Anwendung 1: Histogramm

Dirk Seidensticker/Clemens Schmid

7. Juli 2017

Anwendung 1: Histogramm

ggplot2 und die Daten laden:

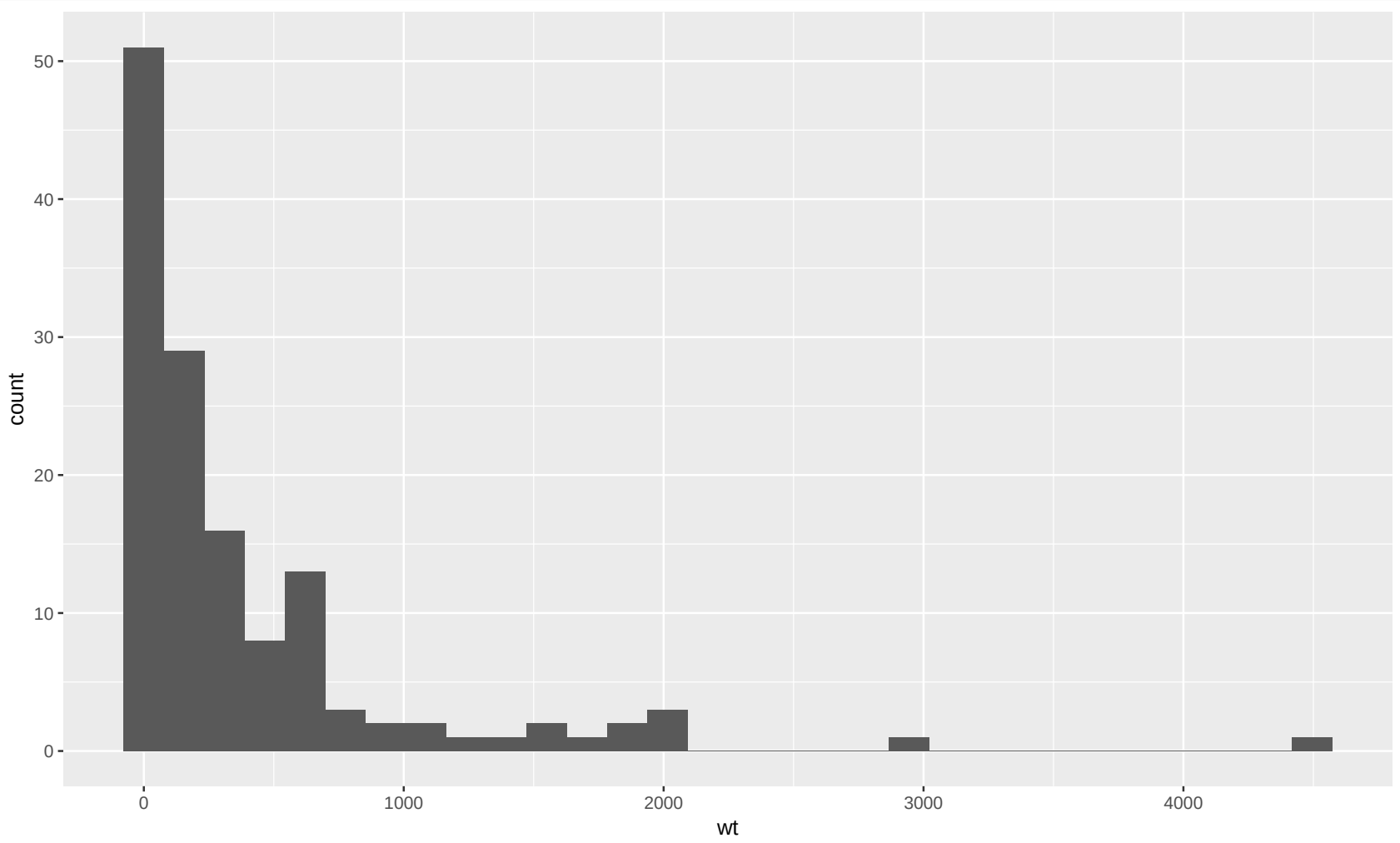
```
library(ggplot2)
```

```
df <- read.csv("https://raw.githubusercontent.com/kacebe/AtlantGIS/master/table
```

Anwendung 1: Histogramm

Einfaches Histogramm

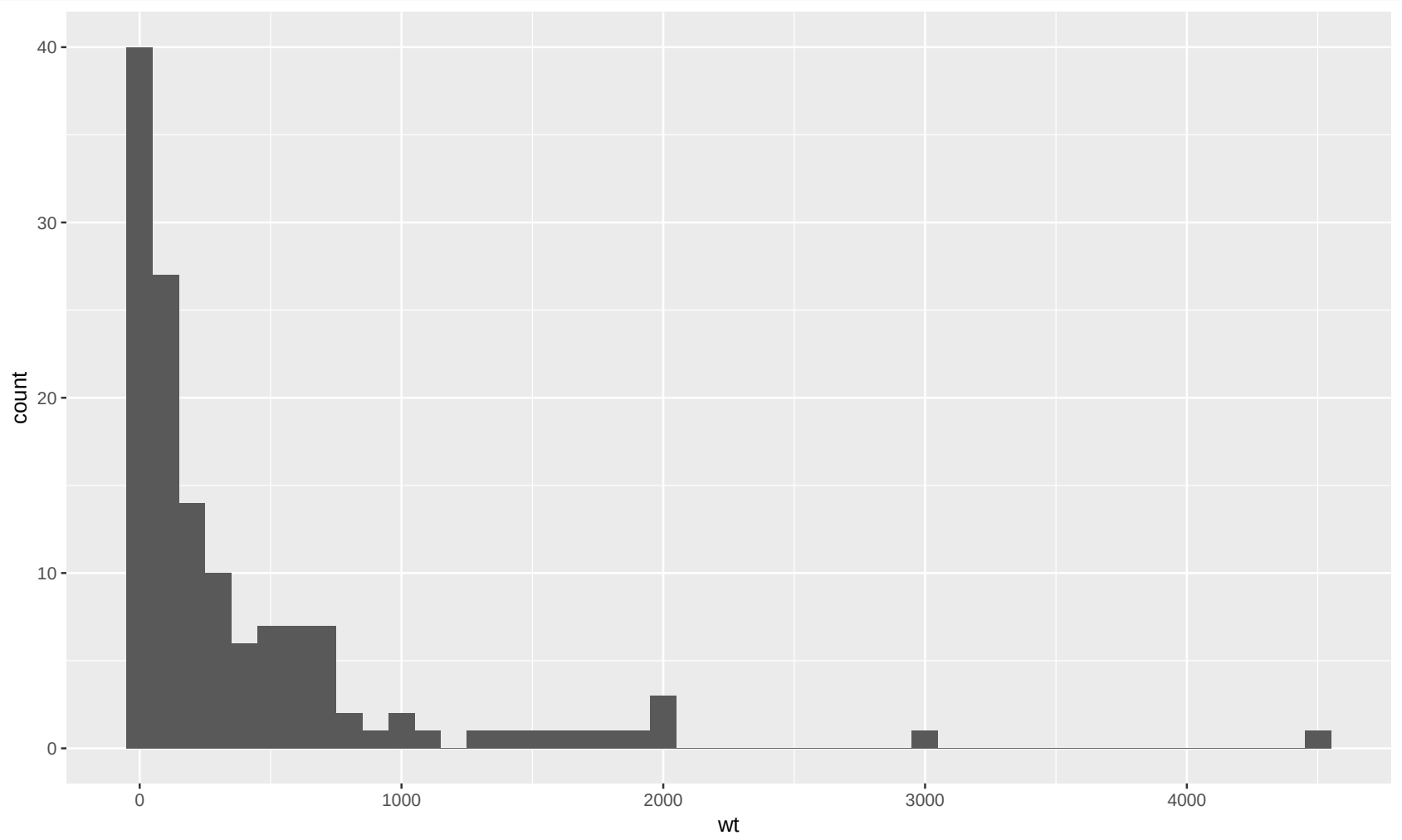
```
ggplot(df, aes(wt)) +
  geom_histogram()
```



Anwendung 1: Histogramm

Klassenbreite

```
ggplot(df, aes(wt)) +
  geom_histogram(binwidth = 100)
```



Anwendung 1: Histogramm

Exkurs: optimale Klassenbreite via Freedman–Diaconis-Regel

siehe:

- `https://stats.stackexchange.com/questions/798/calculating-optimal-number-of-bins-in-a-histogram`
- `https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis_rule`

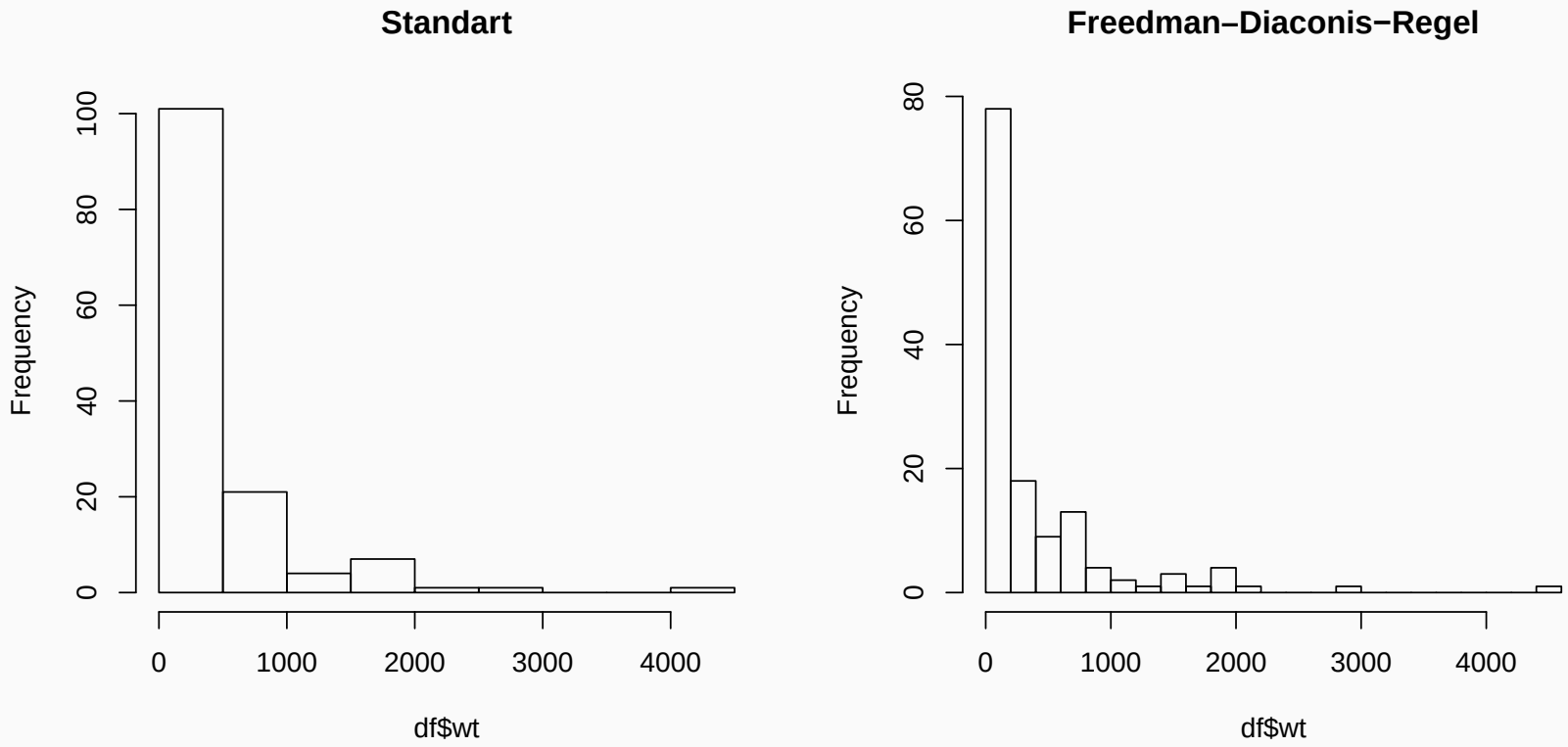
$$Binsize = 2\frac{IRQ(x)}{\sqrt[3]{n}}$$

$IRQ(x) =$ Interquartilabstand

Anwendung 1: Histogramm

Exkurs: optimale Klassenbreite via Freedman–Diaconis-Regel

```
layout(matrix(c(1,2),1,2,byrow=TRUE))
hist(df$wt, main = "Standart")
hist(df$wt, breaks="FD", main = "Freedman-Diaconis-Regel")
```

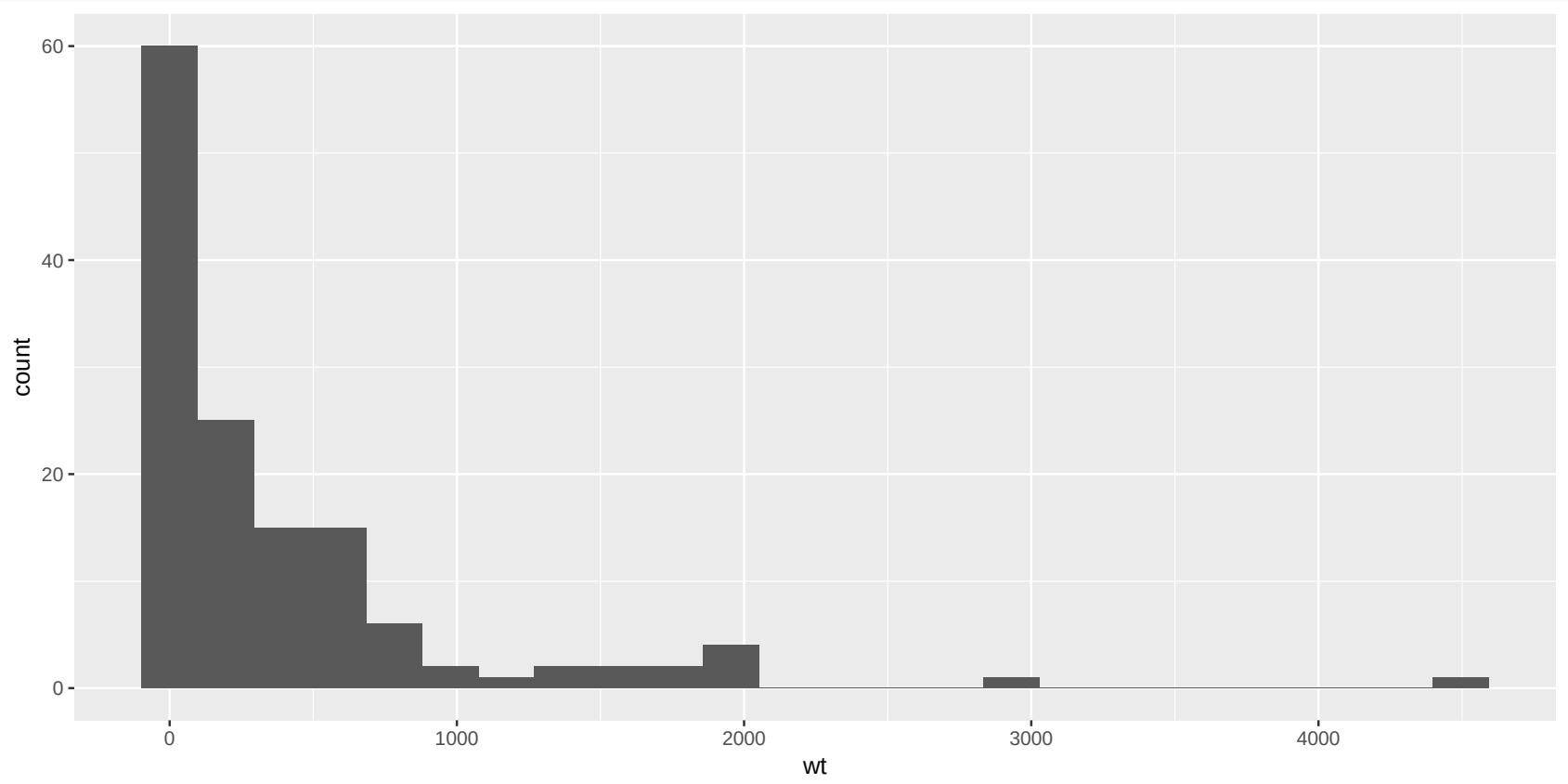


Anwendung 1: Histogramm

Exkurs: optimale Klassenbreite via Freedman–Diaconis-Regel

```
bw <- diff(range(df$wt)) / (2 * IQR(df$wt) / length(df$wt)^(1/3))

ggplot(df, aes(wt)) +
  geom_histogram(bins = bw)
```

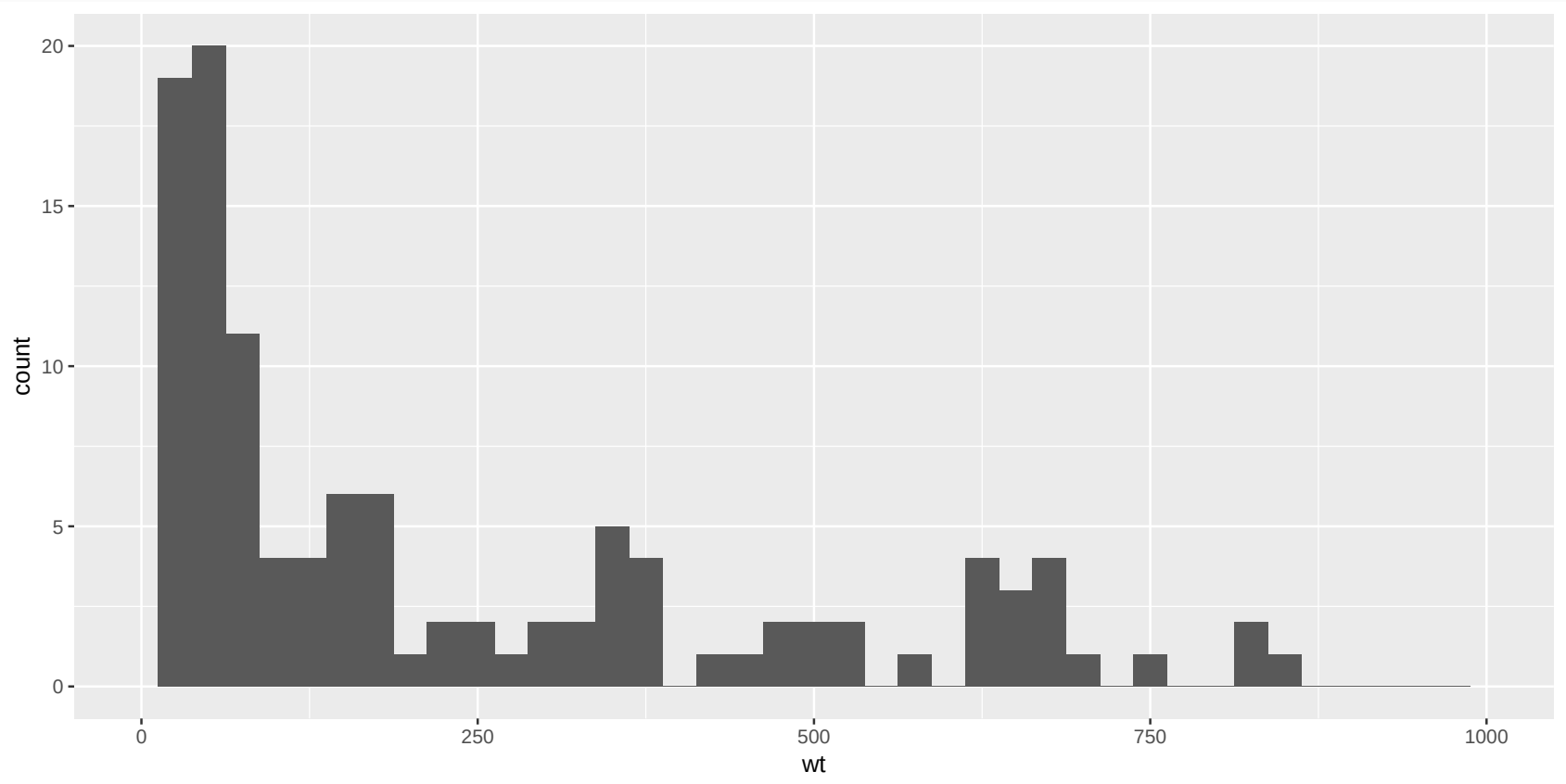


Anwendung 1: Histogramm

Bildausschnitt anpassen

```
ggplot(df, aes(wt)) +
  geom_histogram(binwidth = 25) +
  xlim(0, 1000)
```

```
## Warning: Removed 14 rows containing non-finite values (stat_bin).
```

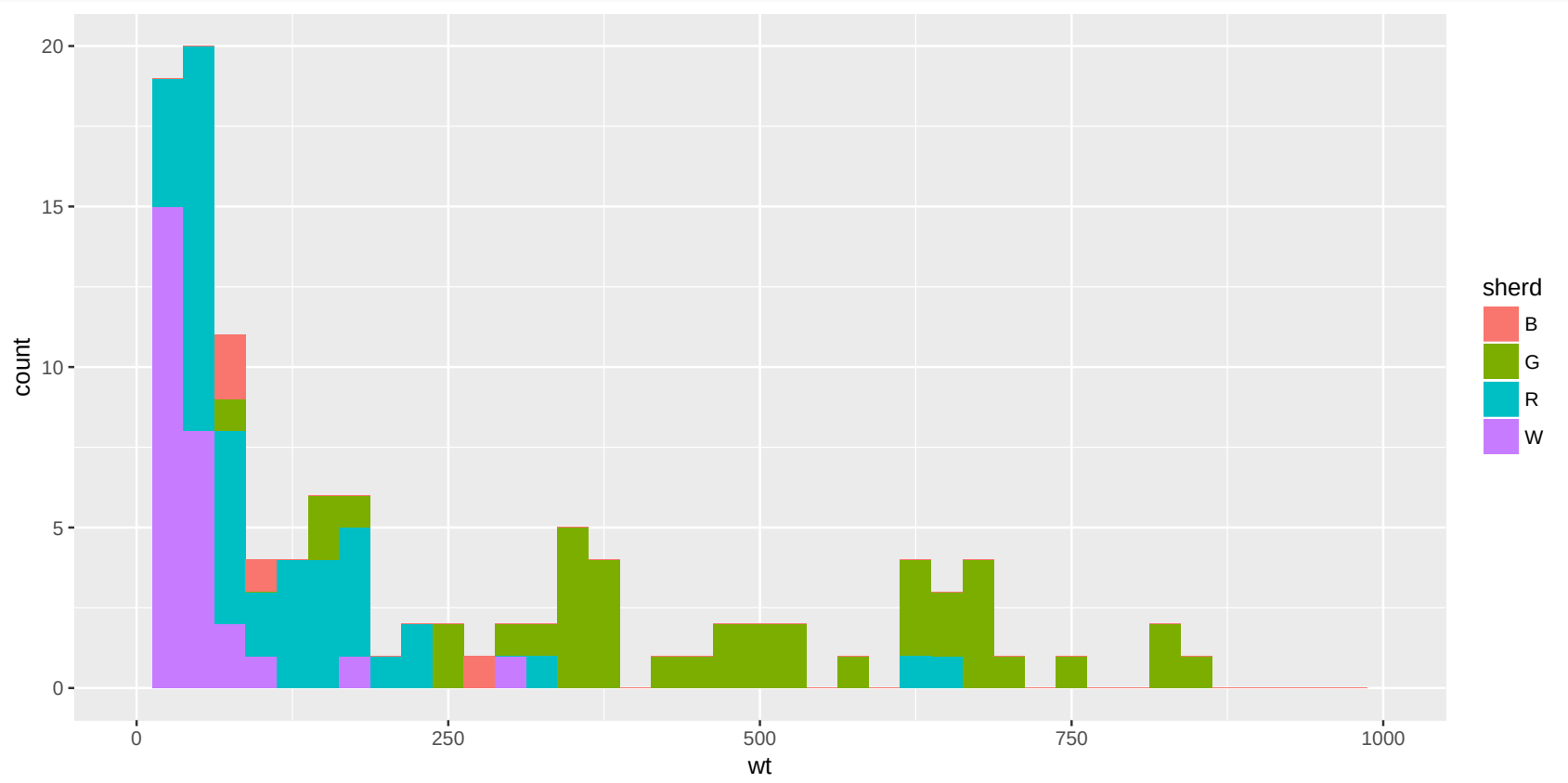


Anwendung 1: Histogramm

Gefülltes Balkendiagramm

```
ggplot(df, aes(wt, fill = sherd)) +
  geom_histogram(binwidth = 25) +
  xlim(0, 1000)
```

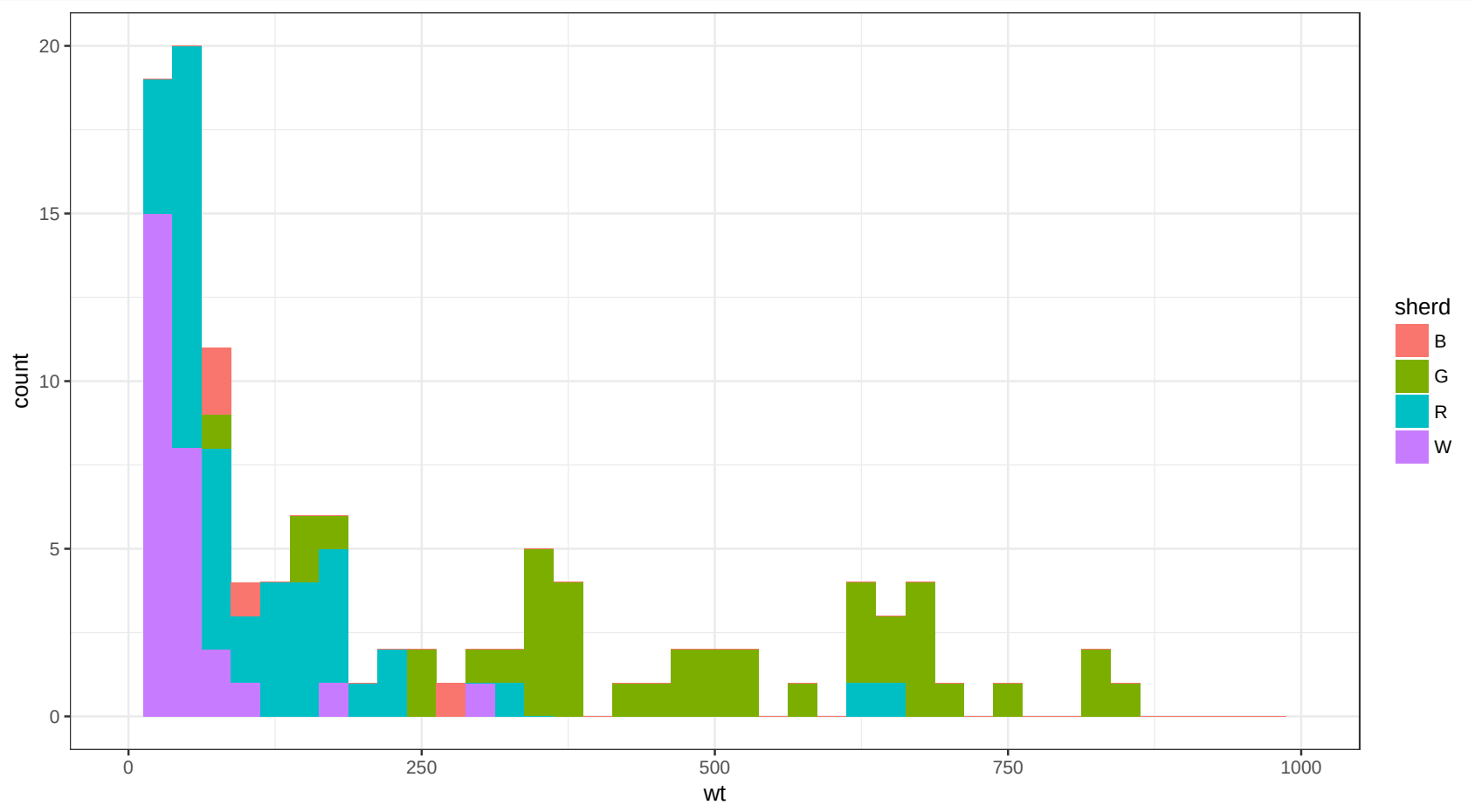
```
## Warning: Removed 14 rows containing non-finite values (stat_bin).
```



Anwendung 1: Histogramm

Alternative Ansicht

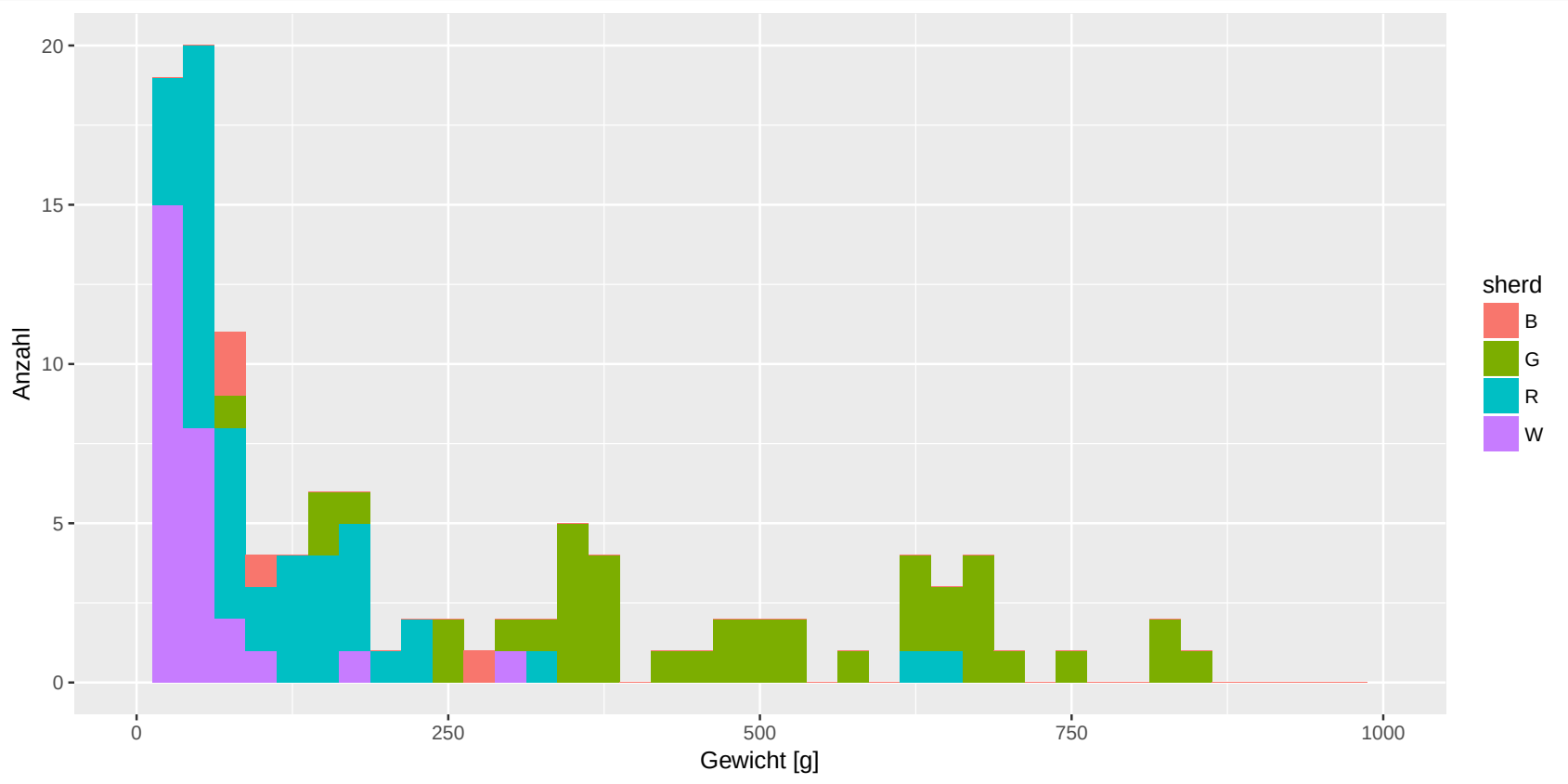
```
ggplot(df, aes(wt, fill = sherd)) +
  geom_histogram(binwidth = 25) +
  xlim(0, 1000) +
  theme_bw()
```



Anwendung 1: Histogramm

Achsenbeschriftung

```
ggplot(df, aes(wt, fill = sherd)) +
  geom_histogram(binwidth = 25) +
  xlim(0, 1000) +
  xlab("Gewicht [g]") +
  ylab("Anzahl")
```



Anwendung 1: Histogramm

Titel und Position der Legende

```
ggplot(df, aes(wt, fill = sherd)) +
  geom_histogram(binwidth = 25) +
  xlim(0, 1000) +
  xlab("Gewicht [g]") +
  ylab("Anzahl") +
  theme(legend.position = c(1,1), legend.justification = c(1,1))
```