ICC\_kline

kaggle

2/29/2020

# Item Stats

|  |  |  |  |
| --- | --- | --- | --- |
| rowname | mean | sd | n |
| SCHB\_GM | 2.20 | 3.636 | 20 |
| SCHB\_GP | 2.65 | 3.031 | 20 |
| SCHB\_RP | 2.60 | 2.583 | 20 |
| SCHB\_AZ | 2.85 | 3.117 | 20 |

|  |  |  |  |
| --- | --- | --- | --- |
| rowname | mean | sd | n |
| PAE\_GM | 0.50 | 0.827 | 20 |
| PAE\_GP | 0.50 | 0.889 | 20 |
| PAE\_RP | 0.35 | 0.933 | 20 |
| PAE\_AZ | 1.25 | 1.293 | 20 |

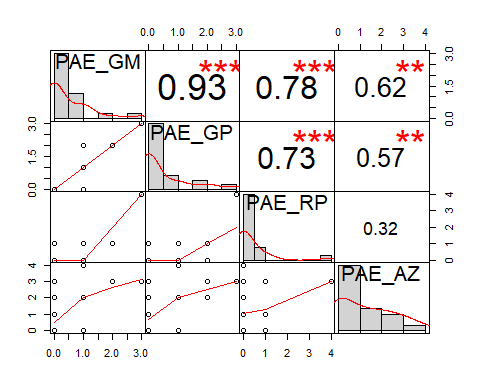
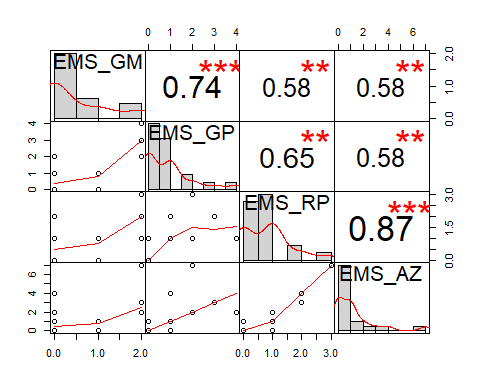
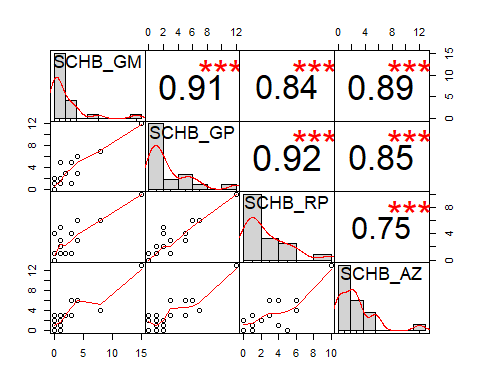
|  |  |  |  |
| --- | --- | --- | --- |
| rowname | mean | sd | n |
| EMS\_GM | 0.50 | 0.761 | 20 |
| EMS\_GP | 0.90 | 1.119 | 20 |
| EMS\_RP | 0.80 | 0.834 | 20 |
| EMS\_AZ | 1.25 | 1.743 | 20 |

# inter-Item correlation matrix

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | SCHB\_GM | SCHB\_GP | SCHB\_RP | SCHB\_AZ |
| SCHB\_GM | 1.000 | 0.909 | 0.838 | 0.885 |
| SCHB\_GP | 0.909 | 1.000 | 0.915 | 0.852 |
| SCHB\_RP | 0.838 | 0.915 | 1.000 | 0.750 |
| SCHB\_AZ | 0.885 | 0.852 | 0.750 | 1.000 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | EMS\_GM | EMS\_GP | EMS\_RP | EMS\_AZ |
| EMS\_GM | 1.000 | 0.742 | 0.581 | 0.575 |
| EMS\_GP | 0.742 | 1.000 | 0.654 | 0.580 |
| EMS\_RP | 0.581 | 0.654 | 1.000 | 0.869 |
| EMS\_AZ | 0.575 | 0.580 | 0.869 | 1.000 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | PAE\_GM | PAE\_GP | PAE\_RP | PAE\_AZ |
| PAE\_GM | 1.000 | 0.931 | 0.784 | 0.615 |
| PAE\_GP | 0.931 | 1.000 | 0.730 | 0.573 |
| PAE\_RP | 0.784 | 0.730 | 1.000 | 0.316 |
| PAE\_AZ | 0.615 | 0.573 | 0.316 | 1.000 |



# Means and correlations summary

## SCHB

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| info | Mean | Min | Max | Range | Max/Min | Variance | N of items |
| Item Means | 2.575 | 2.20 | 2.850 | 0.650 | 1.295 | 0.074 | 4 |
| Inter-Item Correlation | 0.858 | 0.75 | 0.915 | 0.165 | 1.220 | 0.004 | 6 |

## EMS

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| info | Mean | Min | Max | Range | Max/Min | Variance | N of items |
| Item Means | 0.862 | 0.500 | 1.250 | 0.750 | 2.500 | 0.096 | 4 |
| Inter-Item Correlation | 0.667 | 0.575 | 0.869 | 0.294 | 1.511 | 0.014 | 6 |

## PAE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| info | Mean | Min | Max | Range | Max/Min | Variance | N of items |
| Item Means | 0.650 | 0.350 | 1.250 | 0.900 | 3.571 | 0.165 | 4 |
| Inter-Item Correlation | 0.658 | 0.316 | 0.931 | 0.615 | 2.943 | 0.044 | 6 |

# ICC

## SCHB

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :  
## Model failed to converge with max|grad| = 0.003732 (tol = 0.002, component 1)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rowname | type | ICC | F | df1 | df2 | p | lower bound | upper bound |
| Single\_raters\_absolute | ICC1 | 0.845 | 22.839 | 19 | 60 | 0 | 0.727 | 0.927 |
| Single\_random\_raters | ICC2 | 0.845 | 22.839 | 19 | 57 | 0 | 0.727 | 0.927 |
| Single\_fixed\_raters | ICC3 | 0.845 | 22.839 | 19 | 57 | 0 | 0.725 | 0.927 |
| Average\_raters\_absolute | ICC1k | 0.956 | 22.839 | 19 | 60 | 0 | 0.914 | 0.981 |
| Average\_random\_raters | ICC2k | 0.956 | 22.839 | 19 | 57 | 0 | 0.914 | 0.981 |
| Average\_fixed\_raters | ICC3k | 0.956 | 22.839 | 19 | 57 | 0 | 0.913 | 0.981 |

## EMS

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rowname | type | ICC | F | df1 | df2 | p | lower bound | upper bound |
| Single\_raters\_absolute | ICC1 | 0.540 | 5.698 | 19 | 60 | 0 | 0.322 | 0.749 |
| Single\_random\_raters | ICC2 | 0.545 | 6.327 | 19 | 57 | 0 | 0.329 | 0.752 |
| Single\_fixed\_raters | ICC3 | 0.571 | 6.327 | 19 | 57 | 0 | 0.355 | 0.770 |
| Average\_raters\_absolute | ICC1k | 0.825 | 5.698 | 19 | 60 | 0 | 0.655 | 0.923 |
| Average\_random\_raters | ICC2k | 0.828 | 6.327 | 19 | 57 | 0 | 0.663 | 0.924 |
| Average\_fixed\_raters | ICC3k | 0.842 | 6.327 | 19 | 57 | 0 | 0.688 | 0.931 |

## PAE

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rowname | type | ICC | F | df1 | df2 | p | lower bound | upper bound |
| Single\_raters\_absolute | ICC1 | 0.506 | 5.092 | 19 | 60 | 0 | 0.285 | 0.725 |
| Single\_random\_raters | ICC2 | 0.521 | 6.910 | 19 | 57 | 0 | 0.285 | 0.740 |
| Single\_fixed\_raters | ICC3 | 0.596 | 6.910 | 19 | 57 | 0 | 0.384 | 0.786 |
| Average\_raters\_absolute | ICC1k | 0.804 | 5.092 | 19 | 60 | 0 | 0.614 | 0.913 |
| Average\_random\_raters | ICC2k | 0.813 | 6.910 | 19 | 57 | 0 | 0.614 | 0.919 |
| Average\_fixed\_raters | ICC3k | 0.855 | 6.910 | 19 | 57 | 0 | 0.714 | 0.936 |
| Shrout and Fleiss (1979) c | onsider | six case | s of rel | iabili | ty of | ratin | gs done by k r | aters on n targets. |

ICC1: Each target is rated by a different judge and the judges are selected at random. (This is a one-way ANOVA fixed effects model and is found by (MSB- MSW)/(MSB+ (nr-1)\*MSW))

ICC2: A random sample of k judges rate each target. The measure is one of absolute agreement in the ratings. Found as (MSB- MSE)/(MSB + (nr-1)*MSE + nr*(MSJ-MSE)/nc)

ICC3: A fixed set of k judges rate each target. There is no generalization to a larger population of judges. (MSB - MSE)/(MSB+ (nr-1)\*MSE)

Then, for each of these cases, is reliability to be estimated for a single rating or for the average of k ratings? (The 1 rating case is equivalent to the average intercorrelation, the k rating case to the Spearman Brown adjusted reliability.)

ICC1 is sensitive to differences in means between raters and is a measure of absolute agreement.

ICC2 and ICC3 remove mean differences between judges, but are sensitive to interactions of raters by judges. The difference between ICC2 and ICC3 is whether raters are seen as fixed or random effects.

ICC1k, ICC2k, ICC3K reflect the means of k raters.

The intraclass correlation is used if raters are all of the same “class”. That is, there is no logical way of distinguishing them. Examples include correlations between pairs of twins, correlations between raters. If the variables are logically distinguishable (e.g., different items on a test), then the more typical coefficient is based upon the inter-class correlation (e.g., a Pearson r) and a statistic such as alpha or omega might be used. alpha and ICC3k are identical.

If using the lmer option, then missing data are allowed. In addition the lme object returns the variance decomposition. (This is simliar to testRetest which works on the items from two occasions.

The check.keys option by default reverses items that are negatively correlated with total score. A message is issued.

#### Reference

Shrout, Patrick E. and Fleiss, Joseph L. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin, 1979, 86, 420-3428.