

Relatório de publicação dos discentes e egressos do PPG Ecologia: Teoria, Aplicação e Valores da UFBA - quadriênio 2017-2020

Prof. Dr. Bruno Vilela

2022-09-21

Índice

| | | |
|----------|---|----------|
| 1 | Sobre | 1 |
| 2 | Preparação | 1 |
| 2.1 | Dados e scripts | 1 |
| 2.2 | Pacotes necessários | 1 |
| 2.3 | Funções auxiliares | 2 |
| 2.4 | Informações do lattes | 2 |
| 2.5 | Qualis | 2 |
| 2.6 | Discentes e egressos do curso | 2 |
| 3 | Análises dos itens 2.1 e 2.2 | 3 |
| 3.1 | Artigos B3-A1 e A4-A1 por mestre equivalente | 3 |
| 3.2 | Porcentagem de discentes e egressos com A4-A1 | 5 |

1 Sobre

Nesse arquivo, monitoramos os dados de publicações de discentes e egressos no quadriênio e calculamos as métricas 2.1 e 2.2 da ficha de avaliação 2017-2020 para a área de biodiversidade.

2 Preparação

2.1 Dados e scripts

Os dados e scripts disponibilizados aqui podem ser encontrados no seguinte link do github: Para replicar as análises clone o repositório do github e rode os códigos a seguir.

2.2 Pacotes necessários

Instale e carregue os seguintes pacotes para replicar a análise. Note que o pacote NUPEX foi desenvolvido no Instituto de Biologia UFBA para acompanhamento anual da produção de discentes e docentes e está disponível para download no github, conforme instruções abaixo.

```
library(tidyverse)
library(stringi)
library(magrittr)
library(scholar)
library(here)
# Para baixar o pacote NUPEX:
```

```
# devtools::install_github("BrunoVilela/NUPEX")
library(NUPEX)
```

2.3 Funções auxiliares

Para deixar esse relatório mais curto, funções auxiliares foram criadas em um outro script, carregue elas no R usando o código abaixo.

```
source(here("Scripts", "funcoes_auxiliares.R"))
```

2.4 Informações do lattes

Primeiramente, extraímos os dados dos arquivos XML de cada discente e egresso.

```
# lattes
ano_base = 2017:2020
path = here("Dados/lattes")

# Get the lattes
lattes_alunos <- get_lattes_folder(path, progress = FALSE)
```

2.5 Qualis

O qualis de todas as revistas onde discentes e egressos publicaram seus artigos foram classificadas segundo as regras estabelecidas no documento de área. Ainda assim algumas revistas não puderam ser classificadas (símbolo NA nas análises).

```
qualis <- read_csv(here("Dados/qualis", "qualis_revistas_publicadas.csv"))
qualis_estrato <- c("A1", "A2", "A3", "A4", "B1", "B2", "B3", "B4", "C")
```

2.5.1 Combinar artigos com o qualis

A seguir combinamos os artigos com o qualis das revistas.

```
papers <- lattes_alunos$papers %>%
  mutate(TITULO = gsub( "\\s*\\([^\\)]+\\)", "", `TITULO-DO-PERIODICO-OU-REVISTA` ),
         TITULO = trimws(toupper(TITULO))) %>%
  mutate(ANO = `ANO-DO-ARTIGO`) %>%
  filter(ANO < 2021) %>%
  select(NOME, `TITULO-DO-ARTIGO`, TITULO, ISSN, ANO, DOI) %>%
  left_join(y = qualis, by = c("ISSN", "TITULO")) %>%
  distinct()
```

2.6 Discentes e egressos do curso

Aqui carregamos os dados dos egressos do programa e padronizamos os nomes para as análises.

```
egressos <- read.csv(here("Dados/alunos-egressos", "egressos.csv"))
egressos$Egresso <- toupper(stri_trans_general(str = egressos$Egresso,
                                              id = "Latin-ASCII"))
```

Aqui carregamos os dados dos discentes do programa e padronizamos os nomes para as análises.

```
matriculados <- read.csv(here("Dados/alunos-egressos",
                             "matriculados_2017-2020.csv"),
                        encoding = "latin1")
```

```
matriculados$Matriculado.em <- gsub("\\\\.", "", matriculados$Matriculado.em)
matriculados$Nome <- toupper(stri_trans_general(str = matriculados$Nome,
                                              id = "Latin-ASCII"))
```

Aqui padronizamos os nomes dos lattes para as análises.

```
lattes_nomes <- toupper(stri_trans_general(str = lattes_alunos$basic$`NOME-COMPLETO`,
                                          id = "Latin-ASCII"))

## Ajustes específicos para o lattes
### Nome social
murilo <- "MURILLO MEDEIROS CARVALHO"
egressos$Egresso[egressos$Egresso == "MIRELLA MEDEIROS CARVALHO"] <- murilo
matriculados$Nome[matriculados$Nome == "MIRELLA MEDEIROS CARVALHO"] <- murilo

### Hifen
egressos$Egresso[egressos$Egresso == "YUSLENI FIERRO TOSCANO"] <- "YUSLENI FIERRO-TOSCANO"
```

3 Análises dos itens 2.1 e 2.2

3.1 Artigos B3-A1 e A4-A1 por mestre equivalente

Seguindo as instruções da ficha da área. Foram calculados a seguir: - O número de artigos completos B3 ou superior publicados por titulado mestre-equivalente;
- A qualidade da produção intelectual de discentes e egressos, aferida por artigos A4 ou superior publicados por titulado mestre-equivalente.

```
# Número de anos base
n_base <- length(ano_base)

# Vetores e matrizes vazias para guardar o resultados
pub_A4 <- numeric(n_base)
pub_B3 <- numeric(n_base)
qualificado_keep <- matrix(nrow = n_base, ncol = 10)

# Análise para cada ano base
for (i in 1:n_base) {
  # Egressos do ano
  egs <- egressos$Ano.de.Referência. > (ano_base[i] - 5) &
    egressos$Ano.de.Referência. < (ano_base[i])
  egs_tab <- egressos[egs, ]
  # Discentes do ano
  matris <- matriculados$Nome[matriculados$Matriculado.em == ano_base[i]]
  # Alunos completos
  alunos <- c(egs_tab$Egresso, matris)
  alunos <- alunos[!duplicated(alunos)]
  nomes_lattes <- lattes_alunos$basic$`NOME-COMPLETO`[match_names(alunos, lattes_nomes)]
  # Artigos do ano e dos Discentes e egressos do ano
  ano_paper <- papers$ANO == ano_base[i]
  papers_i <- (papers$NOME %in% nomes_lattes) & ano_paper
  papers_ls <- papers[papers_i, ]

  # Calcula produção A1-B3 e A1-A4
  qualificado <- table(factor(papers_ls$ESTRATO,
                             levels = qualis_estrato),
```

```

        useNA = "ifany")
qualificado_keep[i, ] <- qualificado
## Número total de mestre equivalente
docs_egressos <- 2.5 * sum(egs_tab$Nível == "Doutorado")
outros_alunos <-
  length(unique(c(matris,
                  egs_tab$Egresso[egs_tab$Nível != "Doutorado"])))
mest_equiv <- docs_egressos + outros_alunos
## Produção
A4 <- sum(qualificado[1:4])
B3 <- sum(qualificado[1:7])

pub_A4[i] <- A4 / mest_equiv
pub_B3[i] <- B3 / mest_equiv
}

```

3.1.1 Resultado Publicação A1-B3

O número de artigos completos B3 ou superior publicados por titulado mestre-equivalente no quadriênio foi de:

```
sum(pub_B3)
```

```
## [1] 1.21465
```

Note que a maneira de sumarizar esse resultado não está descrito no documento de área, sendo impossível replicar exatamente a análise. Por lógica, pressupõe-se uma soma.

3.1.2 Resultado Publicação A1-A4

A qualidade da produção intelectual de discentes e egressos, aferida por artigos A4 ou superior publicados por titulado mestre-equivalente foi de:

```
sum(pub_A4)
```

```
## [1] 0.9771472
```

Note que a maneira de sumarizar esse resultado não está descrito no documento de área, sendo impossível replicar exatamente a análise. Por lógica, pressupõe-se uma soma.

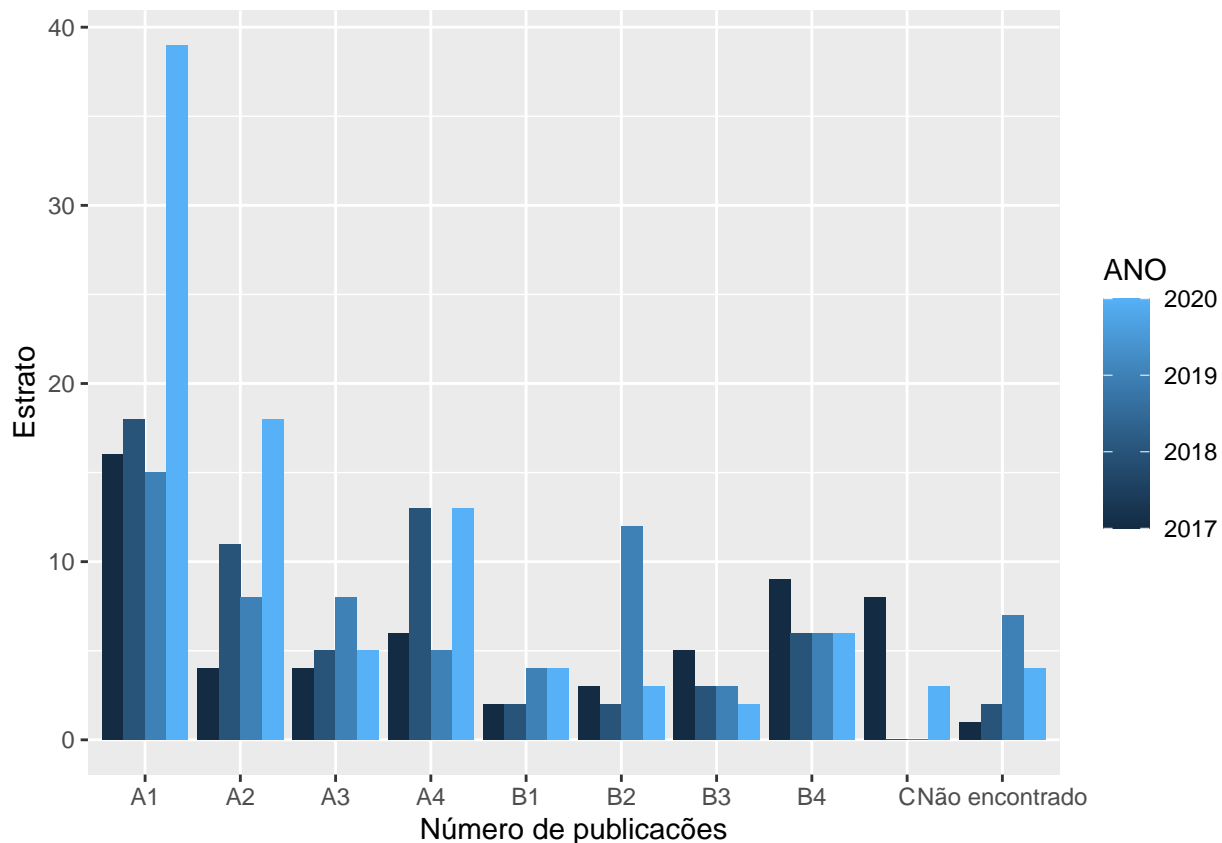
3.1.3 Distribuição de publicações ao longo do quadriênio

A publicação de discentes e egressos referentes a cada ano por estrato pode ser verificada abaixo:

```

colnames(qualificado_keep) <- c(qualis_estrato, "Não encontrado")
qualificado_keep %>%
  as_tibble() %>%
  add_column(ANO = 2017:2020) %>%
  pivot_longer(!ANO) %>%
  ggplot(aes(x = name, y = value, group = ANO, fill = ANO)) +
  geom_col(position = "dodge") +
  ylab("Estrato") +
  xlab("Número de publicações")

```



3.2 Porcentagem de discentes e egressos com A4-A1

A porcentagem de discentes e egressos com publicação A4+ foi calculada da seguinte maneira.

```
NA1.A4 <- papers %>%
  group_by(NOME, ESTRATO) %>%
  summarise(n = n()) %>%
  filter(ESTRATO %in% qualis_estrato[1:4]) %>%
  pull(NOME) %>%
  unique() %>%
  length()
```

```
## `summarise()` has grouped output by 'NOME'. You can override using the
## `.groups` argument.
```

```
Ntotal <- length(lattes_nomes)
porcA1.A4 <- NA1.A4 / Ntotal
```

3.2.1 Resultado

O resultado final obtido foi:

```
porcA1.A4
```

```
## [1] 0.5039683
```