Linear Regression Tutorial Sheet Solutions

MYPEARSON<-read.csv("PearsonLeeSimple.csv")  
Pearson\_child\_parent<-lm(child~parent,data=MYPEARSON)  
summary(Pearson\_child\_parent)

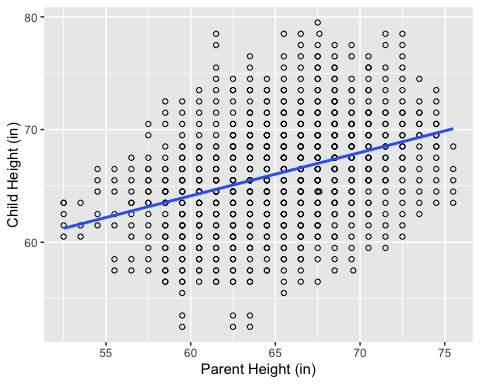
##   
## Call:  
## lm(formula = child ~ parent, data = MYPEARSON)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.9671 -3.5040 0.0329 3.1855 13.8013   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 41.06911 2.41880 16.98 <2e-16 \*\*\*  
## parent 0.38422 0.03711 10.36 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.81 on 744 degrees of freedom  
## Multiple R-squared: 0.126, Adjusted R-squared: 0.1248   
## F-statistic: 107.2 on 1 and 744 DF, p-value: < 2.2e-16

library(ggplot2)  
  
Pearson\_child\_height\_gp\_model<-lm(child~parent+gp,data=MYPEARSON)  
  
summary(Pearson\_child\_height\_gp\_model)

##   
## Call:  
## lm(formula = child ~ parent + gp, data = MYPEARSON)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.5860 -2.9404 0.0624 2.9530 10.1461   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 29.1414 2.3063 12.636 < 2e-16 \*\*\*  
## parent 0.5130 0.0341 15.046 < 2e-16 \*\*\*  
## gpfs 4.7764 0.4011 11.910 < 2e-16 \*\*\*  
## gpmd 2.4194 0.4263 5.675 1.99e-08 \*\*\*  
## gpms 7.6611 0.4285 17.880 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.925 on 741 degrees of freedom  
## Multiple R-squared: 0.4204, Adjusted R-squared: 0.4172   
## F-statistic: 134.3 on 4 and 741 DF, p-value: < 2.2e-16

ggplot(MYPEARSON, aes(x=parent, y=child)) +ylab("Child Height (in)")+xlab("Parent Height (in)")+  
 geom\_point(shape=1) +  
 scale\_colour\_hue(l=50) + # Use a slightly darker palette than normal  
 geom\_smooth(method=lm, # Add linear regression lines  
 se=FALSE)+ # Don't add shaded confidence region  
 scale\_color\_discrete(name = "Parent")

## Scale for 'colour' is already present. Adding another scale for  
## 'colour', which will replace the existing scale.



ggsave("Linear\_Regression.png",dpi=300, width = 4, height = 2)