Appendix E

Exploration of the "Number of reviews" variable

The 'number of reviews' variable shows a strong right skew, with a skewness of 9.75 and kurtosis of 135.21, indicating most products have few reviews, while a small number have many. The data's distribution suggests a logarithmic transformation could normalize it, improving analytical outcomes.

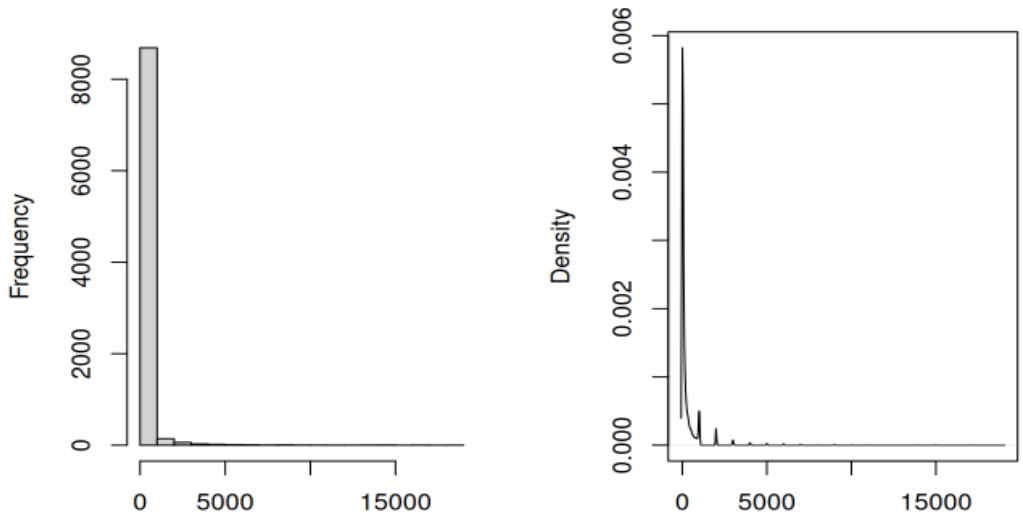
```{r}
# Skewness and kurtosis
skewness(sephora$number_of_reviews)
kurtosis(sephora$number_of_reviews)
```

Skewness: 9.75 **Kurtosis:** 135.21

Histogram and density for value price variable

```{r}
# Exploration number_of_reviews variable
# Set up the plotting layout
par(mfrow = c(1, 2))

# Plot histogram and density hist(sephora$number_of_reviews)
plot(density(sephora$number_of_reviews))
```



QQ and box plots for value variable

```{r}
# QQ plot and boxplot
qqnorm(sephora$number_of_reviews, main = "QQ Plot for value price varible")
qqline(sephora$number_of_reviews)

boxplot(sephora$number_of_reviews, horizontal = TRUE)
```