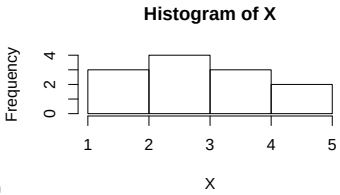
Statistics for the Behavioral Sciences (Chapter 2) Examples

October 29, 2017

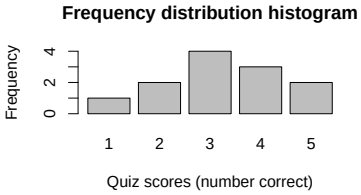
Figure 2.2, page 43

Create vectors with variable names provided by the text.

```
X <- c(5, 5, 4, 4, 4, 3, 3, 3, 3, 2, 2, 1)
hist(X)
```




```
counts <- table(X)
graph <- barplot(counts, main = "Frequency distribution histogram",
    xlab = "Quiz scores (number correct)", ylab = "Frequency")
```



For a bar plot Need to use bar plot to get copy from book since hist() function is not broken down by each value. Found this idea at `https://www.statmethods.net/graphs/bar.html` table() function creates contingency table which shows the distribution of variables.

Figure 2.3, page 43

Create vectors with variable names provided by the text. Need them to be character strings since we do not have the original data.

```
X <- c("30-31", "30-31", "32-33", "32-33", "32-33",
    "32-33", "34-35", "34-35", "34-35", "36-37",
    "36-37", "38-39", "38-39", "38-39", "38-39",
    "38-39", "38-39", "40-41", "40-41", "40-41",
    "42-43", "42-43", "44-45")

# For a bar plot
counts <- table(X)
graph <- barplot(counts, main = "Frequency distribution histogram",
    xlab = "Children's heights (in inches)", ylab = "Frequency")
```

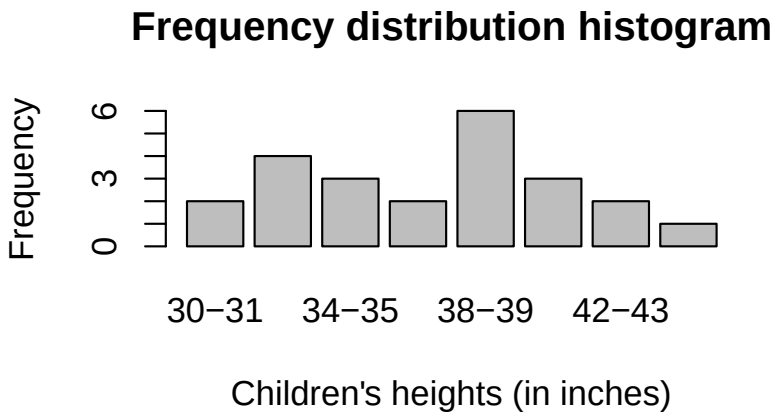
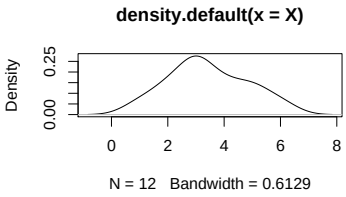


Figure 2.5, page 44

Create vectors with variable names provided by the text and create a density plot.




```
X <- c(6, 5, 5, 4, 4, 3, 3, 3, 3, 2, 2, 1)
plot(density(X))
```

```
density(X)
```

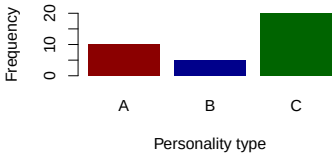
```
##
## Call:
##  density.default(x = X)
##
## Data: X (12 obs.);   Bandwidth 'bw' = 0.6129
##
##        x                y
##  Min.   :-0.8387   Min.   :0.0006123
##  1st Qu.: 1.3306   1st Qu.:0.0271967
##  Median : 3.5000   Median :0.1103603
##  Mean   : 3.5000   Mean   :0.1151039
##  3rd Qu.: 5.6694   3rd Qu.:0.1826300
##  Max.   : 7.8387   Max.   :0.2751871
```

Figure 2.7, page 45

Create vector with variable names provided by the text and shows an example of using *rep* function to replicate elements.

```
X <- c(rep("A", 10), rep("B", 5), rep("C", 20))
```

```
counts <- table(X)
graph <- barplot(counts, xlab = "Personality type",
    ylab = "Frequency", col = c("dark red", "dark blue",
        "dark green"), border = NA)
```



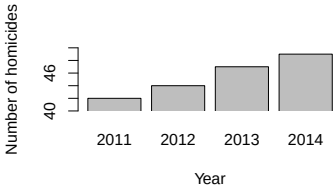
Box 2.1, page 47

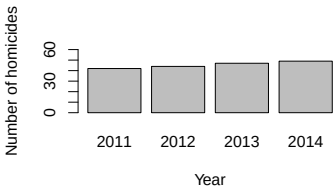
Create *data.frame* with variable names provided by the text.

```
df <- data.frame(Year = c(2011, 2012, 2013, 2014),
    Number_of_Homocides = c(42, 44, 47, 49))
```

Create the two graphs from the text to show how changing the axis alone can change how data is interpreted.

```
barplot(df$Number_of_Homocides, df$Year, xlab = "Year",
    ylab = "Number of homicides", names.arg = df$Year,
    ylim = c(40, 50), xpd = FALSE)
barplot(df$Number_of_Homocides, df$Year, xlab = "Year",
    ylab = "Number of homicides", names.arg = df$Year,
    ylim = c(0, 60))
```





Example 2.8, page 54

Create data.frame with variable names provided by the text.

```
df <- data.frame(X = c(10, 9, 8, 7, 6, 5), f = c(2,
    8, 4, 6, 4, 1), cf = c(25, 23, 15, 11, 5,
    1), cp = c(1, 0.92, 0.6, 0.44, 0.2, 0.04))
knitr::kable(df)
```

X	f	cf	cp
10	2	25	1.00
9	8	23	0.92
8	4	15	0.60
7	6	11	0.44
6	4	5	0.20
5	1	1	0.04

Create data.frame showing range on both scales.

```
df_range <- data.frame(Scores = c(7.5, 7, 6.5),
    Percentages = c(0.44, NA, 0.2), row.names = c("Top",
```

```
        "Intermediate value", "Bottom"))
knitr::kable(df_range)
```

	Scores	Percentages
Top	7.5	0.44
Intermediate value	7.0	NA
Bottom	6.5	0.20

Take difference of the Top Percentage and the Bottom Percentage refering to the data.frame.

```
difference <- df_range["Top", ]$Percentages -
    df_range["Bottom", ]$Percentages
difference
```

```
## [1] 0.24
```

We want take half the difference between the top and bottom and either add or subtract from bottom or top. The first answer taks Top Percentage minus half the difference. The second answer taks *Bottom Percentage* plus half the *difference*.

```
df_range["Top", ]$Percentages - 1/2 * difference
```

```
## [1] 0.32
```

```
df_range["Bottom", ]$Percentages + 1/2 * difference
```

```
## [1] 0.32
```

Example 2.9, page 54

Create data.frame with variable names provided by the text.

```
df <- data.frame(X = c(24, 19, 14, 9, 4), f = c(2,
    3, 3, 10, 2), cf = c(20, 18, 15, 12, 2), cp = c(1,
    0.9, 0.75, 0.6, 0.1))
knitr::kable(df)
```

X	f	cf	cp
24	2	20	1.00
19	3	18	0.90
14	3	15	0.75
9	10	12	0.60
4	2	2	0.10

Create data.frame showing range on both scales.

```
df_range <- data.frame(Scores = c(9.5, NA, 4.5),
    Percentages = c(0.6, 0.4, 0.1), row.names = c("Top",
        "Intermediate value", "Bottom"))
knitr::kable(df_range)
```

	Scores	Percentages
Top	9.5	0.6
Intermediate value	NA	0.4
Bottom	4.5	0.1

Take difference of the Top Percentage and the Bottom Percentage refering to the data.frame.

```
difference <- df_range["Top", ]$Percentages -
    df_range["Bottom", ]$Percentages
difference
```

```
## [1] 0.5
```

We want take half the difference between the top and bottom and either add or subtract from bottom or top. The first answer taks Top Percentage minus half the difference. The second answer taks *Bottom Percentage* plus half the *difference*.

```
df_range["Top", ]$Percentages - 1/2 * difference
```

```
## [1] 0.35
```

```
df_range["Bottom", ]$Percentages + 1/2 * difference
```

```
## [1] 0.35
```

Table 2.3, page 57

Create vector with variable names provided by the text and create a stem and leaf plot. We use scale of 2 becuase scale of 1 will only show 4 stems, but we want a stem for each tenth digit.

```
X <- c(82, 62, 71, 76, 85, 32, 56, 74, 82, 93,
    68, 52, 42, 57, 73, 81, 63, 78, 33, 97, 46,
    59, 74, 76)
stem(X, scale = 2)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   3 | 23
##   4 | 26
##   5 | 2679
##   6 | 238
##   7 | 1344668
##   8 | 1225
##   9 | 37
```

Demonstration 2.1, page 60

Create vector with variable names provided by the text and cut the data up into the given intervals and create a frequency table. Then we make a cumulative sum to get cumulative frequencies and proportions/percentages. Finally combine all into a data frame.

```
X <- c(14, 8, 27, 16, 10, 22, 9, 13, 16, 12, 10,
    9, 15, 17, 6, 14, 11, 18, 14, 11)

f <- table(cut(X, breaks = c(5, 9, 14, 19, 24,
    29)))

cf <- cumsum(f)

cp <- cumsum(prop.table(f))

answer <- data.frame(f = f, cf = cf, cp = cp)

# Remove f.Var1 column since they are
# accounted for in the row names.
answer <- answer[, -1]

# Order columns to match the data presented in
# the text.
answer <- answer[order(-cf), ]

knitr::kable(answer)
```

	f.Freq	cf	cp
(24,29]	1	20	1.00
(19,24]	1	19	0.95
(14,19]	5	18	0.90

	f.Freq	cf	cp
(9,14]	9	13	0.65
(5,9]	4	4	0.20