Disciplina BIE5798: Boas práticas e ferramentas da Ciência Aberta na ecologia

Wilson de Oliveira Souza

2024-07-17

Lendo pacotes necessários para lidar com o dados

```r
library(data.table)
library(tidyverse)
library(lubridate)
library(hms)
library(here)
library(kableExtra)
```

Definindo ambiente dos dados e lendo os dados

```r
path <- here::here("data", "raw", "2021_06_fito.csv") # Para usar o arquivo exemplo upado no github, ut
data <- fread(path)

head(data, n = 10)
```

```
##                                       names           pred
##                                      <char>         <char>
##  1:  2021-06-27_10_35_13.066_3460_923.png coscinodiscus
##  2:   2021-06-27_14_05_51.211_997_305.png coscinodiscus
##  3: 2021-06-29_16_13_35.335_3696_2164.png coscinodiscus
##  4:  2021-06-29_21_14_21.670_1988_852.png coscinodiscus
##  5:  2021-06-26_15_01_51.881_3141_827.png coscinodiscus
##  6: 2021-06-14_21_14_23.207_1620_2742.png coscinodiscus
##  7:  2021-06-23_03_50_55.718_3210_896.png coscinodiscus
##  8: 2021-06-26_15_33_56.993_2732_2619.png coscinodiscus
##  9:   2021-06-26_02_32_28.667_1018_15.png coscinodiscus
## 10: 2021-06-24_13_26_11.072_3135_2518.png coscinodiscus
```

Resumo dos dados brutos

```r
# Definindo a coluna pred como fator
data$pred <- as.factor(data$pred)
```

```
# Resumo por classe
(classe_cont <- data %>%
  dplyr::group_by(pred) %>%
  dplyr::summarize(N_classes = n())) %>%
  dplyr::arrange(N_classes)
```

```
## # A tibble: 9 x 2
##   pred                         N_classes
##   <fct>                            <int>
## 1 guinardia_striata                13442
## 2 chaetoceros                      21275
## 3 rhizosolenia_robusta             35830
## 4 detritos                         37376
## 5 compl_rhizosolenia_proboscia     45719
## 6 compl_guinardia_dactyliosolen    47663
## 7 hemidiscus                       89231
## 8 sombra                          190829
## 9 coscinodiscus                  4640103
```

Organizando os dados

Aqui os dados de data e horário serão formatados

```
# Criando colunas data e hora
df <- data %>%
  as_tibble() %>%
  transmute(class = pred, date_time = str_sub(names, 1, 23) %>% ymd_hms()) %>%
  mutate(cycle_rounded = round_date(date_time, "1 hour"))

# Contabilizando dados por classe agrupados por data e hora
df <- df %>%
  dplyr::group_by(class,
                  cycle_rounded = round_date(cycle_rounded, "1 hour")) %>%
  dplyr::summarize(n = n() %>% as.integer()) %>%
  dplyr::ungroup() %>%
  print()

# Reordenando colunas e substituindo NA por 0, pois neste momento o ciclo existe, mas a classe no ciclo
df <- df %>%
  tidyr::pivot_wider(names_from = class, values_from = n)
df[is.na(df)] <- 0

# Criando colunas para todos os ciclos (horários) existentes
cycles <- tibble(cycle_rounded = seq(ymd_hm("2021-06-01 00:00"),
                                     ymd_hm("2021-06-30 23:59"),
                                     by = "1 hour"))

# Unindo todos os ciclos com os ciclos obtidos
df_full <- dplyr::full_join(cycles, df)
```

Salvando dados organizados

```r
path <- here::here("data", "processed", "2021_06_fito_processed.csv")
data <- fwrite(df_full, path)
```

Resumo dos dados processados

Aqui os dados serão resumidos considerando a data, hora e as classes.

```r
# Resumo dados por dia para cada classe
df_resumo <- df_full %>%
  pivot_longer(!cycle_rounded,
               names_to = "classes", values_to = "N") %>%
  dplyr::group_by(classes,
                  dia = lubridate::day(cycle_rounded) %>% as.factor()) %>%
  dplyr::summarize(`média` = round(mean(N, na.rm = T), 3),
                   mediana = round(median(N, na.rm = T), 0),
                   `desvio padrão` = round(sd(N, na.rm = T), 3),
                   `variância` = round(var(N, na.rm = T), 3),
                   `mínimo` = round(min(N, na.rm = T), 0),
                   `máximo` = round(max(N, na.rm = T), 0),
                   N = sum(N, na.rm = T))

df_resumo %>% kable()
```

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
chaetoceros	1	16.240	16	7.785	60.607	4	31	406
chaetoceros	2	15.292	12	17.046	290.563	3	91	367
chaetoceros	3	19.409	16	15.960	254.729	4	71	427
chaetoceros	4	25.125	19	16.456	270.810	5	74	603
chaetoceros	5	26.042	24	12.418	154.216	5	54	625
chaetoceros	6	29.208	27	10.529	110.868	15	57	701
chaetoceros	7	32.167	31	13.605	185.101	10	62	772
chaetoceros	8	33.375	30	14.702	216.158	8	60	801
chaetoceros	9	28.958	31	13.353	178.303	10	57	695
chaetoceros	10	28.958	25	15.069	227.085	8	68	695
chaetoceros	11	35.625	36	13.223	174.853	8	62	855
chaetoceros	12	35.542	32	17.619	310.433	6	80	853
chaetoceros	13	32.125	32	17.195	295.679	9	90	771
chaetoceros	14	32.167	30	10.573	111.797	13	59	772
chaetoceros	15	23.500	22	11.026	121.565	7	65	564
chaetoceros	16	73.542	57	52.461	2752.172	13	178	1765
chaetoceros	17	64.042	60	22.619	511.607	36	109	1537
chaetoceros	18	37.250	36	13.684	187.239	18	74	894
chaetoceros	19	32.625	34	9.929	98.592	17	55	783
chaetoceros	20	32.333	28	17.357	301.275	15	92	776
chaetoceros	21	21.375	17	13.180	173.723	10	68	513
chaetoceros	22	13.917	14	5.725	32.775	4	24	334
chaetoceros	23	18.542	16	6.534	42.694	6	33	445

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
chaetoceros	24	27.042	20	18.224	332.129	12	91	649
chaetoceros	25	23.292	24	6.389	40.824	10	36	559
chaetoceros	26	24.583	24	8.667	75.123	12	41	590
chaetoceros	27	39.250	35	23.079	532.630	11	100	942
chaetoceros	28	28.917	28	7.569	57.297	17	48	694
chaetoceros	29	21.292	19	8.932	79.781	6	43	511
chaetoceros	30	15.667	15	5.858	34.319	6	32	376
compl_guinardia_dactyliosolen	1	32.640	29	16.464	271.073	12	87	816
compl_guinardia_dactyliosolen	2	23.375	20	10.668	113.810	11	51	561
compl_guinardia_dactyliosolen	3	22.455	22	11.156	124.450	4	54	494
compl_guinardia_dactyliosolen	4	24.208	24	10.384	107.824	9	48	581
compl_guinardia_dactyliosolen	5	25.833	28	8.095	65.536	7	38	620
compl_guinardia_dactyliosolen	6	30.083	32	6.338	40.167	13	39	722
compl_guinardia_dactyliosolen	7	80.208	78	40.977	1679.129	17	157	1925
compl_guinardia_dactyliosolen	8	60.917	54	29.345	861.123	18	118	1462
compl_guinardia_dactyliosolen	9	39.042	36	15.941	254.129	13	68	937
compl_guinardia_dactyliosolen	10	36.000	35	12.728	162.000	16	62	864
compl_guinardia_dactyliosolen	11	42.792	45	15.601	243.389	14	74	1027
compl_guinardia_dactyliosolen	12	45.042	43	14.603	213.259	15	68	1081
compl_guinardia_dactyliosolen	13	43.500	40	11.942	142.609	20	70	1044
compl_guinardia_dactyliosolen	14	48.833	46	16.207	262.667	12	82	1172
compl_guinardia_dactyliosolen	15	33.625	30	11.883	141.201	15	64	807
compl_guinardia_dactyliosolen	16	60.750	56	31.347	982.630	18	144	1458
compl_guinardia_dactyliosolen	17	81.542	82	17.749	315.042	44	112	1957
compl_guinardia_dactyliosolen	18	75.917	78	22.952	526.775	35	120	1822
compl_guinardia_dactyliosolen	19	71.125	77	18.080	326.897	33	102	1707
compl_guinardia_dactyliosolen	20	76.083	76	19.455	378.514	39	106	1826
compl_guinardia_dactyliosolen	21	60.417	54	19.357	374.688	32	107	1450
compl_guinardia_dactyliosolen	22	46.500	48	18.132	328.783	21	82	1116
compl_guinardia_dactyliosolen	23	52.500	49	13.622	185.565	33	83	1260
compl_guinardia_dactyliosolen	24	75.542	74	16.413	269.389	48	102	1813
compl_guinardia_dactyliosolen	25	84.042	83	21.142	446.998	42	124	2017
compl_guinardia_dactyliosolen	26	105.333	100	31.558	995.884	63	178	2528
compl_guinardia_dactyliosolen	27	155.125	150	39.563	1565.245	65	218	3723
compl_guinardia_dactyliosolen	28	177.625	178	30.541	932.766	109	227	4263
compl_guinardia_dactyliosolen	29	163.792	158	42.514	1807.476	100	246	3931
compl_guinardia_dactyliosolen	30	111.625	106	23.969	574.505	79	173	2679
compl_rhizosolenia_proboscia	1	29.640	30	14.136	199.823	4	57	741
compl_rhizosolenia_proboscia	2	45.125	43	22.905	524.636	12	91	1083
compl_rhizosolenia_proboscia	3	48.091	48	15.408	237.420	13	73	1058
compl_rhizosolenia_proboscia	4	24.667	28	9.531	90.841	6	37	592
compl_rhizosolenia_proboscia	5	13.667	14	5.990	35.884	1	26	328
compl_rhizosolenia_proboscia	6	13.042	13	3.770	14.216	6	20	313
compl_rhizosolenia_proboscia	7	14.917	16	5.233	27.384	4	23	358
compl_rhizosolenia_proboscia	8	19.208	18	5.934	35.216	8	32	461
compl_rhizosolenia_proboscia	9	26.458	26	8.851	78.346	8	48	635
compl_rhizosolenia_proboscia	10	34.833	34	13.422	180.145	8	68	836
compl_rhizosolenia_proboscia	11	41.083	42	14.911	222.341	13	69	986
compl_rhizosolenia_proboscia	12	55.625	54	19.877	395.114	21	91	1335
compl_rhizosolenia_proboscia	13	70.417	72	16.513	272.688	32	96	1690
compl_rhizosolenia_proboscia	14	74.083	72	14.984	224.514	50	109	1778

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
compl_rhizosolenia_proboscia	15	86.292	87	29.527	871.868	32	145	2071
compl_rhizosolenia_proboscia	16	100.083	106	25.001	625.036	35	148	2402
compl_rhizosolenia_proboscia	17	147.333	148	43.531	1894.928	77	222	3536
compl_rhizosolenia_proboscia	18	174.083	170	47.671	2272.514	104	253	4178
compl_rhizosolenia_proboscia	19	158.292	162	31.661	1002.389	92	208	3799
compl_rhizosolenia_proboscia	20	160.833	155	40.205	1616.406	95	255	3860
compl_rhizosolenia_proboscia	21	159.958	150	51.695	2672.389	77	269	3839
compl_rhizosolenia_proboscia	22	176.958	158	74.710	5581.607	63	317	4247
compl_rhizosolenia_proboscia	23	87.208	87	24.910	620.520	42	129	2093
compl_rhizosolenia_proboscia	24	56.333	52	18.705	349.884	27	100	1352
compl_rhizosolenia_proboscia	25	27.542	24	11.485	131.911	11	58	661
compl_rhizosolenia_proboscia	26	13.625	12	5.199	27.027	2	24	327
compl_rhizosolenia_proboscia	27	16.125	14	5.636	31.766	9	33	387
compl_rhizosolenia_proboscia	28	12.833	12	5.206	27.101	5	26	308
compl_rhizosolenia_proboscia	29	9.500	10	3.121	9.739	5	15	228
compl_rhizosolenia_proboscia	30	9.875	8	3.675	13.505	5	18	237
coscinodiscus	1	1180.400	954	865.383	748887.500	348	4575	29510
coscinodiscus	2	713.292	744	379.436	143971.520	154	1506	17119
coscinodiscus	3	623.318	626	235.690	55549.942	146	1111	13713
coscinodiscus	4	766.417	683	373.681	139637.471	230	1491	18394
coscinodiscus	5	577.625	508	360.562	130005.027	182	1711	13863
coscinodiscus	6	576.917	498	343.165	117762.514	125	1680	13846
coscinodiscus	7	1012.542	888	609.272	371211.998	234	2733	24301
coscinodiscus	8	2280.750	2020	1115.328	1243955.935	649	5040	54738
coscinodiscus	9	2544.292	2046	1414.534	2000906.303	473	6731	61063
coscinodiscus	10	2071.375	2126	728.448	530637.027	718	3823	49713
coscinodiscus	11	2001.000	1998	700.162	490226.174	632	3530	48024
coscinodiscus	12	2099.292	1834	901.401	812523.607	691	3863	50383
coscinodiscus	13	2510.417	2650	831.229	690940.949	847	4481	60250
coscinodiscus	14	3403.708	3310	1319.907	1742154.389	1465	7057	81689
coscinodiscus	15	3153.375	2869	1264.975	1600160.505	1501	6184	75681
coscinodiscus	16	3920.542	3984	1185.258	1404837.650	2065	7328	94093
coscinodiscus	17	6604.958	6522	1700.047	2890160.911	3644	10051	158519
coscinodiscus	18	8089.833	6860	3412.312	11643873.362	4366	16304	194156
coscinodiscus	19	8017.500	7357	2495.037	6225211.739	4719	15811	192420
coscinodiscus	20	6967.708	6176	2373.745	5634666.303	3965	13372	167225
coscinodiscus	21	8953.583	7994	3440.659	11838137.036	4156	16571	214886
coscinodiscus	22	6235.458	5197	3972.803	15783162.520	2169	17718	149651
coscinodiscus	23	11200.542	10766	4737.365	22442622.955	4493	22249	268813
coscinodiscus	24	14586.125	13702	5143.567	26456278.984	5725	23477	350067
coscinodiscus	25	12243.417	11970	3388.528	11482120.514	6279	20060	293842
coscinodiscus	26	15854.875	15370	5577.513	31108646.027	8802	27863	380517
coscinodiscus	27	18509.250	16490	6259.203	39177617.239	11143	30811	444222
coscinodiscus	28	19244.083	20302	4897.469	23985204.688	9002	25991	461858
coscinodiscus	29	17710.708	18191	4744.755	22512695.781	9090	26485	425057
coscinodiscus	30	9687.083	9156	2637.832	6958158.080	6136	17493	232490
detritos	1	31.440	34	11.072	122.590	4	48	786
detritos	2	32.583	26	20.868	435.471	5	97	782
detritos	3	31.409	32	12.595	158.634	9	53	691
detritos	4	28.833	28	9.907	98.145	15	50	692
detritos	5	29.417	28	10.862	117.993	9	53	706

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
detritos	6	29.833	30	10.433	108.841	12	53	716
detritos	7	33.417	32	11.470	131.558	14	61	802
detritos	8	41.375	40	17.759	315.375	12	91	993
detritos	9	40.583	44	12.683	160.862	14	64	974
detritos	10	35.250	32	11.829	139.935	20	62	846
detritos	11	41.667	45	11.694	136.754	17	57	1000
detritos	12	41.583	42	11.310	127.906	12	58	998
detritos	13	38.792	37	10.950	119.911	20	67	931
detritos	14	45.417	46	8.910	79.384	25	61	1090
detritos	15	37.792	36	12.656	160.172	20	71	907
detritos	16	94.292	75	65.492	4289.259	20	258	2263
detritos	17	72.292	67	18.360	337.085	48	122	1735
detritos	18	63.042	66	14.496	210.129	33	101	1513
detritos	19	54.625	52	14.776	218.332	30	88	1311
detritos	20	53.083	50	14.753	217.645	33	86	1274
detritos	21	49.083	44	15.039	226.167	32	77	1178
detritos	22	42.333	43	15.485	239.797	19	79	1016
detritos	23	68.458	67	20.949	438.868	33	113	1643
detritos	24	83.083	78	31.467	990.167	45	177	1994
detritos	25	70.458	72	17.129	293.389	35	104	1691
detritos	26	73.542	68	15.469	239.303	47	107	1765
detritos	27	88.750	82	29.579	874.891	45	191	2130
detritos	28	83.500	75	26.975	727.652	48	157	2004
detritos	29	70.833	66	23.286	542.232	36	128	1700
detritos	30	51.875	50	16.685	278.375	22	94	1245
guinardia_striata	1	17.200	18	7.555	57.083	6	38	430
guinardia_striata	2	16.375	15	11.313	127.984	2	50	393
guinardia_striata	3	14.227	12	7.892	62.279	3	35	313
guinardia_striata	4	9.708	9	4.850	23.520	1	24	233
guinardia_striata	5	6.583	5	3.922	15.384	1	15	158
guinardia_striata	6	8.250	8	4.099	16.804	1	18	198
guinardia_striata	7	8.292	8	4.408	19.433	1	19	199
guinardia_striata	8	9.458	9	3.822	14.607	3	17	227
guinardia_striata	9	10.625	10	3.865	14.940	2	17	255
guinardia_striata	10	10.208	11	5.340	28.520	2	23	245
guinardia_striata	11	12.583	12	4.951	24.514	5	22	302
guinardia_striata	12	11.625	12	4.362	19.027	5	23	279
guinardia_striata	13	12.917	13	4.138	17.123	2	24	310
guinardia_striata	14	12.875	12	5.136	26.375	4	26	309
guinardia_striata	15	11.792	10	5.548	30.781	4	30	283
guinardia_striata	16	38.792	27	39.612	1569.129	6	167	931
guinardia_striata	17	19.708	17	7.647	58.476	8	39	473
guinardia_striata	18	20.083	20	7.546	56.949	6	37	482
guinardia_striata	19	22.917	20	9.947	98.949	8	47	550
guinardia_striata	20	23.500	24	7.265	52.783	11	44	564
guinardia_striata	21	18.292	17	8.100	65.607	7	35	439
guinardia_striata	22	18.583	17	10.142	102.862	4	42	446
guinardia_striata	23	20.750	19	8.189	67.065	6	40	498
guinardia_striata	24	25.292	22	9.813	96.303	10	44	607
guinardia_striata	25	22.958	22	8.100	65.607	9	39	551
guinardia_striata	26	25.500	24	10.198	104.000	13	46	612

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
guinardia_striata	27	50.042	44	31.819	1012.476	19	170	1201
guinardia_striata	28	32.792	33	8.910	79.389	16	51	787
guinardia_striata	29	29.875	30	10.796	116.549	14	52	717
guinardia_striata	30	18.750	18	7.415	54.978	3	33	450
hemidiscus	1	71.680	68	33.419	1116.810	30	172	1792
hemidiscus	2	69.792	68	33.242	1105.042	22	143	1675
hemidiscus	3	75.545	68	29.358	861.879	33	142	1662
hemidiscus	4	42.500	40	19.691	387.739	11	84	1020
hemidiscus	5	28.750	26	12.159	147.848	9	59	690
hemidiscus	6	30.833	28	13.535	183.188	15	68	740
hemidiscus	7	37.833	35	17.824	317.710	13	77	908
hemidiscus	8	67.667	62	23.737	563.449	37	122	1624
hemidiscus	9	83.625	86	30.367	922.158	22	149	2007
hemidiscus	10	77.042	74	30.069	904.129	30	154	1849
hemidiscus	11	89.792	90	25.480	649.216	44	126	2155
hemidiscus	12	94.208	89	37.128	1378.520	35	171	2261
hemidiscus	13	103.458	104	28.767	827.563	41	174	2483
hemidiscus	14	104.625	102	24.546	602.505	55	159	2511
hemidiscus	15	113.792	108	31.015	961.911	75	197	2731
hemidiscus	16	139.000	142	31.366	983.826	68	207	3336
hemidiscus	17	192.417	202	53.089	2818.428	88	266	4618
hemidiscus	18	223.042	214	51.193	2620.737	147	327	5353
hemidiscus	19	204.875	196	48.260	2328.984	120	293	4917
hemidiscus	20	176.667	172	43.112	1858.667	98	256	4240
hemidiscus	21	154.750	154	49.771	2477.152	83	267	3714
hemidiscus	22	135.667	136	48.947	2395.797	58	209	3256
hemidiscus	23	127.750	123	36.387	1324.022	71	215	3066
hemidiscus	24	152.417	148	28.634	819.906	83	211	3658
hemidiscus	25	165.667	164	47.902	2294.580	89	290	3976
hemidiscus	26	182.958	170	50.744	2574.998	107	287	4391
hemidiscus	27	167.583	164	36.442	1327.993	104	256	4022
hemidiscus	28	201.833	200	30.677	941.101	134	271	4844
hemidiscus	29	224.000	214	46.076	2123.043	151	329	5376
hemidiscus	30	181.500	176	48.307	2333.565	101	303	4356
rhizosolenia_robusta	1	35.680	33	13.777	189.810	14	74	892
rhizosolenia_robusta	2	42.958	36	25.772	664.216	13	123	1031
rhizosolenia_robusta	3	34.455	35	9.405	88.450	14	48	758
rhizosolenia_robusta	4	29.625	30	10.210	104.245	7	47	711
rhizosolenia_robusta	5	34.458	35	11.703	136.955	13	58	827
rhizosolenia_robusta	6	34.083	34	13.266	175.993	13	69	818
rhizosolenia_robusta	7	30.000	28	9.283	86.174	18	49	720
rhizosolenia_robusta	8	29.417	27	11.375	129.384	11	52	706
rhizosolenia_robusta	9	33.250	35	9.312	86.717	13	48	798
rhizosolenia_robusta	10	32.458	32	10.566	111.650	16	59	779
rhizosolenia_robusta	11	40.542	40	12.639	159.737	10	66	973
rhizosolenia_robusta	12	38.417	40	12.826	164.514	8	64	922
rhizosolenia_robusta	13	42.500	40	13.240	175.304	14	87	1020
rhizosolenia_robusta	14	42.500	40	13.204	174.348	22	73	1020
rhizosolenia_robusta	15	44.833	42	12.980	168.493	23	72	1076
rhizosolenia_robusta	16	92.250	76	57.776	3338.109	23	255	2214
rhizosolenia_robusta	17	70.000	69	19.460	378.696	43	117	1680

classes	dia	média	mediana	desvio padrão	variância	mínimo	máximo	N
rhizosolenia_robusta	18	65.250	66	14.441	208.543	44	98	1566
rhizosolenia_robusta	19	73.500	72	14.885	221.565	47	120	1764
rhizosolenia_robusta	20	72.583	71	8.963	80.341	54	87	1742
rhizosolenia_robusta	21	60.125	60	19.016	361.592	30	89	1443
rhizosolenia_robusta	22	52.583	57	15.959	254.688	20	78	1262
rhizosolenia_robusta	23	56.417	53	16.526	273.123	26	88	1354
rhizosolenia_robusta	24	58.625	56	17.468	305.114	25	87	1407
rhizosolenia_robusta	25	56.708	60	13.017	169.433	32	79	1361
rhizosolenia_robusta	26	56.583	52	20.932	438.167	21	101	1358
rhizosolenia_robusta	27	70.875	62	38.628	1492.114	42	231	1701
rhizosolenia_robusta	28	54.458	50	14.554	211.824	32	85	1307
rhizosolenia_robusta	29	57.542	52	19.996	399.824	29	103	1381
rhizosolenia_robusta	30	51.625	48	16.046	257.462	23	79	1239
sombra	1	127.200	110	57.320	3285.583	45	268	3180
sombra	2	147.208	121	88.994	7919.911	19	336	3533
sombra	3	74.409	80	24.488	599.682	34	113	1637
sombra	4	99.583	85	46.259	2139.906	30	168	2390
sombra	5	100.083	100	48.518	2353.993	27	191	2402
sombra	6	169.750	138	108.449	11761.152	38	408	4074
sombra	7	166.875	170	47.772	2282.201	89	246	4005
sombra	8	123.333	133	53.380	2849.449	31	200	2960
sombra	9	96.208	92	38.175	1457.303	24	193	2309
sombra	10	88.458	88	27.187	739.129	31	139	2123
sombra	11	104.208	100	35.483	1259.042	31	193	2501
sombra	12	139.958	133	45.100	2034.042	48	227	3359
sombra	13	181.625	170	56.522	3194.766	104	307	4359
sombra	14	235.292	216	103.041	10617.433	75	476	5647
sombra	15	224.125	226	110.634	12239.940	60	393	5379
sombra	16	241.458	214	131.640	17329.042	69	648	5795
sombra	17	309.125	306	56.486	3190.723	211	427	7419
sombra	18	376.917	328	136.774	18707.123	249	699	9046
sombra	19	537.333	536	106.252	11289.449	358	768	12896
sombra	20	745.708	741	129.995	16898.824	488	1005	17897
sombra	21	516.125	415	354.173	125438.810	65	1187	12387
sombra	22	176.417	154	118.095	13946.428	44	577	4234
sombra	23	233.542	204	110.165	12136.259	94	497	5605
sombra	24	325.583	308	159.619	25478.080	113	732	7814
sombra	25	257.875	256	100.025	10005.071	124	496	6189
sombra	26	369.375	324	190.988	36476.592	147	844	8865
sombra	27	519.583	452	245.233	60139.384	197	971	12470
sombra	28	486.917	427	174.469	30439.558	243	914	11686
sombra	29	510.167	474	182.830	33426.928	261	880	12244
sombra	30	267.667	294	87.054	7578.406	126	417	6424

```r
plot_summary <- function(classe = df_full$chaetoceros,
                  ylab = "Total Chaetoceros/hour",
                  xlab = "Date",
                  xbreak = "3 days",
                  titulo = "chaetoceros") {

  lims <- as.POSIXct(strptime(c("2021-06-01 00:00",
```

```
                        "2021-06-30 00:00"),
                  format = "%Y-%m-%d %H:%M"))

ggplot(data = df_full) +
  geom_point(aes(x = cycle_rounded, y = classe),
             col = "red", size = 1) +
  geom_line(aes(x = cycle_rounded, y = classe),
            col = "black", linewidth = .5) +
  scale_x_datetime(breaks = xbreak) +
  labs(x = xlab, y = ylab,
       title = toupper(titulo)) +
  theme_test() +
  theme(axis.text.x = element_text(size = 15, face = "bold",
                                   angle = 45, hjust = 1,
                                   vjust = 1, color = "black"),
        axis.text.y = element_text(size = 15, face = "bold",
                                   color = "black"),
        axis.title.x = element_text(size = 16, face = "bold",
                                    color = "black"),
        axis.title.y = element_text(size = 16, face = "bold",
                                    color = "black"))

}

plot_summary()
```
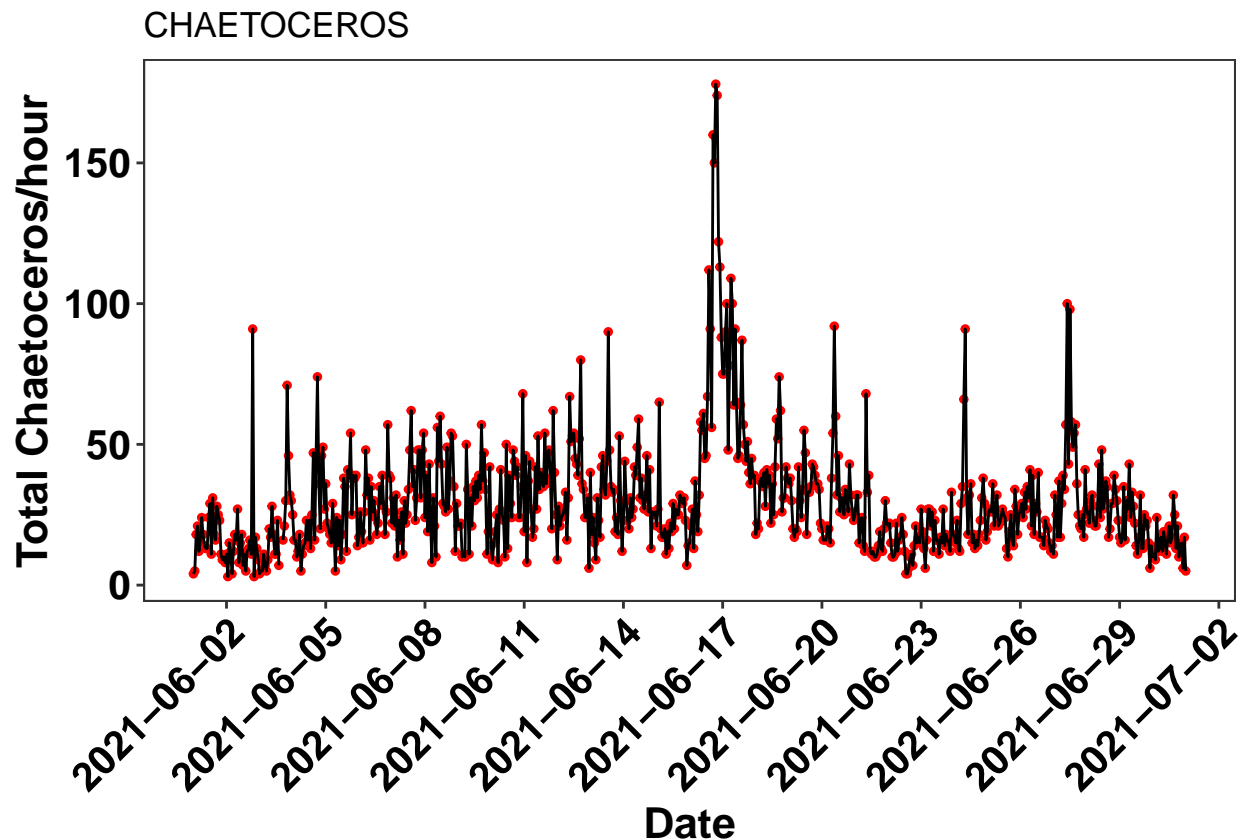
CHAETOCEROS

Análise espectral dos dados

Aqui é realizado uma análise espectral simples.

```r
names(df_full)
```

```
##  [1] "cycle_rounded"                 "chaetoceros"
##  [3] "compl_guinardia_dactyliosolen" "compl_rhizosolenia_proboscia"
##  [5] "coscinodiscus"                 "detritos"
##  [7] "guinardia_striata"             "hemidiscus"
##  [9] "rhizosolenia_robusta"          "sombra"
```

```r
# Cria dataframe com a classe especificada
df_classe <- df_full %>%
  dplyr::filter()

# Preenche NA's com o último valor observado - análise espectral não funciona com NA
# Há inúmeras formas de preencher esses NA's, essa é a mais simples
df_classe <- df_classe %>%
  tidyr::fill(chaetoceros, .direction = "down")

# Espectro com remoção da tendência - span consiste em suavização do espectro
resultado <- spectrum(I(df_classe$chaetoceros - mean(df_classe$chaetoceros)),
                      span = c(3, 3, 3, 3),
                      log = c("no"),
                      plot = F)

# Extração das frequências em uHz (X) e da variância do espectro (Y)
x <- resultado$freq / (3.6 / 1000)
y <- 2 * resultado$spec

# Plot
ggplot() +
  geom_line(aes(x = x, y = y), size = 1) +
  geom_point(aes(x = x, y = y), col = "black", size = .03) +
  scale_x_continuous(n.breaks = 10) +
  labs(x = "Frequency uHz") +
  theme_test() +
  theme(axis.text.x = element_text(size = 12, face = "bold",
                                   color = "black"),
        axis.text.y = element_text(size = 12, face = "bold",
                                   color = "black"),
        axis.title.x = element_text(size = 16, face = "bold",
                                    color = "black"),
        axis.title.y = element_text(size = 16, face = "bold",
                                    color = "black"),
        strip.text = element_text(size = 14, face = "bold",
                                  color = "black"))
```