Lab 14 R Script

Alexander Hernandez

10/18/2022

```
library(UsingR)
```

```
## Loading required package: MASS
## Loading required package: HistData
## Loading required package: Hmisc
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, units
##
## Attaching package: 'UsingR'
## The following object is masked from 'package:survival':
##
##     cancer
```

1) Flue Season in Nevada

```
# Ho: the data fits an equal distribution
# Ha: the data does not fit an equal distribution
flu = c(62,84,17,16,21)
chisq.test(flu)
```

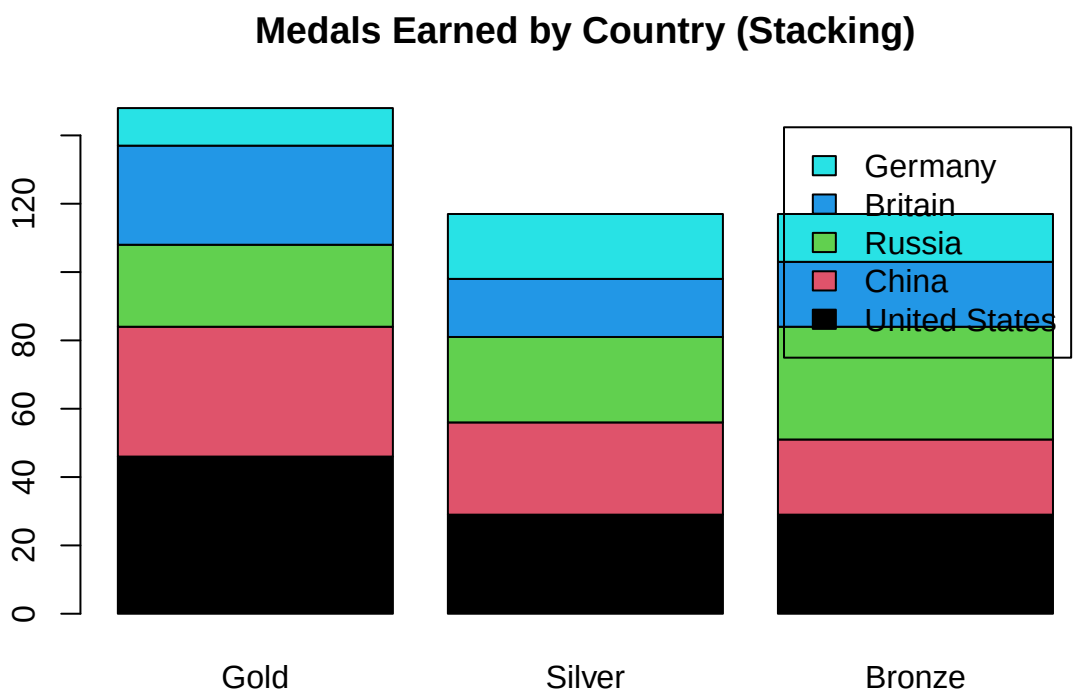
```
##
##  Chi-squared test for given probabilities
##
## data:  flu
## X-squared = 97.15, df = 4, p-value < 2.2e-16
```

```
# With a p-value of 2.2e-16,
# there is enough evidence to reject the null hypothesis
# and claim that the flu season data does not follow an equal distribtuion.
```

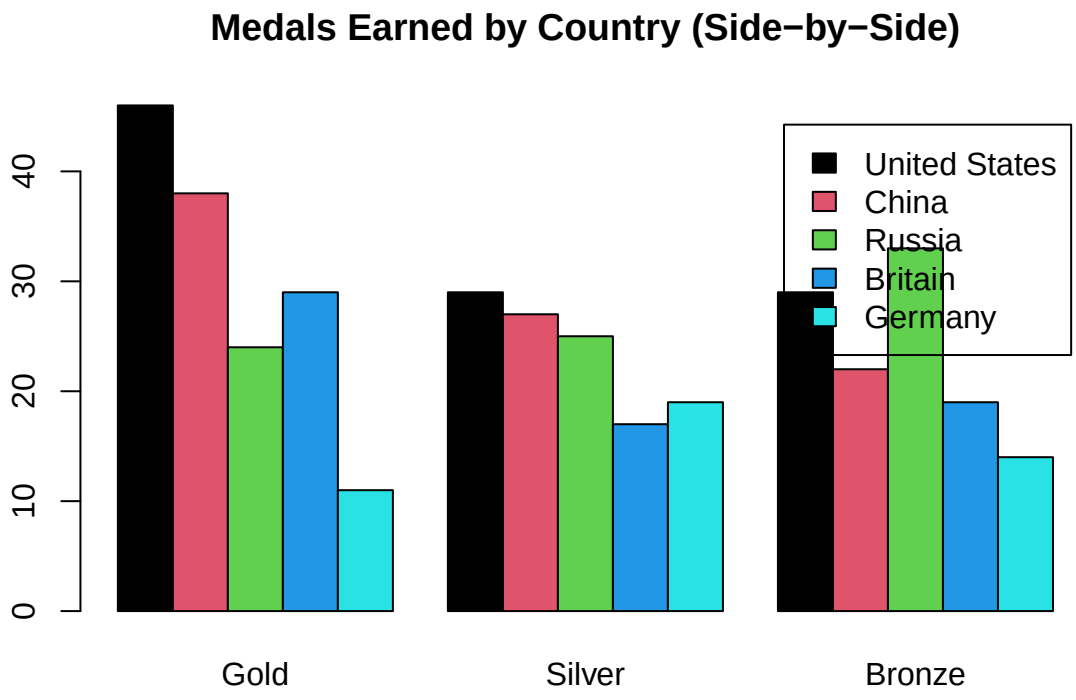
2) Olympic Medal Winners of 2016

```
winners = c(46, 29, 29, 38, 27, 22,
            24, 25, 33, 29, 17, 19,
            11, 19, 14)
winners_t = matrix(data = winners, nrow=5, byrow=TRUE)
colnames(winners_t) = c("Gold", "Silver", "Bronze")
rownames(winners_t) = c("United States", "China", "Russia",
                        "Britain", "Germany")

barplot(winners_t, col=c(1,2,3,4,5),
        legend=rownames(winners_t),
        main = "Medals Earned by Country (Stacking)")
```




```
barplot(winners_t, col=c(1,2,3,4,5), beside = T,
        legend=rownames(winners_t),
        main = "Medals Earned by Country (Side-by-Side)")
```



3) Health and Happiness

```
# Ho: the data follows an equal distribution
# Ha: the data does not follow an equal distribution
happiness = c(271,  261,    82, 20, 247,    567,    231,
              53,  33,  103,    92, 36)
happiness_t = matrix(happiness, nrow=3, byrow=TRUE)
colnames(happiness_t) = c("Excellent", "Good", "Fair", "Poor")
rownames(happiness_t) = c("Very Happy" ,"Pretty Happy", "Not Too Happy")

chisq.test(happiness_t, )
```

```
##
##  Pearson's Chi-squared test
##
## data:  happiness_t
## X-squared = 182.17, df = 6, p-value < 2.2e-16
```

```
# With a p-value of 2.2e-16,
# we have enough evidence to reject the null hypothesis and
# claim that there is a relation between health and happiness
# as the data is not equally distributed
```

4) Seat-Bels in California

```
# Ho: the data follows an equal distribution
# Ha: the data does not follow an equal distribution
seatbelts = c(56, 8, 2, 16)
seatbelts_t = matrix(seatbelts, nrow=2, byrow=TRUE)
colnames(seatbelts_t) = c("Buckled", "Unbuckled")
rownames(seatbelts_t) = c("Buckled", "Unbuckled")
chisq.test(seatbelts_t)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  seatbelts_t
## X-squared = 35.995, df = 1, p-value = 1.978e-09
```

```
# With a p-value of 1.978e-09,
# we have enough evidence to reject the null hypothesis and
# claim that there is a relation between parent and child seatbelt usage
# as the data is not equally distributed.
```

5) M&M Package Colors

```
# Ho: the data(candies) are from the milkchocolate group
# Ha: the data is not from the milkchocolate group
data(mandms)
mandms
```

```
##                   blue   brown   green  orange     red  yellow
## milk chocolate 10.0000 30.0000 10.0000 10.0000 20.0000 20.0000
## Peanut         20.0000 20.0000 10.0000 10.0000 20.0000 20.0000
## Peanut Butter  20.0000 20.0000 20.0000  0.0000 20.0000 20.0000
## Almond         16.6667 16.6667 16.6667 16.6667 16.6667 16.6667
## kid minis      16.6667 16.6667 16.6667 16.6667 16.6667 16.6667
```

```
mms = c(15, 34, 7, 19, 29, 24)
p = c(0.1, 0.3, 0.1, 0.1, 0.2, 0.2)
chisq.test(mms, p=p)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  mms
## X-squared = 7.0651, df = 5, p-value = 0.2158
```

```
# With a p-value of 0.2158,
# we do not have enough evidence to reject the null hypothesis and
# cannot claim that the candies are from a group other than milkchocolate.
```

6) Find the true source of candies

```
pe = c(0.2, 0.2, 0.1, 0.1, 0.2, 0.2)
chisq.test(mms, p=pe)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  mms
## X-squared = 13.328, df = 5, p-value = 0.02049
```

```
pb = c(0.2, 0.2, 0.2, 0, 0.2, 0.2)
chisq.test(mms, p=pb)
```

```
## Warning in chisq.test(mms, p = pb): Chi-squared approximation may be incorrect
```

```
##
##  Chi-squared test for given probabilities
##
## data:  mms
## X-squared = Inf, df = 5, p-value < 2.2e-16
```

```
al = c(0.167, 0.167, 0.167, 0.167, 0.167, 0.165)
chisq.test(mms, p=al)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  mms
## X-squared = 22.401, df = 5, p-value = 0.0004391
```

```
km = c(0.167, 0.167, 0.167, 0.167, 0.167, 0.165)
chisq.test(mms, p=km)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  mms
## X-squared = 22.401, df = 5, p-value = 0.0004391
```

```
# Based upon the p-values from the previous chi-squared tests,
# the null hypothesis can be rejected with evidence obtained from
# the almond and kid minis tests. A p-value of 0.0004391 indicates that
# the mnm data could originate from either almond or kid minis.
```