Week 2 Lab

The purpose of this lab is to introduce the basics of R and the R Markdown files we will be using in Labs across the semester. Use the R Reference as a guide.

Question 1

(a)

Below is an R code chunk, which performs simple arithmetic. Click the Run symbol in the chunk's top right to produce its output.

```
2+3
```

```
## [1] 5
```

(b)

We can assign a value, or the output of some arithmetic, to a variable. In the following code chunk, assign the square of 4 to a variable named `x`.

Then, using this variable, add a second line of code to find the result of $4^2/8$.

```
x <- 4^2
x/8
```

```
## [1] 2
```

Notice how only the second line of code produces an output — the assignment of 4 squared to `x` won't.

(c)

Repeat the operations performed in Section 3.1 of the R Reference.

Question 2

This question will introduce vectors and built-in functions — see Sections 3.2, 3.3, and 3.4 of the R Reference.

(a)

Define `v1` as a vector with 40 values, each spaced equally between 0 and 1.

Hint: See ***`seq(...)`***

```
v1 <- seq(0, 1, length.out = 40)
```

(b)

Using `v1` and the `+` operator, define `v2` as a vector with 40 values, each spaced equally between 1 and 2. Output the value of `v2`.

```
v2 <- v1 + 1
v2
```

```
##  [1] 1.000000 1.025641 1.051282 1.076923 1.102564 1.128205 1.153846 1.179487
##  [9] 1.205128 1.230769 1.256410 1.282051 1.307692 1.333333 1.358974 1.384615
## [17] 1.410256 1.435897 1.461538 1.487179 1.512821 1.538462 1.564103 1.589744
## [25] 1.615385 1.641026 1.666667 1.692308 1.717949 1.743590 1.769231 1.794872
## [33] 1.820513 1.846154 1.871795 1.897436 1.923077 1.948718 1.974359 2.000000
```

(c)

Find both the mean and sum of the values of `v2`.

```
mean(v2)
```

```
## [1] 1.5
```

```
sum(v2)
```

```
## [1] 60
```

(d)

Consider the vector below (you must run the code chunk to define `vals`):

```
vals <- c(1, 1, 5, 3, 6, 1, 2, 1, 5, 2)
```

We can find the number of 1s in the vector using this command:

```
# 'vals == 1' returns a vector of TRUEs and FALSEs at the indexes of vals which are 1.
sum(vals == 1)
```

```
## [1] 4
```

Using a similar technique, find the number of values greater than or equal to 5 in `vals`.

```
sum(vals == 5)
```

```
## [1] 2
```

(e)

Using the appropriate vector, evaluate $\sum_{k=1}^{\infty} \frac{1}{k^2}$.

Hint: In R, we cannot define infinite vectors. Instead we tend to truncate them — here, try approximating this series by defining a vector which contains integers from 1 to 10^6.

```
k_vals <- 1:10^6
sum(1/k_vals^2)
```

```
## [1] 1.644933
```

We can also graph vectors — see Sections 6 & 7 of the R Reference.

Question 3

(a)

Define the function $f(x) = log(x^2 + 3)$ (see Section 12 of the R Reference).

```
f <- function(x) {
  log(x^2+3)
}
```

(b)

Evaluate $\int_1^5 f(x)\,dx$.?

Hint: Enter ***?integrate*** *in the Console for a guide of how to use this function.*

```
integrate(f, 1, 5)
```

```
## 9.747142 with absolute error < 1.3e-13
```

Question 4 [Extension]

The following code chunk defines a bag of coloured balls, each ball represented as a 'character' (see Section 3.3 of the R Reference).

```
bag <- c("blue", "yellow", "indigo", "green", "white")
```

We can simulate grabbing three balls at random from this bag using the `sample(...)` function. Enter `?sample` in the Console for a guide to using this function.

```
sample(bag, 3)
```

```
## [1] "blue"   "white"  "indigo"
```

Now, let us assume that there is an infinite number of balls of each colour in the bag — so we could draw each ball colour multiple times. We add `replace = TRUE` to the function.

```
sample(bag, 3, replace = TRUE)
```

```
## [1] "blue"   "yellow" "white"
```

Finally, suppose that the chance of grabbing a green ball is twice as likely as any other ball. We can also include this distinction in the `sample(...)` function:

```
ordered_probs <- c(1/6, 1/6, 1/6, 2/6, 1/6)
sample(bag, 3, replace = TRUE, prob = ordered_probs)
```

```
## [1] "yellow" "white"  "blue"
```

(a)

At the University of Melbourne, 13% of undergraduate students undertook an international experience in Europe, 11% undertook an international experience outside of Europe, and 76% have not undertaken an international experience at all. Suppose you asked two students at random what overseas opportunities they have taken at the university.

Conduct your own sampling simulation based on this scenario.

```
student <- (c("Europe", "non-Europe", "none"))
probs <- c(13/100, 11/100, 76/100)
sample(student, 2, replace = TRUE, prob = probs)
```

```
## [1] "none" "none"
```