

STA 3100 Programming with Data: Assignment 020

Berkeley Admissions Data

The dataset `UCBAdmissions` is distributed with base R. For the year 1973, applicants to the graduate programs of the six largest academic departments at the University of California Berkeley are cross-classified by admission status, gender, and the department to which they applied.

```
fable(UCBAdmissions, row.vars = c("Admit", "Gender"), col.vars = "Dept")
```

```
##           Dept  A  B  C  D  E  F
## Admit  Gender
## Admitted Male    512 353 120 138  53  22
##           Female    89  17 202 131  94  24
## Rejected Male    313 207 205 279 138 351
##           Female    19   8 391 244 299 317
```

Note: Some of the functions that I used in the class notes have acquired new names. The old names still work, but `margin.table()` has been renamed `marginSums()` and `prop.table()` has been renamed `proportions()`. You can use either, but I would suggest adopting the new function names.

Exercises

1. (5 pts) Display the marginal table cross-classifying admission by gender (ignoring department).

```
margin.table(UCBAdmissions, c("Admit", "Gender"))
```

```
##           Gender
## Admit      Male Female
## Admitted 1198    557
## Rejected 1493    1278
```

2. (5 pts) Display the marginal table giving the proportions admitted and rejected by gender (ignoring department). Express your results as percentages, rounded to one decimal place. For whom is the admission rate higher overall, males or females?

```
round(100 * prop.table(margin.table(UCBAdmissions, c("Admit", "Gender"))), 1)
```

```
##           Gender
## Admit      Male Female
## Admitted   68    32
## Rejected   54    46
```

3. (10 pts) Now display a table giving the proportion admitted by gender within each department. Express your results as percentages, rounded to one decimal place. How do these results compare to what you found in exercise 2?