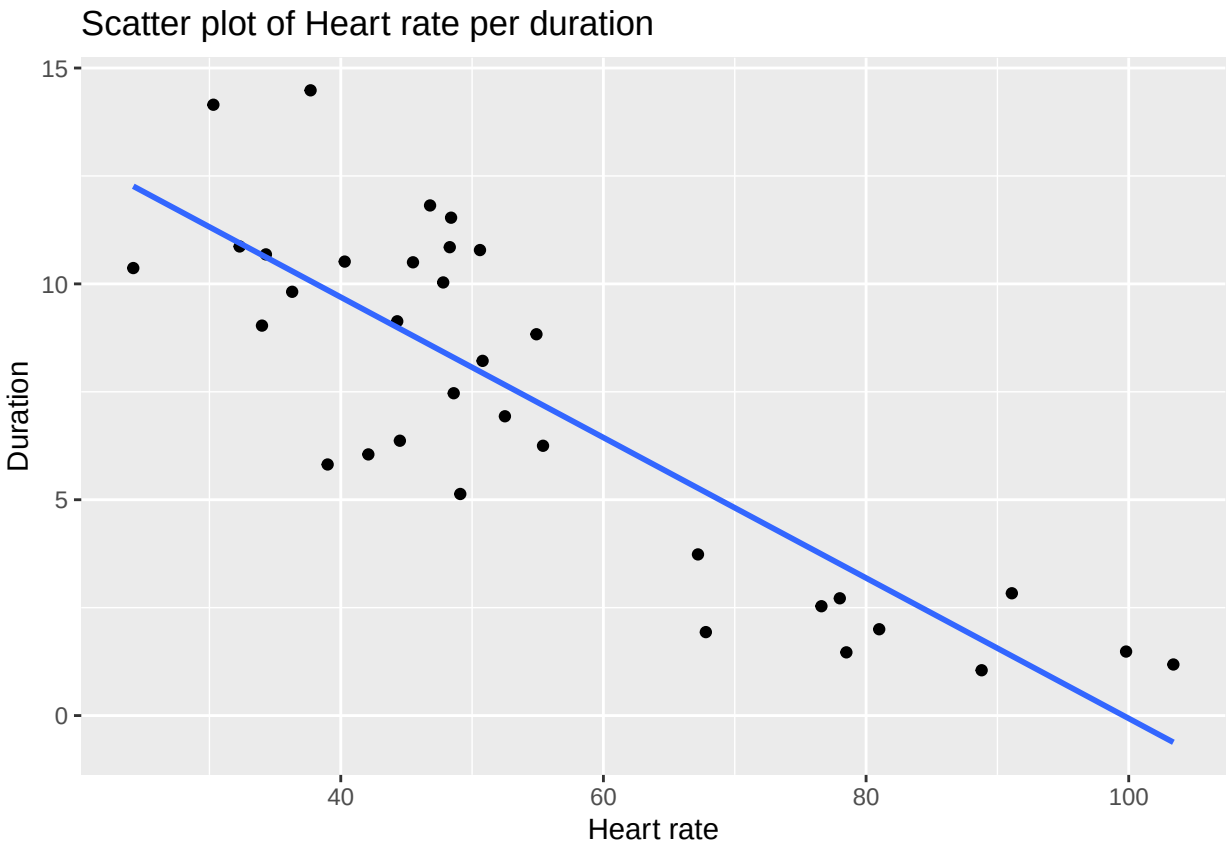
Heart Rate analysis

Contents

Question 1

```
library(ggplot2)
data<-read.csv("penguindives-2.csv")
ggplot(data, aes(x=Heart_Rate, y=Duration)) +
  geom_point()+
  geom_smooth(method=lm, se=FALSE)+ggtitle("Scatter plot of Heart rate per duration")+xlab("Heart rate"
```



Question 2

```
model<-lm(Duration~Heart_Rate,data)
model
```

```
##
## Call:
## lm(formula = Duration ~ Heart_Rate, data = data)
##
## Coefficients:
## (Intercept)   Heart_Rate
##     16.1980      -0.1626
```

Question 3

```
summary(model)
```

```
##
## Call:
## lm(formula = Duration ~ Heart_Rate, data = data)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0384 -1.4531 -0.2779  1.5982  4.4167
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.19805    1.07143  15.118 3.99e-16 ***
## Heart_Rate  -0.16264    0.01826  -8.908 3.55e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.176 on 32 degrees of freedom
## Multiple R-squared:  0.7126, Adjusted R-squared:  0.7036
## F-statistic: 79.35 on 1 and 32 DF,  p-value: 3.55e-10
```

We use the estimates from the summary of the model to write equation.

The equation of the line is Duration=16.19805-0.16264*Heart_Rate

Question 4

Hypothesis to be tested

- Null hypothesis: $\beta_1 = 0$
- Alternative hypothesis: $\beta_1 \neq 0$

Conducting the test

```
summary(model)$coefficients
```

```
##               Estimate Std. Error  t value     Pr(>|t|)
## (Intercept) 16.1980464 1.07142846 15.11818 3.986109e-16
## Heart_Rate  -0.1626403 0.01825862 -8.90759 3.549695e-10
```

```
t_statistic<-(-0.1626403-0)/0.01825862 #t-statistic=(mu-0)/standard error
t_statistic
```

```
## [1] -8.90759
```

The t-test statistic is -8.90759. We compare this value with t-tabulated from Student's t distribution table at 96% confidence interval. This value is 2.054.

P-value

```r
# determine residual degrees of freedom
model$df.residual
```

```
## [1] 32
```

```r
2 * pt(-8.90759, df = 32)
```

```
## [1] 3.549691e-10
```

P-value at 4% significance level is 3.549691e-10. This p-value uses T-statistic value

Decision

With p-value smaller than 0.04 at 4% significance level and again with absolute t-static value being greater than t-tabulated from the table,**we reject the null hypothesis**.

Conclussion

We conclude that ***Heart_Rate*** is important in explaining some of the variation in ***Duration*** at 4% significance level.

Question 5

The model's coefficient of determination is 0.7126. This mean that only 71.26% of the ***Duration*** is explained by ***Heart_Rate***.