Untitled

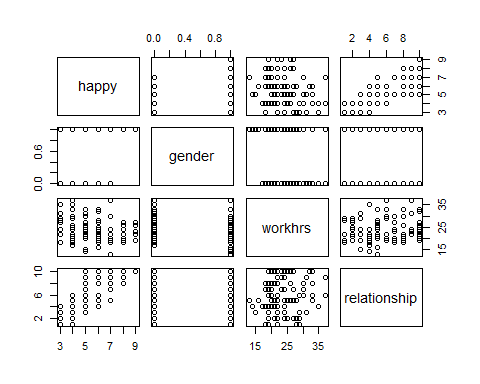
Kendall Brown

November 30, 2017

projdata=read.table("C:/Users/kebro/Desktop/Pstat 126/projdata.txt",header = T)  
fitfm=lm(happy~.,data=projdata)  
summary(fitfm)

##   
## Call:  
## lm(formula = happy ~ ., data = projdata)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.04590 -0.35802 -0.02218 0.37697 1.26763   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.54123 0.28090 12.607 < 2e-16 \*\*\*  
## gender 1.55447 0.10700 14.528 < 2e-16 \*\*\*  
## workhrs -0.07118 0.01082 -6.576 2.52e-09 \*\*\*  
## relationship 0.48538 0.01821 26.649 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5302 on 96 degrees of freedom  
## Multiple R-squared: 0.907, Adjusted R-squared: 0.9041   
## F-statistic: 312.2 on 3 and 96 DF, p-value: < 2.2e-16

plot(projdata)



fitint=lm(happy~.^2,data=projdata)  
summary(fitint)

##   
## Call:  
## lm(formula = happy ~ .^2, data = projdata)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.86671 -0.26448 -0.04598 0.30179 0.86016   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.906103 0.502528 7.773 1.01e-11 \*\*\*  
## gender 0.379229 0.399279 0.950 0.3447   
## workhrs -0.053897 0.020836 -2.587 0.0112 \*   
## relationship 0.401203 0.077494 5.177 1.30e-06 \*\*\*  
## gender:workhrs -0.008898 0.016067 -0.554 0.5810   
## gender:relationship 0.243410 0.027169 8.959 3.26e-14 \*\*\*  
## workhrs:relationship -0.002106 0.003132 -0.672 0.5030   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3933 on 93 degrees of freedom  
## Multiple R-squared: 0.9505, Adjusted R-squared: 0.9473   
## F-statistic: 297.4 on 6 and 93 DF, p-value: < 2.2e-16

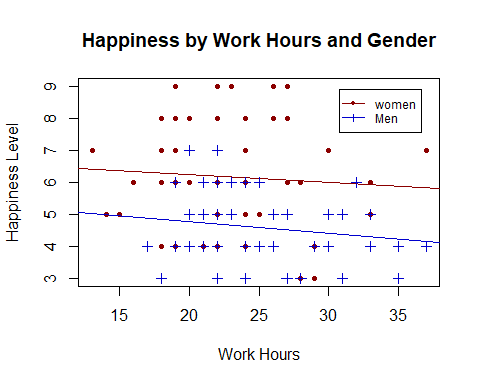
anova(fitfm)

## Analysis of Variance Table  
##   
## Response: happy  
## Df Sum Sq Mean Sq F value Pr(>F)   
## gender 1 61.439 61.439 218.5215 < 2.2e-16 \*\*\*  
## workhrs 1 2.265 2.265 8.0557 0.005536 \*\*   
## relationship 1 199.666 199.666 710.1608 < 2.2e-16 \*\*\*  
## Residuals 96 26.991 0.281   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

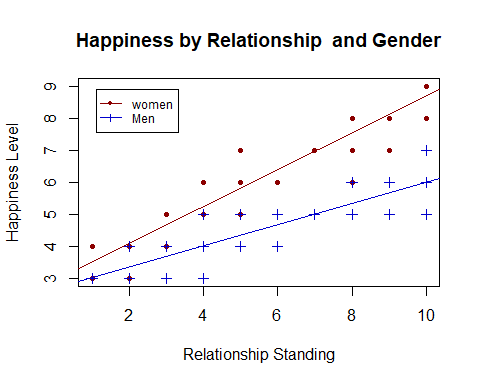
anova(fitint)

## Analysis of Variance Table  
##   
## Response: happy  
## Df Sum Sq Mean Sq F value Pr(>F)   
## gender 1 61.439 61.439 397.2186 < 2.2e-16 \*\*\*  
## workhrs 1 2.265 2.265 14.6433 0.0002352 \*\*\*  
## relationship 1 199.666 199.666 1290.8985 < 2.2e-16 \*\*\*  
## gender:workhrs 1 0.077 0.077 0.4965 0.4827852   
## gender:relationship 1 12.460 12.460 80.5558 3.01e-14 \*\*\*  
## workhrs:relationship 1 0.070 0.070 0.4521 0.5029995   
## Residuals 93 14.384 0.155   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

attach(projdata)  
plot(workhrs[gender==1],happy[gender==1],col="red4",xlab="Work Hours",ylab="Happiness Level", main="Happiness by Work Hours and Gender",pch=20)  
abline(lm(happy[gender==1]~workhrs[gender==1]),col="red4")  
points(workhrs[gender==0],happy[gender==0],col="blue3",pch=3)  
abline(lm(happy[gender==0]~workhrs[gender==0]),col="blue3")  
legend("topright",col=c("red4","blue3"),inset=.05,cex=.75,lty=c(1,1),pch=c(20,3),legend=c("women","Men"))



plot(relationship[gender==1],happy[gender==1],col="red4",xlab="Relationship Standing",ylab="Happiness Level", main="Happiness by Relationship and Gender",pch=20)  
abline(lm(happy[gender==1]~relationship[gender==1]),col="red4")  
points(relationship[gender==0],happy[gender==0],col="blue3",pch=3)  
abline(lm(happy[gender==0]~relationship[gender==0]),col="blue3")  
legend("topleft",col=c("red4","blue3"),inset=.05,cex=.75,lty=c(1,1),pch=c(20,3),legend=c("women","Men"))



fitlow=lm(happy~1,data=projdata)  
step(fitlow,scope=list(lower=fitlow,upper=fitint),direction="forward")

## Start: AIC=108.6  
## happy ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + relationship 1 183.983 106.38 10.182  
## + gender 1 61.439 228.92 86.821  
## + workhrs 1 6.171 284.19 108.447  
## <none> 290.36 108.595  
##   
## Step: AIC=10.18  
## happy ~ relationship  
##   
## Df Sum of Sq RSS AIC  
## + gender 1 67.227 39.150 -87.777  
## + workhrs 1 20.043 86.335 -8.694  
## <none> 106.377 10.182  
##   
## Step: AIC=-87.78  
## happy ~ relationship + gender  
##   
## Df Sum of Sq RSS AIC  
## + gender:relationship 1 12.802 26.348 -125.376  
## + workhrs 1 12.159 26.991 -122.967  
## <none> 39.150 -87.777  
##   
## Step: AIC=-125.38  
## happy ~ relationship + gender + relationship:gender  
##   
## Df Sum of Sq RSS AIC  
## + workhrs 1 11.843 14.506 -183.06  
## <none> 26.348 -125.38  
##   
## Step: AIC=-183.06  
## happy ~ relationship + gender + workhrs + relationship:gender  
##   
## Df Sum of Sq RSS AIC  
## <none> 14.506 -183.06  
## + workhrs:relationship 1 0.073733 14.432 -181.57  
## + gender:workhrs 1 0.051244 14.454 -181.42

##   
## Call:  
## lm(formula = happy ~ relationship + gender + workhrs + relationship:gender,   
## data = projdata)  
##   
## Coefficients:  
## (Intercept) relationship gender   
## 4.28774 0.35210 0.17835   
## workhrs relationship:gender   
## -0.07026 0.24158

step(fitint,scope=list(lower=fitlow,upper=fitint),direction="backward")

## Start: AIC=-179.9  
## happy ~ (gender + workhrs + relationship)^2  
##   
## Df Sum of Sq RSS AIC  
## - gender:workhrs 1 0.0474 14.432 -181.57  
## - workhrs:relationship 1 0.0699 14.454 -181.42  
## <none> 14.384 -179.90  
## - gender:relationship 1 12.4145 26.799 -119.68  
##   
## Step: AIC=-181.57  
## happy ~ gender + workhrs + relationship + gender:relationship +   
## workhrs:relationship  
##   
## Df Sum of Sq RSS AIC  
## - workhrs:relationship 1 0.0737 14.506 -183.06  
## <none> 14.432 -181.57  
## - gender:relationship 1 12.4494 26.881 -121.37  
##   
## Step: AIC=-183.06  
## happy ~ gender + workhrs + relationship + gender:relationship  
##   
## Df Sum of Sq RSS AIC  
## <none> 14.506 -183.06  
## - workhrs 1 11.843 26.348 -125.38  
## - gender:relationship 1 12.485 26.991 -122.97

##   
## Call:  
## lm(formula = happy ~ gender + workhrs + relationship + gender:relationship,   
## data = projdata)  
##   
## Coefficients:  
## (Intercept) gender workhrs   
## 4.28774 0.17835 -0.07026   
## relationship gender:relationship   
## 0.35210 0.24158

step(fitlow,scope=list(lower=fitlow,upper=fitint),direction="both")

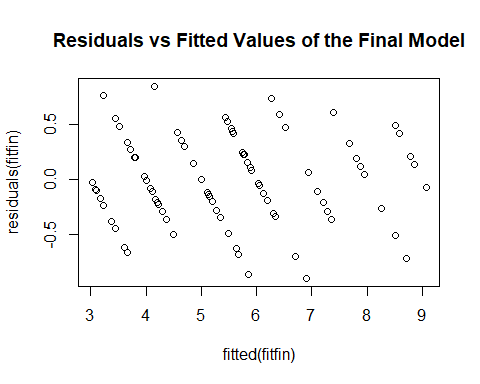
## Start: AIC=108.6  
## happy ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + relationship 1 183.983 106.38 10.182  
## + gender 1 61.439 228.92 86.821  
## + workhrs 1 6.171 284.19 108.447  
## <none> 290.36 108.595  
##   
## Step: AIC=10.18  
## happy ~ relationship  
##   
## Df Sum of Sq RSS AIC  
## + gender 1 67.227 39.150 -87.777  
## + workhrs 1 20.043 86.335 -8.694  
## <none> 106.377 10.182  
## - relationship 1 183.983 290.360 108.595  
##   
## Step: AIC=-87.78  
## happy ~ relationship + gender  
##   
## Df Sum of Sq RSS AIC  
## + gender:relationship 1 12.801 26.348 -125.376  
## + workhrs 1 12.159 26.991 -122.967  
## <none> 39.150 -87.777  
## - gender 1 67.227 106.377 10.182  
## - relationship 1 189.772 228.921 86.821  
##   
## Step: AIC=-125.38  
## happy ~ relationship + gender + relationship:gender  
##   
## Df Sum of Sq RSS AIC  
## + workhrs 1 11.843 14.506 -183.063  
## <none> 26.348 -125.376  
## - relationship:gender 1 12.802 39.150 -87.777  
##   
## Step: AIC=-183.06  
## happy ~ relationship + gender + workhrs + relationship:gender  
##   
## Df Sum of Sq RSS AIC  
## <none> 14.506 -183.06  
## + workhrs:relationship 1 0.0737 14.432 -181.57  
## + gender:workhrs 1 0.0512 14.454 -181.42  
## - workhrs 1 11.8427 26.348 -125.38  
## - relationship:gender 1 12.4853 26.991 -122.97

##   
## Call:  
## lm(formula = happy ~ relationship + gender + workhrs + relationship:gender,   
## data = projdata)  
##   
## Coefficients:  
## (Intercept) relationship gender   
## 4.28774 0.35210 0.17835   
## workhrs relationship:gender   
## -0.07026 0.24158

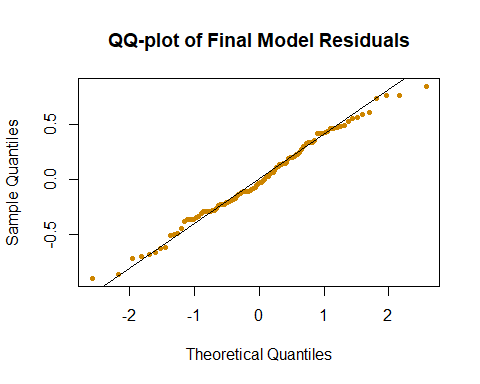
fitfin=lm(happy~gender+workhrs+relationship+gender:relationship)  
summary(fitfin)

##   
## Call:  
## lm(formula = happy ~ gender + workhrs + relationship + gender:relationship)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.89700 -0.26709 -0.02701 0.28099 0.84955   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.287745 0.222865 19.239 < 2e-16 \*\*\*  
## gender 0.178353 0.171396 1.041 0.301   
## workhrs -0.070259 0.007978 -8.807 5.85e-14 \*\*\*  
## relationship 0.352098 0.019935 17.662 < 2e-16 \*\*\*  
## gender:relationship 0.241580 0.026716 9.043 1.84e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3908 on 95 degrees of freedom  
## Multiple R-squared: 0.95, Adjusted R-squared: 0.9479   
## F-statistic: 451.7 on 4 and 95 DF, p-value: < 2.2e-16

plot(fitted(fitfin),residuals(fitfin),main="Residuals vs Fitted Values of the Final Model")



qqnorm(residuals(fitfin),main="QQ-plot of Final Model Residuals",col="orange3",pch=20)  
qqline(residuals(fitfin))



hist(residuals(fitfin),breaks=20,col="chocolate4",main="Histogram of Final Model Residuals")

