Data Analysis

Week 4

Analysis Of Variance - ANOVA

- ANOVA is a hypothesis test that is used to compare mean differences between **two or more** treatment conditions (or populations).

- H_0 = All group population means are equal.
- H_1 = At least one population mean is different from the others.

- This means that ANOVA test can show us if there is statistically significant difference among group means

ANOVA terminology

- FACTOR: an independent variable or quasi-independent variable
- LEVELS OF FACTOR: the groups that make the independent variable (CELLS)
- SINGLE-FACTOR DESIGN: a research study that involves only one factor
- FACTORIAL DESIGN: a study with more than one factor

Why ANOVA?

- Why don't we use multiple t-tests?

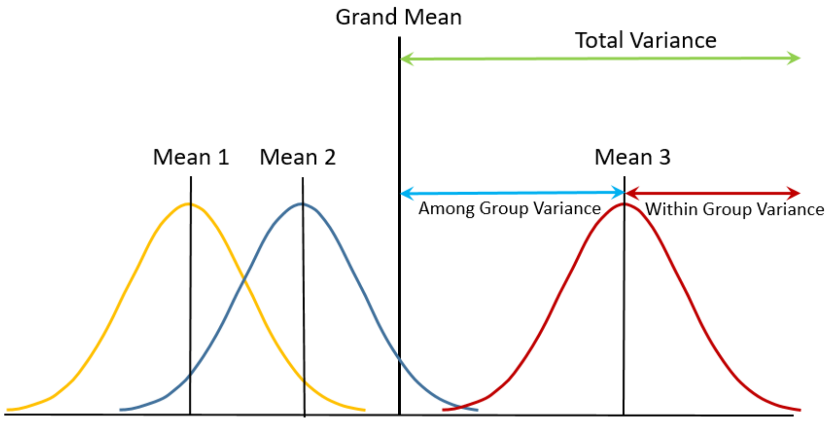
T-test
GROUP 1
GROUP 2
GROUP 3
T-test
T-test

- Each test carries a 5% risk of committing type I error
  - Running multiple tests for a single hypothesis inflates the risk
- [T-TEST] There is no significant difference between group 1 and group 2; group 2 and group 3; and group 1 and group 3
- [ANOVA] There is no significant difference among the three groups (looking at all of them at the same time)

ANOVA

- Analysis of covariance divides the total variability into two basic components:
  - Between-groups variance: variance due to differences between the group means
  - Within-group variance: variance due to differences within the groups

- Total variance is the result of both between-groups variance and within-group variance



Assumptions of One-Way ANOVA

- No significant outliers
- Independence of observations (within each sample)
- The dependent variable (approximately) **normally distributed** for each group of the independent variable (one-way ANOVA is considered "robust" to violations of normality)
  - This can often happen for very large datasets. If the distributions are all similarly skewed, this is not as troublesome. Only strong violations of normality require a non-parametric test.
- Homogeneity of variances between groups (i.e. the variance is equal in each group of your independent variable) – Levene's test > 0.05

Assessing normality

- If your groups are large, perform the S-W test on the whole dataset. Otherwise, it might be a good idea to check each cell of the design.

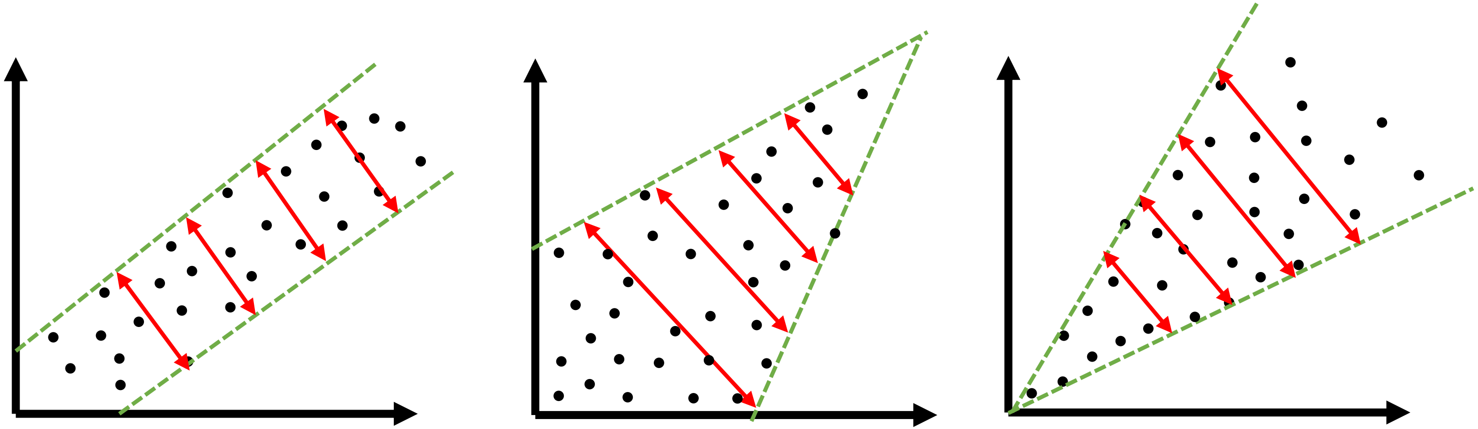
- SPSS: Data ⇢ Split File ⇢ Compare groups, Sort the file by grouping variables

Checking the normality of each cell:

1. Transfer the independent variables into “Groups based on”
2. Check the normality
3. Remove the split effect before proceeding
4. Perform a transformation if needed

Homogeneity of Variance

- Homogeneity of variance is the assumption that the spread of scores is roughly equal in different groups of cases, or more generally that the spread of scores is roughly equal at different points on the predictor variable.



Post-hoc tests

- A series of **pairwise comparisons** designed to compare all different combinations of the treatment groups with the aim of determining which mean differences are significant.
- Pairwise comparisons control the family-wise error by correcting the level of significance for each test such that the overall Type I error rate across all comparisons remains at .05.
- There are many post hoc tests (18 in SPSS!) and they all do the correction differently, but the choice is often based on the nature of the dataset and violations of the ANOVA assumptions.

One-Way ANOVA – Procedure

- SPSS: Analyze ⇢ Compare Means ⇢ One-Way ANOVA
- Options
  - Descriptive
  - Homogeneity of variance test
  - Welch
  - Means plot
- Post Hoc
  - Equal Variances Assumed: Tukey / Bonferroni / Gabriel
  - Equal Variances Not Assumed: Games-Howell

Interpreting results

- When interpreting results, consider these steps:

1. Determining if there are significant outliers
2. Determining if the dependent variable is normally distributed
3. Determining if there is homogeneity of variances
4. Determining if your sample sizes are equal
5. Choosing a post hoc test
6. Determining if there is at least one group with significantly different mean (ANOVA vs. Welch's ANOVA)
7. Determining which group means are different
8. Calculating the effect size

Interpreting results

3. Homogeneity of variances means that the population variances of the dependent variable are equal for all groups of the independent variable (Levene's test > 0.05).

4. Equal sample sizes means that the populations of the dependent variable are equal in size (check descriptives).

Test of Homogeneity of Variances

coping_stress

Levene Statistic	df1	df2	Sig.
2.129	3	27	.120

Descriptives

coping_stress

	N	Mean	Std. Deviation	Std. Error	95% Cor
					Lower Bou
Sedentary	7	4.1513	.77137	.29155	3.4
Low	9	5.8789	1.69131	.56377	4.5
Moderate	8	7.1228	1.57276	.55606	5.8
High	7	7.5054	1.24475	.47047	6.3
Total	31	6.1771	1.84480	.33134	5.5

Interpreting results

5. Choosing a post hoc test:

- If the sample sizes are equal and population variances are similar, use **REGWQ** or **Tukey** as they both have good power and tight control over the Type I error rate.
- If guaranteed control over the Type I error rate is needed, use **Bonferroni** as it is generally considered conservative.
- If sample sizes are slightly different use **Gabriel's** procedure because it has great power, but if sample sizes are very different use **Hochberg's GT2**.
- If there is any doubt that the population variances are equal, use the **Games-Howell** procedure because this offers the best performance. This post hoc test is often run in addition to other tests.

Interpreting results – (homogeneity of variance)

6. If the ANOVA is statistically significant ($p < .05$), it can be concluded that at least one group mean is different to another group mean.

ANOVA

coping_stress

	Sum of Squares	df	Mean Square	F	**Sig.**
Between Groups	49.033	3	16.344	8.316	**.000**
Within Groups	53.066	27	1.965		
Total	102.099	30			

ANOVA

coping_stress

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	49.033	3	16.344	8.316	.000
Within Groups	53.066	27	1.965		
Total	102.099	30			

$F(3, 27) = 8.316, p < .0005$

$F(\text{df1}, \text{df2}) = F\text{-statistic}, p = p\text{-value}$

Interpreting results – (heterogeneity of variance)

6. If the Welch ANOVA is statistically significant ($p < .05$), it can be concluded that at least one group mean is different to another group mean.

- Welch's ANOVA is used when homogeneity of variance is violated.

Robust Tests of Equality of Means

coping_stress

	Statistic[a]	df1	df2	**Sig.**
Welch	14.821	3	14.574	**.000**

a. Asymptotically F distributed.

Robust Tests of Equality of Means

coping_stress

	Statistic[a]	df1	df2	Sig.
Welch	14.821	3	14.574	.000

a. Asymptotically F distributed.

$F(3, 14.574) = 14.821, p < .0005$

$F(\text{df1}, \text{df2}) = F\text{-statistic}, p = p\text{-value}$

Interpreting results

7. The table **Multiple Comparisons** shows the results of a *post hoc* test (in this case Tukey HSD). It gives us all possible combinations of group differences.

- The value of "Mean Difference" is calculated as (I) group – (J) group.
- If significance value is < .05, that means that the two groups have significantly different means.

Multiple Comparisons

Dependent Variable: coping_stress
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Sedentary	Low	-1.72762	.70650	.092	-3.6610	.2058
	Moderate	-2.97153*	.72557	.002	-4.9571	-.9860
	High	-3.35409*	.74936	.001	-5.4048	-1.3034
Low	**Sedentary**	**1.72762**	**.70650**	**.092**	**-.2058**	**3.6610**
	Moderate	-1.24391	.68121	.284	-3.1081	.6203
	High	-1.62647	.70650	.123	-3.5599	.3069
Moderate	**Sedentary**	**2.97153***	**.72557**	**.002**	**.9860**	**4.9571**
	Low	**1.24391**	**.68121**	**.284**	**-.6203**	**3.1081**
	High	-.38256	.72557	.952	-2.3681	1.6030
High	**Sedentary**	**3.35409***	**.74936**	**.001**	**1.3034**	**5.4048**
	Low	**1.62647**	**.70650**	**.123**	**-.3069**	**3.5599**
	Moderate	**.38256**	**.72557**	**.952**	**-1.6030**	**2.3681**

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

Interpreting results

8. Calculating the effect size:

- Omega Squared (ω^2) represents a measure of effect size in the population.
- Partial Eta Squared (η^2) represents a measure of effect size in the sample.

SS_b df_b

ANOVA

coping_stress

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	49.033	3	16.344	8.316	.000
Within Groups	53.066	27	1.965		
Total	102.099	30			

SS_t MS_w

$$\hat{\omega}^2 = \frac{SS_b - (df_b)MS_w}{SS_t + MS_w} = \frac{49.033 - (3)1.965}{102.099 + 1.965} = 0.416$$

Tests of Between-Subjects Effects

Dependent Variable: coping_stress

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	49.033[a]	3	16.344	8.316	.000	.480
Intercept	1165.215	1	1165.215	592.865	.000	.956
group	49.033	3	16.344	8.316	.000	**.480**
Error	53.066	27	1.965			
Total	1284.948	31				
Corrected Total	102.099	30				

a. R Squared = .480 (Adjusted R Squared = .423)

Reporting One-Way ANOVA

A one-way ANOVA was conducted to determine if the ability to cope with workplace-related stress (CWWS score) was different for groups with different physical activity levels. Participants were classified into four groups: sedentary (n = 7), low (n = 9), moderate (n = 8) and high levels of physical activity (n = 7). There were no outliers, as assessed by boxplot; data was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > .05$); and there was homogeneity of variances, as assessed by Levene's test of homogeneity of variances ($p = .120$). Data is presented as mean ± standard deviation. CWWS score was statistically significantly different between different physical activity groups, $F(3, 27) = 8.316$, $p < .001$, $\omega^2 = 0.42$. CWWS score increased from the sedentary (M = 4.15, SD = 0.77) to the low (M = 5.88, SD = 1.69), moderate (M = 7.12, SD = 1.57) and high (M = 7.51, SD = 1.24) physical activity groups, in that order. Tukey post hoc analysis revealed that the mean increase from sedentary to moderate (2.97, 95% CI [0.99, 4.96]) was statistically significant ($p = .002$), as well as the increase from sedentary to high (3.35, 95% CI [1.30, 5.40], $p = .001$), but no other group differences were statistically significant.

Two-Way ANOVA

- Enables to determine whether there is an interaction between two independent variables

- Assumptions and the choice of the post hoc test are the same with One-Way ANOVA (no significant outliers, independence of observations, normality of dependent variable, homogeneity of variance of dependent variable)

- SPSS: Analyze ⇢ General Linear Model ⇢ Univariate
  - Fixed factor(s): your independent variables

Two-Way ANOVA - Procedure

- Plot:
  - factor 1 x factor 2 ⇢ Add
  - factor 2 x factor 1 ⇢ Add
- EM Means
  - Display Means for: add “factor 1 x factor 2”
- Options
  - Descriptive Statistics
  - Estimates of effect size
  - Homogeneity tests
- Save
  - Predicted Values: Unstandardized
  - Residuals: Unstandardized
  - Residuals: Studentized
- Post hoc:
  - Bonferroni

Interpreting results

- When interpreting results, consider these steps:

1. Determining if there are significant outliers
2. Determining if the dependent variable is normally distributed
3. Determining if there is homogeneity of variances
4. Determining if your sample sizes are equal
5. Choosing a post hoc test
6. Determining if there is an interaction between the two independent variables
7. Calculating main effect

Interpreting results

6. Determining if there is an interaction between the two independent variables

- Significance value of interaction effect must be < .05

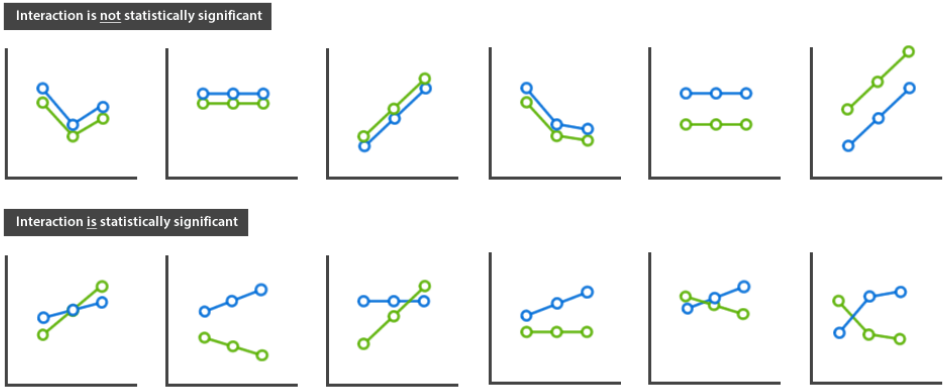
- Plot inspection

Tests of Between-Subjects Effects

Dependent Variable: political_interest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5645.998[a]	5	1129.200	78.538	.000	.883
Intercept	132091.906	1	132091.906	9187.227	.000	.994
gender	8.420	1	8.420	.586	.448	.011
education_level	5446.697	2	2723.348	189.414	.000	.879
gender * education_level	**210.338**	**2**	**105.169**	**7.315**	**.002**	**.220**
Error	747.644	52	14.378			
Total	140265.750	58				
Corrected Total	6393.642	57				

a. R Squared = .883 (Adjusted R Squared = .872)



Interpreting results

7. Calculating main effect I

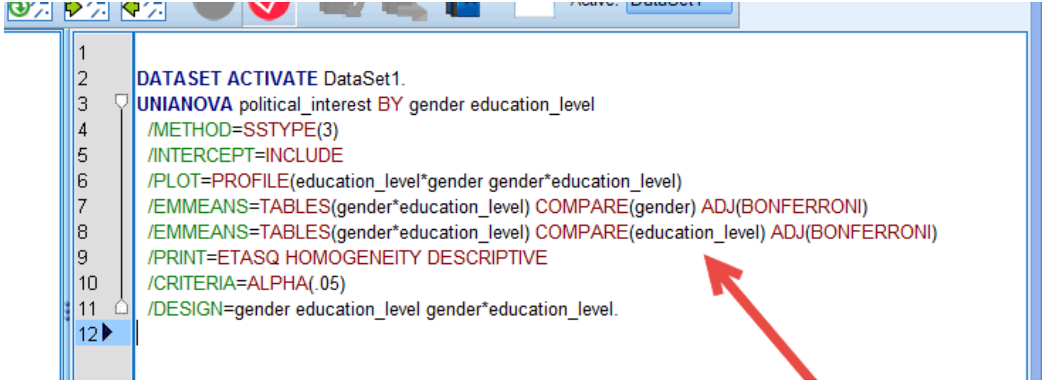
- SPSS: Analyze ⇢ General Linear Model ⇢ Univariate
  - Previously selected options should remain selected

- Save
  - Deselect: Predicted Values: Unstandardized
  - Deselect: Residuals: Unstandardized
  - Deselect: Residuals: Studentized

Interpreting results

7. Calculating main effect II

- Paste
  - Add “COMPARE(factor 1)ADJ(BONFERRONI)” to line 7
  - Copy line 7, paste it below and change the independent variable in the brackets to factor 2

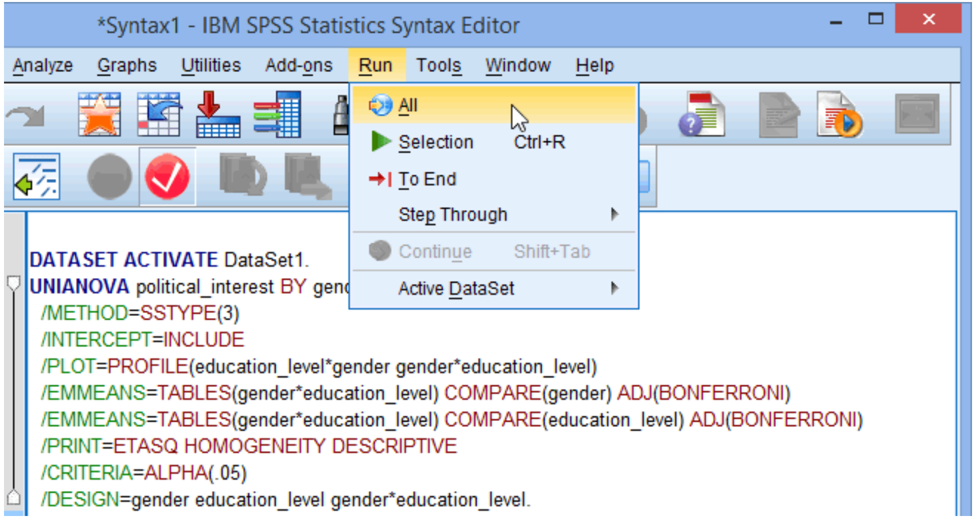


```
1
2  DATASET ACTIVATE DataSet1.
3  UNIANOVA political_interest BY gender education_level
4    /METHOD=SSTYPE(3)
5    /INTERCEPT=INCLUDE
6    /PLOT=PROFILE(education_level*gender gender*education_level)
7    /EMMEANS=TABLES(gender*education_level) COMPARE(gender) ADJ(BONFERRONI)
8    /EMMEANS=TABLES(gender*education_level) COMPARE(education_level) ADJ(BONFERRONI)
9    /PRINT=ETASQ HOMOGENEITY DESCRIPTIVE
10   /CRITERIA=ALPHA(.05)
11   /DESIGN=gender education_level gender*education_level.
12
```

Interpreting results

7. Calculating main effect III

- Go to Run ⇢ All



Interpreting results

7. Calculating main effect IV

- Consult Tests of Between-Subjects Effects to check the significance of each independent variable. For the one(s) with significance level <.05, proceed to do the next step.

Tests of Between-Subjects Effects

Dependent Variable: political_interest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5645.998[a]	5	1129.200	78.538	.000	.883
Intercept	132091.906	1	132091.906	9187.227	.000	.994
gender	8.420	1	8.420	.586	.448	.011
education_level	**5446.697**	**2**	**2723.348**	**189.414**	**.000**	**.879**
gender * education_level	210.338	2	105.169	7.315	.002	.220
Error	747.644	52	14.378			
Total	140265.750	58				
Corrected Total	6393.642	57				

a. R Squared = .883 (Adjusted R Squared = .872)

Interpreting results

7. Calculating main effect V

- If you have unbalanced design (unequal number of participants in each cell), use Estimates and Pairwise Comparisons Table. Otherwise use Descriptive Statistics Table and Multiple Comparisons Table.

Estimates

Dependent Variable: political_interest

education_level	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
School	**38.522**	.871	36.774	40.270
College	**43.772**	.871	42.024	45.520
University	**61.050**	.848	59.349	62.751

Pairwise Comparisons

Dependent Variable: political_interest

(I) education_level	(J) education_level	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
School	College	-5.250*	1.232	.000	-8.2[illegible]	[illegible]
	University	-22.528*	1.216	.000	-25.53[illegible]	-19.520
College	**School**	**5.250***	**1.232**	**.000**	**2.202**	**8.298**
	University	-17.278*	1.216	.000	-20.285	-14.270
University	School	22.528*	1.216	.000	19.520	25.535
	College	17.278*	1.216	.000	14.270	20.285

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Reporting Two-Way ANOVA

A two-way ANOVA was conducted to examine the effects of gender and education level on interest in politics. Residual analysis was performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's normality test for each cell of the design and homogeneity of variances was assessed by Levene's test. There were no outliers, residuals were normally distributed ($p > .05$) and there was homogeneity of variances ($p = .061$).

There was a statistically significant interaction between gender and level of education on interest in politics, $F(2, 52) = 7.315$, $p = .002$, partial $\eta^2 = .220$. Therefore, an analysis of simple main effects for education level was performed with statistical significance receiving a Bonferroni adjustment and being accepted at the $p < .025$ level. There was a statistically significant difference in mean "Political Interest" scores for females educated to either school, college or university level, $F(2, 52) = 62.96$, $p < .0005$, partial $\eta^2 = .708$, as for males, $F(2, 52) = 132.493$, $p < .0005$, partial $\eta^2 = .836$.

All pairwise comparisons were run for each simple main effect with reported 95% confidence intervals and *p*-values Bonferroni-adjusted within each simple main effect. Mean "Political Interest" scores for school-educated, college-educated, and university-educated females were 39.60 ($SD = 3.27$), 44.60 ($SD = 3.27$) and 58.00 ($SD = 6.46$), respectively. School-educated females had a statistically significantly lower mean "Political Interest" score than college-educated females, 5.00, 95% CI [0.81, 9.20], $p = .014$, and university-educated females, 18.40, 95% CI [14.21, 22.60], $p < .001$. College-educated females also had a statistically significantly lower mean "Political Interest" score than university-educated females, 13.40, 95% CI [9.21, 17.60], $p < .001$.

Mean "Political Interest" scores for school-educated, college-educated, and university-educated males were 37.44 ($SD = 2.51$), 42.94 ($SD = 2.34$) and 64.10 ($SD = 3.07$), respectively. School-educated males had a statistically significantly lower mean "Political Interest" score than college-educated males, 5.50, 95% CI [1.08, 9.92], $p = .010$, and university-educated males, 26.66, 95% CI [22.35, 30.97], $p < .001$. College-educated males also had a statistically significantly lower mean "Political Interest" score than university-educated males, 21.16, 95% CI [16.85, 25.47], $p < .001$.

Beyond One-Way and Two-Way ANOVA

One independent variable

When the groups consist of the same participants

When there is more than one dependent variable

One-way
Two-way
Three-way

(Repeated-measures)
(Between-subject f.)
(Mixed)

(M)AN(C)OVA

Interaction of two independent variables

Has both within-subject factor and between-subject factor

When a covariate is added

Interaction of three independent variables

ANOVA family

- One-Way ANOVA
- Two-Way ANOVA
- Three-Way ANOVA
- One-Way repeated measures ANOVA
- Two-Way repeated measures ANOVA
- Three-Way repeated measures ANOVA
- Two-Way mixed ANOVA
- Three-Way mixed ANOVA
- One-Way ANCOVA
- Two-Way ANCOVA
- One-Way MANOVA
- Two-Way MANOVA
- One-Way MANCOVA

References

- Cohen, J. (1988). *Statisticsl power analysis for the behavioral sciences* (2nd ed.). New York: Psychology Press.
- Field, A. (2009). *Discovering statistics using SPSS*. London: Sage.
- Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Prentice Hall
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). New York: Psychology Press
- Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM approach* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Sheskin, D.J. (2011). *Handbook of parametric and nonparametrics statistical procedures* (5th ed.). Boca Raton: Chapman & Hall/CRC Press.