

Solutions: Practical 1

R Refresher

Querying databases; plots and summaries

- (a) 24
- (b) 4847
- (c) 14651; 8517; 18277
- (d) 4013.53; 10620.21
- (e) 927; 739; 16
- (f)
 - (i) Removing all the missing budgets
 - (ii) Partition the plotting space into a 2 by 2 space (giving four plots)
 - (iii) Re-do plots...
 - (iv) Use `plot(Rating, Votes)`
 - (v) Working out the mean for columns 3, 4, 5 and 6
 - (vi) Median = 97.0; IQR = 17.0
 - (vii) `movies2 = movies[,1:2]`

Functions, for loops and if statements

- (a) Answer should be around 0.819
- (b) Need to have `MCintegrate = function(N, l, u)`, and then replace `x = runif(1, -2, 1)` with `x = runif(1, l, u)`.

Simulation

1.
 - (a) `W = rpois(1000, 2.5); barplot(table(W))`
 - (b)
 - (i) `dnorm(3, 12, 0.01)`
 - (ii) `pnorm(2, 12, 0.01)`
 - (iii) `X = rbinon(1000, 12, 0.01); mean(x)` (should be close to $n \times p$); `var(x)` (should be close to $n \times p \times (1 - p)$)
 - (c)
 - (i) No - why have such definite cutoffs, and equal probability in between?
 - (ii) `Y = runif(1000, -15, 45); mean(Y); var(Y)`
 - (iii) `hist(Y, xlab = "Minutes late")`

```
2.
1 wind = c(26,10,10,3,5,33,13,23,5,8,14,24,3,2,7,10,46,15,2,35)
2 hist(wind, freq = FALSE, xlab="Wind speed (mph)", main="Histogram of wind
  speeds", xlim=c(0,80))
3 lambda = 1/mean(wind)
4 x = seq(0, 80, 0.001)
5 d = lambda*exp(-lambda*x)
6 lines(x, d, type="l", col="red")
7 length(wind[wind>70])/length(wind)
```

3. Use `pnorm(1)-pnorm(-2)`.

4. See page 6 of the lecture notes - amend code!