Assignment 5

DSE 512

Assigned 2022-5-3 — Due 2022-5-11 11:59pm

1. (20%) For each of the following, say whether the timer/profiler is *deterministic* or *statistical*

    1. R's `system.time()`
    2. Python's `time.perf_counter()`
    3. R's `Rprof()`
    4. Python's cProfile module

2. (20%) Which has more bytes, 1 MB or 1 MiB?

3. (30%) The numeric computing matrix classes in R and Numpy support double precision. Each number stored in double precision consumes 8 bytes. Suppose you have a numeric matrix with 500 columns stored in double precision (each element is a `double`). How many GiB are consumed by this matrix if it has:

    1. 100,000 rows
    2. 10,000,000 rows
    3. 1,000,000,000 rows

4. (30%) In lecture 24, we saw that the dominant kernel in R's `prcomp()` was `La.svd()`. What is the dominant kernel in R's `lm()` with similarly sized data, i.e.:

```
m = 10000
n = 250
x = matrix(rnorm(m*n), nrow=m, ncol=n)
y = rnorm(m)

lm(y~x)
```