Linear Regression

A linear regression is used to model a linear relationship of the dependent variable y and the regressors x_1, x_2, ...

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...,$$

where β_0, β_1 are the slopes of the regressors.

Height Example

A simple linear regression (correlation) is used to predict the height of 744 children y using the height of their parent x,

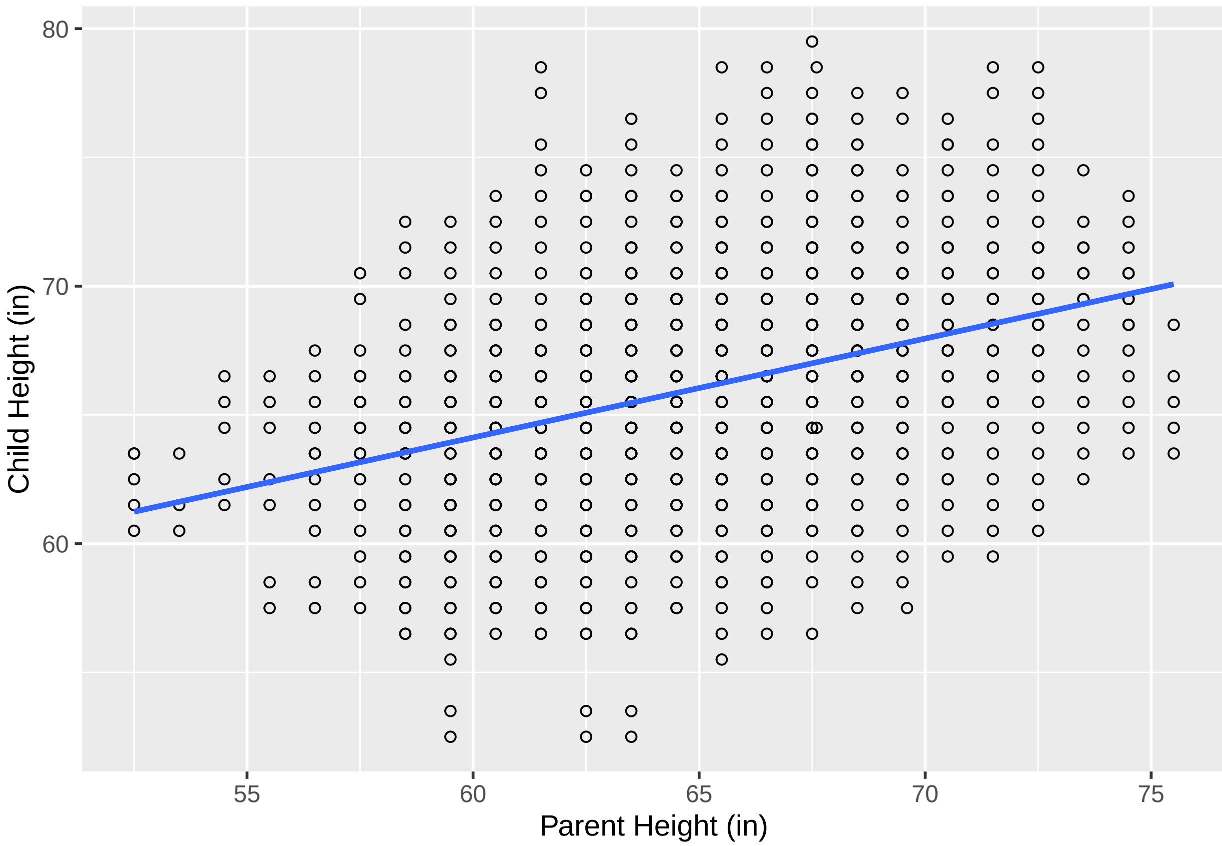
```
MYPEARSON<-read.csv("PearsonLeeSimple.csv")
Pearson_child_parent<-lm(child~parent,data=MYPEARSON)
summary(Pearson_child_parent)
```

```
##
## 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)

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)
```

Interpreting the slope of the regressot, β

1. If β is close to 0 then this would suggest that there is little to no relationship between the variable y and the regressor x, for extreme example, someones height (y) does not correlation with their teachers height x.

2. If β is greater than 0 then this means that there is a positive correlation, for example, a tall person y would have a tall parent x.