PSYR6003 Assignment 2

Daniel Basso

#Load Packages   
library(tidyverse)

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ dplyr 1.1.4 ✔ readr 2.1.5  
✔ forcats 1.0.0 ✔ stringr 1.5.1  
✔ ggplot2 3.5.0 ✔ tibble 3.2.1  
✔ lubridate 1.9.3 ✔ tidyr 1.3.1  
✔ purrr 1.0.2   
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(flexplot)

Attaching package: 'flexplot'  
  
The following object is masked from 'package:ggplot2':  
  
 flip\_data

library(haven)   
library(foreign)   
library(patchwork)   
library(MASS)

Attaching package: 'MASS'  
  
The following object is masked from 'package:patchwork':  
  
 area  
  
The following object is masked from 'package:dplyr':  
  
 select

library(quarto)   
library(apaTables)

#Load in Dataset as an object called "mydata"   
mydata <- read\_spss("PSYR6003.A2.sav")

#View & Inspect dataset   
summary(mydata)

RespondentId sex mps.SPP1.3y mps.SPP2.3y   
 Length:137 Length:137 Min. :1.000 Min. :1.000   
 Class :character Class :character 1st Qu.:3.000 1st Qu.:5.000   
 Mode :character Mode :character Median :5.000 Median :6.000   
 Mean :4.547 Mean :5.314   
 3rd Qu.:6.000 3rd Qu.:6.000   
 Max. :7.000 Max. :7.000   
   
 mps.SPP3.3y mps.SPP4.3y mps.SPP5.3y guilt1.3y   
 Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
 1st Qu.:3.000 1st Qu.:3.000 1st Qu.:3.000 1st Qu.:2.000   
 Median :4.000 Median :4.000 Median :4.000 Median :2.000   
 Mean :4.263 Mean :3.904 Mean :4.022 Mean :2.708   
 3rd Qu.:6.000 3rd Qu.:5.000 3rd Qu.:5.000 3rd Qu.:4.000   
 Max. :7.000 Max. :7.000 Max. :7.000 Max. :5.000   
 NA's :1 NA's :1   
 guilt2.3y guilt3.3y dep1.3y dep2.3y   
 Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
 1st Qu.:1.000 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000   
 Median :2.000 Median :3.000 Median :3.000 Median :3.000   
 Mean :2.219 Mean :2.891 Mean :3.199 Mean :2.728   
 3rd Qu.:3.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000   
 Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000   
 NA's :1 NA's :1   
 dep3.3y fear1.3y fear2.3y fear3.3y   
 Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.000   
 Median :3.000 Median :2.000 Median :2.000 Median :2.000   
 Mean :2.796 Mean :2.409 Mean :2.299 Mean :2.109   
 3rd Qu.:4.000 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:3.000   
 Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000   
   
 host1.3y host2.3y host3.3y tipm.CONS1.3y tipm.CONS2.3y   
 Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000 Min. :1.000   
 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:5.000 1st Qu.:2.000   
 Median :1.000 Median :1.000 Median :2.00 Median :5.000 Median :3.000   
 Mean :1.934 Mean :1.796 Mean :2.38 Mean :5.301 Mean :3.255   
 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:3.00 3rd Qu.:6.000 3rd Qu.:5.000   
 Max. :5.000 Max. :5.000 Max. :5.00 Max. :7.000 Max. :7.000   
 NA's :1

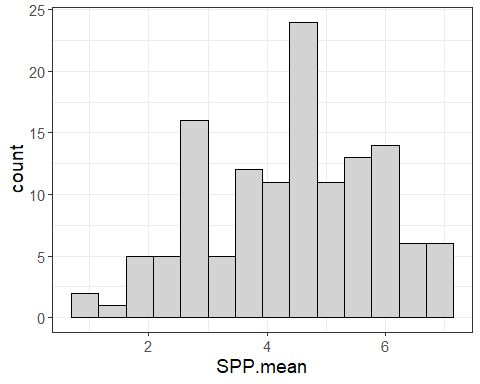
view(mydata)

#reverse coding item tipm.CONS2.3y before scale totalling  
#Recoding sex to be a factor  
#Filtering to only include male and female participants as the other category (n=1) is too small to consider   
#compute scales (mean) totals for socially prescribed perfectionism, conscientiousness, and negative affect   
#subsetting adjustments into object called "clean"   
#any totals with missing values were given NA as a total   
#Omitting Missing values  
  
mydata$tipm.CONS2.3y.REV = (abs(mydata$tipm.CONS2.3y - 8))   
clean <- mydata %>%   
 mutate(RespondentId = RespondentId,   
 sex = fct\_recode(sex, NULL = "Other (specify)", "0" = "Female", "1" = "Male"),   
 SPP.mean = rowMeans(dplyr::select(mydata, mps.SPP1.3y:mps.SPP5.3y), na.rm = F),   
 NEG.AFF.mean = rowMeans(dplyr::select(mydata, guilt1.3y:host3.3y), na.rm = F),   
 CONS.mean = rowMeans(dplyr::select(mydata, tipm.CONS1.3y, tipm.CONS2.3y.REV),   
 na.rm = F),   
 .keep = "used") %>%   
 filter(sex != "Other (specify)") %>%   
 na.omit()

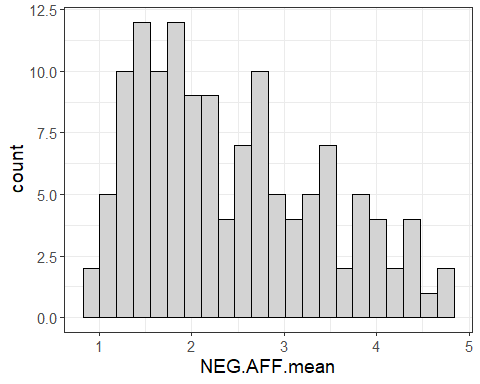
#bivariate correlations and descriptive stats   
correlation <- tibble(clean$SPP.mean, clean$NEG.AFF.mean, clean$CONS.mean, as.numeric(clean$sex))  
apa.cor.table(correlation)

Means, standard deviations, and correlations with confidence intervals  
   
  
 Variable M SD 1 2 3   
 1. clean$SPP.mean 4.38 1.43   
   
 2. clean$NEG.AFF.mean 2.44 0.98 .37\*\*   
 [.21, .51]   
   
 3. clean$CONS.mean 5.03 1.29 -.21\* -.37\*\*   
 [-.37, -.04] [-.51, -.22]   
   
 4. as.numeric(clean$sex) 1.15 0.35 .01 -.16 -.14   
 [-.16, .18] [-.32, .01] [-.30, .03]  
   
  
Note. M and SD are used to represent mean and standard deviation, respectively.  
Values in square brackets indicate the 95% confidence interval.  
The confidence interval is a plausible range of population correlations   
that could have caused the sample correlation (Cumming, 2014).  
 \* indicates p < .05. \*\* indicates p < .01.

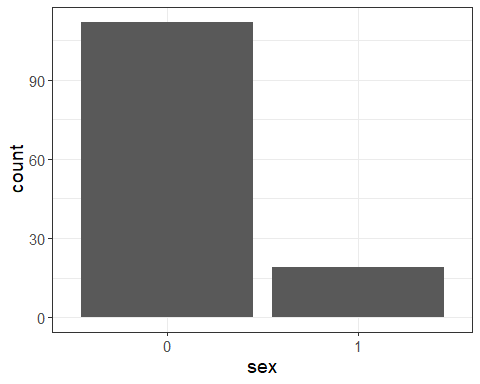
#Visualizing Univariate Distributions   
flexplot(SPP.mean~1, data=clean)



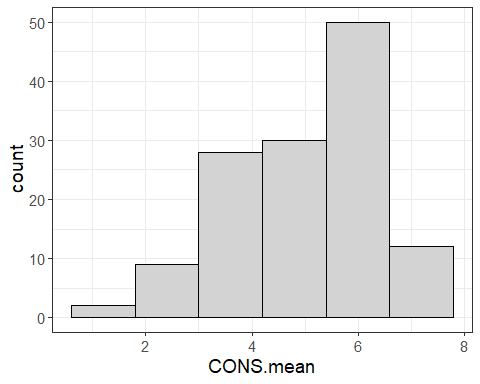
flexplot(NEG.AFF.mean~1, data=clean)



flexplot(sex~1, data=clean)



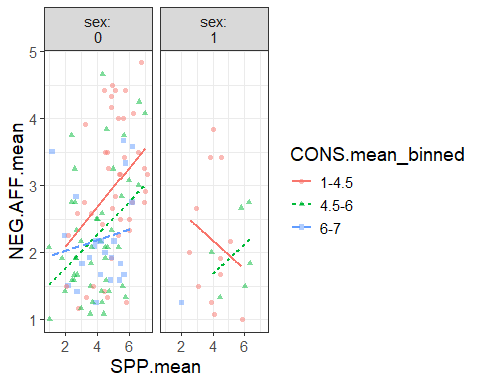
flexplot(CONS.mean~1, data=clean)



#Summary Stats of Variables in Clean   
summary(clean)

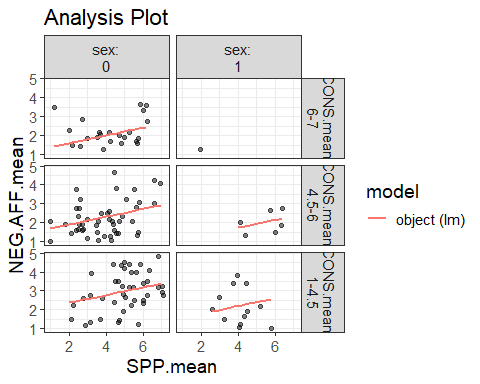
RespondentId sex SPP.mean NEG.AFF.mean CONS.mean   
 Length:131 0:112 Min. :1.000 Min. :1.000 Min. :1.000   
 Class :character 1: 19 1st Qu.:3.200 1st Qu.:1.583 1st Qu.:4.000   
 Mode :character Median :4.400 Median :2.167 Median :5.000   
 Mean :4.385 Mean :2.440 Mean :5.034   
 3rd Qu.:5.600 3rd Qu.:3.167 3rd Qu.:6.000   
 Max. :7.000 Max. :4.833 Max. :7.000

#Model visualization   
flexplot(NEG.AFF.mean ~ SPP.mean + CONS.mean | sex, data=clean, method = "lm")

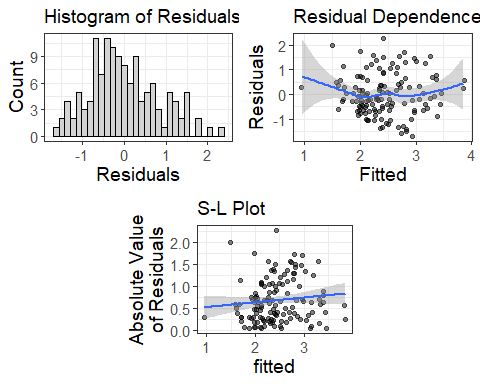


#Visualizing diagnostics   
base <- lm(NEG.AFF.mean ~ SPP.mean + CONS.mean + sex, data=clean)   
visualize(base, plot = "model")

`geom\_line()`: Each group consists of only one observation.  
ℹ Do you need to adjust the group aesthetic?

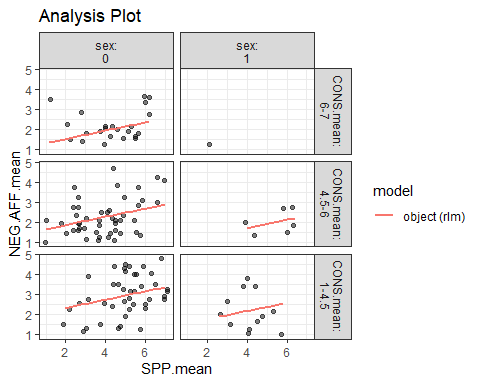


visualize(base, plot = "residuals")

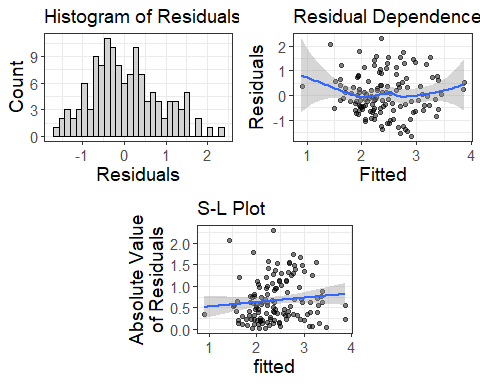


#Robust Model   
robust <- rlm(NEG.AFF.mean ~ SPP.mean + CONS.mean + sex, data = clean)   
visualize(robust, plot="model")

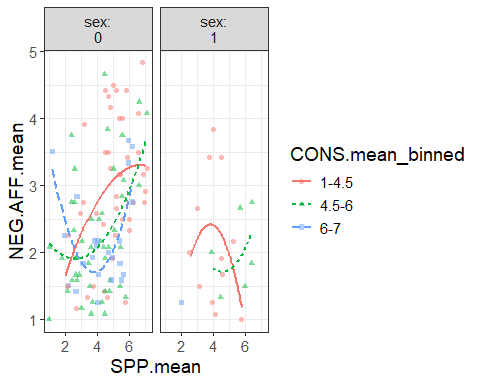
`geom\_line()`: Each group consists of only one observation.  
ℹ Do you need to adjust the group aesthetic?



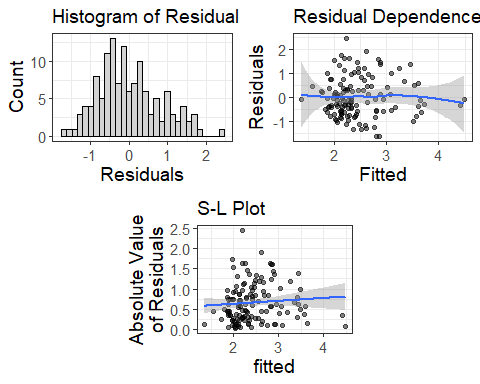
visualize(robust, plot="residuals")



#Quadratic sensitivity analysis   
flexplot(NEG.AFF.mean ~ SPP.mean + CONS.mean | sex, data = clean, method = "quadratic")



quadratic = lm(NEG.AFF.mean~ SPP.mean + CONS.mean + sex + I(SPP.mean^2) + I(CONS.mean^2), data=clean)   
visualize(quadratic, plot="residuals")



#Model Estimates   
estimates(base)

Note: You didn't choose to plot SPP.mean so I am inputting the median

Note: You didn't choose to plot CONS.mean so I am inputting the median

Model R squared:  
0.269 (0.14, 0.4)  
  
Semi-Partial R squared:  
 SPP.mean CONS.mean sex   
 0.134 0.092 0.043   
  
Estimates for Factors:  
 variables levels estimate lower upper  
1 sex 0 2.54 2.38 2.69  
2 1 1.96 1.57 2.35  
  
  
Mean Differences:  
 variables comparison difference lower upper cohens.d  
1 sex 1-0 -0.58 -1.16 0.01 -0.68  
  
  
Estimates for Numeric Variables =   
 variables estimate lower upper std.estimate std.lower std.upper  
1 (Intercept) 2.93 2.09 3.77 0.00 0.00 0.00  
2 SPP.mean 0.20 0.10 0.31 0.30 0.14 0.45  
3 CONS.mean -0.26 -0.37 -0.14 -0.34 -0.49 -0.19

estimates(quadratic)

Note: You didn't choose to plot SPP.mean so I am inputting the median  
  
Note: You didn't choose to plot CONS.mean so I am inputting the median

Model R squared:  
0.299 (0.17, 0.43)  
  
Semi-Partial R squared:  
 SPP.mean CONS.mean sex I(SPP.mean^2) I(CONS.mean^2)   
 0.134 0.092 0.043 0.018 0.012   
  
Estimates for Factors:  
 variables levels estimate lower upper  
1 sex 0 2.33 2.09 2.56  
2 1 1.82 1.42 2.22  
  
  
Mean Differences:  
 variables comparison difference lower upper cohens.d  
1 sex 1-0 -0.51 -1.09 0.07 -0.61  
  
  
Estimates for Numeric Variables =   
 variables estimate lower upper std.estimate std.lower std.upper  
1 (Intercept) 4.92 2.99 6.85 0.00 0.00 0.00  
2 SPP.mean -0.28 -0.81 0.26 -0.41 -1.19 0.38  
3 CONS.mean -0.74 -1.40 -0.09 -0.98 -1.85 -0.11  
4 I(SPP.mean^2) 0.06 -0.01 0.12 0.70 -0.08 1.48  
5 I(CONS.mean^2) 0.05 -0.02 0.12 0.65 -0.22 1.51

#full and reduced   
full <- lm(NEG.AFF.mean~ SPP.mean + CONS.mean + sex, data=clean)   
reduced <- lm(NEG.AFF.mean~ CONS.mean + sex, data=clean)   
model.comparison(full, reduced)

$statistics  
 aic bic bayes.factor p rsq  
full 333.257 347.633 111.432 <2e-16 0.269  
reduced 345.559 357.060 0.009 0.185  
  
$predicted\_differences  
 0% 25% 50% 75% 100%   
0.007 0.073 0.230 0.357 0.644

estimates(full)

Note: You didn't choose to plot SPP.mean so I am inputting the median

Note: You didn't choose to plot CONS.mean so I am inputting the median

Model R squared:  
0.269 (0.14, 0.4)  
  
Semi-Partial R squared:  
 SPP.mean CONS.mean sex   
 0.134 0.092 0.043   
  
Estimates for Factors:  
 variables levels estimate lower upper  
1 sex 0 2.54 2.38 2.69  
2 1 1.96 1.57 2.35  
  
  
Mean Differences:  
 variables comparison difference lower upper cohens.d  
1 sex 1-0 -0.58 -1.16 0.01 -0.68  
  
  
Estimates for Numeric Variables =   
 variables estimate lower upper std.estimate std.lower std.upper  
1 (Intercept) 2.93 2.09 3.77 0.00 0.00 0.00  
2 SPP.mean 0.20 0.10 0.31 0.30 0.14 0.45  
3 CONS.mean -0.26 -0.37 -0.14 -0.34 -0.49 -0.19

estimates(reduced)

Note: You didn't choose to plot CONS.mean so I am inputting the median

Model R squared:  
0.185 (0.07, 0.3)  
  
Semi-Partial R squared:  
CONS.mean sex   
 0.140 0.045   
  
Estimates for Factors:  
 variables levels estimate lower upper  
1 sex 0 2.54 2.37 2.7  
2 1 1.95 1.54 2.35  
  
  
Mean Differences:  
 variables comparison difference lower upper cohens.d  
1 sex 1-0 -0.59 -1.21 0.03 -0.67  
  
  
Estimates for Numeric Variables =   
 variables estimate lower upper std.estimate std.lower std.upper  
1 (Intercept) 4.06 3.43 4.70 0.0 0.00 0.00  
2 CONS.mean -0.31 -0.43 -0.19 -0.4 -0.56 -0.25

summary(full)

Call:  
lm(formula = NEG.AFF.mean ~ SPP.mean + CONS.mean + sex, data = clean)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-1.6989 -0.5926 -0.1454 0.5276 2.2597   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.92875 0.42671 6.864 2.65e-10 \*\*\*  
SPP.mean 0.20323 0.05310 3.828 0.000202 \*\*\*  
CONS.mean -0.25746 0.05937 -4.337 2.92e-05 \*\*\*  
sex1 -0.57578 0.21146 -2.723 0.007381 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.844 on 127 degrees of freedom  
Multiple R-squared: 0.2689, Adjusted R-squared: 0.2517   
F-statistic: 15.57 on 3 and 127 DF, p-value: 1.1e-08

summary(reduced)

Call:  
lm(formula = NEG.AFF.mean ~ CONS.mean + sex, data = clean)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-1.98057 -0.61977 -0.09184 0.61649 2.28292   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 4.06343 0.32287 12.585 < 2e-16 \*\*\*  
CONS.mean -0.30540 0.06105 -5.003 1.83e-06 \*\*\*  
sex1 -0.59120 0.22241 -2.658 0.00886 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.8878 on 128 degrees of freedom  
Multiple R-squared: 0.1846, Adjusted R-squared: 0.1719   
F-statistic: 14.49 on 2 and 128 DF, p-value: 2.127e-06