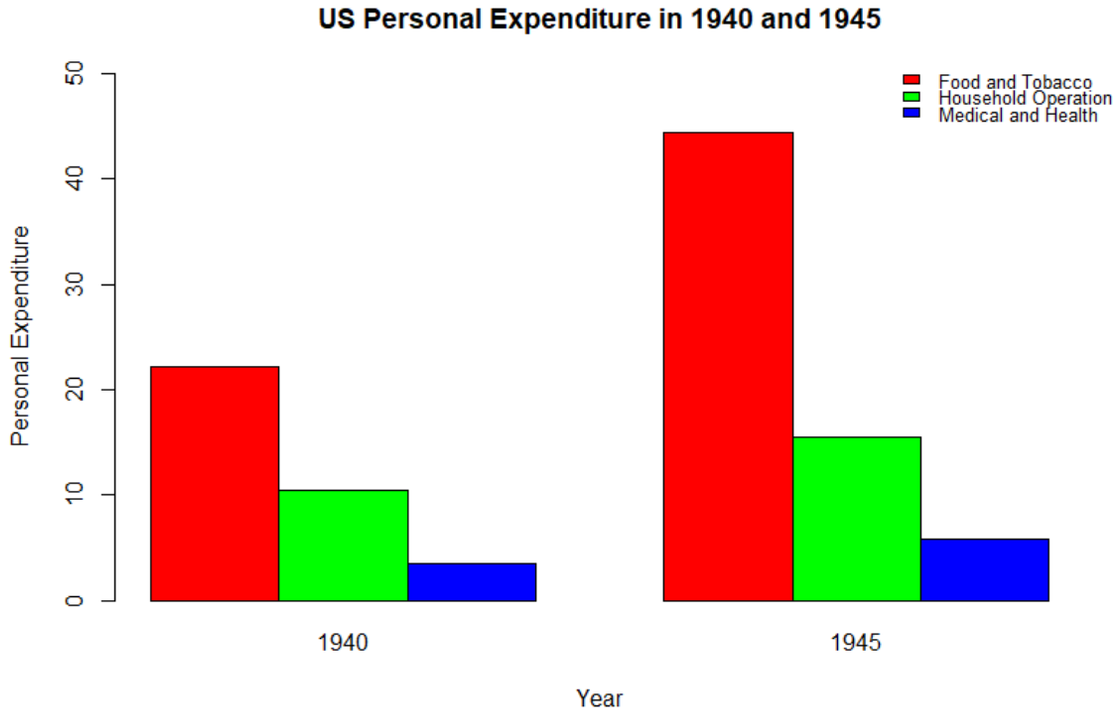
STAT2005 Programming Languages for Statistics
Exercise for Chapter 3

1. Consider the built-in data frame `USPersonalExpenditure`.

(a) Rename `USPersonalExpenditure` into a short form `USPE`. Plot a vertical bar chart to show the US personal expenditure in *Food and Tobacco*, *Household Operation* and *Medical and Health* in 1940. Set the relative font size of the axis names (bar labels) to 0.8, and the range of y-axis to (0,25). You are not required to add title and x-, y-axis labels.

(b) Using the data in `USPE`, plot the bar chart as follows.



Note that the legend should not overlap with the bars. Search for the arguments in `args.legend` to adjust the font size and positon of the legend to avoid the overlap if necessary. Set the bar colors to `rainbow(3)`.

2 The file `ex3_q2.csv` stores the data of a gaming competition in 2023. The competition consists of two rounds. The data in each row represent a team. The variables are

`R1`: Scores of the team in Round 1;

`R2`: Scores of the team in Round 2.

(a) Write R codes to read `ex3_q2.csv` as a `data.frame` object named `data`. Plot two boxplots for each column in `data`.

(b) NRG is a gaming team participating in this competition. They scored 36 in Round 1, and 41 in Round 2. Plot NRG's scores in each round on the two plots as points respectively.

Referring to the plots only, which round did NRG perform better?

(c) The total score of the n^{th} team is given by

$$total_n = \frac{R_{1,n}+R_{2,n}}{2},\ 1 \le n \le 20$$

Plot a normal Q-Q plot for $total$ and add a red reference line to check whether $total$ fits in normal distribution.

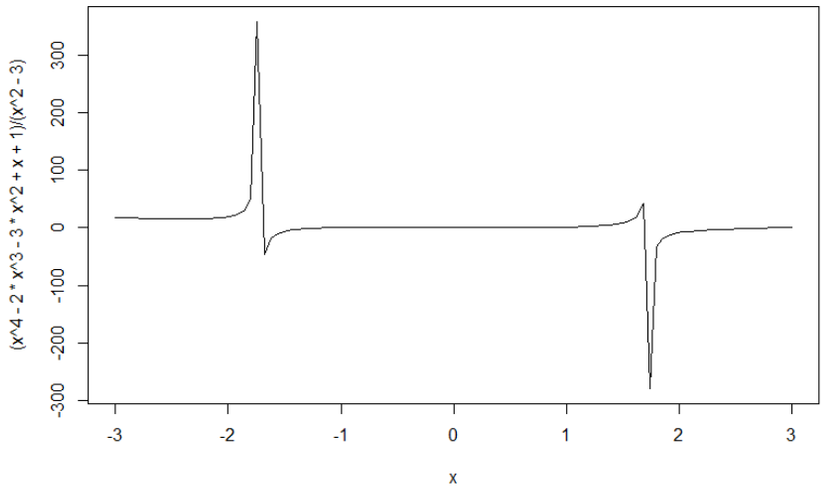
3. (a) Risky tried to write R codes to plot the following curve on $[-3,3]$.

$$y = \frac{x^4 - 2x^3 - 3x^2 + x + 1}{x^2 - 3}$$

This curve has two asymptotes $x = \pm\sqrt{3}$. By executing

```
curve((x^4-2*x^3-3*x^2+x+1)/(x^2-3),-3,3)
```

in R, Risky got the following graph:



However, the curve around $x = \pm\sqrt{3}$ is incorrectly displayed. Please help Risky to plot the correct curve. Set the limit on x-axis to $(-3,3)$ and y-axis to $(-250,250)$ respectively. (Hints: use `seq()`, `plot()`, and `lines()` to plot three parts of the curve one by one.)

(b) Plot two red asymptotes $x = \pm\sqrt{3}$ to the curve.

(c) Using `segments()` function, draw a triangle with vertex $(-1,100)$, $(-1,200)$ and $(0,200)$.

(d) Draw a blue-filled triangle with vertex $(0,100)$, $(1,100)$ and $(1,200)$.

1.

(a)

USPE ← USPersonalExpenditure

barplot(USPE$1940[1:3], cex=0.8, ylim=c(0,25))

cex.names=0.8

USPE[1:3,] ↑ should work, but in fact it is not data from

(b)

barplot(USPE[1:2,1:3], legend.txt = names(USPE$1940)[1:3], args.legend = list(bty="n"), col = rainbow(3))

beside=T

```
# Q1b
b<-barplot(
  USPE[1:3,1:2],
  col=rainbow(3),
  ylim=c(0,50),
  beside=T,
  legend=T,
  args.legend=list(x="topright",bty="n",inset=c(-0.08, -0.02),cex=0.8),
  # x="topright" and inset adjusts the position
  # cex adjusts the font size
  xlab="Year",
  ylab="Personal Expenditure",
  main="US Personal Expenditure in 1940 and 1945"
)
```

2. (a)

data ← read.table("ex3-q2.csv", header=T, sep=",")

~~par(mfrow=c(1,2))~~

~~boxplot(data$R1)~~

~~boxplot(data$R2)~~

boxplot(data)

(b)

for bxt plot, x is 1, 2, ...

points(x=1, 36, pch=19)

points(x=2, 41, pch=19)

```
# Q3
curve((x^4-2*x^3-3*x^2+x+1)/(x^2-3),-3,3)

# Q3a
x1 <- seq(-3,-sqrt(3)-0.001,by=0.001)
x2 <- seq(-sqrt(3)+0.001,sqrt(3)-0.001,by=0.001)
x3 <- seq(sqrt(3)+0.001,3,by=0.001)
y1 <- (x1^4-2*x1^3-3*x1^2+x1+1)/(x1^2-3)
y2 <- (x2^4-2*x2^3-3*x2^2+x2+1)/(x2^2-3)
y3 <- (x3^4-2*x3^3-3*x3^2+x3+1)/(x3^2-3)
plot(x1,y1,type="l",xlim=c(-3,3),ylim=c(-250,250))
lines(x2,y2)
lines(x3,y3)

# Q3b
abline(v=sqrt(3),lty=2,col="red")
abline(v=-sqrt(3),lty=2,col="red")

# Q3c
segments(-1,200,0,200)
segments(-1,100,0,200)
segments(-1,100,-1,200)

# Q3d
polygon(c(0,1,1),c(100,100,200),col="blue")
```

(c)

~~total ← (data$R1 + data$R2)/2~~

~~qqnorm(total)~~

~~qqline(total, col="red")~~

3. (a)

fx ← function(x) {

(x^4 - 2*x^3 - 3*x^2 + x + 1) / (x^2 - 3)

}

plot (-3, -√3):

xs ← seq(-3, -sqrt(3), by = 0.001)

ys ← fx(xs)

plot(xs, ys, type="l", xlim = c(-3,3), ylim = c(-250, 250))

(b)

abline(v = -sqrt(3), col="red", lty=dashed)

abline(v = sqrt(3), col="red", lty=dashed)

(c).

segments(-1, 100, -1, 200, 0, 200)

(d)

polygon(c(0,1,1), c(100,100,200), col="blue")