CYRR 304 **Name:**

Homework 2, Spring 2024

"*Facts are stubborn things, but statistics are pliable.*" MARK TWAIN

Homework 2 has questions 1 through 5 with a total of of 5 points. Your recorded score will be scaled to twenty points.
The point value for each question or part of a question is in the box following each question or part of a question. Neatly write your answers on your own paper, being careful to clearly label each part of each question. Digitize your work as a pdf, and turn it into Canvas. This work is due **Saturday 27 January** at 11:59 PM.

The harmonic mean HM of nonzero numbers $x_1, x_2, \cdots, x_n$ is defined as

$$\mathrm{HM}(x_1, x_2, \cdots, x_n) = \frac{n}{\sum_{k=1}^{n} \frac{1}{x_k}}.$$

For example

$$\mathrm{HM}(5, 45) = \frac{2}{\frac{1}{5} + \frac{1}{45}} = 9.$$

The harmonic mean has the property that for all $\alpha \in \mathbf{R}$ and for all $x_1, x_2, \cdots x_n \in \mathbf{R}_{\neq 0}$, the equation

$$\mathrm{HM}(\alpha x_1, \alpha x_2, \cdots \alpha x_n) = \alpha \mathrm{HM}(x_1, x_2, \cdots, x_n).$$

It also has the property that for all $x \in \mathbf{R}_{\neq 0}$, the equation $\mathrm{HM}(x, x, \cdots, x) = x$ is an identity.

The Julia package StatsBase has a function ' harmmean' that computes the harmonic mean of the members of a vector. For example

```
julia> using StatsBase;

julia> x = 2.0^-1074
5.0e-324

julia> harmmean([x,x])
0.0
```

Given the identity $\mathrm{HM}(x, x, \cdots, x) = x$, a better output is 5.0e-324. It's not mysterious why the output is 0. The reciprocal of the denormalized number 2.0^{-1074} overflows to Inf. But in Julia 1/Inf is zero. So effectively Julia computes 2/(Inf + Inf) = 2/Inf = 0.

Here is a proposed function 'harmonicMean' that will better handle denormalized inputs

```
function harmonicMean(L)
   n = length(L)
   m = max(map(abs, L)...) # maximum abs value
   L = map(x -> m/x,L)
   m*(n / sum(L))
end
```

1. For $x = 2.0^{-1074}$, find the value of 'harmonicMean([x,x])' and 'harmonicMean([x,x])'. Does this version of the harmonic mean resolve the issue for the function from the StatsBase package?

1. In Julia arithmetic 1/Inf = 0. Using this standard, we should have

$$\text{HM}(\infty, 5, 45) = \frac{3}{\frac{1}{\infty} + \frac{1}{5} + \frac{1}{45}} = \frac{3}{\frac{1}{5} + \frac{1}{45}} = \frac{27}{2}.$$

   Evaluate harmonicMean([5, 45, Inf]). The value should be 13.5. Is it? Step through the Julia code and explain why harmonicMean([5, 45, Inf]) returns NaN. **Hint** In Julia arithmetic, we have Inf/Inf = NaN.

1. To fix the defect in harmonicMean when one more input is Inf, one approach is to delete every member of the input that is Inf before computing m. To do this, use

```
filter!(x-> x != Inf,L)  # remove Inf from L
```

   Similarly, remove every input of −Inf. Modify harmonicMean to do this.

1. Test your modified harmonicMean function with the input

$$\text{harmonicMean}[2.0^{-1074}, 2.0^{-1074}, \text{Inf}]) \tag{1}$$

   Is the output correct?

1. Using both harmonicMean and the function harmmean from StatBase, find the harmonic mean of the ten million member vector $[10, 10, 10, \ldots, 10]$. To do this use

```
harmonicMean([10.0 for k=1:10^7])
```

   and

```
harmmean([10.0 for k=1:10^7])
```

   Which function is more accurate for this test? Explain. In Julia