

Independent Samples t-test

We illustrate the independent samples *t*-test using data from *Case Study 9.3*, comparing males and females on their reports of sexual jealousy. The original data is presented below (from Table 9.3).

Data from Table 9.3: Sexual Jealousy scores for Males and Females (N = 30)

	Group 1 (Males)	Group 2 (Females)
	3	3
	3	3
	2	3
	2	3
	3	2
	4	1
	4	1
	3	2
	4	2
	4	3
	4	3
	5	3
	5	5
	4	4
	4	4
N	15	15

This is the data set from the jealousy example from Table 9.3. However, the data must be reorganized before it is entered into the IBM SPSS *Data View* data window and SAS syntax edit screen. Both IBM SPSS and SAS expect that you will have a single column of data indicating your independent variable or grouping variable, and a single column of data indicating your dependent variable or outcome variable. The data should be reorganized into the following format:

Revised Data from Table 9.3: Sexual Jealousy scores by Males and Females ($N = 30$)

Group (Male = 1, Female = 2)	Jealousy
1	3
1	3
1	2
1	2
1	3
1	4
1	4
1	3
1	4
1	4
1	4
1	4
1	5
1	5
1	4
1	4
2	3
2	3
2	3
2	3
2	2
2	1
2	1
2	2
2	2
2	3
2	3
2	3
2	5
2	4
2	4

The above format has a single column indicating Group Membership. Males are coded as "1" while Females are coded as "2". The other column of data indicates the sexual jealousy scores.

Once reorganized, the data may be entered directly into IBM SPSS or SAS. Only enter the numeric data. Once you have entered the two columns of data, in IBM SPSS you can go to the *Variable View* screen and enter the two variable names (Group, Jealousy) and fill-in the variable and value labels. For example, for the variable *Group* you can enter *Gender Group* as the variable label, and enter *1 = Male* and *2 = Female* as the value label. For the variable *Jealousy*, you can enter a variable label *Jealousy Score*.

IBM SPSS syntax. Once data are entered in IBM SPSS, we can now proceed with the analysis using programming syntax. Open a syntax window (go to *File*, then *New*, then *Syntax*). You may directly enter the *t*-test commands shown on the next page in the syntax window. Then, highlight all the syntax, and press the *Play* icon which looks like a play button on a DVD or VCR). Your output should resemble what is illustrated on the next page.

```
T-TEST GROUPS=Group(1 2)
/MISSING=ANALYSIS
/VARIABLES=Jealousy
/CRITERIA=CI(.95).
```

T-Test

Group Statistics					
	Group		Mean	Std. Deviation	Std. Error Mean
	Gender	N			
Jealousy Score	1 Male	15	3.60	.910	.235
	2 Female	15	2.80	1.082	.279

Independent Samples Test															
		Levene's Test for Equality of Variances				t-test for Equality of Means									
		F		Sig.		t		Sig. (2-tailed)		Mean Difference		Std. Error Difference		95% Confidence Interval of the Difference	
												Lower		Upper	
Jealousy Score	Equal variances assumed	.060	.809	2.191	28	.037	.800	.365	.052	1.548					
	Equal variances not assumed			2.191	27.201	.037	.800	.365	.051	1.549					

IBM SPSS first provides the group sample sizes, means and standard deviation values, and standard errors. Levene's test is next offered as an assessment of equality of variances between groups (if used, be conservative with this test at the $p < .001$ level). The *t*-test value is next provided. Next to "Equal variances assumed" is the pooled variance result, showing $t(28) = 2.191, p = .037$ (two-tailed). The mean difference (0.80) is provided, as is a confidence interval surrounding the mean difference.

SAS syntax. We use the exact same data setup noted in Table 9.3. As with most of our SAS examples, we provide the full data setup and output. Once you open SAS, simply go to your *Program Editor* window, and type the following syntax:

```
Data example;
input Group Jealousy;
Datalines;
1      3
1      3
1      2
1      2
1      3
1      4
1      4
1      3
1      4
1      4
1      4
1      5
1      5
1      4
1      4
2      3
2      3
2      3
2      3
2      2
2      1
2      1
2      2
2      2
2      3
2      3
2      3
2      5
2      4
2      4
;
run;

Proc ttest;
  class Group;
  var Jealousy;
run;
```

The SAS System

The TTEST Procedure

Statistics

Variable	Group	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err	Minimum	Maximum
Jealousy	1	15	3.0959	3.6	4.1041	0.6664	0.9103	1.4356	0.235	2	5
Jealousy	2	15	2.2006	2.8	3.3994	0.7924	1.0823	1.7069	0.2795	1	5
Jealousy	Diff (1-2)		0.052	0.8	1.548	0.7936	1	1.3525	0.3651		

T-Tests

Variable	Method	Variances	DF	t Value	Pr > t
Jealousy	Pooled	Equal	28	2.19	0.0369
Jealousy	Satterthwaite	Unequal	27.2	2.19	0.0372

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
Jealousy	Folded F	14	14	1.41	0.5255

SAS provides the group sample sizes, means and standard deviation values (and 95% confidence intervals), and standard errors. Equality of variances is assessed using a Folded F method. Our *t*-test value is located in the center of the output, next to *Pooled*. Below that is the unequal variance *t*-test, labeled *Satterthwaite* to indicate the method of adjustment for unequal variances.

Explore

Notes

Output Created	06-SEP-2019 10:18:17	
Comments		
Input	Data	C: \Users\BIII\Desktop\PILES\Pile35 _Courage\Conferences\SPSP202 0\Data_&_Syntax\courage_new1 9_cleaned.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
Missing Value Handling	N of Rows in Working Data File	307
	Definition of Missing	User-defined missing values for dependent variables are treated as missing.
	Cases Used	Statistics are based on cases with no missing values for any dependent variable or factor used.
Syntax	EXAMINE VARIABLES=pcst BY gender /PLOT BOXPLOT HISTOGRAM NPLOT /COMPARE GROUPS /STATISTICS DESCRIPTIVES /INTERVAL 95 /MISSING LISTWISE /NOTOTAL.	
Resources	Processor Time	00:00:04.57
	Elapsed Time	00:00:04.34

gender Gender

Case Processing Summary

	gender Gender	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
pcst pcsTotal	1 Male	68	90.7%	7	9.3%	75	100.0%
	2 Female	192	88.1%	26	11.9%	218	100.0%

Descriptives

gender Gender			Statistic	Std. Error	
pcst pcsTotal	1 Male	Mean	95.6618	1.71032	
		95% Confidence Interval for Mean	Lower Bound		92.2479
			Upper Bound		99.0756
		5% Trimmed Mean	95.7876		
		Median	94.0000		
		Variance	198.914		
		Std. Deviation	14.10368		
		Minimum	63.00		
		Maximum	126.00		
	Range	63.00			
	Interquartile Range	16.75			
	Skewness	-.102	.291		
	Kurtosis	-.203	.574		
	2 Female	Mean	91.5260	1.00666	
		95% Confidence Interval for Mean	Lower Bound		89.5404
			Upper Bound		93.5116
		5% Trimmed Mean	92.0243		
Median		91.0000			
Variance		194.565			
Std. Deviation		13.94865			
Minimum		30.00			
Maximum		122.00			
Range	92.00				
Interquartile Range	16.00				
Skewness	-.698	.175			
Kurtosis	1.980	.349			

Tests of Normality

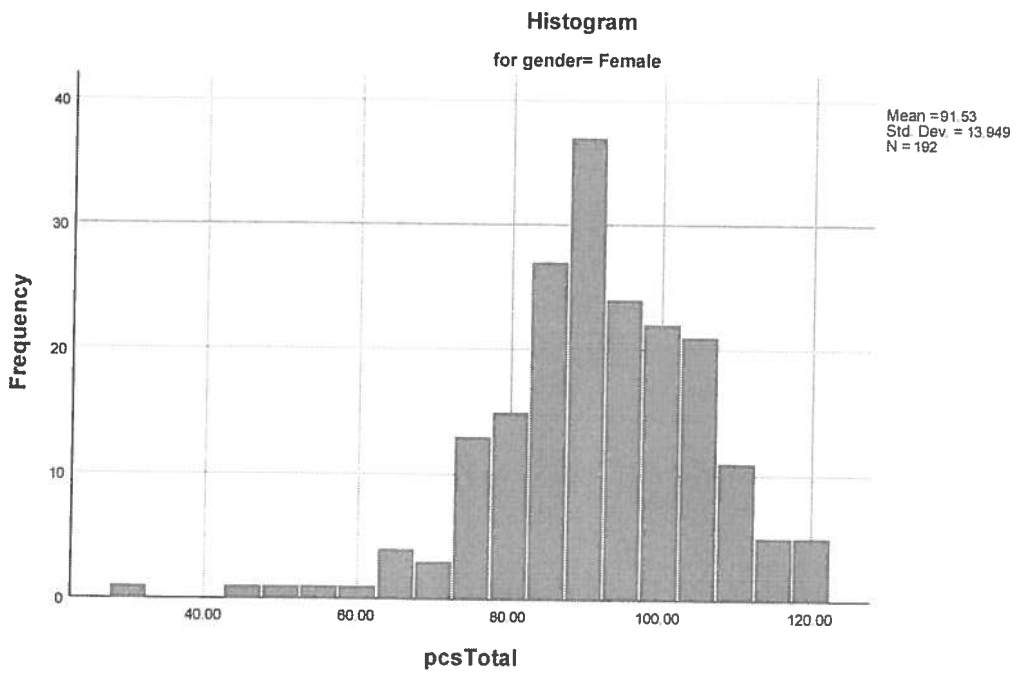
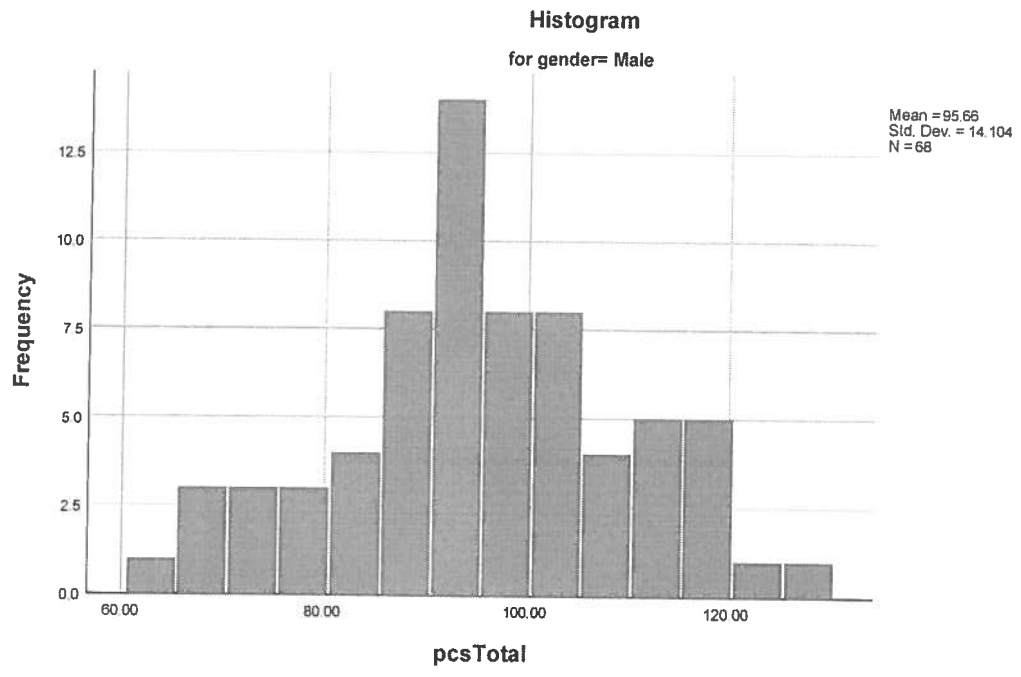
gender Gender	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
pcst pcsTotal 1 Male	.079	68	.200 [*]	.986	68	.626
2 Female	.075	192	.010	.969	192	.000

*. This is a lower bound of the true significance.

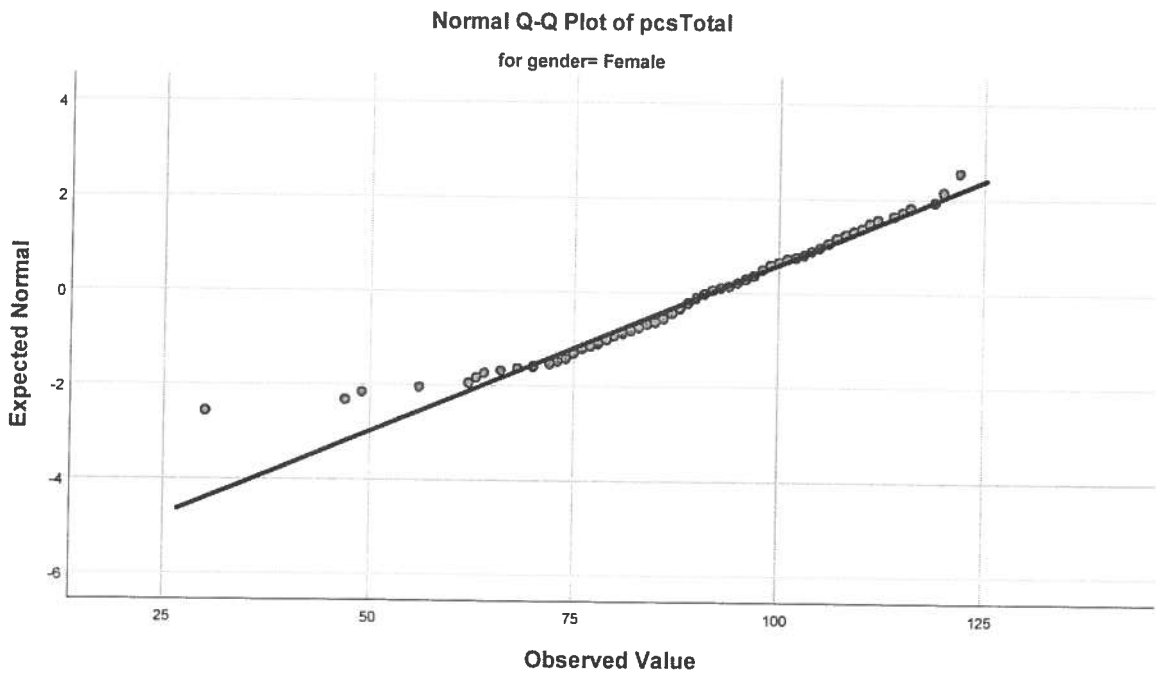
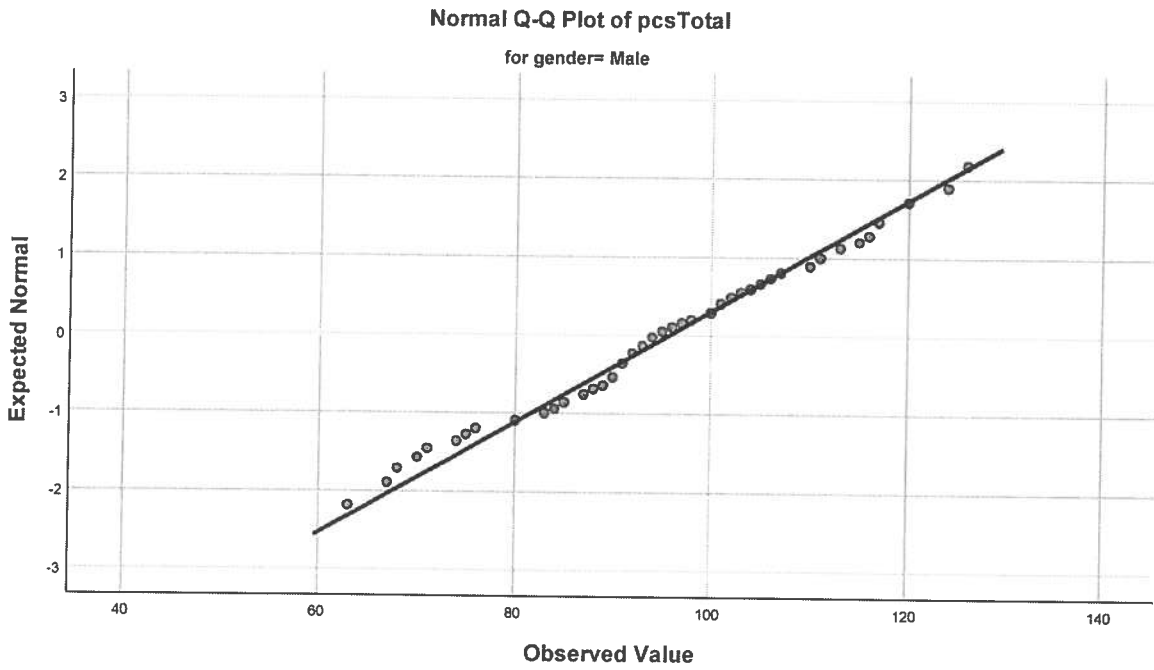
a. Lilliefors Significance Correction

pcst pcsTotal

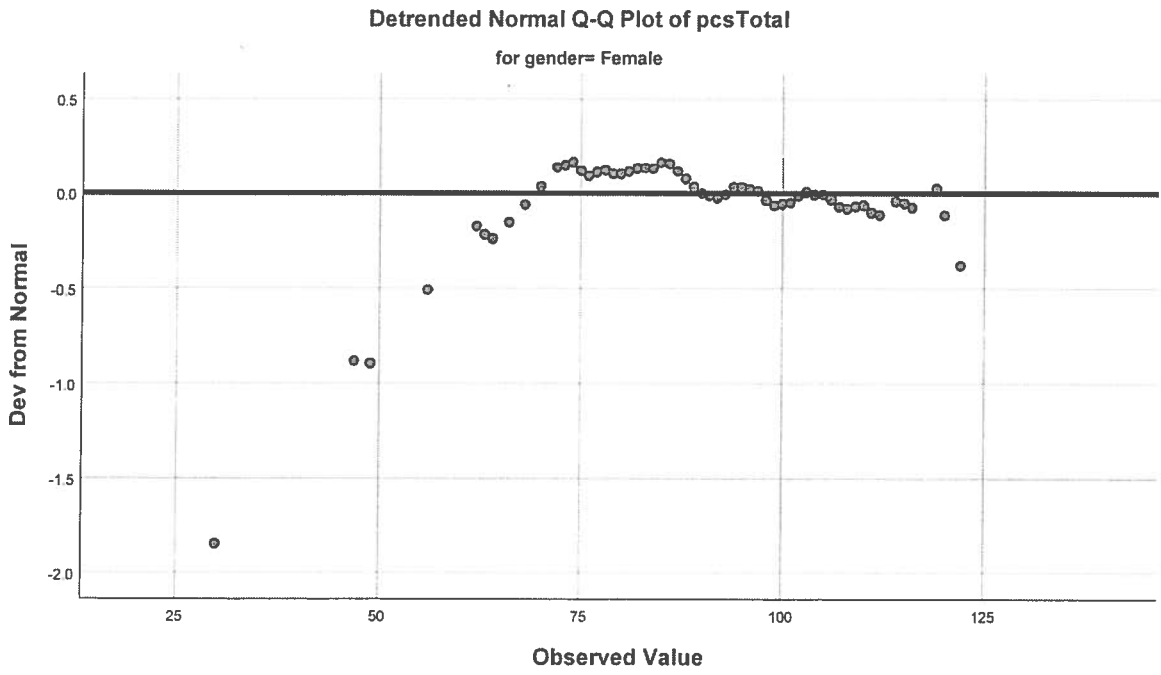
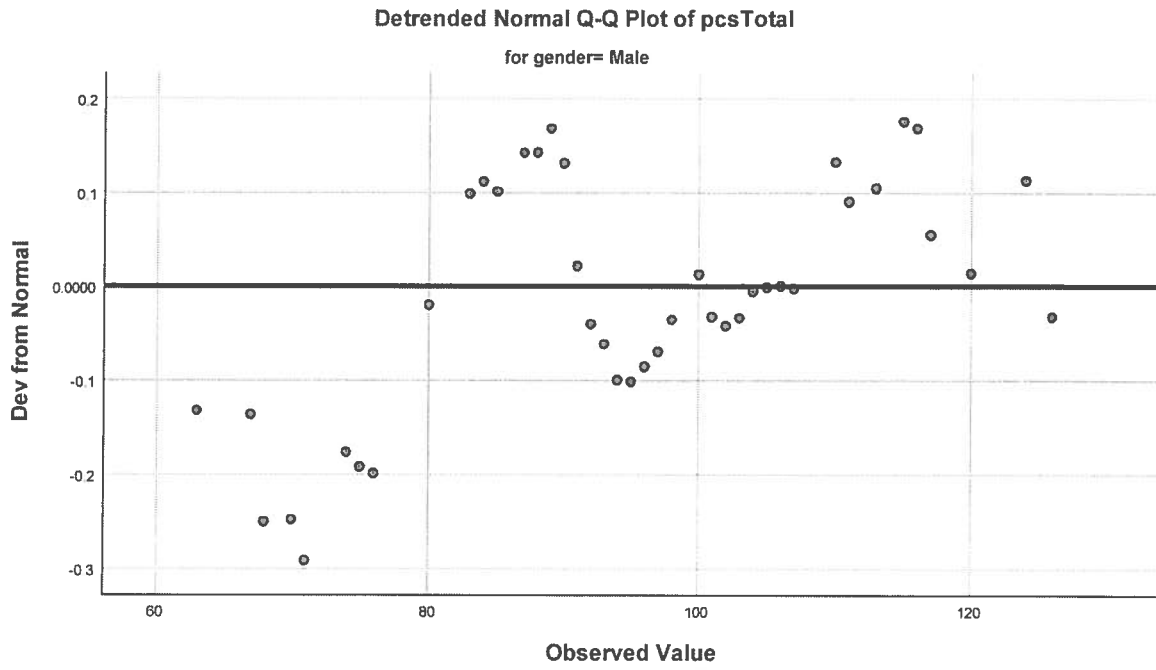
Histograms

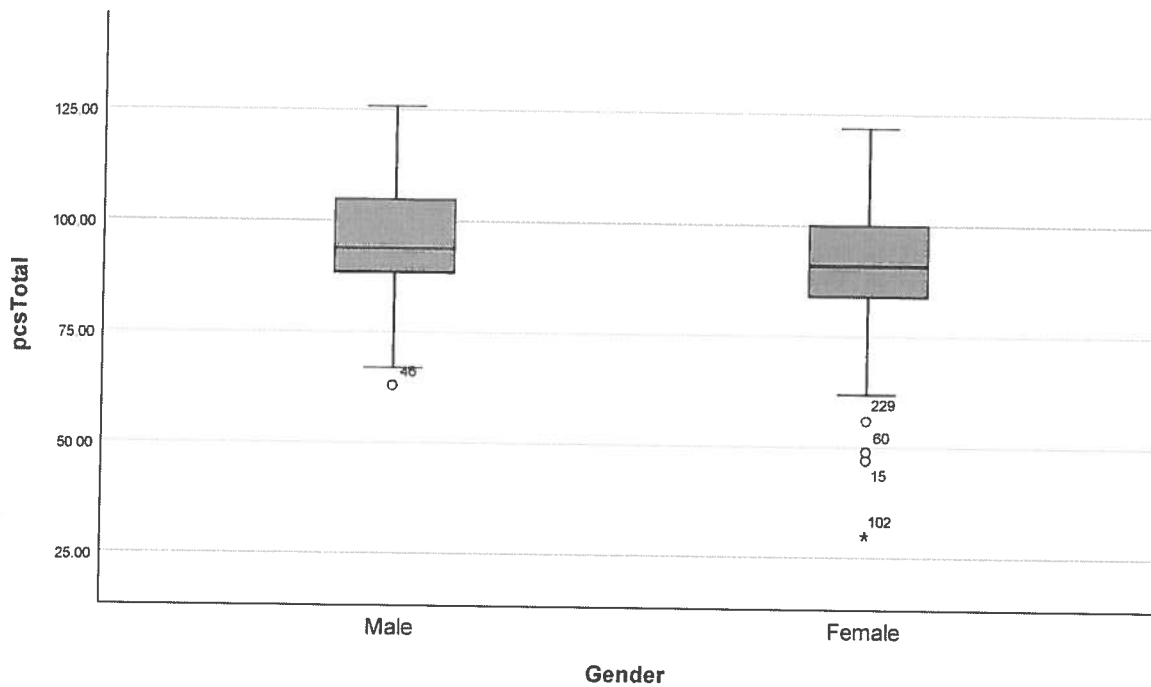


Normal Q-Q Plots



Detrended Normal Q-Q Plots





T-Test

Notes

Output Created	06-SEP-2019 10:21:51	
Comments		
Input	Data	C: \\Users\Bill\Desktop\PILES\Pile35 _Courage\Conferences\SPSP202 0\Data_&_Syntax\courage_new1 9_cleaned.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
Missing Value Handling	N of Rows in Working Data File	307
	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis.
Syntax	T-TEST GROUPS=gender(1 2) /MISSING=ANALYSIS /VARIABLES=pcst /CRITERIA=C(.95).	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.03

Group Statistics

	gender Gender	N	Mean	Std. Deviation	Std. Error Mean
pcst pcsTotal	1 Male	68	95.6618	14.10368	1.71032
	2 Female	192	91.5260	13.94865	1.00666

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
pcsl pcsTotal	Equal variances assumed	.189	.664	2.095	258	.037
	Equal variances not assumed			2.084	116.555	.039

Independent Samples Test

		t-test for Equality of Means			
		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
				Lower	Upper
pcst pcsTotal	Equal variances assumed	4.13572	1.97411	.24831	8.02314
	Equal variances not assumed	4.13572	1.98458	.20521	8.06624

List

Notes

Output Created	06-SEP-2019 10:43:53	
Comments		
Input	Data	C: \Users\Bill\Desktop\desktop\BOO K-t_test_exercise_ben_example. sav
	Active Dataset	DataSet2
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	30
Syntax		List /var all.
Resources	Processor Time	00:00:00.03
	Elapsed Time	00:00:00.05

[DataSet2] C:\Users\Bill\Desktop\desktop\BOOK-t_test_exercise_ben_example.sav

Group Exer_Intent

1.00	5.00
1.00	3.00
1.00	4.00
1.00	5.00
1.00	6.00
1.00	7.00
1.00	4.00
1.00	5.00
1.00	6.00
1.00	4.00
1.00	3.00
1.00	5.00
1.00	5.00
1.00	6.00
1.00	5.00
2.00	3.00
2.00	2.00
2.00	1.00
2.00	4.00
2.00	3.00
2.00	4.00
2.00	3.00
2.00	3.00
2.00	3.00
2.00	4.00
2.00	4.00
2.00	1.00
2.00	2.00
2.00	2.00
2.00	2.00

Number of cases read: 30 Number of cases listed: 30

```
T-TEST GROUPS=Group(1 2)
/MISSING=ANALYSIS
/VARIABLES=Exer_Intent
/CRITERIA=CI(.95).
```

T-Test

Notes

Output Created	06-SEP-2019 10:45:22	
Comments		
Input	Data	C: \\Users\Bill\Desktop\desktop\BOO K-t_test_exercise_ben_example. sav
	Active Dataset	DataSet2
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	30
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis.
Syntax		T-TEST GROUPS=Group(1 2) /MISSING=ANALYSIS /VARIABLES=Exer_Intent /CRITERIA=C(.95).
Resources	Processor Time	00:00:00.05
	Elapsed Time	00:00:00.17

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Exer_Intent Days Intent (# Days)	1.00 Benefits Page	15	4.8667	1.12546	.29059
	2.00 No Benefits Page	15	2.7333	1.03280	.26667

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	df
Exer_Intent Days Intent (# Days)	Equal variances assumed	.002	.969	5.409	28
	Equal variances not assumed			5.409	27.796

Independent Samples Test

		t-test for Equality of Means		
		Sig. (2-tailed)	Mean Difference	Std. Error Difference
Exer_Intent Days Intent (# Days)	Equal variances assumed	.000	2.13333	.39441
	Equal variances not assumed	.000	2.13333	.39441

Independent Samples Test

		t-test for Equality of Means	
		95% Confidence Interval of the Difference	
		Lower	Upper
Exer_Intent Days Intent (# Days)	Equal variances assumed	1.32543	2.94124
	Equal variances not assumed	1.32516	2.94150

ANOUA

Explore

Notes

Output Created	06-SEP-2019 11:20:53	
Comments		
Input	Data	C:\Users\Bill\Desktop\current worklaaa Backup\Teaching\CLASSES\Psy 465-Adv State\SPSS-SAS- RIKIESS_-1.SAV
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	31
Missing Value Handling	Definition of Missing	User-defined missing values for dependent variables are treated as missing.
	Cases Used	Statistics are based on cases with no missing values for any dependent variable or factor used.
Syntax	EXAMINE VARIABLES=prod_dv BY grp /PLOT BOXPLOT HISTOGRAM NPLOT /COMPARE GROUPS /STATISTICS DESCRIPTIVES /INTERVAL 95 /MISSING LISTWISE /NOTOTAL.	
Resources	Processor Time	00:00:05.33
	Elapsed Time	00:00:04.98

grp

Case Processing Summary

	grp	Valid		Cases Missing		Total	
		N	Percent	N	Percent	N	Percent
prod_dv	1.00	10	100.0%	0	0.0%	10	100.0%
	2.00	10	100.0%	0	0.0%	10	100.0%
	3.00	10	100.0%	0	0.0%	10	100.0%

Descriptives

grp			Statistic	Std. Error	
prod_dv	1.00	Mean	92.0000	4.48206	
		95% Confidence Interval for Mean	Lower Bound		81.8609
		Upper Bound	102.1391		
		5% Trimmed Mean	91.7222		
		Median	90.5000		
		Variance	200.889		
		Std. Deviation	14.17353		
		Minimum	72.00		
		Maximum	117.00		
		Range	45.00		
		Interquartile Range	24.00		
		Skewness	.415	.687	
		Kurtosis	-.617	1.334	
	2.00	Mean	Mean	123.0000	4.96432
			95% Confidence Interval for Mean	Lower Bound	
Upper Bound			134.2301		
		5% Trimmed Mean	123.0000		
		Median	123.5000		
		Variance	246.444		
		Std. Deviation	15.69855		
		Minimum	98.00		
		Maximum	148.00		
		Range	50.00		
		Interquartile Range	24.50		
		Skewness	-.092	.687	
		Kurtosis	-.711	1.334	
3.00		Mean	Mean	106.0000	3.94687
			95% Confidence Interval for Mean	Lower Bound	
	Upper Bound		114.9284		
		5% Trimmed Mean	105.4444		
		Median	103.5000		
		Variance	155.778		
		Std. Deviation	12.48110		
		Minimum	92.00		
		Maximum	130.00		
		Range	38.00		
		Interquartile Range	19.75		
		Skewness	.875	.687	
		Kurtosis	-.084	1.334	

Tests of Normality

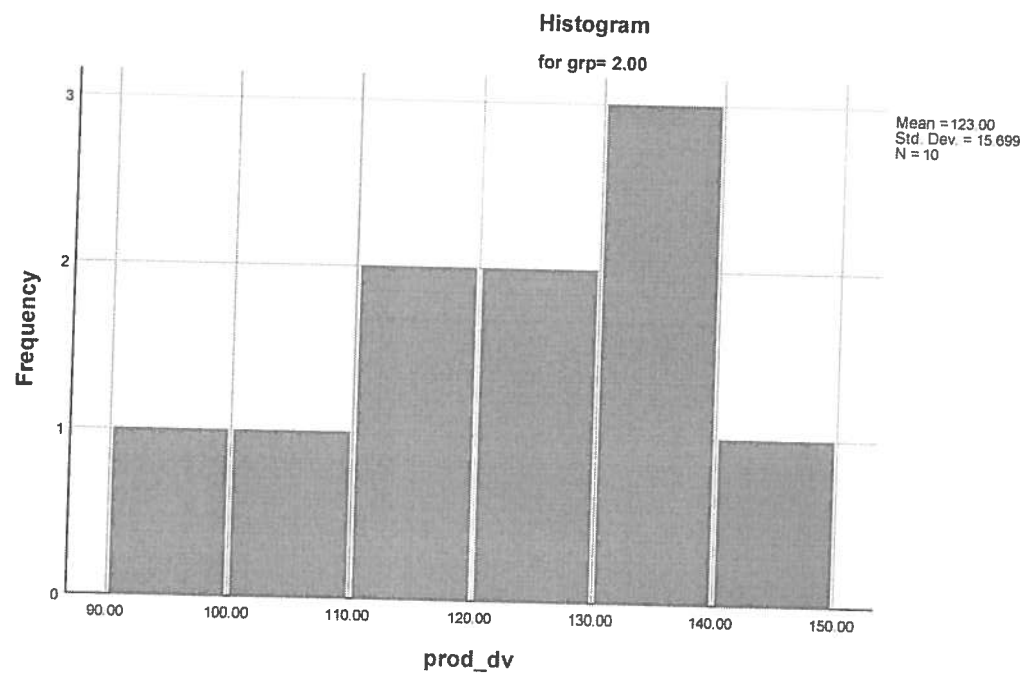
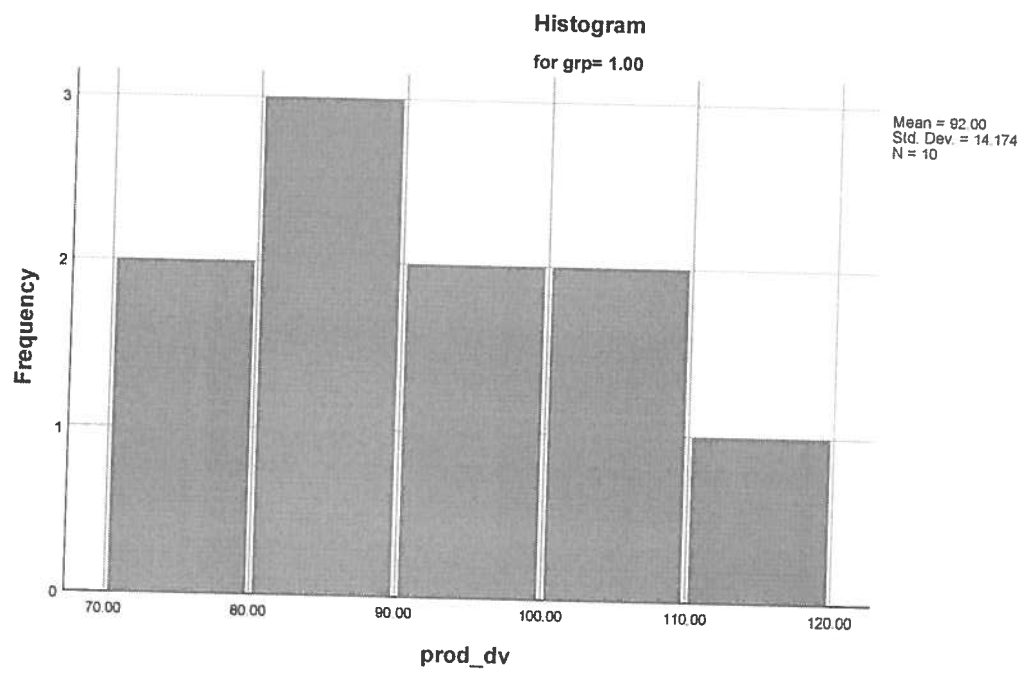
grp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
prod_dv 1.00	.114	10	.200 [*]	.975	10	.936
2.00	.117	10	.200 [*]	.985	10	.987
3.00	.164	10	.200 [*]	.921	10	.361

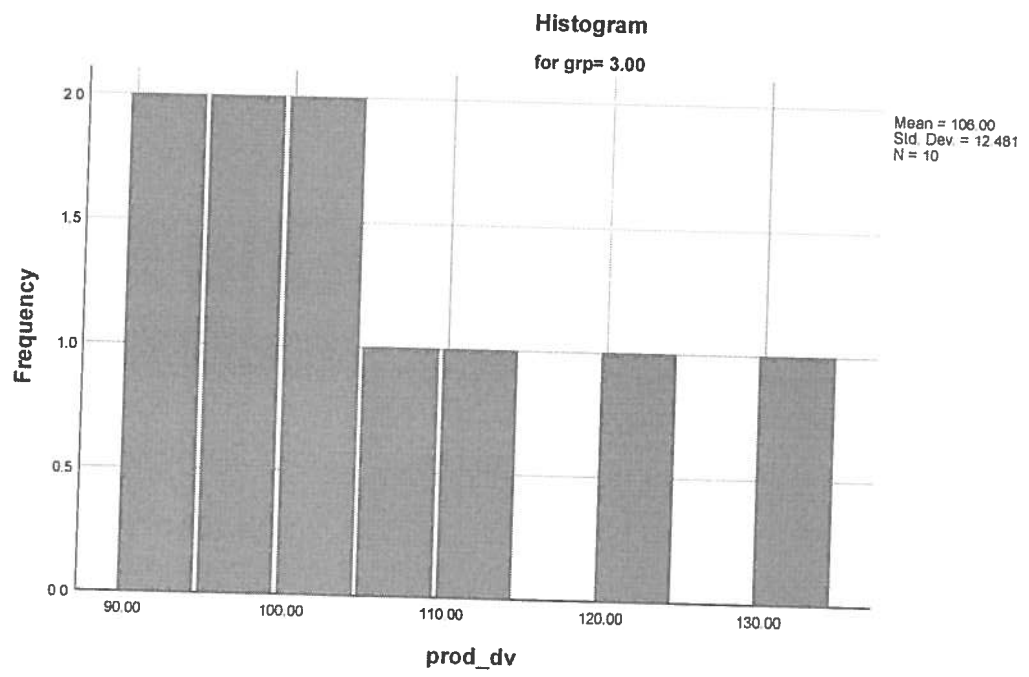
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

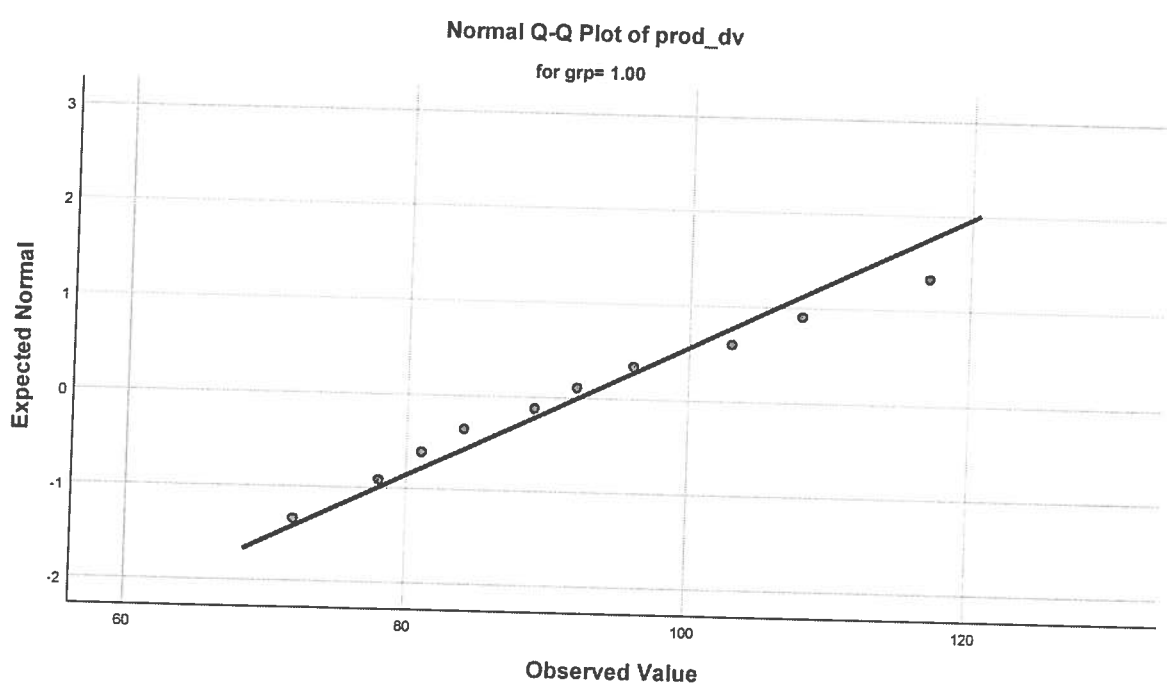
prod_dv

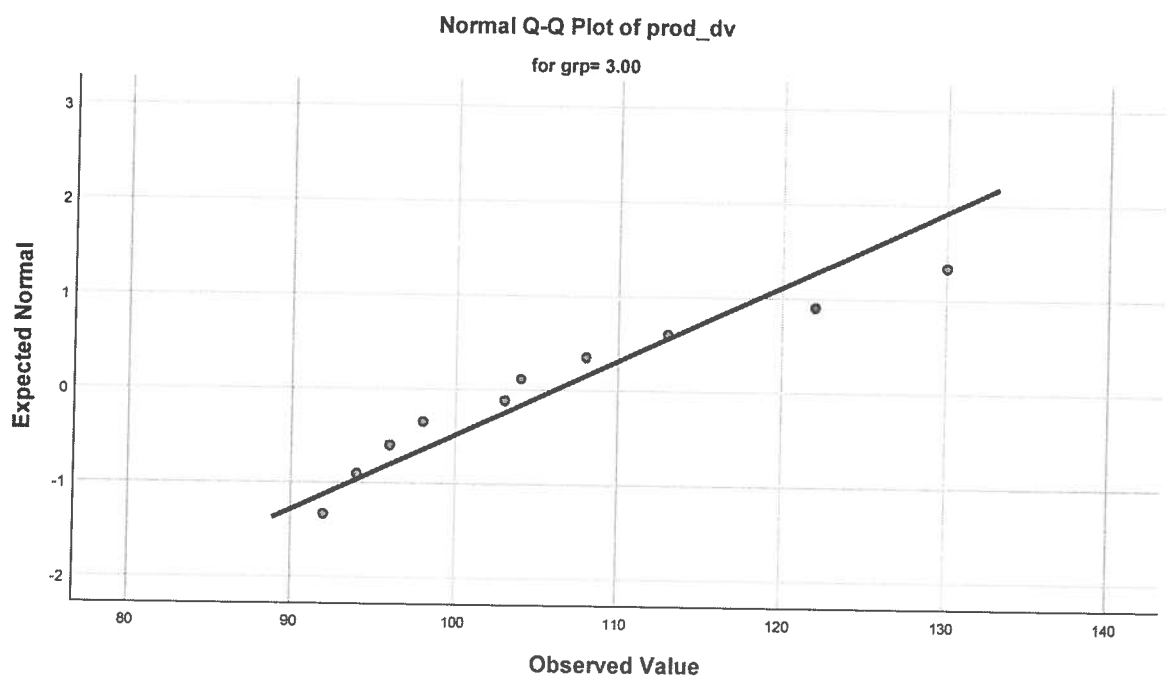
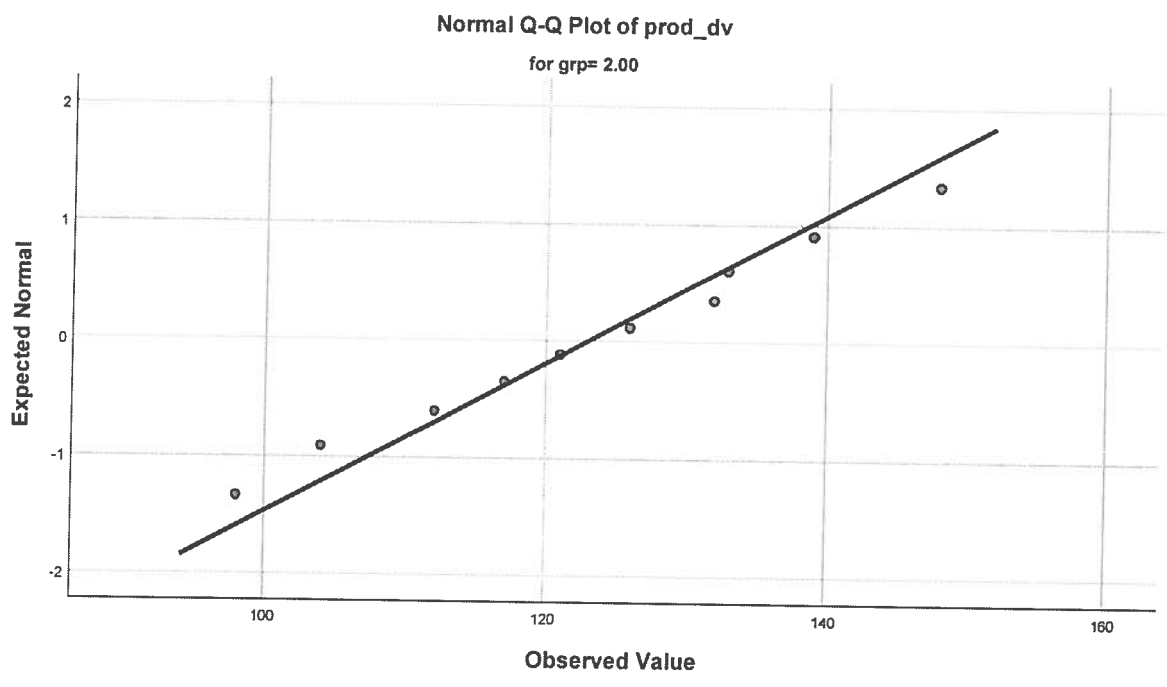
Histograms



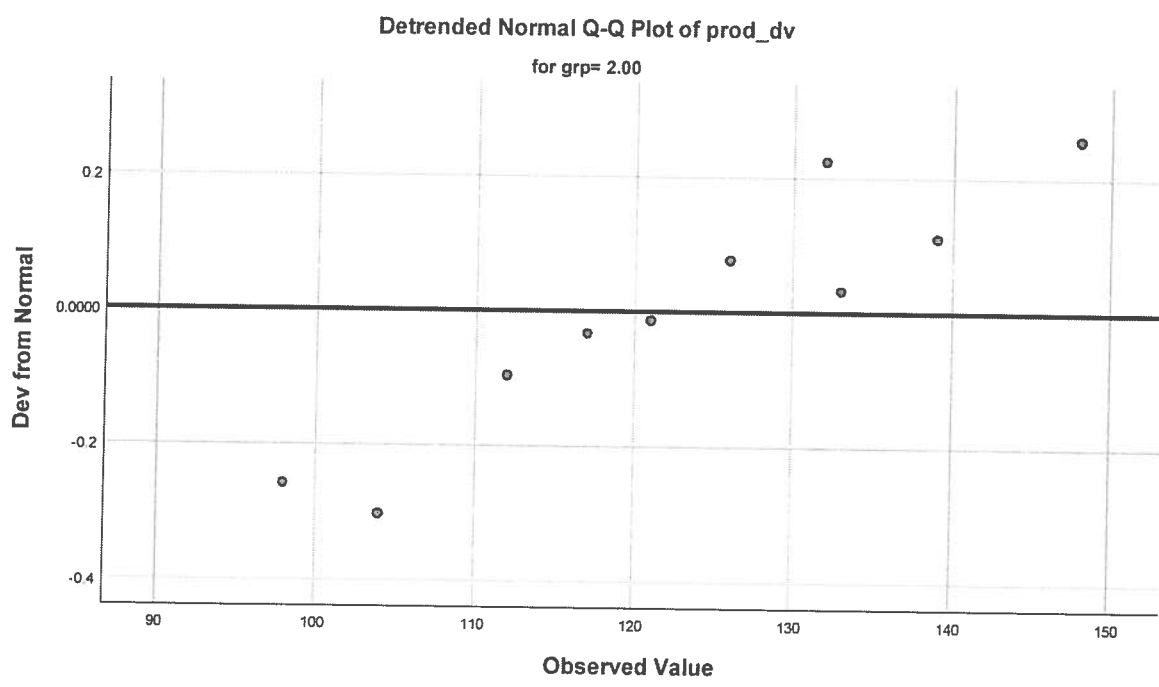
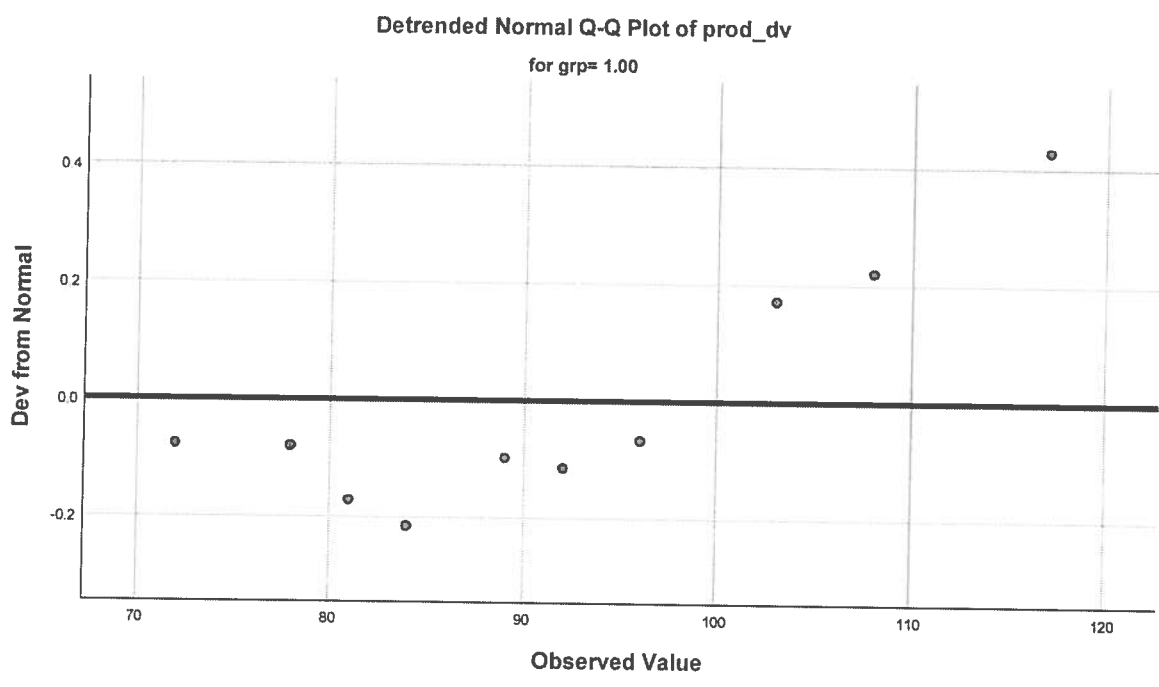


Normal Q-Q Plots

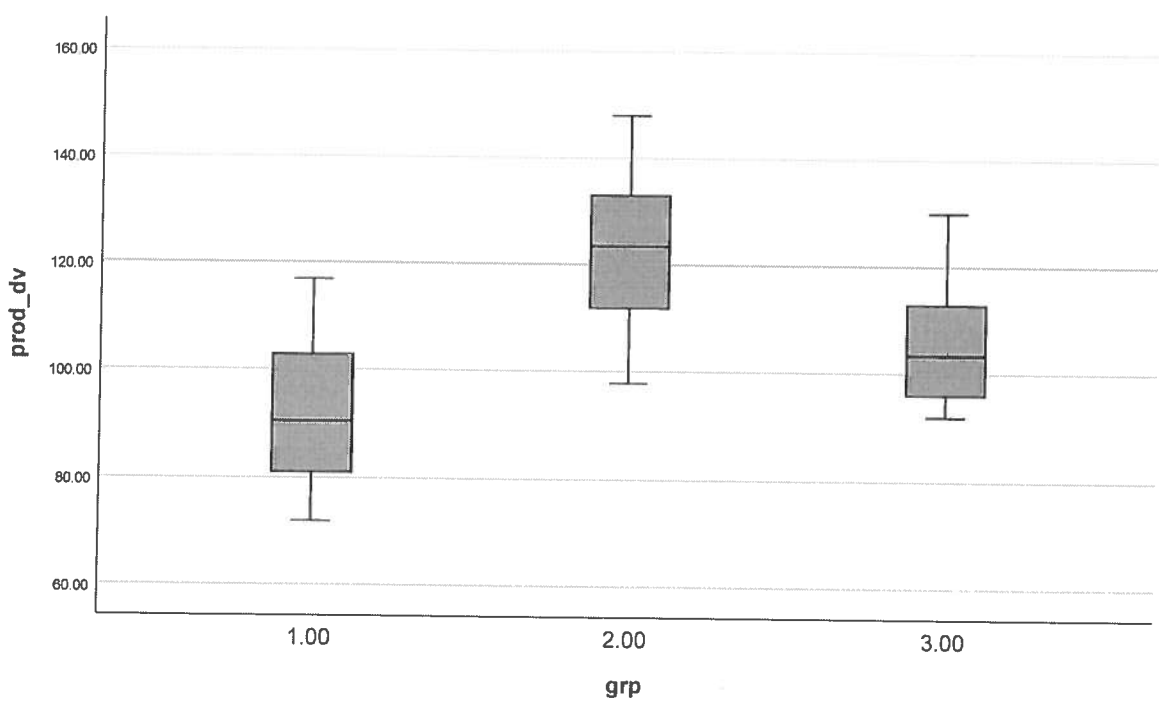
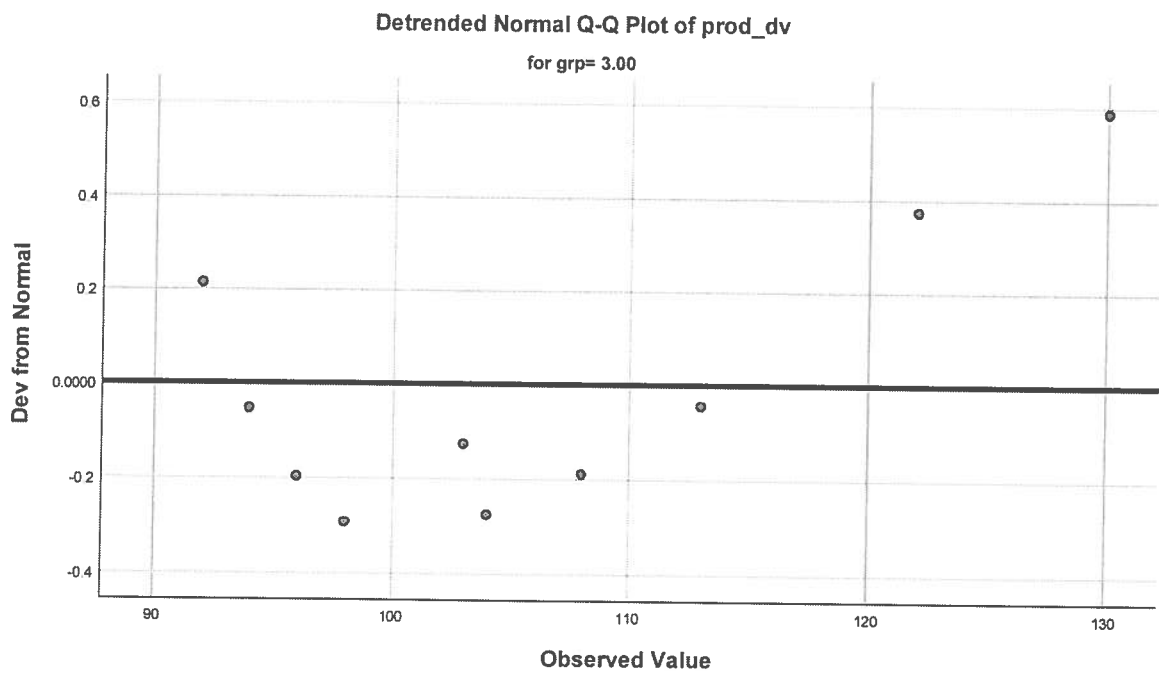




Detrended Normal Q-Q Plots



50



Oneway

Notes

Output Created	06-SEP-2019 11:24:35	
Comments		
Input	Data	C:\Users\Bill\Desktop\current worklaaa Backup\Teaching\CLASSES\Psy 465-Adv Stats\SPSS-SAS-R\KIESS_-1.SAV
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	31
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on cases with no missing data for any variable in the analysis.
Syntax	<pre> ONEWAY prod_dv BY grp /CONTRAST=-2 1 1 /CONTRAST=0 -1 1 /STATISTICS DESCRIPTIVES /PLOT MEANS /MISSING ANALYSIS /POSTHOC=TUKEY SCHEFFE LSD DUNNETT (1) ALPHA(0.05). </pre>	
Resources	Processor Time	00:00:00.87
	Elapsed Time	00:00:00.58

Descriptives

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	10	92.0000	14.17353	4.48206	81.8609	102.1391	72.00	117.00
2.00	10	123.0000	15.69855	4.96432	111.7699	134.2301	98.00	148.00
3.00	10	106.0000	12.48110	3.94687	97.0716	114.9284	92.00	130.00
Total	30	107.0000	18.79839	3.43210	99.9806	114.0194	72.00	148.00

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4820.000	2	2410.000	11.988	.000
Within Groups	5428.000	27	201.037		
Total	10248.000	29			

Contrast Coefficients

Contrast	grp		
	1.00	2.00	3.00
1	-2	1	1
2	0	-1	1

Contrast Tests

prod_dv	Assume equal variances	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
		1	45.0000	10.98281	4.097	27	.000
	Does not assume equal variances	2	-17.0000	6.34093	-2.681	27	.012
		1	45.0000	10.98079	4.098	17.908	.001
		2	-17.0000	6.34210	-2.681	17.130	.016

52

Post Hoc Tests

Multiple Comparisons

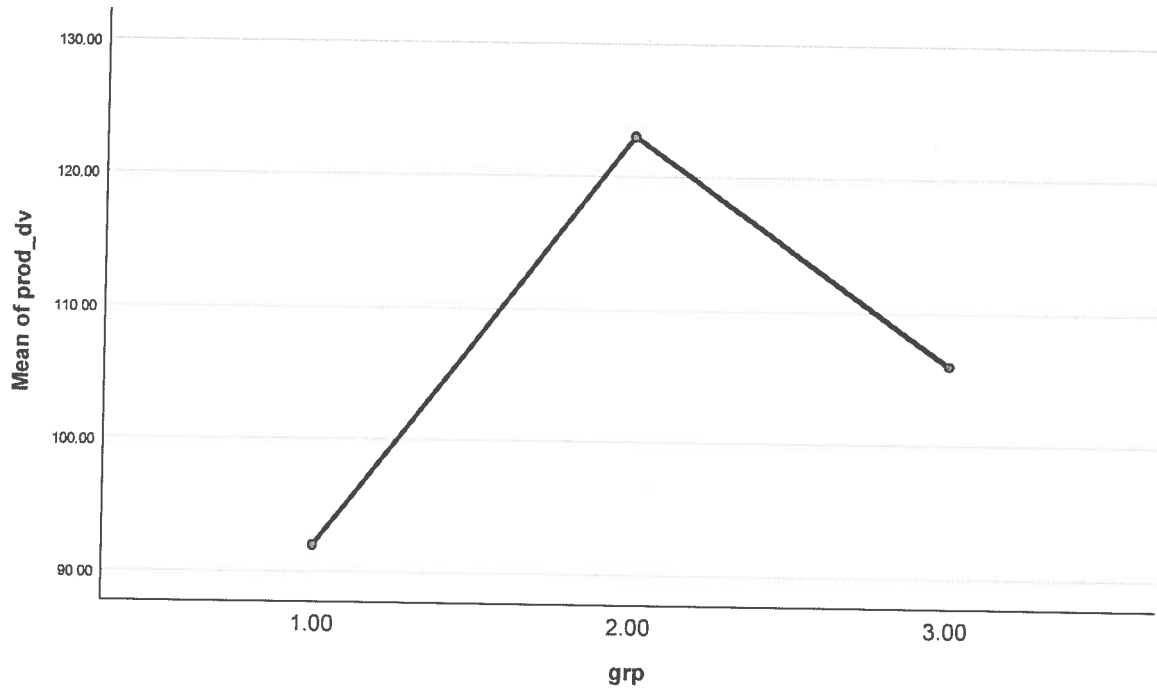
Dependent Variable: prod_dv

	(I) grp	(J) grp	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1.00	2.00	-31.0000*	6.34093	.000	-46.7218	-15.2782
		3.00	-14.0000	6.34093	.088	-29.7218	1.7218
	2.00	1.00	31.0000*	6.34093	.000	15.2782	46.7218
		3.00	17.0000*	6.34093	.032	1.2782	32.7218
	3.00	1.00	14.0000	6.34093	.088	-1.7218	29.7218
		2.00	-17.0000*	6.34093	.032	-32.7218	-1.2782
Scheffe	1.00	2.00	-31.0000*	6.34093	.000	-47.4232	-14.5768
		3.00	-14.0000	6.34093	.106	-30.4232	2.4232
	2.00	1.00	31.0000*	6.34093	.000	14.5768	47.4232
		3.00	17.0000*	6.34093	.041	.5768	33.4232
	3.00	1.00	14.0000	6.34093	.106	-2.4232	30.4232
		2.00	-17.0000*	6.34093	.041	-33.4232	-.5768
LSD	1.00	2.00	-31.0000*	6.34093	.000	-44.0105	-17.9895
		3.00	-14.0000*	6.34093	.036	-27.0105	-.9895
	2.00	1.00	31.0000*	6.34093	.000	17.9895	44.0105
		3.00	17.0000*	6.34093	.012	3.9895	30.0105
	3.00	1.00	14.0000	6.34093	.036	.9895	27.0105
		2.00	-17.0000*	6.34093	.012	-30.0105	-3.9895
Dunnnett t (2-sided) ^b	2.00	1.00	31.0000*	6.34093	.000	16.2040	45.7960
	3.00	1.00	14.0000	6.34093	.065	-.7960	28.7960

*. The mean difference is significant at the 0.05 level.

b. Dunnnett t-tests treat one group as a control, and compare all other groups against it.

Means Plots



```
ods listing;
ods html close;
ods graphics off;
data test;
input prod_dv grp;
datalines;
  72.00      1.00
  89.00      1.00
.....
 122.00      3.00
 130.00      3.00
  92.00      3.00
 108.00      3.00
  98.00      3.00
 104.00      3.00
;
run;
proc glm;
title 'analysis of Kiess data';
classes grp;
model prod_dv=grp;
means grp /tukey;
run;
```

The GLM Procedure
Dependent Variable: prod_dv

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	4820.00000	2410.00000	11.99	0.0002
Error	27	5428.00000	201.03704		
Corrected Total	29	10248.00000			

R-Square Coeff Var Root MSE prod_dv Mean
0.470336 13.25117 14.17875 107.0000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
grp	2	4820.000000	2410.000000	11.99	0.0002

Source	DF	Type III SS	Mean Square	F Value	Pr > F
grp	2	4820.000000	2410.000000	11.99	0.0002

The GLM Procedure

Tukey's Studentized Range (HSD) Test for prod_dv

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	27
Error Mean Square	201.037
Critical Value of Studentized Range	3.50633
Minimum Significant Difference	15.721

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	grp
A	123.000	10	2
B	106.000	10	3
B			
B	92.000	10	1

One-Way ANOVA Example in R

```

> library(foreign)
> kiess.spss<- read.spss("c:/data/kiess.sav", use.value.labels=TRUE)
> print(kiess.spss)
$DV
 [1] 72 89 117 78 108 96 103 84 92 81 117 98 139 121 148 104
132 133 112 126 103 96 113 94 122 130 92 108 98 104
$IV
 [1] 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3

attr("label.table")
attr("label.table")$DV
NULL

attr("label.table")$IV
NULL

attr("codepage")
[1] 1252

> attach(kiess.spss)
> fit1 <- aov(DV~factor(IV))
> summary(fit1)
          Df Sum Sq Mean Sq F value    Pr(>F)
factor(IV)  2   4820    2410  11.99 0.000188 ***
Residuals  27   5428     201
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> TukeyHSD(fit1)
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = DV ~ factor(IV))

$`factor(IV)`
      diff      lwr      upr      p adj
2-1    31  15.278183  46.721817 0.0001184
3-1    14  -1.721817  29.721817 0.0879066
3-2   -17 -32.721817 -1.278183 0.0320958

```