Included Analyses

- Repeated Measures ANOVA for pre_1, pre_2, pre_3, pre_4, and pre_5
- Friedman Test for pre_1, pre_2, pre_3, pre_4, and pre_5

Results

Repeated Measures ANOVA

Introduction

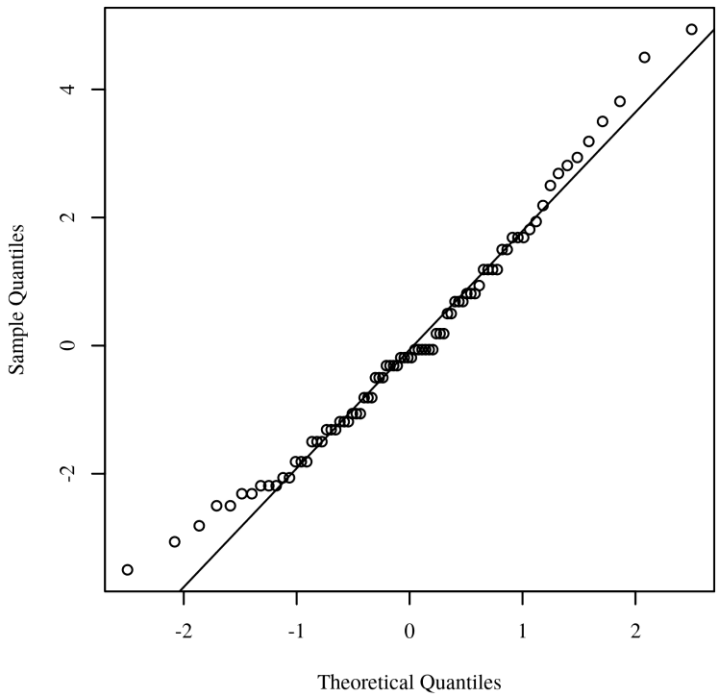
A repeated measures analysis of variance (ANOVA) with one within-subjects factor was conducted to determine whether significant differences exist among pre_1, pre_2, pre_3, pre_4, and pre_5.

Assumptions

Normality. The assumption of normality was assessed by plotting the quantiles of the model residuals against the quantiles of a Chi-square distribution, also called a Q-Q scatterplot (DeCarlo, 1997). For the assumption of normality to be met, the quantiles of the residuals must not strongly deviate from the theoretical quantiles. Strong deviations could indicate that the parameter estimates are unreliable. Figure 1 presents a Q-Q scatterplot of model residuals.

Figure 1

Q-Q scatterplot for normality of the residuals for the regression model.



Sphericity. Mauchly's test was used to assess the assumption of sphericity (Field, 2017; Mauchly, 1940). The results showed that the variances of difference scores between repeated measurements were significantly different from one another based on an alpha of .05, $p < .001$, indicating the sphericity assumption was violated.

Multivariate Outliers. To identify influential points in the residuals, Mahalanobis distances were calculated and compared to a χ^2 distribution (Newton & Rudestam, 2012). An outlier was defined as any Mahalanobis distance that exceeds 20.52, the 0.999 quantile of a χ^2 distribution with 5 degrees of freedom (Kline, 2015). There were no outliers detected in the model.

Results

The results were examined based on an alpha of .05. The *p*-values for the within-subjects factor and the interactions with the within-subjects factor were calculated using the Greenhouse-Geisser correction to adjust for the violation of the sphericity assumption. According to

Greenhouse and Geisser (1959), this is the appropriate way to adjust for violations of the sphericity assumption. The main effect for the within-subjects factor was significant, $F(4, 60) = 5.71$, $p = .007$, indicating there were significant differences between the values of pre_1, pre_2, pre_3, pre_4, and pre_5. Table 1 presents the ANOVA results. The means of the within-subjects factor are presented in Table 2 and Figure 2.

Table 1

Repeated Measures ANOVA Results

Source	*df*	*SS*	*MS*	*F*	*p*	η_p^2
Within-Subjects						
Within Factor	4	21.50	5.37	5.71	.007	0.28
Residuals	60	56.50	0.94			

Table 2

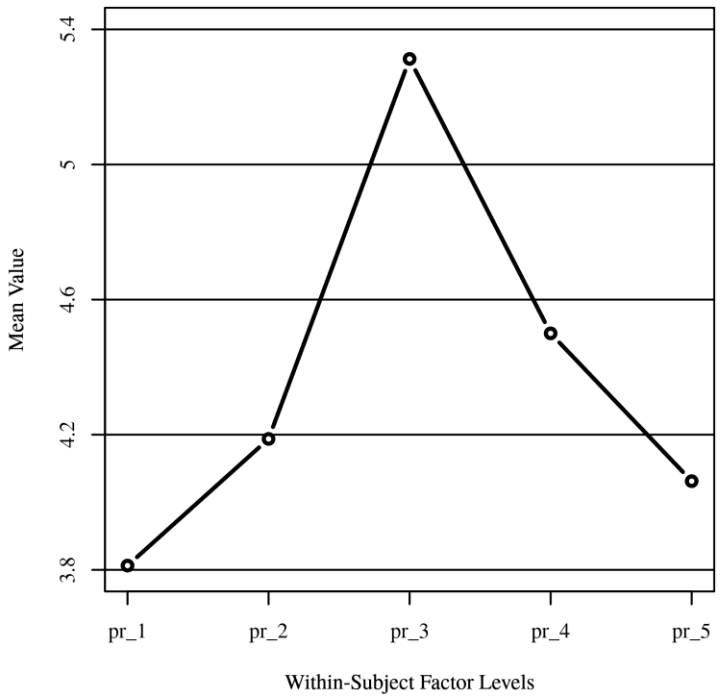
Means Table for Within-Subject Variables

Variable	*M*	*SD*
pre_1	3.81	1.64
pre_2	4.19	1.76
pre_3	5.31	1.49
pre_4	4.50	2.25
pre_5	4.06	1.98

Note. $n = 16$.

Figure 2

Within-subject variable means



Post-hoc. The mean contrasts utilized Tukey comparisons based on an alpha of .05. Tukey comparisons were used to test the differences in the estimated marginal means for each combination of within-subject effects.

Within Effects. pre_1 was significantly less than pre_3, $t(15) = -5.81$, $p < .001$ and pre_2 was significantly less than pre_3, $t(15) = -4.39$, $p = .004$. No other significant differences were found between pre_1, pre_2, pre_3, pre_4, and pre_5. Table 3 presents the marginal means contrasts for the Repeated Measures ANOVA.

Table 3

The Marginal Means Contrasts for each Combination of Within-Subject Variables for the Repeated Measures ANOVA

Contrast	Difference	*SE*	*df*	*t*	*p*
pre_1 - pre_2	-0.38	0.15	15	-2.42	.163
pre_1 - pre_3	-1.50	0.26	15	-5.81	< .001
pre_1 - pre_4	-0.69	0.36	15	-1.90	.359
pre_1 - pre_5	-0.25	0.31	15	-0.81	.924
pre_2 - pre_3	-1.12	0.26	15	-4.39	.004

pre_2 - pre_4	-0.31	0.35	15	-0.89	.895
pre_2 - pre_5	0.12	0.34	15	0.37	.996
pre_3 - pre_4	0.81	0.46	15	1.77	.423
pre_3 - pre_5	1.25	0.50	15	2.52	.138
pre_4 - pre_5	0.44	0.32	15	1.39	.646

Note. Tukey Comparisons were used to test the differences in estimated marginal means.

Friedman Rank Sum Test

Introduction

A Friedman rank sum test was conducted to examine whether the medians of pre_1, pre_2, pre_3, pre_4, and pre_5 were equal. The Friedman test is a non-parametric alternative to the repeated measures one-way ANOVA and does not share the ANOVA's distributional assumptions (Conover & Iman, 1981; Zimmerman & Zumbo, 1993).

Results

The results of the Friedman test were significant based on an alpha value of .05, $\chi^2(4) = 17.42$, $p = .002$, indicating significant differences in the median values of pre_1, pre_2, pre_3, pre_4, and pre_5. Table 4 presents the results of the Friedman rank sum test. Figure 3 presents boxplots of pre_1, pre_2, pre_3, pre_4, and pre_5.

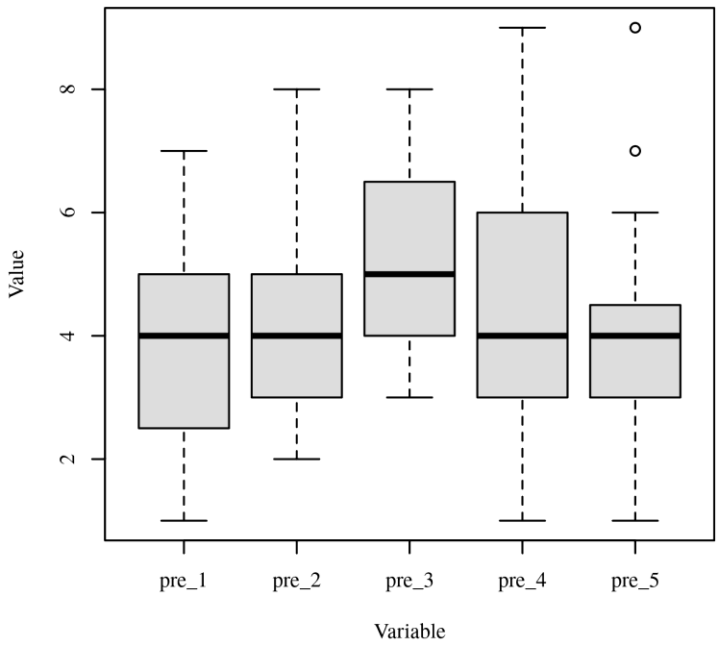
Table 4

Friedman Rank Sum Test

Variable	Mean Rank	χ^2	*df*	*p*
pre_1	2.16	17.42	4	.002
pre_2	2.81			
pre_3	4.25			
pre_4	3.03			
pre_5	2.75			

Figure 3

Boxplots of pre_1, pre_2, pre_3, pre_4, and pre_5



Post-hoc. Pairwise comparisons were examined between each combination of variables. The results of the multiple comparisons indicated significant differences, based on an alpha value of .05, between pre_1-pre_3. Table 5 presents the results of the pairwise comparisons.

Table 5

Pairwise Comparisons for the rank-sums of pre_1, pre_2, pre_3, pre_4, and pre_5

Comparison	Observed Difference	Critical Difference
pre_1-pre_2	10.50	25.11
pre_1-pre_3	33.50	25.11
pre_1-pre_4	14.00	25.11
pre_1-pre_5	9.50	25.11
pre_2-pre_3	23.00	25.11
pre_2-pre_4	3.50	25.11
pre_2-pre_5	1.00	25.11
pre_3-pre_4	19.50	25.11
pre_3-pre_5	24.00	25.11
pre_4-pre_5	4.50	25.11

Note. Observed Differences > Critical Differences indicate significance at the $p < .05$ level.

References

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician, 35*(3), 124-129. https://doi.org/10.1080/00031305.1981.10479327

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods,* 2(3), 292-307. https://doi.org/10.1037/1082-989X.2.3.292

Field, A. (2017). *Discovering statistics using IBM SPSS statistics: North American edition.* Sage Publications

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95-112. https://doi.org/10.1007/BF02289823

Intellectus Statistics [Online computer software]. (2023). Intellectus Statistics. https://statistics.intellectus360.com

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford Publications.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics, 11*(2), 204-209. https://doi.org/10.1214/aoms/1177731915

Newton, R. R., & Rudestam, K. E. (2012). *Your statistical consultant.* Sage Publications.

Zimmerman, D. W., & Zumbo, B. D. (1993). Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *The Journal of Experimental Education, 62*(1), 75-86.

Glossaries

Repeated Measures ANOVA (Analysis of Variance)

The Repeated Measures ANOVA examines differences among repeated measurements on the same subjects.

Fun Fact! *The repeated measures analysis of variance (ANOVA) is commonly mistaken as a multivariate design. This is because in the repeated measures design, each trial represents the measurement of the same characteristic under a different condition.*

Degrees of Freedom (*df*): Refers to the number of values used to compute a statistic; an *F*-test has two values for *df*: the first is determined by the number of groups being compared - 1, and the second is approximately the number of observations in the sample; used with the *F* to determine the *p*-value.

***F* Ratio (*F*):** The ratio of explained variance to error variance; used with the two *df* values to determine the *p*-value.

Normality: Refers to the distribution of the data. The assumption is that the data follows the bell-shaped curve.

Outlier: A data point that is abnormally distant from a set of observations.

Partial Eta Squared (η^2_p): Effect size for the ANOVA and determines the strength of the differences among the groups.

***p*-value:** The probability of obtaining the observed results if the null hypothesis is true.

Residuals: Refers to the difference between the predicted value for the dependent variable and the actual value of the dependent variable.

Sphericity: When there are three or more repeated measurements, the variance of the differences between each pair of measurements must be equal. Sphericity is the term used to describe this measurement.

Type I Error: Rejection of the null hypothesis when the null hypothesis is true; also referred to as a false positive result.

Friedman Test

Friedman test is a non-parametric significance test for more than two dependent samples and is also known as the Friedman two-way analysis of variance; it is used as a null hypothesis test. In other words, it is used to test that there is no significant difference between the size of 'k' dependent samples and the population from which these have been drawn. The Friedman test statistic is distributed approximately as chi-square, with (k - 1) degrees of freedom.

Fun Fact! *The Friedman Test was developed by Milton Friedman, an American economist who was an adviser to President Ronald Reagan and British Prime Minister Margaret Thatcher.*

Chi-Square Test Statistic (χ^2): Refers to the number of values used to compute a statistic. The *df* is determined from the number of groups the nominal variable has; used with χ^2 to compute the *p*-value.

Degrees of Freedom (*df*): Refers to the number of values used to compute a statistic; used with χ^2 to compute the *p*-value.

***p*-value:** The probability of obtaining the observed results if the null hypothesis (no relationship between the independent variable(s) and dependent variable) is true; in most social science research, a result is considered statistically significant if this value is $\leq .05$.

Raw Output

Repeated Measures ANOVA for pre_1, pre_2, pre_3, pre_4, and pre_5

Included Variables:
pre_1, pre_2, pre_3, pre_4, and pre_5

Sample Size (Complete Cases):
N = 16

Sphericity:

Test	W	p
Mauchly's test	0.118	8.201×10^{-04}

Model Results:

Source	df	SS	MS	F	p	η_p^2
Within-Subjects						
Within Factor	4	21.500	5.375	5.708	0.00703	0.276
Residuals	60	56.500	0.942			

Means Table for Within-Subject Variables:

Variable	M	SD
pre_1	3.812	1.642
pre_2	4.188	1.759
pre_3	5.312	1.493
pre_4	4.500	2.251
pre_5	4.062	1.982

Note. n = 16.

The Marginal Means Contrasts for the Repeated Measures ANOVA:

Contrast	Difference	SE	df	t	p
pre_1 - pre_2	-0.375	0.155	15	-2.423	0.163
pre_1 - pre_3	-1.500	0.258	15	-5.809	2.855×10^{-04}
pre_1 - pre_4	-0.688	0.362	15	-1.900	0.359
pre_1 - pre_5	-0.250	0.310	15	-0.808	0.924
pre_2 - pre_3	-1.125	0.256	15	-4.392	0.00408
pre_2 - pre_4	-0.312	0.350	15	-0.892	0.895
pre_2 - pre_5	0.125	0.340	15	0.368	0.996
pre_3 - pre_4	0.812	0.458	15	1.772	0.423
pre_3 - pre_5	1.250	0.496	15	2.521	0.138
pre_4 - pre_5	0.438	0.316	15	1.385	0.646

Note: Tukey Comparisons were used to test the differences in estimated marginal means.

Friedman Rank Sum Test for the Differences in pre_1, pre_2, pre_3, pre_4, and pre_5

Included Variables:
pre_1, pre_2, pre_3, pre_4, and pre_5

Sample Size (Complete Cases):
N = 16

Mean Ranks for Each Variable:

Variable	Mean Rank
pre_1	2.156
pre_2	2.812
pre_3	4.250
pre_4	3.031
pre_5	2.750

Results:
χ^2 = 17.419, df = 4, p = 0.00160

Post-Hoc Tests:

Comparison	Observed Difference	Critical Value
pre_1-pre_2	10.500	25.107
pre_1-pre_3	33.500	25.107
pre_1-pre_4	14.000	25.107
pre_1-pre_5	9.500	25.107
pre_2-pre_3	23.000	25.107
pre_2-pre_4	3.500	25.107
pre_2-pre_5	1.000	25.107
pre_3-pre_4	19.500	25.107
pre_3-pre_5	24.000	25.107
pre_4-pre_5	4.500	25.107

Note: Observed Differences > Critical Differences indicate significance at the p < 0.0500 level.