MANOVA

Learning Centre

JAMES COOK UNIVERSITY SINGAPORE

JAMES COOK UNIVERSITY AUSTRALIA

Cairns
Singapore
Townsville

www.jcu.edu.sg

Table of Contents

Multivariate Analysis of Variance (MANOVA)

JAMES COOK UNIVERSITY SINGAPORE

- The Multivariate Analysis of Variance (MANOVA) is an extension of the ANOVA
- While we only deal with ONE DV in ANOVA, MANOVA accounts for multiple DVs at once
- It wants to know if there are mean differences across groups on multiple DVs; it is suitable to test related DVs – e.g., testing depression, anxiety, and stress across groups at one go

Types of MANOVAs

- Similar to ANOVAs, there are between and within subjects MANOVAs
- If there is one IV, we call it a one-way between/within subjects MANOVA; if there are two IVs, we call it a two-way between/within MANOVA
- A test that mixes both between AND within IVs is called mixed MANOVA

We will focus on 'one-way MANOVA' in the next slides!

Example:
One-Way Between-Subject MANOVA

JAMES COOK UNIVERSITY SINGAPORE

I am interested in finding out if coffee consumption affects anxiety and fatigue levels.

To test this, I shall recruit 100 participants and randomly assign them into 2 groups: an experimental group who will drink a cup of coffee, and a control group who will drink a cup of water.

I will then ask each participant to rate their level of anxiety and fatigue.

Dr Tony Lim
World Class Researcher

Example:
One-Way Between-Subject MANOVA

In this example, we have 1 IV with 2 levels: Coffee vs. Water

We have 2 DVs: Anxiety, Fatigue

Thus, it is appropriate to conduct a one-way between subjects MANOVA

Location of SPSS Data Files

Example SPSS data for practice are available on LearnJCU:

Log in to LearnJCU -> Organisations -> Learning Centre JCU Singapore -> Statistics Support -> Statistics Resources -> SPSS Data for Practice

Assumptions Testing

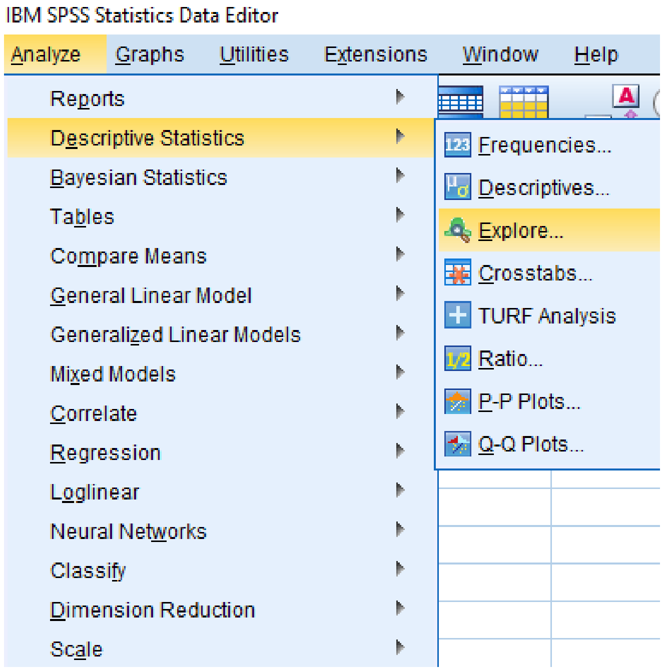
JAMES COOK UNIVERSITY SINGAPORE

1. Normality (Shapiro Wilk)
2. Univariate Outliers (Boxplots)
3. Multivariate Outliers (Mahalanobis Distances)
4. Multicollinearity (Correlation)
5. Linearity (Scatterplot)
6. Homogeneity of variance-covariance matrices (Box's *M*)

*You can meet other criteria before/during data collection, such as independence of observations (each participant can only take part in the study once), and ensuring adequate sample size in each cell (through a power analysis)

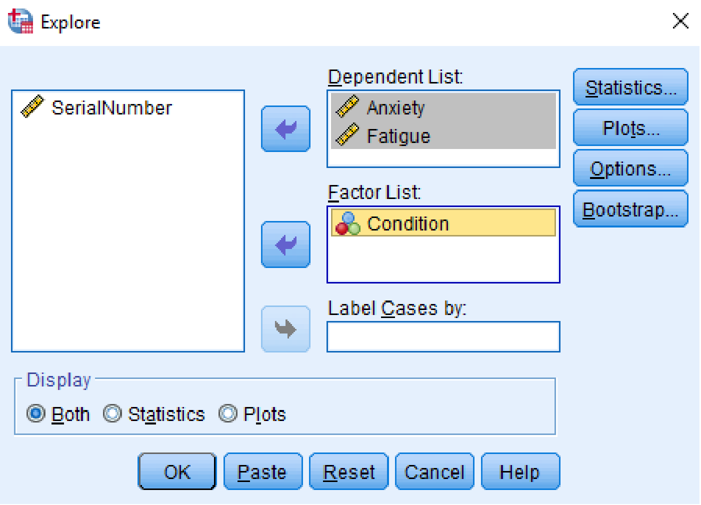
1. Normality

Go Analyze -> Descriptive Statistics -> Explore

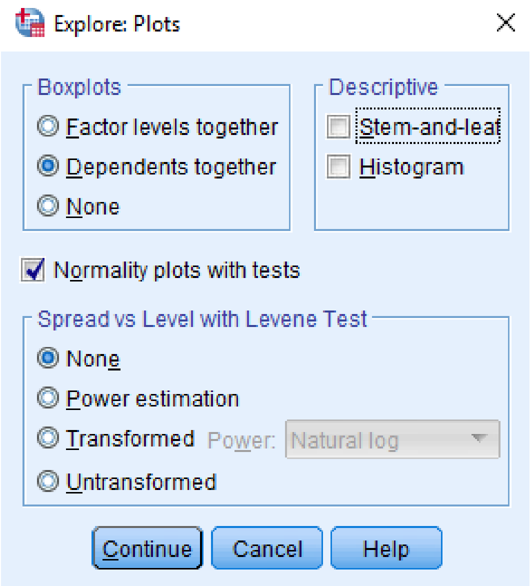


1. Normality

- Move 'Anxiety' and 'Fatigue' to the Dependent List, and 'Condition' to the Factor List



- Tick 'Normality plots with tests
- Continue and OK



1. Normality

Looking at Shapiro-Wilk tests, anxiety data in the one-cup water condition were not normally distributed, $p = .014$.

However, MANOVA is generally robust to a moderate violation of normality, we will continue to do the analysis for now.

Tests of Normality

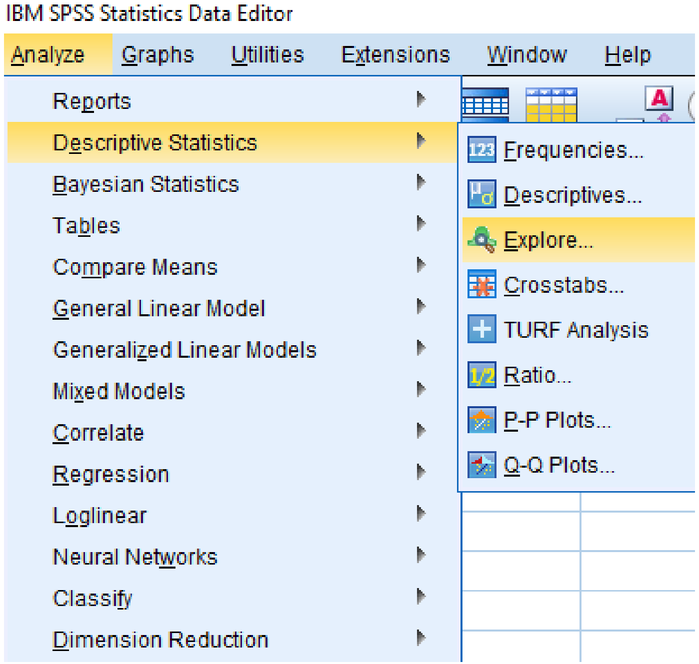
	Condition	Kolmogorov-Smirnov[a] Statistic	Kolmogorov-Smirnov[a] df	Kolmogorov-Smirnov[a] Sig.	Shapiro-Wilk Statistic	Shapiro-Wilk df	Shapiro-Wilk Sig.
Anxiety	1 Cup Coffee	.200	20	.035	.917	20	.088
	1 Cup water	.216	20	.016	.874	20	.014
Fatigue	1 Cup Coffee	.182	20	.080	.924	20	.117
	1 Cup water	.182	20	.082	.911	20	.066

a. Lilliefors Significance Correction

2. Univariate Outliers

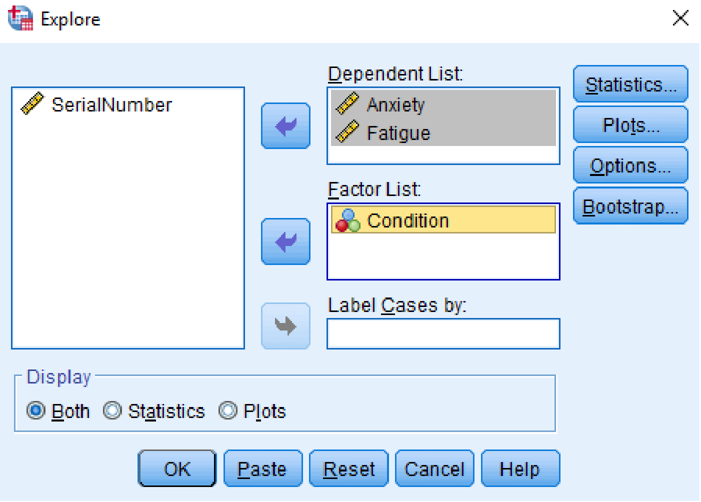
The assumption of univariate outliers can be tested via inspecting boxplots

- Go Analyze -> Descriptive Statistics -> Explore

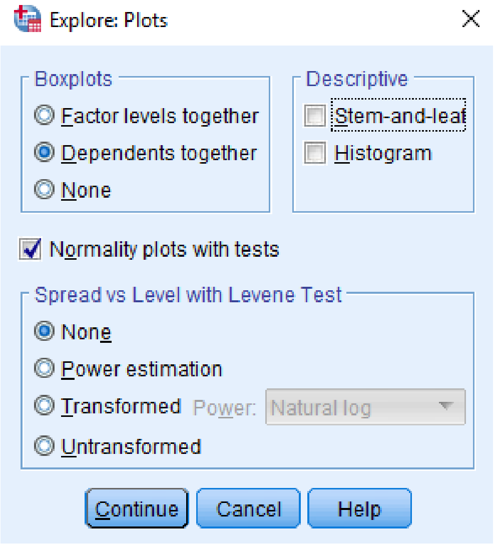


2. Univariate Outliers

- Move ‘Anxiety’ and ‘Fatigue’ to the Dependent List, and ‘Condition’ to the Factor List

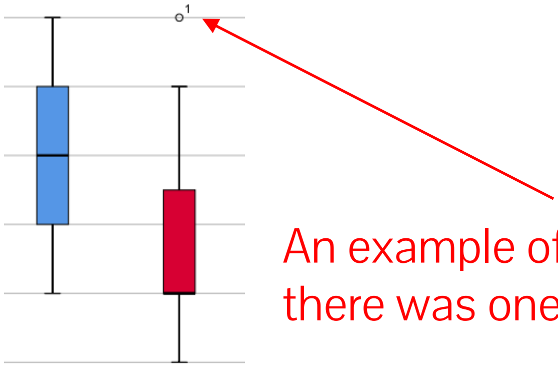


- Under Plots, select ‘Dependents together’
- Continue and OK

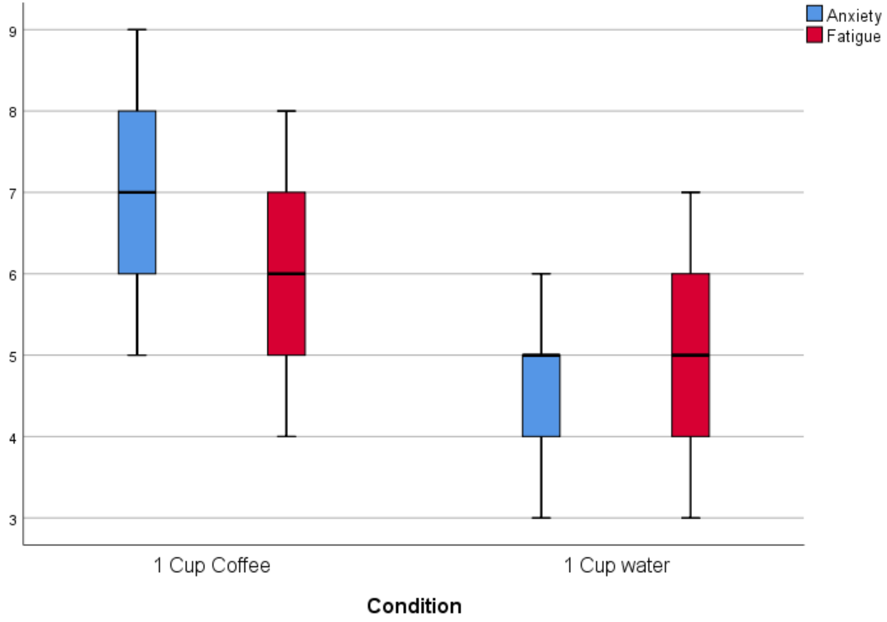


2. Univariate Outliers

Looking at boxplots on the right, we can assume that there are no univariate outliers.



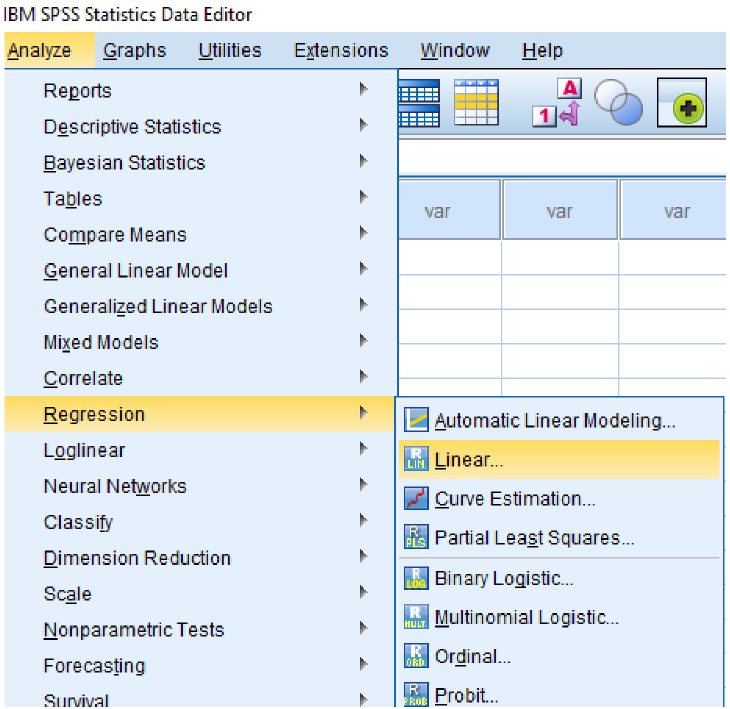
An example of an outlier (if there was one)



3. Multivariate Outliers

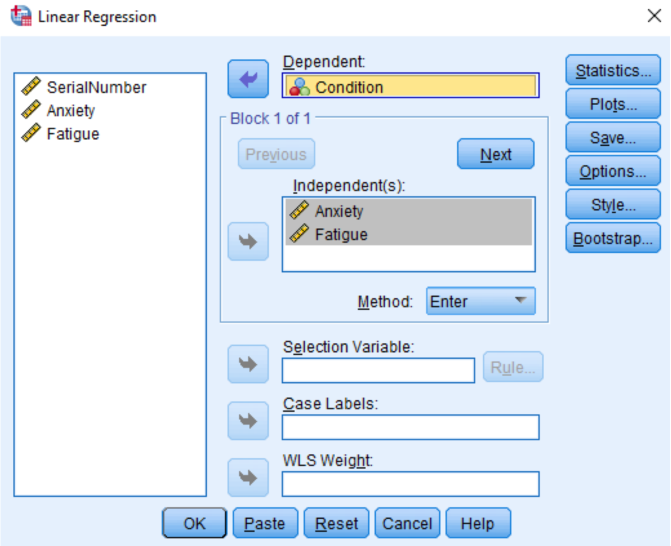
This assumption can be tested via the Mahalanobis Distances

- Analyze -> Regression -> Linear

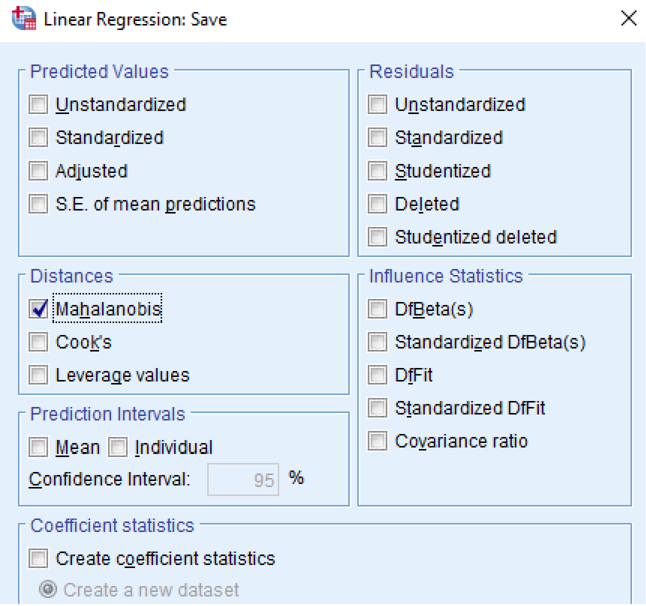


3. Multivariate Outliers

Move ‘Anxiety’ and ‘Fatigue’ to the Independent(s) box, and ‘Condition’ to the Dependent box



In Save, under *Distances*, select ‘Mahalanobis’, continue



3. Multivariate Outliers

Under <u>Residuals Statistics</u>, **Maximum** Malal. Distance = 5.267

This value is *smaller* than the chi-square value at *df* = 2, α = .05, which is 5.991

*Refer to a the critical value in the Chi-Square table; *df = number of DVs*

This indicates no multivariate outlier

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.68	2.31	1.50	.393	40
Std. Predicted Value	-2.074	2.071	.000	1.000	40
Standard Error of Predicted Value	.056	.131	.087	.022	40
Adjusted Predicted Value	.63	2.37	1.50	.397	40
Residual	-.752	.557	.000	.319	40
Std. Residual	-2.295	1.699	.000	.974	40
Stud. Residual	-2.339	1.724	.000	1.006	40
Deleted Residual	-.781	.573	.000	.341	40
Stud. Deleted Residual	-2.500	1.774	-.005	1.027	40
Mahal. Distance	.147	5.267	1.950	1.429	40
Cook's Distance	.000	.240	.023	.040	40
Centered Leverage Value	.004	.135	.050	.037	40

a. Dependent Variable: Condition

4. Multicollinearity

JAMES COOK UNIVERSITY SINGAPORE

The assumption of multicollinearity can be checked via a *correlation analysis*

- Go to Analyze -> Correlate -> Bivariate

*Check out how to run correlation analysis in the **Correlation** slides (JCUS Learning Centre website -> Statistics and Mathematics Support)

4. Multicollinearity

JAMES COOK UNIVERSITY SINGAPORE

In the Correlations table, two DVs are slightly correlated but *not* too strong, *r* = .172 (less than .7)

Therefore, no violation of multicollinearity

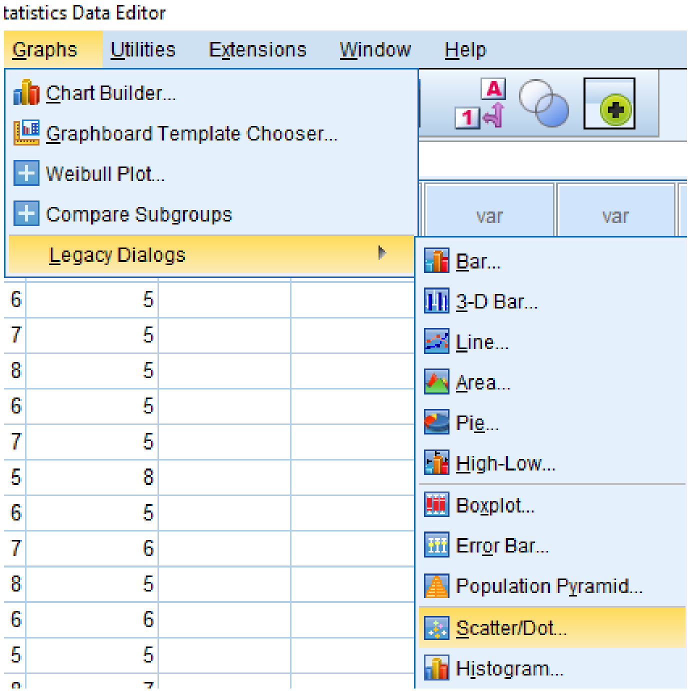
Correlations

		Condition	Anxiety	Fatigue
Pearson Correlation	Condition	1.000	-.766	-.258
	Anxiety	-.766	1.000	.172
	Fatigue	-.258	.172	1.000

5. Linearity

This assumption can be tested using *scatterplots*

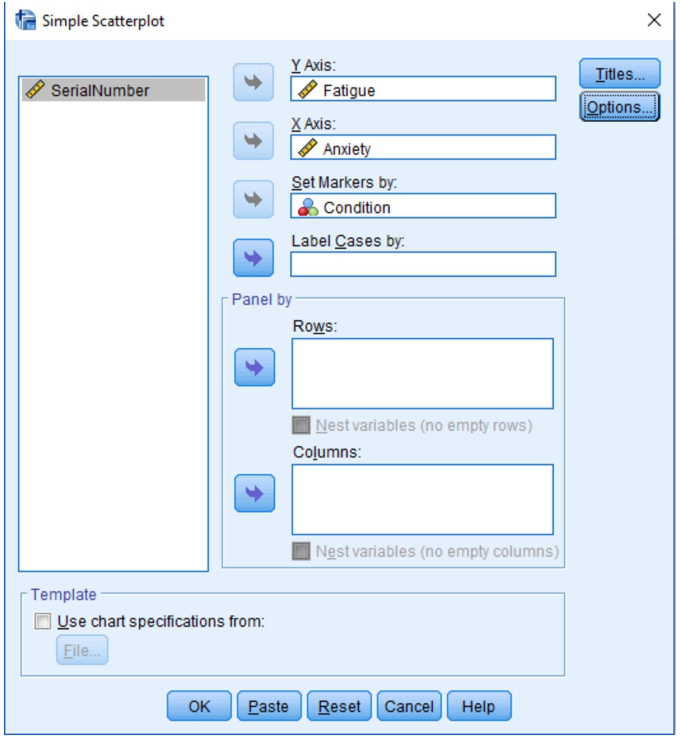
- Graphs -> Legacy Dialogs -> Scatter/Dot -> Simple Scatter -> Define



5. Linearity

Go Graphs -> Legacy Dialogs -> Scatter/Dot -> Simple Scatter -> Define

- Move 'Fatigue' as the Y axis, 'Anxiety' as the X axis, and Set Markers By: 'Condition'
- OK!

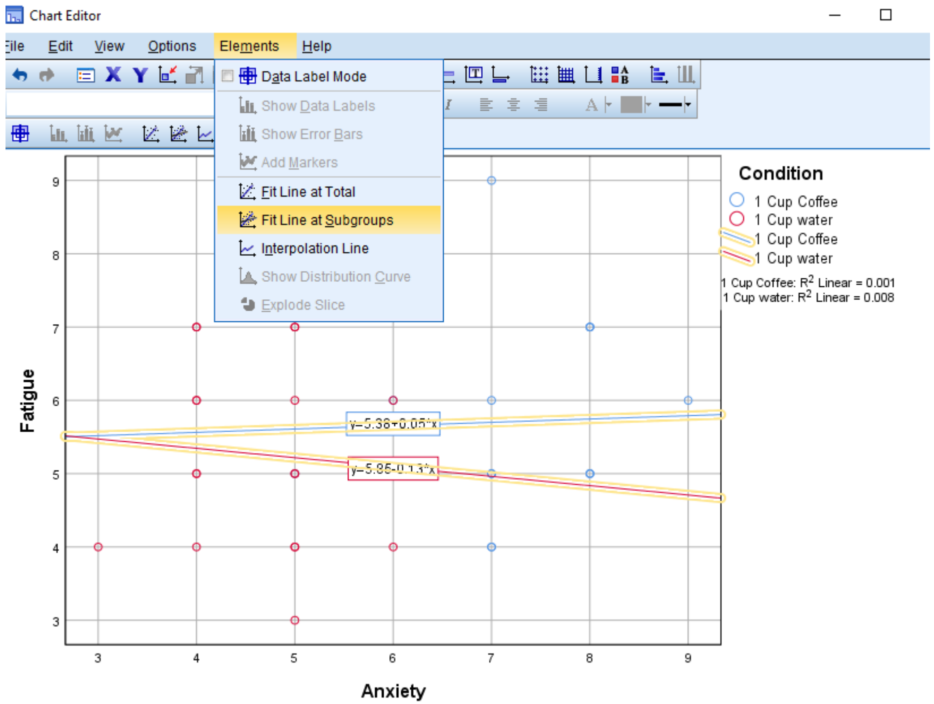


5. Linearity

On the output file, **double click** the scatterplot to open the chart editor

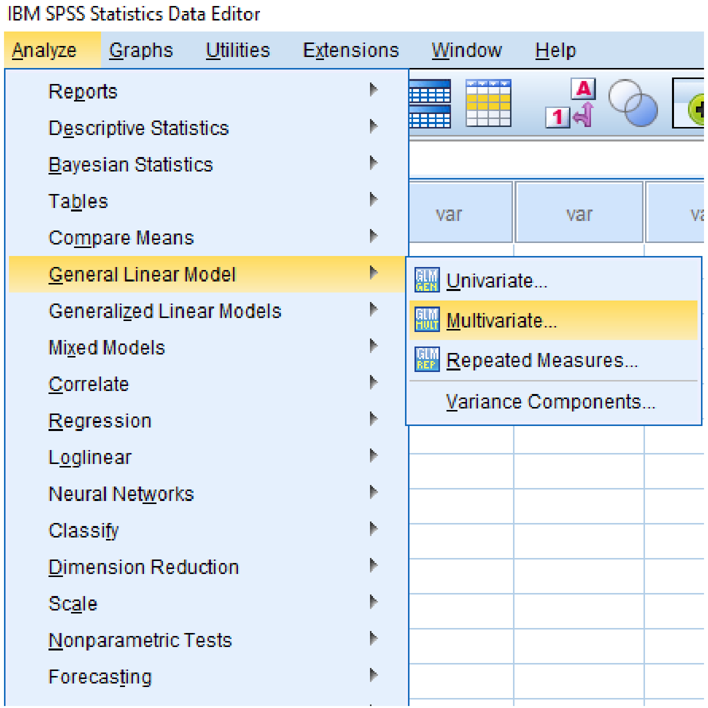
- Click on Elements -> Fit Line at Subgroups
- Ensure that 'Linear' is selected as the Fit Method

If the lines are roughly straight, we conclude that the assumption of linearity is satisfied



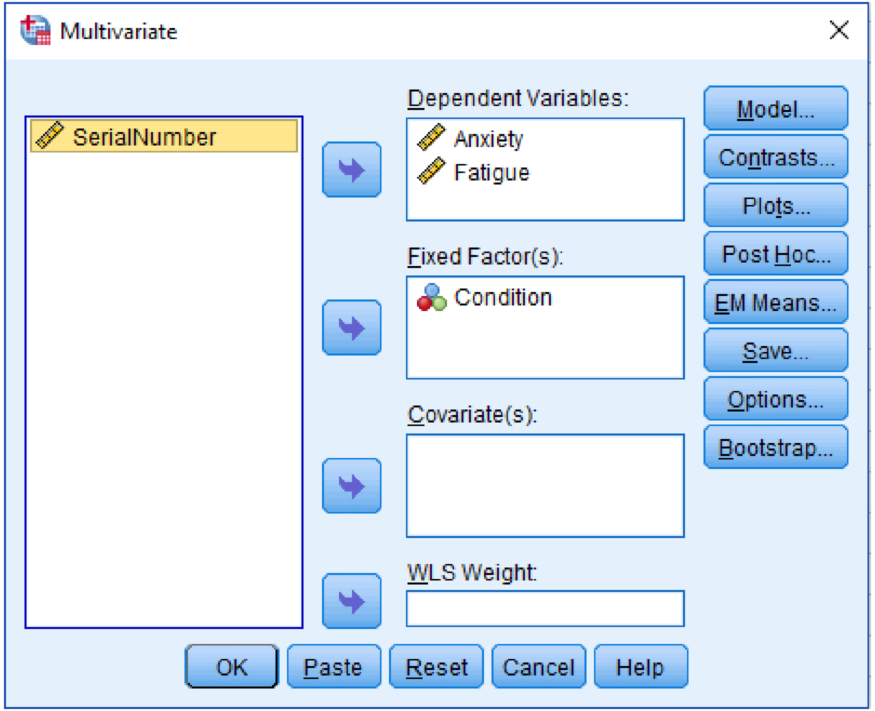
6. Homogeneity of variance-covariance matrices

- Analyze -> General Linear Model -> Multivariate



6. Homogeneity of variance-covariance matrices

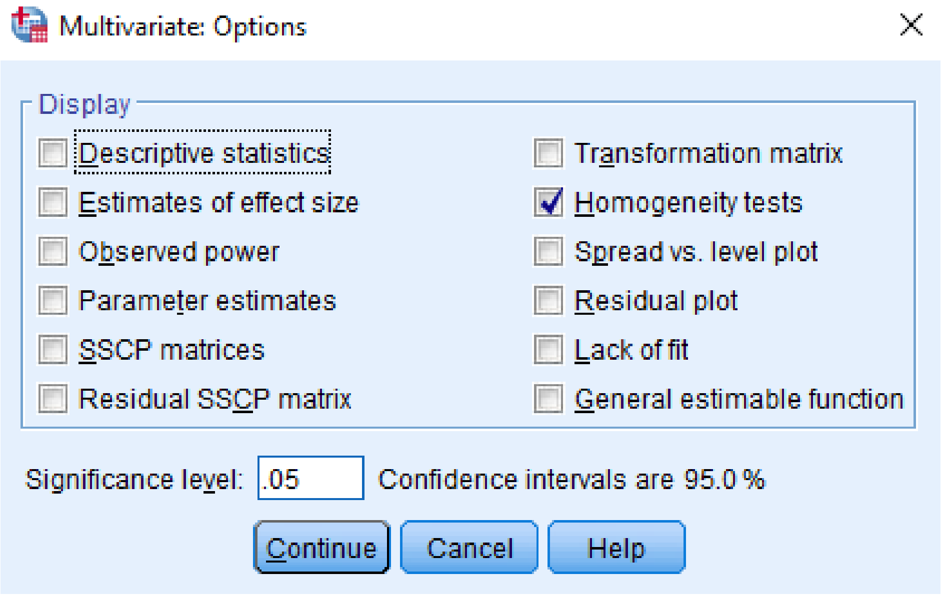
- Move 'Anxiety' and 'Fatigue' to Dependent Variables
- Move 'Condition' to Fixed Factor(s)



6. Homogeneity of variance-covariance matrices

- Under Options, select *Homogeneity tests*

- Continue, and OK



6. Homogeneity of variance-covariance matrices

In order to satisfy this assumption, the Box's *M* value should be *non-significant* at **α = .001**

A significant value of .743 indicates that the assumption has not been violated

Box's Test of Equality of Covariance Matrices[a]

Box's M	1.315
F	.413
df1	3
df2	259920.000
Sig.	.743

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + Condition

6. Homogeneity of variance-covariance matrices

Levene's Test of Equality of Error Variances[a]

		Levene Statistic	df1	df2	Sig.
Anxiety	Based on Mean	.289	1	38	.594
	Based on Median	.543	1	38	.466
	Based on Median and with adjusted df	.543	1	36.954	.466
	Based on trimmed mean	.318	1	38	.576
Fatigue	Based on Mean	.033	1	38	.856
	Based on Median	.026	1	38	.873
	Based on Median and with adjusted df	.026	1	32.666	.873
	Based on trimmed mean	.006	1	38	.941

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Condition

Output shows Levene's Test of Equality of Error Variances, where a *non-significant* Levene Statistic at $\alpha = .05$ would indicate *equality* of variances

Finally… **MANOVA**

How to choose the multivariate test?

*Look at how to conduct MANOVA in Slide 22

Multivariate Tests

Effect		Value	F
Intercept	Pillai's Trace	.983	1101.405[b]
	Wilks' Lambda	.017	1101.405[b]
	Hotelling's Trace	59.535	1101.405[b]
	Roy's Largest Root	59.535	1101.40
Condition	Pillai's Trace	.593	26.961[b]
	Wilks' Lambda	.407	26.961[b]
	Hotelling's Trace	1.457	26.961[b]
	Roy's Largest Root	1.457	26.961[b]

a. Design: Intercept + Condition

b. Exact statistic

Multivariate Test	Robustness					
	Sample Size	Levels of IVs	Uneven Cell Sizes	Unequal variance	Non-normal Data	Collinearity
Pillai's Trace	Small	> 2	Y	Y	Y	Low to medium
Wilk's Lambda	Medium to large	> 2	N	N	N	Low to medium
Hotelling's Trace	Medium to large	= 2	N	N	N	Low to medium
Roy's Largest Root	Medium to large	> 2	N	N	N	Medium to high

MANOVA

JAMES COOK UNIVERSITY SINGAPORE

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.983	1101.405[b]	2.000	37.000	.000
	Wilks' Lambda	.017	1101.405[b]	2.000	37.000	.000
	Hotelling's Trace	59.535	1101.405[b]	2.000	37.000	.000
	Roy's Largest Root	59.535	1101.405[b]	2.000	37.000	.000
Condition	Pillai's Trace	.593	26.961[b]	2.000	37.000	.000
	Wilks' Lambda	.407	26.961[b]	2.000	37.000	.000
	Hotelling's Trace	1.457	26.961[b]	2.000	37.000	.000
	Roy's Largest Root	1.457	26.961[b]	2.000	37.000	.000

a. Design: Intercept + Condition

b. Exact statistic

Looking at Pillai's Trace, $F(2,37) = 26.96$, $p < .001$.

There is a statistically significant difference in anxiety and fatigue across types of drinks.

MANOVA

JAMES COOK UNIVERSITY SINGAPORE

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Anxiety	50.625[a]	1	50.625	53.811	.000
	Fatigue	2.025[b]	1	2.025	1.284	.264
Intercept	Anxiety	1380.625	1	1380.625	1467.517	.000
	Fatigue	1199.025	1	1199.025	760.016	.000
Condition	Anxiety	50.625	1	50.625	53.811	.000
	Fatigue	2.025	1	2.025	1.284	.264
Error	Anxiety	35.750	38	.941		
	Fatigue	59.950	38	1.578		
Total	Anxiety	1467.000	40			
	Fatigue	1261.000	40			
Corrected Total	Anxiety	86.375	39			
	Fatigue	61.975	39			

a. R Squared = .586 (Adjusted R Squared = .575)

b. R Squared = .033 (Adjusted R Squared = .007)

To investigate the effects of each DV, look at the *Tests of Between-Subjects Effects table*

There is a main effect of drinks (coffee or water) on anxiety, $p < .001$, but not fatigue, $p = .264$

Something to note...

This example only contained 1 IV with 2 levels

If we had 3 levels (e.g., 1 cup coffee, 3 cups coffee, 1 cup water), we would have needed to conduct a pairwise comparison test to investigate which level of the IV significantly affected the DV?

This can be done by going to

-> Analyse -> General linear model -> Multivariate -> Post-Hoc -> Moving the IV to 'Post Hoc Tests for:' -> Selecting a preferred post hoc test (common test is Tukey)

Results Write-up

An example write-up can be found on page 167 in

Allen, P., Bennett, K., & Heritage, B. (2019). *SPSS Statistics: A Practical Guide* (4th ed.). Cengage Learning.

Any Questions?

learningcentre-singapore@jcu.edu.au