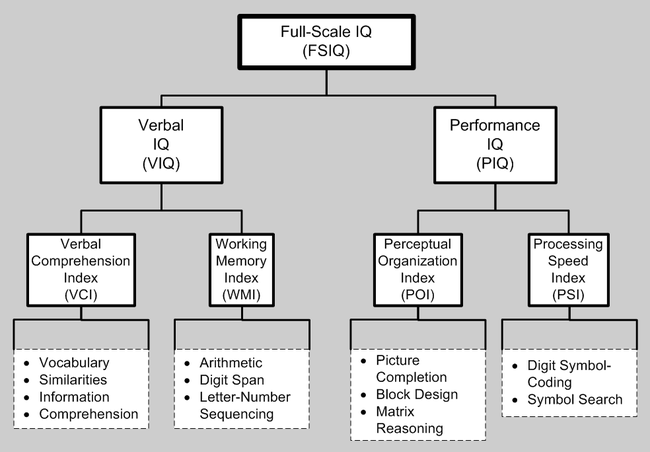
CFA Assignment



Assignment steps:

1. Data screening:
   1. You will have a missing variable, and you should program the model without it.
   2. Assume that the data is accurate.
   3. Missing values:
      1. Include output that shows that there are/not missing values.
      2. Impute the missing values with *mice* if necessary.
   4. Outliers:
      1. Screen the data for outliers.
      2. How many outliers did you have? Include a summary.
      3. Exclude any multivariate outliers.
   5. Multicollinearity
      1. Are any of the variables too correlated?
      2. Run a correlation table (do not need to include).
      3. Use the symnum function to look at the data to see if any of the correlations are too high.
   6. Normality
      1. Include the multivariate histogram.
      2. Is the data normal?
   7. Linearity
      1. Include the QQ plot for linearity.
      2. Is the data linear?
   8. Homogeneity and Homoscedasticity
      1. Include a residual scatterplot of the data.
      2. Is the data homogeneic?
      3. Is the data homoscedastic?
2. Regular CFA – does the factor structure for the WAIS hold up?
   1. Test the following models
      1. Program 4-factor model – Use VCI, WMI, POI, and PSI as latent variables.
      2. Program 2-factor model – with VIQ and PIQ as latent variables (collapse across the indices for those latents)
      3. Program 1-factor model – with IQ as the only latent variable.
      4. NOTE: You will have issues with this set of CFAs – try examining the correlations and collapsing very large correlations.
   2. Be sure to include the following:
      1. Include fit indices.
      2. Difference test against the proposed 4-factor model (use AIC or CFI but be explicit on your reasoning/explanation).
      3. Which model is best?
   3. For the BEST MODEL ONLY:
      1. Check the standardized residuals and see if there are any questions that are not being measured well (i.e. residual Z-score > 2.58).
         1. Use residuals(model.fit, type = "standardized") to get those values.
         2. Report those questions and values.
      2. Should any of the questions be moved? Check the modification indices and report if you moved or double loaded questions.
3. Second order CFA – on a multiple factor model only (if the best model is the one factor, you will need to program the second order on something with multiple factors)
   1. Program a general IQ factor leading to your subscale factors.
   2. List model fit.
   3. Does the data support a second order model?
4. Bi-Factor CFA – on a multiple factor model only (if the best model is the one factor, you will need to program the second order on something with multiple factors)
   1. Program a general IQ factor that is a generalized factor. Include the sub-domains as domain specific factors.
   2. List model fit.
   3. Does the mode support the domain specific areas, after accounting for the generalized latent variable?
   4. NOTE: To fix Heywood cases here: calculate the variance for the items that are giving Heywood problems. Fix those variances to the variance in the dataset one at a time until the negative variances go away.
5. Full SEM – program the following model:
   1. Use modification indices to make a better model (that still makes sense!)
   2. List all fit indices and model differences (CFI/AIC) as you change paths. Only one path at a time please!
   3. Are there any non-significant paths?
   4. How much variance is accounted for by each of the latents (use SMCs)?
6. Be sure to include:
   1. Write up of all the models.
   2. Pictures of all models with standardized estimates.
   3. Fit indices of all models.