Women's Height vs Weight Analysis

Santiago Jaramillo

2024-03-06

Description

In the provided script, we explored the relationship between height and weight in American women using the women dataset. Initially, we visualized this relationship through a scatter plot, highlighting the general trend that taller women tend to weigh more. To quantitatively analyze this trend, we fitted a linear regression model, using height to predict weight. This model helped us understand how weight changes with height, quantified by the model's coefficients. We further tested the model's predictive power by estimating weights for a new set of heights, presenting both the original data and these predictions visually. This approach not only confirmed the model's effectiveness but also demonstrated the practical application of statistical modeling in understanding real-world relationships.

```
# Load and view the initial part of the women dataset
data(women)
head(women)
```

```
##   height weight
## 1     58    115
## 2     59    117
## 3     60    120
## 4     61    123
## 5     62    126
## 6     63    129
```

Women's Height vs Weight
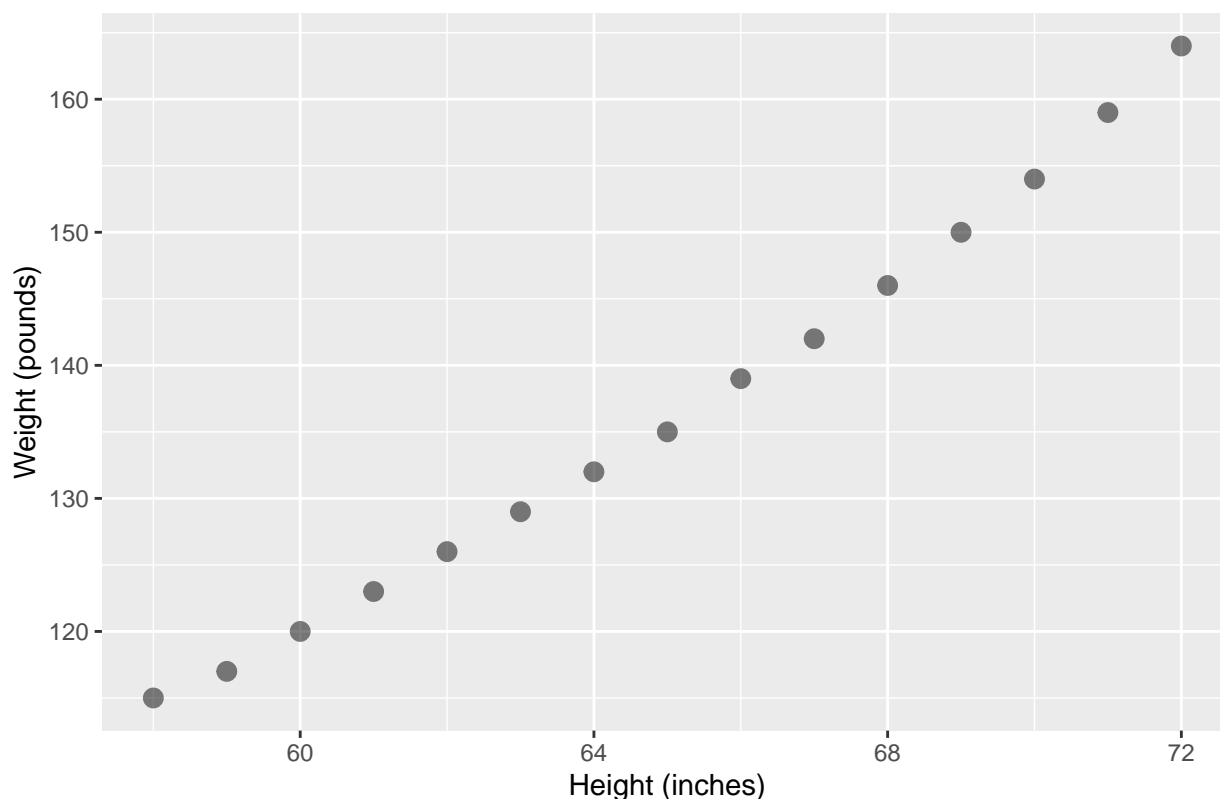

```r
# linear model predicting weight based on height
model1 <- lm(weight ~ height, data = women)
```

```r
# View the summary of the linear model
# The high R-squared value of 0.991 suggests that 99.1% of the variability in weight can be explained by
summary(model1)
```

```
##
## Call:
## lm(formula = weight ~ height, data = women)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7333 -1.1333 -0.3833  0.7417  3.1167
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -87.51667    5.93694  -14.74 1.71e-09 ***
## height        3.45000    0.09114   37.85 1.09e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.525 on 13 degrees of freedom
## Multiple R-squared:  0.991,  Adjusted R-squared:  0.9903
## F-statistic:  1433 on 1 and 13 DF,  p-value: 1.091e-14
```

```r
# Define new data for which predictions are to be made
new_data <- data.frame(height = c(68,70,72,71,63,65,67,57))
```

```
#Predict weights for the new heights using the fitted model
predicted_weights <- predict(model1, new_data)
```

```
# Create a new data frame including both the new heights and their predicted weights
new_data_with_predictions <- data.frame(
  Height = new_data$height,
  PredictedWeight = predicted_weights)
```

```
# Print the rounded predicted weights
print(round(predicted_weights))
```

```
##   1   2   3   4   5   6   7   8
## 147 154 161 157 130 137 144 109
```

Women's Height vs Weight with Predictions