

# Homework 4

Tauqeer Kasam Rumaney

November 13, 2022

Download this R Markdown file, save it on your computer, and perform all the below tasks by inserting your answer in text or by inserting R chunks below. After you are done, upload this file with your solutions on Moodle.

## Exercise 1: Descriptive statistics - example 1

Create a summary table with descriptive statistics of the esoph dataframe in the datasets package. For example by using the summary\_table function in the qwraps2 package.

Alternatively, only compute the descriptive statistics of interest, and put them into a table in Word (or Excel, or another program).

```
# Load the data
library(datasets)

dat1 <- esoph

library(qwraps2)
options(qwraps2_markup = "markdown")

library(table1)
```

```
##
## Attaching package: 'table1'
```

```
## The following objects are masked from 'package:base':
##
```

```
##      units, units<-
```

```
table1::table1(~ alcgp + tobgp + ncases + ncontrols | agegp, data = dat1)
```

```
## Get nicer `table1` LaTeX output by simply installing the `kableExtra` package
```

	25-34	35-44	45-54	55-64	65-74	75+	Overall
	(N=15)	(N=15)	(N=16)	(N=16)	(N=15)	(N=11)	(N=88)
alcgp							
0-39g/day	4 (26.7%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	3 (27.3%)	23 (26.1%)
40-79	4 (26.7%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	3 (20.0%)	4 (36.4%)	23 (26.1%)
80-119	3 (20.0%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	2 (18.2%)	21 (23.9%)
120+	4 (26.7%)	3 (20.0%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	2 (18.2%)	21 (23.9%)

	25-34	35-44	45-54	55-64	65-74	75+	Overall
tobgp							
0-9g/day	4 (26.7%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	4 (36.4%)	24 (27.3%)
10-19	4 (26.7%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	4 (36.4%)	24 (27.3%)
20-29	3 (20.0%)	4 (26.7%)	4 (25.0%)	4 (25.0%)	4 (26.7%)	1 (9.1%)	20 (22.7%)
30+	4 (26.7%)	3 (20.0%)	4 (25.0%)	4 (25.0%)	3 (20.0%)	2 (18.2%)	20 (22.7%)
ncases							
Mean (SD)	0.0667 (0.258)	0.600 (0.986)	2.88 (2.09)	4.75 (2.29)	3.67 (4.10)	1.18 (0.603)	2.27 (2.75)
Median [Min, Max]	0 [0, 1.00]	0 [0, 3.00]	3.00 [0, 6.00]	4.00 [2.00, 9.00]	3.00 [0, 17.0]	1.00 [0, 2.00]	1.00 [0, 17.0]
ncontrols							
Mean (SD)	7.67 (11.1)	12.7 (15.8)	10.4 (12.6)	10.4 (12.8)	7.07 (11.0)	2.82 (4.94)	8.81 (12.1)
Median [Min, Max]	4.00 [0, 40.0]	8.00 [1.00, 60.0]	6.00 [0, 45.0]	6.00 [0, 47.0]	4.00 [0, 43.0]	2.00 [0, 17.0]	4.00 [0, 60.0]

```

dat1$alch <- factor(dat1$alcgp, labels = c("low","medium","high", "very high"))
dat1$tob <- factor(dat1$tobgp, labels = c("low","medium","high", "very high"))

summary1<-
  list(
    "Alcohol Consumption"=
      list(" (0-39) " = ~ qwraps2::n_perc0(alch == "low" , show_symbol = TRUE, na_rm = TRUE),
            " (40-79) " = ~ qwraps2::n_perc0(alch == "medium", show_symbol = TRUE, na_rm = TRUE),
            " (80-119) " = ~ qwraps2::n_perc0(alch == "high", show_symbol = TRUE, na_rm = TRUE),
            " (120+) " = ~ qwraps2::n_perc0(alch == "very high", show_symbol = TRUE, na_rm = TRUE)),
    "Tobacco Consumption"=
      list(" (0-9) " = ~ qwraps2::n_perc0(tob == "low" , show_symbol = TRUE, na_rm = TRUE),
            " (10-19) " = ~ qwraps2::n_perc0(tob == "medium", show_symbol = TRUE, na_rm = TRUE),
            " (20-29) " = ~ qwraps2::n_perc0(tob == "high", show_symbol = TRUE, na_rm = TRUE),
            " (30+) " = ~ qwraps2::n_perc0(tob == "very high", show_symbol = TRUE, na_rm = TRUE)),
    "Ncases" =
      list("Mean (SD)" = ~ qwraps2::mean_sd(ncases, denote_sd = "paren", na_rm = TRUE, show_n = "never"),
            "Median (IQR)" = ~ qwraps2::median_iqr(ncases, na_rm = TRUE, show_n = "never")),
    "Ncontrols" =
      list("Mean (SD)" = ~ qwraps2::mean_sd(ncontrols, denote_sd = "paren", na_rm = TRUE, show_n = "never"),
            "Median (IQR)" = ~ qwraps2::median_iqr(ncontrols, na_rm = TRUE, show_n = "never"))
  )

table2<- summary_table(dplyr::group_by(dat1, agegp), summary1)
table2_overall <- summary_table(dat1, summary1)
table2_both <- cbind(table2_overall, table2)

print(table2_both,

```

```
rtitle = "Summary Statistics",
cnames = c("Overall \\ (N = 88)", "Age grp (25-34) \\ (N = 15)", "Age grp (35-44) \\ (N = 15)"
```

Summary Statistics	Overall (N = 88)	Age grp (25-34) (N = 15)	Age grp (35-44) (N = 15)	Age grp (45-54) (N = 16)	Age grp (55-64) (N = 16)	Age grp (65-74) (N = 15)	Age grp (75+) (N = 11)
<b>Alcohol Consumption</b>							
(0-39)	23 (26%)	4 (27%)	4 (27%)	4 (25%)	4 (25%)	4 (27%)	3 (27%)
(40-79)	23 (26%)	4 (27%)	4 (27%)	4 (25%)	4 (25%)	3 (20%)	4 (36%)
(80-119)	21 (24%)	3 (20%)	4 (27%)	4 (25%)	4 (25%)	4 (27%)	2 (18%)
(120+)	21 (24%)	4 (27%)	3 (20%)	4 (25%)	4 (25%)	4 (27%)	2 (18%)
<b>Tobacco Consumption</b>							
(0-9)	24 (27%)	4 (27%)	4 (27%)	4 (25%)	4 (25%)	4 (27%)	4 (36%)
(10-19)	24 (27%)	4 (27%)	4 (27%)	4 (25%)	4 (25%)	4 (27%)	4 (36%)
(20-29)	20 (23%)	3 (20%)	4 (27%)	4 (25%)	4 (25%)	4 (27%)	1 (9%)
(30+)	20 (23%)	4 (27%)	3 (20%)	4 (25%)	4 (25%)	3 (20%)	2 (18%)
<b>Ncases</b>							
Mean (SD)	2.27 (2.75)	0.07 (0.26)	0.60 (0.99)	2.88 (2.09)	4.75 (2.29)	3.67 (4.10)	1.18 (0.60)
Median (IQR)	1.00 (0.00, 4.00)	0.00 (0.00, 0.00)	0.00 (0.00, 1.00)	3.00 (1.00, 4.25)	4.00 (3.00, 6.00)	3.00 (1.00, 4.50)	1.00 (1.00, 1.50)
<b>Ncontrols</b>							
Mean (SD)	8.81 (12.14)	7.67 (11.14)	12.67 (15.83)	10.44 (12.63)	10.38 (12.82)	7.07 (11.02)	2.82 (4.94)
Median (IQR)	4.00 (1.00, 10.00)	4.00 (1.50, 7.00)	8.00 (2.50, 13.00)	6.00 (1.75, 14.00)	6.00 (1.75, 13.50)	4.00 (1.00, 7.50)	2.00 (0.00, 3.00)

```
#Converting to Excel
#install.packages("writexl")
library(writexl)
#write_xlsx(table2_both, "/Users/taugeerrumaney/BioStat/ abc.xlsx") #Error: Argument x must be a data f
write_xlsx(dat1, "/Users/taugeerrumaney/BioStat/ abc.xlsx")
```

## Exercise 2: Descriptive statistics - example 2 (optional, more advanced)

Create a summary table with descriptive statistics of the UKLungDeaths data in the datasets package. For example by using the summary\_table function in the qwraps2 package.

Alternatively, only compute the descriptive statistics of interest, and put them into a table in Word (or Excel, or another program).

```
# Load the data  
library(datasets)  
?UKLungDeaths
```

Hint: the data consists of three datasets.