Minitab

Lecture: Statistical tests, normality test

- Minitab provides a wide range of **hypothesis testing tools**. A critical step before many parametric tests is to **check the assumption of normality**.

استخدام أوامر Minitab للاختبارات t و z

أقدم أدناه شرحًا مفصلاً عمليًا لكيفية إجراء اختبارات **t** و **z** في **Minitab** بطريقتين: **مباشرة**) أدخل البيانات يدوياً أو عبر SET ثم استخدم القوائم) و **غير مباشرة / مبرمجة**) استيراد من ملف، أو استخدام أوامر الجلسة/الماكروز أو البيانات الملخّصة). أبدأ بمبسّط لكل اختبار ثم أمثلة عملية وأوامر جاهزة للتجريب.

ما هو اختبار t ومتى يُستخدم؟ / What is the t-test (short)

اختبار t يقارن المتوسطات. أهم أنواعه:

- **1-Sample t**: يختبر ما إذا كان متوسط العينة مختلفًا عن قيمة مفترضة (μ0) عندما لا نعرف انحراف المجتمع σ.
- **2-Sample t (Independent)**: يقارن متوسطين لمجموعتين مستقلتين؛ يمكن افتراض تساوي التباينات (pooled) أو تركها مختلفة (Welch).
- **Paired t**: يقارن فروق أزواج مترابطة (قبل/بعد).

t-tests compare sample means when the population standard deviation is unknown.

Types

: 1-sample t

2-sample (pooled or Welch), paired

Methods and formulas for 1-Sample t

In This Topic

- Confidence interval (CI)
- T-value
- P-value

1-Confidence interval (CI)

Formula

$$\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Notation

Term	Description
$\bar{x}$	sample mean
$t_{\alpha/2}$	inverse cumulative probability of a t distribution with n-1 degrees of freedom at 1-α/2
α	1 – confidence level/100
s	sample standard deviation
n	sample size

Dr.Mahmood M Taher

2-T-value
Formula

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

3-P-value
Formula
The calculation for the p-value depends on the alternative hypothesis.

Alternative Hypothesis	P-value
$H_1: \mu > \mu_0$	$P(t_{n-1} \geq t \mid \mu = \mu_o)$
$H_1: \mu < \mu_0$	$P(t_{n-1} \leq t \mid \mu = \mu_o)$
$H_1: \mu \neq \mu_0$	$2 \times P(t_{n-1} \geq \lvert t \rvert \mid \mu = \mu_o)$

التنفيذ العملي في Minitab — الطريقة الغير مباشرة (inDirect)

Manual data entry (Copy/Paste or SET)

1- **Paste from Excel: Copy the columns and then paste into a Minitab sheet; each column in Excel becomes C1, C2…**
2- **Use SET for quick entry from the command window**

```
SET C1
DATA 78 85 92 67 74
END
```

تشغيل 1-Sample t (raw data)

1. بعد إدخال البيانات في عمود (مثلاً C1 اسمه :(`Grades`
   Stat > Basic Statistics > 1-Sample t....
2. اختر "One or more samples, each in a column" ثم أدخل العمود. فعل "Perform hypothesis test" وحدد `Hypothesized mean` ثم OK.

```
MTB > OneT C1;
SUBC>   Test 70;
SUBC>   Confidence 95.0;
SUBC>   Alternative 0.
MTB >
```

Descriptive Statistics

Dr.Mahmood M Taher

Mahmood81_tahr@uomosul ed.iq

N	Mean	StDev	SE Mean	95% CI for μ
5	79.20	9.68	4.33	(67.18, 91.22)

μ: mean of C1

Test

Null hypothesis	H_0: μ = 70
Alternative hypothesis	H_1: μ ≠ 70

T-Value	P-Value
2.13	0.101

```
MTB > OneT C1;
SUBC>   Test 70;
SUBC>   Confidence 95.0;
SUBC>   Alternative -1.
```

Descriptive Statistics

N	Mean	StDev	SE Mean	95% Upper Bound for μ
5	79.20	9.68	4.33	88.43

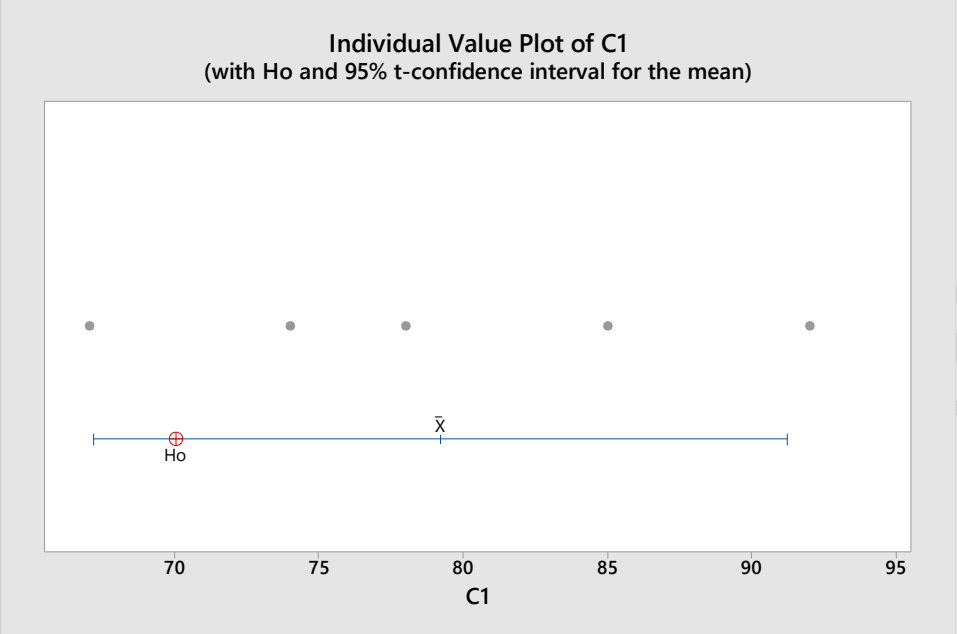
μ: mean of C1

Test

Null hypothesis	H_0: μ = 70
Alternative hypothesis	H_1: μ < 70

T-Value	P-Value
2.13	0.950

Dr. Mahmood M Taher



شرح

النقاط الرمادية في الأعلى

هذه النقاط تمثل القيم الفردية لكل ملاحظة في العينة أي كل قيمة من C1. مثلاً إذا كان لديك 5 مشاهدات، فكل نقطة تمثل قيمة واحدة منها.

- الخط الأفقي الأزرق

يمثل فترة الثقة (Confidence Interval) للمتوسط الحسابي.
الطرفان الأيمن والأيسر يوضحان حدود فترة الثقة (الحد الأدنى والأعلى).
الرمز $\bar{X}$ في المنتصف هو المتوسط الحسابي للعينة.

- الرمز الأحمر (Ho)

يمثل الفرضية الصفرية H0
أي القيمة المفترضة لمتوسط المجتمع (Population mean) التي يتم اختبارها في اختبار t.
مثلاً:

$H0:\mu=70$

التفسير الإحصائي

من الرسم، يمكن ملاحظة ما يلي:

القيمة المفترضة H0 تقع داخل فترة الثقة (الخط الأزرق). هذا يعني أن الفرق بين المتوسط الفعلي والعينة ليس دالًا إحصائيًا عند مستوى ثقة 95%.
بعبارة أخرى، لا نرفض الفرضية الصفرية H0 لأن القيمة المفترضة تقع ضمن فترة الثقة.

Dr.Mahmood M Taher

2-Sample t (pooled or Welch)

1. ادخل عمودين (كل مجموعة في عمود) أو عمود واحد مع عمود تعريف المجموعة.
2. **Stat > Basic Statistics > 2-Sample t....**
3. في النافذة: اختر "Samples in different columns" أو "Both samples in one column" حسب بياناتك.
4. لتطبيق pooled (تساوي التباينات) فعّل **Assume equal variances**، وإلا اتركها للتحليل بدون افتراض (Welch).

1. Prepare two columns (each group in one column) or one column with a group identification column.
2. Stat > Basic Statistics > 2-Sample t...
3. In the window: Choose "Samples in different columns" or "Both samples in one column" according to your data.
4. To apply pooled (equal variances) values, enable Assume equal variances; otherwise, leave the analysis without an assumption (Welch).

```
MTB > TwoT C1 C2;
SUBC>   Confidence 95.0;
SUBC>   Test 0.0;
SUBC>   Alternative 0.
MTB >
```

Two-Sample T-Test and CI: C1, C2

Method

μ_1: mean of C1 when C2 = 1

μ_2: mean of C1 when C2 = 2

Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics: C1

C2	N	Mean	StDev	SE Mean
1	3	85.00	7.00	4.0
2	3	67.00	7.00	4.0

Estimation for Difference

Difference	95% CI for Difference
18.00	(2.13, 33.87)

Mahmood M Taher

Mahmood81_tahr@uomosul ed.iq

Test

Null hypothesis	$H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
3.15	4	0.035

```
MTB > TwoT C1 C2;
SUBC>   Confidence 95.0;
SUBC>   Test 60;
SUBC>   Alternative 0.
MTB >
```

Two-Sample T-Test and CI: C1, C2

Method

μ_1: mean of C1 when C2 = 1

μ_2: mean of C1 when C2 = 2

Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics: C1

C2	N	Mean	StDev	SE Mean
1	3	85.00	7.00	4.0
2	3	67.00	7.00	4.0

Estimation for Difference

Difference	95% CI for Difference
18.00	(2.13, 33.87)

Test

Null hypothesis	$H_0: \mu_1 - \mu_2 = 60$
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq 60$

Mahmood M Taher

T-Value	DF	P-Value
-7.35	4	0.002

اختبار t للازواج المترابطة

MTB > Paired C1 C3;

SUBC> Confidence 95.0;

SUBC> Test 0.0;

SUBC> Alternative 0.

MTB >

Paired T-Test and CI: C1, C3

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
C1	6	76.00	11.68	4.77
C3	6	56.33	23.74	9.69

Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for μ_difference
19.67	23.67	9.66	(-5.17, 44.51)

μ_difference: mean of (C1 - C3)

Test

Null hypothesis	H_0: μ_difference = 0
Alternative hypothesis	H_1: μ_difference ≠ 0

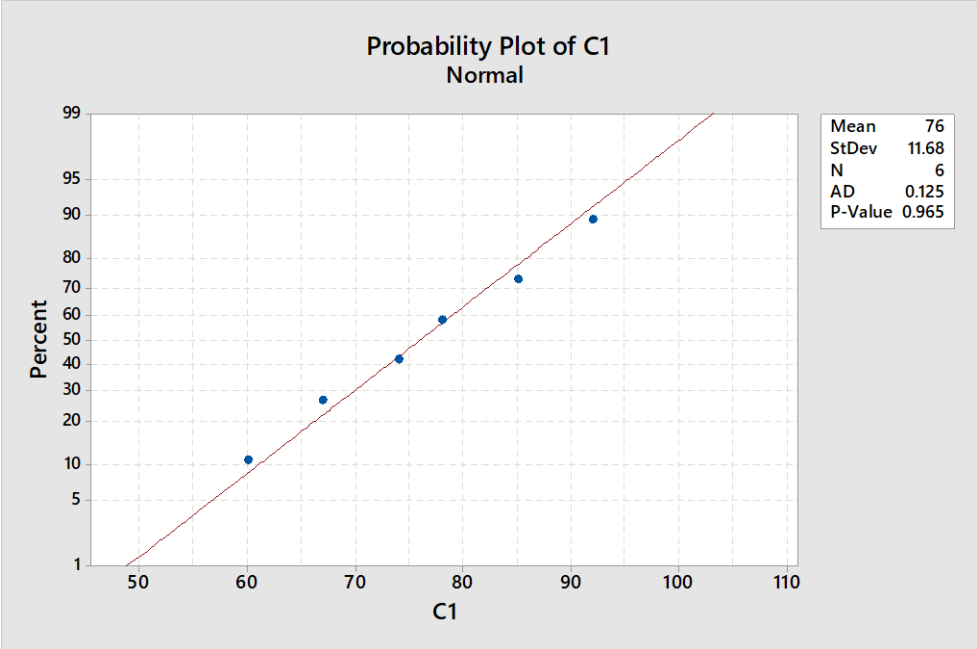
T-Value	P-Value
2.04	0.097

ما هو اختبار z ومتى يُستخدم؟What is the z-test (short)

اختبار z يستخدم عندما **نعرف** انحراف المجتمع σ أو عندما يكون حجم العينة كبيرًا جدًا بحيث يمكن تقريب t بالـ normal. عادة نستخدم Sample Z (Minitab 1)-يوفرها (لاختبار متوسط ضد μ0 مع σ معروف .

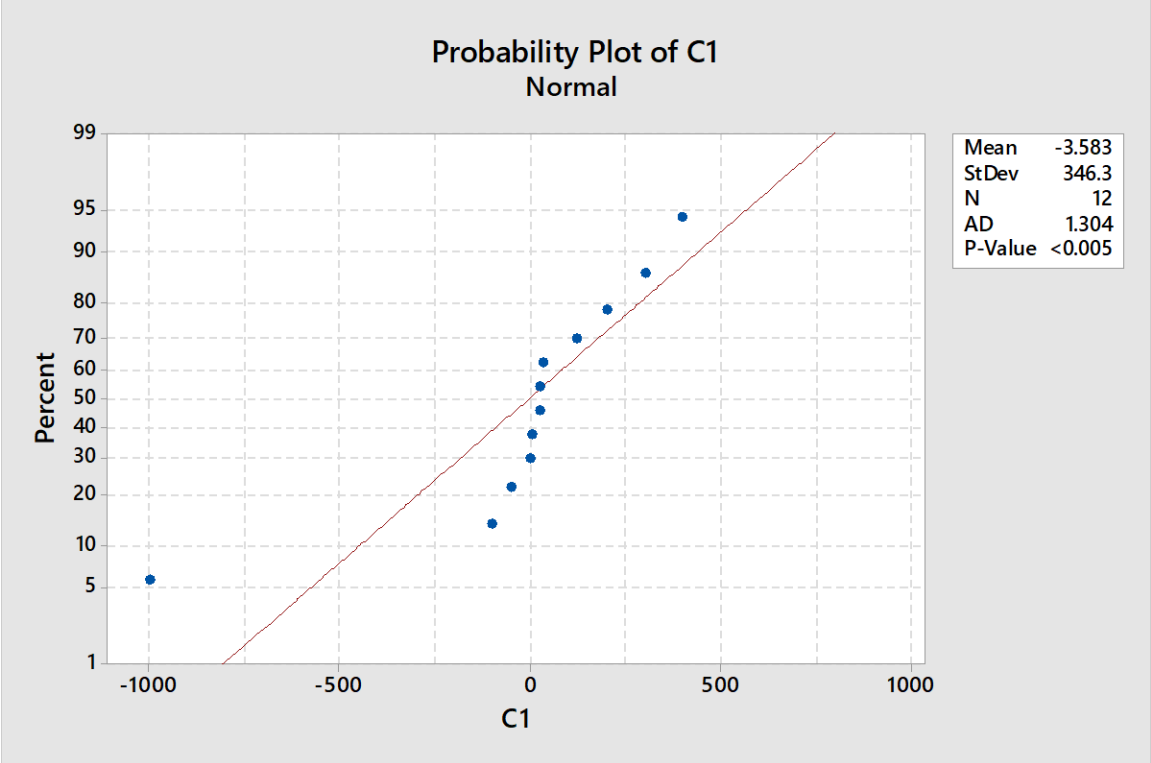
The z-test is used when we know the population skewness σ or when the sample size is so large that t can be approximated by normal. We typically use 1-Sample Z (provided by Minitab) to test the mean against μ0 with σ known.

MTB > NormTest C1.
78
85
92
67
74
60



25
50-
120
300
1000-
200
400
33
100-
24
5
0

Dr.Mahmood Taher



MTB > Outlier C1;

SUBC> Grubbs;

SUBC> Alpha 0.05;

SUBC> Alternative 0;

SUBC> NoDefault;

SUBC> TMethod;

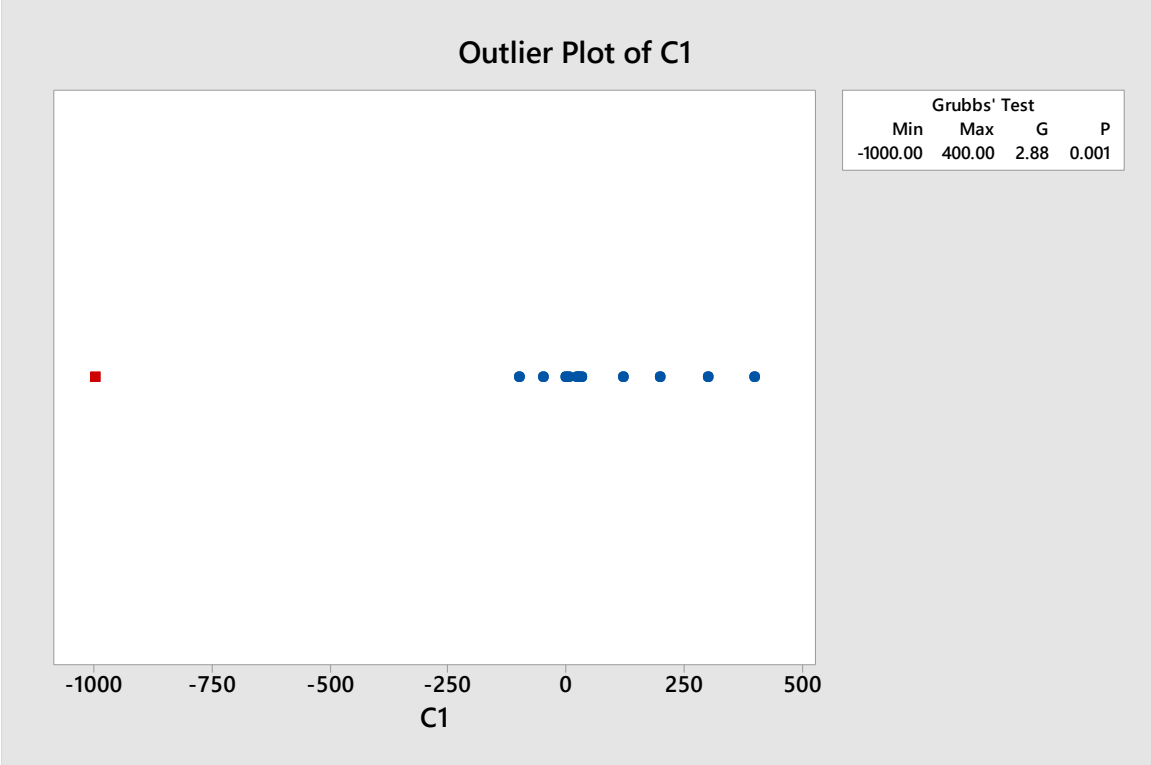
SUBC> TTest;

SUBC> TOutlier;

SUBC> GOutlierplot.

MTB >

Dr Mahmood



Outlier Test: C1

Method

Null hypothesis	All data values come from the same normal population
Alternative hypothesis	Smallest or largest data value is an outlier
Significance level	α = 0.05

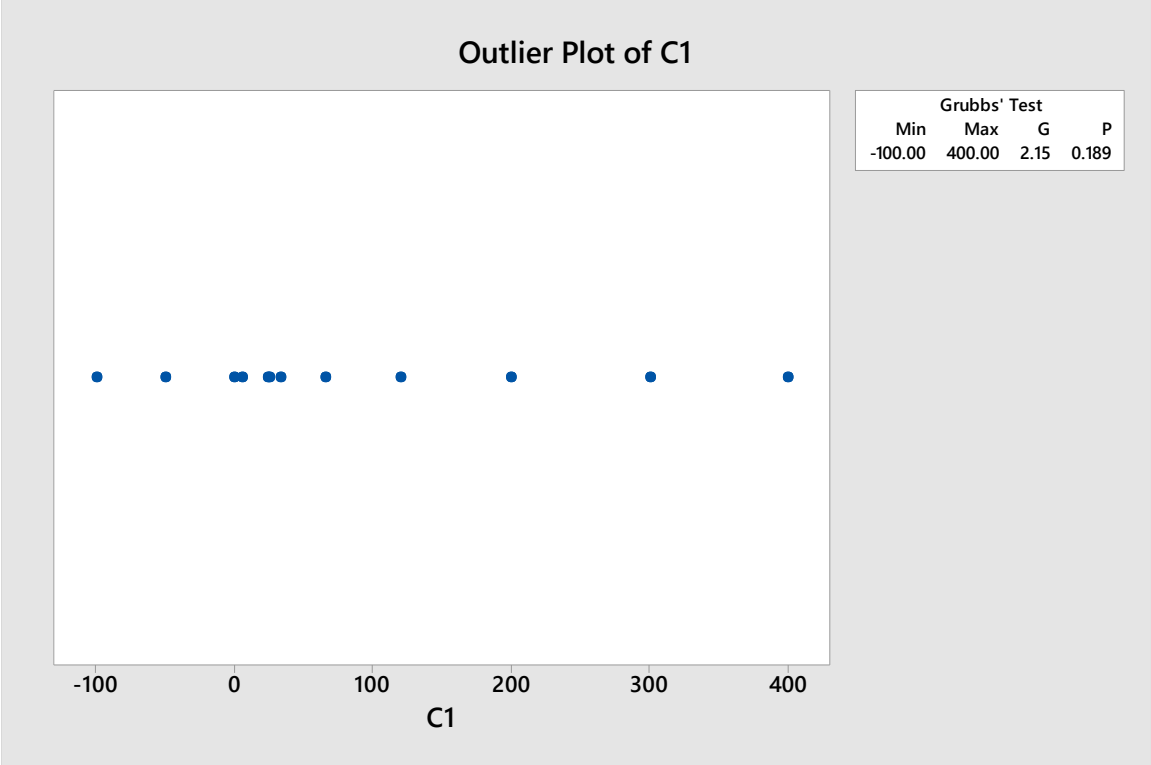
Grubbs' Test

Variable	N	Mean	StDev	Min	Max	G	P
C1	12	-4	346	-1000	400	2.88	0.001

Outlier

Variable	Row	Outlier
C1	5	-1000

استبدل قيمة -1000 الى 65



Outlier Test: C1

Method

Null hypothesis	All data values come from the same normal population
Alternative hypothesis	Smallest or largest data value is an outlier
Significance level	α = 0.05

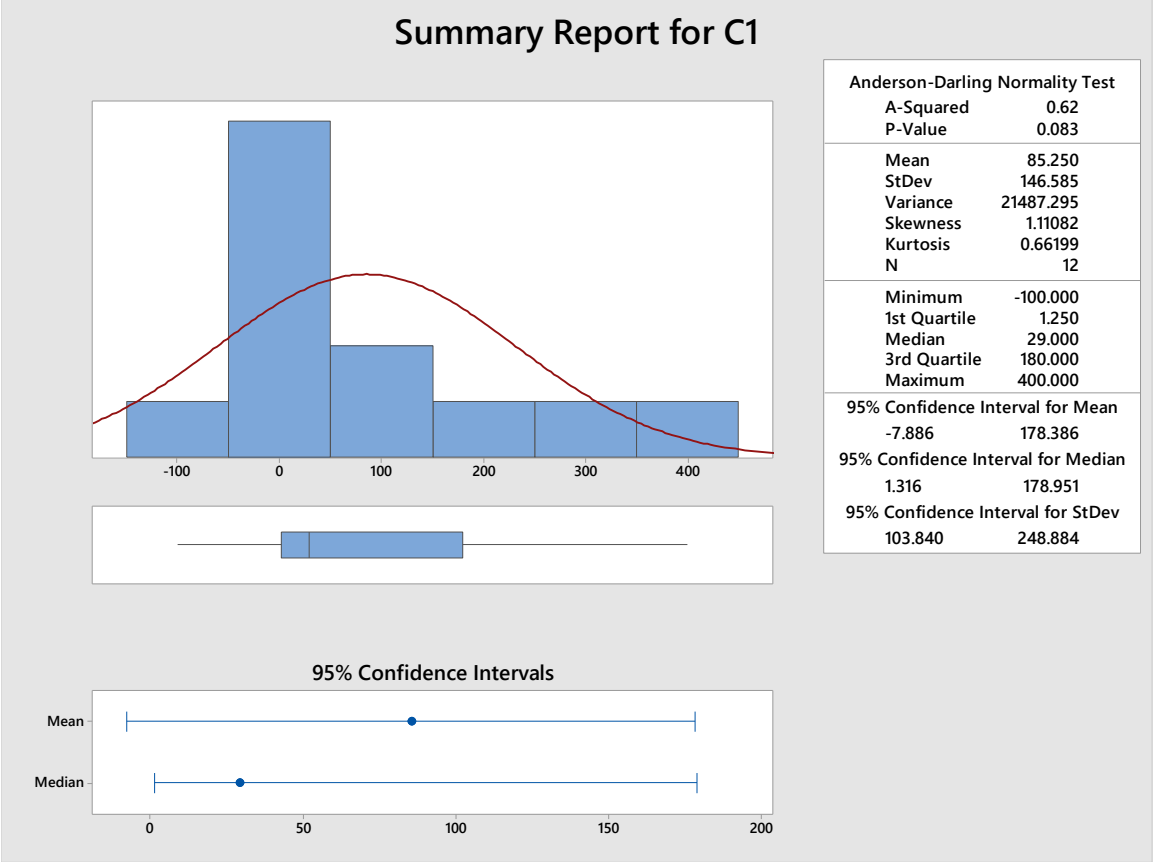
Grubbs' Test

Variable	N	Mean	StDev	Min	Max	G	P
C1	12	85.3	146.6	-100.0	400.0	2.15	0.189

*** NOTE * No outlier at the 5% level of significance**

MTB > GSummary C1.

Mahmood81_tahr@uomosul ed.iq



Anderson-Darling Normality Test	
A-Squared	0.62
P-Value	0.083
Mean	85.250
StDev	146.585
Variance	21487.295
Skewness	1.11082
Kurtosis	0.66199
N	12
Minimum	-100.000
1st Quartile	1.250
Median	29.000
3rd Quartile	180.000
Maximum	400.000
95% Confidence Interval for Mean	
-7.886	178.386
95% Confidence Interval for Median	
1.316	178.951
95% Confidence Interval for StDev	
103.840	248.884

Dr.Mahm