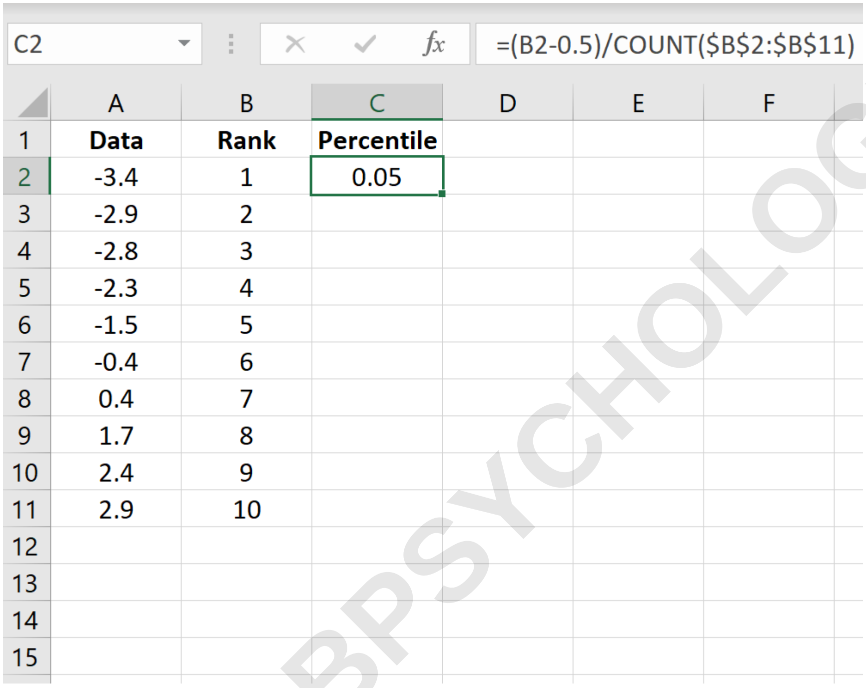
## Step 3: Calculating Empirical Percentiles (Plotting Positions)

The empirical percentile, also known as the plotting position, represents the estimated cumulative probability associated with each data point. This value dictates where the data point sits relative to the entire distribution. We use a standardized formula to calculate this position to account for the discrete nature of the sample data.

Next, use the following formula in Column C to calculate the empirical percentile of the first value (using the rank located in B2):

**=(B2-0.5)/COUNT($B$2:$B$11)**

This formula employs the common adjustment `(Rank - 0.5) / N`, where N is the total count of observations. The `-0.5` correction is applied to ensure that the empirical quantiles do not include 0 or 1, which are problematic for the subsequent inverse normal calculation. The `COUNT($B$2:$B$11)` function efficiently determines the sample size (N).


C2 =(B2-0.5)/COUNT($B$2:$B$11)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Data | Rank | Percentile | | | |
| 2 | -3.4 | 1 | 0.05 | | | |
| 3 | -2.9 | 2 | | | | |
| 4 | -2.8 | 3 | | | | |
| 5 | -2.3 | 4 | | | | |
| 6 | -1.5 | 5 | | | | |
| 7 | -0.4 | 6 | | | | |
| 8 | 0.4 | 7 | | | | |
| 9 | 1.7 | 8 | | | | |
| 10 | 2.4 | 9 | | | | |
| 11 | 2.9 | 10 | | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |

Ensure the range for the `COUNT` function is fixed using dollar signs. Copy this formula down to all of the other cells in the column to generate the empirical percentiles for every observation.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | **Data** | **Rank** | **Percentile** | | |
| 2 | -3.4 | 1 | 0.05 | | |
| 3 | -2.9 | 2 | 0.15 | | |
| 4 | -2.8 | 3 | 0.25 | | |
| 5 | -2.3 | 4 | 0.35 | | |
| 6 | -1.5 | 5 | 0.45 | | |
| 7 | -0.4 | 6 | 0.55 | | |
| 8 | 0.4 | 7 | 0.65 | | |
| 9 | 1.7 | 8 | 0.75 | | |
| 10 | 2.4 | 9 | 0.85 | | |
| 11 | 2.9 | 10 | 0.95 | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |

## Step 4: Calculating Theoretical Quantiles (Standard Normal Z-Scores)

The final critical calculation involves determining the theoretical quantiles, which are the standard z-scores corresponding to the calculated empirical percentiles. If the data were perfectly normally distributed, the observed value would align with this theoretical z-score.

Use the following formula in Column D to calculate the theoretical z-score for the first data value (using the percentile from C2):

**=NORM.S.INV(C2)**

The `NORM.S.INV` function in Excel takes a probability (our calculated percentile) and returns the corresponding quantile (or z-score) from the standard normal distribution. This column of values represents the theoretical distribution against which your data will be plotted.

D2 =NORM.S.INV(C2)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Data | Rank | Percentile | Z-Score | | |
| 2 | -3.4 | 1 | 0.05 | -1.64485 | | |
| 3 | -2.9 | 2 | 0.15 | | | |
| 4 | -2.8 | 3 | 0.25 | | | |
| 5 | -2.3 | 4 | 0.35 | | | |
| 6 | -1.5 | 5 | 0.45 | | | |
| 7 | -0.4 | 6 | 0.55 | | | |
| 8 | 0.4 | 7 | 0.65 | | | |
| 9 | 1.7 | 8 | 0.75 | | | |
| 10 | 2.4 | 9 | 0.85 | | | |
| 11 | 2.9 | 10 | 0.95 | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |

Copy this formula down to all of the other cells in the column D. You now have the two essential variables required for plotting: the theoretical quantiles (X-axis) and the observed data values (Y-axis).

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Data | Rank | Percentile | Z-Score | |
| 2 | -3.4 | 1 | 0.05 | -1.64485 | |
| 3 | -2.9 | 2 | 0.15 | -1.03643 | |
| 4 | -2.8 | 3 | 0.25 | -0.67449 | |
| 5 | -2.3 | 4 | 0.35 | -0.38532 | |
| 6 | -1.5 | 5 | 0.45 | -0.12566 | |
| 7 | -0.4 | 6 | 0.55 | 0.125661 | |
| 8 | 0.4 | 7 | 0.65 | 0.38532 | |
| 9 | 1.7 | 8 | 0.75 | 0.67449 | |
| 10 | 2.4 | 9 | 0.85 | 1.036433 | |
| 11 | 2.9 | 10 | 0.95 | 1.644854 | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |

## Step 5: Generating the Q-Q Scatter Plot

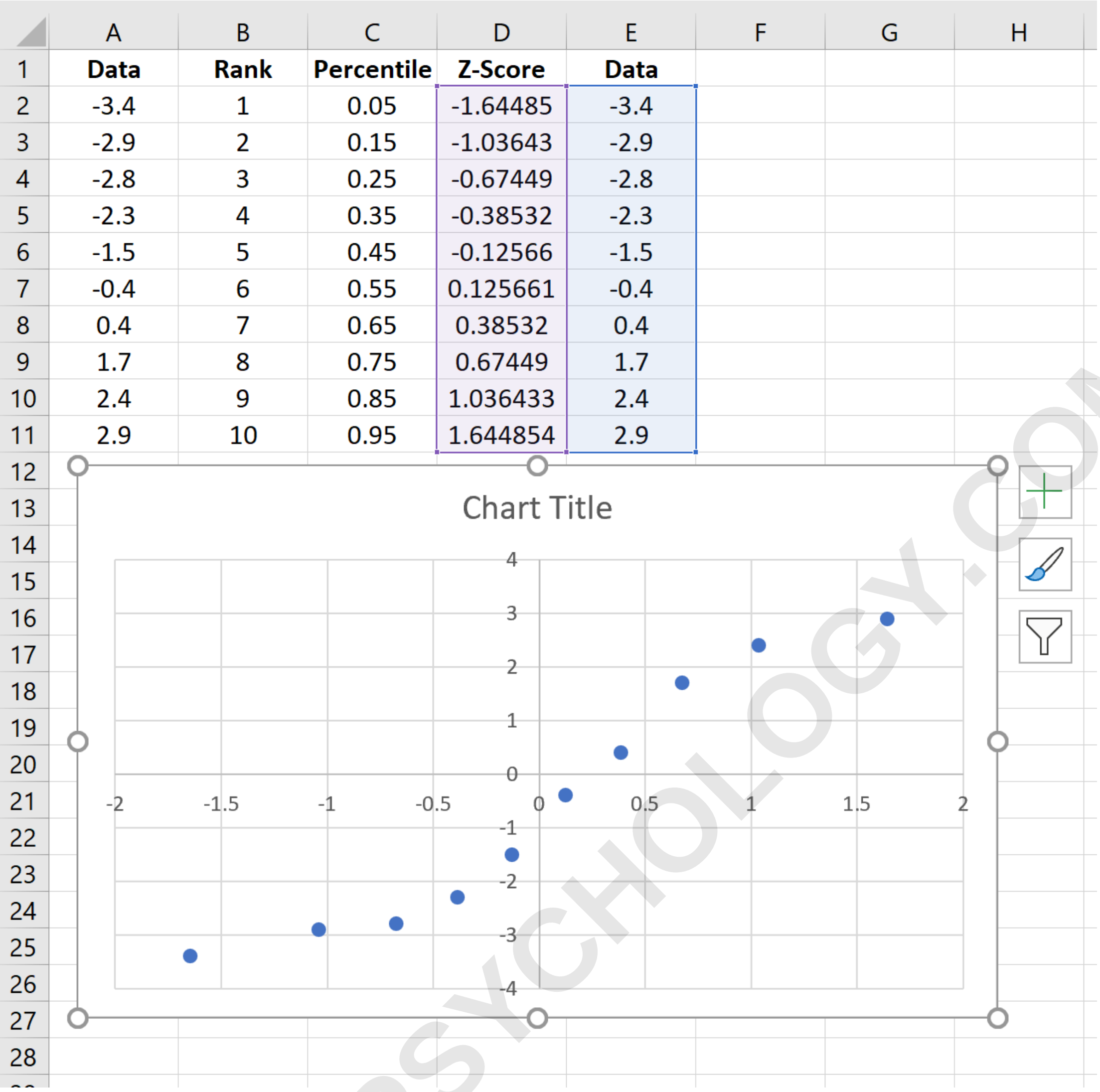
With the theoretical quantiles calculated, we can proceed to visualize the relationship between the sample data and the standard normal distribution.

First, if necessary, copy the original data from column A into a new column, such as column E, to keep the plotting columns adjacent. Then, highlight the data in column D (Theoretical Quantiles/Z-scores) and column E (Observed Data Values).

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | **Data** | **Rank** | **Percentile** | **Z-Score** | **Data** | |
| 2 | -3.4 | 1 | 0.05 | -1.64485 | -3.4 | |
| 3 | -2.9 | 2 | 0.15 | -1.03643 | -2.9 | |
| 4 | -2.8 | 3 | 0.25 | -0.67449 | -2.8 | |
| 5 | -2.3 | 4 | 0.35 | -0.38532 | -2.3 | |
| 6 | -1.5 | 5 | 0.45 | -0.12566 | -1.5 | |
| 7 | -0.4 | 6 | 0.55 | 0.125661 | -0.4 | |
| 8 | 0.4 | 7 | 0.65 | 0.38532 | 0.4 | |
| 9 | 1.7 | 8 | 0.75 | 0.67449 | 1.7 | |
| 10 | 2.4 | 9 | 0.85 | 1.036433 | 2.4 | |
| 11 | 2.9 | 10 | 0.95 | 1.644854 | 2.9 | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |



Navigate to the **Insert** tab on the top ribbon. Within the **Charts** group, select **Insert Scatter (X, Y)** and choose the option labeled **Scatter** (points only). This action will generate the fundamental Q-Q plot visualization:



| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Data | Rank | Percentile | Z-Score | Data |
| 2 | -3.4 | 1 | 0.05 | -1.64485 | -3.4 |
| 3 | -2.9 | 2 | 0.15 | -1.03643 | -2.9 |
| 4 | -2.8 | 3 | 0.25 | -0.67449 | -2.8 |
| 5 | -2.3 | 4 | 0.35 | -0.38532 | -2.3 |
| 6 | -1.5 | 5 | 0.45 | -0.12566 | -1.5 |
| 7 | -0.4 | 6 | 0.55 | 0.125661 | -0.4 |
| 8 | 0.4 | 7 | 0.65 | 0.38532 | 0.4 |
| 9 | 1.7 | 8 | 0.75 | 0.67449 | 1.7 |
| 10 | 2.4 | 9 | 0.85 | 1.036433 | 2.4 |
| 11 | 2.9 | 10 | 0.95 | 1.644854 | 2.9 |

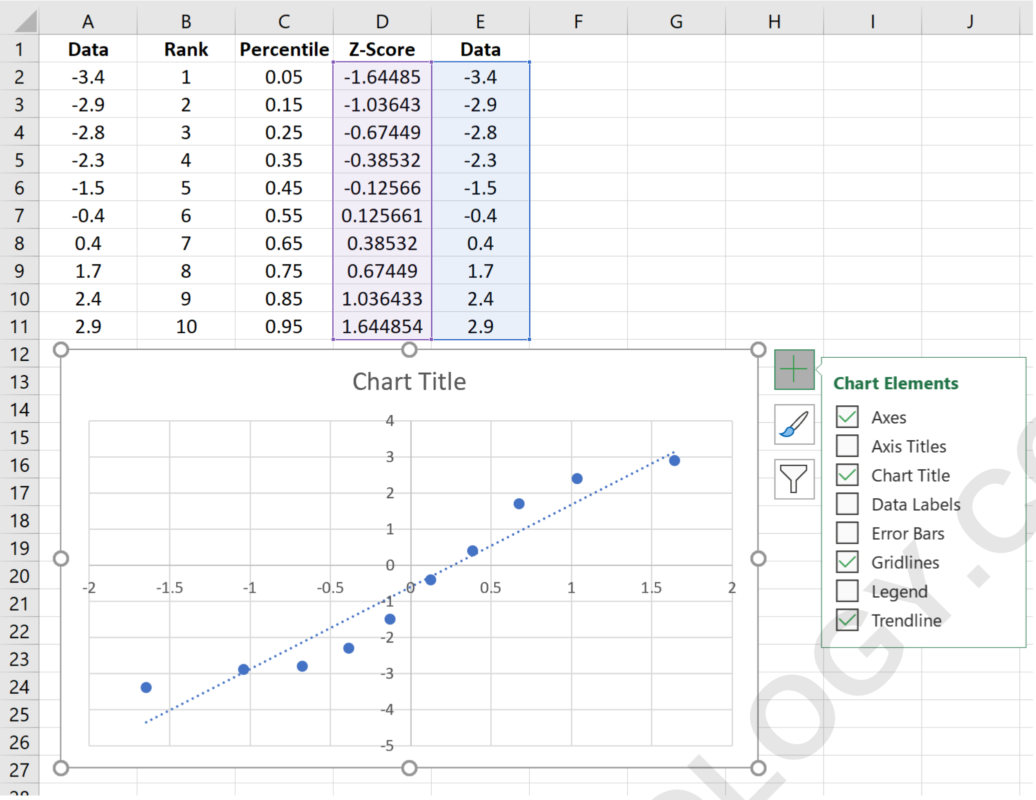




## Step 6: Adding the Reference Trendline

For proper visual assessment, a reference line must be added to the chart. This line, often a Trendline, represents the expected path if the data were perfectly normally distributed.

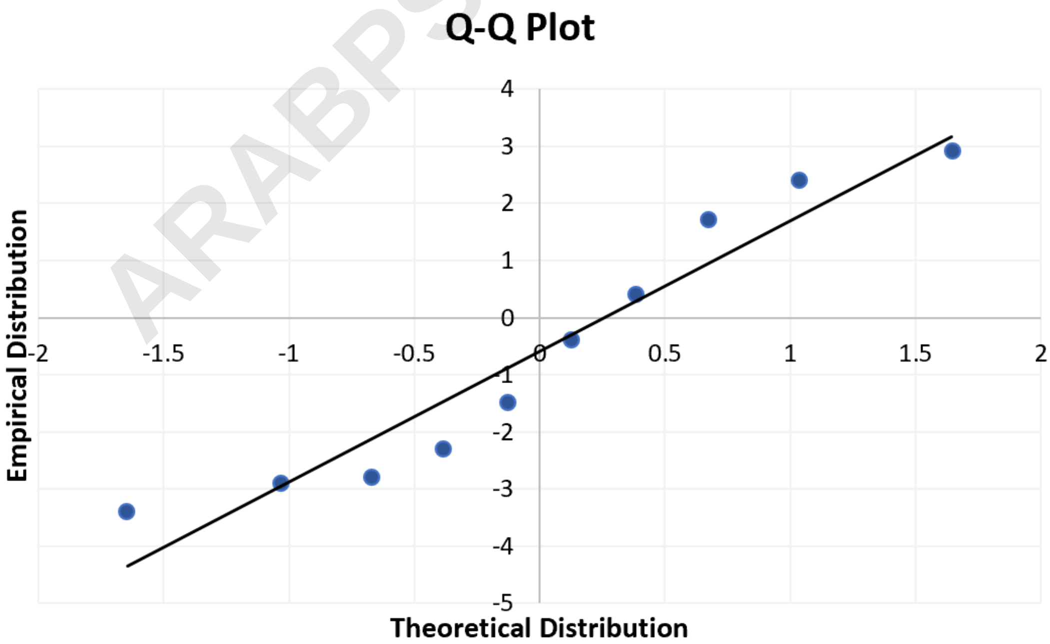
To add this crucial reference, click the plus sign (+) icon located on the top right-hand corner of the generated graph. Check the box next to **Trendline**. This will automatically fit a linear line to the data points, which serves as our visual benchmark.



| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Data | Rank | Percentile | Z-Score | Data |
| 2 | -3.4 | 1 | 0.05 | -1.64485 | -3.4 |
| 3 | -2.9 | 2 | 0.15 | -1.03643 | -2.9 |
| 4 | -2.8 | 3 | 0.25 | -0.67449 | -2.8 |
| 5 | -2.3 | 4 | 0.35 | -0.38532 | -2.3 |
| 6 | -1.5 | 5 | 0.45 | -0.12566 | -1.5 |
| 7 | -0.4 | 6 | 0.55 | 0.125661 | -0.4 |
| 8 | 0.4 | 7 | 0.65 | 0.38532 | 0.4 |
| 9 | 1.7 | 8 | 0.75 | 0.67449 | 1.7 |
| 10 | 2.4 | 9 | 0.85 | 1.036433 | 2.4 |
| 11 | 2.9 | 10 | 0.95 | 1.644854 | 2.9 |



For professional presentation, enhance the graph by adding clear titles and axis labels. The X-axis should be labeled "Theoretical Quantiles" (or "Z-scores") and the Y-axis should be labeled "Observed Data Values."



## Interpreting the Q-Q Plot for Normality

The interpretation of the Q-Q plot is straightforward and highly visual: if the data values fall closely along the straight reference line (which typically forms a roughly 45-degree angle), then the data can be reasonably considered normally distributed.

Deviations from this straight line indicate departures from normality. Analyzing the pattern of deviation can offer clues about the nature of the non-normality:

**Curvature:** If the points form an S-shape, the distribution may have heavier or lighter tails than the normal distribution.
**Deviation at the Ends:** If the central points align but the points at the extreme ends (the tails) stray away from the line, this suggests the presence of outliers or issues with skewness or kurtosis.

In the example Q-Q plot shown above, we can clearly observe that the data values tend to deviate significantly from the straight reference line, particularly at both the lower and upper tail ends. This visual evidence strongly suggests that this specific data set is **not normally distributed**. Although a Q-Q plot does not constitute a formal statistical hypothesis test, it provides an invaluable and easy-to-use graphical confirmation of distributional assumptions before proceeding with parametric statistical modeling.