PSY.308c.DA3

Daniel Pinedo

**Prompt** You now work for a nationwide transitional living program. They serve homeless individuals between the age of 16-25. The various services they are providing include: (1) a service that helps youth get a job (2) a literacy program (3) a program to help youth graduate from high school and (4) shelter to live in. They’ve collected data on the youth in their program on income, illiteracy, high school graduation, and safety levels (see below for an explanation of the variables). The program ultimately cares about what best predicts the successful transition of these youth. According to the literature, high school graduation and safety levels explain successful transitioning best. The CEO thinks that from her experience income is actually an important factor to consider as well.

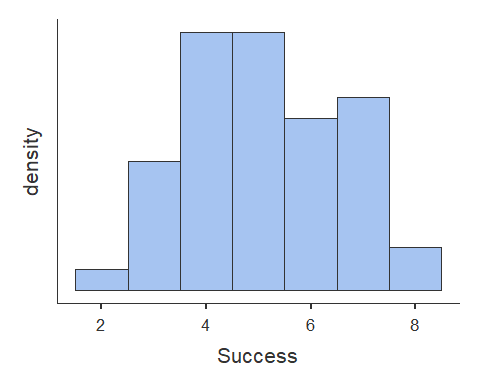
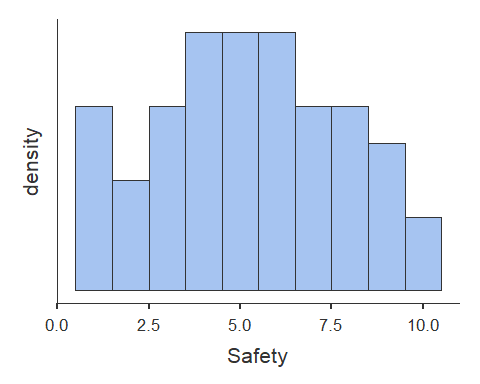
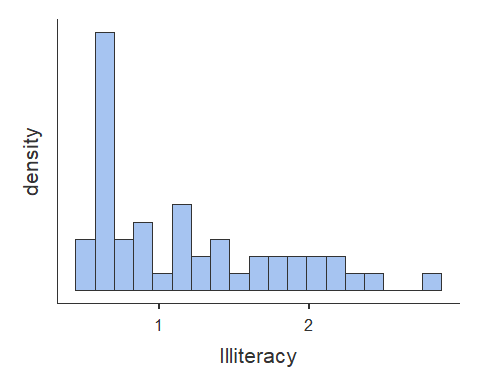
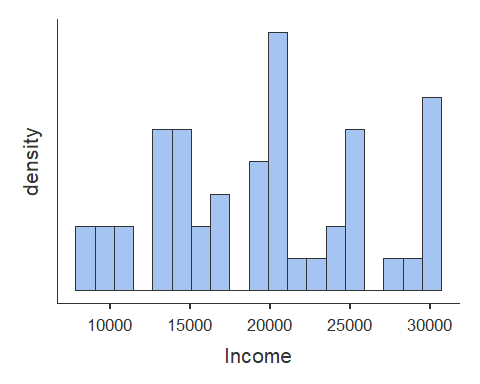
**Variables** Income: annual income Illiteracy: level of illiteracy on a scale of 0-3 (higher being more illiterate) Safety: safety levels in the area in which they live on a scale of 0-10 HS.Grad: whether they graduated on time (18 yrs old; D2), later (any age after 18; D1), or did not graduate at all (intercept / B0) Success (outcome): successful transition scaled based on a variety of factors compiling to an ultimate score between the values of 0-10

**Hypotheses** H0: no relationship between variables Ha: income + illiteracy + safety + graduation status predict success N = 50

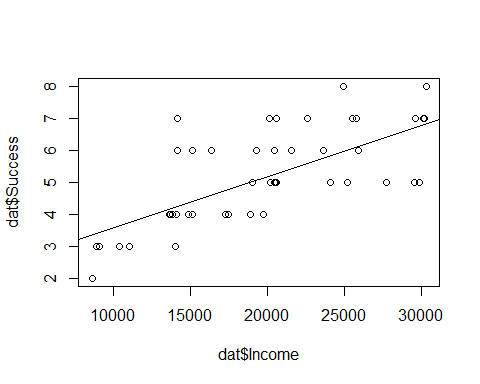
**Descriptive Statistics and Assumptions**

# Prerequisitites  
 # 1. Variables are measured on the continuous level  
  
# Assumptions  
 # 1. Normal Distribution for X and Y (Product) [i.e. histogram, skew +-3, kurtosis +-10]  
 # Histogram for Income appears normal  
 # Histogram for Illitaracy appears unimodal and skewed positively  
 # Histogram for Safety appears normal  
 # Histogram for Success appears normal  
 # Skewness - ALL PASS  
 # Kurtosis - ALL PASS   
   
 # 2. Linear Relationship beween X and Y  
 # Visual inspection of scatterplot and prediction model line indicate a linear relationship  
 # 3. Homoscedasticity  
 # a. Visual inspection of scatterplots indicate:   
 # possible lower variance at lower end of Income  
 # possible lower variance at upper end of Illiteracy  
 # likely equal variance across Safety  
 # b. non-constant variance test - H0 = TRUE (PASS)  
   
 # 4. [Examine residuals (e = Y - Y~predicted~) to understand 2 and 3 mathematically]  
  
# Descriptives [Assumption 1]  
desc <- descriptives(data = dat,   
 vars = c('Income', 'Illiteracy', 'Safety', 'Success'),   
 hist = TRUE,   
 sd = TRUE,   
 range = TRUE,   
 skew = TRUE,   
 kurt = TRUE)  
desc

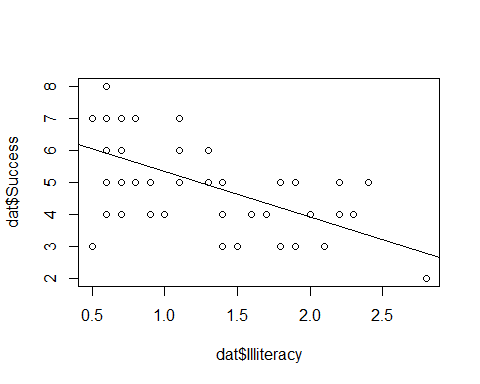
##   
## DESCRIPTIVES  
##   
## Descriptives   
## --------------------------------------------------------------------   
## Income Illiteracy Safety Success   
## --------------------------------------------------------------------   
## N 50 50 50 50   
## Missing 0 0 0 0   
## Mean 19483 1.17 5.24 5.10   
## Median 19938 0.950 5.00 5.00   
## Standard deviation 6432 0.610 2.54 1.47   
## Range 21681 2.30 9 6   
## Minimum 8603 0.500 1 2   
## Maximum 30284 2.80 10 8   
## Skewness 0.144 0.870 0.0142 0.0995   
## Std. error skewness 0.337 0.337 0.337 0.337   
## Kurtosis -0.947 -0.276 -0.849 -0.797   
## Std. error kurtosis 0.662 0.662 0.662 0.662   
## --------------------------------------------------------------------



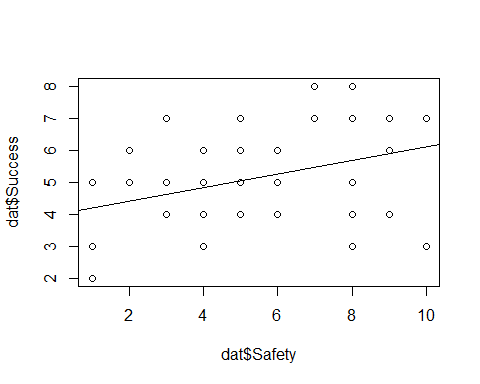
# Scatterplots [Assumption 2 and 3a]  
plot(dat$Income, dat$Success, abline(lm(dat$Success ~ dat$Income)))



plot(dat$Illiteracy, dat$Success, abline(lm(dat$Success ~ dat$Illiteracy)))



plot(dat$Safety, dat$Success, abline(lm(dat$Success ~ dat$Safety)))



# Homoscedasticity [Assumption 3b]  
  
# non-constant variance Chi-squared test [Chi-squared (df) = ##.##, p = .###]  
# H0 = homoscedastic - TRUE  
# Ha = heteroscedastic  
  
ncvTest(lm(Success ~ Income + Illiteracy + Safety, data = dat))

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 3.103699, Df = 1, p = 0.078115

**Correlations**

# Correlation  
cortable <- corrMatrix(data = dat,   
 vars = c('Income', 'Illiteracy', 'Safety', 'Success'),   
 flag = TRUE)  
cortable

##   
## CORRELATION MATRIX  
##   
## Correlation Matrix   
## --------------------------------------------------------------------------   
## Income Illiteracy Safety Success   
## --------------------------------------------------------------------------   
## Income Pearson's r  -0.415 0.196 0.699   
## p-value  0.003 0.172 < .001   
##   
## Illiteracy Pearson's r  -0.691 -0.589   
## p-value  < .001 < .001   
##   
## Safety Pearson's r  0.363   
## p-value  0.009   
##   
## Success Pearson's r    
## p-value    
## --------------------------------------------------------------------------   
## Note. \* p < .05, \*\* p < .01, \*\*\* p < .001

**Center the continuous predictor variables**

# c = x - M  
# Centering only quantitatively changes the intercept for regression equation  
# Center Income, Illiteracy, Safety  
dat$Income.c <- dat$Income - mean(dat$Income)  
dat$Illiteracy.c <- dat$Illiteracy - mean(dat$Illiteracy)  
dat$Safety.c <- dat$Safety - mean(dat$Safety)

**Simple Regression of centered continuous predictor variables**

# Simple regression  
# R = correlation between observed scores and predicted scores  
# R squared = percentage of variance explained  
# t = Estimate / SE  
# df = N - k - 1 [k is number of predictors]  
# H0: B0 = 0; H0; R squared = 0  
  
model1 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c'),  
 blocks = list('Income.c'),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
model1

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.699 0.489 45.9 1 48 < .001   
## -----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 5.10 0.151 4.80 5.40 33.87 < .001   
## Income.c 1.60e-4 2.37e-5 1.13e-4 2.08e-4 6.78 < .001 0.699   
## ------------------------------------------------------------------------------------------------

model2 <- linReg(data = dat,   
 dep = 'Success',  
 covs = c('Illiteracy.c'),  
 blocks = list('Illiteracy.c'),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
model2

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.589 0.347 25.5 1 48 < .001   
## -----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ----------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ----------------------------------------------------------------------------------------------   
## Intercept 5.10 0.170 4.76 5.442 29.97 < .001   
## Illiteracy.c -1.43 0.282 -1.99 -0.858 -5.05 < .001 -0.589   
## ----------------------------------------------------------------------------------------------

model3 <- linReg(data = dat,   
 dep = 'Success',  
 covs = c('Safety.c'),  
 blocks = list('Safety.c'),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
model3

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## ----------------------------------------------------------   
## Model R R² F df1 df2 p   
## ----------------------------------------------------------   
## 1 0.363 0.132 7.31 1 48 0.009   
## ----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## --------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## --------------------------------------------------------------------------------------------   
## Intercept 5.100 0.1962 4.7054 5.495 25.99 < .001   
## Safety.c 0.211 0.0779 0.0540 0.367 2.70 0.009 0.363   
## --------------------------------------------------------------------------------------------

**Multiple regression with dummy codes for Categorical Variable (Graduation Status [3 levels])**

# Model comparison  
# D1 is predicted difference between D1 (Graduated later) and reference group (Did not graduate) for a 1 unit change in Y (Success)  
# D2 is predicted difference between D2 (Graduated normal) and reference group (did not graduate) for 1 unit change in Y (Success)  
model4 <- linReg(data = dat,   
 dep = 'Success', #outcome  
 covs = c('D1', 'D2'), #predictors  
 blocks = list(c('D1', 'D2')), #order matters here if separate blocks of variables are provided  
 modelTest = TRUE,   
 stdEst = TRUE,   
 ciStdEst = TRUE,   
 r2Adj = TRUE)  
model4

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -------------------------------------------------------------------------   
## Model R R² Adjusted R² F df1 df2 p   
## -------------------------------------------------------------------------   
## 1 0.484 0.234 0.201 7.18 2 47 0.002   
## -------------------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## -------------------------------------------------------------------------------------------   
## Predictor Estimate SE t p Stand. Estimate Lower Upper   
## -------------------------------------------------------------------------------------------   
## Intercept 4.07 0.340 11.96 < .001   
## D1 1.23 0.467 2.63 0.012 0.398 0.0937 0.703   
## D2 1.71 0.461 3.72 < .001 0.563 0.2580 0.868   
## -------------------------------------------------------------------------------------------

Model 1 is best fit for simple regression Income predicts 49% of variance for Success

**Model 1 Comparison with Illiteracy added**

# Model comparison  
# H0 = delta of R squared = 0  
compare5 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c', 'Illiteracy.c'),  
 blocks = list(  
 list('Income.c'),  
 list('Illiteracy.c')),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
compare5

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.699 0.489 45.9 1 48 < .001   
## 2 0.773 0.597 34.8 2 47 < .001   
## -----------------------------------------------------------   
##   
##   
## Model Comparisons   
## ----------------------------------------------------------------   
## Model Model <U+0394>R² F df1 df2 p   
## ----------------------------------------------------------------   
## 1 - 2 0.108 12.6 1 47 < .001   
## ----------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 5.10 0.151 4.80 5.40 33.87 < .001   
## Income.c 1.60e-4 2.37e-5 1.13e-4 2.08e-4 6.78 < .001 0.699   
## ------------------------------------------------------------------------------------------------   
##   
##   
## MODEL 2  
##   
## Model Coefficients   
## ---------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ---------------------------------------------------------------------------------------------------   
## Intercept 5.100 0.135 4.83 5.372 37.74 < .001   
## Income.c 1.26e-4 2.33e-5 7.90e-5 1.73e-4 5.40 < .001 0.549   
## Illiteracy.c -0.874 0.246 -1.37 -0.379 -3.55 < .001 -0.361   
## ---------------------------------------------------------------------------------------------------

Model 5 is a good fit for multiple regression Income and Illiteracy predict 60% of variance for Success

**Model 1 Comparison with Safety added**

# Model comparison  
# H0 = delta of R squared = 0  
compare6 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c', 'Safety.c'),  
 blocks = list(  
 list('Income.c'),  
 list('Safety.c')),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
compare6

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.699 0.489 45.9 1 48 < .001   
## 2 0.736 0.542 27.8 2 47 < .001   
## -----------------------------------------------------------   
##   
##   
## Model Comparisons   
## ----------------------------------------------------------------   
## Model Model <U+0394>R² F df1 df2 p   
## ----------------------------------------------------------------   
## 1 - 2 0.0533 5.47 1 47 0.024   
## ----------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 5.10 0.151 4.80 5.40 33.87 < .001   
## Income.c 1.60e-4 2.37e-5 1.13e-4 2.08e-4 6.78 < .001 0.699   
## ------------------------------------------------------------------------------------------------   
##   
##   
## MODEL 2  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 5.100 0.1440 4.8103 5.390 35.41 < .001   
## Income.c 1.50e-4 2.31e-5 1.03e-4 1.96e-4 6.49 < .001 0.653   
## Safety.c 0.136 0.0583 0.0191 0.254 2.34 0.024 0.235   
## ------------------------------------------------------------------------------------------------

Model 6 is not best fit for multiple regression Income and Safety predict 54% of variance for Success

**Model 1 Comparison with Graduation added**

# Model comparison  
# H0 = delta of R squared = 0  
# D1 is predicted difference between D1 (Graduated later) and reference group (Did not graduate) for a 1 unit change in Y (Success)  
# D2 is predicted difference between D2 (Graduated normal) and reference group (did not graduate) for 1 unit change in Y (Success)  
compare7 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c', 'D1', 'D2'),  
 blocks = list(  
 list('Income.c'),  
 list('D1', 'D2')),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
compare7

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.699 0.489 45.9 1 48 < .001   
## 2 0.753 0.567 20.0 3 46 < .001   
## -----------------------------------------------------------   
##   
##   
## Model Comparisons   
## ----------------------------------------------------------------   
## Model Model <U+0394>R² F df1 df2 p   
## ----------------------------------------------------------------   
## 1 - 2 0.0775 4.11 2 46 0.023   
## ----------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 5.10 0.151 4.80 5.40 33.87 < .001   
## Income.c 1.60e-4 2.37e-5 1.13e-4 2.08e-4 6.78 < .001 0.699   
## ------------------------------------------------------------------------------------------------   
##   
##   
## MODEL 2  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------------   
## Intercept 4.568 0.272 4.020 5.12 16.79 < .001   
## Income.c 1.42e-4 2.39e-5 9.39e-5 1.90e-4 5.94 < .001 0.619   
## D1 0.465 0.377 -0.294 1.22 1.23 0.224 0.151   
## D2 1.038 0.368 0.297 1.78 2.82 0.007 0.341   
## ------------------------------------------------------------------------------------------------

Model 7 is not best fit for multiple regression Income and Graduation predict 57% of variance for Success

Model 5 is most parsimonious fit for multiple regression Income and Illiteracy predict 60% of variance for Success

**Model 5 Comparison with Safety added**

# Model comparison  
# H0 = delta of R squared = 0  
compare8 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c', 'Illiteracy.c', 'Safety.c'),  
 blocks = list(  
 list('Income.c', 'Illiteracy.c'),  
 list('Safety.c')),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
compare8

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.773 0.597 34.8 2 47 < .001   
## 2 0.773 0.597 22.7 3 46 < .001   
## -----------------------------------------------------------   
##   
##   
## Model Comparisons   
## --------------------------------------------------------------------   
## Model Model <U+0394>R² F df1 df2 p   
## --------------------------------------------------------------------   
## 1 - 2 6.81e-5 0.00778 1 46 0.930   
## --------------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ---------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ---------------------------------------------------------------------------------------------------   
## Intercept 5.100 0.135 4.83 5.372 37.74 < .001   
## Income.c 1.26e-4 2.33e-5 7.90e-5 1.73e-4 5.40 < .001 0.549   
## Illiteracy.c -0.874 0.246 -1.37 -0.379 -3.55 < .001 -0.361   
## ---------------------------------------------------------------------------------------------------   
##   
##   
## MODEL 2  
##   
## Model Coefficients   
## -----------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## -----------------------------------------------------------------------------------------------------   
## Intercept 5.10000 0.1366 4.825 5.375 37.3418 < .001   
## Income.c 1.26e-4 2.38e-5 7.83e-5 1.74e-4 5.3014 < .001 0.5506   
## Illiteracy.c -0.85351 0.3410 -1.540 -0.167 -2.5031 0.016 -0.3529   
## Safety.c 0.00669 0.0758 -0.146 0.159 0.0882 0.930 0.0115   
## -----------------------------------------------------------------------------------------------------

Model 8 is not a parsimonious fit for multiple regression Income, Illiteracy, and Safety predict 60% of variance for Success (no added account for variance)

**Model 5 Comparison with Graduation added**

# Model comparison  
# H0 = delta of R squared = 0  
# D1 is predicted difference between D1 (Graduated later) and reference group (Did not graduate) for a 1 unit change in Y (Success)  
# D2 is predicted difference between D2 (Graduated normal) and reference group (did not graduate) for 1 unit change in Y (Success)  
compare9 <- linReg(data = dat,   
 dep = 'Success',   
 covs = c('Income.c', 'Illiteracy.c', 'D1', 'D2'),  
 blocks = list(  
 list('Income.c', 'Illiteracy.c'),  
 list('D1', 'D2')),   
 modelTest = TRUE,   
 stdEst = TRUE,   
 ci = TRUE)  
compare9

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.773 0.597 34.8 2 47 < .001   
## 2 0.785 0.616 18.0 4 45 < .001   
## -----------------------------------------------------------   
##   
##   
## Model Comparisons   
## ----------------------------------------------------------------   
## Model Model <U+0394>R² F df1 df2 p   
## ----------------------------------------------------------------   
## 1 - 2 0.0185 1.08 2 45 0.347   
## ----------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ---------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ---------------------------------------------------------------------------------------------------   
## Intercept 5.100 0.135 4.83 5.372 37.74 < .001   
## Income.c 1.26e-4 2.33e-5 7.90e-5 1.73e-4 5.40 < .001 0.549   
## Illiteracy.c -0.874 0.246 -1.37 -0.379 -3.55 < .001 -0.361   
## ---------------------------------------------------------------------------------------------------   
##   
##   
## MODEL 2  
##   
## Model Coefficients   
## ----------------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ----------------------------------------------------------------------------------------------------   
## Intercept 4.9715 0.309 4.349 5.594 16.095 < .001   
## Income.c 1.27e-4 2.36e-5 8.00e-5 1.75e-4 5.410 < .001 0.5559   
## Illiteracy.c -0.7456 0.311 -1.372 -0.119 -2.398 0.021 -0.3083   
## D1 -0.0557 0.420 -0.901 0.790 -0.133 0.895 -0.0181   
## D2 0.4097 0.438 -0.472 1.291 0.936 0.354 0.1347   
## ----------------------------------------------------------------------------------------------------

Model 9 is not a parsimonious fit for multiple regression Income, Illiteracy, and Graduation predict 62% of variance for Success (no significant added account for prior predicted variance of 60%)

Based on prior literature, Graduation and Safety are best predictors of success. In this case, neither graduation nor safety accounted for a significantly greater amount of variance when added to Income and Illiteracy, Income accounted for highest amount of overall variance by itself, and Income and Illiteracy accounted for the most parsimonious model overall.

Thus, Model 5 is best, most parsimonious fit for multiple regression Income and Illiteracy predict 60% of variance for Success

**Transform Normalized Illiteracy to Literacy on a scale of 0-3 (higher being more literate)**

dat$Literacy.t <- 3 - dat$Illiteracy.c

**Model 5 with normalized Literacy transform**

# Multiple regression [Success ~ Income.c + Literacy.t]  
# Y = B0 + B1\*Income + B2\*Literacy + residuals [B0 = 2.48, B1 = 12,600, B2 = 0.87]  
# Accounting for error (Sum of Y - Y predicted / N - standard error in gray below):   
 #with average income and literacy, Y is 2.48 {low success}  
  
transform5 <- linReg(data = dat,   
 dep = 'Success', #outcome  
 covs = c('Income.c', 'Literacy.t'), #predictors  
 blocks = list(c('Income.c', 'Literacy.t')), #order matters here if separate blocks of variables are provided  
 modelTest = TRUE,   
 stdEst = TRUE,   
 ciStdEst = TRUE,   
 r2Adj = TRUE)  
transform5

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## --------------------------------------------------------------------------   
## Model R R² Adjusted R² F df1 df2 p   
## --------------------------------------------------------------------------   
## 1 0.773 0.597 0.580 34.8 2 47 < .001   
## --------------------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## --------------------------------------------------------------------------------------------   
## Predictor Estimate SE t p Stand. Estimate Lower Upper   
## --------------------------------------------------------------------------------------------   
## Intercept 2.478 0.751 3.30 0.002   
## Income.c 1.26e-4 2.33e-5 5.40 < .001 0.549 0.345 0.754   
## Literacy.t 0.874 0.246 3.55 < .001 0.361 0.157 0.566   
## --------------------------------------------------------------------------------------------

**Visualization with Centered and Transformed Data**

# plotting a multiple regression model based on:   
 # Model 5 Transform: Success.c ~ Income.c + Literacy.t [centered predictors]  
  
# create predicted values from predictors and save in object  
model5 <- lm(Success ~ Income.c + Literacy.t, data = dat)  
summary(model5)

##   
## Call:  
## lm(formula = Success ~ Income.c + Literacy.t, data = dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.89881 -0.71826 0.06009 0.66334 1.98364   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.478e+00 7.507e-01 3.301 0.001846 \*\*   
## Income.c 1.259e-04 2.333e-05 5.398 2.17e-06 \*\*\*  
## Literacy.t 8.741e-01 2.461e-01 3.551 0.000884 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9555 on 47 degrees of freedom  
## Multiple R-squared: 0.5971, Adjusted R-squared: 0.58   
## F-statistic: 34.83 on 2 and 47 DF, p-value: 5.275e-10

model\_p <- ggpredict(model5, terms = c('Income.c', 'Literacy.t'), full.data = TRUE, pretty = FALSE)  
  
# plot predicted line  
plot <- ggplot(model\_p, aes(x, predicted)) +  
 geom\_smooth(method = "lm", se = TRUE, fullrange=TRUE) + xlab("Score") + ggtitle("Plot of Model of Income and Literacy Predicting Success") + ylab("Success") +  
 geom\_point() + theme\_minimal()  
  
plot

