UNIT 4B: Supplemental Learn By Doing Materials

PART A - ANALYSES ARE PRODUCED FROM PULSE DATA which was an experiment, conducted on students in a statistics class. Students recorded their resting pulse and approximately half of students then ran for 1 minute where the other half continued sitting in their seats. Then each student measured their pulse a second time. Other variables were recorded. Information can be found in Topic 0B – Dataset Information.

[Case CQ – Two Dependent Samples] Paired T-Test comparing Pulse 2 to Pulse 1 for individuals who Sat

SAS Output

The TTEST Procedure

Difference: PULSE2 - PULSE1

Treatment = Sat

N	Mean	Std Dev	Std Err	Minimum	Maximum
62	-1.0323	3.9462	0.5012	-12.0000	8.0000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-1.0323	-2.0344	-0.0301	3.9462	3.3534	4.7957

DF	t Value	Pr > |t|
61	-2.06	0.0437

SPSS Output

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pulse after Treatment (bpm)	74.13	62	9.264	1.177
Resting Pulse (bpm)	75.16	62	10.309	1.309

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pulse after Treatment (bpm) - Resting Pulse (bpm)	-1.032	3.946	.501	-2.034	-.030	-2.060	61	.044

Question Set A1: Among individuals who sat, conduct a paired t-test to determine if the mean change in pulse rate differs from zero using the output above.

SHOW	Set up the hypotheses being tested and define the parameter used.	HIDE
SHOW	What is the value of the test statistic given in the output?	HIDE
SHOW	Are the conditions satisfied to conduct this test?	HIDE
SHOW	What is the value of the p-value given in the output?	HIDE
SHOW	What is your conclusion in context?	HIDE
SHOW	What type of error (Type I or Type II) could you have made and why? What would this error mean in context?	HIDE
SHOW	Interpret the confidence interval given in the output for the population mean difference in context. **NOTE: When statistically significant, confidence interval interpretations for differences should always clearly indicate which is larger/smaller and by how much (per the confidence interval estimate).**	HIDE
SHOW	Instructor Comments:	HIDE

[Case CQ – Two Dependent Samples] Paired T-Test comparing Pulse 2 to Pulse 1 for individuals who Ran

SAS Output

The TTEST Procedure

Difference: PULSE2 - PULSE1

Treatment = Ran

N	Mean	Std Dev	Std Err	Minimum	Maximum
44	52.4091	20.5984	3.1053	12.0000	94.0000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
52.4091	46.1466	58.6716	20.5984	17.0189	26.0987

DF	t Value	Pr > |t|
43	16.88	<.0001

SPSS Output

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pulse after Treatment (bpm)	126.00	44	25.310	3.816
Resting Pulse (bpm)	73.59	44	11.384	1.716

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper			
Pulse after Treatment (bpm) - Resting Pulse (bpm)	52.409	20.598	3.105	46.147	58.672	16.877	43	.000

Question Set A2: Among individuals who ran, conduct a paired t-test to determine if the mean change in pulse rate differs from zero using the output above.

Set up the hypotheses being tested and define the parameter used.

SHOW

HIDE

SHOW	What is the value of the test statistic given in the output?	HIDE
SHOW	Are the conditions satisfied to conduct this test?	HIDE
SHOW	What is the value of the p-value given in the output?	HIDE
SHOW	What is your conclusion in context?	HIDE
SHOW	Interpret the confidence interval given in the output for the population mean difference in context. NOTE: When statistically significant, confidence interval interpretations for differences should always clearly indicate which is larger/smaller and by how much.	HIDE
SHOW	What non-parametric test(s) would be appropriate as an alternative to the paired t-test in both Part A and Part B?	HIDE
SHOW	Instructor Comments:	HIDE

WHAT'S NEXT?

We have already conducted the primary tests of interest for this experiment, however, from this point we will investigate other questions in this data, some of which are discussed below.

- There is some concern that some students chose not to run even though they were randomized to do so (the instructor was not always aware of the assignment such as when the student tossed their own coin). You can see from the sample sizes in the two groups that **around 40% ran and 60% sat.** However, if the **distribution of all of the other variables is similar between the two treatment groups then** even if there was such a problem, it will be **less likely to introduce bias** into our results. **Thus we may investigate if the treatment variable – whether the student ran or sat – is associated with any of our other variables**. If there are associations this could cause reason for concern.
- Although we studied the change in pulse rate and not the actual pulse rates measured, it may be of interest to see if there the **resting (or baseline) pulse rate is associated with any of our other variables**. This could be due to random chance or the impact of a certain group of students being more likely to disregard their assignment to run and instead remain sitting. Thus we may use the resting pulse rate as the basis for some of our work.
- Otherwise, we will look at any results that illustrate additional possibilities of interest.

[Case CQ – Two Independent Samples] Two sample t-test

- **Comparing <u>between the treatment groups (ran or sat) for multiple variables simultaneously.</u>**

SAS and SPSS use different default tests for equality of variances so there is the possibility of different conclusions to the test for equality of variances and thus the possibility that different t-tests will be used to compare means. You should have the correct conclusion and use the correct tests based upon your software.

SPSS Output

If you have one primary binary variable and wish to run multiple t-tests for a number of other quantitative variables, you can do so in SPSS by putting multiple variables in the dependent variables box. You can only choose one dependent variable at a time. The results are combined as in the results that follow.

- In all of these tests there is a **large p-value for the test for equal variances** (first highlighted column). This **results in continuing to assume equal variances** and using the first row for the t-test results.
- The **t-tests (assuming equal variances as we just mentioned above)** also all have **p-values larger than 0.05** and thus our **conclusion**: We find <u>**no significant differences in the means between those who ran or sat**</u> for any of these variables (**resting pulse, age, BMI, weight, and height**).

Group Statistics

	Whether the student ran or sat	N	Mean	Std. Deviation	Std. Error Mean
Resting Pulse (bpm)	Ran	44	73.59	11.384	1.716
	Sat	62	75.16	10.309	1.309
Age (years)	Ran	44	20.18	2.863	.432
	Sat	62	20.90	4.626	.588
Body Mass Index (kg/sq. m)	Ran	44	22.1080	3.35056	.50512
	Sat	62	21.9409	3.32566	.42236
Weight (kg)	Ran	44	66.90	14.736	2.222
	Sat	62	66.44	15.116	1.920
Height (cm)	Ran	44	173.23	11.189	1.687
	Sat	62	173.21	9.534	1.211

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Resting Pulse (bpm)	Equal variances assumed	.503	.480	-.740	104	.461	-1.570	2.122	-5.779	2.638
	Equal variances not assumed			-.728	86.871	.469	-1.570	2.159	-5.861	2.720
Age (years)	Equal variances assumed	1.590	.210	-.917	104	.361	-.721	.787	-2.282	.839
	Equal variances not assumed			-.990	102.332	.325	-.721	.729	-2.167	.725
Body Mass Index (kg/sq. m)	Equal variances assumed	.344	.559	.254	104	.800	.16715	.65759	-1.13687	1.47118
	Equal variances not assumed			.254	92.332	.800	.16715	.65843	-1.14048	1.47479
Weight (kg)	Equal variances assumed	.186	.667	.157	104	.876	.462	2.949	-5.386	6.310
	Equal variances not assumed			.157	94.177	.875	.462	2.936	-5.367	6.292
Height (cm)	Equal variances assumed	.956	.331	.009	104	.993	.018	2.021	-3.989	4.025
	Equal variances not assumed			.008	83.160	.993	.018	2.076	-4.112	4.147

SAS Output

If you have one primary binary variable and wish to run multiple t-tests for a number of other quantitative variables, you can do so in SAS by putting multiple variables in the VAR statement in PROC TTEST. You can only choose one dependent variable at a time (using the CLASS statement). The results are not combined, we have removed the first table in the t-test results with descriptive statistics to minimize the needed output.

- In all of these tests except that for AGE, there is a **large p-value for the test for equal variances** (last table, labeled Equality of Variances). This **results in continuing to assume equal variances** and using the POOLED row for the t-test results. For AGE we use the SATTERTHWAITE row for UNEQUAL variances.
- The **t-tests (assuming equal variances for all tests except AGE as we just mentioned above)** all have **p-values larger than 0.05** and thus our **conclusion**: We find **no significant differences in the means between those who ran or sat** for any of these variables (**resting pulse, age, BMI, weight, and height**).

The TTEST Procedure

Variable: PULSE1 (First pulse measurement (bpm))

TRT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Ran		73.5909	70.1298	77.0520	11.3840	9.4057	14.4239
Sat		75.1613	72.5434	77.7792	10.3087	8.7599	12.5278
Diff (1-2)	Pooled	-1.5704	-5.7789	2.6381	10.7663	9.4809	12.4581
Diff (1-2)	Satterthwaite	-1.5704	-5.8608	2.7201			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	-0.74	0.4610
Satterthwaite	Unequal	86.871	-0.73	0.4689

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	43	61	1.22	0.4709

Variable: AGE (Age (years))

TRT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Ran		20.1818	19.3113	21.0524	2.8633	2.3658	3.6279
Sat		20.9032	19.7284	22.0780	4.6260	3.9310	5.6219
Diff (1-2)	Pooled	-0.7214	-2.2822	0.8393	3.9927	3.5161	4.6201
Diff (1-2)	Satterthwaite	-0.7214	-2.1674	0.7246			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	-0.92	0.3615
Satterthwaite	Unequal	102.33	-0.99	0.3247

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	61	43	2.61	0.0012

Variable: BMI (Body Mass Index)

TRT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Ran		22.1080	21.0894	23.1267	3.3506	2.7683	4.2452
Sat		21.9409	21.0963	22.7854	3.3257	2.8260	4.0416
Diff (1-2)	Pooled	0.1672	-1.1369	1.4712	3.3360	2.9377	3.8602
Diff (1-2)	Satterthwaite	0.1672	-1.1405	1.4748			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	0.25	0.7998
Satterthwaite	Unequal	92.332	0.25	0.8002

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	43	61	1.02	0.9450

Variable: WEIGHT (Weight (kg))

TRT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Ran		66.8977	62.4175	71.3780	14.7363	12.1754	18.6712
Sat		66.4355	62.5968	70.2742	15.1158	12.8448	18.3698
Diff (1-2)	Pooled	0.4622	-5.3856	6.3101	14.9601	13.1740	17.3108
Diff (1-2)	Satterthwaite	0.4622	-5.3673	6.2918			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	0.16	0.8757
Satterthwaite	Unequal	94.177	0.16	0.8752

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	61	43	1.05	0.8704

Variable: HEIGHT (Height (cm))

TRT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Ran		173.2	169.8	176.6	11.1894	9.2449	14.1773
Sat		173.2	170.8	175.6	9.5336	8.1013	11.5859
Diff (1-2)	Pooled	0.0176	-3.9894	4.0246	10.2507	9.0269	11.8614
Diff (1-2)	Satterthwaite	0.0176	-4.1122	4.1474			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	0.01	0.9931
Satterthwaite	Unequal	83.16	0.01	0.9933

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	43	61	1.38	0.2476

[Case CQ – Two Independent Samples] Two sample t-test

- **Comparing Body Mass Index between Males and Females**

In the results below, notice that we would **reject the null hypotheses of equal variances.**

- The test for equal variances has a **p-value of 0.035 from SPSS or 0.0466 from SAS**.
- This leads to looking in the **row for the t-test results for which we DO NOT assume equal variances (labeled as SATTERTHWAITE in SAS)**.

In this analysis, the p-value for both t-tests are given as 0.000 by SPSS but **there are differences in the t-value, df, std. error, and confidence interval values between the equal and unequal variances rows for the t-test results**. In SAS the p-values of the two tests are slightly different and we would use 0.0001 from the SATTERTHWAITE row.

Here the p-value of the t-test is small so we reject the null hypothesis that the mean BMI is equal between males and females.

T-test Conclusion: There is enough evidence that the population mean BMI differs between males and females.

Confidence Interval interpretation: The 95% confidence interval estimates that the population mean BMI for males is between 1.18276 and 3.54222 larger than the population mean BMI for females.

Note: We know that males are larger than females because of the order used and the sign of the resulting confidence interval values. In both SAS and SPSS we see that Males are listed first and Females second. We have constructed the confidence interval for the difference in means as Males – Females. The confidence interval values are both positive and thus Males must be larger than Females for the difference to be positive.

SAS Output

Variable: BMI (Body Mass Index)

GENDER	N	Mean	Std Dev	Std Err	Minimum	Maximum
Male	57	23.1024	3.4847	0.4616	16.7968	32.1402
Female	49	20.7399	2.6261	0.3752	16.5889	29.0006
Diff (1-2)		2.3625	3.1179	0.6074		

GENDER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Male		23.1024	22.1778	24.0270	3.4847	2.9420	4.2748
Female		20.7399	19.9856	21.4942	2.6261	2.1900	3.2808
Diff (1-2)	**Pooled**	2.3625	1.1580	3.5670	3.1179	2.7457	3.6079
Diff (1-2)	**Satterthwaite**	2.3625	1.1828	3.5422			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	3.89	0.0002
Satterthwaite	Unequal	102.33	3.97	0.0001

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	56	48	1.76	0.0466

SPSS Output

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Body Mass Index (kg/sq. m)	Male	57	23.1024	3.48470	.46156
	Female	49	20.7399	2.62611	.37516

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Body Mass Index (kg/sq. m)	Equal variances assumed	4.571	.035	3.889	104	.000	2.36249	.60742	1.15796	3.56702
	Equal variances not assumed			3.972	102.329	.000	2.36249	.59480	1.18276	3.54222

[Case CQ – Two Independent Samples] Two sample t-test

- **Comparing Resting Pulse Rates between Males and Females**

SAS Output

The TTEST Procedure
Variable: PULSE1 (First pulse measurement (bpm))

GENDER	N	Mean	Std Dev	Std Err	Minimum	Maximum
Male	57	72.6667	10.0380	1.3296	49.0000	92.0000
Female	49	76.6531	11.2334	1.6048	47.0000	104.0
Diff (1-2)		-3.9864	10.6065	2.0663		

GENDER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Male		72.6667	70.0032	75.3301	10.0380	8.4746	12.3141
Female		76.6531	73.4264	79.8797	11.2334	9.3678	14.0339
Diff (1-2)	**Pooled**	-3.9864	-8.0839	0.1111	10.6065	9.3402	12.2731
Diff (1-2)	**Satterthwaite**	-3.9864	-8.1224	0.1496			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	104	-1.93	0.0564
Satterthwaite	Unequal	97.241	-1.91	0.0587

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	48	56	1.25	0.4158

SPSS Output

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Resting Pulse (bpm)	Male	57	72.67	10.038	1.330
	Female	49	76.65	11.233	1.605

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Resting Pulse (bpm)	Equal variances assumed	.431	.513	-1.929	104	.056	-3.986	2.066	-8.084	.111
	Equal variances not assumed			-1.913	97.241	.059	-3.986	2.084	-8.122	.150

Question Set A3: Conduct a two-independent samples t-test comparing the population mean resting pulse rates between males and females using the output above.

SHOW	Set up the hypotheses being tested and define the parameters involved.	HIDE
SHOW	Are the conditions satisfied to conduct this test?	HIDE
SHOW	Should we assume equal variances?	HIDE
SHOW	What is the value of the appropriate p-value **OF THE T-TEST**?	HIDE
SHOW	What is your conclusion **OF THE T-TEST** in context?	HIDE
SHOW	Interpret the appropriate confidence interval given in the output for the difference in population means in context.	HIDE
SHOW	What type of error (Type I or Type II) could you have made and why? What would this error mean in context?	HIDE
SHOW	What non-parametric test(s) would be appropriate?	HIDE

[Case CQ – Two Independent Samples] Two sample t-test

- **Comparing <u>CHANGE in Pulse Rate (After – Before) vs. Treatment (Ran or Sat)</u>**

We certainly expect a difference in the means and we also find a significant difference in the variation. Both make biological sense. I created a **very small sample of size 17** for this t-test. Since there is such a huge effect on pulse rates from running vs. sitting, we will still find significance but **we will need to investigate whether the distribution of the change in pulse rate is normally distributed in our two treatment groups or not**.

The SAS output is not available for this example. The only difference is the location of the results in the output. The processes are the same.

SPSS Output

Group Statistics

	Whether the student ran or sat	N	Mean	Std. Deviation	Std. Error Mean
Change in Pulse (After - Before)	Ran	6	39.3333	17.90717	7.31057
	Sat	11	-.5455	4.45788	1.34410

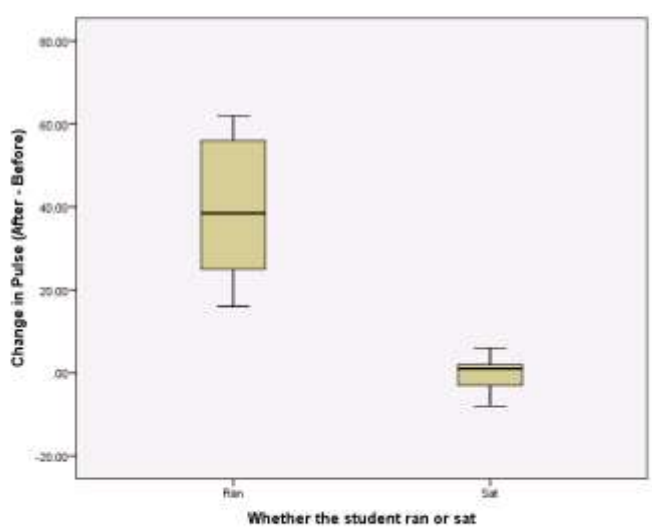
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Change in Pulse (After - Before)	Equal variances assumed	19.400	.001	7.169	15	.000	39.87879	5.56278	28.02201	51.73556
	Equal variances not assumed			5.365	5.341	.002	39.87879	7.43311	21.13225	58.62533

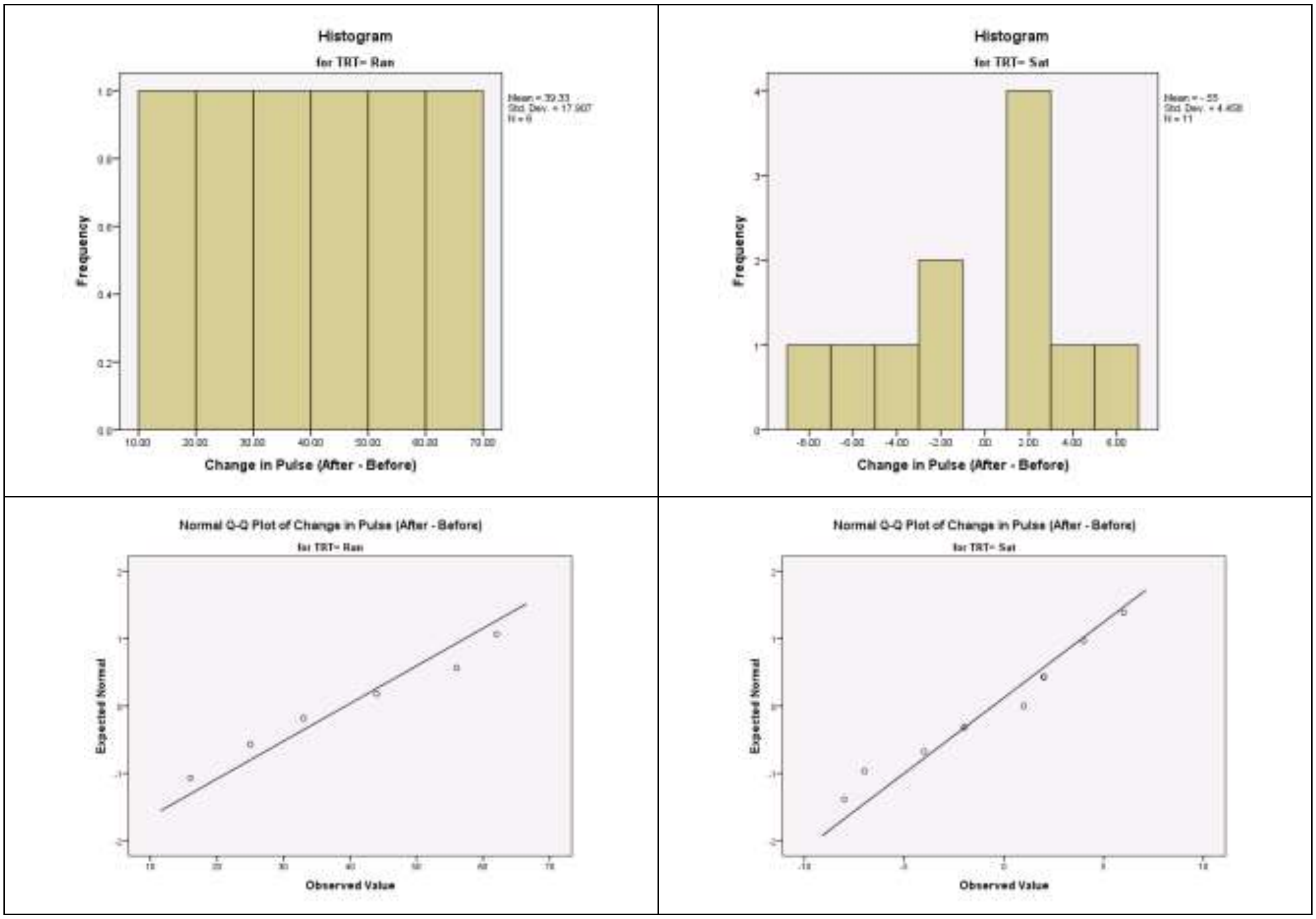
Since the sample sizes are small in our two groups, we will need to investigate the assumption of normality of the distributions of the differences our two groups.

Investigate the Distribution of Change in Pulse Rate between the Two Groups Graphically

- **Side-by-Side Boxplots:**
- Clearly shows both the difference in means and the difference in variation between the two treatment groups.



- **Histograms and QQ-Plots for each group:** Do you see any clear concerns given the small sample sizes?



Our Answer: Neither distribution is clearly normally distributed however with such a small sample size, we have no clear evidence of skewness or extreme outliers so we are left without a very clear answer here. In practice we would likely continue with the t-test with caution. Alternatively, we could always use a non-parametric alternative where we would not have the same assumption of normality.

Question Set A4: Conduct a two-independent samples t-test comparing the population mean change in pulse rate after 1 minute between those who ran and those who sat.

SHOW

Set up the hypotheses being tested and define the parameters involved.

HIDE

SHOW

Are the conditions satisfied to conduct this test?

HIDE

SHOW

Should we assume equal variances?

HIDE

SHOW

What is the value of the appropriate p-value **OF THE T-TEST**?

HIDE

SHOW

What is your conclusion **OF THE T-TEST** in context?

HIDE

SHOW

Interpret the appropriate confidence interval given in the output for the difference in population means in context.

NOTE: When statistically significant, confidence interval interpretations for differences should always clearly indicate which is larger/smaller and by how much.

HIDE

SHOW

What non-parametric test(s) would be appropriate?

HIDE

SHOW

Instructor Comments:

HIDE

[Case CQ – More than Two Independent Samples] One-Way ANOVA (Analysis of Variance)

- **Comparing Resting Pulse Rate among Exercise Groups (3-Levels)**

SPSS Output

This first table comes from requesting descriptives. From this we can see the sample sizes, how the sample means compare between the groups, and how the standard deviations compare.

Descriptives

Resting Pulse (bpm)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
High	14	68.64	12.689	3.391	61.32	75.97	49	96
Moderate	56	74.23	10.657	1.424	71.38	77.09	47	104
Low	36	77.22	9.302	1.550	74.07	80.37	52	92
Total	106	74.51	10.743	1.043	72.44	76.58	47	104

This ANOVA TABLE is the default output and contains the main results required.

ANOVA

Resting Pulse (bpm)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	751.072	2	375.536	3.403	.037
Within Groups	11367.419	103	110.363		
Total	12118.491	105			

SAS Output

The GLM Procedure
Dependent Variable: PULSE1 First pulse measurement (bpm)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	751.07192	375.53596	3.40	0.0371
Error	103	11367.41865	110.36329		
Corrected Total	105	12118.49057			

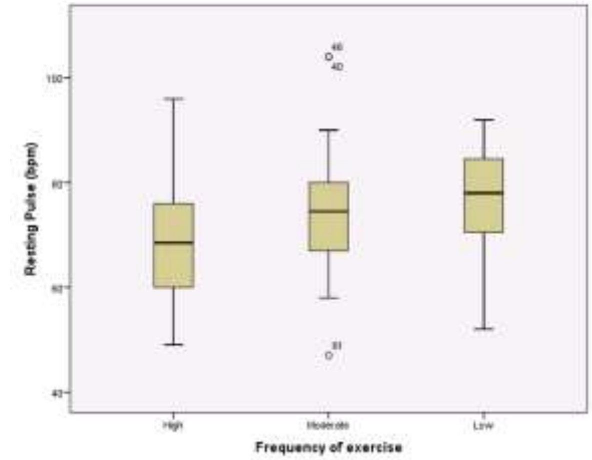
R-Square	Coeff Var	Root MSE	PULSE1 Mean
0.061977	14.09941	10.50539	74.50943

Source	DF	Type I SS	Mean Square	F Value	Pr > F
EXERCISE	2	751.0719152	375.5359576	3.40	0.0371

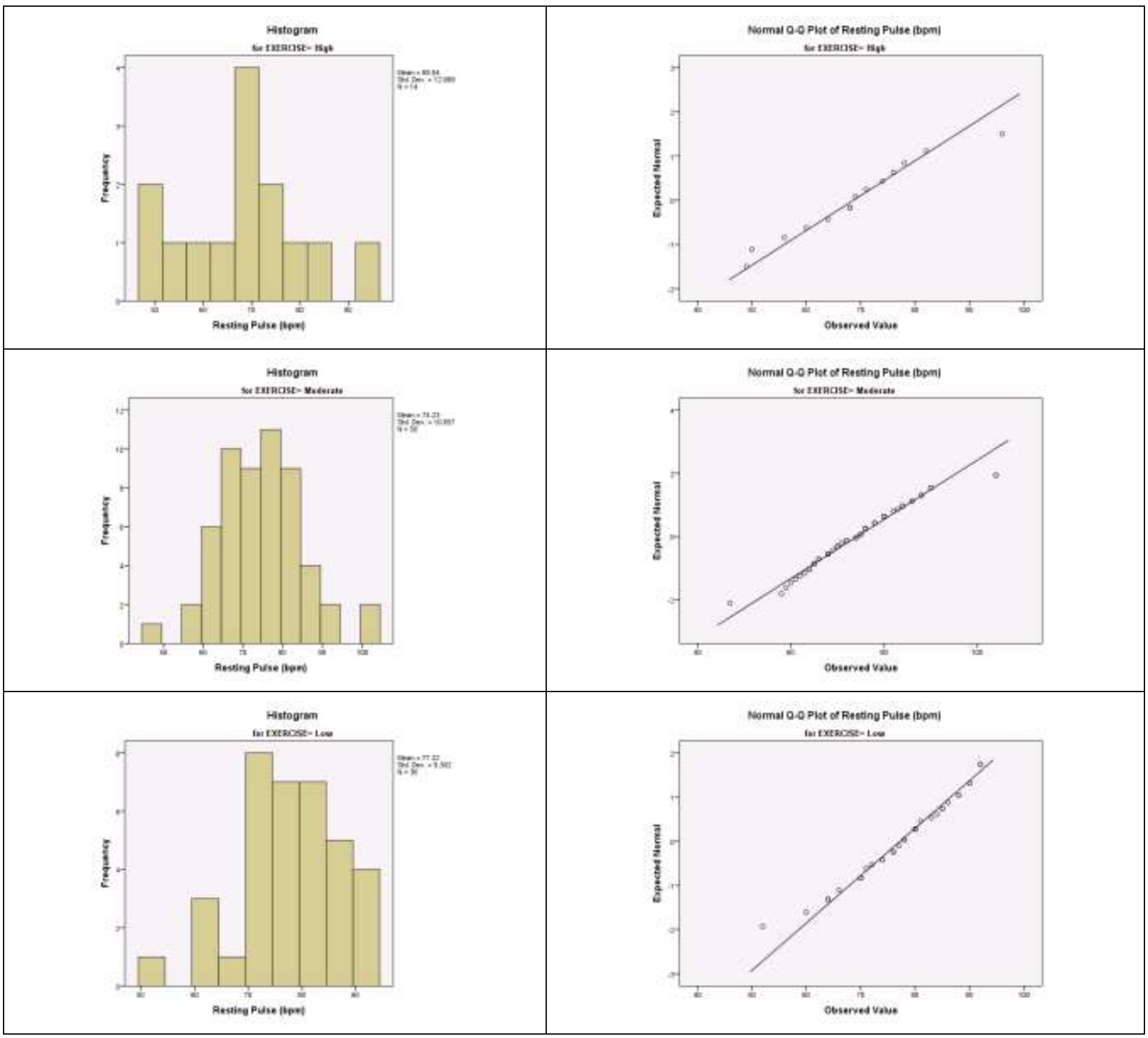
Source	DF	Type III SS	Mean Square	F Value	Pr > F
EXERCISE	2	751.0719152	375.5359576	3.40	0.0371

Investigate the Distribution between the groups graphically: Since the sample sizes in the groups are not all larger than 30, we will investigate the distributions.

- **Side-by-Side Boxplots:** Notice the order is alphabetical (the x-axis goes from High to Low) but the trend makes sense. The less you exercise, the higher your resting pulse rate. The variation seems similar between the boxplots so there is not immediate cause for concern about the assumption of equal variances required for the one-way ANOVA test. We do have two outliers in the moderate exercise group.



- **Histograms and QQ-Plots for each group:** Do you see any clear concerns given the small sample sizes?



Question Set A5: Conduct a one-way ANOVA test comparing the population mean resting pulse rate between the three exercise groups (Low, Moderate, High)

SHOW HIDE

Set up the hypotheses being tested:

SHOW HIDE

Are the conditions satisfied to conduct this test?

SHOW HIDE

What is the value of the p-value given in the output?

SHOW HIDE

What is your conclusion in context?

SHOW HIDE

What type of error (Type I or Type II) could you have made and why? What would this error mean in context?

SHOW HIDE

What non-parametric test(s) would be appropriate?

[Case CQ – More than Two Independent Samples] One-Way ANOVA (Analysis of Variance)

- **Comparing Age between the weight groups (5-Levels)**

SPSS Output

This first table comes from requesting descriptives.

Descriptives

Age (years)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
<= 55	23	20.13	3.429	.715	18.65	21.61	18	34
56 - 60	26	19.69	1.436	.282	19.11	20.27	18	23
61 - 67	16	20.31	2.626	.656	18.91	21.71	18	28
68 - 79	21	21.76	5.214	1.138	19.39	24.14	18	41
80+	20	21.35	5.833	1.304	18.62	24.08	18	45
Total	106	20.60	3.990	.388	19.84	21.37	18	45

This ANOVA TABLE is the default output and contains the main results required.

ANOVA

Age (years)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	67.414	4	16.854	1.061	.380
Within Groups	1603.944	101	15.881		
Total	1671.358	105			

SAS Output

Dependent Variable: AGE Age (years)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	67.414310	16.853577	1.06	0.3797
Error	101	1603.944181	15.880635		
Corrected Total	105	1671.358491			

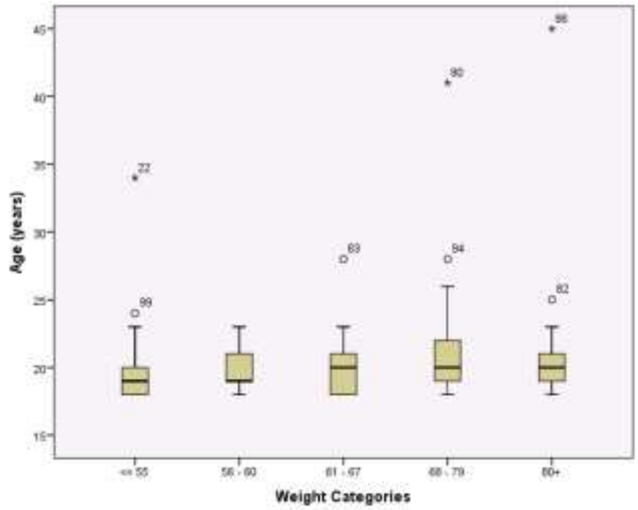
R-Square	Coeff Var	Root MSE	AGE Mean
0.040335	19.34137	3.985051	20.60377

Source	DF	Type I SS	Mean Square	F Value	Pr > F
WTGroups	4	67.41430957	16.85357739	1.06	0.3797

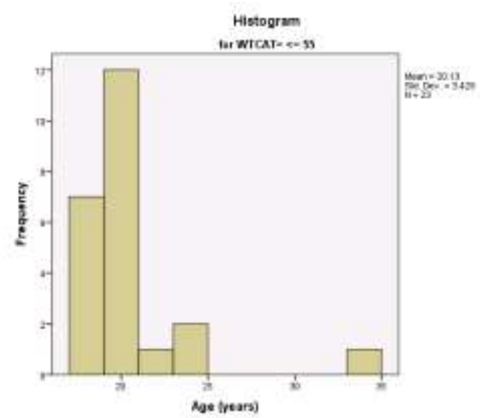
Source	DF	Type III SS	Mean Square	F Value	Pr > F
WTGroups	4	67.41430957	16.85357739	1.06	0.3797

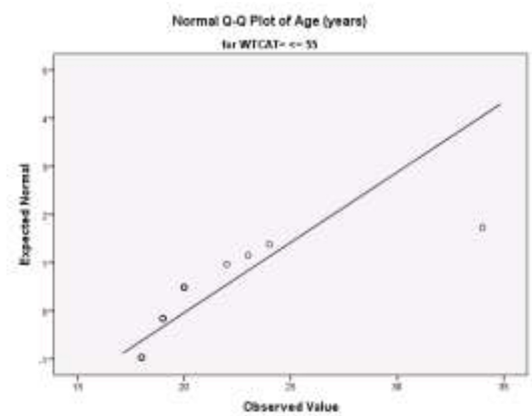
Investigate the Distribution between the groups graphically: Since the sample sizes in the groups are all smaller than 30, we will investigate the distributions.

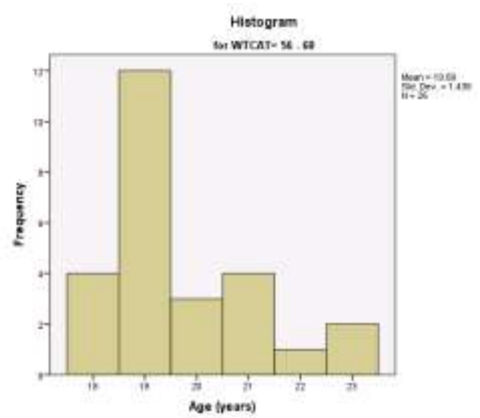
- **Side-by-Side Boxplots:** Notice that generally we have skewness and/or outliers in these distributions.

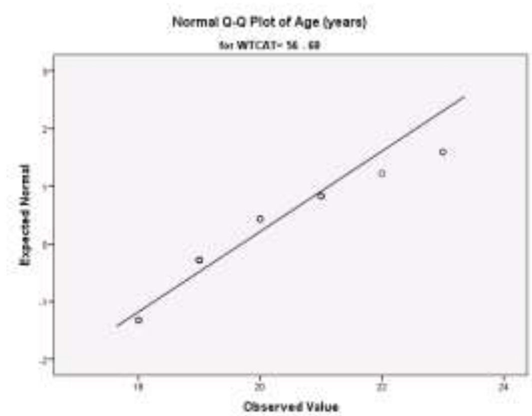


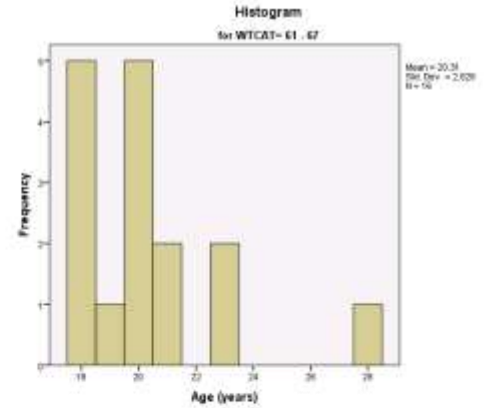
- **Histograms and QQ-Plots for each group:** Do you see any clear concerns given the small sample sizes?

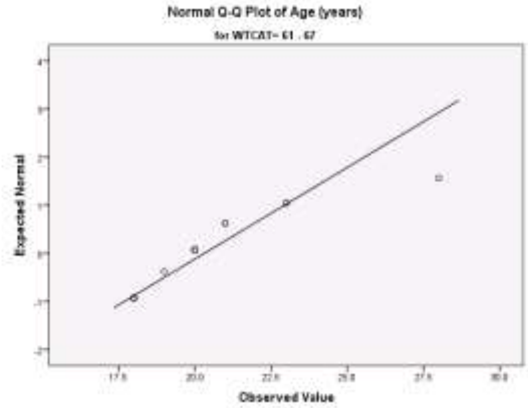


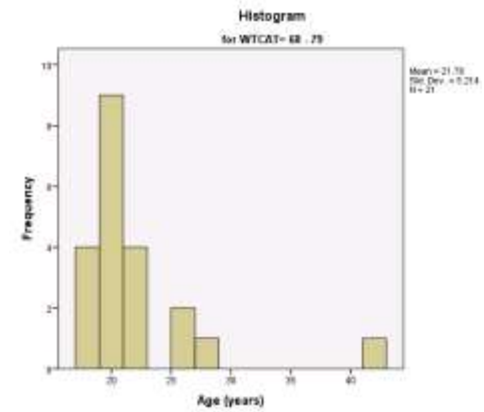


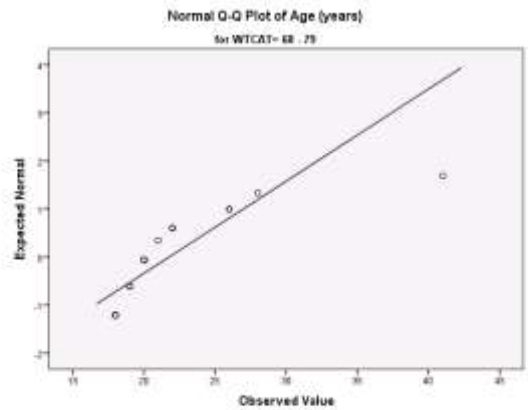


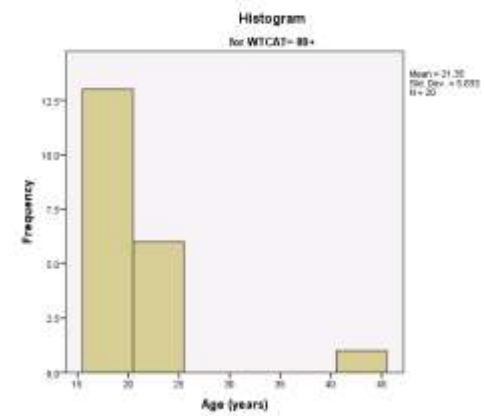


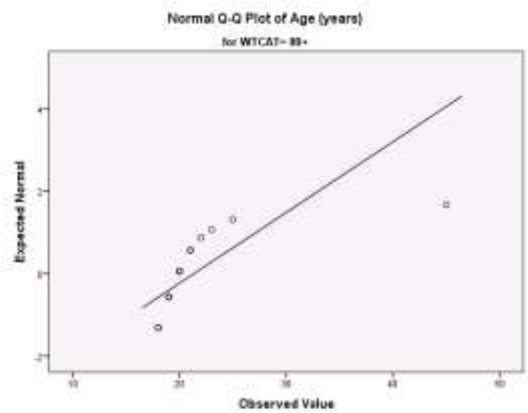

Histogram
for WTCAT= 61 . 67
Frequency
Age (years)


Normal Q-Q Plot of Age (years)
for WTCAT= 61 . 67
Expected Normal
Observed Value


Histogram
for WTCAT= 68 . 79
Frequency
Age (years)


Normal Q-Q Plot of Age (years)
for WTCAT= 68 . 79
Expected Normal
Observed Value


Histogram
for WTCAT= 80+
Frequency
Age (years)


Normal Q-Q Plot of Age (years)
for WTCAT= 80+
Expected Normal
Observed Value

Question Set A6: Conduct a one-way ANOVA test comparing the population mean age between the five weight groups.

Set up the hypotheses being tested:
Are the conditions satisfied to conduct this test?
What is the value of the p-value given in the output?
What non-parametric test(s) would be appropriate?

SHOW HIDE

SHOW HIDE

SHOW HIDE

SHOW HIDE

[Case CQ – More than Two Independent Samples] Kruskal-Wallis (Non-Parametric Analysis of Variance)

Since there are concerns with the standard one-way ANOVA, we could use the Kruskal-Wallis non-parametric test. Here, the p-value is 0.44 resulting in the same conclusion. We would fail to reject the null hypothesis.

Conclusion: There is not enough evidence of an association between age and the multi-level weight groups.

SPSS Output

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Age (years) is the same across categories of Weight Categories.	Independent-Samples Kruskal-Wallis Test	.439	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

SAS Output

The NPAR1WAY Procedure

Kruskal-Wallis Test	
Chi-Square	3.7600
DF	4
Pr > Chi-Square	0.4395

[Case CC] Chi-Square Test (Pearson's or Continuity Corrected)

- **Is there an association between the treatment (ran or sat) and gender?**

SPSS Output

Crosstab

			Gender		Total
			Male	Female	
Whether the student ran or sat	Ran	Count	22	22	44
		% within Whether the student ran or sat	50.0%	50.0%	100.0%
	Sat	Count	35	27	62
		% within Whether the student ran or sat	56.5%	43.5%	100.0%
Total		Count	57	49	106
		% within Whether the student ran or sat	53.8%	46.2%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.431[a]	1	.512		
Continuity Correction[b]	.210	1	.646		
Likelihood Ratio	.431	1	.512		
Fisher's Exact Test				.557	.323
Linear-by-Linear Association	.427	1	.514		
N of Valid Cases	106				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 20.34.

b. Computed only for a 2x2 table

Note: The appropriate Fisher's Exact test p-value is highlighted in green for those interested in learning more about this test. This result is automatically generated for 2x2 tables in some versions of SPSS.

SAS Output

Table of TRT by GENDER

TRT(Whether the student ran or sat)	GENDER(Gender)		
Frequency Percent Row Pct Col Pct	Male	Female	Total
Ran	22 20.75 50.00 38.60	22 20.75 50.00 44.90	44 41.51
Sat	35 33.02 56.45 61.40	27 25.47 43.55 55.10	62 58.49
Total	57 53.77	49 46.23	106 100.00

Statistic	DF	Value	Prob
Chi-Square	1	0.4309	0.5115
Likelihood Ratio Chi-Square	1	0.4308	0.5116
Continuity Adj. Chi-Square	1	0.2105	0.6464
Mantel-Haenszel Chi-Square	1	0.4269	0.5135
Phi Coefficient		-0.0638	
Contingency Coefficient		0.0636	
Cramer's V		-0.0638	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	22
Left-sided Pr <= F	0.3231
Right-sided Pr >= F	0.8035
Table Probability (P)	0.1266
Two-sided Pr <= P	0.5568

Sample Size = 106

Question Set A7: Conduct the appropriate chi-square test for independence to test for an association between the treatment (ran or sat) and gender?

Set up the hypotheses being tested:
Do you have any concerns about using the chi-square test? Explain.
What is the p-value of the appropriate chi-square test given in the output?
What is your conclusion in context?
Compare the distribution of gender between the treatment groups using the appropriate conditional percentages.

SHOW HIDE

SHOW HIDE

SHOW HIDE

SHOW HIDE

SHOW HIDE

[Case CC] Chi-Square Test (Pearson's or Continuity Corrected)

- **Is there an association between gender and our binary body mass index variable?**

SPSS Output

Crosstab

			Biinary Body Mass Index		Total
			Normal and Underweight	Overweight or Obese	
Gender	Male	Count	41	16	57
		% within Gender	71.9%	28.1%	100.0%
	Female	Count	46	3	49
		% within Gender	93.9%	6.1%	100.0%
Total		Count	87	19	106
		% within Gender	82.1%	17.9%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.627[a]	1	.003		
Continuity Correction[b]	7.200	1	.007		
Likelihood Ratio	9.449	1	.002		
Fisher's Exact Test				.004	.003
Linear-by-Linear Association	8.546	1	.003		
N of Valid Cases	106				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 8.78.

b. Computed only for a 2x2 table

Note: The appropriate Fisher's Exact test p-value is highlighted in green for those interested in learning more about this test. This result is automatically generated for 2x2 tables in some versions of SPSS.

SAS Output

Table of GENDER by BinaryBMI

GENDER(Gender)	BinaryBMI(Binary Body Mass Index)		
Frequency Percent Row Pct Col Pct	< 25	25+	Total
Male	41 38.68 71.93 47.13	16 15.09 28.07 84.21	57 53.77
Female	46 43.40 93.88 52.87	3 2.83 6.12 15.79	49 46.23
Total	87 82.08	19 17.92	106 100.00

Statistic	DF	Value	Prob
Chi-Square	1	8.6275	0.0033
Likelihood Ratio Chi-Square	1	9.4486	0.0021
Continuity Adj. Chi-Square	1	7.2001	0.0073
Mantel-Haenszel Chi-Square	1	8.5461	0.0035
Phi Coefficient		-0.2853	
Contingency Coefficient		0.2743	
Cramer's V		-0.2853	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	41
Left-sided Pr <= F	0.0028
Right-sided Pr >= F	0.9996
Table Probability (P)	0.0024
Two-sided Pr <= P	0.0044

Sample Size = 106

Question Set A8: Conduct the appropriate chi-square test for independence to test for an association between the Binary BMI and gender?

SHOW	Set up the hypotheses being tested:	HIDE
SHOW	Do you have any concerns about using the chi-square test? Explain.	HIDE
SHOW	What is the p-value of the appropriate chi-square test given in the output?	HIDE
SHOW	What is your conclusion in context?	HIDE
SHOW	Compare the distribution of binary BMI between the males and females using the appropriate conditional percentages.	HIDE

[Case CC] Chi-Square Test (Pearson's or Continuity Corrected)

- **Is there an association between gender and the multi-level body mass index variable?**

SPSS Output

Crosstab

			Body Mass Index Categories				Total
			Underweight	Normal	Overweight	Obese	
Gender	Male	Count	5	36	14	2	57
		% within Gender	8.8%	63.2%	24.6%	3.5%	100.0%
	Female	Count	10	36	3	0	49
		% within Gender	20.4%	73.5%	6.1%	0.0%	100.0%
Total		Count	15	72	17	2	106
		% within Gender	14.2%	67.9%	16.0%	1.9%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.239[a]	3	.017
Likelihood Ratio	11.590	3	.009
Linear-by-Linear Association	9.598	1	.002
N of Valid Cases	106		

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is .92.

SAS Output

Table of GENDER by BMIGroups

GENDER(Gender)	BMIGroups(Body Mass Index Categories)				
Frequency Percent Row Pct Col Pct	< 18.5	[18.5, 25)	[25, 30)	30+	Total
Male	5 4.72 8.77 33.33	36 33.96 63.16 50.00	14 13.21 24.56 82.35	2 1.89 3.51 100.00	57 53.77
Female	10 9.43 20.41 66.67	36 33.96 73.47 50.00	3 2.83 6.12 17.65	0 0.00 0.00 0.00	49 46.23
Total	15 14.15	72 67.92	17 16.04	2 1.89	106 100.00

Statistic	DF	Value	Prob
Chi-Square	3	10.2389	0.0166
Likelihood Ratio Chi-Square	3	11.5903	0.0089
Mantel-Haenszel Chi-Square	1	9.5979	0.0019
Phi Coefficient		0.3108	
Contingency Coefficient		0.2968	
Cramer's V		0.3108	

WARNING: 25% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 106

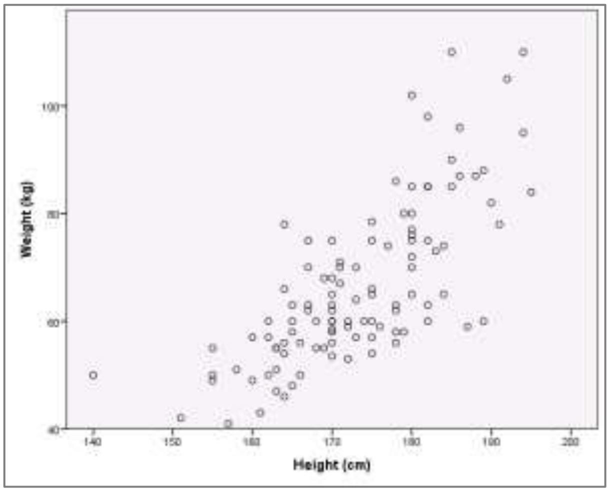
Question Set A9: Conduct the appropriate chi-square test for independence to test for an association between gender and BMI groups?

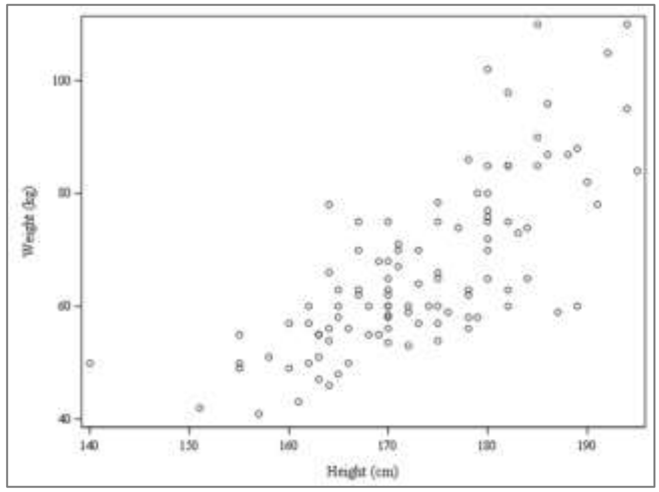
SHOW	Set up the hypotheses being tested:	HIDE
SHOW	Do you have any concerns about using the chi-square test? Explain.	HIDE
SHOW	What is the p-value of the appropriate chi-square test given in the output?	HIDE
SHOW	What is your conclusion in context? Answer this question completely regardless of your answer regarding concerns about this test. (On the course project we will generally ask you to provide conclusions and interpretations even if you have concerns in order to assess your understanding of these tests).	HIDE
SHOW	What non-parametric test(s) would be appropriate?	HIDE

[Case QQ] Scatterplots, Correlation, and Regression

- **Is there a linear relationship between height and weight in the entire sample?**

Scatterplot: (SPSS on left, SAS on right)





Pearson's Correlation Coefficient: (SPSS on left, SAS on right)

Correlations

		Height (cm)	Weight (kg)
Height (cm)	Pearson Correlation	1	.741**
	Sig. (2-tailed)		.000
	N	106	106
Weight (kg)	Pearson Correlation	.741**	1
	Sig. (2-tailed)	.000	
	N	106	106

**. Correlation is significant at the 0.01 level (2-tailed).

Pearson Correlation Coefficients, N = 106
Prob > |r| under H0: Rho=0

	HEIGHT	**WEIGHT**
HEIGHT Height (cm)	1.00000	0.74052 <.0001
WEIGHT Weight (kg)	0.74052 <.0001	1.00000

Correlation is listed first in the cell in the table and the p-value (Prob > |r|) is second.

Note: These are usually presented paired in this way where the information is repeated – this helps when you are looking at many variables from numerous angles. The repeated information is highlighted in green above.

Useful Linear Regression Tables:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.741[a]	.548	.544	10.055

a. Predictors: (Constant), Height (cm)

b. Dependent Variable: Weight (kg)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12766.571	1	12766.571	126.276	.000[b]
	Residual	10514.460	104	101.101		
	Total	23281.031	105			

a. Dependent Variable: Weight (kg)

b. Predictors: (Constant), Height (cm)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-120.595	16.689		-7.226	.000	-153.690	-87.499
	Height (cm)	1.081	.096	.741	11.237	.000	.890	1.272

a. Dependent Variable: Weight (kg)

Residuals Statistics[a]

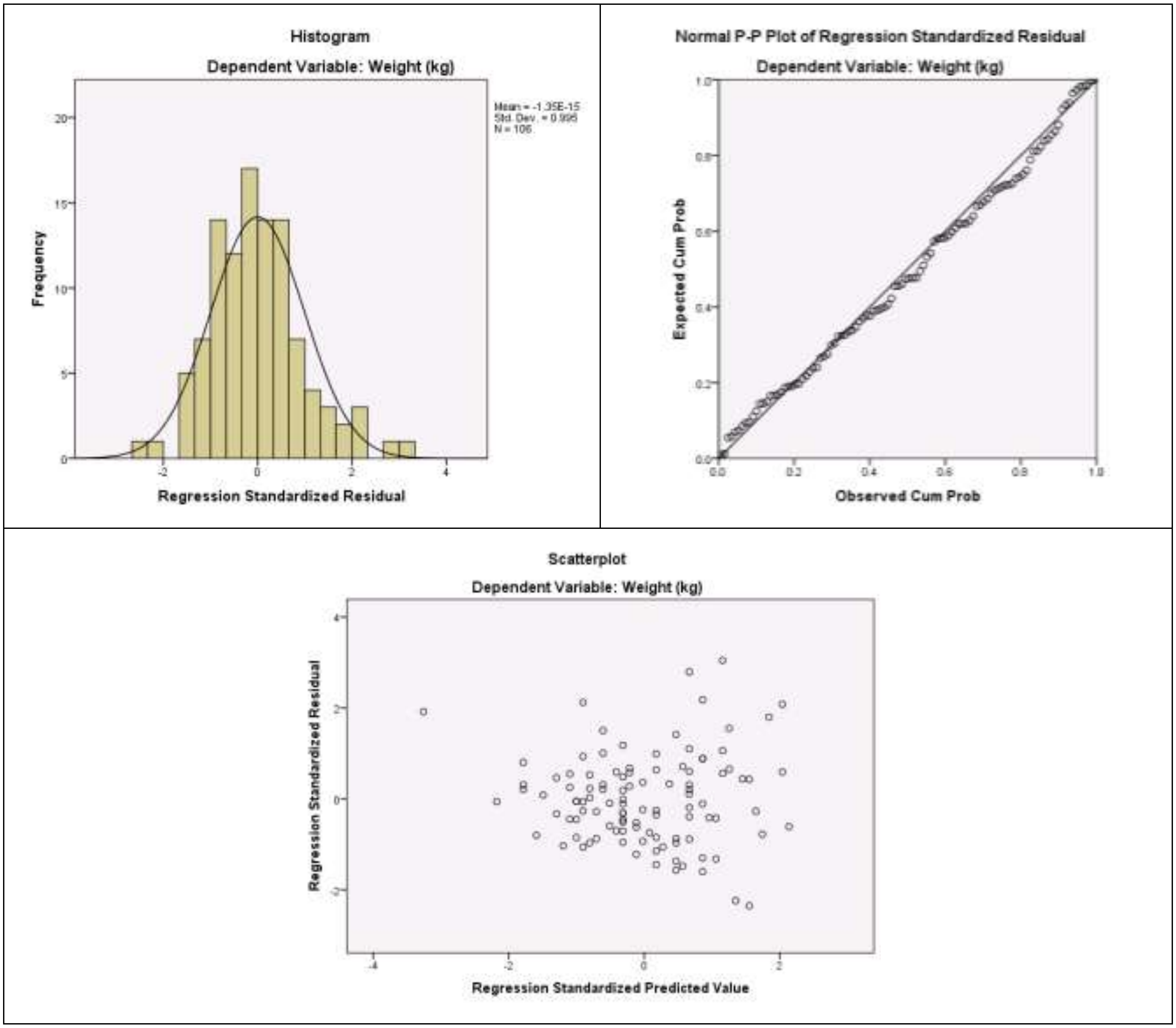
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	30.72	90.17	66.63	11.027	106
Residual	-23.686	30.637	.000	10.007	106
Std. Predicted Value	-3.256	2.135	.000	1.000	106
Std. Residual	-2.356	3.047	.000	.995	106

a. Dependent Variable: Weight (kg)

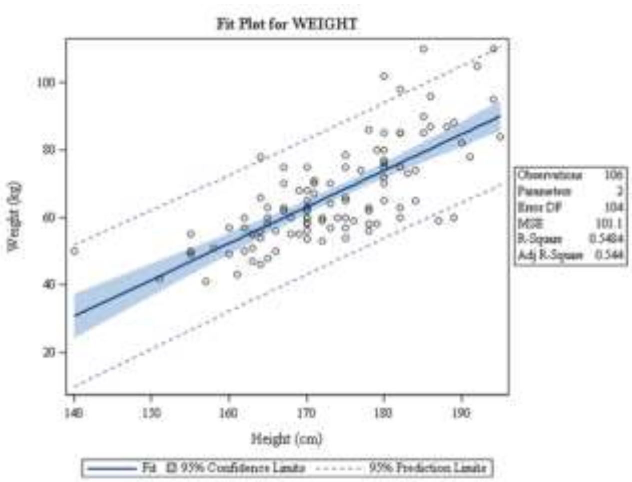
Note: The highlighted rows and columns contain possibly useful information about the predicted values and residuals for our model.

- Here the minimum predicted Weight (kg) is 30.72 and the maximum is 90.17. This can be helpful to see if the model predictions go outside reasonable values.
- The largest negative and positive residuals are off by 23.686 and 30.637 kg respectively which are fairly large considering that the weight values ranged from around 40 kg to around 120 kg. This means for these individuals (whose actual – predicted values result in the largest and smallest residuals), our model prediction is off by 23.7 kg and 30.6 kg respectively.
- The standardized values can help determine if any of the predicted values or residuals is very extreme. Here we do have values above 3 which are somewhat outside the range of what we would like to see in a perfect scenario (values between around 2 and -2 with an occasional value slightly outside that range, especially for large sample sizes.)

Linear Regression Diagnostic Graphs:



SAS Output



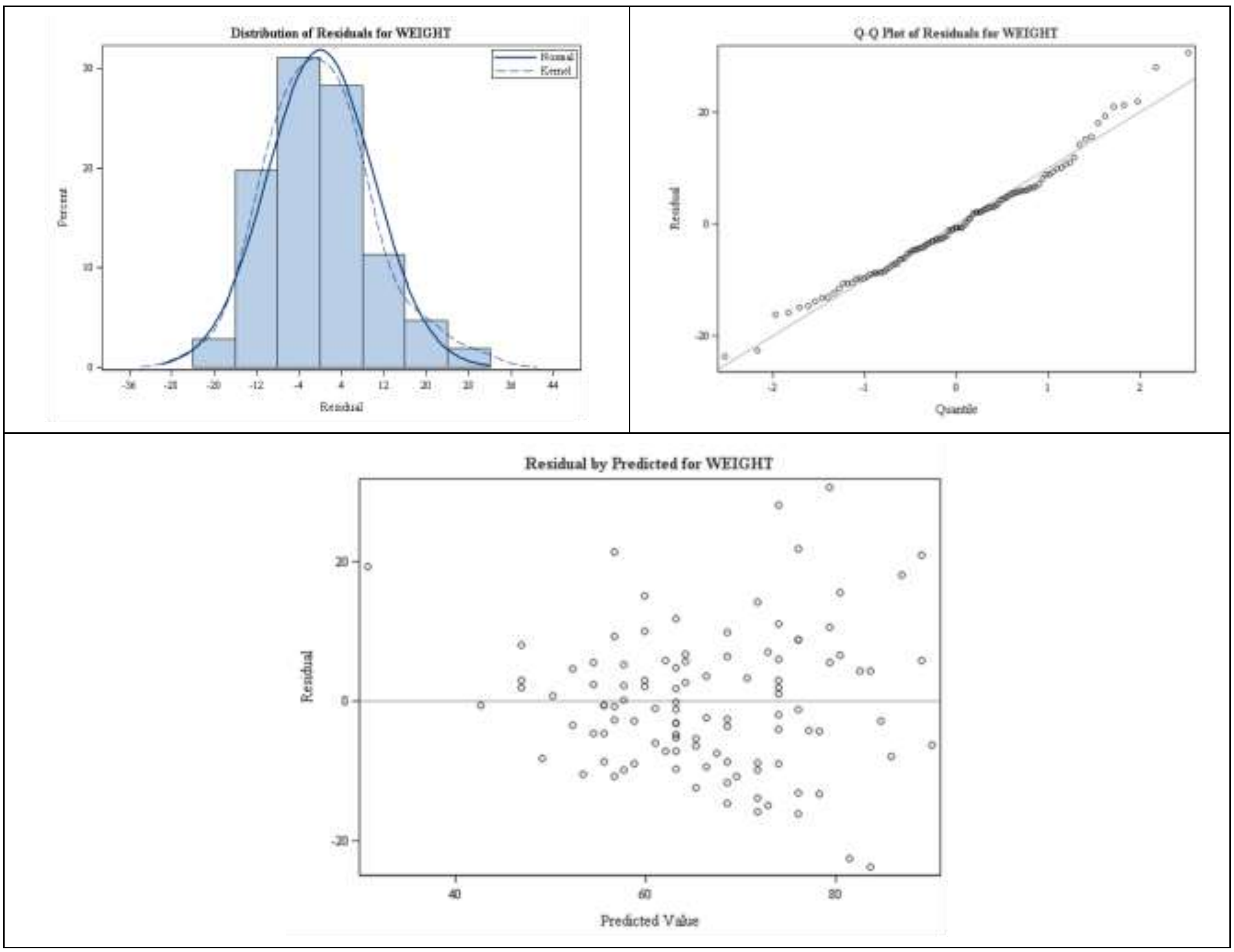
The REG Procedure
Model: MODEL1
Dependent Variable: WEIGHT Weight (kg)

Number of Observations Read	106
Number of Observations Used	106

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	12767	12767	126.28	<.0001
Error	104	10514	101.10058		
Corrected Total	105	23281			

Root MSE	10.05488	**R-Square**	0.5484
Dependent Mean	66.62736	**Adj R-Sq**	0.5440
Coeff Var	15.09122		

Parameter Estimates								
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**95% Confidence Limits**	
Intercept	Intercept	1	-120.59472	16.68943	-7.23	<.0001	-153.69049	-87.49895
HEIGHT	Height (cm)	1	1.08085	0.09618	11.24	<.0001	0.89011	1.27159



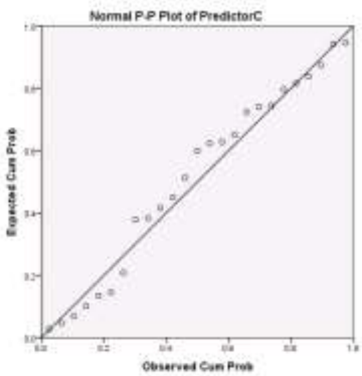
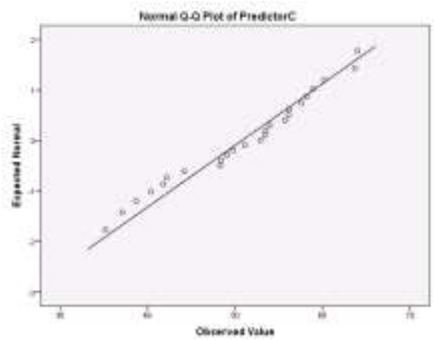
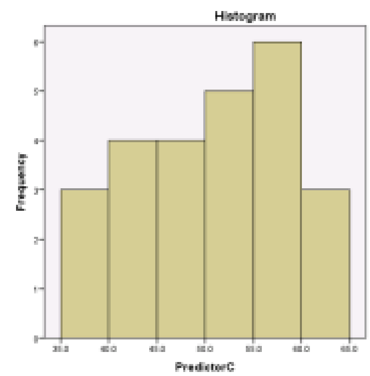
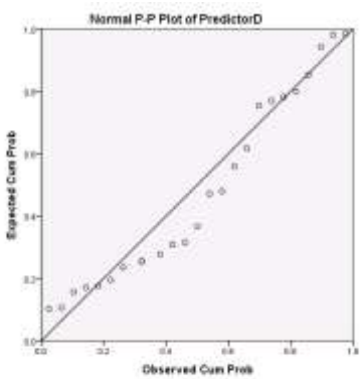
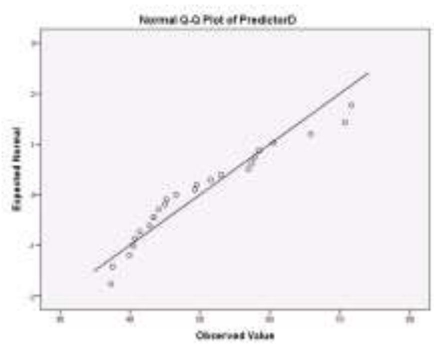
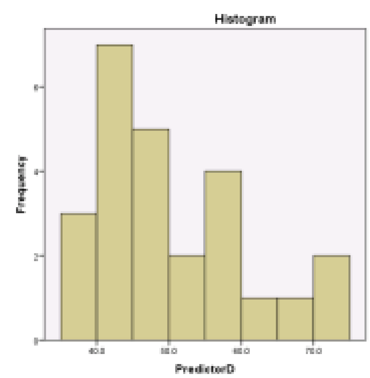
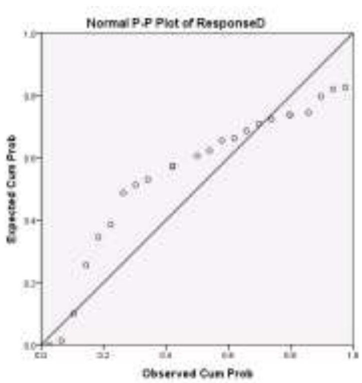
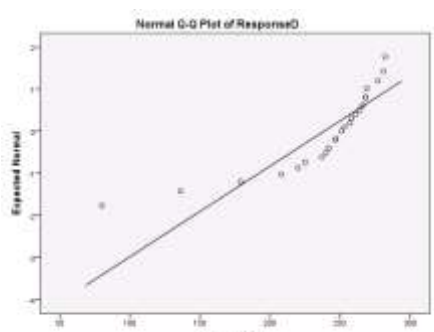
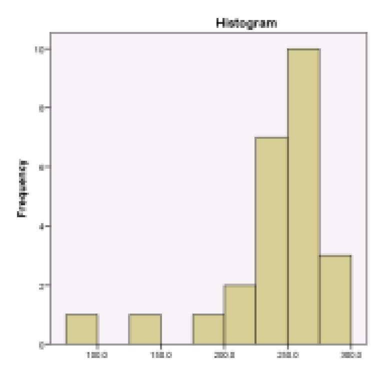
Question Set A10: Answer the following questions regarding the relationship between height and weight in the entire sample.

SHOW	Based upon the scatterplot, is linear regression and correlation a reasonable analysis?	HIDE
SHOW	What is the value of Pearson's Correlation Coefficient and it's two-sided p-value?	HIDE
SHOW	Interpret the value of the correlation coefficient.	HIDE
SHOW	The regression equation is:	HIDE
SHOW	Interpret R-square in context.	HIDE
SHOW	Interpret the slope in context.	HIDE
SHOW	Are there any concerns with the assumption that the error term is normally distributed?	HIDE
SHOW	Are there any concerns with the assumption of constant variance around the regression line?	HIDE
SHOW	State the p-value for the slope and provide a conclusion to the associated test in context.	HIDE

PP-Plots: These are similar to QQ-plots in that they help investigate normality.

- **The difference between QQ-plots and PP-plots is that PP-plots will always start and end at (0,0) and (1,1) on the line drawn since the plot is based upon probabilities. The QQ-plot is based upon quantiles/percentiles and points can deviate from the line at the ends.**
- **For both, any systematic deviations from the line indicate potential non-normality. The more severe deviations, the more concern we have.**

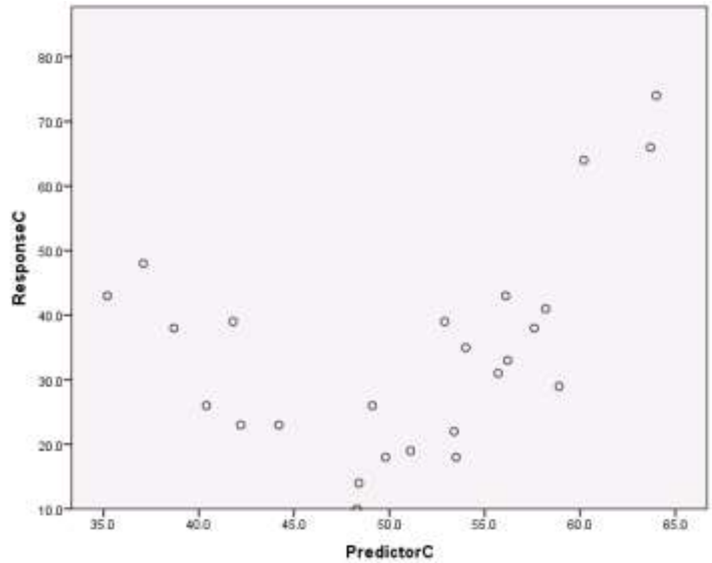
Here is a pp-plot, qq-plot, and histogram for a few simulated variables.

PP-PLOT	QQ-PLOT	HISTOGRAM
		
		
		

[Case QQ] Scatterplots, Correlation, and Regression

- **Simulated Data for Regression**

Scatterplot: (SPSS Only)



Pearson's Correlation Coefficient: (SPSS Only)

Correlations

		PredictorC	ResponseC
PredictorC	Pearson Correlation	1	.374
	Sig. (2-tailed)		.066
	N	25	25
ResponseC	Pearson Correlation	.374	1
	Sig. (2-tailed)	.066	
	N	25	25

Linear Regression Tables:

Descriptive Statistics

	Mean	Std. Deviation	N
ResponseC	34.400	16.1658	25
PredictorC	50.828	8.1795	25

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.374[a]	.140	.102	15.3173

a. Predictors: (Constant), PredictorC

b. Dependent Variable: ResponseC

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	875.741	1	875.741	3.733	.066[b]
	Residual	5396.259	23	234.620		
	Total	6272.000	24			

a. Dependent Variable: ResponseC

b. Predictors: (Constant), PredictorC

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3.137	19.669		-.159	.875
	PredictorC	.739	.382	.374	1.932	.066

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	-43.826	37.552
	PredictorC	-.052	1.529

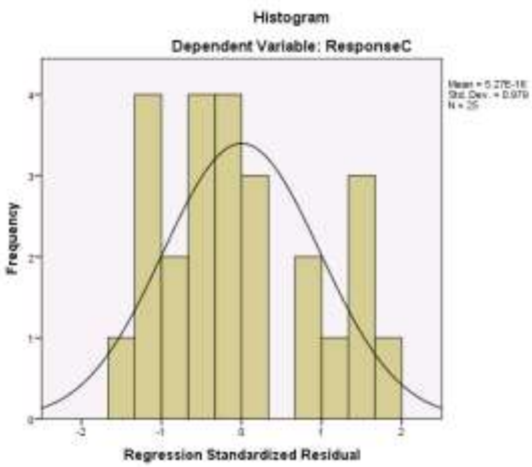
a. Dependent Variable: ResponseC

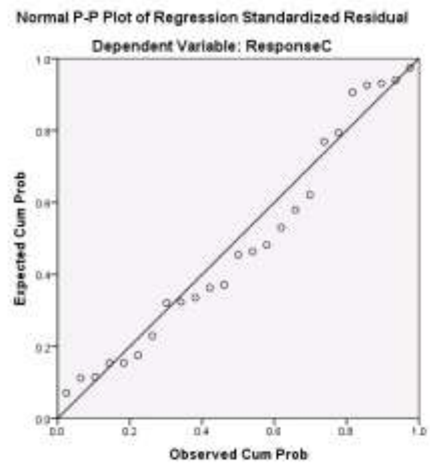
Residuals Statistics[a]

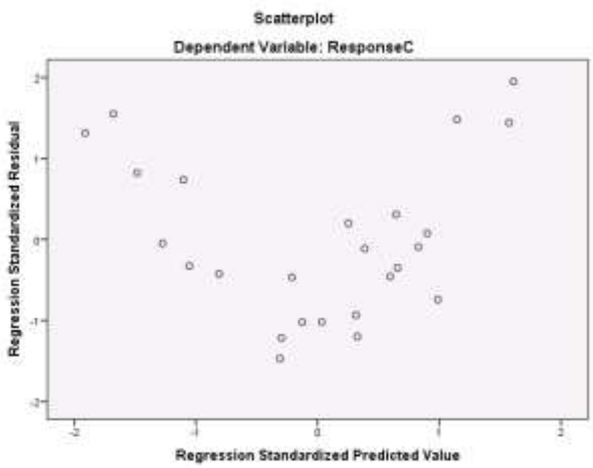
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	22.859	44.128	34.400	6.0406	25
Residual	-22.5330	29.8723	.0000	14.9948	25
Std. Predicted Value	-1.911	1.610	.000	1.000	25
Std. Residual	-1.471	1.950	.000	.979	25

a. Dependent Variable: ResponseC

Linear Regression Diagnostic Graphs:







SAS Output

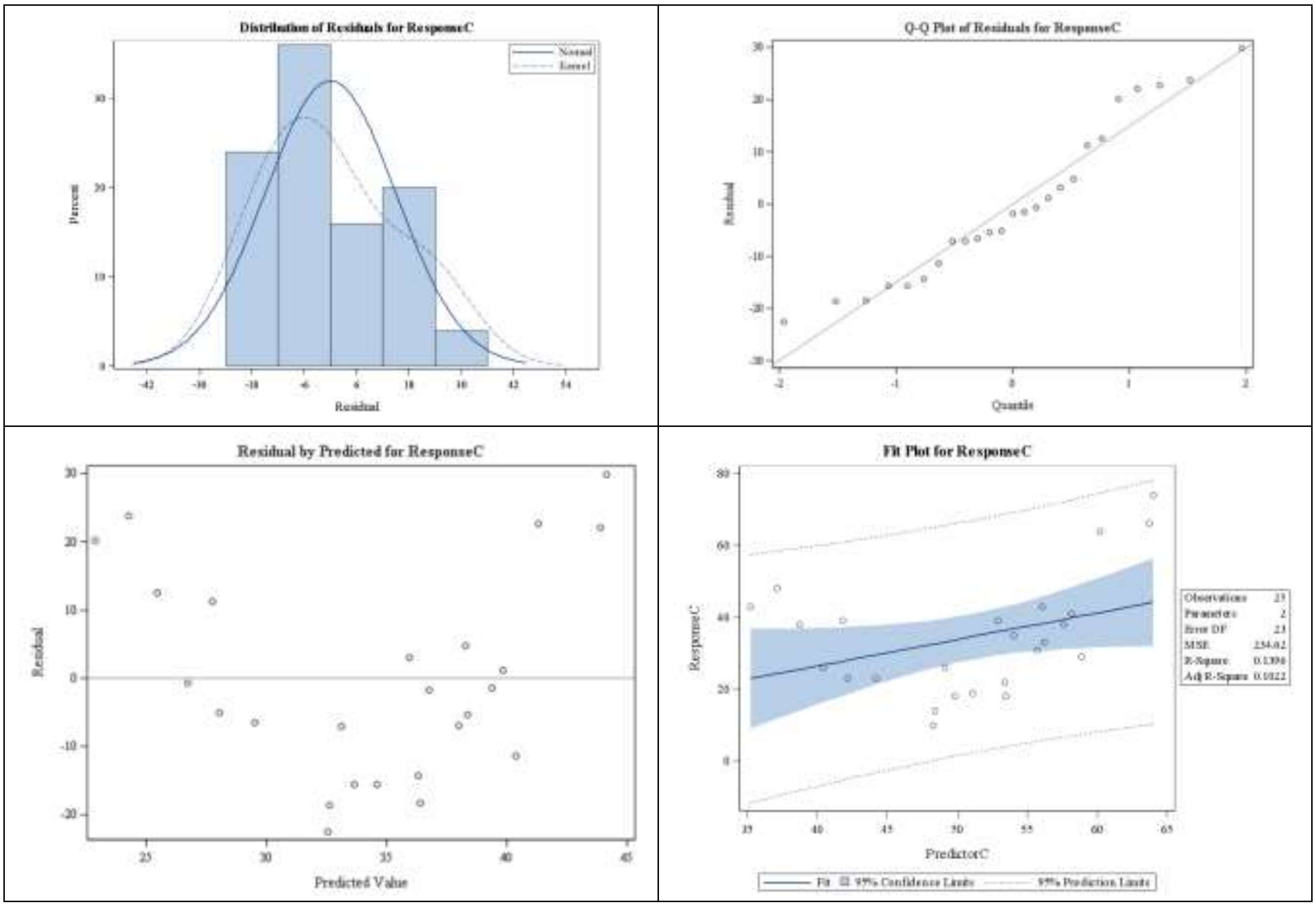
The REG Procedure
Model: MODEL1
Dependent Variable: ResponseC

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	875.74110	875.74110	3.73	0.0658
Error	23	5396.25890	234.61995		
Corrected Total	24	6272.00000			

Root MSE	15.31731	**R-Square**	0.1396
Dependent Mean	34.40000	**Adj R-Sq**	0.1022
Coeff Var	44.52706		

Parameter Estimates							
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**95% Confidence Limits**	
Intercept	1	-3.13704	19.66922	-0.16	0.8747	-43.82591	37.55184
PredictorC	1	0.73851	0.38225	1.93	0.0658	-0.05224	1.52926



Question Set A11: Answer the following questions regarding the relationship between Predictor C and Response C.

SHOW Based upon the scatterplot, is linear regression and correlation a reasonable analysis? HIDE

SHOW What is the value of Pearson's Correlation Coefficient and it's two-sided p-value. HIDE

SHOW Interpret the value of the correlation coefficient. HIDE

SHOW Instructor Comments: HIDE

Note: Spearman's Rank Correlation

Here are two scatterplots which illustrate non-linear associations for which Spearman's rank correlation **would be** appropriate.

On the left, we have a non-linear but decreasing trend and on the right a non-linear but increasing trend.