Self-compassion and Gratitude Scales and Their Association with Demographics

Cheng Peng

03/12/2022

Table of Contents

Introduction

The goal of this research project is to measure the level of self-compassion as well as the self-care of BSW and MSW students in a Social Work Program at a regional University. We will be using the following two reliable and validated instruments to measure their level of self-compassion and self-care as they immerse themselves in the helping profession. We hope to see how the SC of the students correlates to other independent variables i.e. undergrad/grad program social work, age, education level, religiosity, spirituality, gender, etc.

1. The Self-Compassion Scale
2. The Gratitude Questionnaire

The purpose of our research is to study the perception of self-compassion in social work students and how it can link to self-care as well as success in the social work program and field. This will help provide students with self-care practices during their training to thrive in the profession in the future.

Data Management and Analyzing Survey Instruments

The original survey data have three components, self-compassion scale and gratitude questionnaire instruments, and some demographic questions.

The three components have different portions of missing values. We split the original data set into three subsets of data and impute the missing values related to the self-compassion and gratitude data based on the survey instruments. Since there are only a few missing values, we replace the missing values in each survey question with the mode of the associated survey item. We create indexes of the two instruments separately to aggregate the information in the two survey data sets.

Since R does not have a function to find the model of a given data set, I write the following function to find the model of a data set.

We will perform both principal component analysis (PCA) and exploratory factor analysis (EFA).

Handling Missing Values Self-Compassion Instrument

This instrument contains only the data associated with the 12 items in the survey instrument. In the original data file, the 12 variables are named as Q2_1, Q2_2, …, Q2_12. We impute the missing value by replacing the missing value in each of the 12 items with the mode of the corresponding survey items. Since there are only a few missing values in this instrument, this imputation will not impact the subsequent PCA and EFA.

Handling Missing Values in Gratitude Scale

The gratitude questionnaire contains only the variables associated with gratitude questions. The variables used in the original data file are Q3_1, Q3_2, ..., Q3_6. We use the same mode imputation method to fill in the missing values as used in the above self-compassion survey data. The gratitude questionnaire has even fewer missing values. Any imputation will not impact any subsequent analysis.

Since Likert scales of the Q3_3 and Q3_6 were in reverse order in the design. We transform back the usual order and create a new dataset using the same variable names.

Handling Demographic Variables

The demographic variables have two issues: missing values and imbalance categories. Since the size of the data set is slightly close to 120, imputing missing values in a meaningful way is crucial to maintain the sample size and the statistical power of all subsequent association analysis. About 15 records in the data sets do not have demographic information. Therefore these records were deleted in the final data.

A few missing values occurred in the years of education and employment that are imputed using the auxiliary information in the variables of age, the years of education, and the length of employment.

The major issue of these categorical variables is the imbalance category. The following modifications to the original demographic variables are utilized.

```
# replace missing values with 99.
demographic00[is.na(demographic00)] <- 99
# Create a frequency table for collapsing categories
list(Q8.1=table(demographic00$Q8_1),
     Q8.2=table(demographic00$Q8_2),
     Q8.3=table(demographic00$Q8_3),
     Q8.5=table(demographic00$Q8_5),
     Q8.6=table(demographic00$Q8_6),
     Q.9=table(demographic00$Q9),
     Q.11=table(demographic00$Q11),
     Q.13=table(demographic00$Q13),
     Q.14=table(demographic00$Q14),
     Q.15=table(demographic00$Q15),
     Q.16=table(demographic00$Q16),
     Q.17=table(demographic00$Q17),
     Q.18=table(demographic00$Q18),
     Q.19=table(demographic00$Q19),
     Q.20=table(demographic00$Q20)
     )

grp.age = Q8_1:  1 = (3,23], 2 = [24, 30],  3 = [31, 59]
grp.edu = Q8_2:    1 = [0,15] associate, 2 = [15.5,18.5] bachelor, 3 =
```

```
[19, 25]  advdegree
grp.empl = Q8_3:   1 = [0,5]  entry, 2 = [5.5,10] junior, 3 = [10.5,
35]   senior
kid.num = Q8_5: 1 = (0) No child, 2 = at least one child
home.size = Q8_6:  1 = (1), 2 = (2), 3 = 3 or more

gender = Q9:  1 = (2) female, 2 = (1,3,5,7) male
race = Q11:   1 = (1)   white,   2 = (2,3) other
marital.st = Q13: 1 = (1) single, 2 = (2) married/civil Partner, 3 =
other
disability = Q14:  1 = yes,  2 = No
religion = Q15:  1 = 9 (no religion), 2 = (1,2,3,6,7,8,10,11,99)
(religion)
Sexual.orient = Q16:  1 = (4)  heterosexual,  2 = (1,2,3,5,6,7,10)
other
poli.affil = Q17: 1 = (4)ind, 2 = (5,6,7)  democrats, 3 = (1,2,3, 99)
other
SW.Program = Q18: 1 = (1) Business Student, 2 = (2,99) SWK Students
Urbanity = Q19:  1 = (1) urban, 2 = (2) rural, 3 = (3) suburban
Spirituality = Q20: 1 = (1,2,3) low, 2 = (4)   moderate,  3 = (5,6,7)
high
```

We re-code the demographic variables based on the above modification. The modified demographic variables will be used in subsequent modeling.

PCA and EFA for Survey Instruments

We perform both principal component analysis (PCA) and the exploratory factor analysis (EFA). The number We next try several exploratory factor analysis models to find the best model that explains the most variation of the data.

Different methods for finding the number of components/factors to retain in an exploratory principal component or factor analysis are used in this analysis. The classical ones are the Kaiser rule, the parallel analysis, and the usual scree test. Non-graphical solutions by Raiche et al. (2013) to the Cattell subjective scree test are also proposed: an acceleration factor (AF) and the index of the optimal coordinate OC. The acceleration factor indicates the turning point of the scree plot. It corresponds to the acceleration of the curve, i.e. the second derivative. The optimal coordinates are the extrapolated coordinates of the previous eigenvalue that allow the observed eigenvalue to go beyond this extrapolation.

Our objective is to aggregate the information to a few principle components that carry as much information as possible, then use simple statistical analysis to characterize the behavior of the resulting PCA or EFA scores and relevant association analyses. Therefore, the standard orthogonal rotation is used to avoid the potential correlation between the resulting PCA or EFA scores.

Whether PCA or EFA will be used and how many of the PCAs or EFAs will be selected will be based on the well-received analytical methods and subjective judgment on the proportion of variables explained by the selected PCAs or EFAs. In case, both PCA and EFA give similar results in terms of the proportion of the total variation, we will go with PCA since PCA's estimation of loading does not assume distribution of normality to define the maximum likelihood objection. This will reduce the risk of model misspecification.

Some R Functions For Extracting Information from PCA and EFA

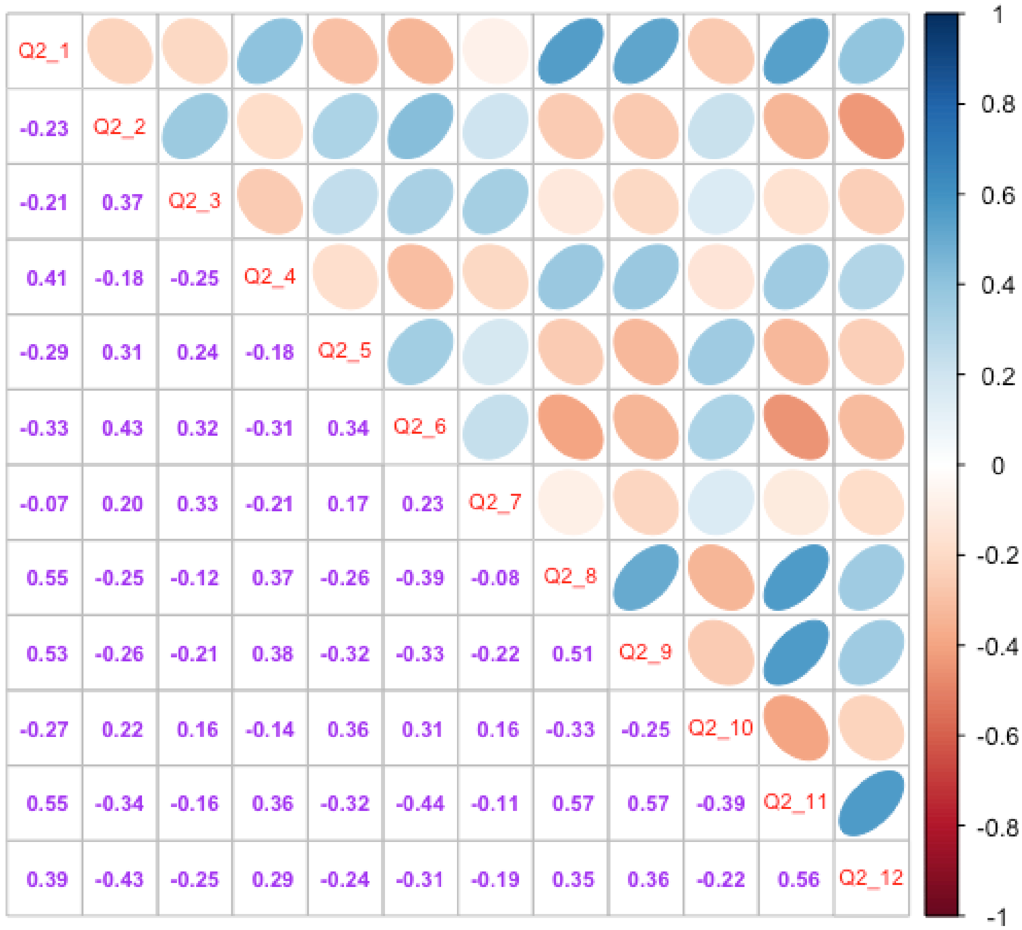
We defined and modified some R functions to extract specific information from the PCA or EFA analyses to address the analytic questions outlined in the previous section.

My.plotnScree produce a plot that summarized the results of four commonly used methods used for identifying the number of components or factors to be retained for exploratory analyses. This graphics function was modified from the *plotnScree()* from library **nFactors**.

My.loadings.var produce a list of two objects: factor loadings and proportion variance explained by each factor. There are no existing R functions that can be used to extract the proportion of variance from the output of *factanal()*. The function can also extract similar information from the output of a PCA but we need to specify the method in the argument.

Self-compassion Index

We start with some correlation plot to see the relevance of the PCA procedure on the self-compassion data.

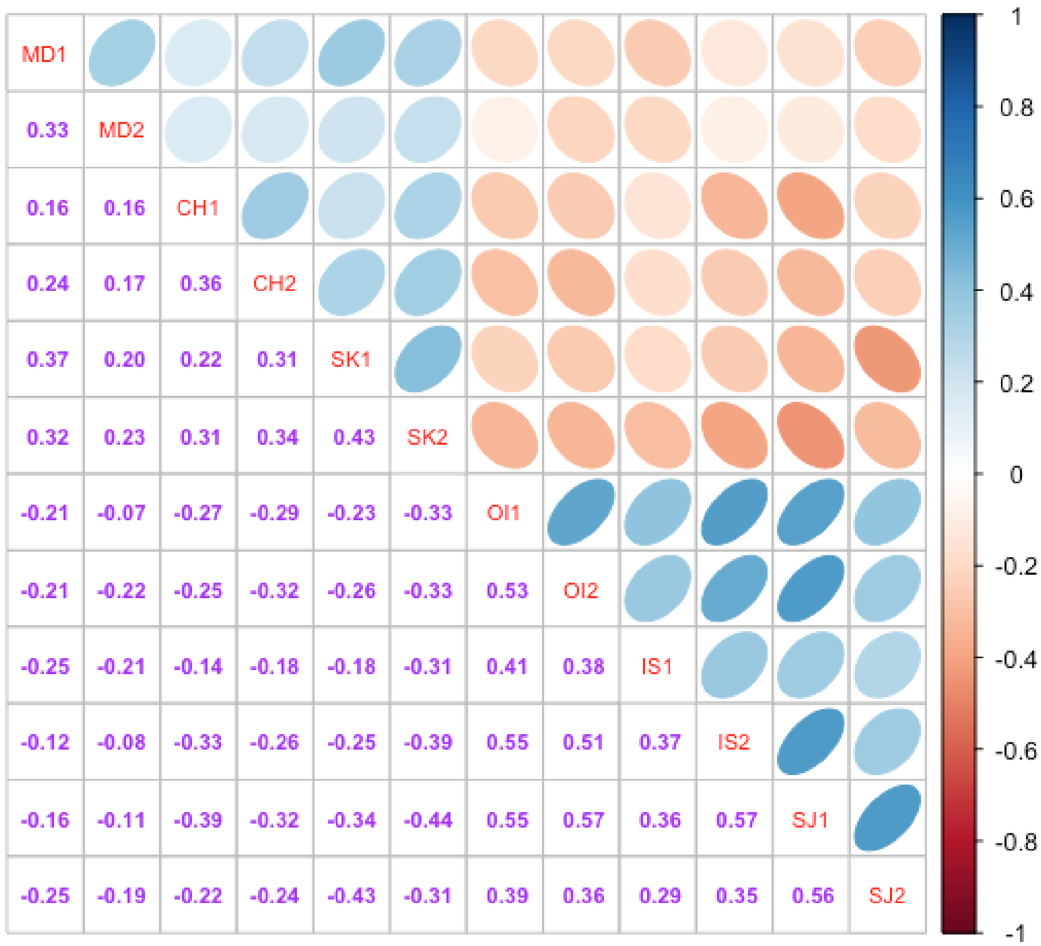


The pairwise correlation plot reveals the potential relevance of PCA. The shape of an ellipse represents the correlation. The skinnier the ellipse, the higher the correlation. The direction reflects whether a correlation is positive or negative. The off-diagonal direction implies s positive correlation while the main diagonal direction implies a negative association.

Figure 1 shows the moderate association between individual survey items. This implies that the PCA is relevant in aggregating the information in the survey items.

Similar to López et. al. (2018), we use the six positive items associated with mindfulness, self-kindness and the sense of common humanity to define the self-compassion index.

From the pair-wise correlation plot we can see that some of the items are reversely scored. We next change back the scale so that all items are positively correlated.



The pairwise correlation plot reveals the potential relevance of PCA. Group items based on self-compassion and self-coldness

Next, we make the following heatmap to illustrate the pairwise correlation between the items in the survey instrument based on the positive adjustment of the scale.

Pairwise correlation based on negatively adjusted subscales. The shape of an ellipse represents the correlation. The skinnier the ellipse, the higher the correlation. The direction reflects whether a correlation is positive or negative. The off-diagonal direction implies s positive correlation while the main diagonal direction implies a negative association.

Intern Consistency

With the adjusted the scores in the self-compassion instrument, we calculate the one of the commonly used internal consistency reliability Cronbach alpha as follows.

Next we find the Cronbach alpha and it 95% confidence interval.

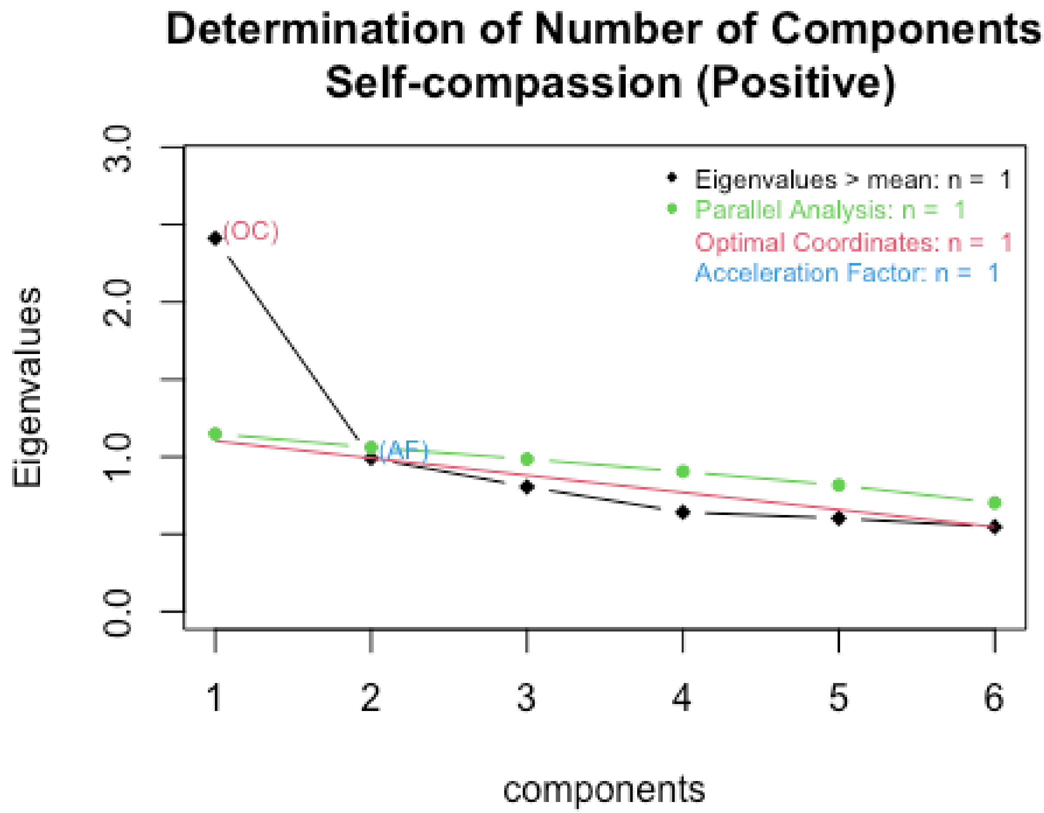
Confodence Interval of Cranbach Alpha

LCI	alpha	UCI
0.6330792	0.7002141	0.7586652

We can see that the Cronbach's alpha is 0.77 with 95% confidence interval (0.70, 0.84) suggesting that the items in the self-compassion instrument have relatively high internal consistency.

PCA Extraction and Number of PCA Determination

The number of PCAs selected for the future exploratory analyses is the key issue and is also the first question we need to address before we move to any further analysis with the PCA scores. Raiche et al (2013) simulation-based test and Scree plot indicate that it is sufficient to choose the first principle component for future analysis. For exploratory purposes, we will choose the first two principal components for both PCA and EFA procedures and use them for association analysis.



Different methods of identification of the number of principal components to be retained in exploratory analysis: Kaiser's eigenvalue rule, Raiche et al Monte Carlo simulation method (parallel analysis), optimal coordinate (OC) index, and accelerate factor (AF) method.

Figure 2 indicates that it is sufficient to retain the first principle component for the subsequent analysis. In the following, we will extract the first two PCAs. The PCA factor loadings and the proportion of variance explained by the retained PCAs are summarized in the following tables.

Factor loadings of the first few PCAs and the cumulative proportion of variation explained by the corresponding PCAs in the self-compassion survey.

	PC1	PC2
MD1	0.414	-0.470
MD2	0.329	-0.556
CH1	0.363	0.546
CH2	0.415	0.397
SK1	0.447	-0.066
SK2	0.466	0.096

Cumulative and proportion of variances explained by each principal component in the self-compassion survey.

	PC1	PC2
Standard deviation	1.553	0.996
Proportion of Variance	0.402	0.165
Cumulative Proportion	0.402	0.567

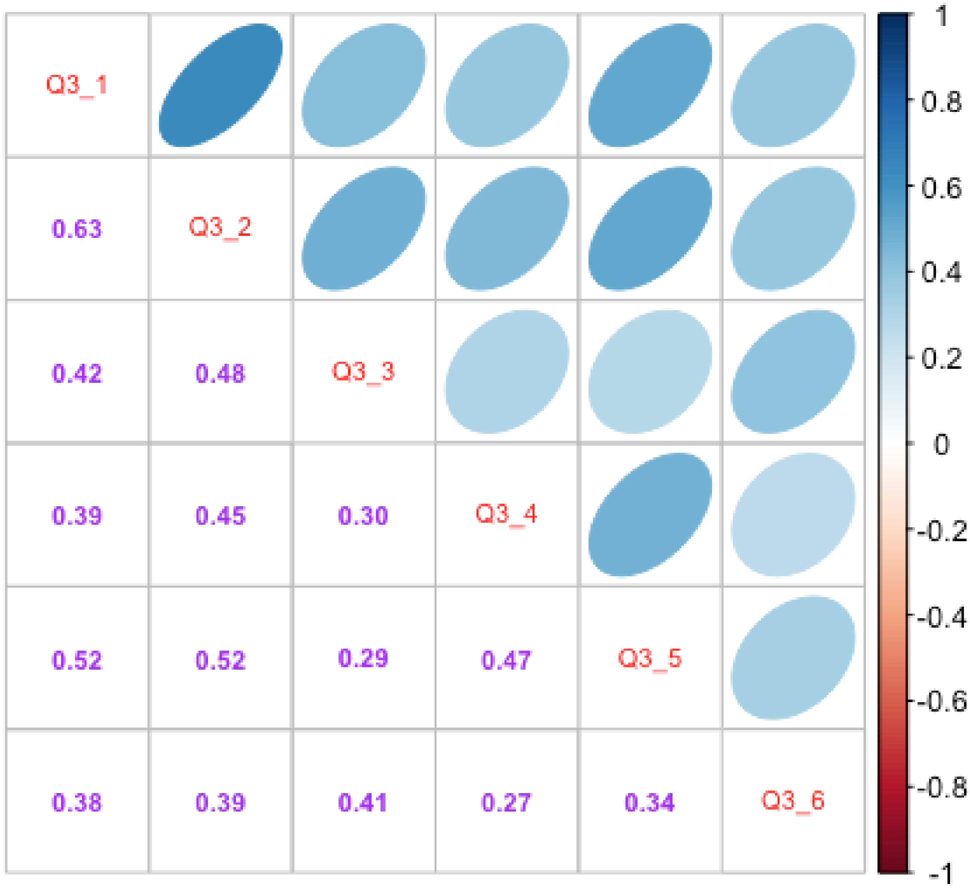
We also conduct the same analysis using EFA. The Scree type of test also suggests retaining a single factor. The proportion of total variation is lower than that of PCA. we decide using PCA method and extract the first two principle components for the future analysis. Table 1 shows the factor loadings of the first two principle components. We can see that each of the original items contribute to the two PCAs evenly in terms of the magnitude. The first PCA counts about 41.3% of the total variation and the second PCA counts 10.9% of total variation. We can simply call the first PCA as *self-compassion index*, denoted by sc.idx.

Next we extract the self-compassion index in the following code

Histogram of the first principle component extract from the self-compassion survey.

Gratitude Questionaire

In this section, we perform the same analysis on the gratitude survey instrument. First of all, we present a pairwise correlation plot to display the correlation between individual survey items in the gratitude survey instrument.



The pairwise correlation plot reveals the potential relevance of PCA for the gratitude instrument. The shape of an ellipse represents the correlation. The skinnier the ellipse, the higher the correlation. The direction reflects whether a correlation is positive or negative. The off-diagonal direction implies s positive correlation while the main diagonal direction implies a negative association.

Figure 3 shows that a moderate correlation exists between individual variables. This implies the PCA and EFA can be used to aggregate the information in the set original survey items. Next, we estimate the number of PCAs or EFAs to be retained for the subsequent analysis using the commonly used procedures and summarize the results in the following figure 4.

Cronbach's Alpha

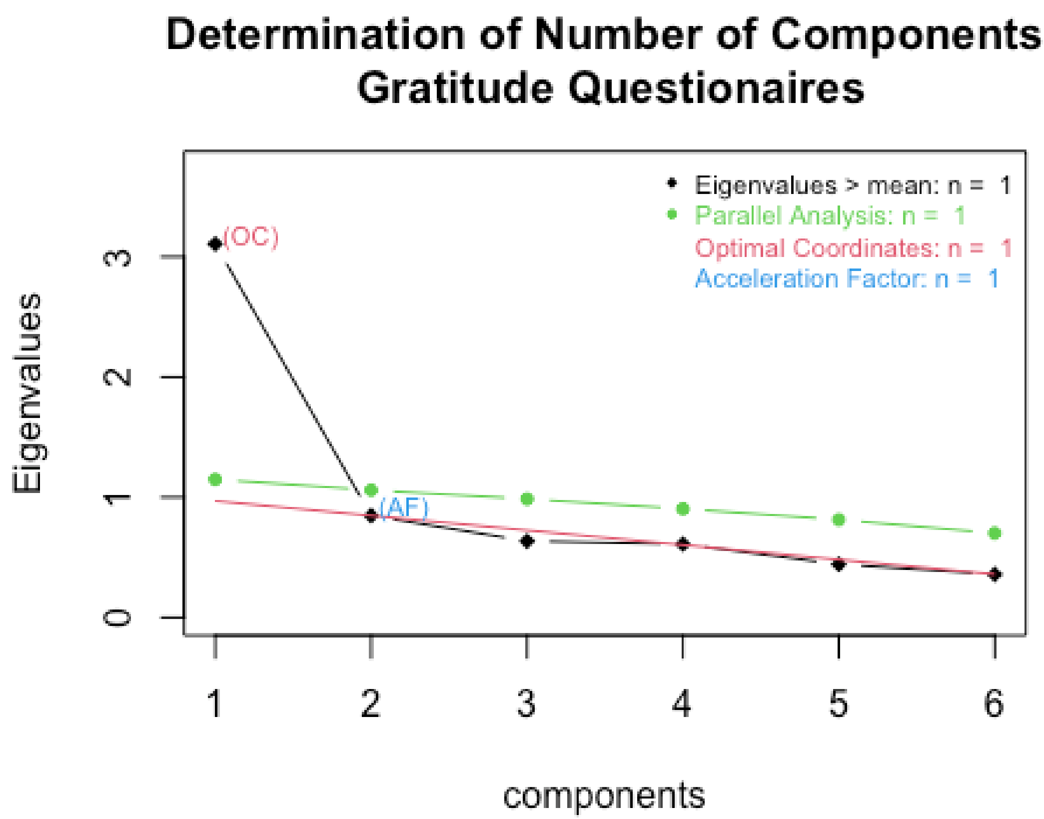
The internal consistency measure, Cronbach's alpha, of the gratitude instruments calculated below

Confodence Interval of Cranbach Alpha

LCI	alpha	UCI
0.7389249	0.7866934	0.828283

We can see that the Cronbach's alpha is 0.8 with a 95% confidence interval (0.74, 0.86) also suggesting that the items in the Gratitude rating instrument have relatively high internal consistency.

PCA Extraction and Number of PAC Determination



Different methods of identification of number of principle components to be retained in exploratory analysis for the gratitude survey instrument: Kaiser's eigenvalue rule, Raiche et al Monte Carlo simulation method (parallel analysis), optimal coordinate (OC) index, and accelerate factor (AF) method.

Figure 5 indicates retaining one PCA is sufficient for future exploratory analyses. As we did in the self-compassion survey instrument, we extract the first two principal components for potential analysis. The factor loadings of the two principal components and the corresponding proportion of variation of each component are summarized in the following two tables.

Factor loadings of the first few PCAs and the cumulative proportion of variation explained by the corresponding PCAs in the Gratitude Questionaire Survey.

	PC1	PC2
Q3_1	0.450	-0.044

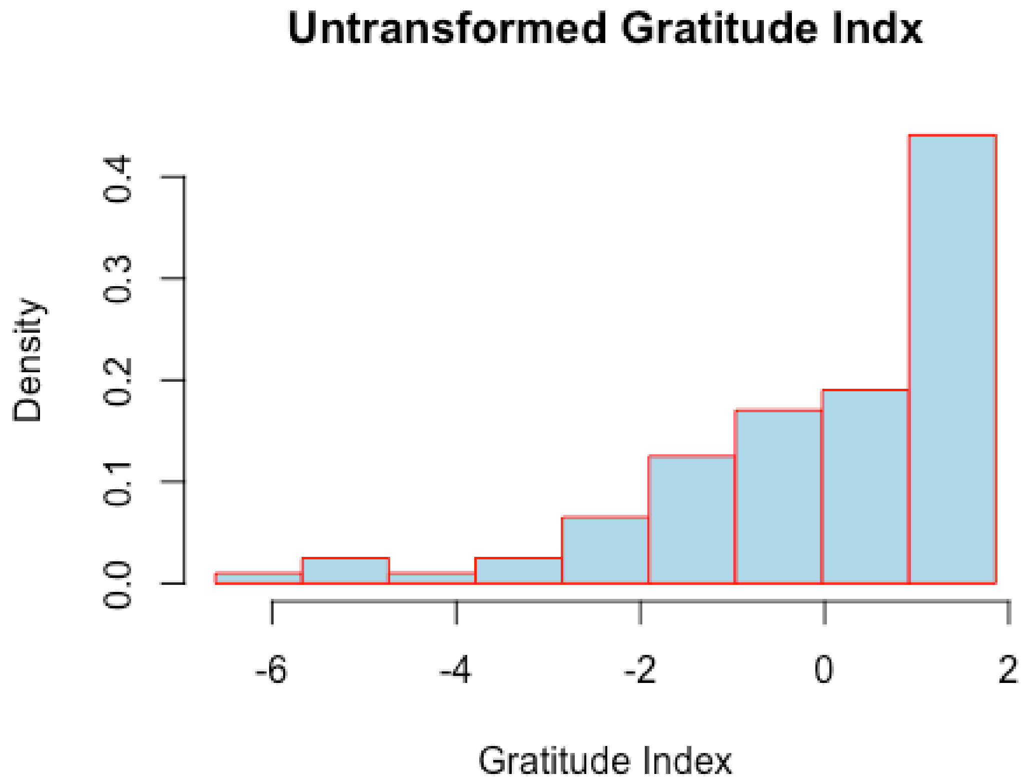
Q3_2	0.467	-0.035
Q3_3	0.375	0.538
Q3_4	0.374	-0.489
Q3_5	0.416	-0.418
Q3_6	0.354	0.542

Cumulative and proportion of variances explained by each principle component from the Gratitude Questionaire Survey.

	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	1.763	0.920	0.797	0.781	0.665	0.599
Proportion of Variance	0.518	0.141	0.106	0.102	0.074	0.060
Cumulative Proportion	0.518	0.659	0.765	0.866	0.940	1.000

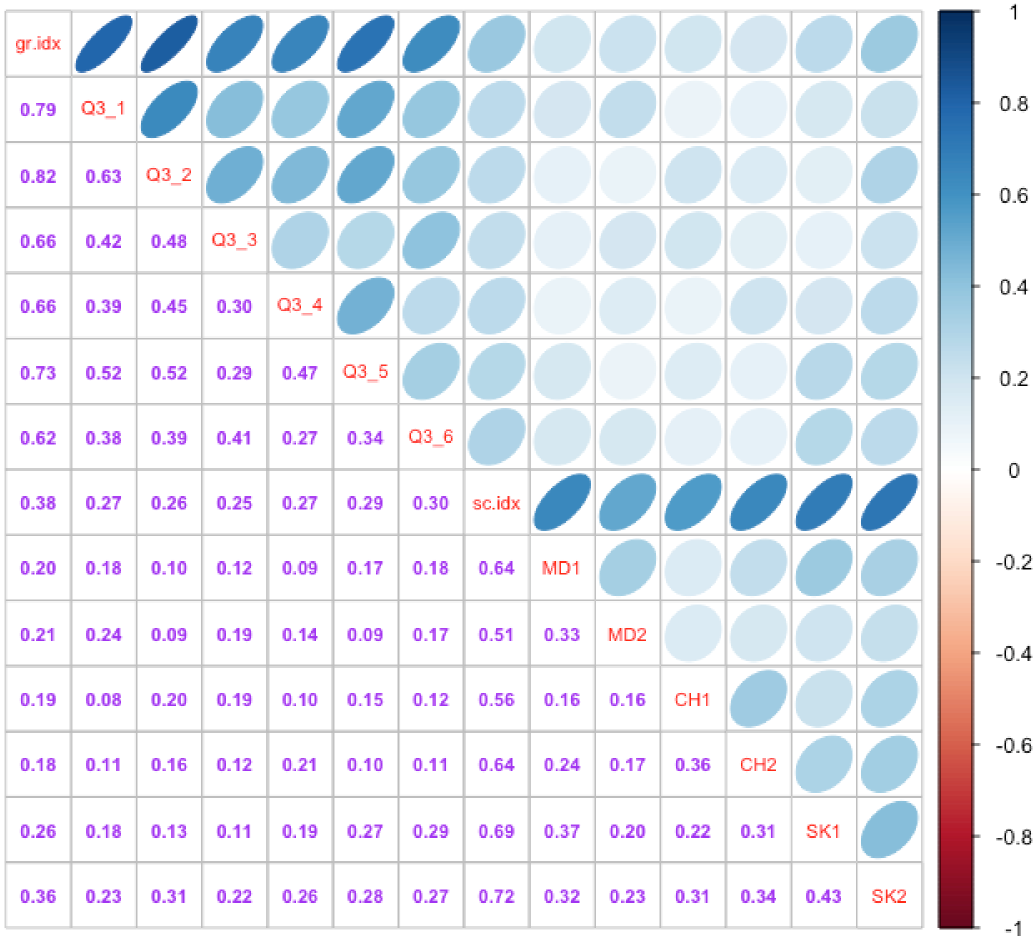
A similar analysis was also conducted using EFA. The result indicates that the first principle factors is sufficient for exploratory analysis. The proportion of the total variation explained by the first factor is about 40% which is about 10% less than that in the first principle component. We will use the first PCA and call it as gratitude index.

Next we extract the first PCA scores.



Histograms of the gratitude index scores. The distribution of the gratitude index is skewed to the left. We will perform a Box-Cox transformation to fix the distributional issue in the regression residuals.

To define a meaningful index of self-compassion, we want to make sure that the proposed index is positively correlated to the individual item. Since the principle component analysis algorithm is essentially a orthogonal rotation (transformation), we can adjust the direction of the coordinate system to make the PCA scores meaningful index for subsequent association analysis. The is the plot of the pairwised association between the individual items and the two new PCA scores.

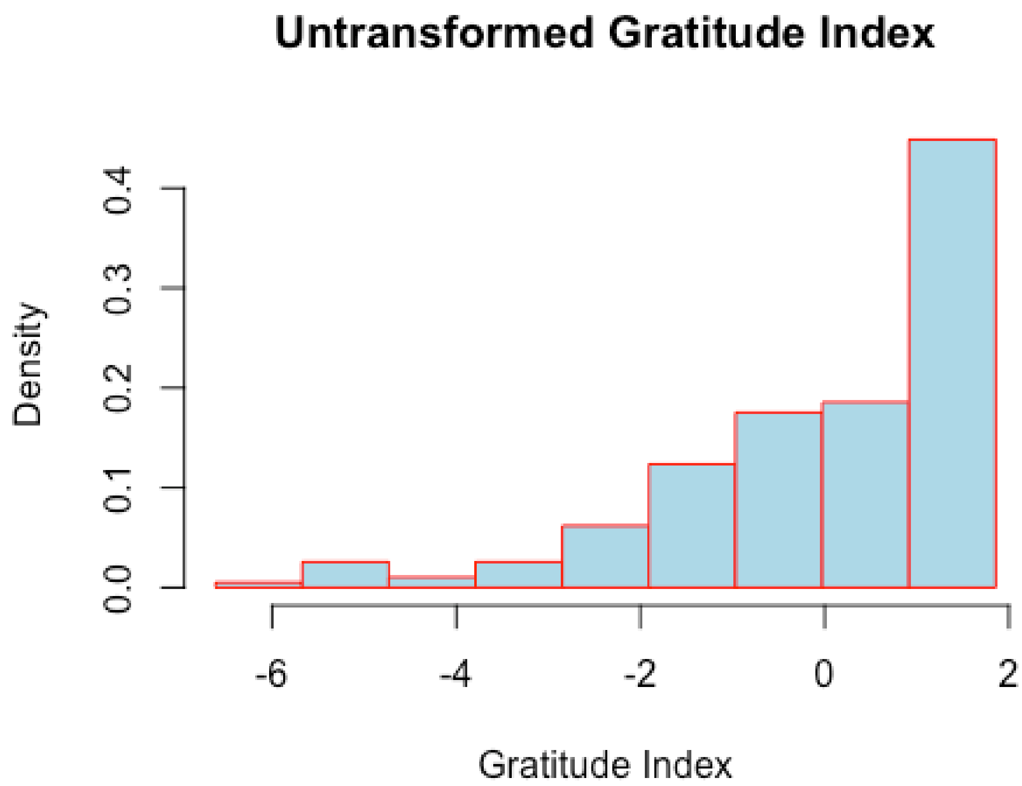


Pair-wise correlation plot between individual survey items and the two first PCA scores extract from the two instruments.

From the above plot that the first PCA of self-compassion is negatively associated with individual items in the instrument. This implies that the negative scores of the PCA are appropriate indexes for the self-compassion. This index based on the negative PCA will be added to the final data set in the next section.

Final Analytic Data Set

We now create a dataframe that contains PCAs based on self-compassion and gratitude indexes and the demographics.



Histograms of the Box-Cox transformed gratitude index scores.

In next section, we focus on building regression models to address the research questions.

Association Aanalyses

Three families of linear regression models will be built in this section. The final models in each of the three families will be presented to address the association between self-compassion and other variables of interests.

Association between Self-Compassion (raw Sum) and Demographics

The following R function reports the common goodness of fit measures to model selection.

Association of Between Self-compassion (mean item scores) and Demographics

The response variable will be the mean item scores.

Summary Table of Goodness of Fit Measures : Response = Mean

Predictor	Mallows_cp	R.square	SBIC	Adj.Rsq	RMSE	ACI	SBC
gr.idx	1.589	0.133	153.406	0.129	1.437	737.961	747.945
spirituality	-1.677	0.156	150.189	0.147	1.422	734.563	747.874
urbanity	-4.857	0.178	147.032	0.161	1.41	733.111	753.078
home.size	-7.445	0.181	142.285	0.16	1.396	719.452	742.644

Best Subset Selection

In this section, we build and search the best regression model to detect the potential association between the self-compassion index and the demographics. we use an automatic variable selection procedure to find the final model that contains significant variables. The subset selection methods allows us to select top predictors.

```
## [1] 6

##      (Intercept)      home.size3+ urbanitysuburban  spiritualitylow
##        0.2847187        0.4522247       -0.5437709       -0.5437250
##           gr.idx
##        0.2659892

##
## Call:
## lm(formula = sc.idx ~ home.size + urbanity + spirituality + gr.idx,
##     data = final.data.step)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8152 -0.7530  0.1739  0.9622  2.7134
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.28243    0.26764   1.055   0.2927
## home.size3+       0.45262    0.21003   2.155   0.0324 *
## urbanitysuburban -0.54151    0.26451  -2.047   0.0420 *
## urbanityurban     0.00430    0.31398   0.014   0.9891
## spiritualitylow  -0.54431    0.22157  -2.457   0.0149 *
## gr.idx            0.26599    0.06343   4.193 4.25e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.414 on 186 degrees of freedom
## Multiple R-squared:  0.1933, Adjusted R-squared:  0.1716
## F-statistic: 8.912 on 5 and 186 DF,  p-value: 1.322e-07
```

Conclusion: Both forward and backward selections yield the same results.

Best SubSet Selection

We use the following best subset selection to select top 10 strong variables to include in the final model.

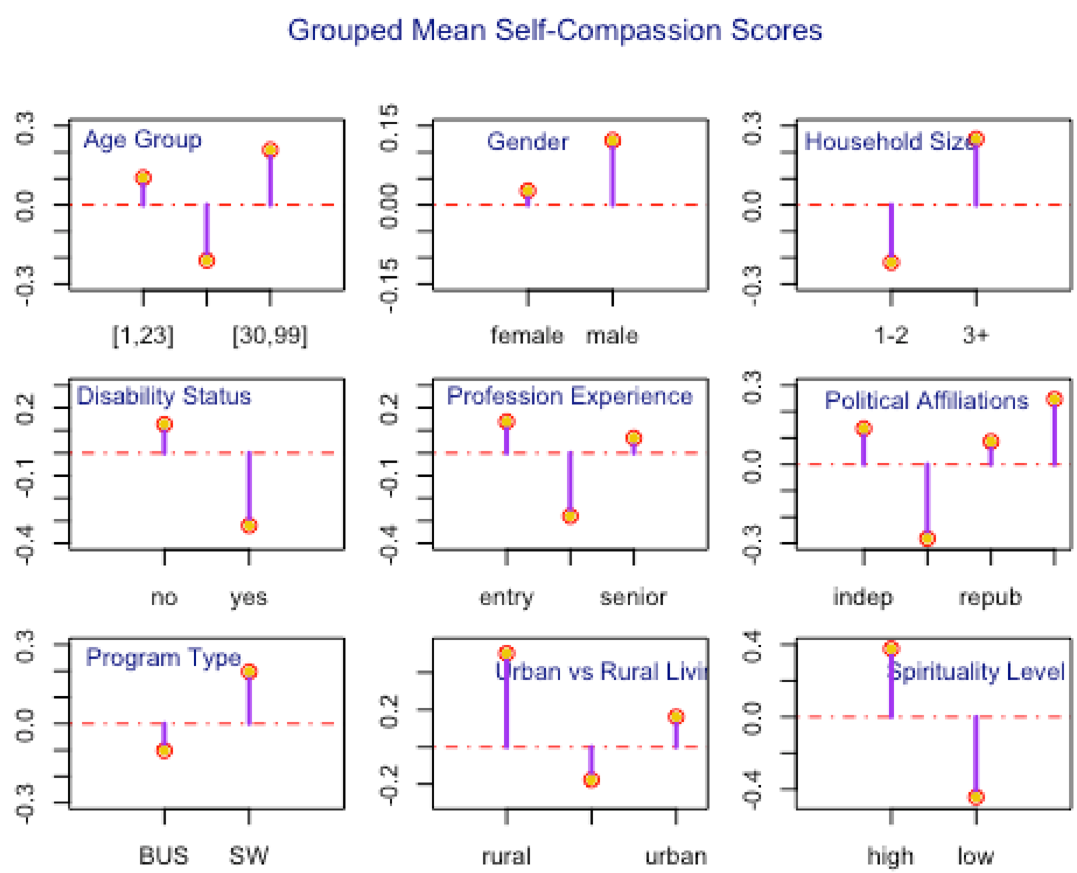
```
## [1] 6

##                (Intercept)          grp.eduBachelor
home.size3+
##                 -0.1742360               -0.1946215
0.4814853
##                 gendermale            disabilityyes
religionreligion
##                  0.2936649               -0.4269504
0.5538808
##         sexual.orientother poli.affilStrongDemocrats
urbanitysuburban
##                 -0.2168022                0.4973184                         -
0.4962881
##            spiritualitylow                   gr.idx
##                 -0.5836672                0.2820066
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.3929667	0.3789044	1.0371130	0.3011060
grp.age[24,30]	-0.2536078	0.2853333	-0.8888123	0.3753171
grp.age[30,99]	-0.0163567	0.3496067	-0.0467860	0.9627368
grp.empljunior	-0.2016477	0.2915303	-0.6916869	0.4900454
grp.emplsenior	-0.3351703	0.3609911	-0.9284725	0.3544341
home.size3+	0.3685568	0.2292640	1.6075648	0.1097230
gendermale	0.3359827	0.2757270	1.2185336	0.2246521
disabilityyes	-0.4895000	0.2777946	-1.7620936	0.0797895
poli.affilModDemocrats	-0.1774009	0.2886863	-0.6145110	0.5396713
poli.affilRepublican	-0.2024552	0.2883966	-0.7020028	0.4836047
poli.affilStrongDemocrats	0.3329456	0.3186990	1.0447023	0.2975933
programSW	0.2192850	0.2554585	0.8583975	0.3918405
urbanitysuburban	-0.4563993	0.2791656	-1.6348692	0.1038638
urbanityurban	-0.0153606	0.3397518	-0.0452112	0.9639902
spiritualitylow	-0.5501549	0.2294145	-2.3980822	0.0175269

gr.idx 0.2777384 0.0681646 4.0745252 0.0000697

```
library(scales)
par(mfrow = c(2,2))
#plot(forward.model, which=c(1,1), pch= 19, col = alpha("purple", 0.6))
plot(forward.model, pch= 19, col = alpha("navy", 0.6))
#points(forward.model, pch= 16, col = alpha("purple", 0.6))
```

Group means



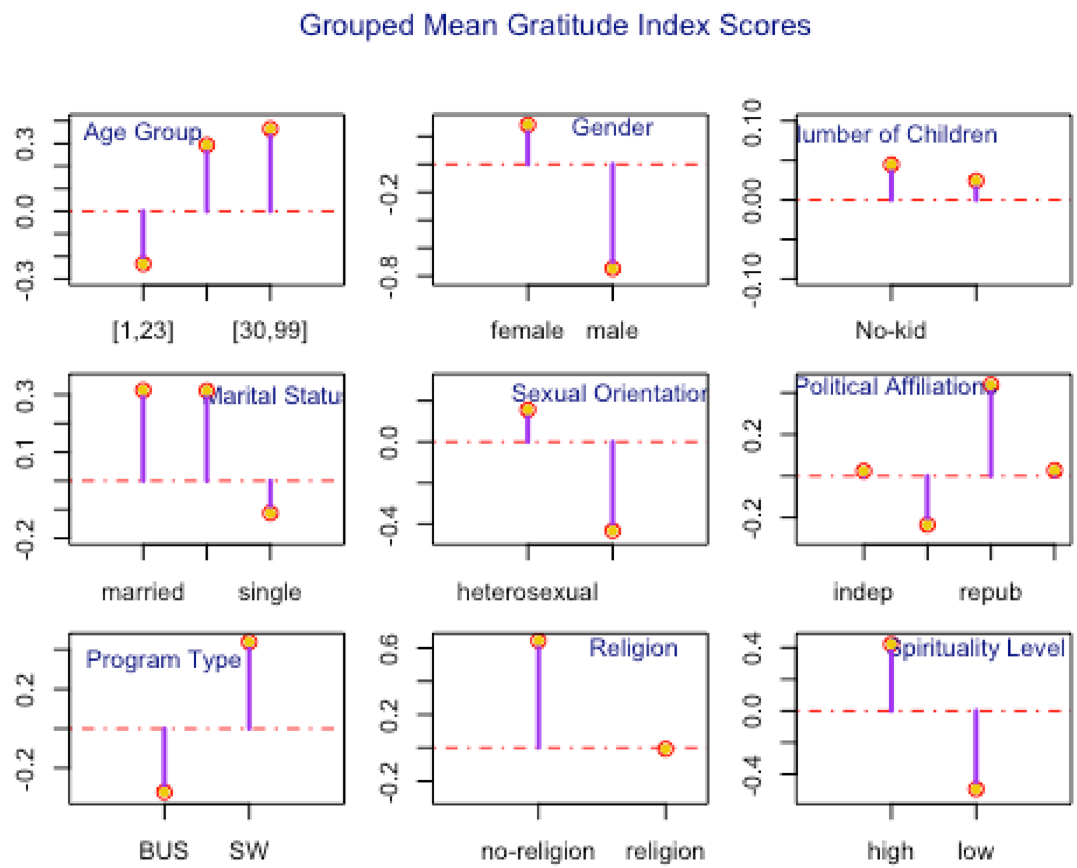
Gratitude Index versus Demographic Adjusted by Self-compassion Index

The linear regression model will be based on the index scores created based on the principle component analysis.

```
## [1] 6

##                  (Intercept)            grp.age[24,30]
kid.numWith-kid
##                    1.0940333                 0.3208304                       -
0.3272531
##                   gendermale             marital.stother
religionreligion
##                   -0.8516187                 0.4248014                       -
0.6353602
```

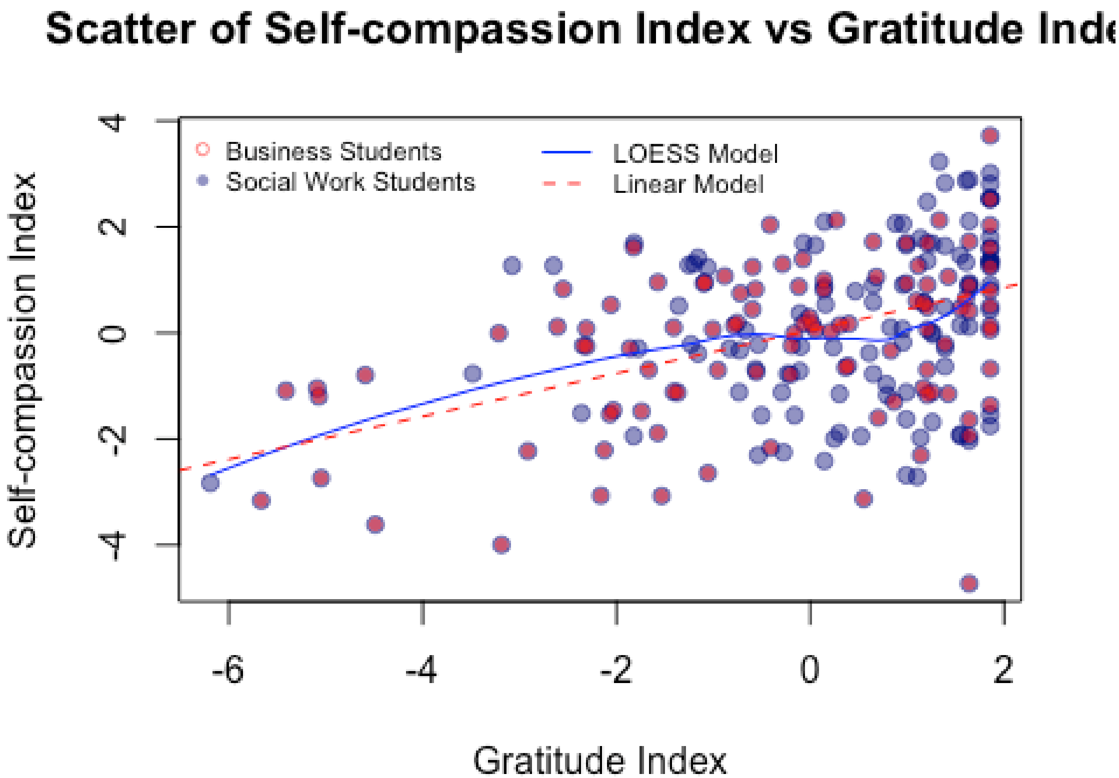
```
##        sexual.orientother    poli.affilModDemocrats
poli.affilStrongDemocrats
##                -0.3382331                -0.6713880                -
0.4689883
##                 programSW            spiritualitylow
sc.idx
##                 0.4229411                -0.6363173
0.3039110
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.8957641	0.5941073	1.5077481	0.1333819
grp.age[24,30]	0.4562476	0.2919431	1.5627961	0.1198668
grp.age[30,99]	0.3571904	0.3571109	1.0002226	0.3185530
kid.numWith-kid	-0.4155868	0.2876402	-1.4448148	0.1502582
gendermale	-0.9315584	0.2820473	-3.3028448	0.0011557
marital.stother	0.3549690	0.3679038	0.9648419	0.3359257
marital.stsingle	-0.0493897	0.3509846	-0.1407177	0.8882512
religionreligion	-0.6184788	0.4136720	-1.4950947	0.1366506
sexual.orientother	-0.2927963	0.2912073	-1.0054565	0.3160343
poli.affilModDemocrats	-0.5159720	0.2923029	-1.7651961	0.0792360
poli.affilRepublican	0.2965943	0.3022141	0.9814046	0.3277179
poli.affilStrongDemocrats	-0.4023794	0.3332566	-1.2074162	0.2288649
programSW	0.3587515	0.2571647	1.3950261	0.1647365
spiritualitylow	-0.5846721	0.2456986	-2.3796312	0.0183816
sc.idx	0.3385007	0.0727446	4.6532763	0.0000063




```
library(scales)
par(mfrow = c(2,2))
#plot(forward.model, which=c(1,1), pch= 19, col = alpha("purple", 0.6))
plot(final.gr.model, pch= 19, col = alpha("navy", 0.6))
#points(forward.model, pch= 16, col = alpha("purple", 0.6))
```

Relationship between Self-compassion and Gratitude



Three variables were retained in the final model based on the step-wise model selection. The diagnostic plots of the above linear model are given in the following.

This is a simple linear regression model.

The marginal association between self-compassion and gratitude index

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.1468	0.1558	-0.9426	0.3470
poly(gr.idx, 3, raw = TRUE)1	0.2406	0.0889	2.7069	0.0074
poly(gr.idx, 3, raw = TRUE)2	0.0927	0.0639	1.4511	0.1483
poly(gr.idx, 3, raw = TRUE)3	0.0218	0.0127	1.7082	0.0891

References

Raiche, G., Walls, T. A., Magis, D., Riopel, M. and Blais, J.-G. (2013). Non-graphical solutions for Cattell's scree test. *Methodology*, 9(1), 23-29.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS One*, 12(6), e0174035.

Franklin, Scott B., Gibson, David J., Robertson, Philip A., Pohlmann, John T. and Fralish, James S. (1995). Parallel Analysis: a Method for Determining Significant Principal Components, *Journal of Vegetation Science*, Vol. 6(1): 99-106

Kotera, Y & Sheffield, D. (2020). Revisiting the Self-compassion Scale-Short Form: Stronger Associations with Self-inadequacy and Resilience, *SN Comprehensive Clinical Medicine*, 2: 761-769.

López, A, Sanderman, R., Ranchor, A. V. & Schroevers, M. J. (2018). Compassion for Others and Self-Compassion: Levels, Correlates,and Relationship with Psychological Well-being. *Mindfulness*, 9: 325 - 331.