```
```{r setup, include=FALSE}
library(tidyverse)
library(naniar)
library(visdat)
Billionaires <- read.csv("forbes_billionaires.csv", header = TRUE, na.strings = c("NA","N/A",""))
head(Billionaires)
glimpse(Billionaires)
sum(duplicated(Billionaires))
Billionaires %>%
 count(Name) %>%
 filter (n > 1)
sum(is.na(Billionaires))
miss_var_summary(Billionaires)
Billionaires %>%
 arrange(NetWorth) %>%
 vis_miss()
```{r cars}
Billionaires_cleaned <- na.omit(Billionaires) %>%
 distinct(Name, .keep_all = TRUE)
sum(is.na(Billionaires_cleaned))
```

```
Billionaires_cleaned %>%

count(Name) %>%

filter (n > 1)

print(Billionaires$NetWorth)
```

Billionaires_cleaned %>%

ggplot(aes(x = NetWorth)) + geom_histogram(binwidth = 30, color = "red", fill = "white") + labs(title =
"Net Worth")

...

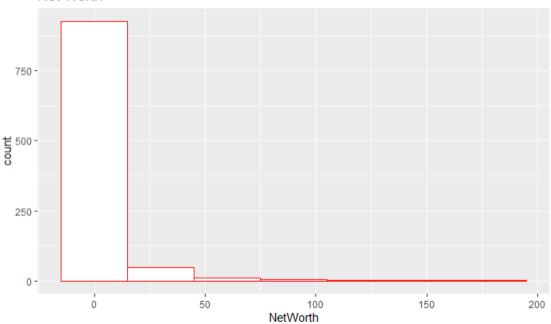
[757] 3.9 3.9 3.9 3.9 [757] 3.9 3.9 3.9 3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	•
3.9 3.9 3.9 3.9 3.9 [775] 3.8 3.8 3.8 3.8 3.8 3.8 3.8 3.8 3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	
[793] 3.8 3.8 3.8 3.8 3.8 3.7 3.7 3.7 3.7	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	
[811] 3.7 3.7 3.7 3.7 3.7 3.7 3.7 3.7	3.7	3.7	3.7	3.7	3.7	3.7	3.7	3.7	3.7	
[829] 3.7 3.7 3.6 3.6 3.6 3.6 3.6 3.6 3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	
[847] 3.6 3.6 3.6 3.6 3.5 3.5 3.5 3.5 3.5	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.5	
[865] 3.5 3.5 3.5 3.5 3.5 3.5 3.5 3.5 3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	
[883] 3.5 3.5 3.5 3.5 3.4 3.4 3.4 3.4 3.4	3.5	3.5	3.5	3.5	3.4	3.4	3.4	3.4	3.4	
[901] 3.4 3.4 3.4 3.4 3.4 3.4 3.4 3.4 3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.4	
[919] 3.4 3.4 3.4 3.4 3.3 3.3 3.3 3.3 3.3	3.4	3.4	3.3	3.3	3.3	3.3	3.3	3.3	3.3	
[937] 3.3 3.3 3.3 3.3 3.3 3.3 3.3 3.3 3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	
[955] 3.3 3.2 3.2 3.2 3.2 3.2 3.2 3.2 3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	
[973] 3.2 3.2 3.2 3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	
3.1 3.1 3.1 3.1 3.1 [991] 3.1 3.1 3.1 3.1 3.1						3.1				į
[reached getOption("max.print') (omicted	1 1/00	entrie	:5]					Ψ.











```{r pressure, echo=FALSE}

NetWorthLabel = c("1 - 88.5", "88.6 - 177")

 $Billionaires\_cleaned \$ NetWorth\_Group = cut(Billionaires\_cleaned \$ NetWorth, breaks = c(1, 88.5, Inf), labels = NetWorthLabel, right = FALSE)$ 

ggplot(data = Billionaires\_cleaned, mapping = aes(x = NetWorth\_Group, y = Age)) + geom\_boxplot(alpha = 0, color = "red", fill = "white") + geom\_jitter(color = "blue", alpha = 0.5) + labs(title = "Forbes Billionaires 2021 by Age")

...

## Forbes Billionaires 2021 by Age

