
Question1 (5 points)

Use `:`, `seq()`, `rep()` and possibly other operators/functions, but definitely not `c()`, to create the following sequences.

(a) (1 point) `2.0 2.3 2.6 2.9 3.2 3.5 3.8 4.1 4.4`

(b) (1 point) `"ax" "ay" "by" "bz" "az" "az"`

(c) (1 point) `TRUE TRUE FALSE FALSE FALSE FALSE`

(d) (1 point) `1 22 333 4444 55555 666666`

(e) (1 point) `0 1 2 3 0 2 4 6 0 3 6 9 0 4 8 12`

Question2 (6 points)

Consider the following magic trick for kids.

- Ask him to think of a number, n, and carry out the following steps in his head
  1. Compute $n_1 = 3n$ and announce whether n_1 is even or odd.
  2. If n_1 is even, compute the number n_2 that is half of n_1.

     If n_1 is odd, compute the number n_2 that is half of $n_1 + 1$.
  3. Compute $n_3 = 3n_2$.
  4. Divide n_3 by 9 and reveal the quotient k, discard any remainder.
- You compute the original number n as

  $2k$ if n_1 was even.

  $2k + 1$ if n_1 was odd.
- Reveal the original number n. Voila!

(a) (5 points) Write an R function (name it as `hints.func`) that take a single integer n, works through the steps 1-4 above and returns a list containing the value of k and an indication of whether n_1 was even or odd.

(b) (1 point) Write an R function (name it as `guess.func`) that takes the list produced by `hints.func` and returns the reconstructed value of n.

Question3 (7 points)

Given a sequence consisting of 0's and 1's, the chunk of consecutive 0's between a pair of 1's is known as a gap in the sequence. We are interested in finding the frequencies of gap lengths ($= 0, 1, 2, \ldots$). Zeros at the beginning and end of the sequence are ignored, e.g.

`0 1 0 0 0 0 1 1 0 0 1 1 1 0 0`

has five gaps, being of length 4, 0, 2, 0, 0, respectively. The gap lengths can thus be summarised in a one-way frequency table as follows:

```
> gap.freq(x, m = 3)
```

```
0 1 2 3+
3 0 1 1
```

where the gap lengths $\geq$ `m` are treated as one group.

(a) (1 point) Consider the following

```
> set.seed(572); u = runif(1e6)
```

Write a single R statement that converts `u` into a vector that consists only 0's and 1's by treating the values in `u` as 1's if they are strictly between 0 and 0.3, and 0's if otherwise.

(b) (6 points) Write the R function `gap.freq()` that, given `x` and `m`, returns the frequency table of gap lengths in the same format above. You need to demonstrate that your function works properly for the long sequence created in (a). Use `m = 10` in your demonstration.

Question4 (5 points)

The dataset `islands` in R contains the areas in thousands of square miles of the 48 largest landmasses in the world. Use R statements or functions to find

(a) (1 point) the area of the largest landmass.

(b) (1 point) the number of landmasses with areas between 100 and 1000 square miles.

(c) (1 point) the ranking of the area of the North Island of New Zealand (`New Zealand (N)`) in the world.

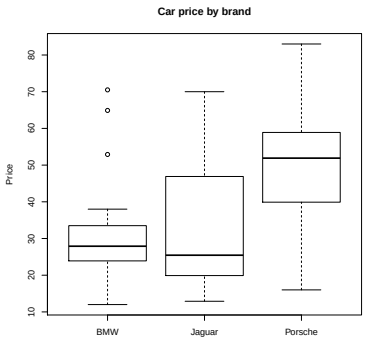
(d) (1 point) the name of the landmass that has the most similar area to New Zealand (North and South Islands).

(e) (1 point) the names of the top 10 largest landmasses.

Question5 (2 points)

(a) (1 point) Read `three_cars.csv` into R as a data frame, and name it `threecars.df`.

(b) (1 point) Produce the following plot in R and output it into a pdf file.



Question6 (3 points)

(a) (1 point) Read `integers_letters.dbf` into R as a data frame, and name it `il.df`.

(b) (1 point) Write R statement/s to find the integer/s between 1 and 26 that is missing.

(c) (1 point) Write R statement/s to find the letter/s that is/are capitalised.

Question7 (2 points)

(a) (1 point) Read whisky price into R. [Hint: You may find `readLines` and `scan` useful.]

(b) (1 point) Use R statement/s or functions to perform a two sample t-test on the dataset.