ANOVA

(Between Groups/Subjects)

Analysis Of Variance

Narine Gevorgyan & Eric McNally

What is ANOVA?

- ***Like T test, compares the means of populations/groups***

- ***Unlike T Test, ANOVA can compare two or more populations/groups***

- ***ANOVA is part of family of F-tests, and compares means relative to variation within groups***

Assumptions for ANOVA

-Normal data distribution
-Interval Data
-Levine's test P value .035

-Normal data distribution
-Ordinal Data
-Levine's test P value .564

-Abnormal data distribution
-Ratio Data
-Levine's test P value .342

-Normal data distribution
-Interval Data
-Levine's test P value .254

Assumptions for ANOVA

-Normal data distribution
-Interval Data
-Levine's test P value .254

- ***Interval or Ratio Data***

- ***Data must be NORMAL***

- ***Homogenous Variance***

ONE-WAY ANOVA

“A one-way ANOVA is used when we have 1 IV with >3 levels, independent samples, and the same DV” -Miguel

ONE-WAY ANOVA

	1	2	3	4
Participant Test Scores (%)	72	86	72	76
	52	69	94	86
	94	71	65	69
	71	92	85	74
	82	71	73	94

Participant Groups

- **No Exercise**
- **Walking**
- **Running**
- **Cycling**

Study testing the effect of different forms of morning exercise on cognitive function (Measured using standardized problem-solving quiz)

ONE-WAY ANOVA

	4 GROUPS TOTAL			
Participant Test Scores (%)	72	86	72	76
	52	69	94	86
	94	71	65	69
	71	92	85	74
	82	71	73	94

More than 2 groups, T-Test CANNOT be used.

HYPOTHOSES: ANOVA VS T-TEST

H^0 for T-Test

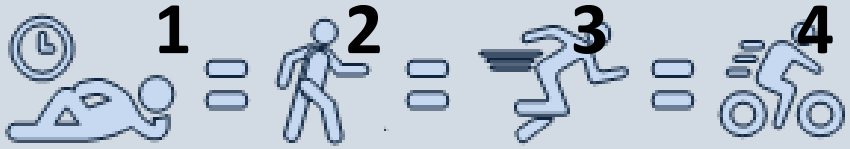
$\mu^1=\mu^2$

H^1 for T-Test

$\mu^1\neq\mu^2$

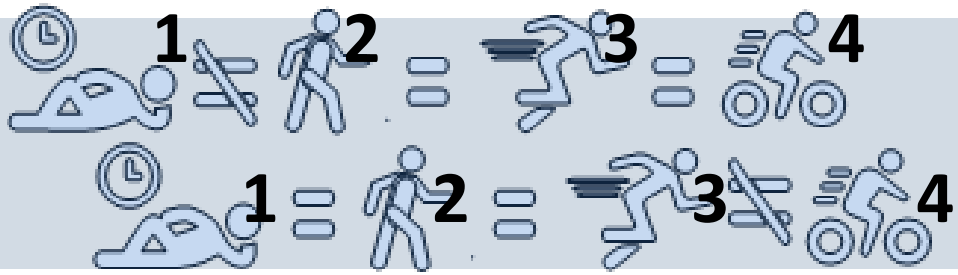
H^0 One-Way ANOVA

$\mu 1 = \mu 2 = \mu 3 = \mu 4$



H^11 One-Way ANOVA

$\mu^* \neq \mu^*$ (*=any)

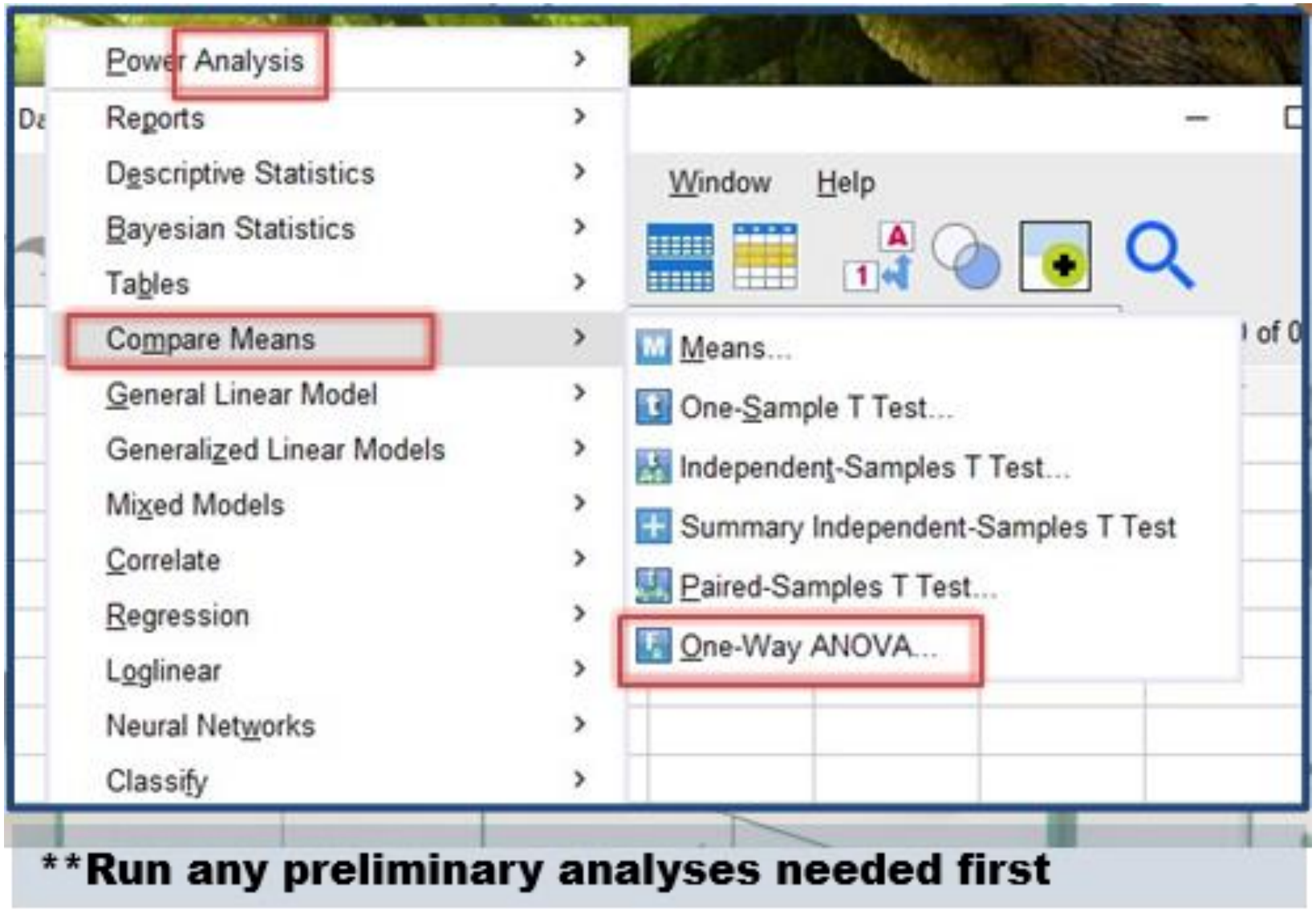


WARNING ABOUT ANOVA

IMPORTANT! The ONE-WAY ANOVA is not intended to give details about bilateral relationships between groups.

It simply presents a boolean True/False value relating to the ANOVA null hypothesis.

Accessing One-Way ANOVA in SPSS

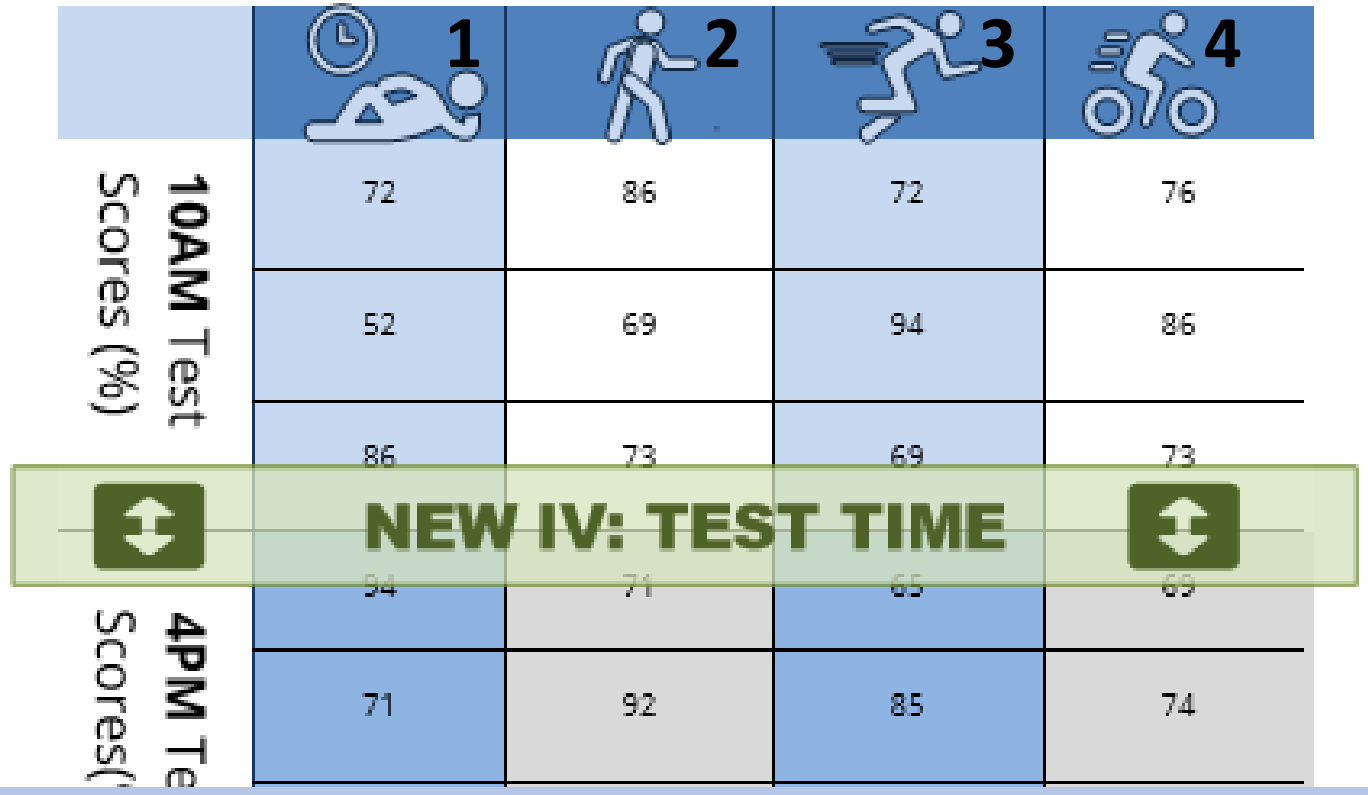


**Ensure data normality

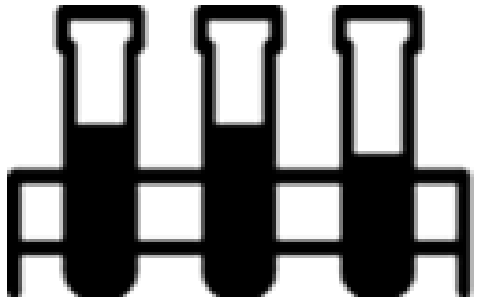
**Deal with outliers

**Ensure data is interval or ratio

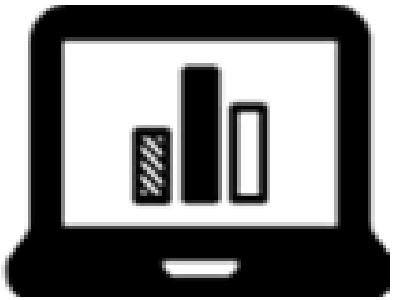
When to use TWO-WAY ANOVA



In an expanded version of the study, time of day that the test is taken is also a factor. One-Way ANOVA cannot be used.



TWO-WAY ANOVA



A TWO-WAY ANOVA tests the effect of multiple groups of two independent variables on a dependent variable and on each other.

The two independent variables in a TWO-WAY ANOVA are called FACTORS. Each factor will have two or more levels within it.

TREATMENT GROUPS

Treatment groups are formed by making all possible combinations of two factors.

For example, if the first factor has 3 levels and the second factor has 2 levels, then there will be 3 x 2=6 different treatment groups.

ANOVA

As a follow-up study, Lorraine wants to test the effect of the time of the day of the test in addition to the type of exercises performed (TWO-WAY ANOVA).

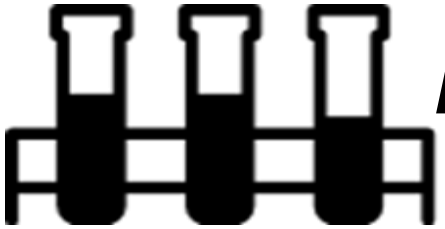
		IV: TIME OF THE DAY			
Test Scores	10 AM	72	86	72	76
		52	69	94	86
		86	73	69	73
Test Scores	4 PM	94	71	65	69
		71	92	85	74
		82	71	73	94

- ***First factor: Type of exercise assigned to the participants; levels: none, walking, running, cycling (4 levels)***

- ***Second factor: Time of the day of the test; levels: 10:00 AM and 4:00 PM (2 levels)***

- ***Hence, Lorraine will have 4x2=8 different treatment groups***

10: 00 AM	65	86	72	80
	74	69	94	94
	90	82	65	87
	71	92	85	74
	82	83	73	86
	86	80	90	92
4:00 PM	77	80	90	76
	66	72	85	86
	72	71	73	73
	80	92	60	74
	68	63	68	94

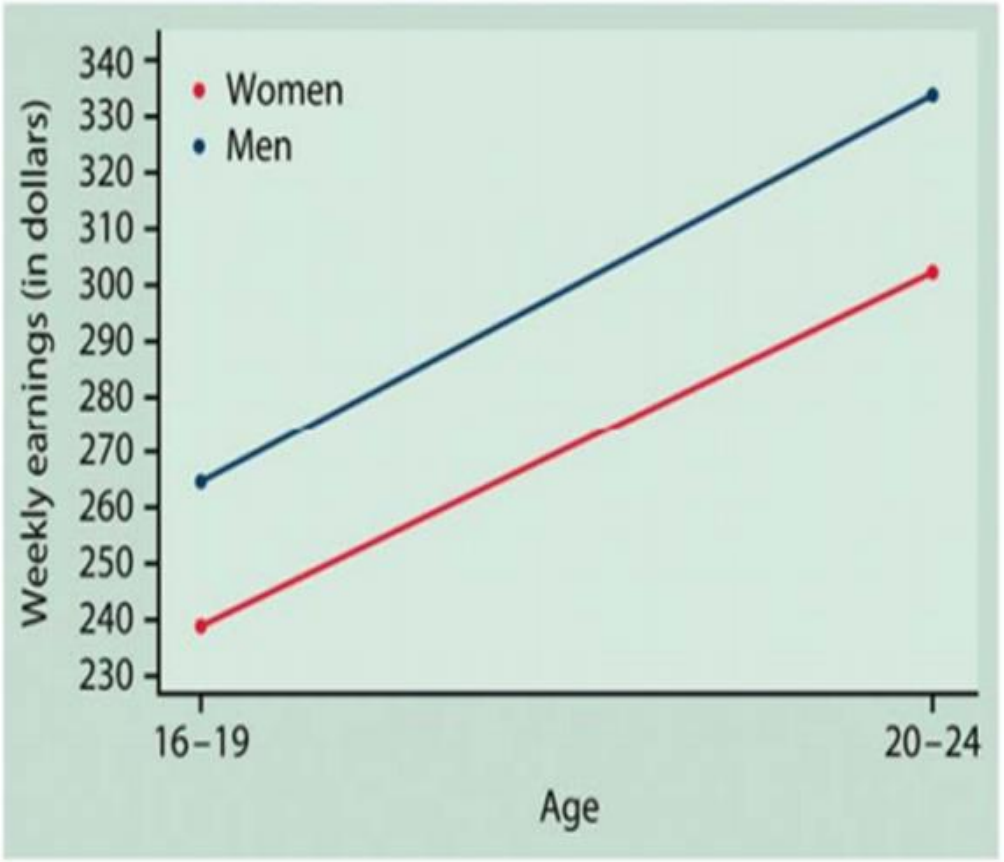


MAIN EFFECT AND INTERACTION

Main effect *is the direct effect of a single factor or independent variable on the dependent variable.*

Interaction effect *is the combined effect of two or more factors on the dependent variable.*

Interpreting Main Effects and Interactions Plots

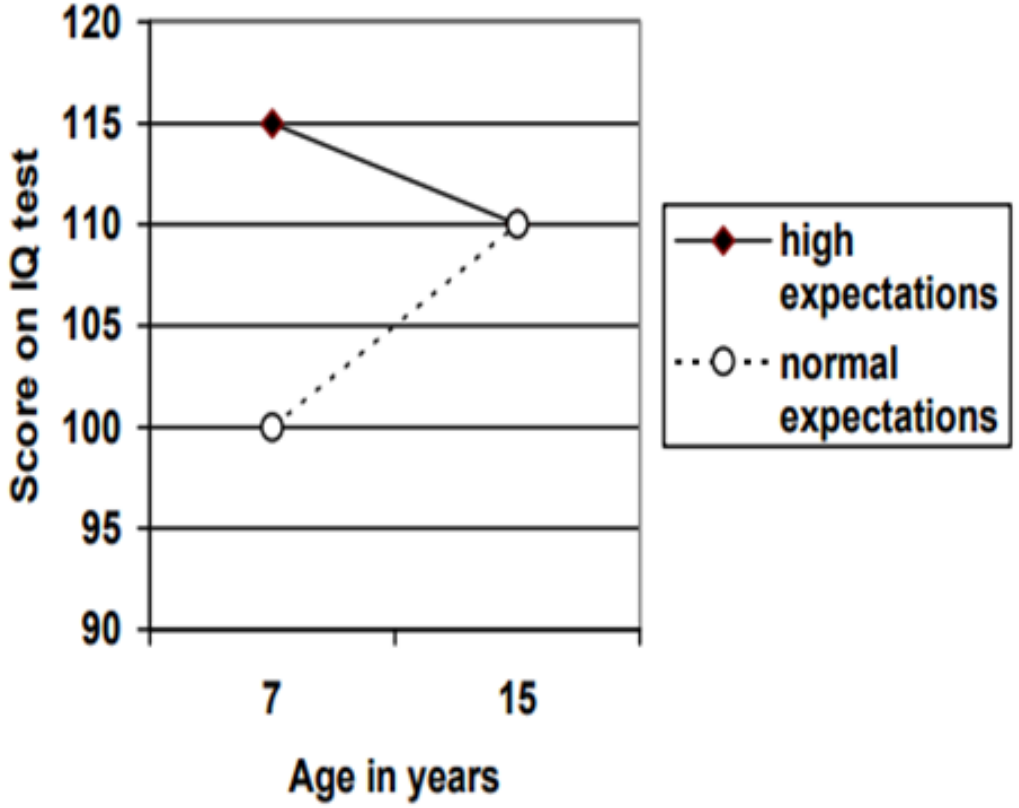


There are only two main effects that are present: an effect due to age and gender.

Parallel lines in plot usually indicate little or no interaction.

Retrieved from https://www2.slideshare.net/musadoto/2-way-anovaanalysis-of-variance?from_action=save

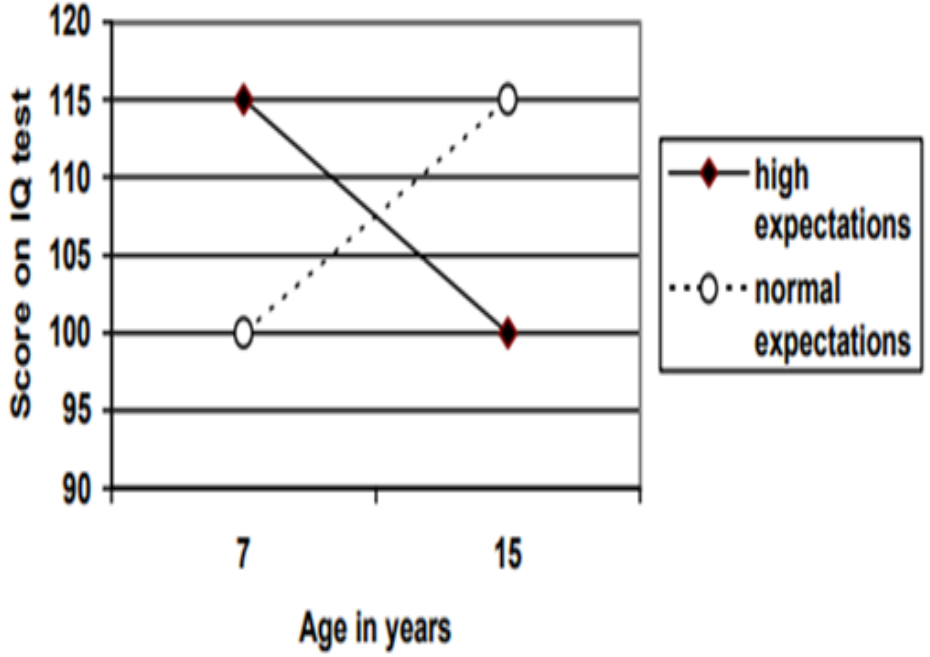
Interpreting Main Effects and Interactions through Figures



Effects of age and teacher expectations on IQ scores

The less parallel the lines are, the more likely there is to be a significant interaction. So, in this case, we would expect an interaction.

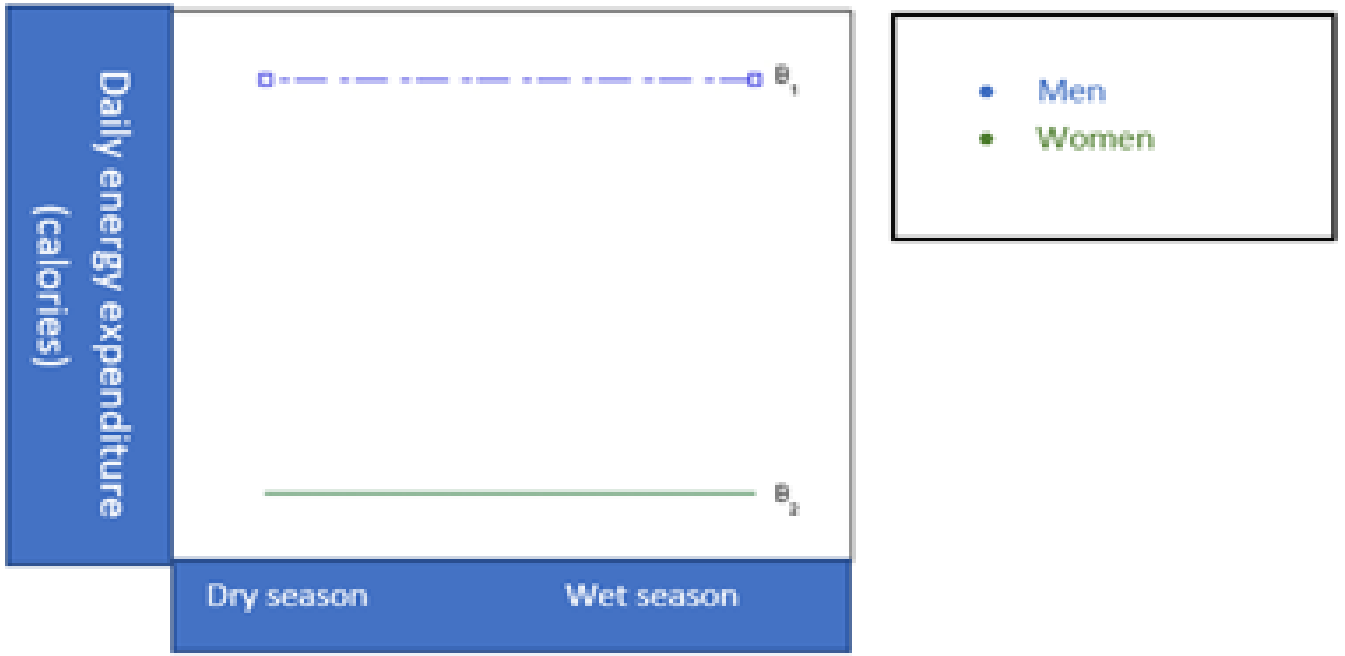
Interpreting Main Effects and Interactions through Figures



The lines are not parallel, so there is likely to be an interaction. Because the lines intersect, this type of interaction is sometimes called a cross-over interaction.

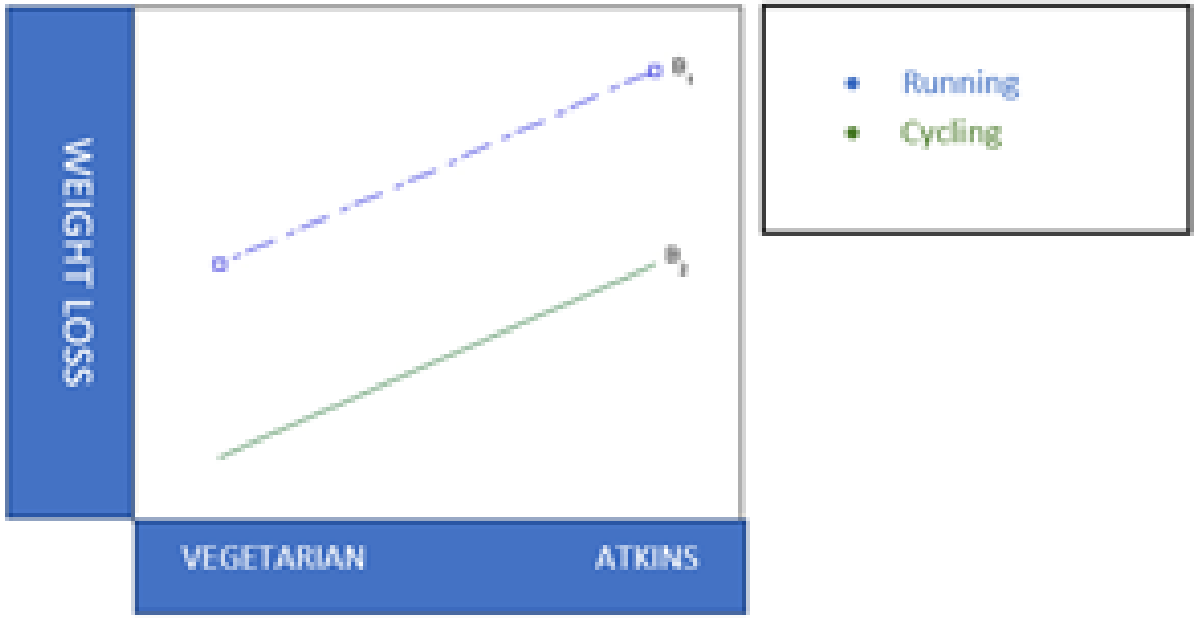
A cross-over interaction

Interpret the Following Plots: Main Effects and Interaction



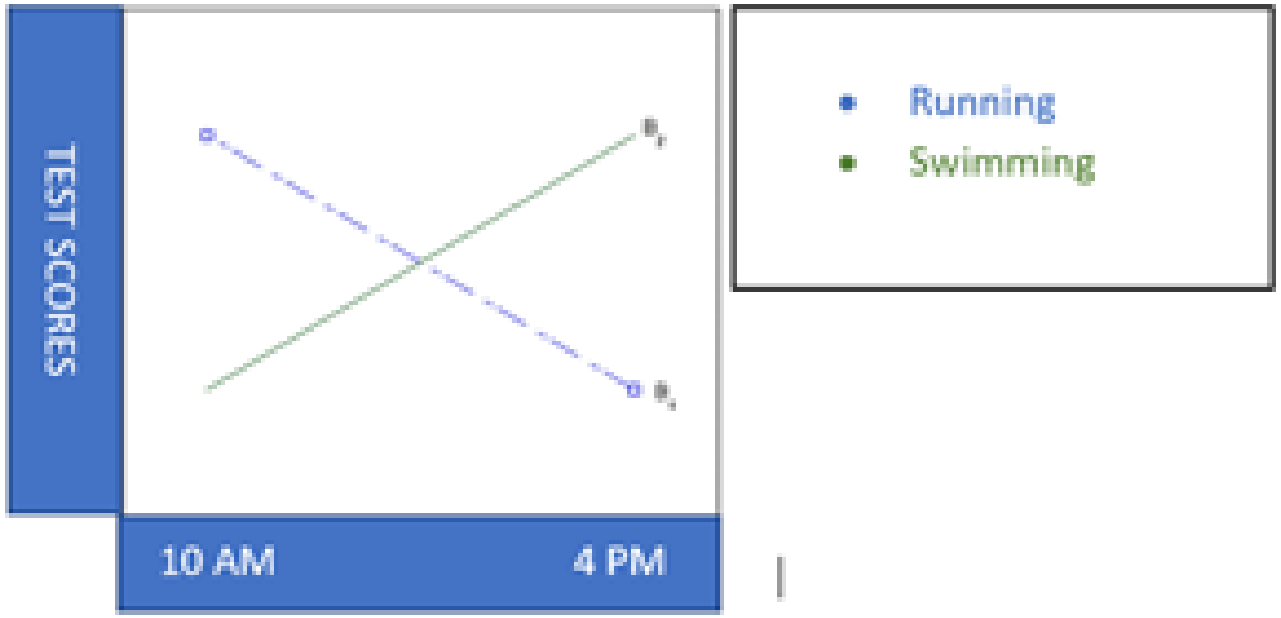
- **A) Main Effect Factor Gender**
- **B) Main Effect Factor Season**
- **C) Interaction of Both Factors**

Interpret the Following Plots: Main Effects and Interaction



- **A) Main Effect Factor Diet**
- **B) Main Effect Factor Exercise**
- **C) Interaction of Both Factors**

Interpret the Following Plots: Main Effects and Interaction



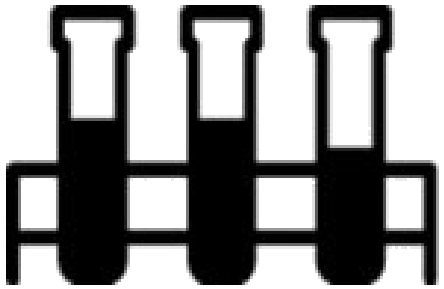
- **A) Main Effect Factor Time**
- **B) Main Effect Factor Exercise**
- **C) Interaction of Both Factors**

TWO-WAY ANOVA: ASSUMPTIONS CHECKLIST

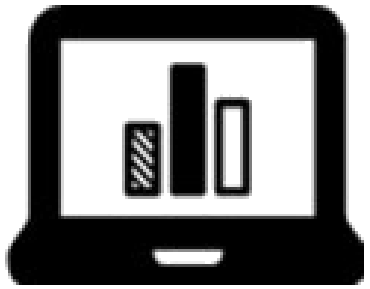
- ***Assumption #1:*** *Your dependent variable should be measured at the continuous level (i.e., they are interval or ratio variables).*

- ***Assumption #2:*** *Your two independent variables should each consist of two or more categorical, independent groups*

- ***Assumption #3:*** *You should have independence of observations*



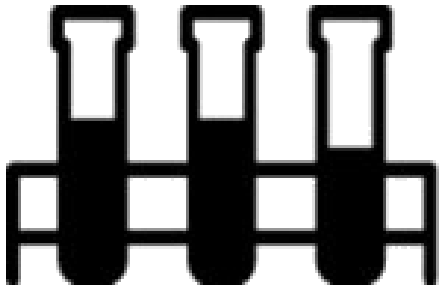
TWO-WAY ANOVA: ASSUMPTIONS CHECKLIST



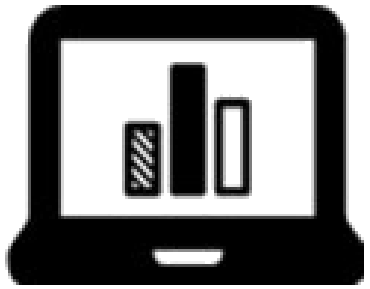
- **Assumption #4:** There should be no significant outliers.

- **Assumption #5:**
Your dependent variable should be approximately normally distributed for each combination of the groups of the two independent variables

- **Assumption #6:**
There needs to be homogeneity of variances for each combination of the groups of the two independent variables.



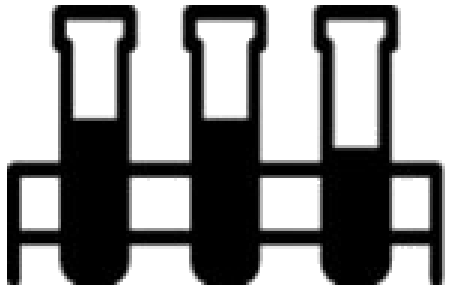
HYPOTHESES



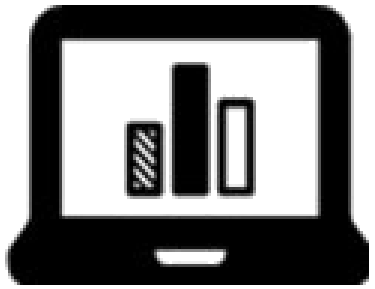
- ***Ho1:The sample means of the first factor are equal.***

- ***Ho2:The sample means of the second factor are equal.***

- ***Ho3: There is no interaction between the two factors.***

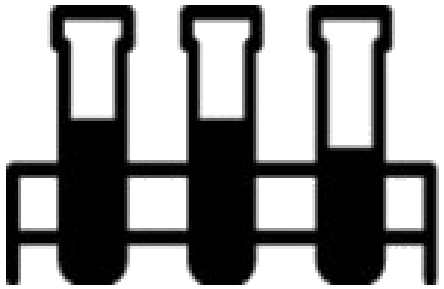


TWO-WAY ANOVA IN SPSS

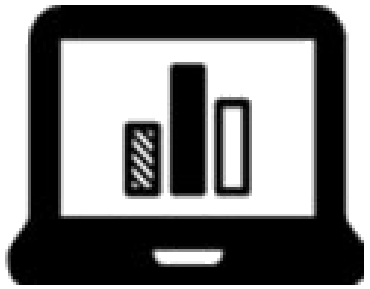


Lorraine wants to conduct a research study to see if the individual's weight is affected by the type of the diet and the type of the exercise.

In order to find out, 180 participants were randomly assigned to one of the three types of diet (None, Vegetarian and Atkins) and three types of exercises (None, 30 min a day and 60 min. a day)



HYPOTHESES

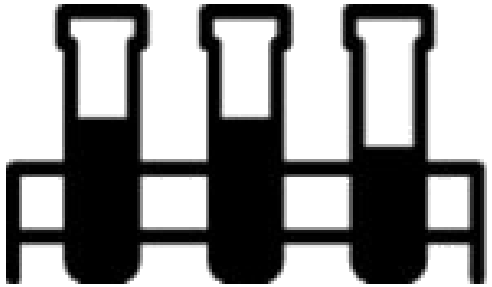


H0= The means of all diet groups are equal
H1=The mean of at least one diet group is different.

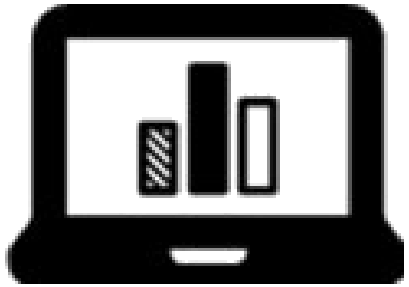
H0= The means of the exercise groups are equal
H1=The means of the exercise groups are different.

H0= There is no interaction between the diet and exercise

H1=There is an interaction between the diet and exercise



TREATMENT GROUPS



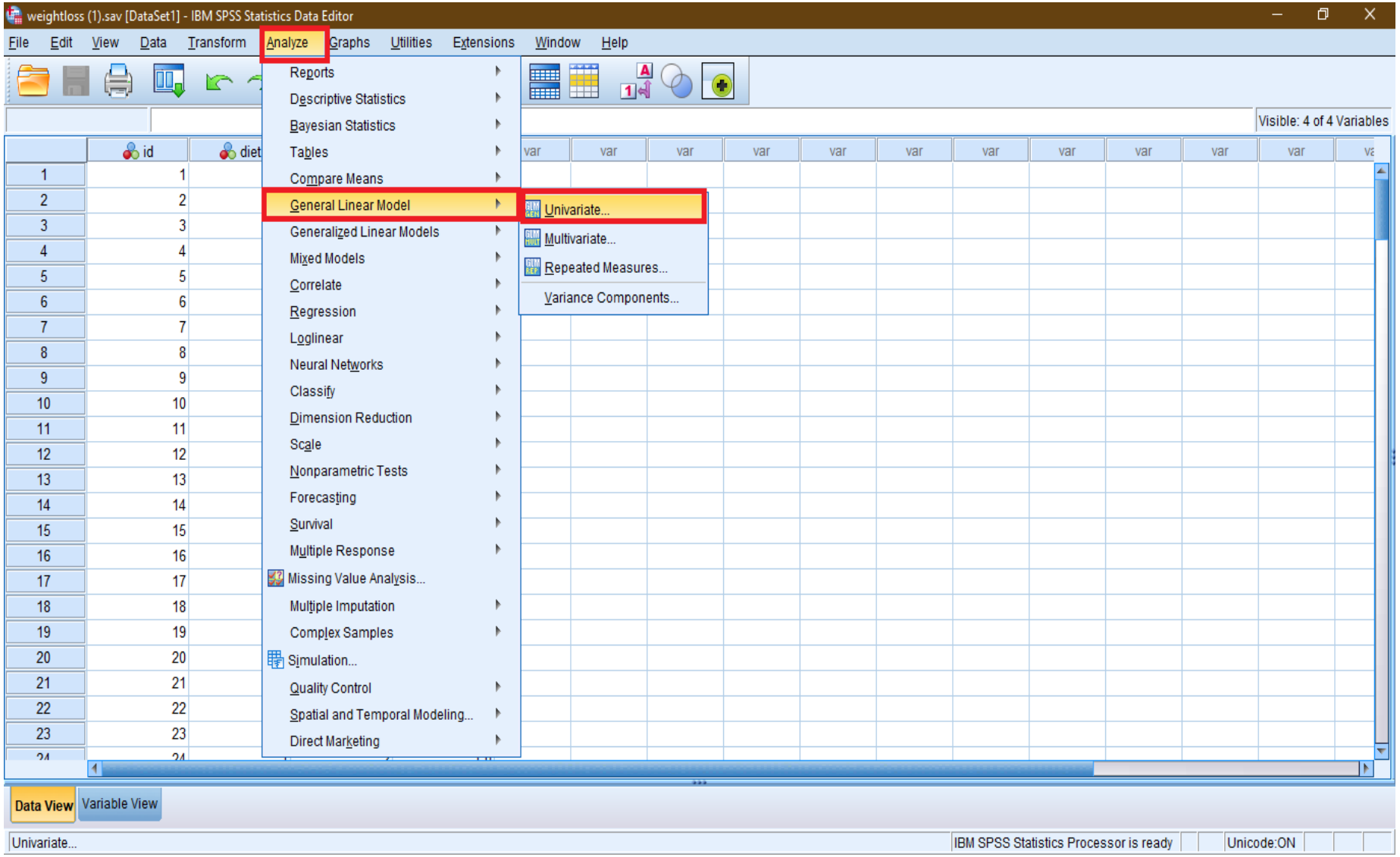
- ***First factor: Type of the exercise assigned to the participants; levels: none, 30 min a day, 60 min a day (3 levels)***

- ***Second factor: Type of the diet assigned to the participants; levels: None, Vegetarian and Atkins (3 levels)***

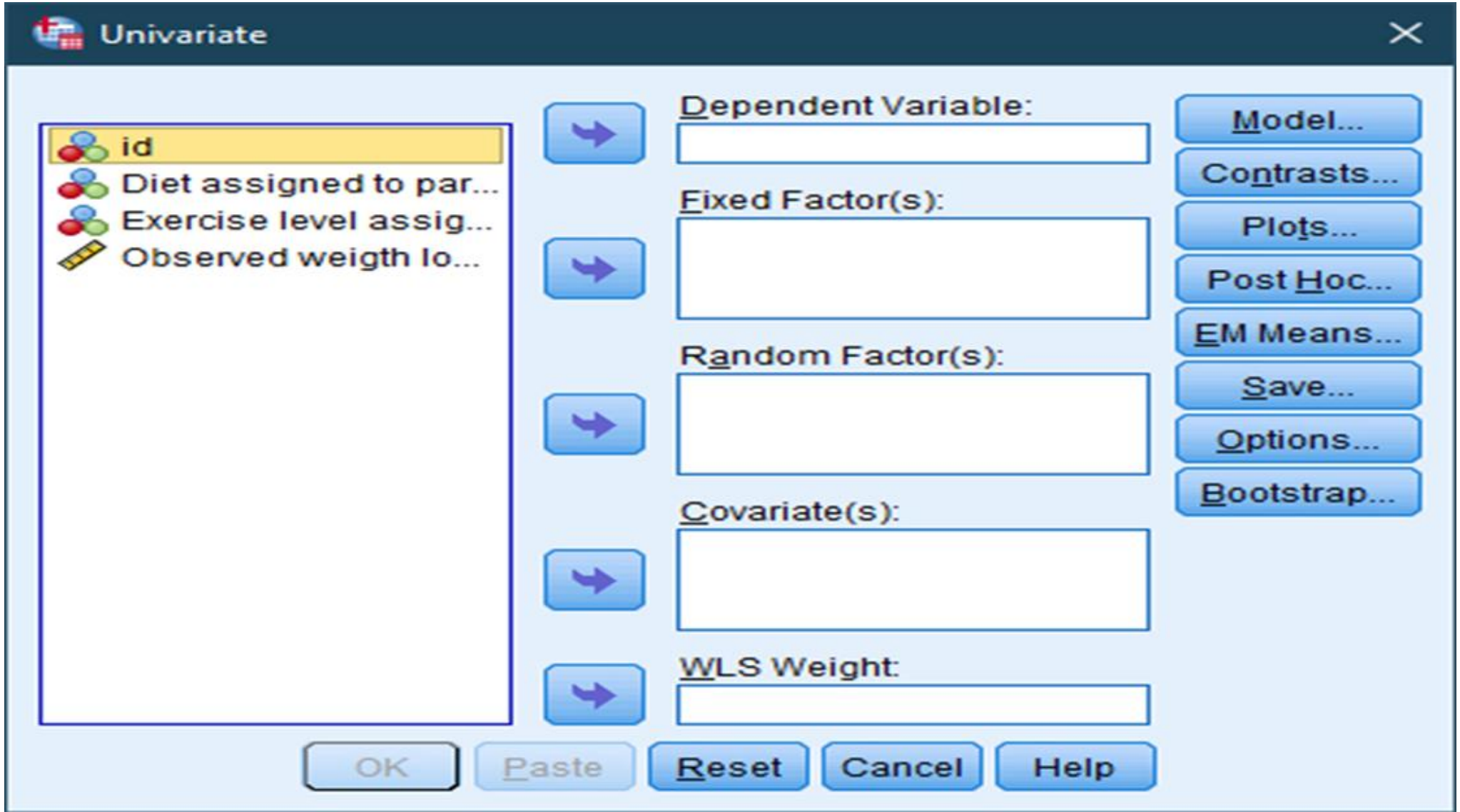
- ***Hence, Lorraine will have 3x3=9 different treatment groups***

	None	Vegetarian	Atkins
None	20	20	20
30 min a day	20	20	20
60 min a day	20	20	20

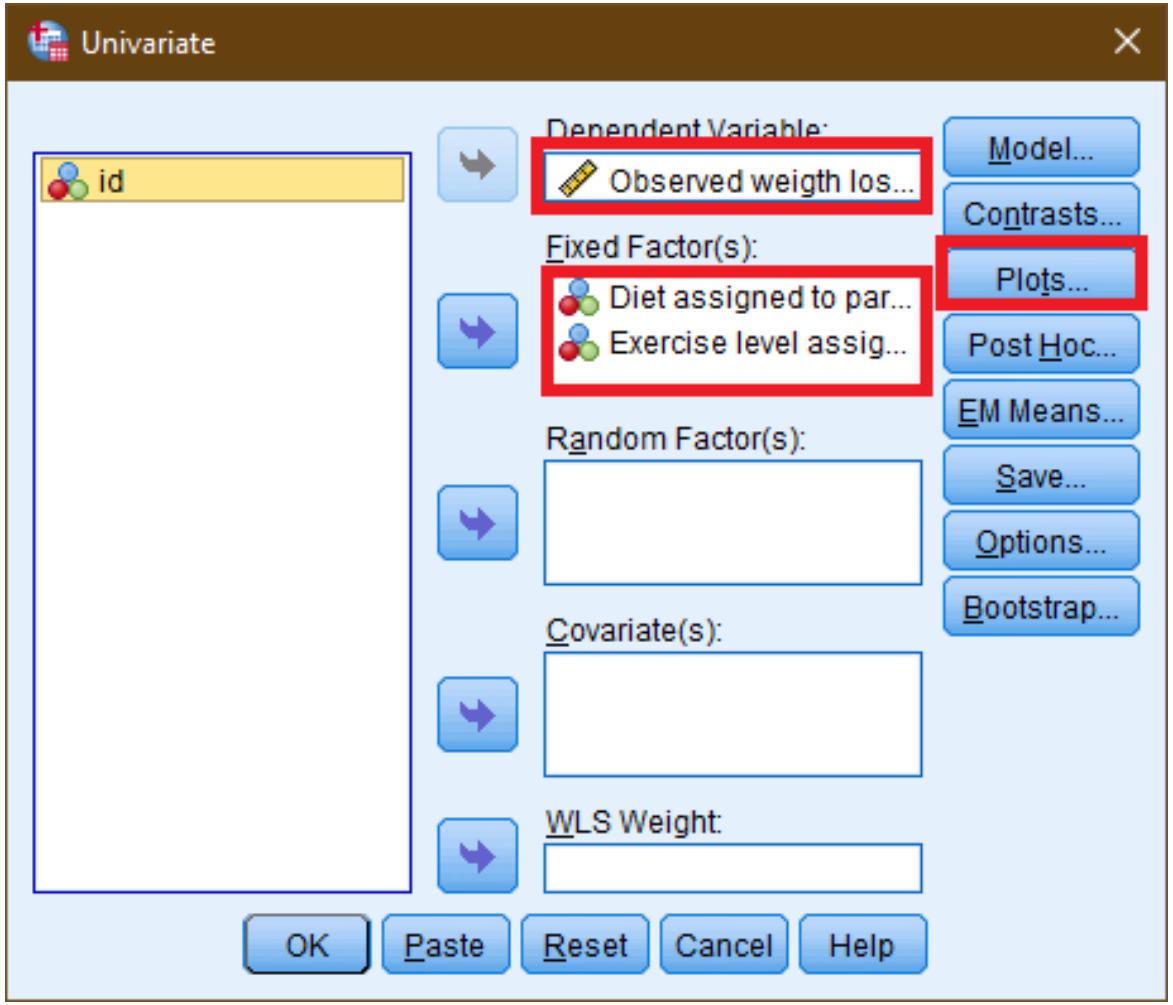
1. Analyze >> General Linear Model >> Univariate...



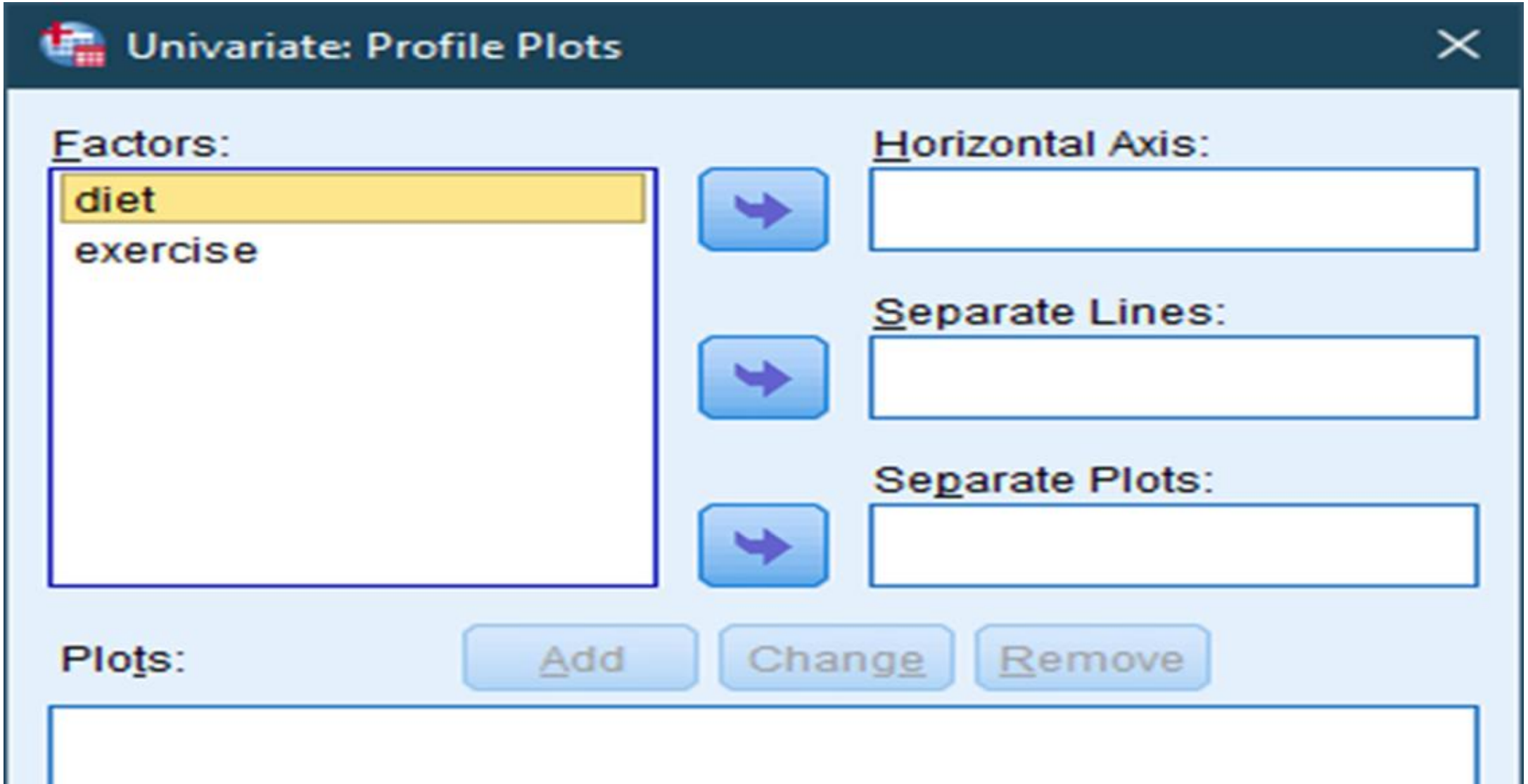
2. You will be presented with the "UNIVARIATE", as shown below:



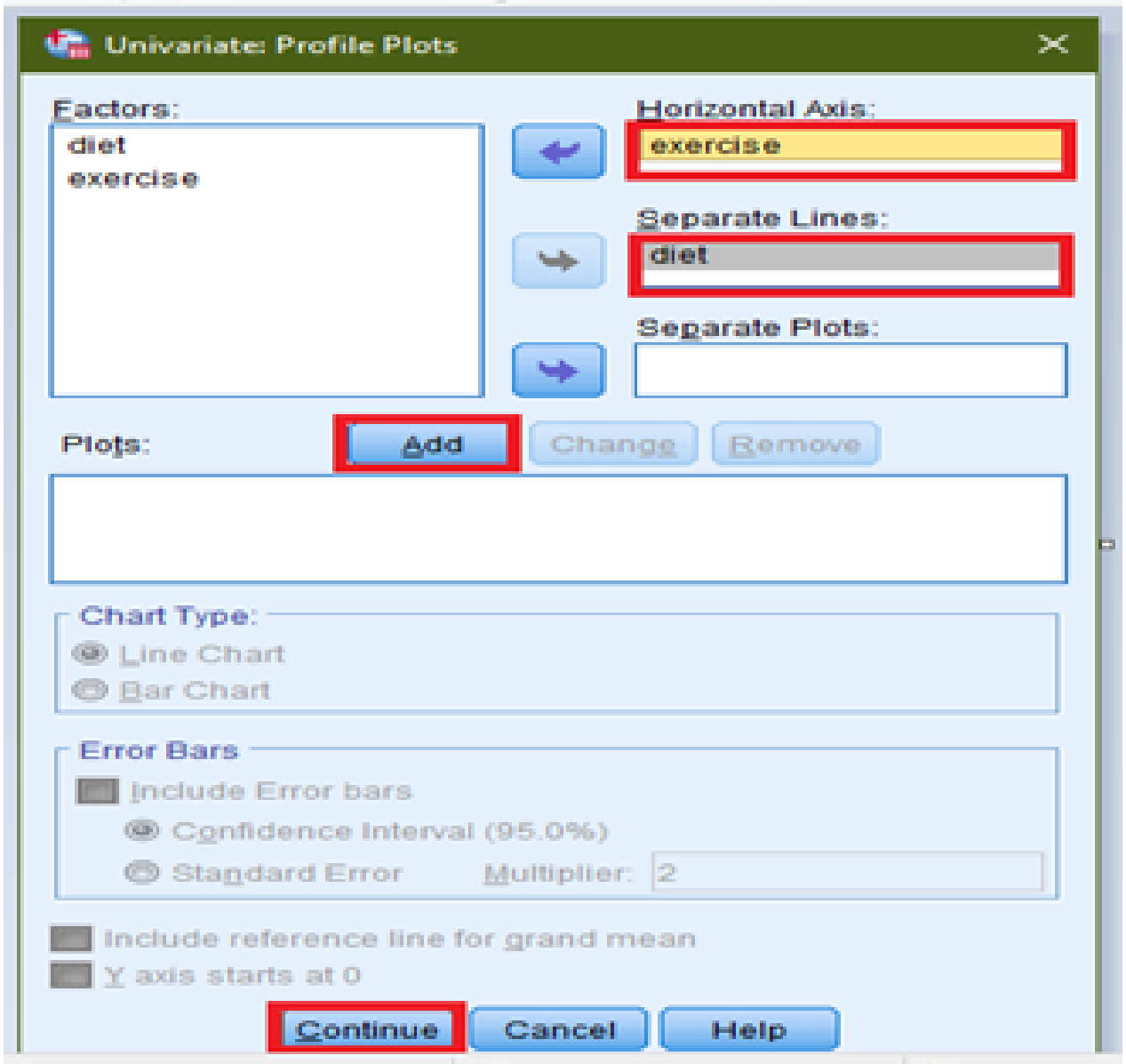
3. Add “Weight loss” into the “Dependent Variable” and Diet and Exercise into the "Fixed Factor(s)". After click on Plots



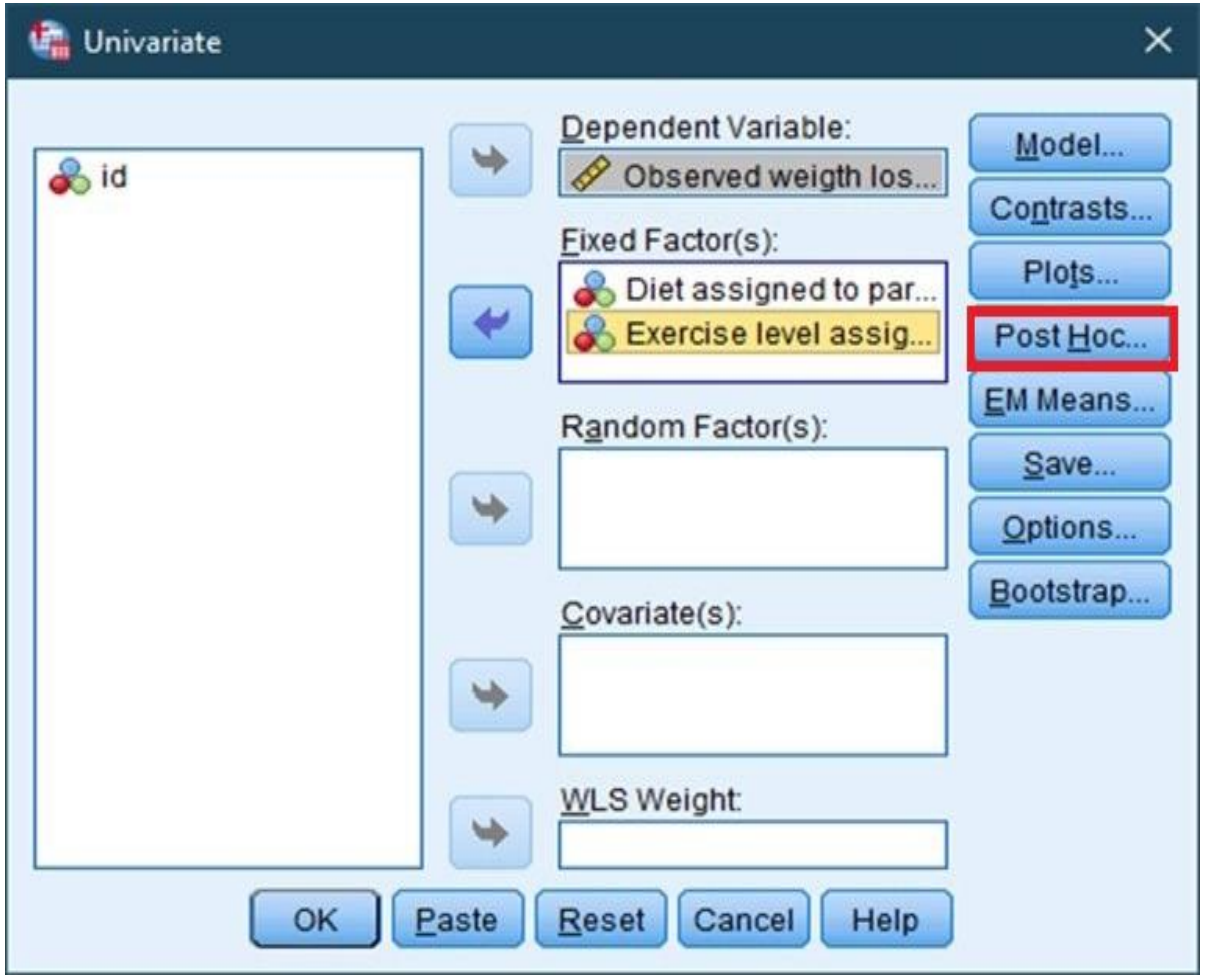
4. You'll be presented with the "Univariate: Profile Plots" window, as shown below:



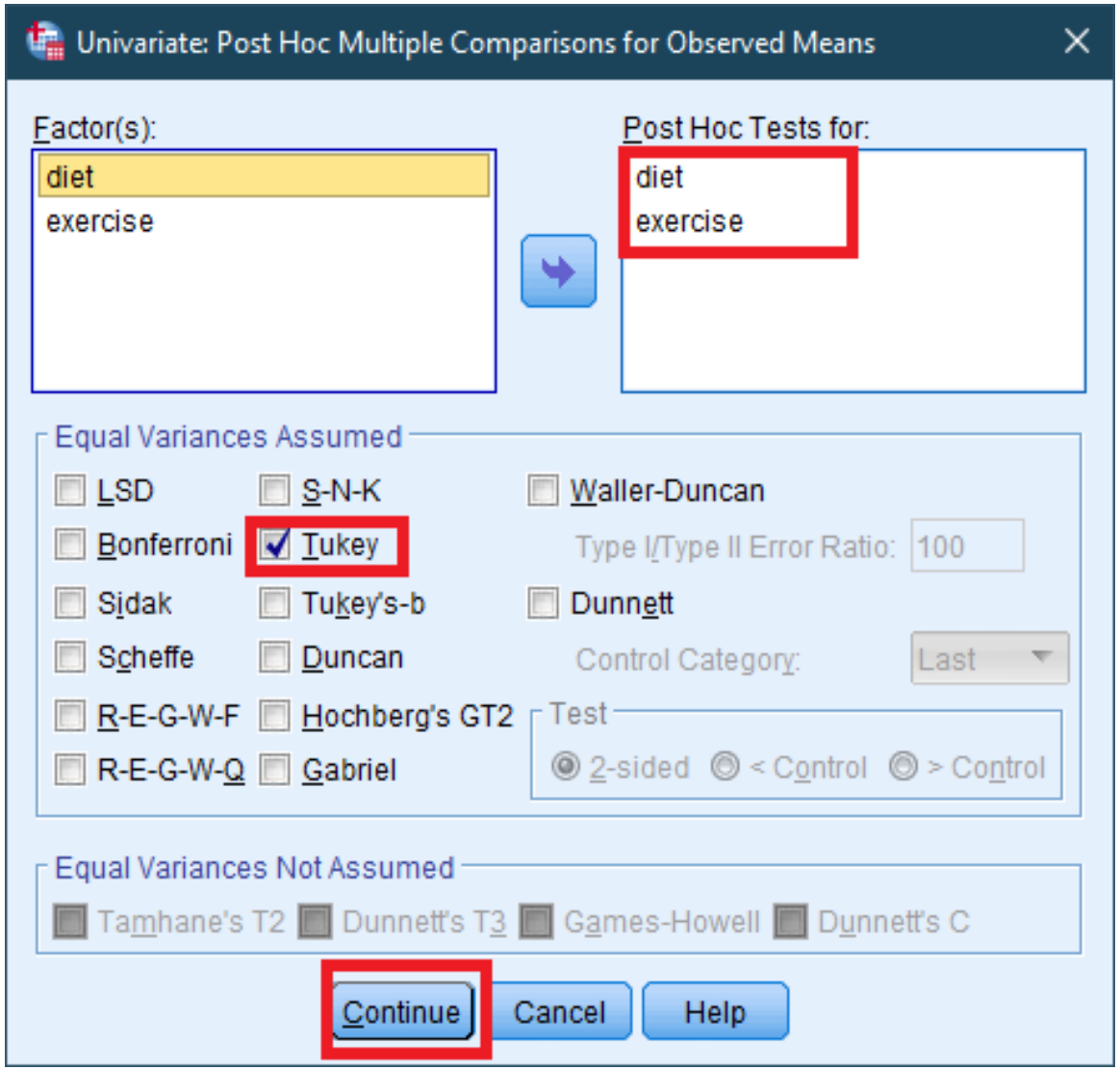
5. Transfer “Diet” and “Exercise” to the "Horizontal Axis" and "Separate lines". Click on "Continue" and then you’ll be returned to the Univariate dialogue box.



6. Click on "Post Hoc"

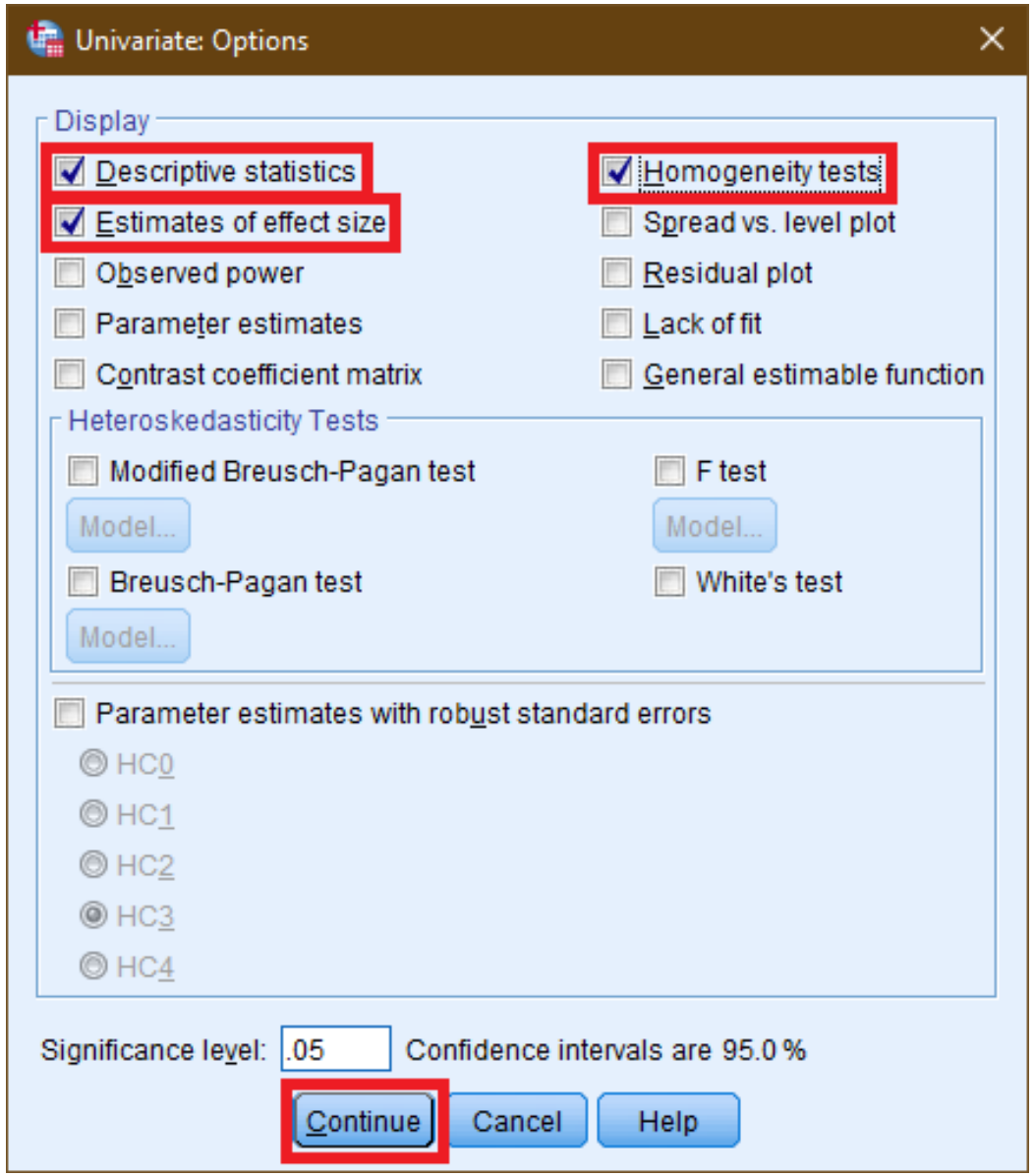


7.Transfer the factors to the "Post Hoc Tests for", after check "Tukey" >>Continue, which will bring you back to "Univariate" window

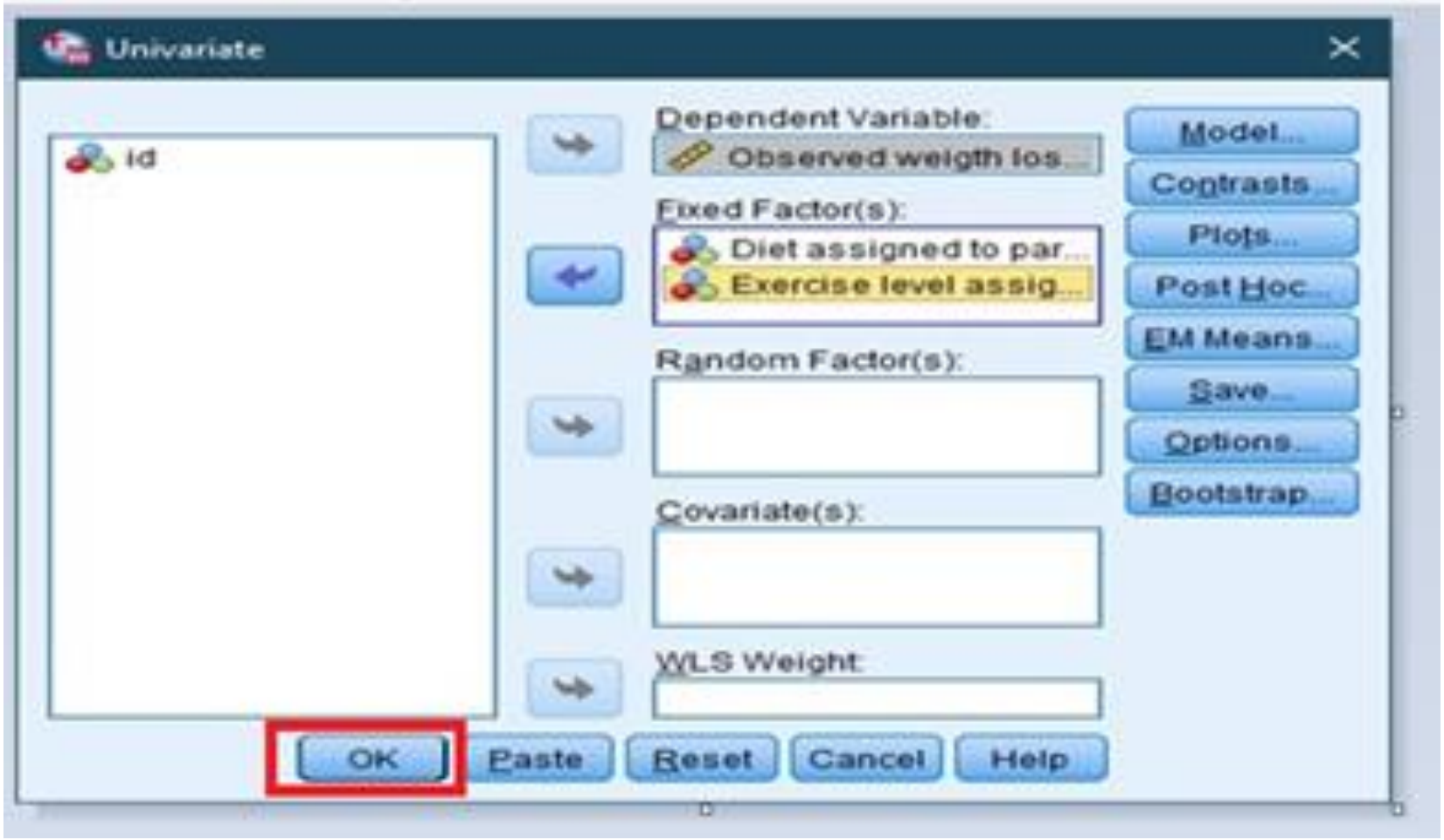


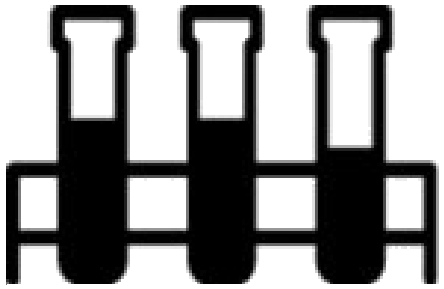
8. Click on the "Options"

9. “Options” >> “Descriptive Statistics” >> “Estimates of effect size” >> “Homogeneity tests” >> Continue>>OK

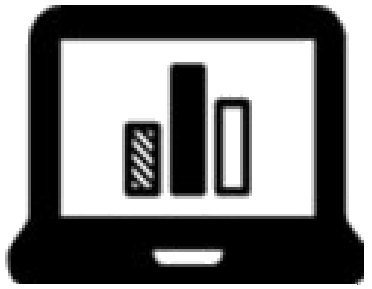


10. You will be back to "Univariate" window again. Click "OK" to run the analysis.





TESTS OF BETWEEN SUBJECT EFFECTS



Tests of Between-Subjects Effects

Dependent Variable: Observed weight loss in kilos over last 2 months

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1905.037[a]	8	238.130	27.647	.000
Intercept	3624.868	1	3624.868	420.844	.000
diet	363.904	2	181.952	21.124	.000
exercise	1508.859	2	754.429	87.589	.000
diet * exercise	32.274	4	8.069	.937	.444
Error	1472.879	171	8.613		
Total	7002.784	180			
Corrected Total	3377.916	179			

a. R Squared = .564 (Adjusted R Squared = .544)

Main Effects for Diet

$F(2, 171) = 21.124, p = .000$

Main effects for Exercise

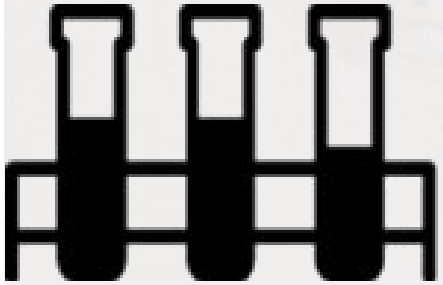
$F(2, 171) = 87,589,\ p = .000$

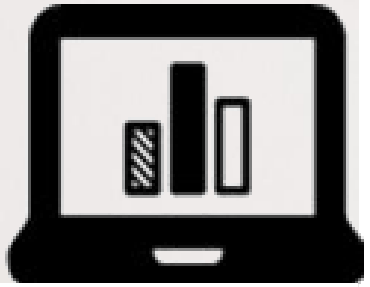
No interaction between diet and exercise

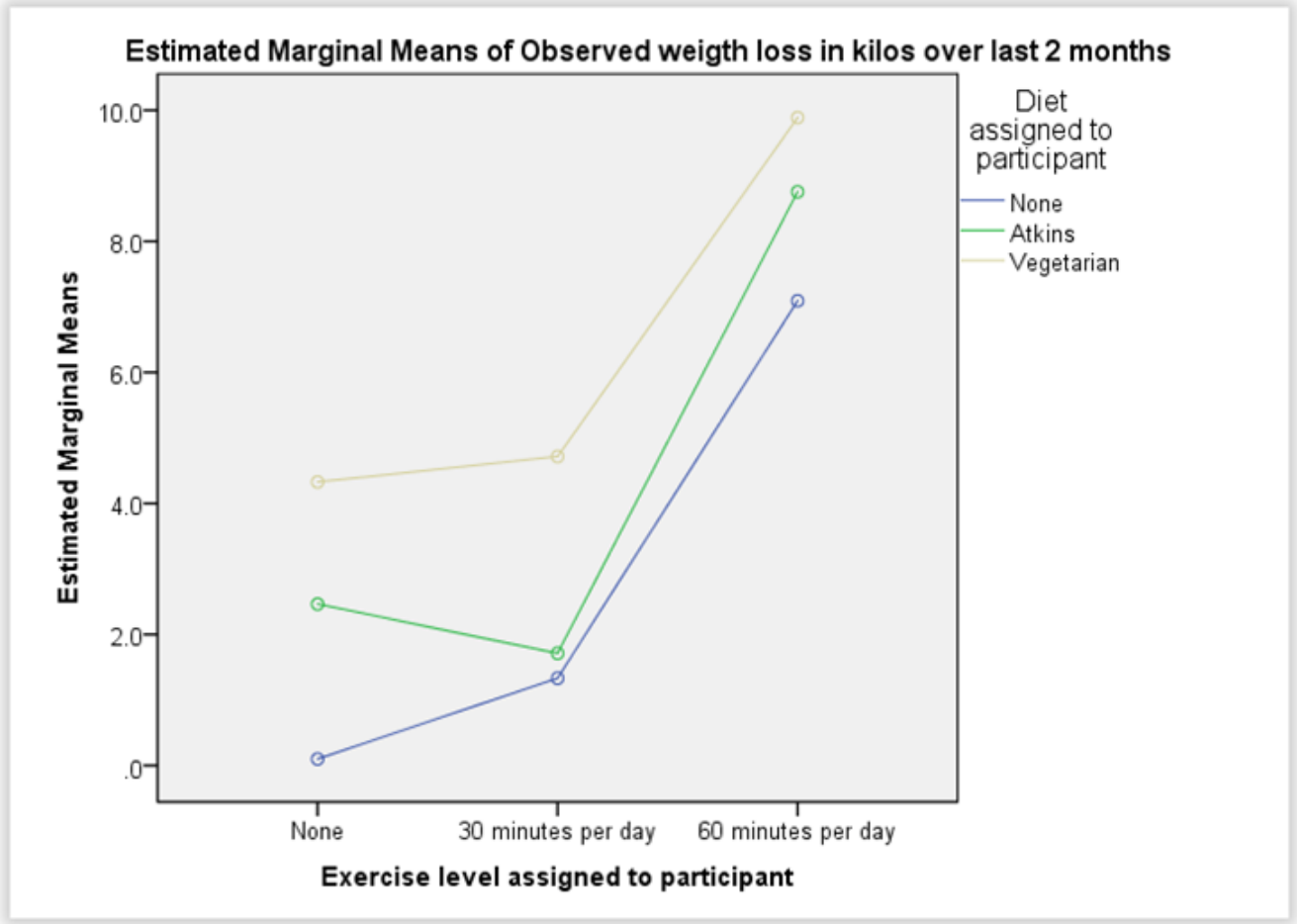
$F(4, 171) = .937, p = .444$

Retrieved from https://www.spss-tutorials.com/spss-two-way-anova-basics-tutorial/

TWO-WAY ANOVA OUTPUT:







Each line rises steeply between 30 to 60 minutes of exercise per day.

A vegetarian diet resulted in more weight loss than the other diets.

Both diet and exercise seem to have a ***main effect*** on weight loss.

Since these lines look pretty similar, the plot doesn't show much of an interaction effect.

Retrieved from https://www.spss-tutorials.com/spss-two-way-anova-basics-tutorial/

MULTIPLE COMPARISONS

Multiple Comparisons

Dependent Variable: Observed weight loss in kilos over last 2 months

Tukey HSD

(I) Exercise level assigned to participant	(J) Exercise level assigned to participant	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
None	30 minutes per day	-.289	.5358	.852	-1.556	.978
	60 minutes per day	-6.281*	.5358	.000	-7.548	-5.014
30 minutes per day	None	.289	.5358	.852	-.978	1.556
	60 minutes per day	-5.992*	.5358	.000	-7.259	-4.725
60 minutes per day	None	6.281*	.5358	.000	5.014	7.548
	30 minutes per day	5.992*	.5358	.000	4.725	7.259

Based on observed means.

The error term is Mean Square(Error) = 8.613.

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

The difference in weight loss between no exercise and 30 minutes is 0.29 kg.

The difference between no exercise and 60 min exercise is very big, about 6.28 kg. Also, p value shows that the difference is statistically significant.

Retrieved from https://www.spss-tutorials.com/spss-two-way-anova-basics-tutorial/

References

Two-way ANOVA in SPSS Statistics. (n.d.). Retrieved December 12, 2020, from https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics.php

ANOVA Test: Definition, Types, Examples. (2020, September 17). Retrieved December 12, 2020, from https://www.statisticshowto.com/probability-and-statistics/hypothesis-testing/anova/

MacKenzie, R. (2018, July 23). One-Way vs Two-Way ANOVA: Differences, Assumptions and Hypotheses. Retrieved December 12, 2020, from https://www.technologynetworks.com/informatics/articles/one-way-vs-two-way-anova-definition-differences-assumptions-and-hypotheses-306553

SPSS Two Way ANOVA – Basics Tutorial. (n.d.). Retrieved December 12, 2020, from https://www.spss-tutorials.com/spss-two-way-anova-basics-tutorial/

Musadoto Follow. (2017, November 16). 2 way ANOVA(Analysis Of VAriance. Retrieved December 12, 2020, from https://www2.slideshare.net/musadoto/2-way-anovaanalysis-of-variance?from_action=save

For a study testing three different fertilizers and their effects on the same type of plant, which test would be used to compare the means of the three groups?

Which test would you use for a study testing average running speed of people wearing two different types of shoes?

Which test would you use for a study testing the effect of yoga and tai chi on the blood pressure depending on the time of the day?