

FA1

Espiritu, Joseph Raphael, Harneyyer, Clores

2026-01-26

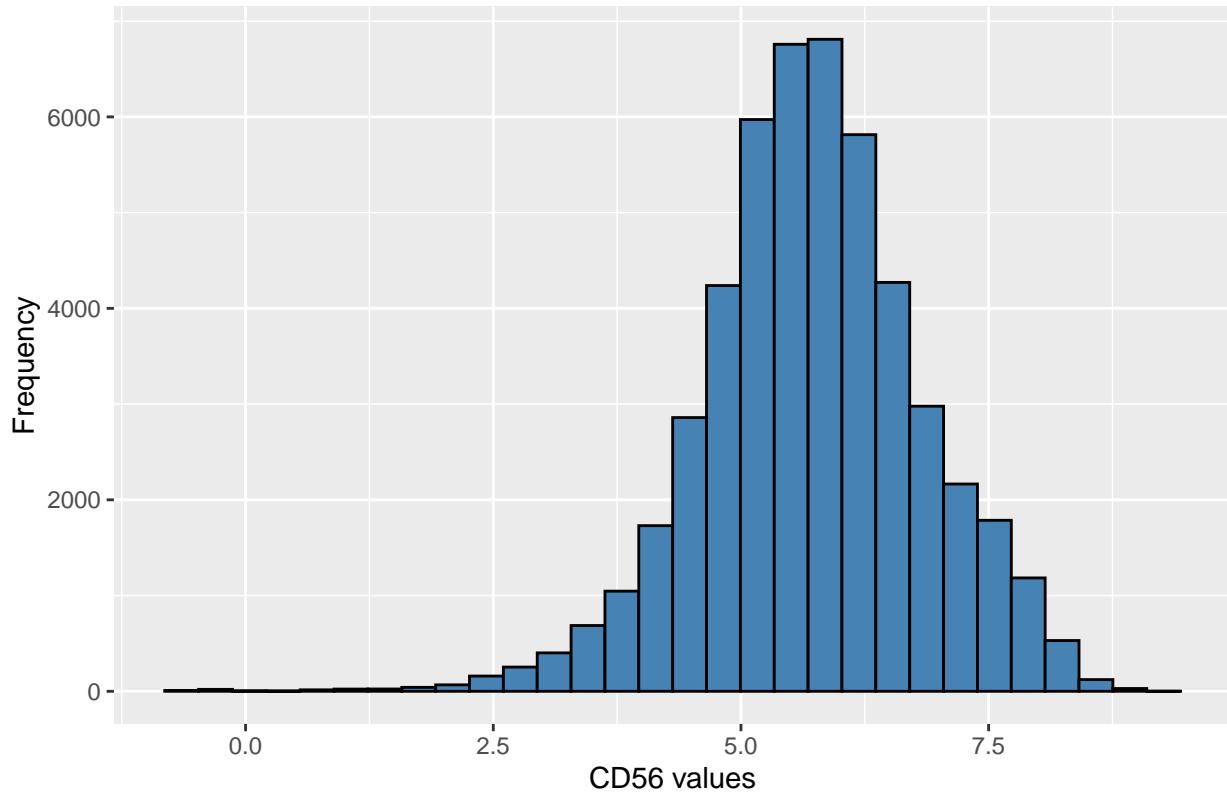
```
knitr::opts_chunk$set(echo = TRUE, message = FALSE, warning = FALSE)
library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.6
## v forcats   1.0.1     v stringr   1.6.0
## v ggplot2   4.0.1     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr    1.3.1
## v purrr    1.2.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
library(here)

## here() starts at D:/JR Studies/College - FEU/EDA
cytof <- read.csv("cytof_one_experiment.csv")

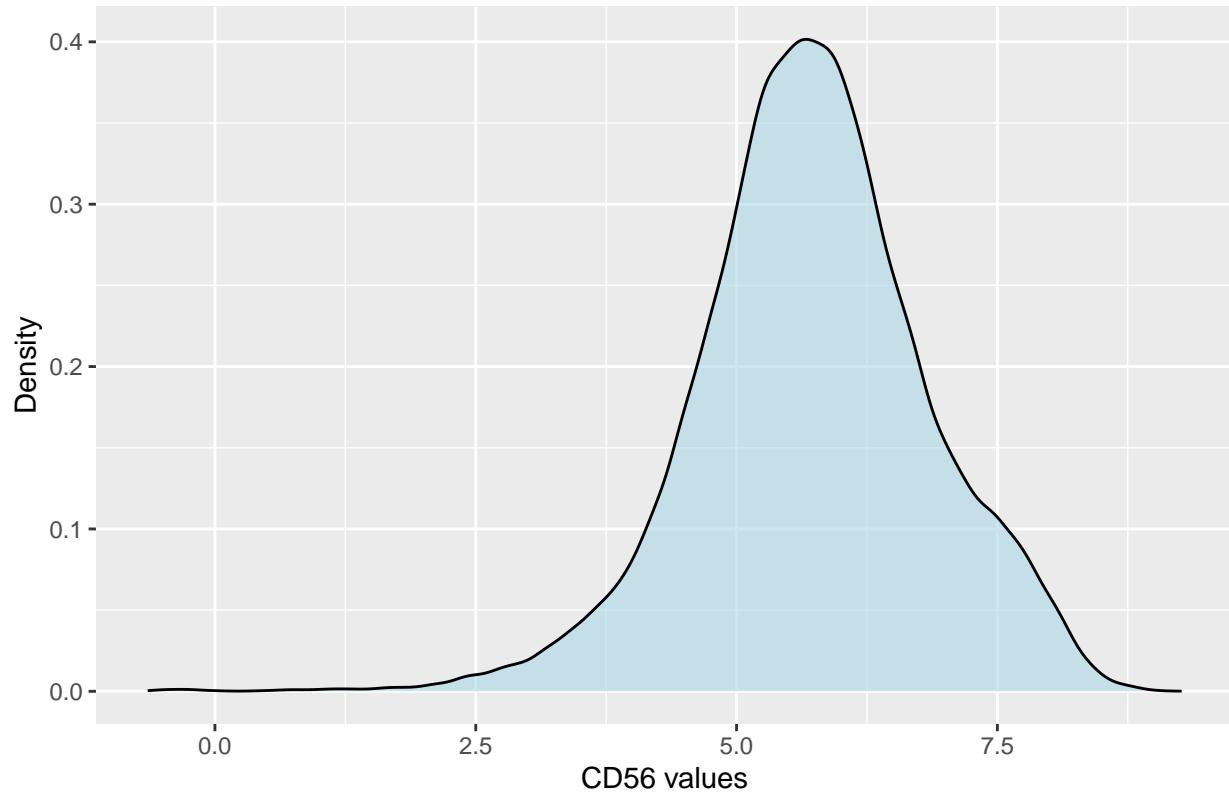
ggplot(cytof, aes(x = CD56)) +
  geom_histogram(bins = 30, fill = "steelblue", color = "black") +
  labs(
    title = "Histogram of CD56 Expression",
    x = "CD56 values",
    y = "Frequency"
  )
```

Histogram of CD56 Expression



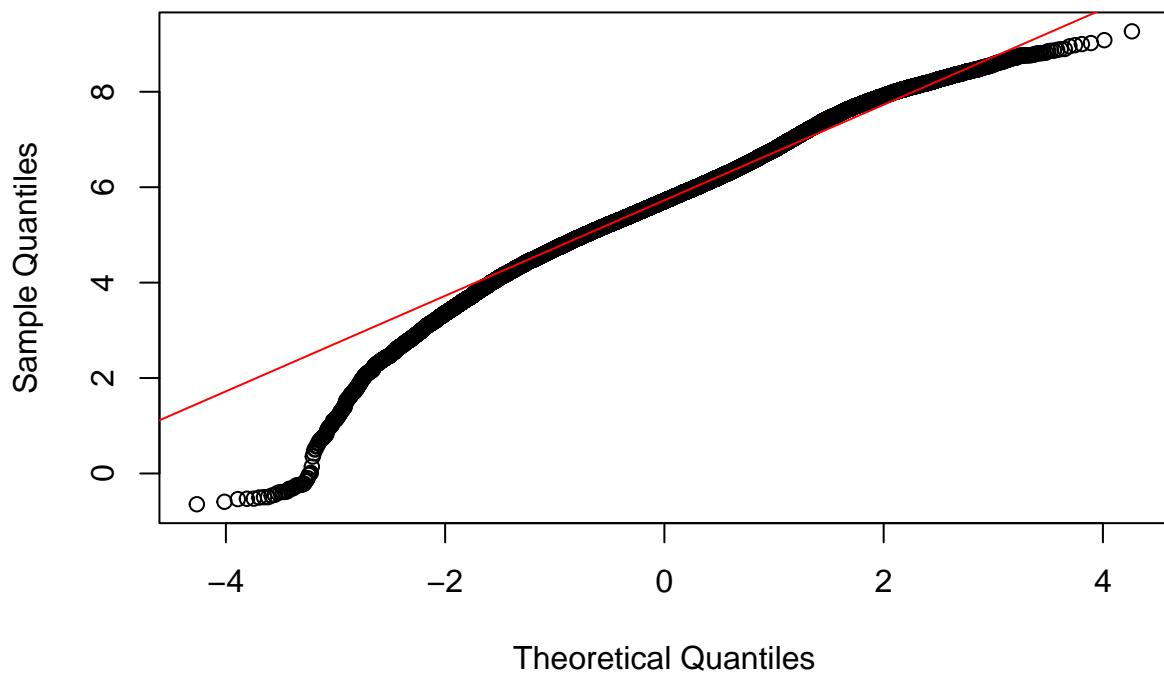
```
ggplot(cytof, aes(x = CD56)) +  
  geom_density(fill = "lightblue", alpha = 0.6) +  
  labs(  
    title = "Density Estimate of CD56 Expression",  
    x = "CD56 values",  
    y = "Density"  
)
```

Density Estimate of CD56 Expression



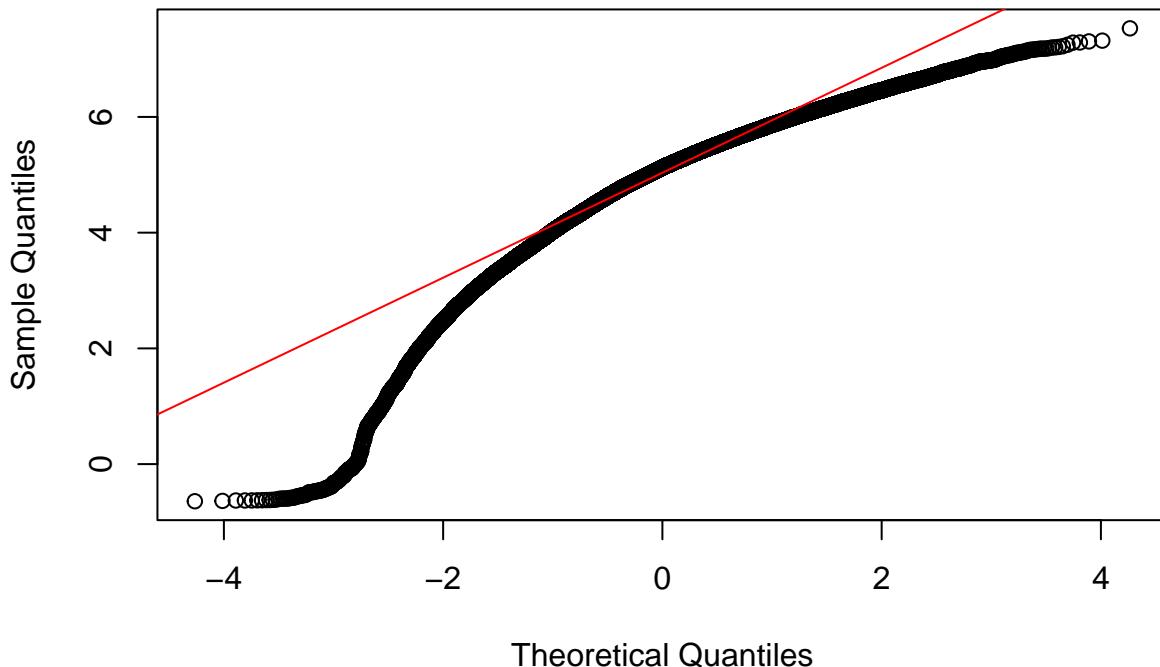
```
### Q-Q Plot for CD56
qqnorm(cytof$CD56, main = "Q-Q Plot of CD56")
qqline(cytof$CD56, col = "red")
```

Q-Q Plot of CD56



```
### Q-Q Plot for CD16
qqnorm(cytof$CD16, main = "Q-Q Plot of CD16")
qqline(cytof$CD16, col = "red")
```

Q-Q Plot of CD16



Results Distribution of CD56 Expression

The distribution of CD56 expression was examined using a histogram and a kernel density estimate. The distribution was unimodal, with most values concentrated in the midrange (approximately 5–6). Frequencies decreased toward both extremes, with a longer lower tail, indicating the presence of relatively low expression values. This pattern suggests that the majority of observations exhibited moderate CD56 expression, while fewer observations showed very low or very high expression levels.

The density estimate supported these findings, showing a single prominent peak and a slight asymmetry in the distribution. The concentration of values around the central peak, combined with gradual decline toward the tails, indicates that CD56 expression does not follow a perfectly normal distribution and exhibits mild skewness.

Comparison of CD56 and CD16 Distributions

Quantile-quantile plots were used to assess the distributions of CD56 and CD16 relative to a theoretical normal distribution. For CD56, deviations from the reference line were observed at both tails, particularly in the lower tail, indicating non-normality and heavier tails than expected under normal assumptions. CD16 displayed a similar pattern of deviation; however, differences in curvature suggested variation in dispersion and tail behavior between the two variables.

Although both markers showed departures from normality, the extent of deviation differed. CD56 demonstrated greater variability and more extreme values compared to CD16, suggesting differences in the underlying distributional properties of the two markers.

Summary

Overall, both CD56 and CD16 exhibited non-normal distributions characterized by skewness and heavy tails. While their distributional shapes were broadly similar, differences in spread and tail behavior indicate distinct expression patterns across the two markers.