

List 1

Init and Dataset import

```
library(tidyverse)
library(jmv)
hand1994_babies <- read_csv("data/hand1994_babies.csv",
  col_types = cols(Survival = col_factor(levels = c("no", "yes"))))

hand1994_elderheight <- read_lines("data/hand1994_elderheight.csv") %>%
  str_split_fixed(" ", n = Inf) %>%
  as.numeric %>%
  na.omit()

hand1994_maternaldiet <- read_csv("data/hand1994_maternaldiet.csv")
```

Question 1 - Slide 15

Make a Two-way table of the hand1994_babies dataset.

```
hand1994_babies # Dataset used
```

```
## # A tibble: 50 x 2
##   Weight Survival
##   <dbl> <fct>
## 1  1050 no
## 2  2500 no
## 3  1890 no
## 4  1760 yes
## 5  2830 yes
## 6  1175 no
## 7  1030 no
## 8  1940 no
## 9  1930 yes
## 10 1410 yes
## # ... with 40 more rows
```

```
summary(hand1994_babies) # Dataset summary
```

```
##      Weight      Survival
## Min.   :1030    no :27
## 1st Qu.:1432    yes:23
## Median :1855
## Mean   :1975
## 3rd Qu.:2485
## Max.   :3640
```

```
options(digits = 4)
result <- hand1994_babies %>%
  mutate(Weight = cut_interval(Weight, 5)) %>% # Bin Weight in 5 intervals
  count(Survival, Weight) %>%
  contTables(formula = n ~ Weight:Survival, pcRow = T) # Set Weight as exp. variable

result$freqs
```

```
##
## Contingency Tables
## -----
##      Weight                no      yes      Total
## -----
##      [1.03e+03,1.55e+03]  Observed      13      2      15
##                          % within row  86.67    13.33  100.00
##
##      (1.55e+03,2.07e+03]  Observed      7      8      15
##                          % within row  46.67    53.33  100.00
##
##      (2.07e+03,2.6e+03]   Observed      6      5      11
##                          % within row  54.55    45.45  100.00
##
##      (2.6e+03,3.12e+03]   Observed      1      5      6
##                          % within row  16.67    83.33  100.00
##
##      (3.12e+03,3.64e+03]  Observed      0      3      3
##                          % within row   0.00   100.00  100.00
##
##      Total                Observed     27     23     50
##                          % within row  54.00    46.00  100.00
## -----
```

Question 4 - Slide 16

Analyze hand1994_elderheight dataset.

```
hand1994_elderheight # Dataset used
```

```
##      [1] 142 145 145 145 146 147 147 147 147 148 148 149 150 150 150 150 150 150
##      [19] 151 151 151 151 151 151 152 152 152 152 152 152 152 152 152 152 152 152
##      [37] 153 153 153 153 153 153 153 153 153 153 153 153 153 153 153 153 153 154
##      [55] 154 154 154 154 154 154 154 154 154 154 155 155 155 155 155 155 155 155
##      [73] 155 155 155 155 155 155 155 155 155 155 155 155 155 156 156 156 156 156
##      [91] 156 156 156 156 156 156 156 156 156 156 156 156 156 156 156 157 157 157
##     [109] 157 157 157 157 157 157 157 157 157 157 157 157 157 157 157 157 157 158
##     [127] 158 158 158 158 158 158 158 158 158 158 158 158 158 158 158 158 158 158
##     [145] 158 158 158 158 158 158 158 158 158 158 158 159 159 159 159 159 159 159
##     [163] 159 159 159 159 159 159 159 159 159 159 159 160 160 160 160 160 160 160
##     [181] 160 160 160 160 160 160 160 160 160 160 160 160 160 160 160 161 161 161
##     [199] 161 161 161 161 161 161 161 161 161 161 161 161 161 161 161 161 162 162
##     [217] 162 162 162 162 162 162 162 162 162 162 162 162 162 162 162 162 163 163
##     [235] 163 163 163 163 163 163 163 163 163 163 163 163 163 163 163 163 163 163
```

```
## [253] 163 163 163 163 163 163 163 163 163 163 164 164 164 164 164 164 164
## [271] 164 164 164 164 164 164 164 164 164 164 165 165 165 165 165 165 165
## [289] 165 165 165 165 165 165 165 165 165 166 166 166 166 166 166 166 166
## [307] 166 166 167 167 167 167 167 167 167 168 168 168 168 168 168 169 169
## [325] 169 169 169 169 169 170 170 170 170 170 170 170 170 170 170 171 171
## [343] 171 173 173 173 174 176 177 178 178
## attr("na.action")
## [1] 120 150 180 210 240 270 300 330 360
## attr("class")
## [1] "omit"
```

```
summary(hand1994_elderheight) # Dataset summary
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      142    156     160     160    164     178
```

```
length(hand1994_elderheight) # Total incidence
```

```
## [1] 351
```

```
hand1994_elderheight %>% # Binned incidence
  data.frame() %>%
  transmute(Height = cut_width(.,5)) %>% # Bin Height in 5 cm intervals
  count(Height)
```

```
##      Height    n
## 1 [138,142]    1
## 2 (142,148]    8
## 3 (148,152]   27
## 4 (152,158]   89
## 5 (158,162]  107
## 6 (162,168]   83
## 7 (168,172]   28
## 8 (172,178]    6
## 9 (178,182]    2
```

The Height distribution of the 351 subjects appears to have a normal-like distribution, with mean and median in 160 cm, which is inside the 158 – 162 (here, represented as (158, 162]) bin. The 5 cm interval bin appears to be sufficient to show the distribution.

Question 6 - Slide 18

Complete the prevalence two-way table of the hand1994_maternaldiet dataset, with control and cases as exp. var. Analyze and discuss.

```
hand1994_maternaldiet # Dataset used
```

```
## # A tibble: 3 x 3
##   'Dieta materna' Casos Controles
##   <chr>          <dbl>    <dbl>
## 1 Boa           34      43
## 2 Razoavel     110     48
## 3 Pobre        100     32
```

```

results <- hand1994_maternaldiet %>%
  pivot_longer(-`Dieta materna`, names_to = "Exp_pop") %>%
  contTables(formula = value ~ Exp_pop: `Dieta materna`, pcRow = T)

results$freqs

```

```

##
## Contingency Tables
## -----
##   Exp_pop          Boa      Pobre      Razoavel      Total
## -----
##   Casos      Observed      34      100      110      244
##              % within row  13.93  40.98  45.08  100.00
##
##   Controles  Observed      43      32      48      123
##              % within row  34.96  26.02  39.02  100.00
##
##   Total      Observed      77      132      158      367
##              % within row  20.98  35.97  43.05  100.00
## -----

```

The group of mothers of babies with spina bifida tend to have a poorer diet in comparison with the control group. However, the control group itself do not have any apparent tendency in diet quality.