Hypothesis Testing in Healthcare: Drug Safety

A pharmaceutical company GlobalXYZ has just completed a randomized controlled drug trial. To promote transparency and reproducibility of the drug's outcome, they (GlobalXYZ) have presented the dataset to your organization, a non-profit that focuses primarily on drug safety.

The dataset provided contained five adverse effects, demographic data, vital signs, etc. Your organization is primarily interested in the drug's adverse reactions. It wants to know if the adverse reactions, if any, are of significant proportions. It has asked you to explore and answer some questions from the data.

The dataset drug_safety.csv was obtained from **Hbiostat** courtesy of the Vanderbilt University Department of Biostatistics. It contained five adverse effects: headache, abdominal pain, dyspepsia, upper respiratory infection, chronic obstructive airway disease (COAD), demographic data, vital signs, lab measures, etc. The ratio of drug observations to placebo observations is 2 to 1.

For this project, the dataset has been modified to reflect the presence and absence of adverse effects adverse_effects and the number of adverse effects in a single individual num_effects.

The columns in the modified dataset are:

Column	Description
sex	The gender of the individual
age	The age of the individual
week	The week of the drug testing
trx	The treatment (Drug) and control (Placebo) groups
wbc	The count of white blood cells
rbc	The count of red blood cells
adverse_effects	The presence of at least a single adverse effect
num_effects	The number of adverse effects experienced by a single individual

Question 1:

<u>Determine if the proportion of adverse effects differs significantly between the Drug and Placebo groups.</u>

Null Hypothesis (H0): The proportion of adverse effects is the same in the Drug and Placebo groups.

Alternative Hypothesis (H1): The proportion of adverse effects is different between the Drug and Placebo groups.

Chi-square test is a statistical test used to compare observed results with expected results. The purpose of this test is to determine if a difference between observed data and expected data is due to chance, or if it is due to a relationship between the variables you are studying.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
trx_group * adverse_effects_group	16103	100.0%	0	0.0%	16103	100.0%

trx_group * adverse_effects_group Crosstabulation

Count

		adverse_effects_group		
		No	Yes	Total
trx_group	Placebo	4864	512	5376
	Drug	9703	1024	10727
Total		14567	1536	16103

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.002[a]	1	.964		
Continuity Correction[b]	.000	1	.987		
Likelihood Ratio	.002	1	.964		
Fisher's Exact Test				.976	.493
Linear-by-Linear Association	.002	1	.964		
N of Valid Cases	16103				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 512.79.

b. Computed only for a 2x2 table

Analysis Summary

p-value = 0.964; significance level = 0.05

A chi-square test was used to compare the rates of side effects between two groups: the Drug group and the Placebo group. The test gave a p-value of 0.964, while the chosen significance level was 0.05.

Since the p-value (0.964) is higher than the significance level (0.05), the decision is to **not reject the null hypothesis**. This result suggests that there is no meaningful difference in the rates of side effects between the Drug and Placebo groups.

What it means: With a p-value of 0.964, the results show there is no significant difference in the rates of side effects between the Drug and Placebo groups. The small difference we see in the data is likely just by chance, rather than showing a real difference between the groups.

In simple terms, the chi-square test did not find enough evidence to conclude that the rates of side effects are different between the two groups. The null hypothesis, which assumes no difference in the rates, cannot be rejected based on the available data.

Question 2:

<u>Find out if the number of adverse effects is independent of the treatment and control groups</u>

Null Hypothesis (H0): There is no relationship between the number of adverse effects and treatment and control groups

Alternative Hypothesis (H1): There is a relationship between the two variables.

The Chi-Square Test of Independence determines whether there is an association between categorical variables (i.e., whether the variables are independent or related). It is a nonparametric test. This test is also known as: **Chi-Square Test of Association**.

num_effects * trx_group Crosstabulation

% within num_effects

		trx_group		
		Placebo	Drug	Total
num_effects	0	33.4%	66.6%	100.0%
	1	33.7%	66.3%	100.0%
	2	28.4%	71.6%	100.0%
	3	16.7%	83.3%	100.0%
Total		33.4%	66.6%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.800[a]	3	.615
Likelihood Ratio	1.922	3	.589
Linear-by-Linear Association	.138	1	.710
N of Valid Cases	16103		

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 2.00.

I used a chi-square test to check if the number of adverse effects was related to the treatment group or the control group. The test gave a **p-value of 0.615**, and the **significance level was set at 0.05**.

Since the p-value (0.615) is higher than the significance level (0.05), we cannot reject the null hypothesis. This means the data does not show a clear connection between the number of adverse effects and whether someone was in the treatment group or the control group.

In simple terms, based on this test, **the adverse effects seem to happen regardless of which group someone was in - the treatment group or the control group.**

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
trx_group * num_effects	Between Groups	(Combined)	.400	3	.133	.600	.615
		Linearity	.031	1	.031	.138	.710
		Deviation from Linearity	.370	2	.185	.831	.436
	Within Groups		3580.818	16099	.222		
	Total		3581.218	16102			

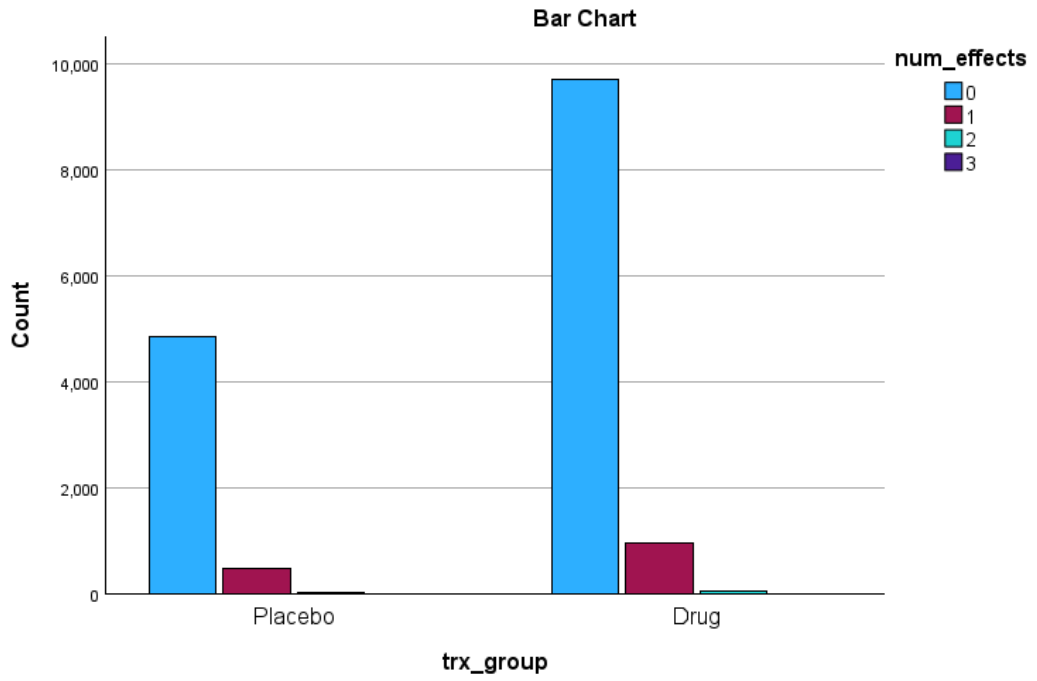
Deviation from linearity test: This test is part of the **ANOVA procedure** and specifically examines the linearity assumption. In the current case, **p = 0.436 > 0.05**, it means there is not enough evidence to reject the null hypothesis of linearity. Therefore, the **relationship between the variables trx_group and num_effects can be assumed to be linear**.

Correlations

			num_effects	trx_group
Spearman's rho	num_effects	Correlation Coefficient	1.000	.001
		Sig. (2-tailed)	.	.941
		N	16103	16103
	trx_group	Correlation Coefficient	.001	1.000
		Sig. (2-tailed)	.941	.
		N	16103	16103

A **Spearman's correlation coefficient value of 0.001** is very close to 0, which implies a negligible or almost non-existent correlation between the two variables being analyzed.

A p-value of 0.941, which is greater than the typical significance level of 0.05, indicates that there is **no statistically significant correlation between the two variables**. In other words, the null hypothesis of no correlation cannot be rejected, suggesting that the variables are nearly independent or have a very weak monotonic relationship.



Question 3

<u>Examine if there is a significant difference between the ages of the Drug and Placebo groups.</u>

I use the Kolmogorov-Smirnov test of normality and check normality for the Drug and Placebo groups.

H0: the sample data follows the **normal** distribution

H1: the sample data **NOT** follows the **normal** distribution

Descriptives

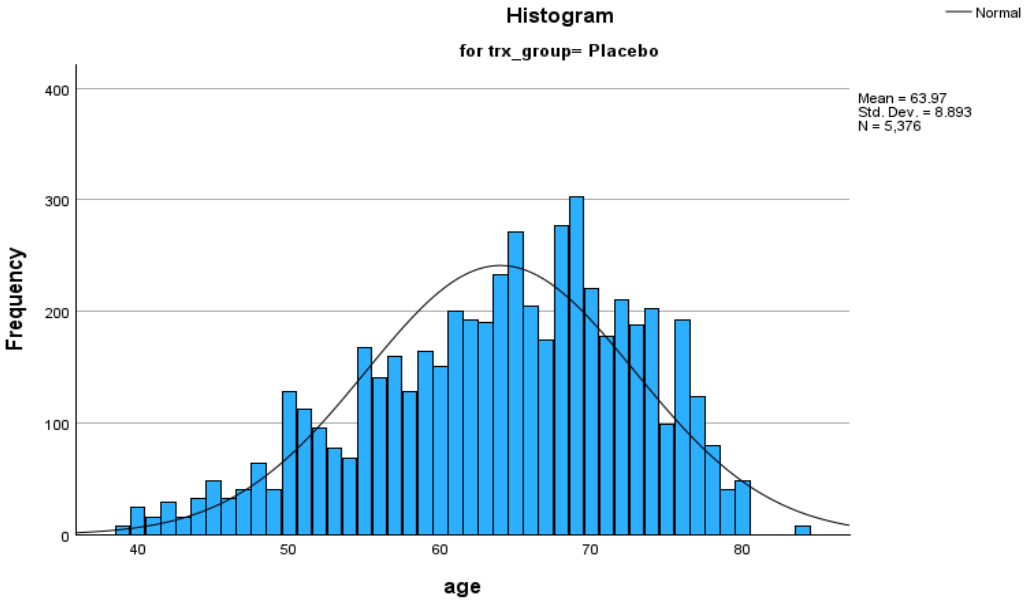
	trx_group			Statistic	Std. Error
age	Placebo	Mean		63.97	.121
		95% Confidence Interval for Mean	Lower Bound	63.73	
			Upper Bound	64.21	
		5% Trimmed Mean		64.27	
		Median		65.00	
		Variance		79.086	
		Std. Deviation		8.893	
		Minimum		39	
		Maximum		84	
		Range		45	
		Interquartile Range		13	
		Skewness		-.454	.033
		Kurtosis		-.396	.067
	Drug	Mean		64.19	.084
		95% Confidence Interval for Mean	Lower Bound	64.03	
			Upper Bound	64.36	
		5% Trimmed Mean		64.50	
		Median		65.00	
		Variance		76.163	
		Std. Deviation		8.727	
		Minimum		39	
		Maximum		82	
		Range		43	
		Interquartile Range		13	
		Skewness		-.474	.024
		Kurtosis		-.266	.047

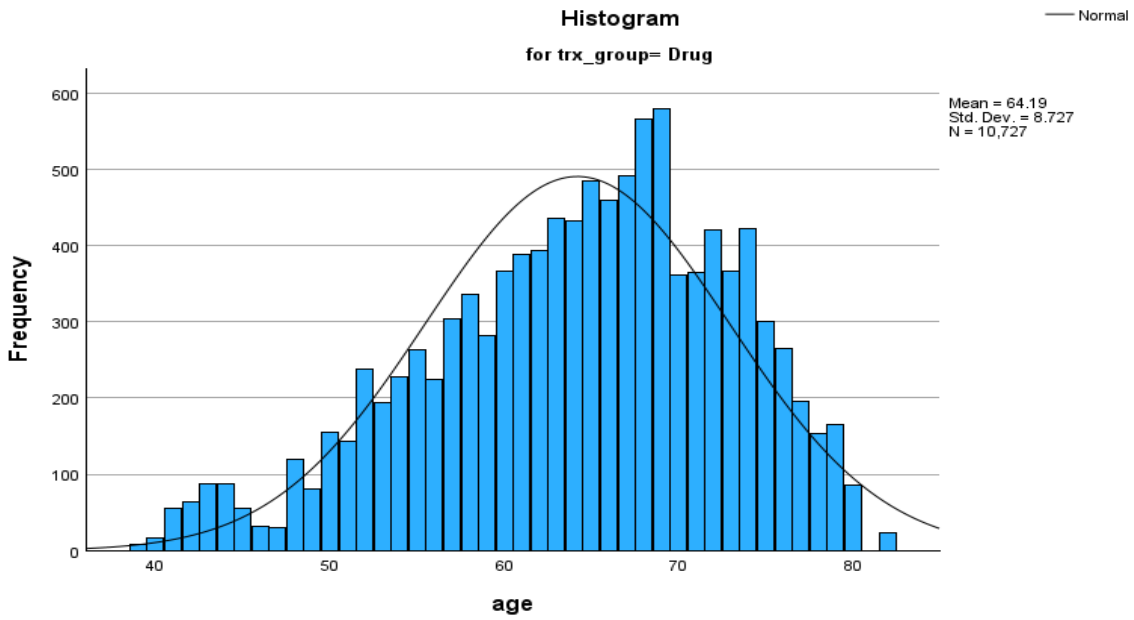
Tests of Normality

		Kolmogorov-Smirnov[a]		
	trx_group	Statistic	df	Sig.
age	Placebo	.078	5376	<.001
	Drug	.070	10727	<.001

a. Lilliefors Significance Correction

The p-value = <.0.001 is less than the chosen significance level 0.05, I **reject the null hypothesis**. This means that **the data does NOT follow the normal distribution.**

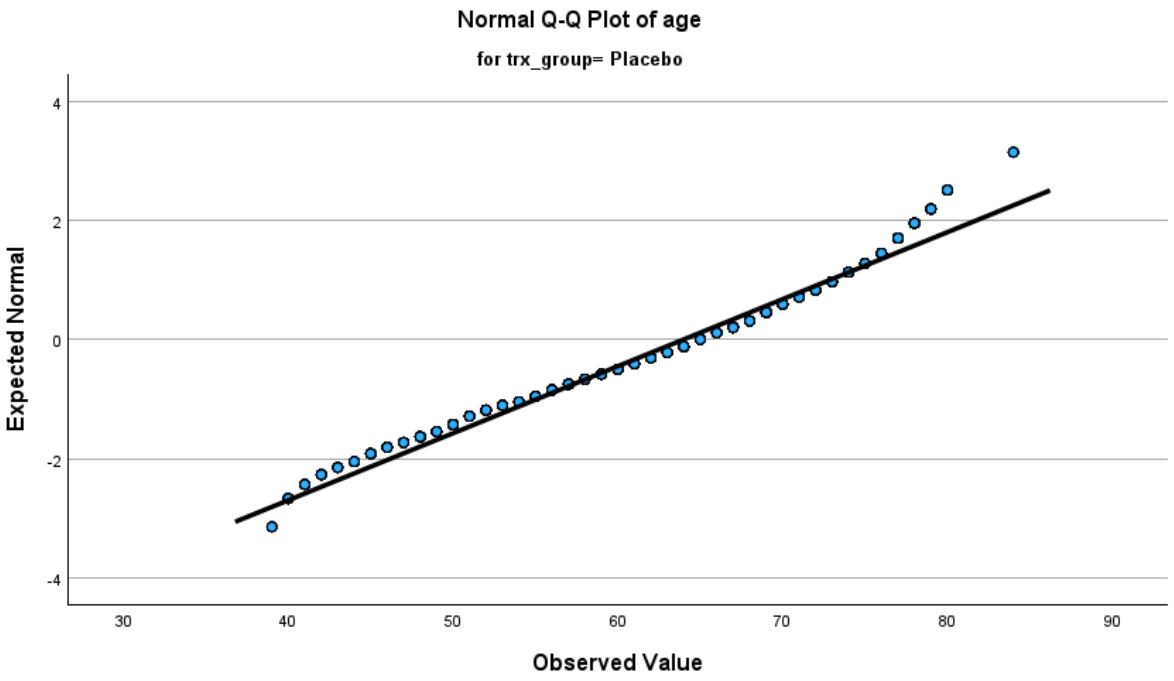


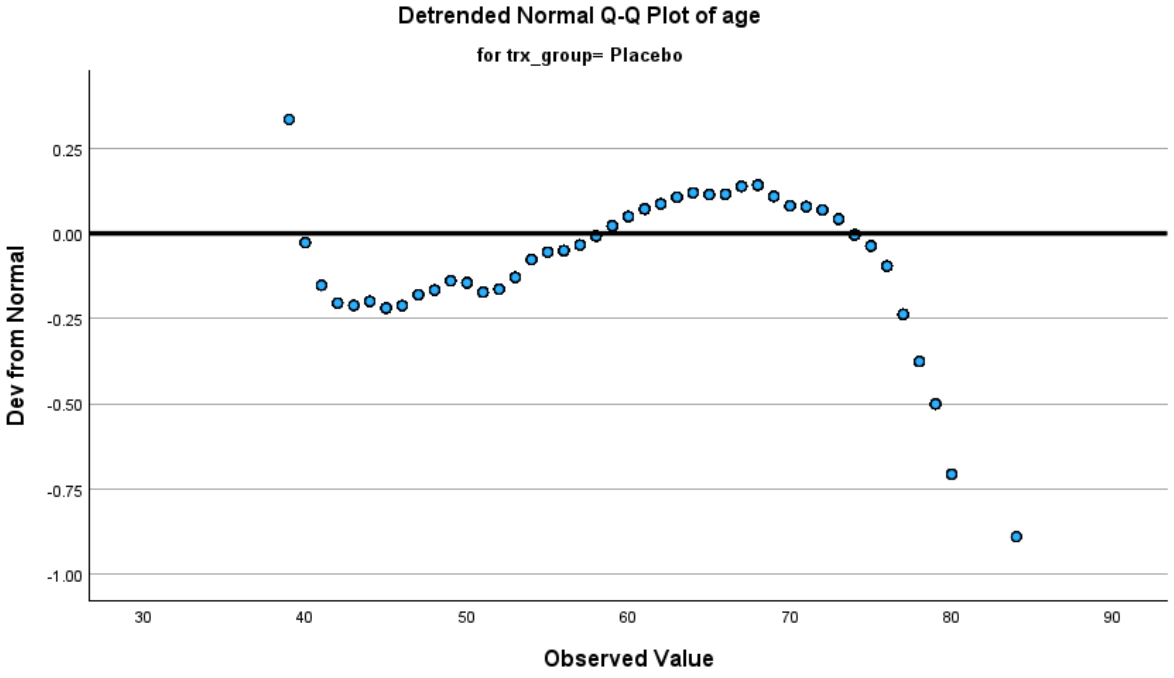


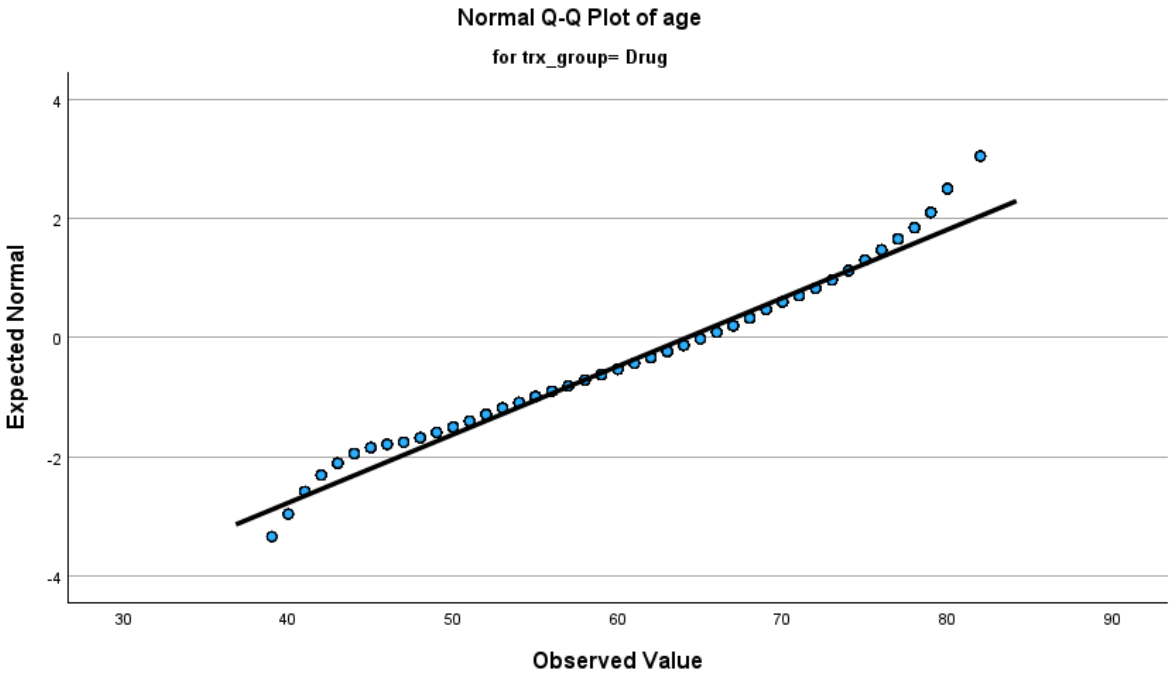
A **normal Q-Q (quantile-quantile) plot** is a graphical method used to assess whether a data set follows a normal distribution or not.

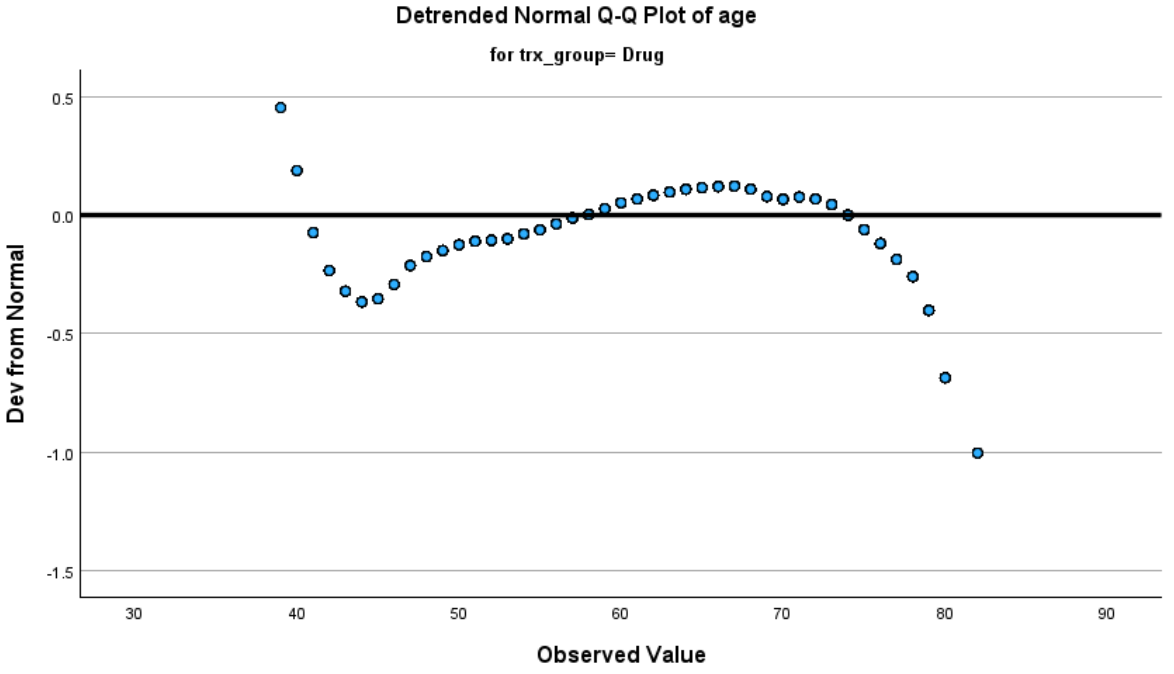
If the points on the Q-Q plot make an almost straight line, it means the data is likely normally distributed. The closer the points are to the straight line, the stronger the evidence that the data follows a normal distribution.

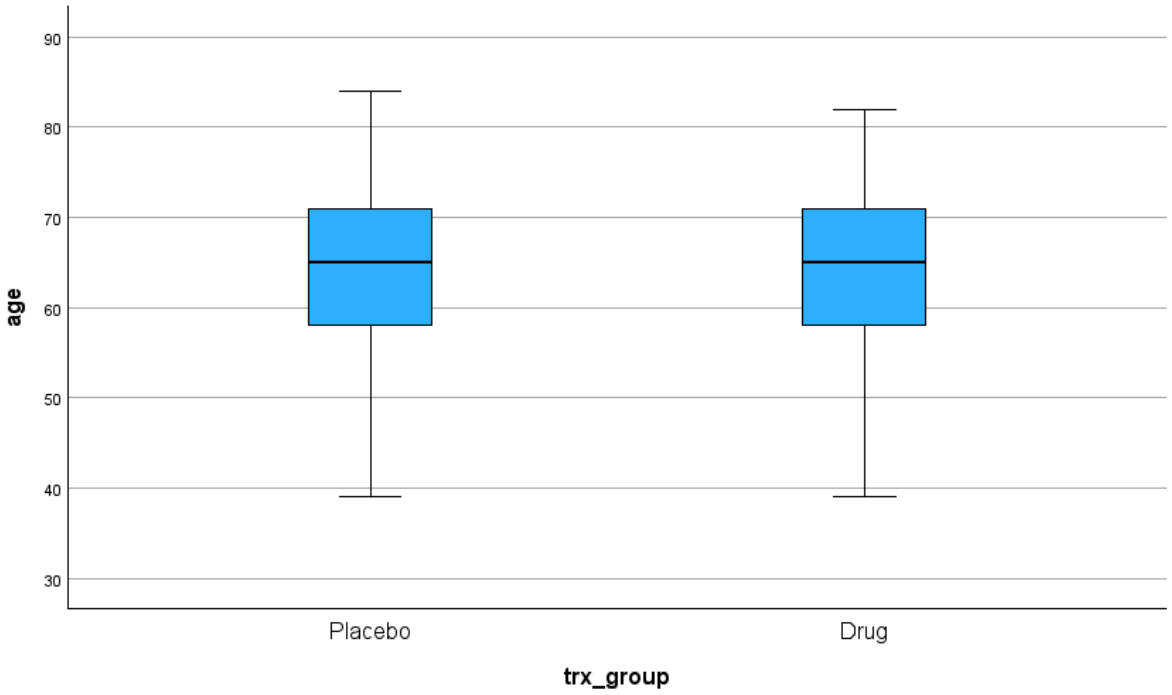
A **detrended Q-Q plot** is a variation of the normal Q-Q plot used to assess normality in a dataset. The main difference is that a detrended Q-Q plot accounts for and removes any potential trend or shift in the data before evaluating its normality.


Normal Q-Q Plot of age
for trx_group= Placebo
Expected Normal
4
2
0
-2
-4
30
40
50
60
70
80
90
Observed Value


Detrended Normal Q-Q Plot of age
for trx_group= Placebo
Dev from Normal
0.25
0.00
-0.25
-0.50
-0.75
-1.00
30
40
50
60
70
80
90
Observed Value


Normal Q-Q Plot of age
for trx_group= Drug
Expected Normal
4
2
0
-2
-4
30
40
50
60
70
80
90
Observed Value


Detrended Normal Q-Q Plot of age
for trx_group= Drug
Dev from Normal
0.5
0.0
-0.5
-1.0
-1.5
30
40
50
60
70
80
90
Observed Value



Summarization:
Based on the visual inspection of the graphics for the Placebo and Drug groups, we can see that the data is not normally distributed.

Independent-Samples Mann-Whitney U Test
The Independent-Samples Mann-Whitney U Test is a **non-parametric** statistical test used to compare the distributions of two **independent samples**.

H0: There is **NO significant difference** between the age distributions of the Drug and Placebo groups

H1:There is significant difference between the age distributions of the Drug and Placebo groups

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The distribution of age is the same across categories of trx_group.	Independent-Samples Mann-Whitney U Test	.257	Retain the null hypothesis.

a. The significance level is .050.
b. Asymptotic significance is displayed.

Summarization:
The p-value of 0.257 is higher than the significance level of 0.05. The Mann-Whitney U test result shows there is **no important difference in the ages between the Drug group and the Placebo group**. This means age is likely not an important factor separating the two treatment groups.