

HB1103 Bongo Data

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Libraries

```
library(tidyverse)
library(lubridate)
```

Reading in data.

```
dta_bg <- read_csv("bongo_20180604_edit.csv")
```

Showing all column names and checking date format.

```
names(dta_bg)
```

```
## [1] "opsid" "zplk_pk_seq"
## [3] "taxa_004" "cruise_name"
## [5] "station" "event_number"
## [7] "net_number" "zoo_aliquot"
## [9] "zooplankton_count" "conc_10m2"
## [11] "conc_100m3" "vial_number"
## [13] "zoo_stage_000" "zoo_stage_013"
## [15] "zoo_stage_020" "zoo_stage_021"
## [17] "zoo_stage_022" "zoo_stage_023"
## [19] "zoo_stage_024" "zoo_stage_028"
## [21] "zoo_stage_029" "zoo_stage_030"
## [23] "zoo_stage_050" "zoo_stage_051"
## [25] "zoo_stage_054" "zoo_stage_999"
## [27] "net_pk_seq" "gear"
## [29] "gear_volume_filtered" "sample_jar_count"
## [31] "sample_type" "flowmeter_calibration_factor"
## [33] "flowmeter_revolutions" "flag_region"
## [35] "flag_clogging" "net_comment"
## [37] "event_pk_seq" "operation"
## [39] "event_date" "formatted_event_date"
## [41] "latitude" "longitude"
## [43] "region" "region_nvl"
## [45] "plankton_stratum" "groundfish_stratum"
## [47] "surface_temperature" "surface_temperature_nvl"
## [49] "tow_profile" "tow_protocol"
## [51] "tow_maximum_depth" "bottom_depth_max_wire_out"
## [53] "bio_volume_zoo" "bio_volume_zoo_1m2"
## [55] "bio_volume_zoo_100m3" "bio_volume_large_conc_1m2"
## [57] "bio_volume_large_conc_100m3" "total_zoo_count"
## [59] "total_large_zoo_count" "total_large_zoo_length"
## [61] "large_zoo_mean_length" "taxa_name"
```

And checking the data format.

```
head(dta_bg$formatted_event_date)
```

```
## [1] "6/4/2011 9:07" "6/4/2011 9:07" "6/4/2011 9:07" "6/4/2011 9:07"
## [5] "6/4/2011 9:07" "6/4/2011 9:07"
```

Initial cleaning.

```
dta_bg <- dta_bg %>%
  mutate(DT.UTC = mdy_hm(formated_event_date), tz = "UTC",
         Date = date(mdy_hm(formated_event_date))) %>%
  arrange(DT.UTC) %>% # to put in time-order
  mutate( leg = factor(case_when(Date <= ymd("2011-06-22") ~ 1,
                                Date >= ymd("2011-06-27") & Date <= ymd("2011-07-15") ~ 2,
                                Date >= ymd("2011-07-20") ~ 3)))

dta_bg %>%
  group_by(leg) %>%
  summarize(`Number of Stations per Leg` = n_distinct(station)) %>%
  ungroup()
```

```
## # A tibble: 3 x 2
##   leg `Number of Stations per Leg`
##   <fctr>                <int>
## 1     1                   30
## 2     2                   30
## 3     3                   25
```

Reading in functional group data. I initially grouped the critters and Nadine checked them for me.

```
fnct_grp <- read_csv("plankton_fnct_grp_from_Nadine.csv")
```

```
## Parsed with column specification:
## cols(
##   taxa_name = col_character(),
##   functional_group = col_character(),
##   notes = col_character()
## )
```

Joining the functional group file to the data file.

```
dta_bg_jn <- dta_bg %>%
  left_join(fnct_grp, by = "taxa_name")

dta_bg_jn %>%
  group_by(functional_group) %>%
  summarise(cnt = n()) %>%
  ungroup()
```

```
## # A tibble: 11 x 2
##   functional_group cnt
##   <chr> <int>
## 1 Calanus 191
## 2 Centropages 148
## 3 Chaetognaths 108
## 4 crap 19
## 5 Euphausiids 113
## 6 fish 51
## 7 gelatinous 233
## 8 OIT 499
## 9 Other copepods 1359
## 10 Pteropod 80
## 11 <NA> 5
```

Oops... where did the NAs come from? (Check the Next or '2' button if you don't see the NAs)

```
dta_bg_jn %>%  
  filter(is.na(functional_group)) %>%  
  select(taxa_name, event_number)
```

```
## # A tibble: 5 x 2  
##   taxa_name event_number  
##   <chr>      <int>  
## 1 Paguridea      301  
## 2 Paguridea      421  
## 3 Paguridea      681  
## 4 Paguridea      721  
## 5 Paguridea      961
```

Well, at least they are all the same taxa name.

Fixing the NAs in the taxa name.

```
dta_bg_jn <- dta_bg_jn %>%  
  mutate(functional_group = ifelse(taxa_name == 'Paguridea', 'OIT', functional_group))  
  
dta_bg_jn %>%  
  group_by(functional_group) %>%  
  summarise(cnt = n()) %>%  
  ungroup()
```

```
## # A tibble: 10 x 2  
##   functional_group cnt  
##   <chr> <int>  
## 1 Calanus 191  
## 2 Centropages 148  
## 3 Chaetognaths 108  
## 4 crap 19  
## 5 Euphausiids 113  
## 6 fish 51  
## 7 gelatinous 233  
## 8 OIT 504  
## 9 Other copepods 1359  
## 10 Pteropod 80
```