Longitudinal Analysis of Marital Therapy's Impact on Intimacy in Heterosexual Couples

Yaa Amankwah Awuah and Mostafa Zahed, Department of Mathematics and Statistics, East Tennessee State University, TN, USA

ABSTRACT

This research study examines the effects of marital treatment on the levels of closeness among wives of 50 heterosexual married couples utilizing longitudinal data from a simulated study. Participants were randomly assigned to a wait-list control condition or a 16-week marriage therapy treatment. Data were gathered at various times in time, resulting in a comprehensive longitudinal dataset that made it possible to examine individual and group differences in the development of intimacy through time.
In-depth descriptive statistics and visualization techniques were used in the research to carefully examine the data's overall distribution as well as any trends, patterns, or probable outliers. Mixed-effect models were used for additional statistical testing, which included sensitivity analysis, interaction terms, and random effects for intercepts and slopes. This enabled a thorough evaluation of variations over time within and across groups.

The Akaike Information Criterion (AIC) was used to compare the Random Intercept (RI) model to the Random Time Slope (RT) model. The RT model showed a better match and provided a meaningful representation of the data. Additionally, residual diagnostics were employed to evaluate the general model fit and to validate the model's premise.
The findings supported the efficacy of marital treatment by showing a significant rise in closeness levels over the course of the 16-week period. The RT model demonstrated an anticipated rise in intimacy each week, indicating that therapy sessions had a long-term beneficial effect. The findings also showed significant individual variation, which may mean that different types of partnerships had different therapeutic responses.

In summary, this study supports the use of a 16-week marital therapy intervention to improve intimacy between married couples. It also emphasizes how crucial it is to consider individual differences and particular relationship dynamics when analyzing the results of such interventions. Future studies might concentrate on finding variables that indicate therapy success, allowing for a more specialized method of marital therapy.

INTRODUCTION

Longitudinal data analysis is an indispensable approach in the field of statistics, offering unparalleled insights into the temporal dynamics of various outcomes across a myriad of disciplines (TP & K, 2017). This methodology captures changes within individual units over time, unveiling complex patterns of temporal evolution in ways that cross-sectional studies cannot match (Tan, Shiyko, Li, & Dierker, 2012). It plays a crucial role in disciplines such as psychology, sociology, epidemiology, and economics, where understanding changes, predicting future trends based on past data, and deciphering individual differences in trajectories are of prime importance.

This study gives an in-depth examination of a longitudinal dataset that includes measurements of wives' levels of closeness over a 16-week period from 50 heterosexual married couples (Niall Bolger, 2013). These couples took part in an intervention trial and were paired up randomly to either receive marital counseling treatment or remain on the waiting list. Understanding the progression of intimate change over time throughout therapy and examining how this change varies between people and treatment groups are the main goals of this analysis(Niall Bolger, 2013).

To extract valuable insights from the longitudinal data, our analytical strategy incorporates several complex statistical techniques (SL & KY, 1992, M & CY, 1999), and Locascio & Atri, 2011). To acquire a preliminary grasp of the variable distributions and spot any early patterns or trends, we first conduct an exploratory analysis using descriptive statistics and visualizations.

As shown by a thorough survey of the literature, in our analysis of longitudinal data, we critically assessed the traditional methods of repeated measures ANOVA and MANOVA. Due to their capacity to take into consideration within-subject correlation, these conventional techniques are frequently employed when working with repeated measurements from the same subjects across time (Rowell and Walters 1976, Mead et al. 1993, Hand and Crowder 1996, Senn et al. 2000), (Piepho & Richter, 2004).

However, that these methods have certain inherent shortcomings, most notably their inability to handle missing data and the variation in change rates among individuals. This study's overarching goals were to comprehend the trajectories of intimacy change through time and to offer light on the efficacy of marital treatment. We seek to guide the design of future interventions and offer insights that might improve relationship satisfaction among couples by considering individual differences within groups. The

importance of longitudinal data analysis in uncovering intricate temporal patterns and individual differences in psychological and social consequences is emphasized in this research, underscoring its potential to revolutionize our understanding of how people behave and interact. (R., D., J., & (2020), 2020), (Han, et al., 2023)

This work instead uses random intercept and random slope models considering these drawbacks. Given that these models support both random variables (intercept, slope) and fixed effects (time, treatment), our decision to use them was motivated by our desire to present a more accurate and nuanced portrayal of the data. These models can therefore effectively handle the complexity and individual variability seen in longitudinal data.

To further corroborate our findings, we employed robustness tests using techniques such as sensitivity analysis. This guaranteed the accuracy of our findings and enabled us to explore any significant group differences through post hoc (Nosek, et al., 2022). This guaranteed the accuracy of our findings and enabled us to explore any significant group differences through post hoc analysis. To uncover intricate linkages in our data and improve our comprehension overall, we also used data mining tools.

The thorough analysis of residuals, a step essential to confirming the underlying premises of our selected models and assessing how well they fit our data, was a vital component of our study strategy. Our study has shown a viable alternative to traditional ANOVA and MANOVA procedures by implementing these methodologies, which include random intercept and random slope models, advanced analytics, and data mining. These methodologies have produced relevant information on longitudinal data.

The data showed that marital therapy significantly increased wife closeness, with apparent individual differences in therapy response. The research underlined the need for customized interventions that take into consideration the dynamics of each person and each relationship (Niall Bolger, 2013). This will improve therapy outcomes and, in turn, increase couple relationships' level of satisfaction.

DESCRIPTION OF DATA

The longitudinal dataset analyzed in this report consists of data collected from individuals at multiple time points. The dataset includes the variables id, time, time01, intimacy, and treatment.
The dataset contains simulated data for an intervention study that measures changes in intimacy levels over time. The study involved 50 heterosexual married couples who were randomly assigned to a 16-week marital therapy treatment condition or a 16-week wait-list condition. The goal of the study was to investigate both group differences and individual differences within each group.
The dataset includes five variables:
1. Time01" variable represents a continuous measure of time, with 0 indicating the first time point and 1 indicating the final time point.
2. ID, which is an integer variable that ranges from 1 to 50 and is used to identify each participant.
3. Time, an integer variable that ranges from 0 to 15 and represents the 16-week intervention period when the intimacy level was measured.
4. Intimacy, a continuous variable that ranges from 0 to 10 and measures the level of intimacy reported by participants at each time point. It is expected to differ between the treatment and control groups; and
5. Treatment, a binary variable that takes the value 0 for the control group and 1 for the intervention group, indicating the participant's randomly assigned group.

Response variable: intimacy (1 to 10)
Predictor variables: time, time01, id (1 to 50), treatment (0 or 1). All variables are imputed in SAS as shown by the table in Fig 1

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
1	id	Num	8	BEST12.	BEST32.
4	intimacy	Num	8	BEST12.	BEST32.
2	time	Num	8	BEST12.	BEST32.
3	time01	Num	8	BEST12.	BEST32.
5	treatment	Num	8	BEST12.	BEST32.

Figure 1 - Variables used.

Data was collected using web diaries completed by participants. The web diaries were completed weekly by the treatment group on the evening before therapy sessions and by the control group on a fixed day each week.
The dataset is longitudinal because it measures the same individuals at multiple time points, and it is hierarchical because it includes multiple observations for each participant.

PROPOSED ANALYSIS

The proposed analysis for a longitudinal dataset with variables id, time, time01, Intimacy, and treatment will involve using several visualization techniques, including time plot, spaghetti plot, interaction plot, residual plot, and variogram.

An analysis of the longitudinal data could involve the following steps:

- Descriptive statistics like means, standard deviations, and frequencies should be calculated for all four variables: id, time, intimacy, and treatment. These statistics will provide an overview of the distribution of the data and help identify any missing data, or other issues that need to be addressed before proceeding to more complex analyses.
- Visual inspection: The data should be visually inspected to look for patterns or trends in the time course of the outcome variable (intimacy) across the two groups (treatment vs. control). This can be done using line plots or scatterplots of intimacy over time for each group.
- Group differences: The primary research question in this study is whether there are group differences in the time course of intimacy over the 16-week period. To test this hypothesis, a mixed-effects model should be used with time and treatment as predictors and intimacy as the outcome variable. The model should include random effects for both intercepts and slopes to account for subject-level differences within each group. The model should also include an interaction term between time and treatment to test for differential treatment effects over time. The model should be fit using maximum likelihood estimation and the significance of the fixed effects should be tested using likelihood ratio tests or Wald tests.
- Sensitivity analysis: To ensure the robustness of the results, sensitivity analysis can be performed by testing different model specifications, such as using different covariance structures for the random effects or including additional covariates to control for potential confounding factors.
- Post hoc analysis: If significant group differences are found, post hoc analyses can be conducted to explore the nature of these differences. For example, subgroup analyses can be performed based on demographic or clinical characteristics to determine whether the treatment effect is consistent across different subgroups. Alternatively, mediation analysis can be performed to investigate the mechanisms underlying the treatment effect.

Overall, the proposed analysis of the longitudinal data involves a combination of descriptive, visual, and statistical techniques to examine group differences in the time course of the outcome variable and identify potential factors that may influence these differences.

ANALYSIS OF RESULTS

A conditional model is appropriate for longitudinal data analysis when the focus is on modeling the within-individual change over time, while accounting for the correlation among repeated measurements from the same individual.

For conditional models, we assume that the relationship between the response variable and the predictor variables may differ between individuals and estimate individual-specific parameters that capture the within-subject variation. In our case, we are using the conditional longitudinal model and assume that in the treatment group, each wife completed a web diary each week prior to the day of the therapy session and in the control group, each wife completed a web diary on a fixed day of the week.

In assessing the patterns and trends in the dataset, a time plot (Figure-2) was run on the dataset set. We can see that most of the data were concentrated in the middle, with a few outliers. Also, from [Figure- 3], the Spaghetti plot which is another method of visualizing the possible flows in the data, we can see that although it looks compact, the variations go up over time.
An interaction plot was done, plotting outcome against time which was grouped by all the variables of interest. From Figure 4., we can see that there is an interaction between the treatment 0 and 1.
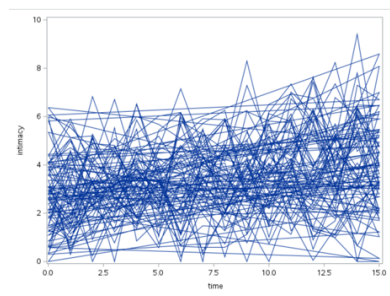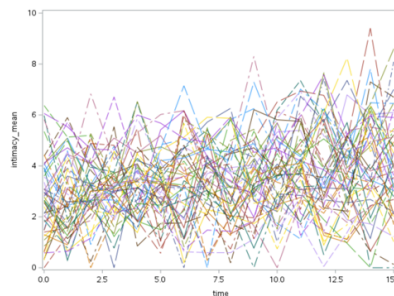

| Figure 2-Time Plot. | Figure 3- Spaghetti Plot. | Figure 4- Interaction Plot |
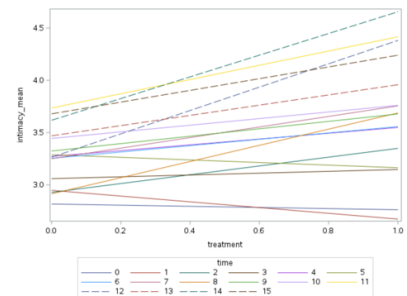
A variogram is a tool to describe the data spatial continuity. When several data samples are available, the analyst can use a variability measure between each pair of the samples at different distances to calculate the experimental variogram function considering the best type of the variogram fitted (CV & AG, 1998).

Number of Observations Read	800
Number of Observations Used	800

Spatial Distribution of residuals Observations

The VARIOGRAM Procedure
Dependent Variable: residuals

Empirical Semivariogram			
Lag Class	Pair Count	Average Distance	Semivariance
0	22832	3.09	2.359
1	122400	9.95	2.661
2	86900	19.44	2.748
3	55476	29.37	2.560
4	27858	39.03	2.545
5	4134	46.70	2.350
6	0	.	.
7	0	.	.
8	0	.	.
9	0	.	.
10	0	.	.

The VARIOGRAM Procedure
Dependent Variable: residuals

Empirical Semivariogram for residuals

Figure 5 - Variogram

The variogram in (Figure. 5) shows that the horizontal line represents the maximum possible level of spatial autocorrelation between two locations in the dataset being analyzed. It indicates a theoretical line for the determination of strengths or weakness of association. The graph indicates a weak association at the beginning and then moves closer to the horizontal line, which shows strong association. The means and variances of the intimacy of women increases and decreases overtime. This indicates that the pattern is not continuously decreasing or increasing. The means of the intimacy of women increases and decreases overtime. The pattern is not continuously decreasing or increasing.
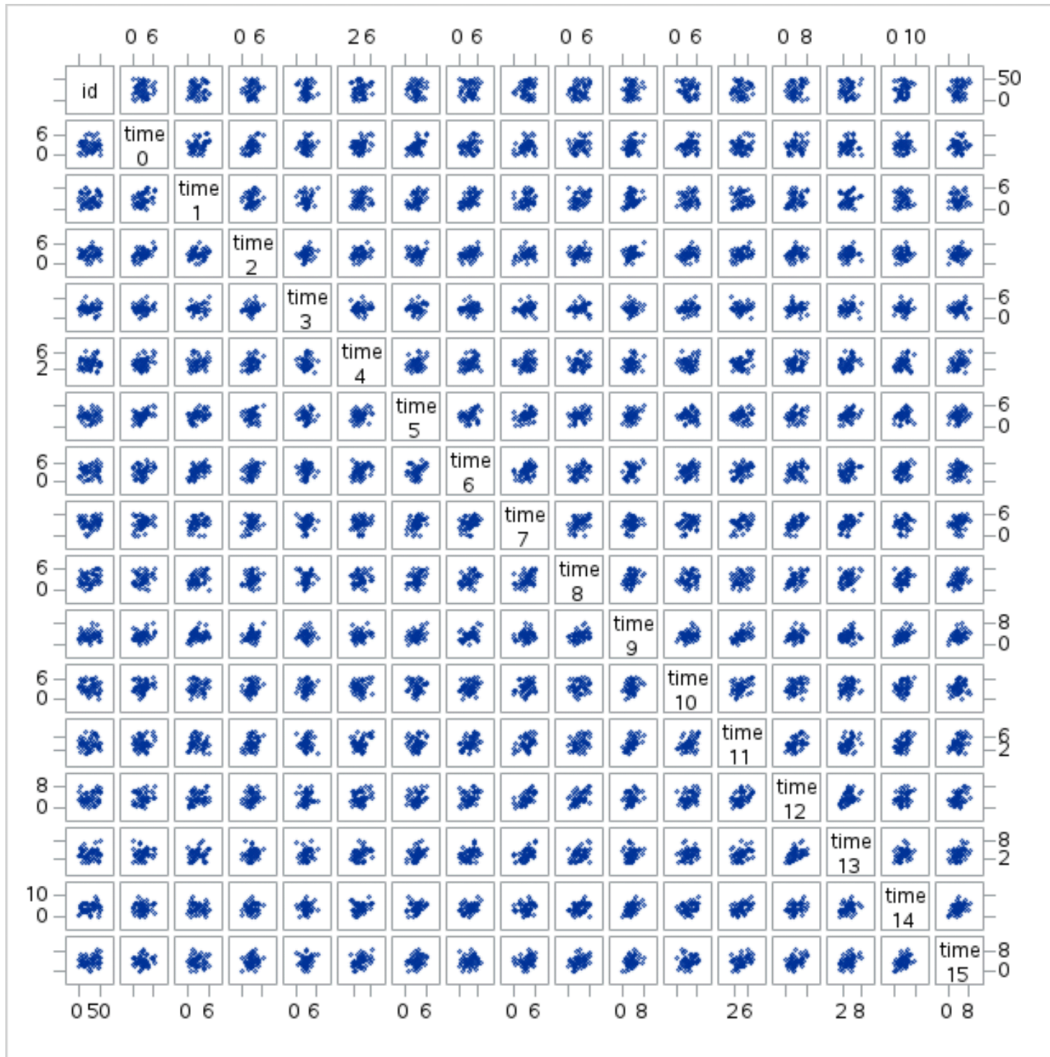
Figure 6 – Scatter Plot

From the scatter plot matrix, Figure. 6, there is randomness, and the association gets weaker over time. The association, however, is fairly moderate since the highest association or relationship between id and time is 0.61. The residual correlation matrix in Figure. 7 also aligns with this result as most of the residuals are within the same range as the scatter plot matrix.

17 Variables:	id time0 time1 time2 time3 time4 time5 time6 time7 time8 time9 time10 time11 time12 time13 time14 time15

| Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
id	50	0	1.00000	0	-1.68069	1.68069
time0	50	0	1.00000	0	-1.76961	2.27155
time1	50	0	1.00000	0	-1.61541	1.98051
time2	50	0	1.00000	0	-2.35448	2.76885
time3	50	0	1.00000	0	-2.56530	2.98160
time4	50	0	1.00000	0	-1.90712	2.32331
time5	50	0	1.00000	0	-2.13875	2.24955
time6	50	0	1.00000	0	-1.94692	2.14061
time7	50	0	1.00000	0	-2.45427	1.68639
time8	50	0	1.00000	0	-2.01737	1.91881
time9	50	0	1.00000	0	-2.13685	2.92938
time10	50	0	1.00000	0	-2.32741	1.87875
time11	50	0	1.00000	0	-1.86757	2.07577
time12	50	0	1.00000	0	-1.90022	1.92238
time13	50	0	1.00000	0	-1.69156	2.78094
time14	50	0	1.00000	0	-1.99156	2.53887
time15	50	0	1.00000	0	-2.09285	2.44740

Pearson Correlation Coefficients, N = 50
Prob > |r| under H0: Rho=0

	id	time0	time1	time2	time3	time4	time5	time6	time7	time8	time9	time10	time11	time12	time13	time14	time15
id	1.00000	0.02939 0.8394	0.05015 0.7294	0.05353 0.7120	0.03426 0.8133	0.13213 0.3604	-0.03394 0.8150	0.09248 0.5230	0.05408 0.7091	0.27292 0.0552	0.13694 0.3430	-0.13351 0.3553	0.07818 0.5894	0.22608 0.1144	0.06843 0.6368	0.25491 0.0740	0.15663 0.2774
time0	0.02939 0.8394	1.00000	0.36462 0.0092	0.42114 0.0023	0.12240 0.3971	0.31020 0.0284	0.51424 0.0001	0.22140 0.1223	0.29756 0.0358	0.21830 0.1278	0.28457 0.0452	0.18201 0.2059	0.17658 0.2199	0.19397 0.1771	0.08359 0.5639	0.00699 0.9616	0.09790 0.4988
time1	0.05015 0.7294	0.36462 0.0092	1.00000	0.16841 0.2424	0.05803 0.6890	0.17483 0.2246	0.41653 0.0026	0.53557 <.0001	0.19495 0.1749	0.35712 0.0109	0.25084 0.0789	0.13209 0.3605	-0.06021 0.6779	0.23789 0.0962	0.17428 0.2261	-0.08056 0.5781	-0.00956 0.9475
time2	0.05353 0.7120	0.42114 0.0023	0.16841 0.2424	1.00000	0.22664 0.1135	0.30784 0.0296	0.08930 0.5374	0.39918 0.0041	0.34820 0.0132	0.11864 0.4119	0.29733 0.0360	0.25446 0.0745	0.35616 0.0111	0.22246 0.1205	0.18325 0.2027	0.12068 0.4038	0.16476 0.2529
time3	0.03426 0.8133	0.12240 0.3971	0.05803 0.6890	0.22664 0.1135	1.00000	0.03570 0.8056	0.06178 0.6700	0.11092 0.4432	0.15343 0.2874	0.07907 0.5852	0.09075 0.5308	0.14533 0.3139	0.00158 0.9913	0.10845 0.4534	0.14353 0.3200	0.06678 0.6450	0.18419 0.2004
time4	0.13213 0.3604	0.31020 0.0284	0.17483 0.2246	0.30784 0.0296	0.03570 0.8056	1.00000	0.24078 0.0921	0.29246 0.0393	0.16357 0.2564	0.15908 0.2698	0.15267 0.2899	0.09973 0.4908	0.29798 0.0356	0.24918 0.0810	0.33300 0.0181	0.27587 0.0525	0.37737 0.0069
time5	-0.03394 0.8150	0.51424 0.0001	0.41653 0.0026	0.08930 0.5374	0.06178 0.6700	0.24078 0.0921	1.00000	0.25944 0.0688	0.35607 0.0112	0.31970 0.0236	0.35303 0.0119	-0.02769 0.8486	0.14382 0.3190	0.22899 0.1097	0.34731 0.0135	0.14797 0.3051	0.20081 0.1620
time6	0.09248 0.5230	0.22140 0.1223	0.53557 <.0001	0.39918 0.0041	0.11092 0.4432	0.29246 0.0393	0.25944 0.0688	1.00000	0.34442 0.0143	0.40413 0.0036	0.38431 0.0059	0.32252 0.0224	0.37124 0.0079	0.31202 0.0274	0.37466 0.0073	0.36483 0.0092	0.09982 0.4904
time7	0.05408 0.7091	0.29756 0.0358	0.19495 0.1749	0.34820 0.0132	0.15343 0.2874	0.16357 0.2564	0.35607 0.0112	0.34442 0.0143	1.00000	0.41341 0.0028	0.15140 0.2939	0.23625 0.0986	0.37669 0.0070	0.62576 <.0001	0.53034 <.0001	0.18373 0.2015	0.22625 0.1141
time8	0.27292 0.0552	0.21830 0.1278	0.35712 0.0109	0.11864 0.4119	0.07907 0.5852	0.15908 0.2698	0.31970 0.0236	0.40413 0.0036	0.41341 0.0028	1.00000	0.34481 0.0142	0.12540 0.3856	0.28297 0.0465	0.52864 <.0001	0.30917 0.0289	0.31062 0.0281	0.41561 0.0027
time9	0.13694 0.3430	0.28457 0.0452	0.25084 0.0789	0.29733 0.0360	0.09075 0.5308	0.15267 0.2899	0.35303 0.0119	0.38431 0.0059	0.15140 0.2939	0.34481 0.0142	1.00000	0.24271 0.0895	0.51997 0.0001	0.24680 0.0840	0.18615 0.1956	0.49629 0.0002	0.31205 0.0274
time10	-0.13351 0.3553	0.18201 0.2059	0.13209 0.3605	0.25446 0.0745	0.14533 0.3139	0.09973 0.4908	-0.02769 0.8486	0.32252 0.0224	0.23625 0.0986	0.12540 0.3856	0.24271 0.0895	1.00000	0.41323 0.0029	0.26833 0.0595	0.16578 0.2499	0.22926 0.1093	0.14582 0.3123
time11	0.07818 0.5894	0.17658 0.2199	-0.06021 0.6779	0.35616 0.0111	0.00158 0.9913	0.29798 0.0356	0.14382 0.3190	0.37124 0.0079	0.37669 0.0070	0.28297 0.0465	0.51997 0.0001	0.41323 0.0029	1.00000	0.42938 0.0019	0.32370 0.0218	0.47022 0.0006	0.36102 0.0100
time12	0.22608 0.1144	0.19397 0.1771	0.23789 0.0962	0.22246 0.1205	0.10845 0.4534	0.24918 0.0810	0.22899 0.1097	0.31202 0.0274	0.62576 <.0001	0.52864 <.0001	0.24680 0.0840	0.26833 0.0595	0.42938 0.0019	1.00000	0.46112 0.0008	0.32958 0.0194	0.40732 0.0033
time13	0.06843 0.6368	0.08359 0.5639	0.17428 0.2261	0.18325 0.2027	0.14353 0.3200	0.33300 0.0181	0.34731 0.0135	0.37466 0.0073	0.53034 <.0001	0.30917 0.0289	0.18615 0.1956	0.16578 0.2499	0.32370 0.0218	0.46112 0.0008	1.00000	0.32465 0.0214	0.39474 0.0046
time14	0.25491 0.0740	0.00699 0.9616	-0.08056 0.5781	0.12068 0.4038	0.06678 0.6450	0.27587 0.0525	0.14797 0.3051	0.36483 0.0092	0.18373 0.2015	0.31062 0.0281	0.49629 0.0002	0.22926 0.1093	0.47022 0.0006	0.32958 0.0194	0.32465 0.0214	1.00000	0.43492 0.0016
time15	0.15663 0.2774	0.09790 0.4988	-0.00956 0.9475	0.16476 0.2529	0.18419 0.2004	0.37737 0.0069	0.20081 0.1620	0.09982 0.4904	0.22625 0.1141	0.41561 0.0027	0.31205 0.0274	0.14582 0.3123	0.36102 0.0100	0.40732 0.0033	0.39474 0.0046	0.43492 0.0016	1.00000

Figure 7 – Correlation Matrix

Two models were run, the Random Intercept model (the reduced model) and the Random Time Slope model (full model). A summary of the 2 models is found below:
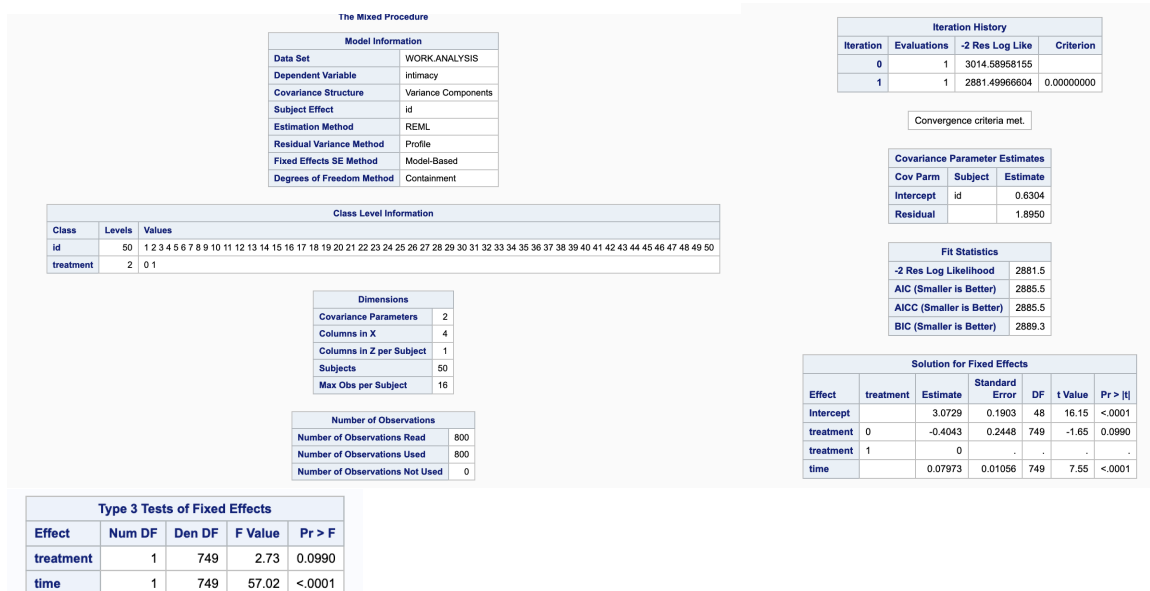
The Mixed Procedure

Model Information	
Data Set	WORK.ANALYSIS
Dependent Variable	intimacy
Covariance Structure	Variance Components
Subject Effect	id
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
id	50	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
treatment	2	0 1

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z per Subject	1
Subjects	50
Max Obs per Subject	16

Number of Observations	
Number of Observations Read	800
Number of Observations Used	800
Number of Observations Not Used	0

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	3014.58958155	
1	1	2881.49966604	0.00000000

Convergence criteria met.

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
Intercept	id	0.6304
Residual		1.8950

Fit Statistics	
-2 Res Log Likelihood	2881.5
AIC (Smaller is Better)	2885.5
AICC (Smaller is Better)	2885.5
BIC (Smaller is Better)	2889.3

Solution for Fixed Effects						
Effect	treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		3.0729	0.1903	48	16.15	<.0001
treatment	0	-0.4043	0.2448	749	-1.65	0.0990
treatment	1	0	.	.	.	.
time		0.07973	0.01056	749	7.55	<.0001

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
treatment	1	749	2.73	0.0990
time	1	749	57.02	<.0001

Figure 8 – Random Intercept Model

The Random Intercept (RI) model is:

$$\hat{Y}_{it} = 3.0729 - 0.4043\ treatment_{it} + 0.07973\ time_{it} + \mu_i + \varepsilon_{it}$$

Terms:

1. $\hat{Y}_{it}$: This is the dependent variable, representing the predicted outcome. In the context of the study, this would likely represent the wives' level of intimacy.

2. $treatment_{it}$: The treatment condition (marital counseling or control/waiting list) is indicated by this independent variable. It is a binary variable, usually recorded as 0 for "control" or "no treatment received" and 1 for "treatment received." The coefficient -0.4043 indicates that, in comparison to the control group, receiving the therapy (marital counseling) is linked to a decline in the anticipated result (degree of closeness).

3. $time_{it}$: This is also an independent variable, denoting the time point at which the measurement was taken (from 1 to 16 weeks). The positive coefficient of 0.07973 indicates that as time increases, the predicted intimacy level increases.

4. 3.0729: This is the regression equation's constant or intercept. When all the independent variables (treatment and time) are equal to zero, it symbolizes the anticipated result (degree of intimacy). Practically speaking, this would be the anticipated level of intimacy for study participants in the control group at the outset.

5. μ_i: This is the random subject effect with variation σ_μ^2

6. ε_{it}: This is the error with the variation σ_ε^2

The Mixed Procedure

Model Information	
Data Set	WORK.ANALYSIS
Dependent Variable	intimacy
Covariance Structure	Unstructured
Subject Effect	id
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	3014.58958155	
1	2	2843.68208308	0.00000002
2	1	2843.68207214	0.00000000

Convergence criteria met.

Class Level Information		
Class	Levels	Values
id	50	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
treatment	2	0 1

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
UN(1,1)	id	0.7055
UN(2,1)	id	-0.03810
UN(2,2)	id	0.009105
Residual		1.6925

Dimensions	
Covariance Parameters	4
Columns in X	4
Columns in Z per Subject	2
Subjects	50
Max Obs per Subject	16

Fit Statistics	
-2 Res Log Likelihood	2843.7
AIC (Smaller is Better)	2851.7
AICC (Smaller is Better)	2851.7
BIC (Smaller is Better)	2859.3

Number of Observations	
Number of Observations Read	800
Number of Observations Used	800
Number of Observations Not Used	0

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
3	170.91	<.0001

Solution for Fixed Effects						
Effect	treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		3.0070	0.1886	48	15.94	<.0001
treatment	0	-0.2726	0.2345	700	-1.16	0.2454
treatment	1	0	.	.	.	.
time		0.07973	0.01678	49	4.75	<.0001

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
treatment	1	700	1.35	0.2454
time	1	49	22.57	<.0001

Figure 9 – Random Time Model

The Random Time (RT) model is:

$$\hat{Y}_{it} = 3.0070 - 0.2726\ treatment_{it} + +0.07973\ time_{it} + \mu_{0i} + \mu_{1i} * t_{it} + \varepsilon_{it}$$

1. $\hat{Y}_{it}$: This is the dependent variable representing the predicted outcome. In the context of the study, this likely represents the predicted wives' level of intimacy at a particular time for a particular individual.
2. i: This is an index for the individual or subject in the study. In this case, 'i' would denote a specific wife among the 50 wives participating in the study. It allows us to account for the individual differences among wives, which is essential in a longitudinal study.
3. t: This is an index for time, denoting the specific time point at which the measurement was taken (across the 16 weeks). It allows the model to account for changes over time.
4. μ_{0i}: This is the fluctuation in the baseline.
5. $\mu_{1i:}$ This is the fluctuation in the time effect.

The coefficients and variables have the same interpretation as mentioned in the previous response:

$treatment_{it}$: This is an independent variable, denoting the treatment condition (marital counseling or control/waiting list). The coefficient -0.2726 suggests that receiving the treatment (marital counseling) is associated with a decrease in the predicted outcome (level of intimacy) as compared to the control group.

1   time: This is also an independent variable, denoting the time point at which the measurement was taken. The positive coefficient of 0.07973 indicates that as time increases, the predicted intimacy level also increases.

The criterion for selecting the best model is the Akaike Information Criterion (AIC). The Random Time Slope model was selected since it had the smallest AIC (**2851.7**) (Figure 9) compared to the Random Intercept model (**2885.5**). (Figure 8 and Figure 9). Based on AIC, the selected model is the **Random Slope Model.**

Also, a likelihood ratio test was done to determine the most appropriate covariance structure and to check whether the larger model (RT model) explains significantly greater information about the intimacy for wives from 50 heterosexual married couples who participated in a 16-week marital therapy treatment or a wait-list condition or not since we chose it as our best model.
The null and alternative hypotheses were defined as follows:

$$H_0: \text{A larger model does not explain significantly greater information.}$$
$$H_1: \text{A larger model explains significantly greater information.}$$

From the data analysis of the likelihood ratio test (Fig. 10), the RT model has a p-value that is less than 0.001. This indicates that the RT model explains significantly greater information.

For further analysis, we move on to find the residual diagnostics to help determine the overall model fit. Although the RT model was chosen as the better model, we ran a residual diagnostic on it to ensure that it satisfies the model assumptions.
Initially, the adequacy of the standard random intercept model should be assessed. From the results, the residuals have a range of (3.41 and 3.23), indicating that observed values are larger than those predicted by the model. Furthermore, the residual plot displays large values and distinct "groups" of observations that indicate the model may not be capturing certain factors. The Shapiro-Wilk normality test is highly significant, with a p-value of 0.3, and the normal probability plot follows a straight line. These results suggest that the RT Model does not fit very well.
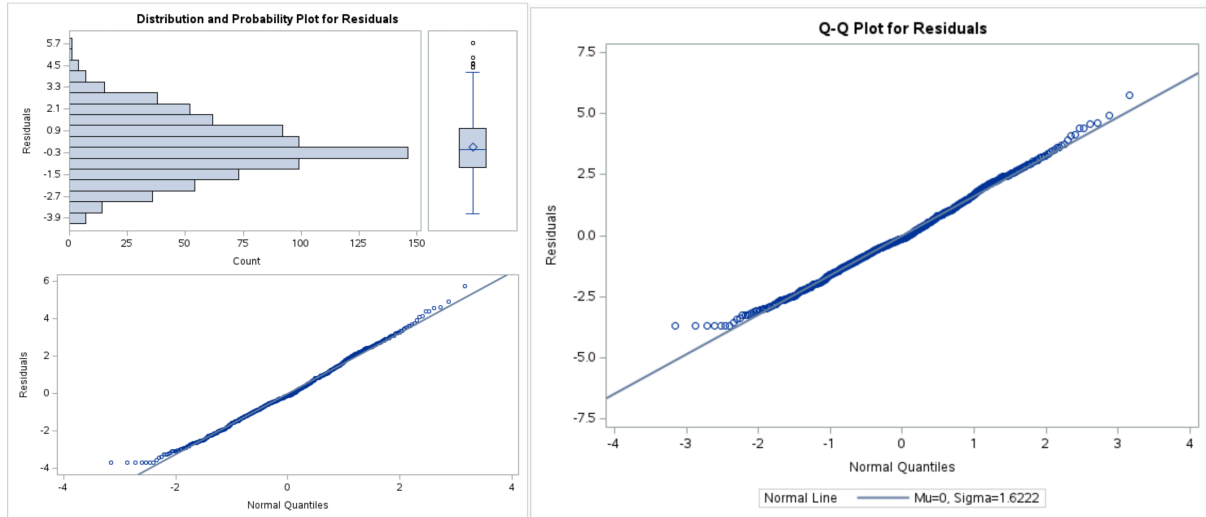
Figure 10 – Residual Plots

Also, considering the fit of the standard random intercept model. The residuals have a range of (-3.95, 4.41), suggesting observed values much larger than predicted using the model. The plot of residuals versus predicted values shows large values, and clear "groups" of observations that suggest something is not being captured by the model. The Shapiro-Wilk test of normality is highly significant, with a p-value of 0.7 and the normal probability plot shows a fairly normal distribution.
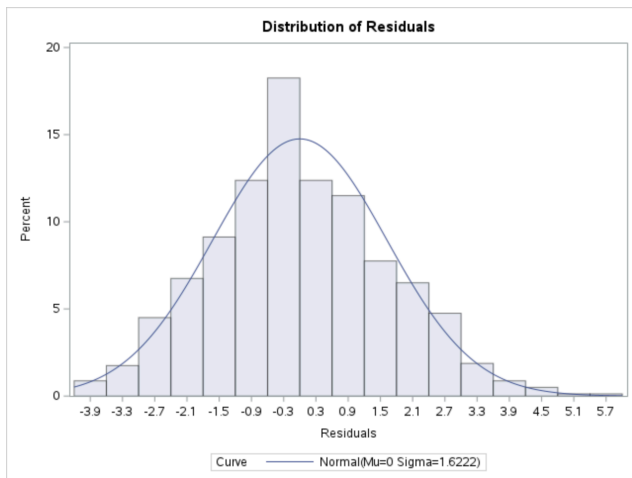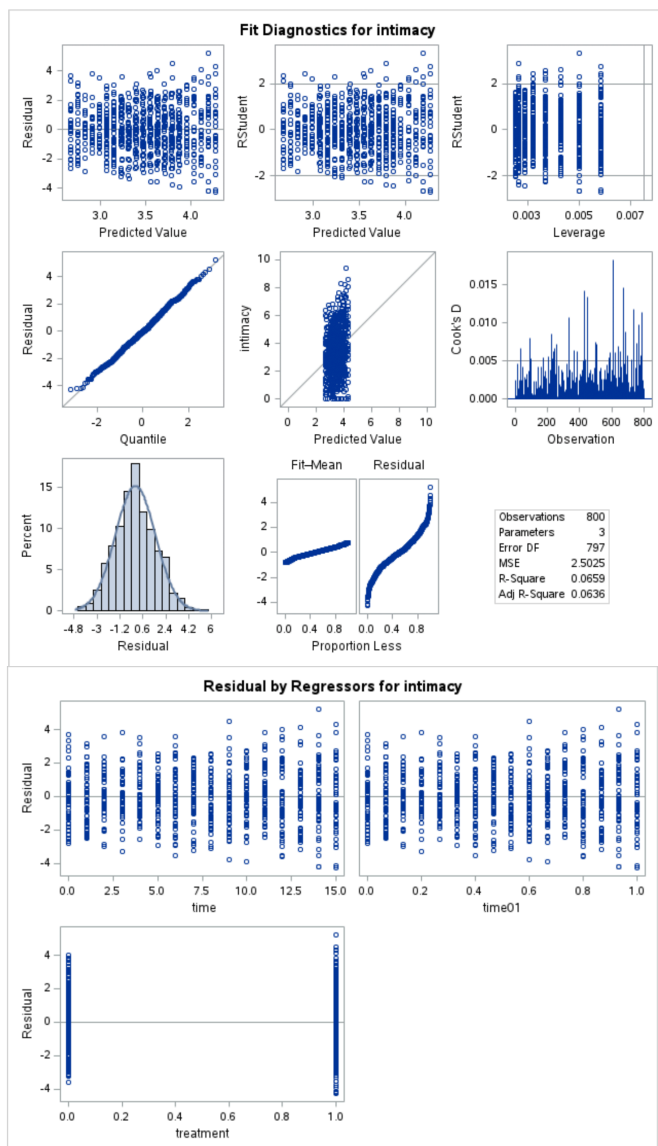
Figure 11 Residuals for Intimacy

A semi-variogram plot is used to examine the relationship between the spatial lag distance (i.e., the distance between pairs of points) and the semi-variance (i.e., the average of the squared differences between pairs of points). Fig 14 shows a high association for the RT model, but Fig 10 shows that the association for the RI model moves from higher to weak over time.

Based on the above results, the RT Model is the best model for explaining the intimacy for wives from the 50 heterosexual married couples who participated in the 16-week marital therapy treatment.

CONCLUSION

In conclusion, this longitudinal dataset provides simulated data on the time course of an outcome, representing a study of wives from 50 heterosexual married couples randomly assigned to a 16-week marital therapy treatment condition or a wait-list condition. The dataset includes four variables: id, time, intimacy, and treatment. The intervention is expected to produce group differences in the time course and subject-level differences within each group. The data may be used to investigate the effects of marital therapy on intimacy in married couples over time. The dataset is a valuable resource for researchers interested in studying the effectiveness of marital therapy and the dynamics of intimate relationships within couples.

10

REFERENCES

Adam, H. (2021). Retrieved from https://www.investopedia.com/terms/l/longitudinaldata.asp.

Banaee, H., Ahmed, M. U., & Loutfi, a. A. (2013). Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges.

Black, A. C., Harel, O., & Matthews, G. (2012). *Techniques for analyzing intensive longitudinal data with missing values. In M. R. Mehl & T. S. Conner (Eds.), Handbook of research methods for studying daily life.* The Guilford Press.

CV, D., & AG, J. (1998). *Geostatistical software library and user's guide.* New York: Oxford University Press.

D., S., & M, T. D. (2013). *Managing and Mining of Sensor Data in Healthcare: A Survey. In: Aggarwal C.C.* Berlin: Springer.

Han, E., Miller, M. R., DeVeaux, C., Jun, H., Nowak, K. L., Hancock, J. T., . . . Bailenson, a. N. (2023). People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse . *Journal of Computer-Mediated Communication*.

JOHNSON, R. A., & WICHERN, D. W. (2007). *Applied Multivariate Statistical Analysis.*

Locascio, J. J., & Atri, A. (2011). An Overview of Longitudinal Data Analysis Methods for Neurological Research . *Dement Geriatr Cogn Disord Extra* .

M, P., & CY, L. (1999). Latent variables, measurement error, and methods for analyzing longitudinal binary and ordinal data. *Stat Med*.

Niall Bolger, J.-p. L. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research.* The Guilford Press new york London.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., . . . Scherer, a. L. (2022). Replicability, Robustness, and Reproducibility in Psychological Science.

Piepho, H. P., & Richter, A. B. (2004). *A Mixed Modelling Approach for Randomized Experiments with Repeated Measures.* Berlin: Blackwell Verlag.

R., R., D., B., J., L. B., & (2020), G. L. (2020). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. . *Media Psychology*.

SL, Z., & KY, L. (1992). An overview of methods for the analysis of longitudinal data. *Stat Med*.

Tan, X., Shiyko, M. P., Li, Y., & Dierker, L. (2012). *A time-varying effect model for intensive longitudinal data.*

TP, G., & K, M. (2017). Statistical Approaches to Longitudinal Data Analysis in Neurodegenerative Diseases: Huntington's Disease as a Model. . *Curr Neurol Neurosci Rep.*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Yaa Amankwah Awuah
awuahy@etsu.edu

Mostafa Zahed
zahedm@etsu.edu

APPENDIX

```
FILENAME REFFILE '/home/u58919495/time.csv';

PROC IMPORT DATAFILE=REFFILE
       DBMS=CSV
       OUT=Analysis;
       GETNAMES=YES;
RUN;

PROC CONTENTS DATA=Analysis; RUN;

/*view mean and standard deviation of dataset*/

proc means data=Analysis Mean StdDev ndec=3;
run;

/*normalize the dataset*/
proc stdize data=Analysis out=normalized_Analysis;
 var id intimacy time time01 treatment ;
run;

/*print normalized dataset*/
proc print data=normalized_Analysis;

/*view mean and standard deviation of normalized dataset*/
proc means data=normalized_Analysis Mean StdDev ndec=2;
 var id time time01 intimacy treatment;
run;
```

```
/* Calculate the min and max values */
proc sql;
    select min(intimacy), max(intimacy)
    into :min_intimacy, :max_intimacy
    from Analysis;
quit;

/* Apply the min-max normalization */
data Analysis_Norm;
    set Analysis;
    intimacy_normalized = (intimacy - &min_intimacy) / (&max_intimacy - &min_intimacy);
run;

/* Histogram */
proc sgplot data=Analysis;
    histogram intimacy;
run;

/* Time Plot */
proc sgplot data=Analysis;
    series x=time y=intimacy;
run;

/* Spagheti Plot */
PROC SQL;
    CREATE TABLE Analysis_means AS
    SELECT time, id, MEAN(intimacy) AS intimacy_mean
    FROM Analysis
    GROUP BY time, id;
QUIT;
PROC SGPLOT DATA=Analysis_means;
    SERIES X=time Y=intimacy_mean / GROUP=id;
RUN;

/* Interaction Plot */
PROC SQL;
```

```
   CREATE TABLE Analysis_means AS
   SELECT time, treatment, MEAN(intimacy) AS intimacy_mean
   FROM Analysis
   GROUP BY time, treatment;
QUIT;
PROC SGPLOT DATA=Analysis_means;
   SERIES X=treatment Y=intimacy_mean / GROUP=time;
RUN;


/* VARIOGRAM */
PROC REG DATA=Analysis;
   MODEL intimacy=treatment;
   OUTPUT OUT=Analysis_residuals RESIDUAL = residuals;
RUN;


PROC VARIOGRAM DATA=Analysis_residuals;
   COMPUTE MAXLAGS=10 LAGDIST=10;
   COORDINATES xc=id yc=time;
   VAR residuals;
RUN;


PROC SGPLOT DATA=Analysis_residuals;
   SERIES x=id y=time / LINEATTRS=(color=blue thickness=2);
RUN;



/* Means and Variances*/
PROC MEANS DATA=Analysis;
   VAR intimacy;
   BY time;
   OUTPUT OUT=Means_By_Time MEAN=Mean_Intimacy;
RUN;


PROC MEANS DATA=Analysis ;
   VAR intimacy;
   BY time;
```

```
    OUTPUT OUT=Variances_By_Time VAR=Variance_Intimacy;
RUN;


/* Residuals*/
DATA Analysis_residuals;
    SET Analysis ;
RUN;
PROC CONTENTS DATA = Analysis_residuals;
RUN;
DATA ResidualData;
    SET Analysis_residuals;
    RENAME residuals=residualData;
RUN;
PROC CONTENTS DATA = Analysis_residuals;
RUN;



PROC TRANSPOSE DATA=Analysis_residuals OUT=Analysis_Wide(drop=NAME)
PREFIX=time;
    BY id;
    ID time;
    VAR intimacy;
RUN;


/*Scatter Plot*/
PROC SGSCATTER DATA=Analysis_Wide;
    MATRIX _NUMERIC_;
RUN;



PROC STDIZE DATA=Analysis_Wide OUT=Analysis_Wide NOMISS;
RUN;

DATA Analysis_Wide ;
    SET Analysis_Wide ;
    IF CMISS(of ALL) > 0 THEN DELETE;
RUN;
```

```
/*RESIDUAL CORRELATION MATRIX */
OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;


PROC CORR DATA=Analysis_Wide OUTP=CorrOut;
   VAR _NUMERIC_;
RUN;


/*Conditional Longitudinal Models-RI-Model*/


PROC MIXED DATA=Analysis;
   CLASS id treatment;
   MODEL intimacy = treatment time / SOLUTION;
   RANDOM intercept / SUBJECT=id TYPE=VC;
RUN;
/*Random Slope Model*/
PROC MIXED DATA=Analysis;
   CLASS id treatment;
   MODEL intimacy = treatment time / SOLUTION;
   RANDOM intercept time / SUBJECT=id TYPE=UN;
RUN;
/*Comparing Models*/
PROC MIXED DATA=Analysis;
   CLASS id treatment;
   MODEL intimacy = treatment;
   RANDOM intercept / SUBJECT=id TYPE=UN;
   RUN;


PROC MIXED DATA=Analysis;
   CLASS id treatment;
   MODEL intimacy = treatment time;
   RANDOM intercept time / SUBJECT=id TYPE=UN;
   RUN;


/*Likelihood Ratio Test*/
PROC MIXED DATA=Analysis METHOD=ML;
   CLASS id treatment;
```

```
   MODEL intimacy = treatment;
   RANDOM intercept / SUBJECT=id TYPE=UN;
   RUN;


PROC GLM DATA=Analysis;
   CLASS treatment;
   MODEL intimacy = treatment;
   OUTPUT OUT=Resid_Pred P=Predicted R=Residuals;
   RUN;
PROC UNIVARIATE DATA=Resid_Pred ;
   VAR Residuals;
   OUTPUT OUT=RangeResid RANGE=RangeResid;
   RUN;


PROC PRINT DATA=RangeResid;
   VAR RangeResid;
   RUN;
/*This print calculates the range of reasiduals*/


/*Normality plot */
PROC UNIVARIATE DATA=Resid_Pred PLOT;
   VAR Residuals;
   QQPlot Residuals / NORMAL (MU=EST SIGMA=EST);
   RUN;


/*Shapiro-Wilk Test*/
PROC UNIVARIATE DATA=Resid_Pred;
   VAR Residuals;
   HISTOGRAM / NORMAL;
RUN;


/*computation of residuals and predicted values from the random time model (RTModel)*/


proc reg data=Analysis;
   model intimacy = time time01 treatment;
   output out=resid_data r=residual p=predicted;
```

```
run;

/* To see the range of residuals, you might use proc means */
proc means data=resid_data min max;
   var residual;
run;

proc reg data=Analysis outest=estimates ;
   model intimacy = time time01 treatment;
run;

proc score data=Analysis score=estimates out=scored_data;
   var time time01 treatment;
run;

/*Q-plot*/
PROC REG DATA=Analysis;
    MODEL intimacy = time time01 treatment;
    OUTPUT OUT=RegOut PREDICTED=RIPred RESIDUAL=RIResid;
RUN;
QUIT;
PROC SGPLOT DATA=RegOut;
    SCATTER x=RIPred y=RIResid;
RUN;
QUIT;

/*Residuals of RT moDEL*/
/* Create dataset with residuals */
DATA Residuals;
    SET Analysis;
    RTResid = RESIDUAL;
RUN;

/* Perform Shapiro-Wilk test and create Q-Q plot */
PROC UNIVARIATE DATA=Residuals PLOTS;
    VAR RTResid;
```

```
    HISTOGRAM / NORMAL;
    QQPLOT / NORMAL (MU=EST SIGMA=EST);
    PROBPLOT / NORMAL;
RUN;
QUIT;
```