# problems_analysis2

Daniela Linero

27/5/2020

Load database

```
diversity <- readRDS("./output/cleaned_data/01_Filter_data_PREDICTS_Frugivores_and_Endoplants.rds")
```

Correct sampling effort

```
diversity <- yarg::CorrectSamplingEffort(diversity)
```

```
## Correcting 0 missing sampling effort values
## Rescaling sampling effort
## Correcting 84929 values for sensitivity to sampling effort
```

Calculate diversity metrics

```
sites <- diversity %>%

  # add Diversity_metric_is_valid column
  mutate(Diversity_metric_is_valid = TRUE) %>%

  # Calculate diversity metrics
  yarg::SiteMetrics(extra.cols = c("SSB", "SSBS", "Predominant_land_use"))
```

```
## Computing site metrics for 99617 measurements
## The data contain 99 sources, 125 studies and 4991 sites
## TODO fix within-site uniqueness check
## Computing site-level values
## Computing total abundance
## Computing species richness
## Computing Simpson's diversity
## Computing Chao
## Computing Rarefied Species Richness
## Assembling site-level values
```

I'm getting this warning message: **In min(x[x > 0]) : no non-missing arguments to min; returning Inf**

This are the results of this function for SSS: BS1_2010___Page 1 1

Table 1: Results for BS1_2010___Page 1 1

|  | SSS | Total_abundance | Species_richness | Simpson_diversity |
|---|---|---|---|---|
| 1671 | BS1_2010___Page 1 1 | 10 | 10 | 10 |

The correct results should be:

```r
Correct <- diversity %>%

  filter(SSS == "BS1_2010__Page 1 1") %>%

  select(Taxon_name_entered, Measurement, SSS) %>%

  mutate(Total_Abundance = sum(Measurement), Species_richness = length(unique(Taxon_name_entered)))
```

Table 2: Correct results for BS1_2010___Page 1 1

| Taxon_name_entered | Measurement | Total_Abundance | Species_richness |
|---|---|---|---|
| Caryota urens | 2 | 10 | 3 |
| Ficus hispida | 1 | 10 | 3 |
| Xanthophyllum flavescens | 7 | 10 | 3 |

**However**, if I do all of the steps with a subset, I get the correct results and I don't get the warning message

```r
# Subset database
diversity_subset <- diversity %>%
  subset(SSS == "BS1_2010__Page 1 1") %>%
  droplevels()

# Correct sampling effort
diversity_subset <- yarg::CorrectSamplingEffort(diversity_subset)
```

```
## Correcting 0 missing sampling effort values
## Rescaling sampling effort
## Correcting 3 values for sensitivity to sampling effort
```

```r
# Calculate diversity metrics
sites_subset <- diversity_subset %>%

  # add Diversity_metric_is_valid column
  mutate(Diversity_metric_is_valid = TRUE) %>%

  # Calculate diversity metrics
  yarg::SiteMetrics(extra.cols = c("SSB", "SSBS", "Predominant_land_use"))
```

```
## Computing site metrics for 3 measurements
## The data contain 1 sources, 1 studies and 1 sites
## TODO fix within-site uniqueness check
## Computing site-level values
```

```
## Computing total abundance
## Computing species richness
## Computing Simpson's diversity
## Computing Chao
## Computing Rarefied Species Richness
## Assembling site-level values
```

Table 3: Results if I calculate the metrics using a subset of the database

| SSS | Total_abundance | Species_richness | Simpson_diversity |
|---|---|---|---|
| BS1_2010___Page 1 1 | 10 | 3 | 1.851852 |