

Lab 1 - Linear Regression in R

Haroon Popal
February 12, 2021

Research Question

Interpersonal relationships have significant implications for perception, emotion, memory, motivation, and decision-making. In a long term perspective, having strong social relationships are important for personal success, health and well-being. One theory proposes that humans have a unique ability to track a large number of social relationships, which have allowed us to live in large groups and societies. As humans enjoy very diverse social lives across their family life, careers, friends, communities, and social media, the research question this project attempts to address is what is the impact of a decrease in the typical number and variety of social relationships on mental health.

Data Source

Participants were recruited from Amazon Mechanical Turk (mTurk). mTurk is hosted by Amazon and provides an online platform for individuals (mTurkers) to partake in surveys. Use of mTurk has become popular in psychology research, as it allows for larger sample sizes than what can be typically collected in a lab setting. A link to the survey was posted on mTurk using turkprime.com. The use of “bots” to pose as real participants has been an issue with mTurk, and turkprime has a system that catches and excludes some of these bots. Participants from across the United States were eligible to take the survey. An additional eligibility criteria was that participants must have learned English as their first language. This is due to cultural differences in social relationships, which we believe exist but are unable to directly address. In total, 768 participants completed the survey at the first time point which will be analyzed here.

Dropped Cases

After excluding six participants who failed attention check questions (questions which have an obvious answer such as selecting which of five items is typically red), and 34 participants who did not follow directions in how to answer questions (participants need to respond in a specific syntax so that their responses could be easily captured), 728 participants from the first time point were included for the analysis.

Variables

This study includes variables that can be grouped as predictors, outcomes, and nuisance variables. The outcome we measured was a latent factor of mental well-being, derived from measures of anxiety, behavioral/emotional control, depression, calm-positive affect, and

happy/positive affect. These five variables were measured by asking participants whether they felt each factor much less, less, about the same, more, or much more in the past week. The predictors of interest were the number of social relationships, the variety of social relationship, and the quality of social relationships they experienced in the past week. Participants were asked to report each person they interacted with in the past week, and their relationship to that person, results in their number of relationships (i.e. 10 relationships), and variety of relationship (i.e. 5 unique relationships) reported. They were also asked to rate how positive or negative each relationship was on a scale of one to five indicating very negative, negative, neutral, positive, or very positive. Nuisance variables were environmental richness, measured by a report of how much they traveled outside of their home in the past week, and physical well-being, measured by a report of how they feel similar to the mental well-being questions).

Proposed Path Analysis

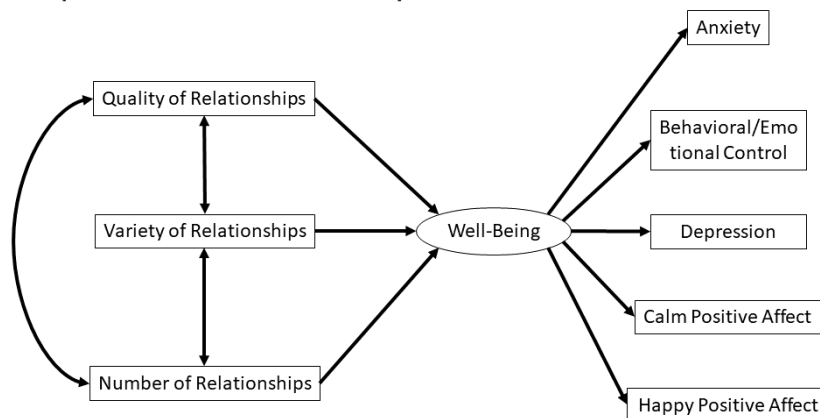


Figure 1. Proposed path analysis We hypothesize that a latent factor of mental well-being will be significantly predictive of anxiety, behavioral/emotional control, depression, calm positive affect, and happy positive affect. We further hypothesize that features of social relationships that individuals experience at home will be predictive of the latent factor of mental well-being.

Variable definitions

Table 1. Questionnaire items. Variables come from two major domains: social and well-being. In the survey, questions were asked that probed aspects of these domains. From these questions a series of variables were derived. Multiple variables could be derived from a single question, as some item-level responses were free responses which provided rich detail.

Domain	Question	Abbreviation(s)	Item-level responses
Social	How many individuals live in your household, not including yourself?	Immediate network size (INS)	Numeric response

	Please list the relationship that each individual in your household has with you and indicate the amount of time you have spent with them (in hours), during the past week.	IN_URels	Numeric response for each relationship (i.e. mother: 40, father: 40, brother1: 60, brother2: 20)
	Please list the relationship that each individual in your household has with you and indicate how positive or negative that relationship has been, during the past week	IN_valw	Very positive: 5, Positive: 4, Neutral: 3, Negative: 2, Very negative: 1 Each relationship was given its own rating (i.e. mother: 3, father: 3, brother1: 5, brother1: 1)
Well-being (Wellb)	How much of the time, during the past week, have you been a nervous person?	anxiety	Much less than usual, Less than usual, About the same as usual, More than usual, Much more than usual
	How much of the time, during the past week, have you felt calm and peaceful?	pos_affect_calm	Much less than usual, Less than usual, About the same as usual, More than usual, Much more than usual
	How much of the time, during the past week, have you felt downhearted and blue?	depression	Much less than usual, Less than usual, About the same as usual, More than usual, Much more than usual
	How much of the time, during the past week, have you been a happy person?	pos_affect_happy	Much less than usual, Less than usual, About the same as usual, More than usual, Much more than usual
	How much of the time, during the past week, have you felt so down in the dumps that nothing could cheer you up?	behav_emo_control	Much less than usual, Less than usual, About the same as usual, More than usual, Much more than usual

We also created a latent variable of well-being from the five mental health questions. This latent variable served as a predictor of the mental health outcomes, and an outcome of the social relationship variables.

Scatterplot Matrix

Scatterplot matrix for outcome variable of well-being and predictors.

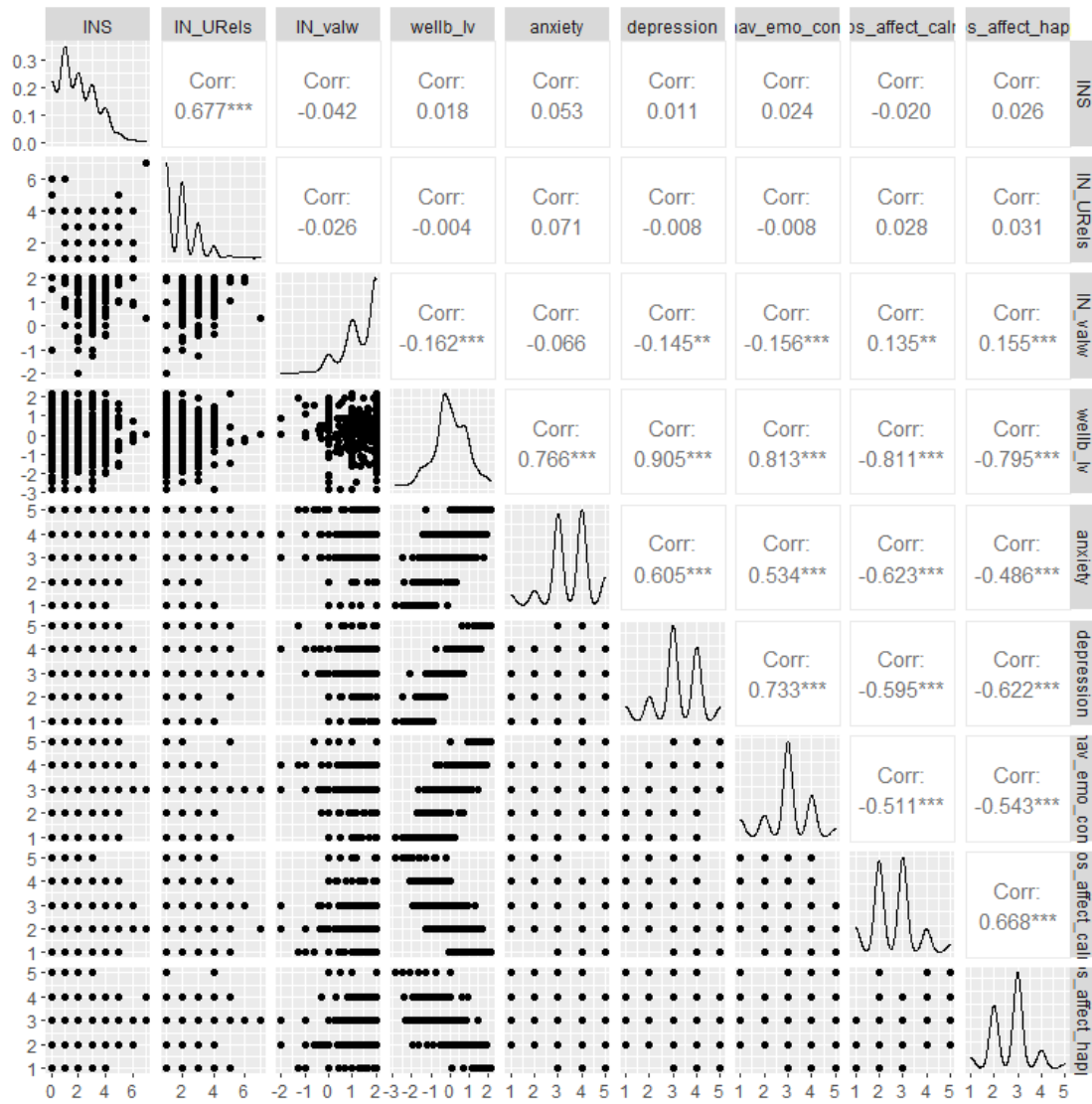


Figure 2. Scatterplot matrix of outcome and predictors.

Distribution of Outcome

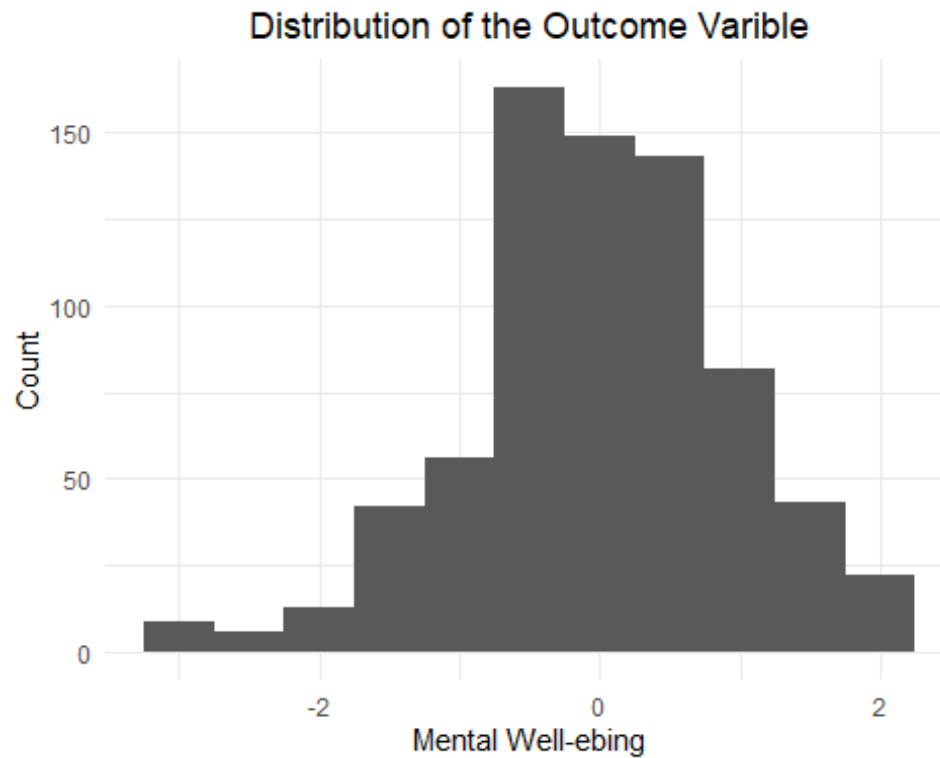


Figure 3. Histogram of the mental well-being outcome variable.

Summary Statistics

Table 2. Summary statistics.

Measure	Mean	Standard Deviation
INS	1.84	1.42
IN_URels	1.90	1.00
IN_valw	1.34	0.75
Anxiety	3.48	0.93
Depression	3.22	0.95
Behavioral/emotional control	3.01	0.94
Positive affect calm	2.58	0.92
Positive affect happy	2.70	0.80
Well-being	0	0.94
Boredom	3.50	0.95
Interests	3.21	1.13
Regulation	2.83	0.67

Routine	2.11	0.89
Health	3.76	0.79

Regression Table

Table 3. Path analysis results.

Dependent Variable	Independent Variable	β	SE	<i>p</i>
Anxiety	Well-being	0.39	0.05	< .001
Positive calm affect	Well-being	-0.42	0.05	< .001
Depression	Well-being	0.46	0.06	< .001
Positive happy affect	Well-being	-0.38	0.05	< .001
Behavioral/emotional control	Well-being	0.43	0.06	< .001
Well-being	INS	0.03	0.06	0.636
Well-being	IN_URels	0.00	0.08	0.978
Well-being	IN_valw	-0.38	0.14	0.001

lavaan plot

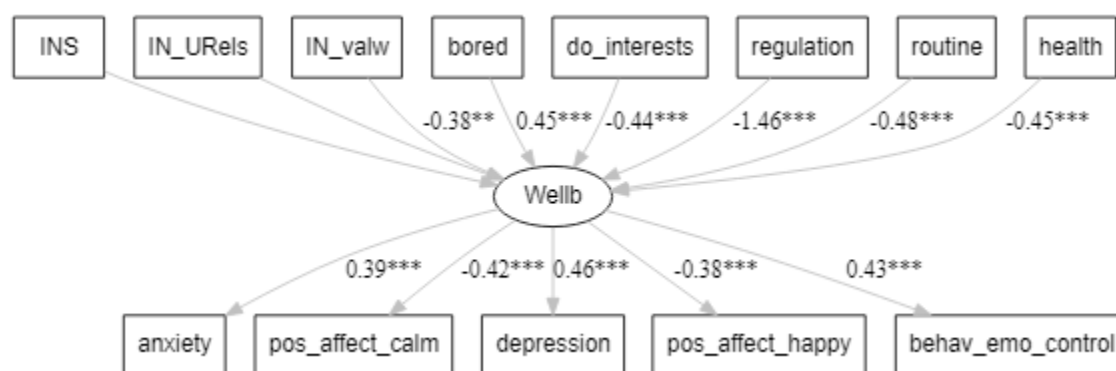


Figure 4. Path analysis results indicating significant paths.

Write-up

Immediate Network Relationship Valence Predicts Well-being

A path analysis was done to test whether changes in social relationships due to social distancing measures was related to a latent variable of well-being. Overall, the model fit for this analysis was poor, with a significant chi-square ($\chi^2 = 307.88$, $p < .001$), low CFI (0.87), low TLI (0.84), and a high RMSEA (0.09, CI [0.08 0.10]).

The latent variable of well-being was significantly related to each of the five mental health items in the model. Higher well-being was related to higher levels of anxiety, depression, and behavioral/emotional control issues, and lower levels of calm and happy positive affect. The well-being latent variable was then used as a dependent variable for the predictors of immediate network size, number of unique relationships, time-weighted immediate network valence, and the nuisance variables of boredom, ability to do interests, regulation of emotions, ability to do normal routine, and general physical health. Immediate network size and number of unique relationships were not significant predictors of well-being, but time-weighted immediate network valence was, despite the five nuisance variables also significantly predicting mental well-being (Table 3; Fig. 4). More time spent with negatively valenced relationships was predictive of higher levels of the well-being latent factor, which was related to higher levels of anxiety, depression, behavioral/emotional control issues, and lowers levels of calm and happy positive affect (Fig. 5). Specifically, a 0.38 decrease in the average valence of relationships in one's immediate network was indicative of a one unit increase in worse mental well-being.

Graph

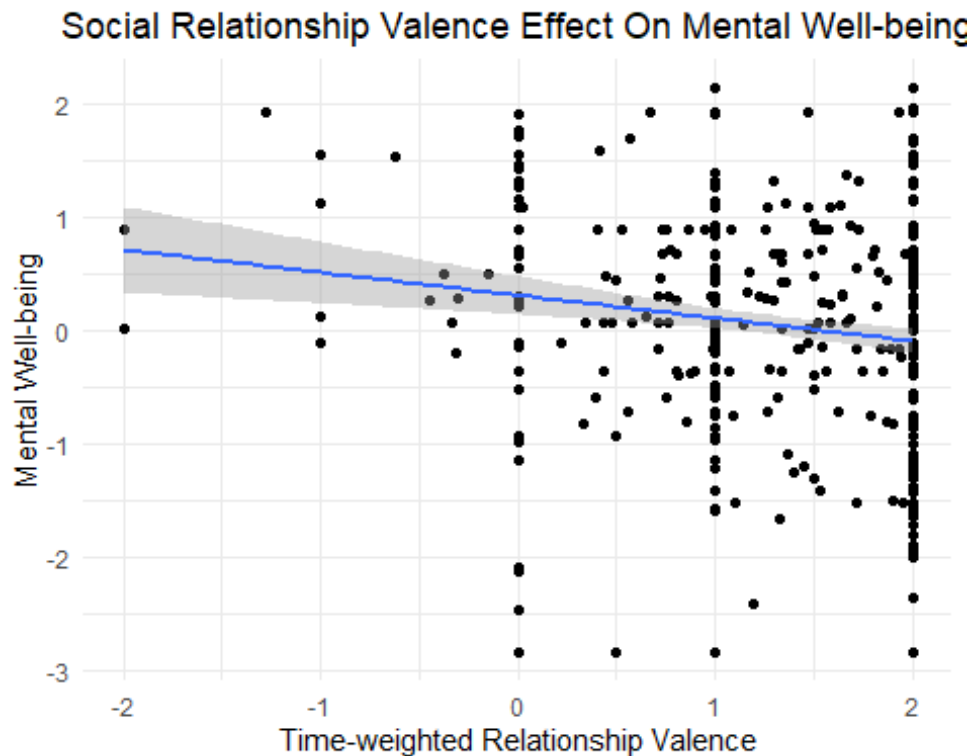


Figure 5. Relationship between immediate network relationship valence and mental well-being.

Appendix

```
# Create latent variable
```

```
wellb_mod_mg <- 'Wellb =~ NA*anxiety + pos_affect_calm + depression + pos_affect_happy +  
behav_emo_control  
Wellb ~~ 1*Wellb'
```

```
wellb_configural_res <- sem(wellb_mod_mg, model1_data, estimator='wlsmv')  
model1_data$wellb_lv <- as.numeric(predict(wellb_configural_res))
```

```
# Create variable scatterplots
```



```
ggpairs(model1_data[c('INS', 'IN_URels', 'IN_valw',
                     'wellb_lv',
                     'anxiety', 'depression', 'behav_emo_control', 'pos_affect_calm',
                     'pos_affect_happy')], missing='exclude')
```

```
# Plot outcome distribution
```

```
ggplot(model1_data, aes(x=wellb_lv)) + geom_histogram(binwidth = 0.5) +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5)) +
  ggtitle("Distribution of the Outcome Variable") +
  xlab("Mental Well-being") + ylab("Count")
```

```
Wellb_mod_mg <- 'Wellb =~ NA*anxiety + pos_affect_calm + depression + pos_affect_happy +
behav_emo_control
```

```
    Wellb =~ 1*Wellb
```

```
    Wellb ~ INS + IN_URels + IN_valw + bored + do_interests + regulation + routine +
health
```

```
    INS =~ INS
```

```
    IN_URels =~ IN_URels
```

```
    IN_valw =~ IN_valw
```

```
    bored =~ bored
```

```
    do_interests =~ do_interests
```

```
    regulation =~ regulation
```

```
    routine =~ routine
```

```
    health =~ health'
```

```
Wellb_Configural_res <- sem(Wellb_mod_mg, model1_data, estimator='wlsmv')
Wellb_Configural_res
```

Model fit information

```
trim_fit(Wellb_Configural_res, Wellb_Configural_res_fit)
```

#View model fit information

```
Wellb_Configural_res_fit
```

```
##          index  value
## npar          npar 26.000
## chisq          chisq 307.876
## df            df 65.000
## pvalue          pvalue 0.000
## cfi            cfi 0.870
## tli            tli 0.844
## rmsea          rmsea 0.089
## rmsea.ci.lower rmsea.ci.lower 0.079
## rmsea.ci.upper rmsea.ci.upper 0.099
## rmsea.pvalue   rmsea.pvalue 0.000
```

Model fit is poor, but not horrible. Results should be cautiously considered

Parameter Estimates

```
parameterestimates(Wellb_Configural_res, ci = F, standardized = T)
```

```
##      lhs op      rhs  est  se    z pvalue std.lv
## 1 Wellb =~      anxiety 0.390 0.051 7.605 0.000 0.648
## 2 Wellb =~ pos_affect_calm -0.423 0.054 -7.905 0.000 -0.703
## 3 Wellb =~      depression 0.457 0.061 7.458 0.000 0.759
## 4 Wellb =~ pos_affect_happy -0.380 0.049 -7.765 0.000 -0.632
## 5 Wellb =~ behav_emo_control 0.430 0.056 7.658 0.000 0.715
## 6 Wellb ~~      Wellb 1.000 0.000  NA   NA 0.362
## 7 Wellb ~      INS 0.030 0.064 0.473 0.636 0.018
## 8 Wellb ~      IN_URels -0.002 0.083 -0.027 0.978 -0.001
## 9 Wellb ~      IN_valw -0.380 0.135 -2.815 0.005 -0.229
```

```

## 10      Wellb ~      bored 0.455 0.133 3.409 0.001 0.274
## 11      Wellb ~ do_interests -0.443 0.120 -3.698 0.000 -0.267
## 12      Wellb ~      regulation -1.464 0.243 -6.015 0.000 -0.881
## 13      Wellb ~      routine -0.482 0.144 -3.342 0.001 -0.290
## 14      Wellb ~      health -0.452 0.122 -3.692 0.000 -0.272
## 15      INS ~~      INS 1.479 0.100 14.746 0.000 1.479
## 16      IN_URels ~~      IN_URels 1.006 0.094 10.658 0.000 1.006
## 17      IN_valw ~~      IN_valw 0.567 0.048 11.852 0.000 0.567
## 18      bored ~~      bored 0.920 0.060 15.354 0.000 0.920
## 19      do_interests ~~      do_interests 1.203 0.077 15.694 0.000 1.203
## 20      regulation ~~      regulation 0.444 0.039 11.536 0.000 0.444
## 21      routine ~~      routine 0.730 0.049 14.911 0.000 0.730
## 22      health ~~      health 0.629 0.038 16.497 0.000 0.629
## 23      anxiety ~~      anxiety 0.426 0.041 10.285 0.000 0.426
## 24      pos_affect_calm ~~      pos_affect_calm 0.329 0.031 10.661 0.000 0.329
## 25      depression ~~      depression 0.311 0.033 9.484 0.000 0.311
## 26      pos_affect_happy ~~      pos_affect_happy 0.225 0.023 9.694 0.000 0.225
## 27      behav_emo_control ~~      behav_emo_control 0.391 0.037 10.459 0.000 0.391
##      std.all std.nox
## 1      0.705 0.705
## 2      -0.775 -0.775
## 3      0.806 0.806
## 4      -0.800 -0.800
## 5      0.753 0.753
## 6      0.362 0.362
## 7      0.022 0.022
## 8      -0.001 -0.001
## 9      -0.172 -0.172
## 10     0.263 0.263
## 11     -0.293 -0.293
## 12     -0.587 -0.587
## 13     -0.248 -0.248
## 14     -0.216 -0.216

```

```
## 15  1.000  1.000
## 16  1.000  1.000
## 17  1.000  1.000
## 18  1.000  1.000
## 19  1.000  1.000
## 20  1.000  1.000
## 21  1.000  1.000
## 22  1.000  1.000
## 23  0.503  0.503
## 24  0.399  0.399
## 25  0.351  0.351
## 26  0.361  0.361
## 27  0.433  0.433
```

```
# Create lavaan plot
```

```
wellb_conf_fig <- lavaanPlot(model = Wellb_Configural_res,
  node_options = list(shape = "box", fontname = "Helvetica"),
  edge_options = list(color = "grey"),
  graph_options = list(layout = "dot", rankdir = "TB"),
  coefs = T, stand = F, sig = 0.05, covs = T, stars = c("regress", "latent"))
wellb_conf_fig
```

```
# Regression graph
```

```
ggplot(model1_data, aes(x=IN_valw, y=wellb_lv)) +
  geom_point() +
  geom_smooth(method='lm') +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5)) +
  ggtitle("Social Relationship Valence Effect On Mental Well-being") +
  xlab("Time-weighted Relationship Valence") + ylab("Mental Well-being")
```

```
## `geom_smooth()` using formula 'y ~ x'
```