Warehouse Wars: The Secret Life of Shipment Volumes and Handling Times

Avery Holloman

2024-08-11

```
# Libraries
library(magrittr)
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
library(ggthemes)
library(earth)
```

```
## Loading required package: Formula
```

```
## Loading required package: plotmo
```

```
## Loading required package: plotrix
```

```
library(car)
```

```
## Loading required package: carData
```

```
library(plotmo)
library(broom)
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:car':
##
##     recode
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(vtreat)
```

```
## Loading required package: wrapr
```

```
##
## Attaching package: 'wrapr'
```

```
## The following object is masked from 'package:dplyr':
##
##     coalesce
```

```
## The following object is masked from 'package:car':
##
##     bc
```

```
library(sjmisc)
```

```
# Anscombe's Quartet custom dataset I created
data(anscombe)
```

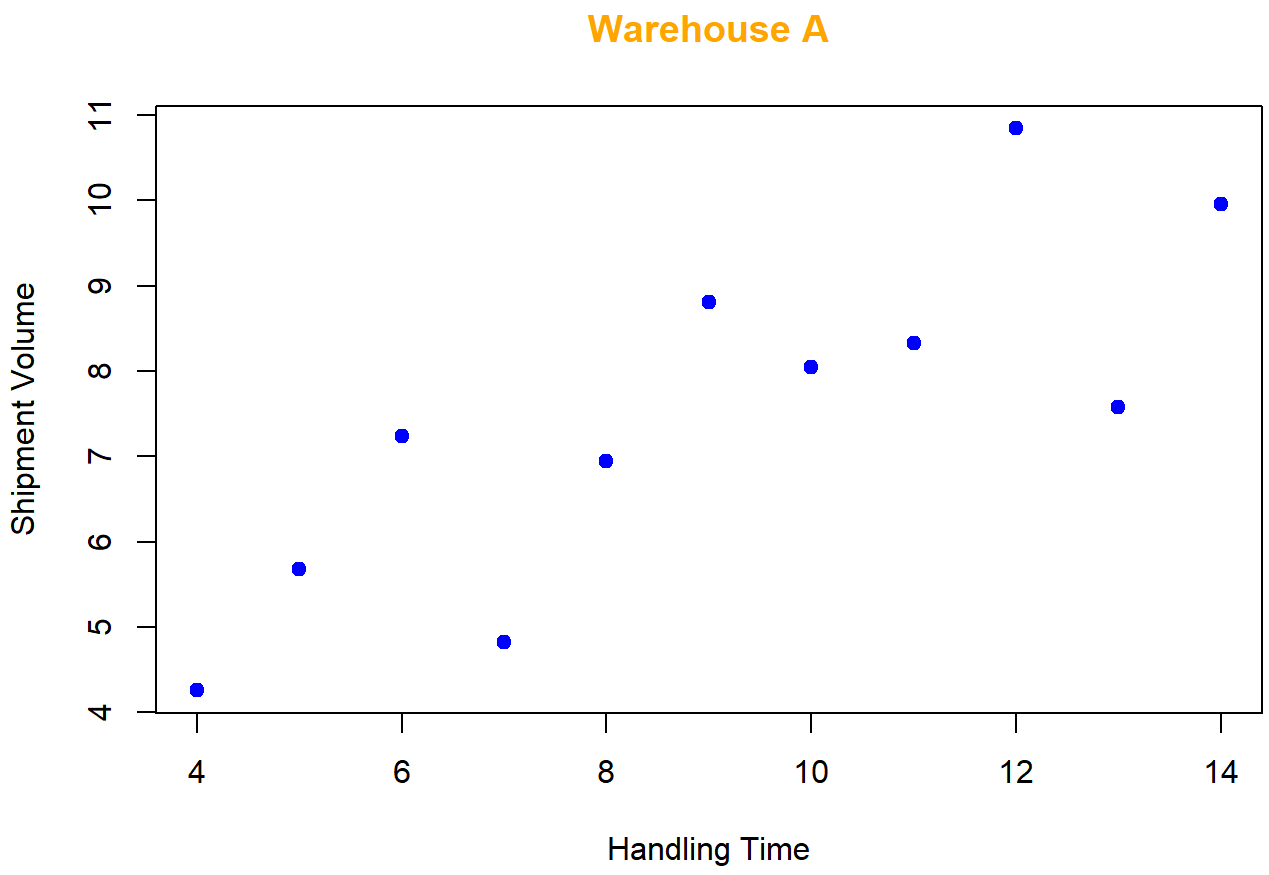
```
# I renamed variables to reflect warehouse logistics
warehouse_data <- anscombe %>%
  rename(
    Handling_Time_A = x1,
    Shipment_Volume_A = y1,
    Handling_Time_B = x2,
    Shipment_Volume_B = y2,
    Handling_Time_C = x3,
    Shipment_Volume_C = y3,
    Handling_Time_D = x4,
    Shipment_Volume_D = y4
  )
```

```
# Next I wanted to calculate the correlations between the data
cor_A <- cor(warehouse_data$Handling_Time_A, warehouse_data$Shipment_Volume_A) # Correlation for Warehouse A
cor_B <- cor(warehouse_data$Handling_Time_B, warehouse_data$Shipment_Volume_B) # Correlation for Warehouse B
```

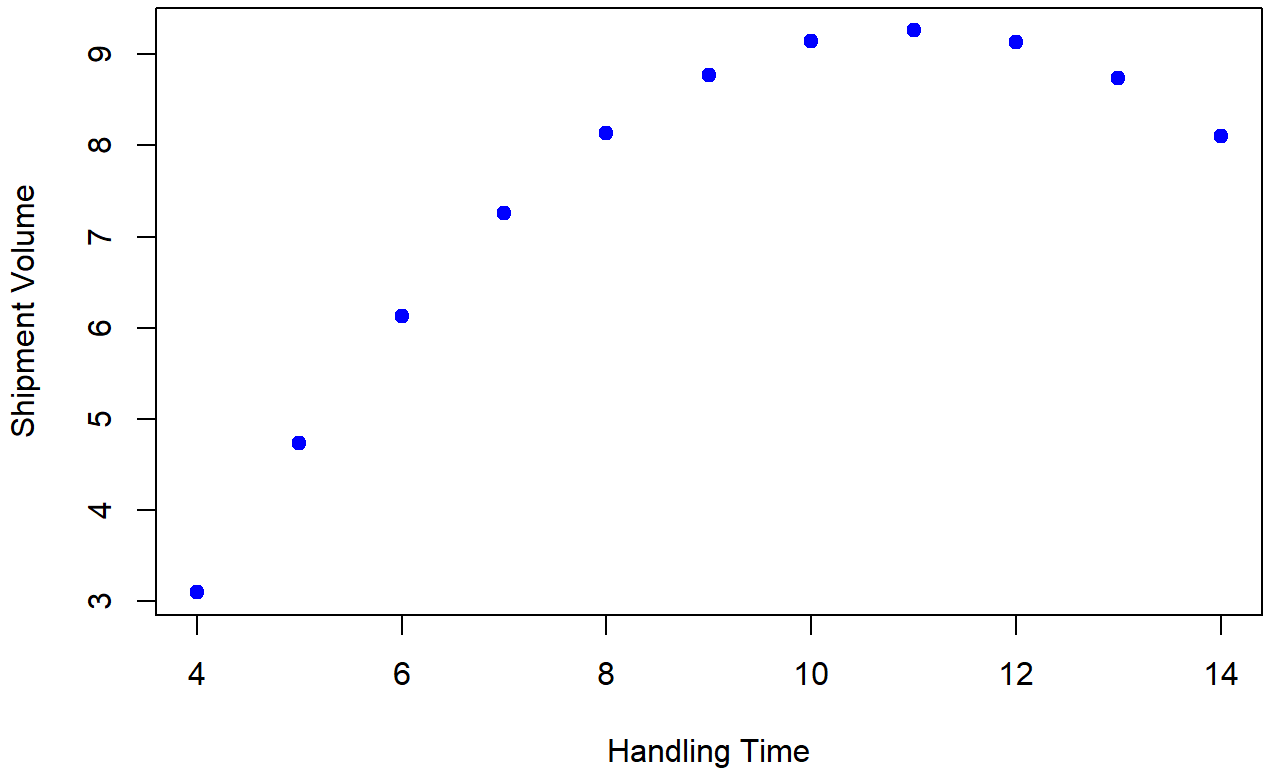
```
# Plotting Warehouse Logistics Data with custom colors
par(mfrow = c(2, 2)) # 2x2 grid for plots
```

```
# Changed the colors for my plot to blue and orange
point_color <- "blue"
title_color <- "orange"
```

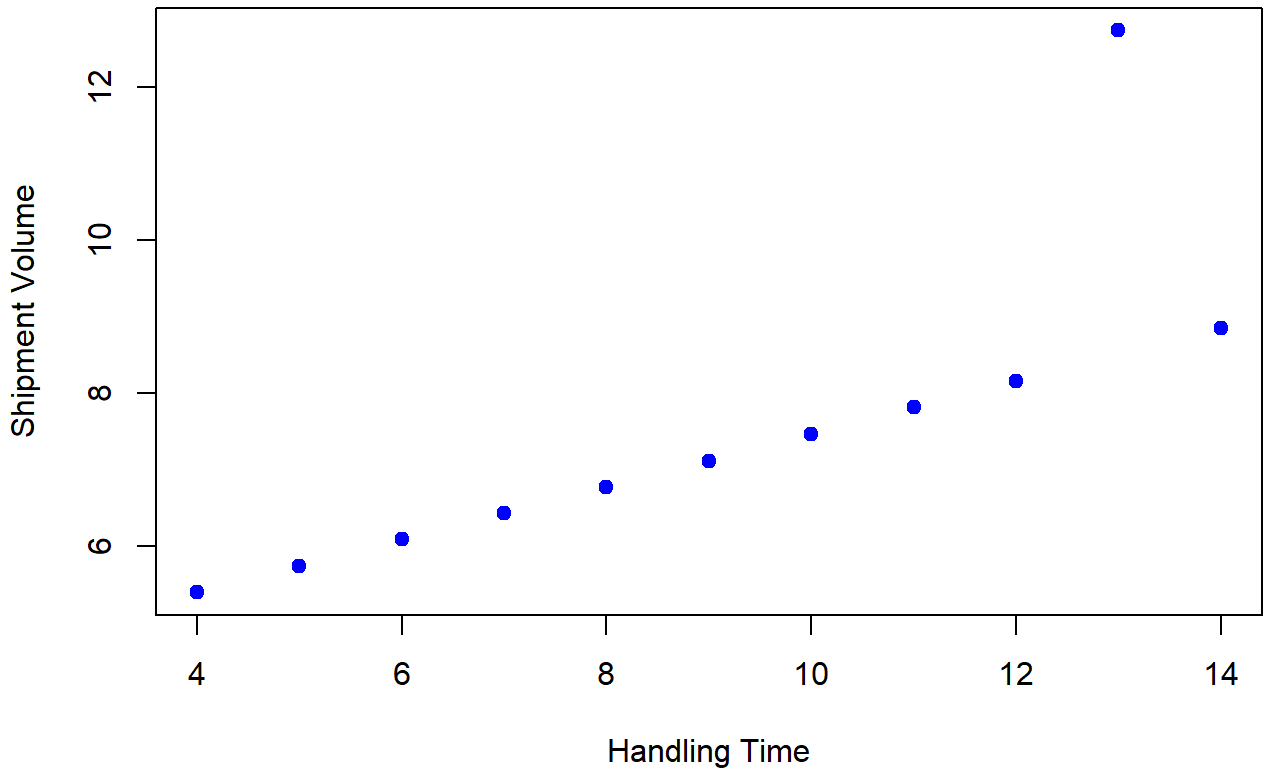
```
#Plotting with Title name and x and y labels
plot(warehouse_data$Handling_Time_A, warehouse_data$Shipment_Volume_A,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
title(main = "Warehouse A", col.main = title_color)
```




```
plot(warehouse_data$Handling_Time_B, warehouse_data$Shipment_Volume_B,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
title(main = "Warehouse B", col.main = title_color)
```

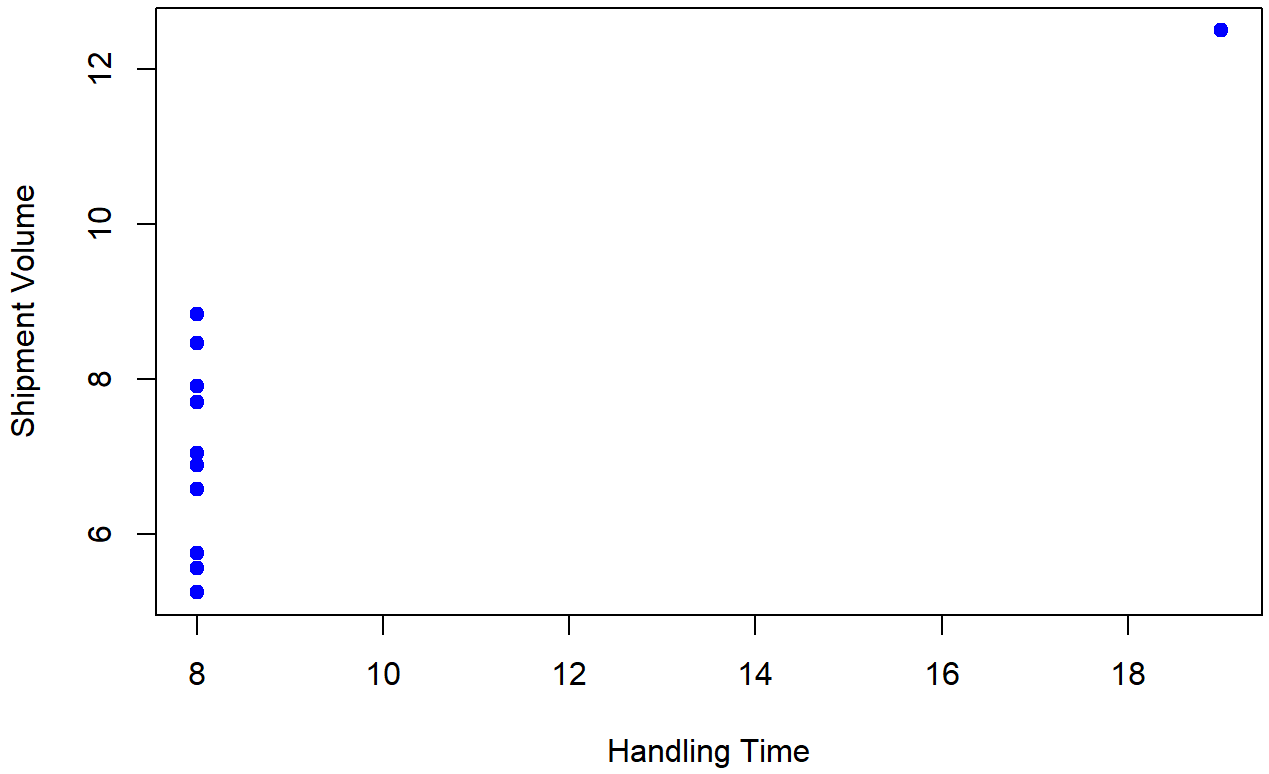



```
plot(warehouse_data$Handling_Time_C, warehouse_data$Shipment_Volume_C,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
title(main = "Warehouse C", col.main = title_color)
```




```
plot(warehouse_data$Handling_Time_D, warehouse_data$Shipment_Volume_D,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
title(main = "Warehouse D", col.main = title_color)
```

Warehouse D




```
# Making a few fitted linear models
model_A <- lm(Shipment_Volume_A ~ Handling_Time_A, data = warehouse_data)
model_B <- lm(Shipment_Volume_B ~ Handling_Time_B, data = warehouse_data)
model_C <- lm(Shipment_Volume_C ~ Handling_Time_C, data = warehouse_data)
model_D <- lm(Shipment_Volume_D ~ Handling_Time_D, data = warehouse_data)
```

```
# I was curious to see the statisical summary
summary(model_A)
```

```
##
## Call:
## lm(formula = Shipment_Volume_A ~ Handling_Time_A, data = warehouse_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.92127 -0.45577 -0.04136  0.70941  1.83882
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.0001     1.1247   2.667  0.02573 *
## Handling_Time_A   0.5001     0.1179   4.241  0.00217 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.237 on 9 degrees of freedom
## Multiple R-squared:  0.6665, Adjusted R-squared:  0.6295
## F-statistic: 17.99 on 1 and 9 DF,  p-value: 0.00217
```

```
summary(model_B)
```

```
##
## Call:
## lm(formula = Shipment_Volume_B ~ Handling_Time_B, data = warehouse_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9009 -0.7609  0.1291  0.9491  1.2691
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.001      1.125   2.667  0.02576 *
## Handling_Time_B    0.500      0.118   4.239  0.00218 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.237 on 9 degrees of freedom
## Multiple R-squared:  0.6662, Adjusted R-squared:  0.6292
## F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002179
```

```
summary(model_C)
```

```
##
## Call:
## lm(formula = Shipment_Volume_C ~ Handling_Time_C, data = warehouse_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1586 -0.6146 -0.2303  0.1540  3.2411
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.0025     1.1245   2.670  0.02562 *
## Handling_Time_C   0.4997     0.1179   4.239  0.00218 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.236 on 9 degrees of freedom
## Multiple R-squared:  0.6663, Adjusted R-squared:  0.6292
## F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002176
```

```
summary(model_D)
```

```
##
## Call:
## lm(formula = Shipment_Volume_D ~ Handling_Time_D, data = warehouse_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.751 -0.831  0.000  0.809  1.839
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.0017     1.1239   2.671  0.02559 *
## Handling_Time_D   0.4999     0.1178   4.243  0.00216 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.236 on 9 degrees of freedom
## Multiple R-squared:  0.6667, Adjusted R-squared:  0.6297
## F-statistic:    18 on 1 and 9 DF,  p-value: 0.002165
```

```r
# For better clarity I need some regression lines my plots with red
par(mfrow = c(2, 2)) # Reset to 2x2 grid for plots

plot(warehouse_data$Handling_Time_A, warehouse_data$Shipment_Volume_A,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
abline(model_A, col = "red")
title(main = "Warehouse A with Regression Line", col.main = title_color)

plot(warehouse_data$Handling_Time_B, warehouse_data$Shipment_Volume_B,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
abline(model_B, col = "red")
title(main = "Warehouse B with Regression Line", col.main = title_color)

plot(warehouse_data$Handling_Time_C, warehouse_data$Shipment_Volume_C,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
abline(model_C, col = "red")
title(main = "Warehouse C with Regression Line", col.main = title_color)

plot(warehouse_data$Handling_Time_D, warehouse_data$Shipment_Volume_D,
     xlab = "Handling Time", ylab = "Shipment Volume", col = point_color, pch = 19)
abline(model_D, col = "red")
title(main = "Warehouse D with Regression Line", col.main = title_color)
```