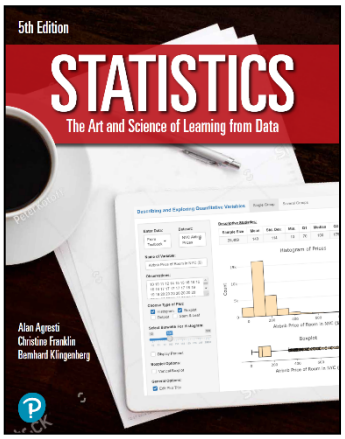


R Code for Examples in the book
"Statistics: The Art and Science of Learning from Data"
by Agresti, Franklin and Klingenberg, 5th edition

Chapter 2

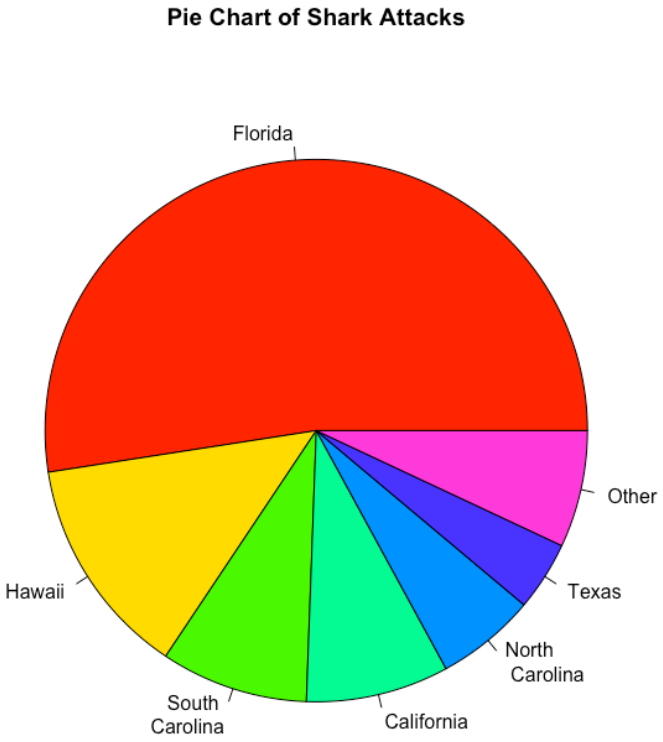
Example 3: Shark Attacks in the U.S. – Bar and Pie Charts

Create dataset:

```
state <- c('Florida', 'Hawaii', 'South \n Carolina',
           'California', 'North \n Carolina', 'Texas', 'Other')
frequency <- c(203, 51, 34, 33, 23, 16, 27)
```

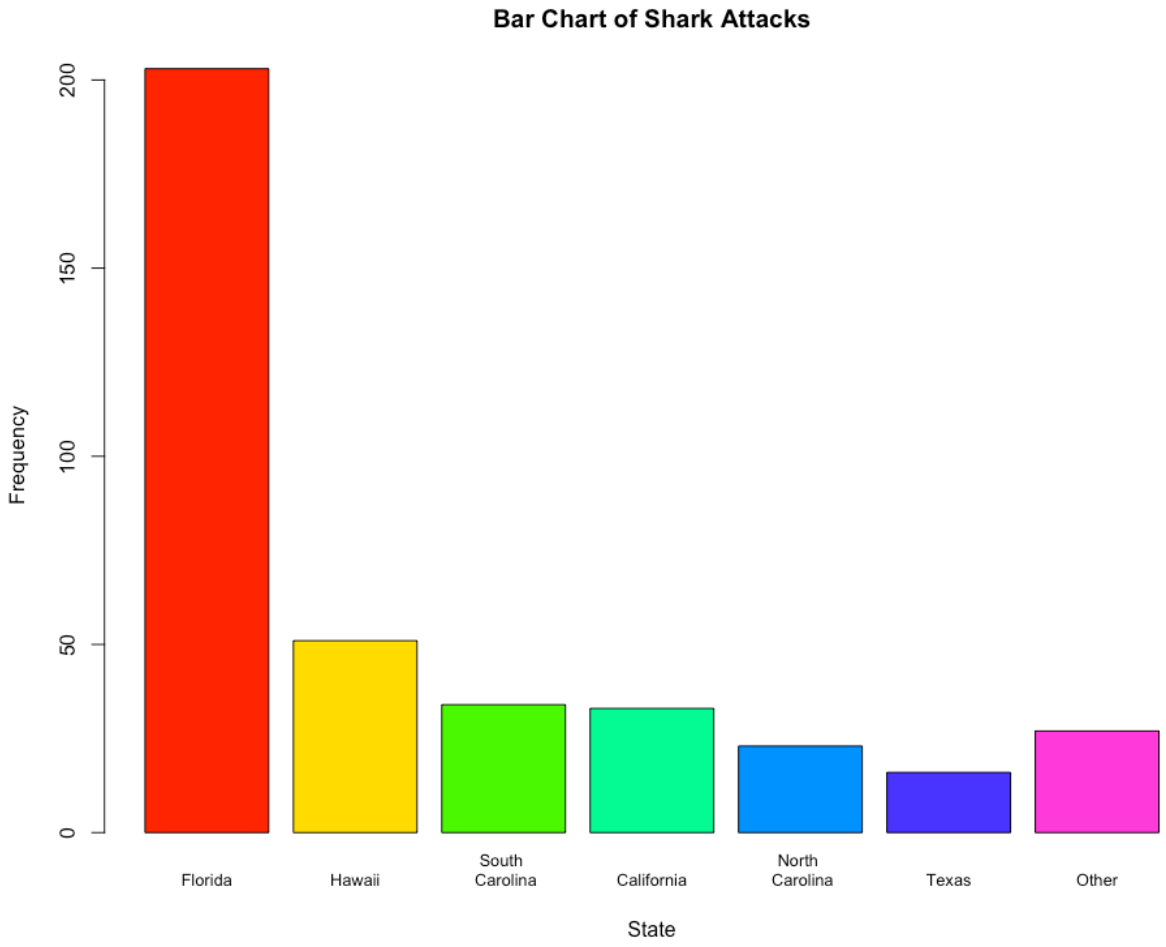
Create basic piechart:

```
pie(frequency, labels = state, col = rainbow(7),
    main = 'Pie Chart of Shark Attacks')
```



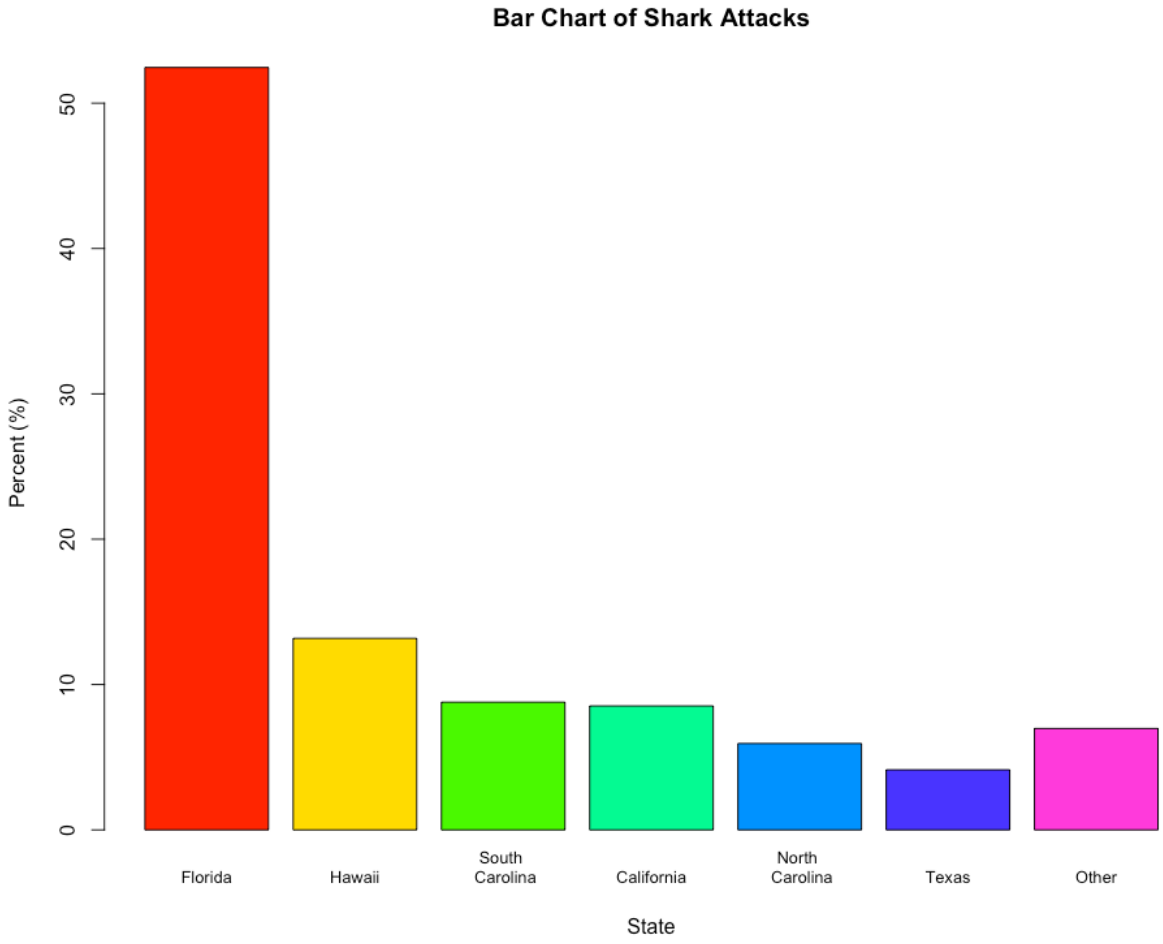
Create basic bar graph showing counts:

```
barplot(frequency, names.arg = state, cex.names=0.8, col=rainbow(7),
        main='Bar Chart of Shark Attacks',
        xlab='State', ylab='Frequency')
```



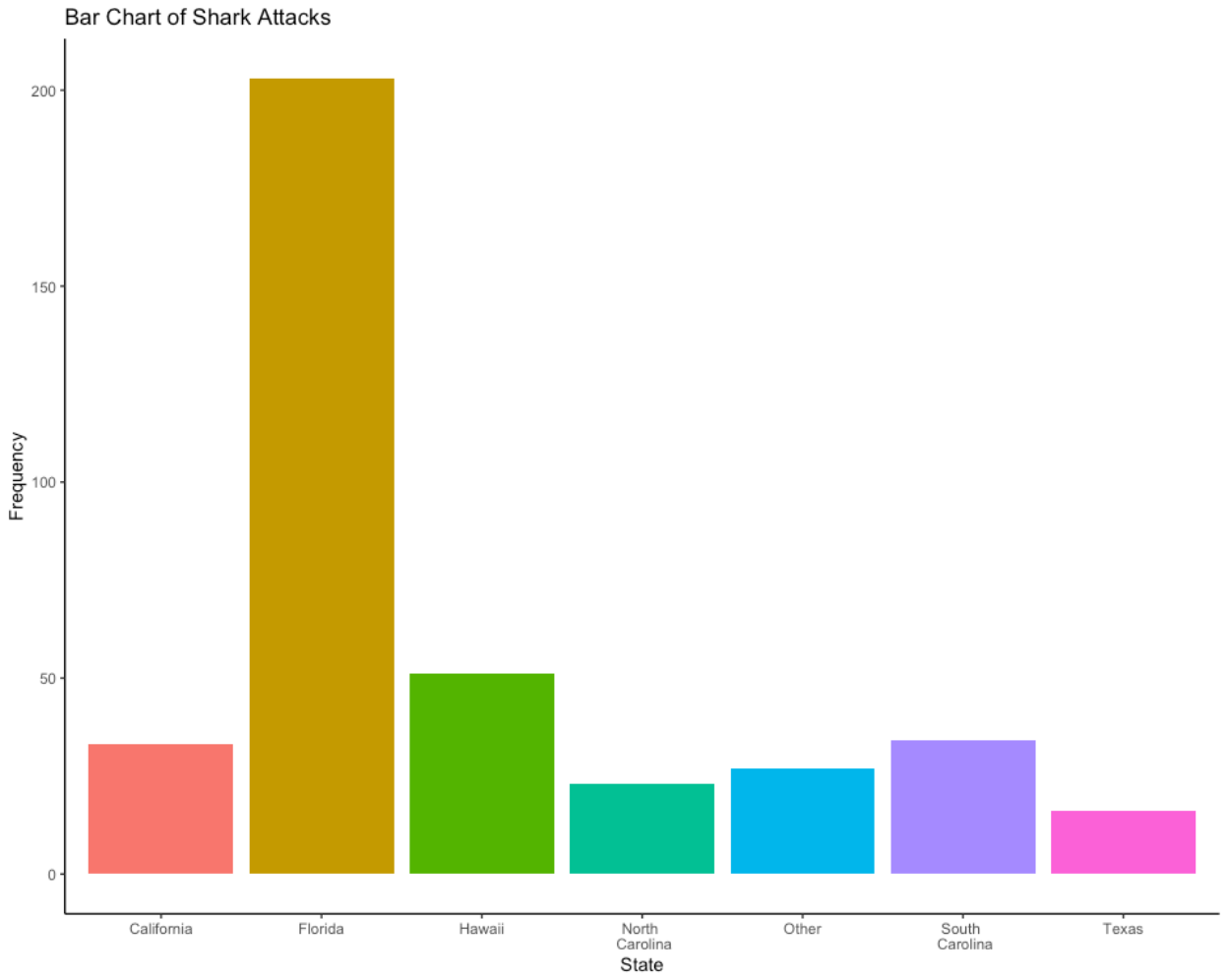
Create basic bar graph showing percentages:

```
percent <- 100 * (frequency / sum(frequency))
barplot(percent, names.arg = state, cex.names = 0.8, col = rainbow(7),
        main='Bar Chart of Shark Attacks',
        xlab='State', ylab='Percent (%)')
```



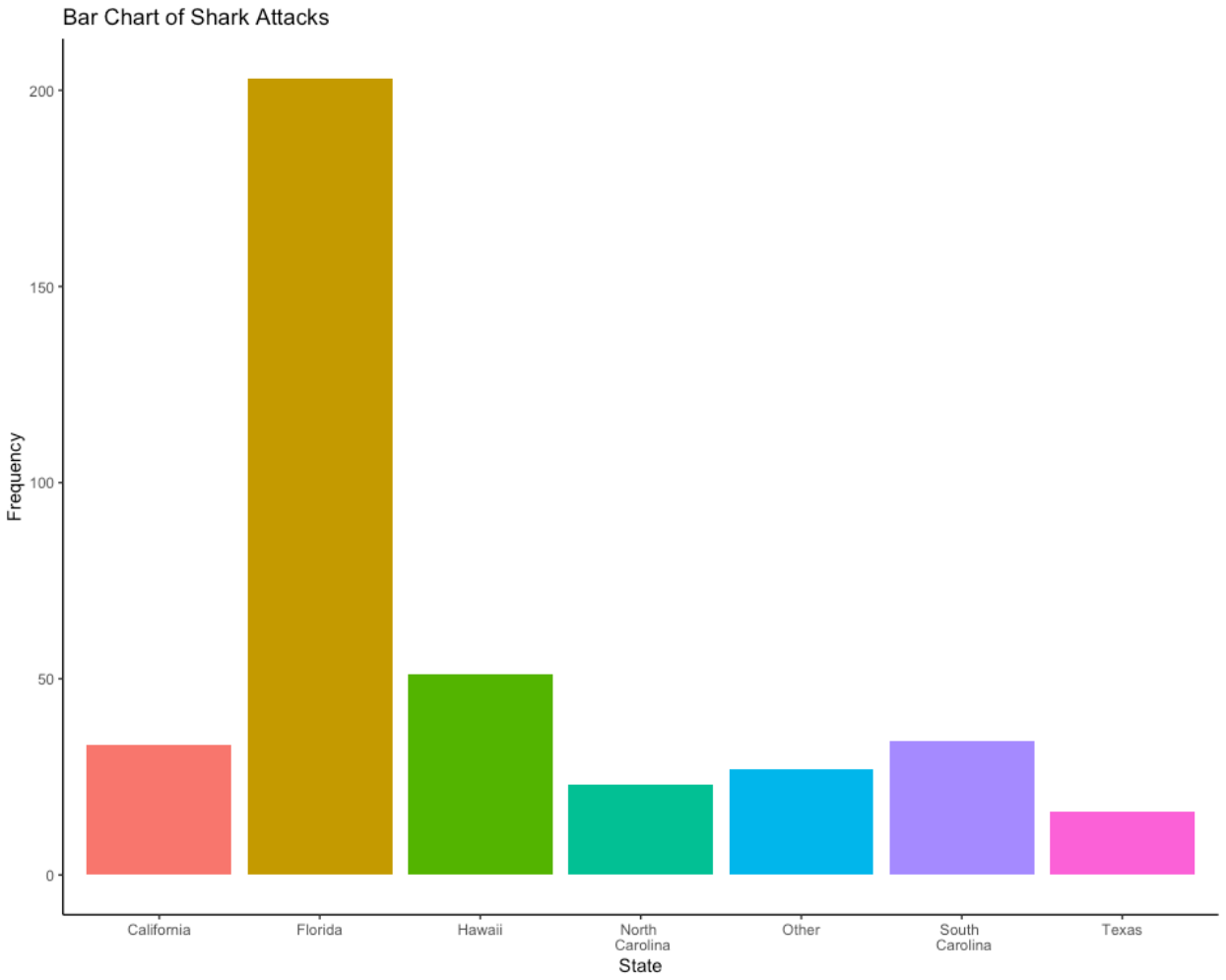
A better-looking bar graph can be obtained with the ggplot2 library. To install the ggplot2 library, use `install.packages('ggplot2')`

```
library(ggplot2)
plotdata = data.frame(state, frequency)
ggplot(data = plotdata,
       aes(x = state,
           y = frequency,
           fill = state)) +
  geom_col(show.legend = FALSE) +
  labs(title = 'Bar Chart of Shark Attacks',
       x = 'State', y = 'Frequency') +
  theme_classic()
```



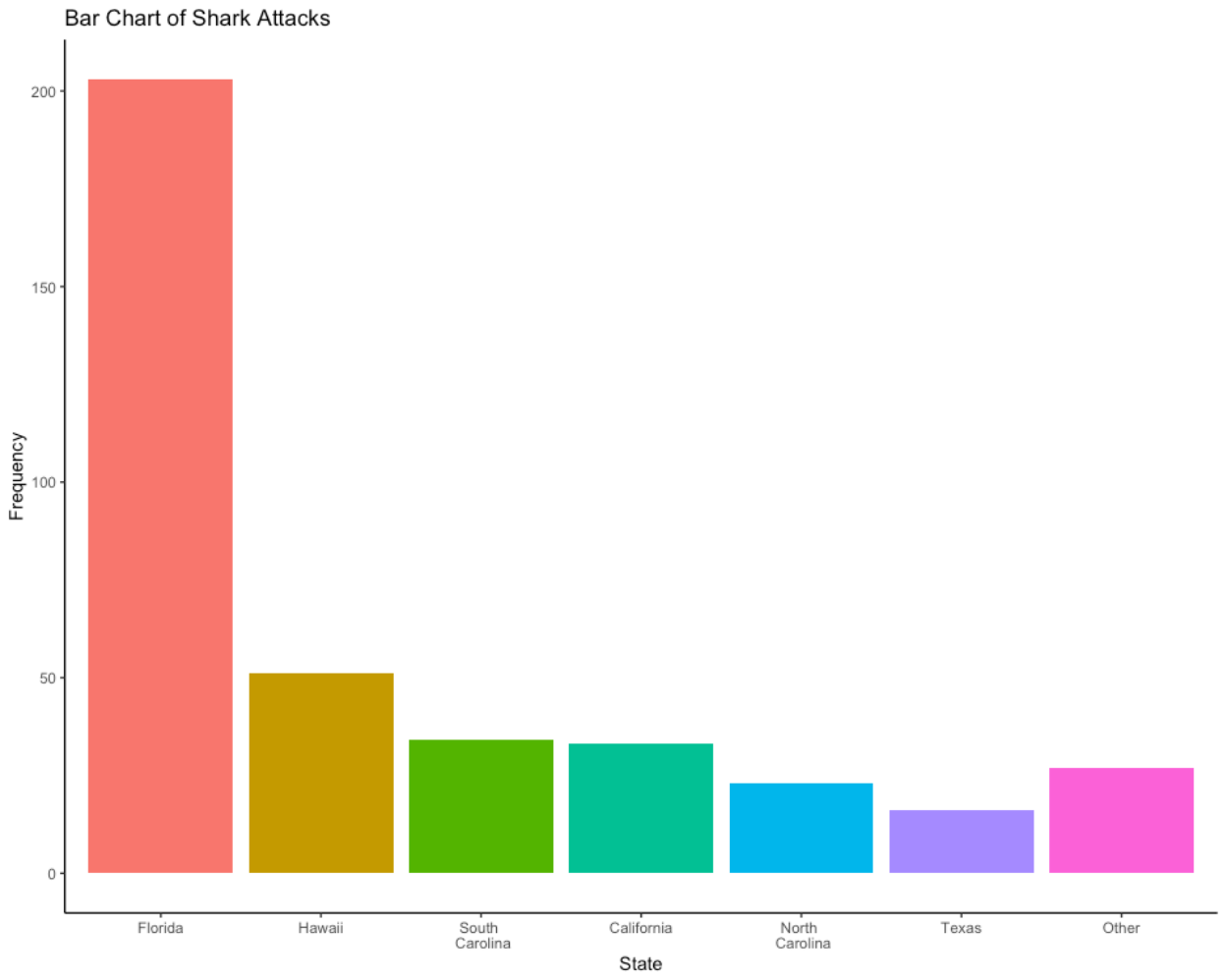
Or use the `geom_bar` function with `stat = identity`

```
ggplot(data = plotdata,
       aes(x = state,
           y = frequency,
           fill = state)) +
  geom_bar(stat = 'identity',
           show.legend = FALSE) +
  labs(title = 'Bar Chart of Shark Attacks',
       x = 'State', y = 'Frequency') +
  theme_classic()
```



To manually sort in decreasing order but with `Other` as the last category you can use the `mutate()` function from the dplyr package along with the `fct_relvel()` function from the forcats package. To install both packages, use `install.packages('dplyr')` and `install.packages('forcats')`.

```
library(dplyr)
library(forcats)
plotdata1 = plotdata %>%
  mutate(state =
           fct_relevel(state, c('Florida', 'Hawaii', 'South \n Carolina',
                                'California', 'North \n Carolina',
                                'Texas', 'Other')))
ggplot(data = plotdata1,
       aes(x = state,
           y = frequency,
           fill = state)) +
  geom_col(show.legend = FALSE) +
  labs(title = 'Bar Chart of Shark Attacks',
       x = 'State', y = 'Frequency') +
  theme_classic()
```



Note that the ggplot2, dplyr, and forcats packages are all conveniently included within the tidyverse package. So you can instead use `install.packages('tidyverse')` and then use the `library(tidyverse)` function.