"*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. This work is due **Saturday 3 Feb** at 11:59 PM. For this assignment, turn in a Jupyter notebook (a IPYNB file) to Canvas

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 lots of applications to statistics, engineering, and finance. One simple example is that if Suzi drives from Kearney to North Platte at 60 mph and returns home at 70 mph, her average speed for the trip is the harmonic, not arithmetic mean of 60 and 70.

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

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

is an identity. Additionally, 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
julia> using StatsBase;

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

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

Given the identity $\text{HM}(x, x) = x$, a better output is 5.0e-324. But 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 (miss) computes

$$2/(1/5.0 \times 10^{-324} + 1/5.0 \times 10^{-324}) = 2/(\text{Inf} + \text{Inf}) = 2/\text{Inf} = 0. \tag{1}$$

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

1. 2. In Julia arithmetic $1/\text{Inf} = 0$. 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 $\text{harmonicMean}([5, 45, \text{Inf}])$. 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. 3. To fix the defect in harmonicMean when one or 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
```

The Julia function filter! modifies its argument. Similarly, remove every input of $-\text{Inf}$. Modify harmonicMean to do this.

1. 4. Test your modified harmonicMean function with the input

$$\text{harmonicMean}[2.0^{-1074}, 2.0^{-1074}, \text{Inf}]) \tag{2}$$

Is the output correct?

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

```
harmonicMean([10.0 for k=1:10^3])
```

and

```
harmmean([10.0 for k=1:10^3])
```

Which function is more accurate for this test? Why? Congratulations, you have built a better Julia function for the harmonic mean.