

Pollution In-Class

Bhavik Singhal

2025-09-04

```
pollution <- read.csv("ny_pollution.csv.gz")

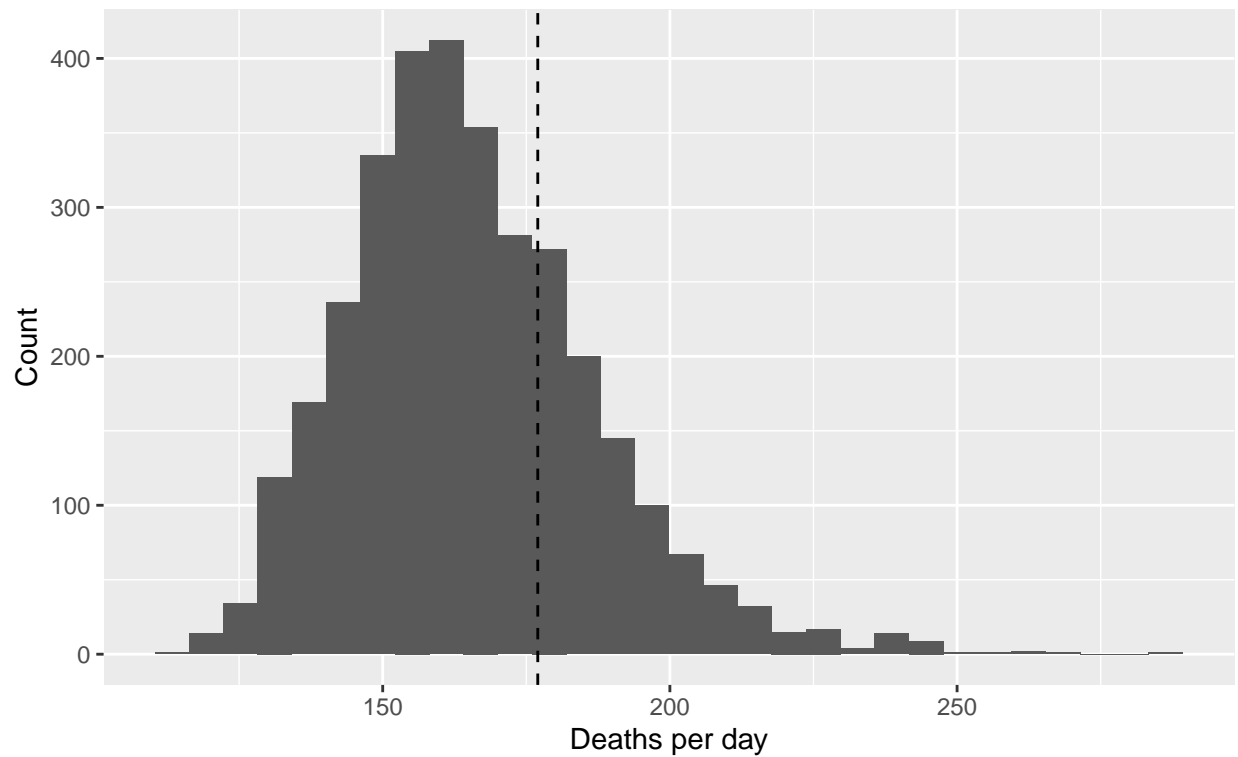
q3 <- quantile(pollution$death, 0.75, na.rm = TRUE)
q3

## 75%
## 177

ggplot(pollution, aes(x = death)) +
  geom_histogram(bins = 30) +
  geom_vline(xintercept = q3, linetype = "dashed") +
  labs(
    title = "Distribution of Daily Deaths",
    subtitle = paste0("75th percentile = ", round(q3, 1)),
    x = "Deaths per day",
    y = "Count"
  )
```

Distribution of Daily Deaths

75th percentile = 177



Statement: The 75th percentile of daily deaths in New York from 1997-2005 is 177

Supporting Premise: I used a histogram and graphed the 75th percentile.