Comparing Means, samples t-tests, and P-values in SPSS

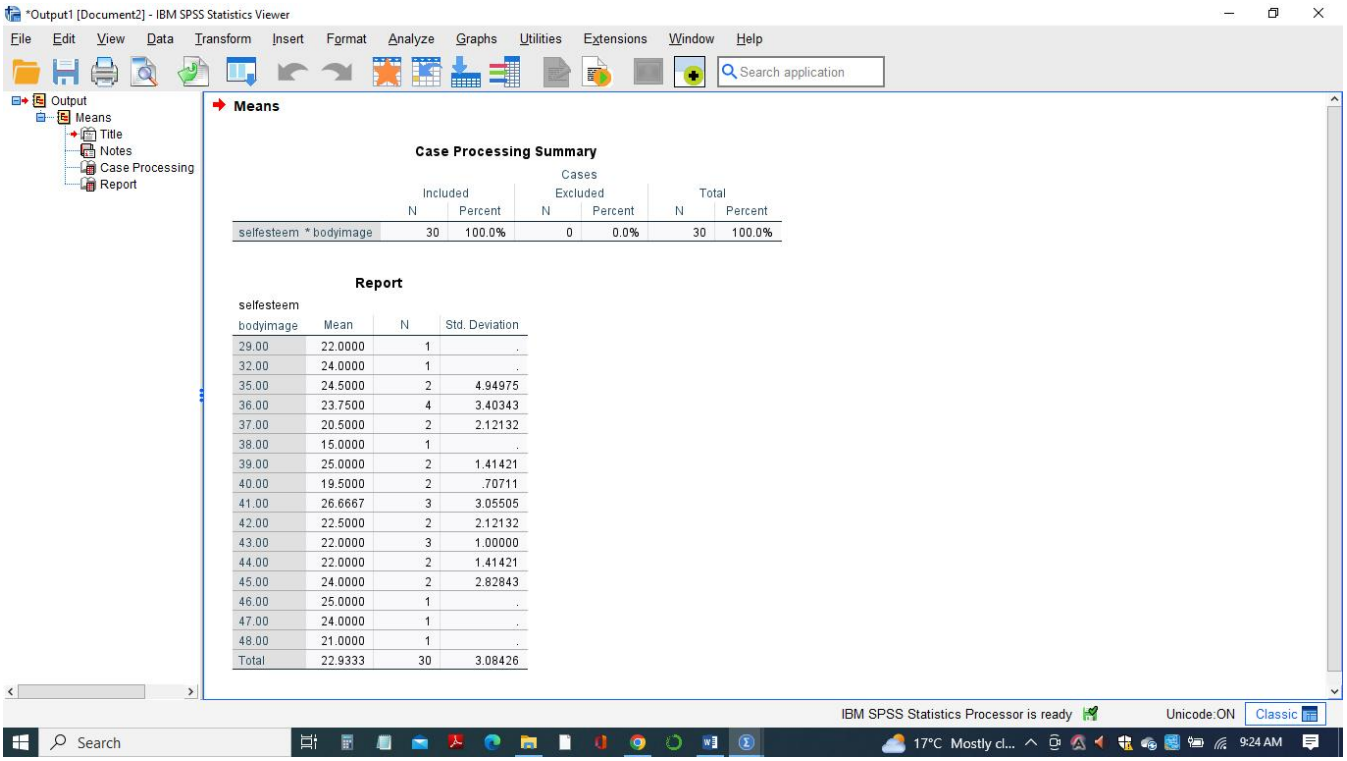
Vincent Were

Comparing Means, samples t-tests, and P-values in SPSS

Introduction

In many situations, you will wish to compare the means, samples t-tests, and p-values of two populations or samples in statistics. The method one employs to calculate and compare the mean relies on the kind of data an individual has and how it is organized. A statistical test t-test is used to compare the means of the two groups. It frequently appears in hypothesis testing. While the P-value calculates the likelihood of getting outcomes observed, presuming that the null hypothesis is correct (Park,2019).

1a) Compare Means function on the Body Image and Self Esteem variables



Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
selfesteem * bodyimage	30	100.0%	0	0.0%	30	100.0%

Report

selfesteem

bodyimage	Mean	N	Std. Deviation
29.00	22.0000	1	.
32.00	24.0000	1	.
35.00	24.5000	2	4.94975
36.00	23.7500	4	3.40343
37.00	20.5000	2	2.12132
38.00	15.0000	1	.
39.00	25.0000	2	1.41421
40.00	19.5000	2	.70711
41.00	26.6667	3	3.05505
42.00	22.5000	2	2.12132
43.00	22.0000	3	1.00000
44.00	22.0000	2	1.41421
45.00	24.0000	2	2.82843
46.00	25.0000	1	.
47.00	24.0000	1	.
48.00	21.0000	1	.
Total	22.9333	30	3.08426

1b) The differences in means comparisons

Using the compare means function between the Body Image and Self Esteem variables on the Food Consumption SPSS data set. Body Image is placed as the independent variable, while the Self Esteem as the dependent variable for analysis. Among the total N=30 Samples, cases included were 30 samples translated to 1000%, while the cases excluded were zero 0.

From the total 30 samples, the least average was 15.0000, and the highest mean average was 26.6667. Secondly, the lowest body image, 29.00, has a mean value of 22.0000, while the highest Body Image variable, 48.00, has an average mean value of 22.9333. There were nearly the same mean average among the 32.00, 35.00, and 47.00 Body Image variables. The 32.00 and

47.00 Body Image variables have the same mean average values. The standard deviations within the Body Image variables are not relatively close (Liang,2019). Finally, among the total samples, the difference in mean or average of the total 30 samples (n = 30) was just about 2.0000.

2a) (i) Independent samples t-tests between Dieting Status and Body Image variables

Body image is the test variable, while Dieting status is the grouping variable.

Group Statistics

	dietingstatus	N	Mean	Std. Deviation	Std. Error Mean
bodyimage	not dieting	15	41.9333	3.97252	1.02570
	dieting	15	38.0667	4.30061	1.11041

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
bodyimage	Equal variances assumed	.026	.874	2.558	28	.008	.016	3.86667	1.51165	.77020	6.96314
	Equal variances not assumed			2.558	27.826	.008	.016	3.86667	1.51165	.76932	6.96401

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
bodyimage	Cohen's d	4.13982	.934	.171	1.683
	Hedges' correction	4.25499	.909	.166	1.637
	Glass's delta	4.30061	.899	.098	1.674

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

2a) (ii) Independent samples t-tests between Dieting Status and Self-Esteem variables

Self-esteem is the test Variable, while Dieting Status is the grouping Variable.

Group Statistics

	dietingstatus	N	Mean	Std. Deviation	Std. Error Mean
selfesteem	not dieting	15	23.8667	2.92445	.75509
	dieting	15	22.0000	3.04725	.78680

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
selfesteem	Equal variances assumed	.041	.841	1.712	28	.049	.098	1.86667	1.09051	-.36714	4.10047
	Equal variances not assumed			1.712	27.953	.049	.098	1.86667	1.09051	-.36731	4.10064

Independent Samples Effect Sizes

				95% Confidence Interval	
		Standardizer[a]	Point Estimate	Lower	Upper
selfesteem	Cohen's d	2.98648	.625	-.114	1.354
	Hedges' correction	3.06956	.608	-.111	1.317
	Glass's delta	3.04725	.613	-.147	1.353

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

2b)(i) P-values and the relationship between Dieting status and Body image

Correlations

		dietingstatus	bodyimage
dietingstatus	Pearson Correlation	1	-.435*
	Sig. (2-tailed)		.016
	N	30	30
bodyimage	Pearson Correlation	-.435*	1
	Sig. (2-tailed)	.016	
	N	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

The test is significant since the p-value of 0.016 (p=0.016) is less than the 0.05 two-tailed significant level; thus, there is a significant relationship between the Dieting status and the Body Image variables.

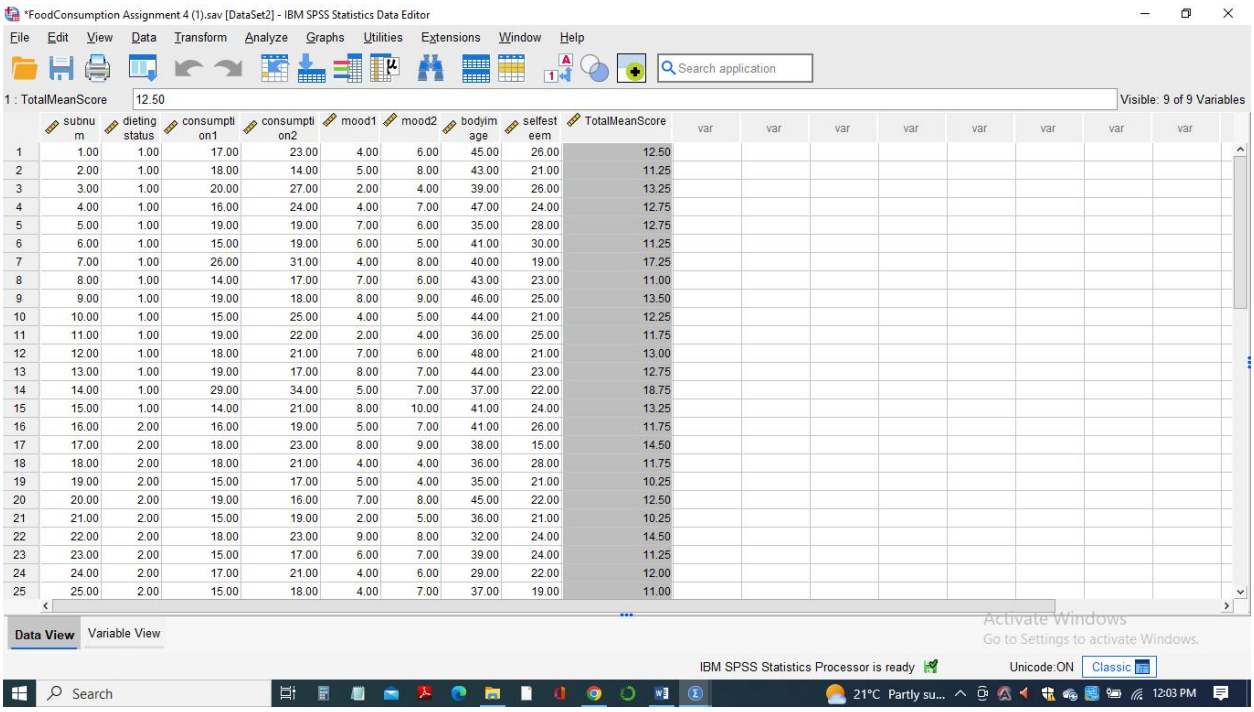
2b)(ii) P-values and the relationship between Dieting Status and Self-esteem

Correlations

		dietingstatus	selfesteem
dietingstatus	Pearson Correlation	1	-.308
	Sig. (2-tailed)		.098
	N	30	30
selfesteem	Pearson Correlation	-.308	1
	Sig. (2-tailed)	.098	
	N	30	30

The p-value of 0.098 (p=0.098) suggests that the test is insignificant. This is because the p-value obtained is greater than the 0.05 two-tailed significant level; thus, there is no significant relationship between the Dieting Status and Self-esteem variables.

3a) Means for Consumption1, Consumption2, Mood1, and Mood2



	subnum	dieting status	consumption1	consumption2	mood1	mood2	bodyimage	selfesteem	TotalMeanScore
1	1.00	1.00	17.00	23.00	4.00	6.00	45.00	26.00	12.50
2	2.00	1.00	18.00	14.00	5.00	8.00	43.00	21.00	11.25
3	3.00	1.00	20.00	27.00	2.00	4.00	39.00	26.00	13.25
4	4.00	1.00	16.00	24.00	4.00	7.00	47.00	24.00	12.75
5	5.00	1.00	19.00	19.00	7.00	6.00	35.00	28.00	12.75
6	6.00	1.00	15.00	19.00	6.00	5.00	41.00	30.00	11.25
7	7.00	1.00	26.00	31.00	4.00	8.00	40.00	19.00	17.25
8	8.00	1.00	14.00	17.00	7.00	6.00	43.00	23.00	11.00
9	9.00	1.00	19.00	18.00	8.00	9.00	46.00	25.00	13.50
10	10.00	1.00	15.00	25.00	4.00	5.00	44.00	21.00	12.25
11	11.00	1.00	19.00	22.00	2.00	4.00	36.00	25.00	11.75
12	12.00	1.00	18.00	21.00	7.00	6.00	48.00	21.00	13.00
13	13.00	1.00	19.00	17.00	8.00	7.00	44.00	23.00	12.75
14	14.00	1.00	29.00	34.00	5.00	7.00	37.00	22.00	18.75
15	15.00	1.00	14.00	21.00	8.00	10.00	41.00	24.00	13.25
16	16.00	2.00	16.00	19.00	5.00	7.00	41.00	26.00	11.75
17	17.00	2.00	18.00	23.00	8.00	9.00	38.00	15.00	14.50
18	18.00	2.00	18.00	21.00	4.00	4.00	36.00	28.00	11.75
19	19.00	2.00	15.00	17.00	5.00	4.00	35.00	21.00	10.25
20	20.00	2.00	19.00	16.00	7.00	8.00	45.00	22.00	12.50
21	21.00	2.00	15.00	19.00	2.00	5.00	36.00	21.00	10.25
22	22.00	2.00	18.00	23.00	9.00	8.00	32.00	24.00	14.50
23	23.00	2.00	15.00	17.00	6.00	7.00	39.00	24.00	11.25
24	24.00	2.00	17.00	21.00	4.00	6.00	29.00	22.00	12.00
25	25.00	2.00	15.00	18.00	4.00	7.00	37.00	19.00	11.00

3b) How scores on Consumption and Mood changed from Time 1 to Time 2

Using a new variable named TotalMeanScore, the means of the Consumption1, Consumption2, Mood1, and Mood2 variables are computed, and the mean score for Consumption and Mood is found to have slightly increased from the first *Time 1* session to the second session *Time 2* in both cases.

3c) Dependent samples t-tests on the Consumption variables and the Mood variables

3c)(i) Dependent samples t-tests on the Consumption1 and Mood1 variables

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	consumption1	16.7667	30	3.94517	.72029
	mood1	5.5000	30	2.04686	.37370

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	consumption1 & mood1	30	-.181	.169	.337

Paired Samples Test

		Paired Differences					t	df	Significance One-Sided p	Significance Two-Sided p
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper				
Pair 1	consumption1 - mood1	11.26667	4.76288	.86958	9.48818	13.04516	12.956	29	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Pair 1	consumption1 - mood1	Cohen's d	4.76288	2.366	1.656	3.063
		Hedges' correction	4.89065	2.304	1.613	2.983

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

3c)(ii) Dependent samples t-tests on the Consumption2 and Mood2 variables

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	consumption2	19.8667	30	5.09045	.92938
	mood2	6.5667	30	1.77499	.32407

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	consumption2 & mood2	30	-.037	.423	.845

Paired Samples Test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	consumption2 - mood2	13.30000	5.45293	.99556	11.26384	15.33616	13.359	29	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	consumption2 - mood2	Cohen's d	5.45293	2.439	1.714	3.153
		Hedges' correction	5.59921	2.375	1.669	3.071

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

3d) P-values on the Consumption and Mood variables and conclusion about the changes in consumption and mood.

3d(i) P-value of Consumption1 and Mood1

Correlations

		consumption1	mood1
consumption1	Pearson Correlation	1	-.181
	Sig. (2-tailed)		.337
	N	30	30
mood1	Pearson Correlation	-.181	1
	Sig. (2-tailed)	.337	
	N	30	30

The p-value of 0.337(p=0.337) suggests that the test is insignificant. This is because the p-value obtained is greater than the 0.05 two-tailed significant level; thus, there is no significant relationship between the Consumption1 and Mood1 variables.

3d(ii P-value of Consumption2 and Mood2

Correlations

		consumption2	mood2
consumption2	Pearson Correlation	1	-.037
	Sig. (2-tailed)		.845
	N	30	30
mood2	Pearson Correlation	-.037	1
	Sig. (2-tailed)	.845	
	N	30	30

Also, in the second case, the p-value of 0.845 (p=0.845) suggests that the test is insignificant. This is because the p-value is greater than the 0.05 two-tailed significant level; thus, there is no significant relationship between the Consumption2 and Mood2 variables.

In conclusion, the p-value between consumption and mood variables increased from 0.337 to 0.845 in the two sessions amongst the 30 participants.

4a) Research hypothesis (unrelated to the Food Consumption dataset)

A research hypothesis is assessed using a dependent samples t-test on a hypothetical data set containing the height and weight of 10 individuals. The heights and weights Variables are used as independent variables, and the entry values as the dependent variables for both cases.

4b) Dependent samples t-test on the hypothesis data in SPSS

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	height	49.9000	10	11.80819	3.73408
	weight	69.9000	10	13.69874	4.33192

Paired Samples Correlations

				Significance	
		N	Correlation	One-Sided p	Two-Sided p
Pair 1	height & weight	10	.700	.012	.024

Paired Samples Test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	height - weight	-20.00000	10.03328	3.17280	-27.17737	-12.82263	-6.304	9	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	height - weight	Cohen's d	10.03328	-1.993	-3.074	-.880
		Hedges' correction	10.97883	-1.822	-2.809	-.804

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

4c) Description of the research hypothesis and support from the hypothetical data

From the hypothetical data containing the height and weight Variables, it is noted that the dependent t-test can look for "differences" between means when the 10 participants are measured on the same dependent variable under different conditions, thereby supporting the research hypothesis. In addition, the one-sided p-value of 0.012 (p=0.012) and the two-sided p-value of 0.024(p=0.024) are less than the 0.05 two-tailed significant level suggesting the test is significant; thus, there is a significant relationship between the height and the weight variables (Cleophas et al.,2010).

Conclusion

In conclusion, the Compare Means procedure might be helpful when comparing variations in descriptive statistics across one or more factors or categorical variables. In addition, two tables will be produced by the Compare Means procedures: the Case Processing Summary, which includes details about the number of valid cases used to generate the statistics, and the Report table, which includes the descriptive statistics themselves (Plonsky,2015). Therefore when comparing multiple numerical variables to one or more categorical variables, Compare

Means is the method of choice even when summarizing numerical variables simultaneously across categories; it is helpful (Orcan,2020).

References

Cleophas, T. J., & Zwinderman, A. H. (2010). Unpaired Continuous Data (Unpaired t-Tests, Mann–Whitney)(20 Patients). In *SPSS for Starters* (pp. 11–13). Springer, Dordrecht.

Liang, G., Fu, W., & Wang, K. (2019). Analysis of t-test misuses and SPSS operations in medical research papers. *Burns & trauma*, *7*.

Or, F. (2020). Parametric or non-parametric: Skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education*, *7*(2), 255-265.

Park, H. M. (2019). Comparing group means t-tests and one-way ANOVA using Stata, SAS, R, and SPSS.

Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A "back-to-basics" approach to advancing quantitative methods in L2 research. *Advancing quantitative methods in second language research*, pp. 23–45.