Regression Analysis: Distance from CO Source and Asthma Severity

Marc

2025-10-04

Introduction

Air pollution, especially **carbon monoxide (CO)**, has been linked to a variety of negative respiratory health outcomes.
In this analysis, we test whether **distance from a CO source** can predict **asthma severity** among households.

We expect that people living **closer to a CO source** may experience **more severe asthma symptoms** than those living farther away.

Data

The following table presents the study data:

Household	Distance_from_CO (m)	Asthma_Severity (1–10)
1	50	9.0
2	75	8.5
3	100	8.0
4	125	7.4
5	150	7.0
6	200	6.5
7	250	6.0
8	300	5.5
9	350	5.2
10	400	5.0

Regression Model

We will fit a **simple linear regression model** to examine whether *distance from CO source* predicts *asthma severity*.

Model:

$$\text{Asthma Severity} = \beta_0 + \beta_1(\text{Distance from CO}) + \epsilon$$

Create the dataset

```r
household <- 1:10
distance <- c(50, 75, 100, 125, 150, 200, 250, 300, 350, 400)
asthma <- c(9.0, 8.5, 8.0, 7.4, 7.0, 6.5, 6.0, 5.5, 5.2, 5.0)

data <- data.frame(Household = household,
                   Distance_from_CO = distance,
                   Asthma_Severity = asthma)

# Fit linear regression model
model <- lm(Asthma_Severity ~ Distance_from_CO, data = data)

# Show model summary
summary(model)
```

```
##
## Call:
## lm(formula = Asthma_Severity ~ Distance_from_CO, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3767 -0.2558 -0.0600  0.2275  0.4900
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       9.0766667  0.2169961   41.83 1.18e-10 ***
## Distance_from_CO -0.0113333  0.0009415  -12.04 2.09e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3411 on 8 degrees of freedom
## Multiple R-squared:  0.9477, Adjusted R-squared:  0.9411
## F-statistic: 144.9 on 1 and 8 DF,  p-value: 2.093e-06
```

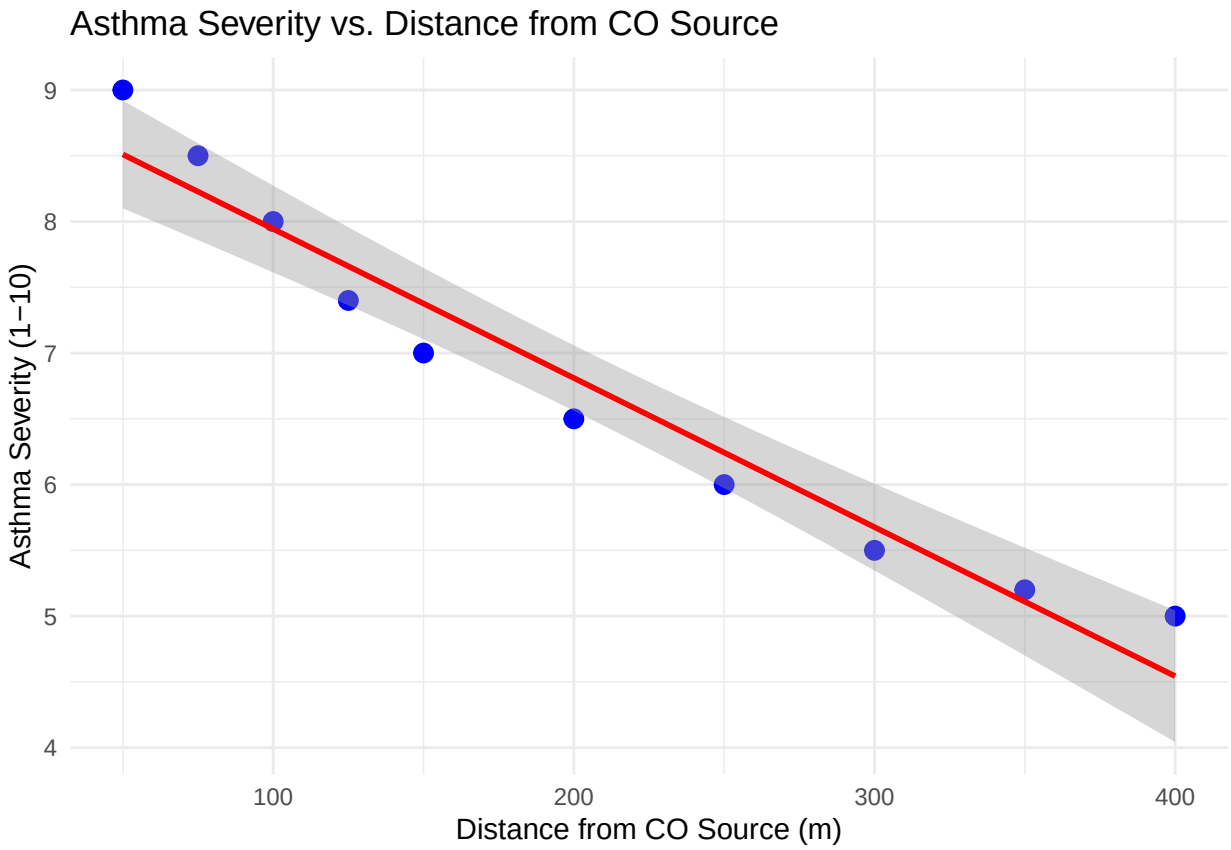
```r
coef(model)
```

```
##      (Intercept) Distance_from_CO
##       9.07666667      -0.01133333
```

```r
library(ggplot2)

ggplot(data, aes(x = Distance_from_CO, y = Asthma_Severity)) +
  geom_point(size = 3, color = "blue") +
  geom_smooth(method = "lm", se = TRUE, color = "red") +
  labs(title = "Asthma Severity vs. Distance from CO Source",
       x = "Distance from CO Source (m)",
       y = "Asthma Severity (1-10)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```


Asthma Severity vs. Distance from CO Source
Asthma Severity (1–10)
9
8
7
6
5
4
100
200
300
400
Distance from CO Source (m)

“‘