**Univariate vs Multivariate – Normality Check.**

**Using the Q-Q Plot (Graphical Method)**

**For Univariate Data**

Steps includes,

1. Arrange the values in ascending order (Data > Sort)
2. Compute the ranks of the values in the dataset (using RANK.EQ function)

Manually change the ranks of the set of values having the same rank.

1. Calculate the cumulative probability of the ranks of the dataset, formula below.

Assigned or Fitted Cumulative Probability = ((i – 0.5)/n)

Where n is the size (count) of the dataset, and i is the rank of data value.

1. Calculate the normal predicted cumulative probability for the dataset based on empirical or theoretical normal distribution density function (using NORM.DIST function).
2. Finally, plot a scatter plot with the Predicted Normal Cumulative Probability on the x-axis and the Assigned Cumulative Probability on the y-axis.
3. If the resulting graph is approximately a straight line, the dataset can be assumed to be normally distributed; otherwise, the dataset is not normally distributed.
4. One can also plot quantile vs sorted data set and the resulting graph is expected to be a straight line, if the dataset can be assumed to be normally distributed.

Reference

[Test for Normality (Normal Dist.) – Excel and Google Sheets - Automate Excel](https://www.automateexcel.com/stats/test-for-normality-normal-dist/)

[Normality Test Using Microsoft Excel : Intact Prolink Blog (inprolink.com)](https://www.inprolink.com/blog/normality-test-using-microsoft-excel/) based on chi-square test

[Empirical Distribution Function / Empirical CDF - Statistics How To](https://www.statisticshowto.com/empirical-distribution-function/)

[Excel NORM.DIST function | Exceljet](https://exceljet.net/functions/norm.dist-function)

**Multivariate Data**

Steps includes,

1. Calculate Squared Mahalanobis distance (T-square) for each observation in dataset.
2. Sort the Squared Mahalanobis distance (T-square) in ascending order.
3. Compute the ranks of the values in the dataset (using RANK.EQ function)
4. Calculate the cumulative probability of the ranks of the dataset (Assigned based on rank).
5. Calculate the chi-square predicted cumulative probability (as T-square is expected to follow Chi-square distribution).
6. Finally, plot a scatter plot with the Predicted Cumulative Probability on the x-axis and the Assigned Cumulative Probability on the y-axis.
7. If the resulting graph is approximately a straight line, the dataset can be assumed to be Multivariate normal distributed; otherwise, the dataset is not normally distributed.

Reference

[BIVARIATE QQ-PLOTS AND SPIDER WEB PLOTS on JSTOR](https://www.jstor.org/stable/24306465#:~:text=qq%2Dplot%20then%20draws%20an,of%20the%20spherical%20normal%20distribution.)

[4.4 - Multivariate Normality and Outliers | STAT 505 (psu.edu)](https://online.stat.psu.edu/stat505/lesson/4/4.4)