Advanced Bioinformatics 2019 assessment

8946

30/05/2019

# **Task 1**

sum(5:55)

## [1] 1530

# **Task 2**

sumfun <- function(n) {sum(5:n)}  
sumfun(10)

## [1] 45

sumfun(20)

## [1] 200

sumfun(100)

## [1] 5040

# **Task 3**

n <- 12  
fib <- numeric(n)  
fib[1] <- 1  
fib[2] <- 1  
for (i in 3:n)  
{  
fib[i] <- fib[i-1]+fib[i-2]   
3  
}  
fib

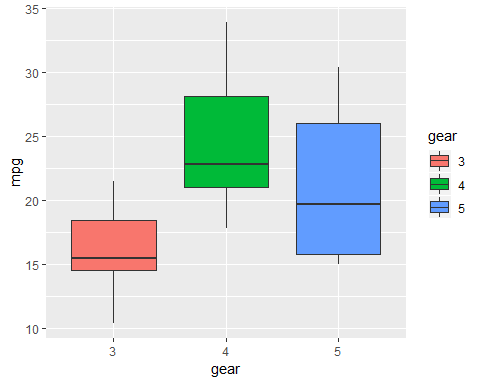
## [1] 1 1 2 3 5 8 13 21 34 55 89 144

# **Task 4**

library(ggplot2)

## Registered S3 methods overwritten by 'ggplot2':  
## method from   
## [.quosures rlang  
## c.quosures rlang  
## print.quosures rlang

data<-mtcars  
data$gear<-as.factor(mtcars$gear)  
p<-ggplot(data, aes(x=gear, y=mpg, fill=gear )) + geom\_boxplot()  
p



# **Task 5**

cars.lm <- lm(dist ~ speed, data = cars)  
cars.lm2 <- lm(dist ~ 0 + speed, data = cars)  
summary(cars.lm)

##   
## Call:  
## lm(formula = dist ~ speed, data = cars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -29.069 -9.525 -2.272 9.215 43.201   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -17.5791 6.7584 -2.601 0.0123 \*   
## speed 3.9324 0.4155 9.464 1.49e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.38 on 48 degrees of freedom  
## Multiple R-squared: 0.6511, Adjusted R-squared: 0.6438   
## F-statistic: 89.57 on 1 and 48 DF, p-value: 1.49e-12

summary(cars.lm2)

##   
## Call:  
## lm(formula = dist ~ 0 + speed, data = cars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.183 -12.637 -5.455 4.590 50.181   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## speed 2.9091 0.1414 20.58 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 16.26 on 49 degrees of freedom  
## Multiple R-squared: 0.8963, Adjusted R-squared: 0.8942   
## F-statistic: 423.5 on 1 and 49 DF, p-value: < 2.2e-16

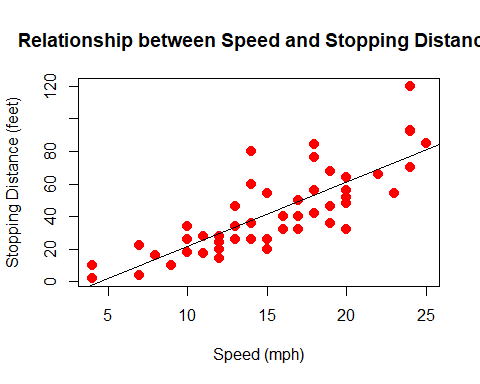
# Fitted slope = 3.9324

# Intercept of line = -17.5791

# Standard deviation = 0.4155

# **Task 6**

plot(cars, col='red',pch=20, cex=2, main="Relationship between Speed and Stopping Distance ", "Speed (mph)", "Stopping Distance (feet)")  
abline(-17.5791, 3.9324)



# **Task 7**

new\_speed <- cars$speed \* (5280/3600)  
line<-lm(dist ~ 0+ new\_speed + I(new\_speed^2),cars)  
summary(line)

##   
## Call:  
## lm(formula = dist ~ 0 + new\_speed + I(new\_speed^2), data = cars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.836 -9.071 -3.152 4.570 44.986   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## new\_speed 0.84479 0.38180 2.213 0.03171 \*   
## I(new\_speed^2) 0.04190 0.01366 3.067 0.00355 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.02 on 48 degrees of freedom  
## Multiple R-squared: 0.9133, Adjusted R-squared: 0.9097   
## F-statistic: 252.8 on 2 and 48 DF, p-value: < 2.2e-16

library(ggplot2)  
data <- cars  
reg\_data <- lm(dist ~ speed + I(speed^2), data=data)  
reg\_data

##   
## Call:  
## lm(formula = dist ~ speed + I(speed^2), data = data)  
##   
## Coefficients:  
## (Intercept) speed I(speed^2)   
## 2.47014 0.91329 0.09996

reg\_data$coefficients[3]

## I(speed^2)   
## 0.0999593

data$new\_speed <- data$speed \* (5280/3600)  
reg\_data <- lm(dist ~ 0 + new\_speed + I(new\_speed^2), data=data)  
summary(reg\_data)

##   
## Call:  
## lm(formula = dist ~ 0 + new\_speed + I(new\_speed^2), data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.836 -9.071 -3.152 4.570 44.986   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## new\_speed 0.84479 0.38180 2.213 0.03171 \*   
## I(new\_speed^2) 0.04190 0.01366 3.067 0.00355 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.02 on 48 degrees of freedom  
## Multiple R-squared: 0.9133, Adjusted R-squared: 0.9097   
## F-statistic: 252.8 on 2 and 48 DF, p-value: < 2.2e-16

ggplot(data, aes(x = new\_speed, y = dist)) +   
 geom\_point(color='Black', size = 1) + xlim(c(0,40)) +  
 geom\_smooth(method = "lm", formula = "y ~ 0 + x + I(x^2)", color="green", fullrange='TRUE') + labs(title= 'Average Breaking Reaction Time ', y = 'Distance (feet)', x = "Speed (secs)")

