Elaborazioni Gambero2025

journal

Gianandrea LP

2025-07-01

Gambero 2025

Serie storica triennale su *Procambarus clarkii* del lago Trasimeno. Anni di riferimento: luglio 2018 - maggio 2021

```
# import data
pc <- read.xlsx(here("input/Gamberi morfometria 3 anni_.xlsx"),
                detectDates = T)
glimpse(pc)
```

```
Rows: 3,737
Columns: 13
$ data            <date> 2018-07-04, 2018-07-04, 2018-07-04, 2018-07-04, 2018-~
$ n               <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,~
$ sesso           <chr> "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M",~
$ lt.carapace     <dbl> 43.28, 54.11, 52.73, 81.55, 72.00, 52.21, 54.89, 59.21~
$ peso.tot        <dbl> 16.90, 34.50, 30.20, 35.40, 28.50, 33.60, 37.00, 39.10~
$ peso.addome     <dbl> 5.0, 8.2, 7.0, 8.2, 6.0, 7.5, 7.7, 7.8, 8.2, 6.0, 6.4,~
$ peso.epato      <dbl> 0.7, 1.7, 1.8, 1.9, 0.9, 2.1, 1.3, 1.7, 1.6, 1.4, 1.1,~
$ Hiw             <dbl> 4.142012, 4.927536, 5.960265, 5.367232, 3.157895, 6.25~
$ TwB             <dbl> 29.58580, 23.76812, 23.17881, 23.16384, 21.05263, 22.3~
$ SA              <chr> "SI", "SI", "SI", "SI", "SI", "SI", "SI", "SI", "SI", ~
$ muta            <chr> "NO", "NO", "NO", "NO", "NO", "NO", "NO", "NO", "NO", ~
$ maturità.gonadi <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ peso.gonadi     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
```

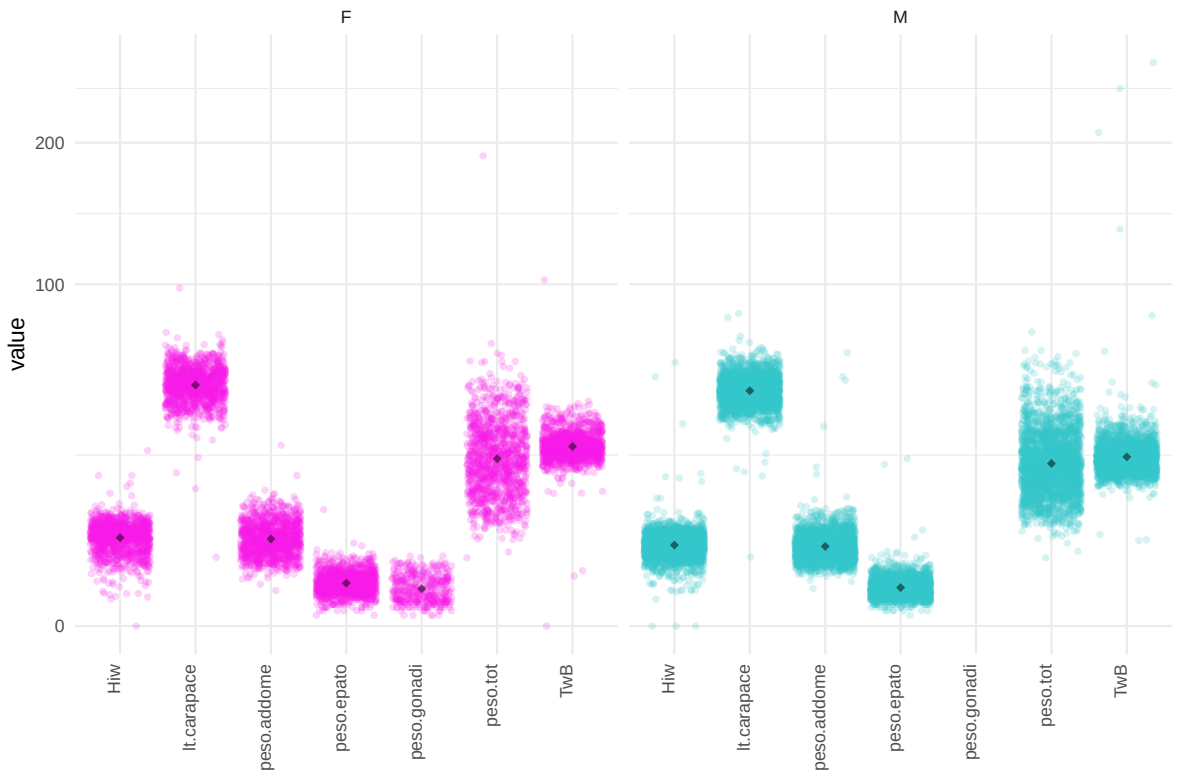
Statistiche descrittive

```
pc %>%
    select(where(is.numeric),sesso,-n) %>%
    group_by(sesso) %>%
    get_summary_stats(type = 'five_number') %>%
    kable(digits = 2)
```

sesso	variable	n	min	max	q1	median	q3
F	lt.carapace	1243	4.06	97.83	45.38	50.03	54.81
F	peso.tot	1243	4.00	189.30	16.85	23.50	31.50
F	peso.addome	1242	1.09	28.00	4.80	6.30	8.30
F	peso.epato	1242	0.10	11.60	1.10	1.50	2.20
F	Hiw	1243	0.00	26.36	5.38	7.01	8.33
F	TwB	1243	0.00	102.56	24.91	27.16	30.02
F	peso.gonadi	409	0.10	4.20	0.60	1.10	2.00
M	lt.carapace	2494	4.09	83.62	43.60	47.40	51.45
M	peso.tot	2493	4.00	74.10	16.80	21.90	28.70
M	peso.addome	2494	1.20	64.00	4.30	5.30	6.56
M	peso.epato	2491	0.10	24.00	0.90	1.20	1.60
M	Hiw	2493	0.00	59.42	4.65	5.75	6.74
M	TwB	2493	6.25	271.94	21.65	24.10	26.67

```
pc %>%
    select(where(is.numeric),sesso) %>%
    pivot_longer(lt.carapace:peso.gonadi,
                 names_to = 'var', values_to = 'value') %>%
    filter(value<800) %>% # <- per eliminare outliers
    ggplot(aes(var,value))+
    geom_jitter(aes(col=sesso),alpha=0.2, size=2)+
    scale_y_sqrt()+

    stat_summary(
        aes(group = sesso),
        fun = mean,
        geom = "point",
        shape = 18, size = 3, col='black', alpha=0.5,
        position = position_dodge(width = 0.3) # per separare se necessario
    ) +
```

```
    labs(x='')+
    facet_grid(~sesso)+
    theme(legend.position = 'none',
          axis.text.x = element_text(angle = 90, vjust = 0, hjust=1))
```




```
# add month and year
pc <- pc %>%
    mutate(month=month(data),
           year=year(data),
           date=floor_date(data, unit = 'month'))
```

Sex ratio

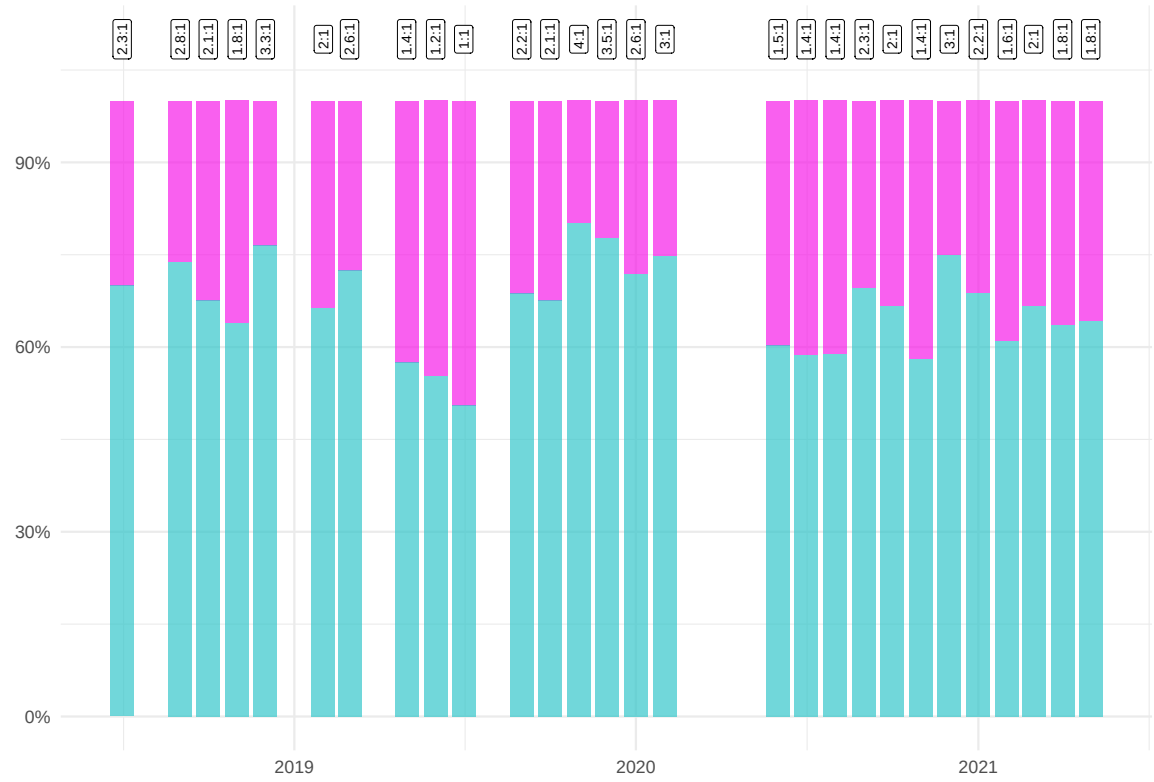
```
clark.sex <- pc %>%
    group_by(date,sesso) %>%
```

```
    summarise(n=n()) %>%
    mutate(tot=cumsum(n))
```

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

```
clark.sexratioF <- clark.sex %>%
    filter(sesso=='F')
clark.sexratioM <- clark.sex %>%
    filter(sesso=='M')
sexratio <- cbind(data.frame(ratio=round(clark.sexratioM$n/clark.sexratioF$n,1)),
                  distinct(pc,date))
```

```
pc %>%
    count(sesso,date) %>%
    group_by(date) %>%
    mutate(perc=n/sum(n)) %>%
    ggplot(aes(date,perc,group = sesso))+
    geom_bar(stat = 'identity', alpha=0.7, aes(fill=sesso))+
    scale_y_continuous(labels = scales::percent_format())+
    labs(x='',y='')+
    geom_label(data=sexratio,
               aes(x=date,y=1.1,
                   label=paste0(ratio,':1')),
               inherit.aes = FALSE,
               angle=90)+
    theme(legend.position = 'none')
```



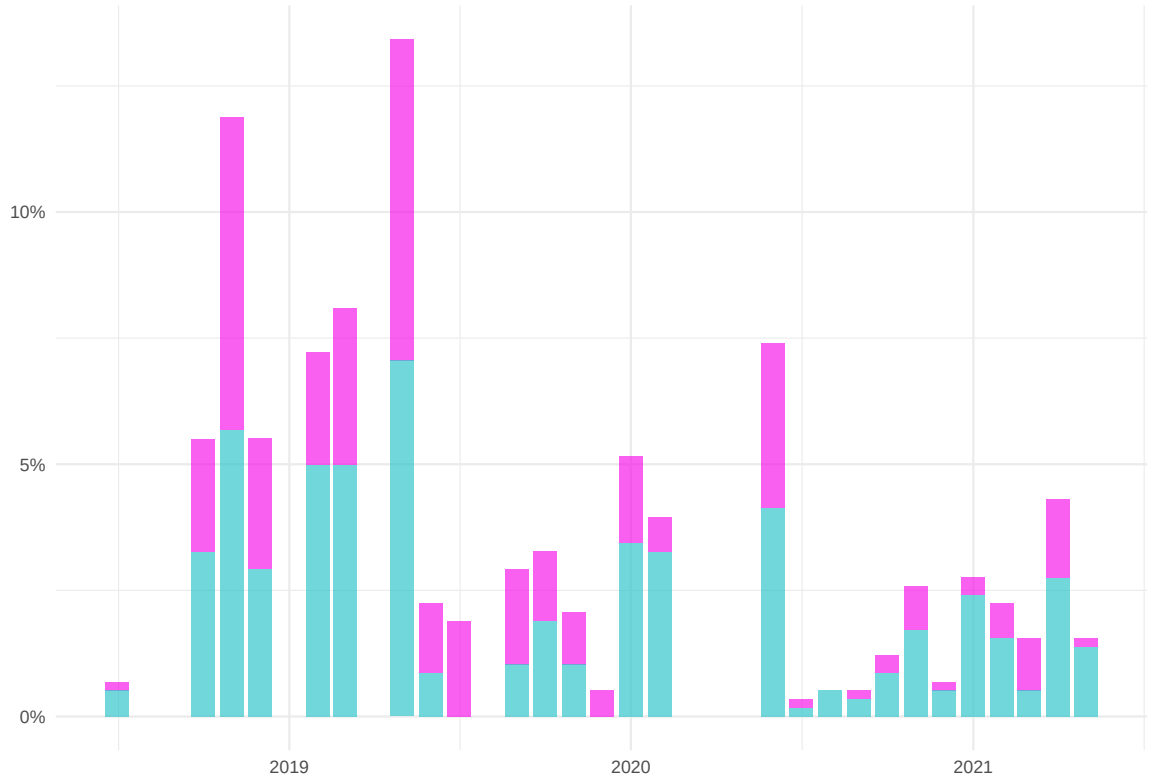
Mute

```
mute <- pc %>%
    filter(muta=='SI') %>%
    count(sesso,date) %>%
    mutate(perc=n/sum(n))

mute %>%
    mutate(perc=round(perc*100,2)) %>%
    pivot_wider(id_cols=-n,names_from = sesso,values_from = perc) %>%
    kable()
```

date	F	M
2018-07-01	0.17	0.52
2018-10-01	2.24	3.27
2018-11-01	6.20	5.68

date	F	M
2018-12-01	2.58	2.93
2019-02-01	2.24	4.99
2019-03-01	3.10	4.99
2019-05-01	6.37	7.06
2019-06-01	1.38	0.86
2019-07-01	1.89	NA
2019-09-01	1.89	1.03
2019-10-01	1.38	1.89
2019-11-01	1.03	1.03
2019-12-01	0.52	NA
2020-01-01	1.72	3.44
2020-02-01	0.69	3.27
2020-06-01	3.27	4.13
2020-07-01	0.17	0.17
2020-09-01	0.17	0.34
2020-10-01	0.34	0.86
2020-11-01	0.86	1.72
2020-12-01	0.17	0.52
2021-01-01	0.34	2.41
2021-02-01	0.69	1.55
2021-03-01	1.03	0.52
2021-04-01	1.55	2.75
2021-05-01	0.17	1.38
2020-08-01	NA	0.52

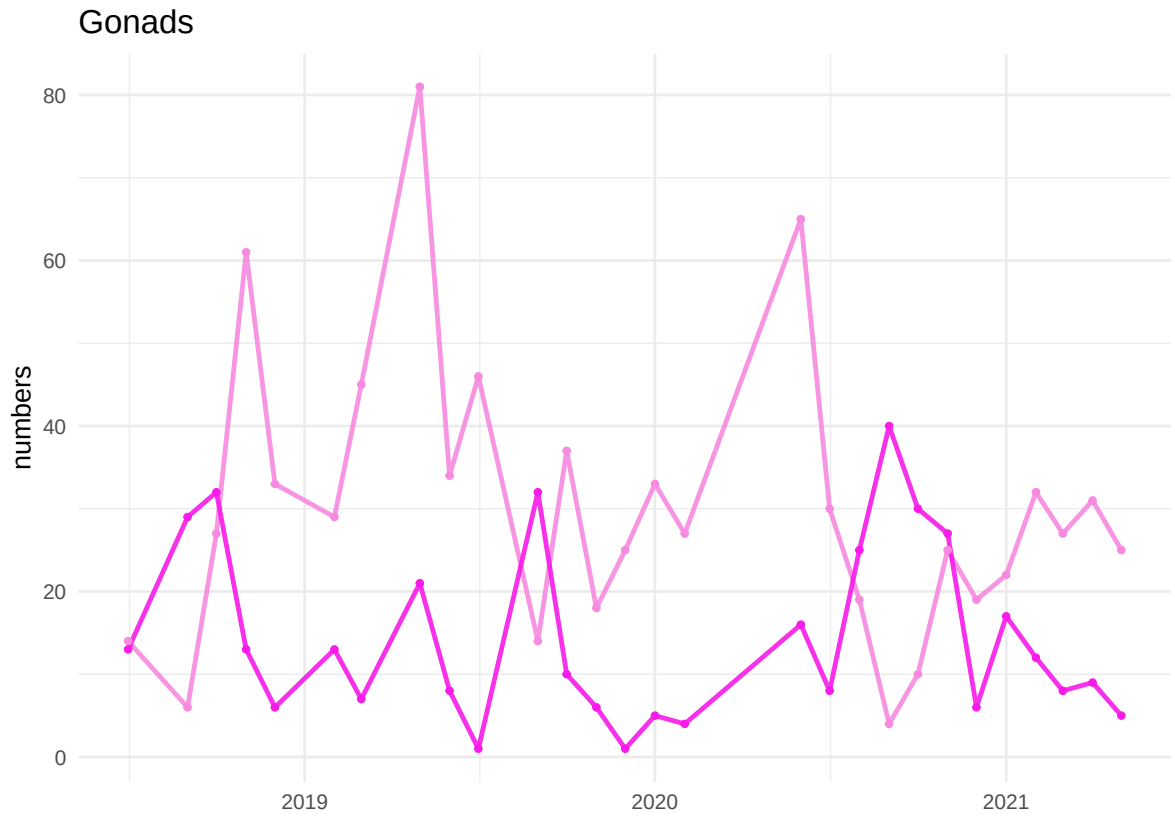
```
mute %>% ggplot(aes(date,perc,group = sesso))+
    geom_bar(stat = 'identity', alpha=0.7, aes(fill=sesso))+
    scale_y_continuous(labels = scales::percent_format())+
    labs(x='',y='')+
    theme(legend.position = 'none')
```



Gonadi

```
# corregge i nomi
pc <- pc %>% mutate(maturità.gonadi=str_trim(maturità.gonadi))

# crea una nuova variabile con stadio gonadi: 0, 1
pc.f <- pc %>%
    filter(sesso=='F') %>%
    mutate(gonadBin=ifelse(maturità.gonadi=='MATURE','1','0'))

pc.f %>%
    count(date,gonadBin) %>%
    ggplot(aes(x=date,y=n, group=gonadBin,col=gonadBin)) +
    geom_line(alpha=0.9, lwd=1.5) +
    geom_point(size=2) +
    ggtitle('Gonads') + ylab('numbers')+ xlab('')+
```

```
    scale_color_manual(values = c('#f78ae0','#f71ae8'))+
    theme(legend.position = 'none')
```

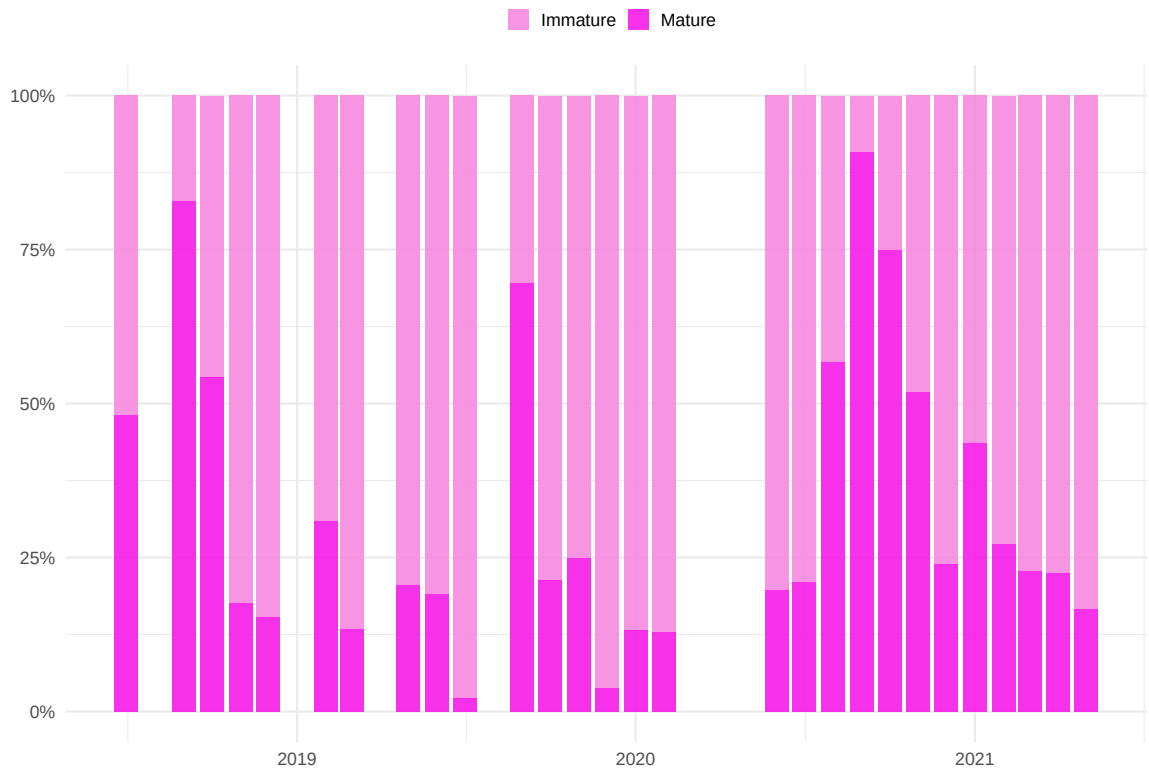



```
pc.f %>%
    count(date,gonadBin) %>%
    group_by(date) %>%
    mutate(perc=round(n/sum(n)*100,2)) %>%
    pivot_wider(id_cols = date,
                values_from = perc,
                names_from = gonadBin) %>%
    kable()
```

date	0	1
2018-07-01	51.85	48.15
2018-09-01	17.14	82.86

date	0	1
2018-10-01	45.76	54.24
2018-11-01	82.43	17.57
2018-12-01	84.62	15.38
2019-02-01	69.05	30.95
2019-03-01	86.54	13.46
2019-05-01	79.41	20.59
2019-06-01	80.95	19.05
2019-07-01	97.87	2.13
2019-09-01	30.43	69.57
2019-10-01	78.72	21.28
2019-11-01	75.00	25.00
2019-12-01	96.15	3.85
2020-01-01	86.84	13.16
2020-02-01	87.10	12.90
2020-06-01	80.25	19.75
2020-07-01	78.95	21.05
2020-08-01	43.18	56.82
2020-09-01	9.09	90.91
2020-10-01	25.00	75.00
2020-11-01	48.08	51.92
2020-12-01	76.00	24.00
2021-01-01	56.41	43.59
2021-02-01	72.73	27.27
2021-03-01	77.14	22.86
2021-04-01	77.50	22.50
2021-05-01	83.33	16.67

```
pc.f %>%
    count(date,gonadBin) %>%
    group_by(date) %>%
    mutate(perc=n/sum(n)) %>%
    ggplot(aes(date,perc,group = gonadBin, fill=gonadBin))+
    geom_bar(stat = 'identity', alpha=0.9)+
    scale_y_continuous('',labels = scales::percent_format())+
    scale_fill_manual('',labels=c('Immature','Mature'),
                      values = c('#f78ae0','#f71ae8'))+
    labs(x='',y='')+
    theme(legend.position = 'top')
```



Attività maschi

```
unique(pc$SA) #!malefici spazi!!!!
```

```
[1] "SI"  "NO"  NA    "NO " "SI "
```

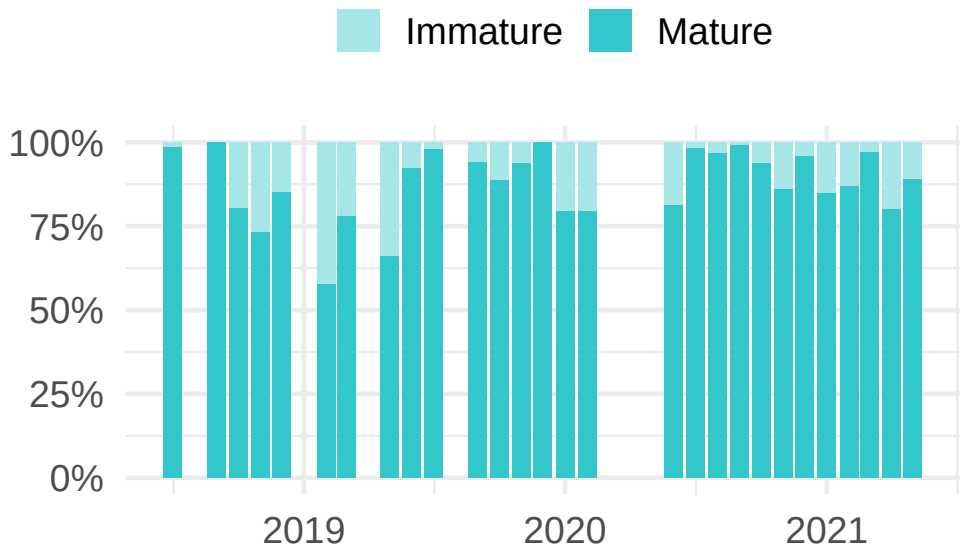
```
pc <- pc %>% mutate(SA=str_trim(SA))

pc.m <- pc %>%
    filter(sesso=='M', !is.na(SA)) %>%
    group_by(date,SA) %>%
    count(date,SA) %>%
    group_by(date) %>%
    mutate(perc=n/sum(n))

pc.m %>%
```

```
    pivot_wider(id_cols = date,
                values_from = perc,
                names_from = SA) %>%
    kable()
```

date	NO	SI
2018-07-01	0.0158730	0.9841270
2018-09-01	NA	1.0000000
2018-10-01	0.1951220	0.8048780
2018-11-01	0.2671756	0.7328244
2018-12-01	0.1496063	0.8503937
2019-02-01	0.4216867	0.5783133
2019-03-01	0.2189781	0.7810219
2019-05-01	0.3405797	0.6594203
2019-06-01	0.0769231	0.9230769
2019-07-01	0.0208333	0.9791667
2019-09-01	0.0594059	0.9405941
2019-10-01	0.1122449	0.8877551
2019-11-01	0.0618557	0.9381443
2019-12-01	NA	1.0000000
2020-01-01	0.2061856	0.7938144
2020-02-01	0.2065217	0.7934783
2020-06-01	0.1869919	0.8130081
2020-07-01	0.0185185	0.9814815
2020-08-01	0.0317460	0.9682540
2020-09-01	0.0099010	0.9900990
2020-10-01	0.0625000	0.9375000
2020-11-01	0.1388889	0.8611111
2020-12-01	0.0400000	0.9600000
2021-01-01	0.1529412	0.8470588
2021-02-01	0.1304348	0.8695652
2021-03-01	0.0285714	0.9714286
2021-04-01	0.2000000	0.8000000
2021-05-01	0.1111111	0.8888889

```
pc.m %>%
    ggplot(aes(date,perc, fill=SA))+
    geom_col()+
    scale_fill_manual('',labels=c('Immature','Mature'),
                      values=c("#A6E6E7","#33C6CA"))+
```

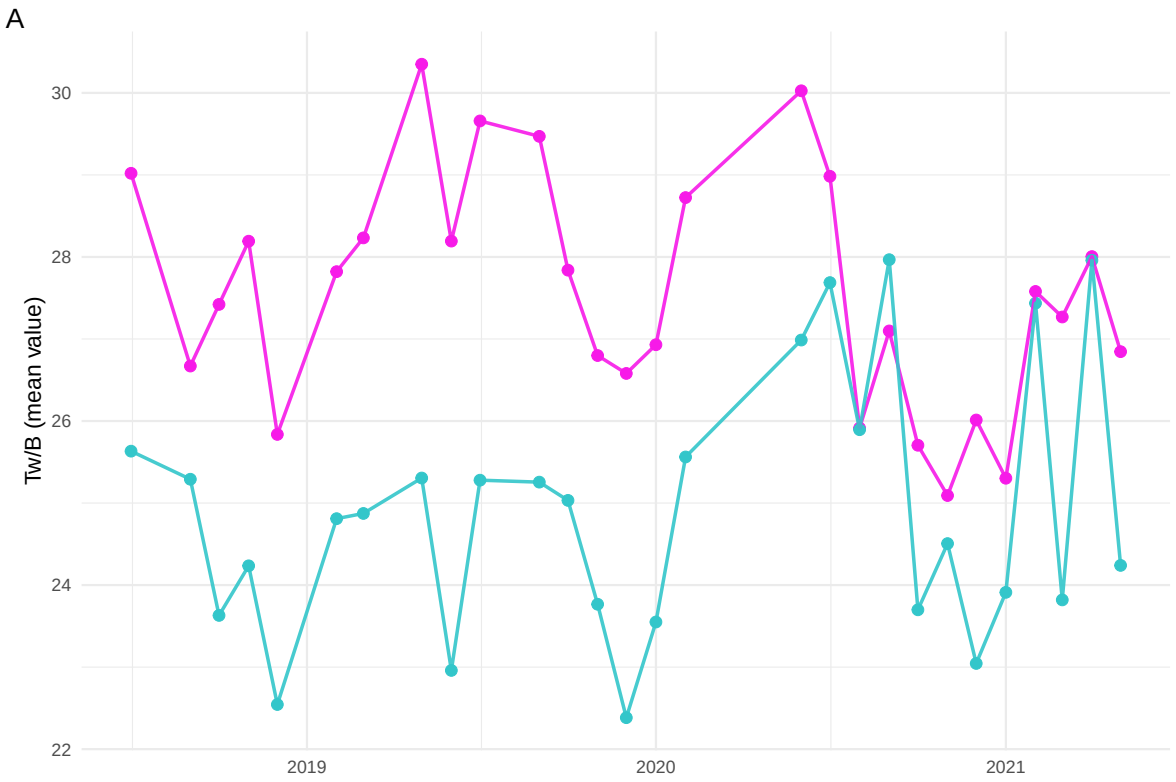
```
    scale_y_continuous('',labels = scales::percent_format())+
    labs(x='',y='')+
    theme(legend.position = 'top')
```



Tw/B e Hiw

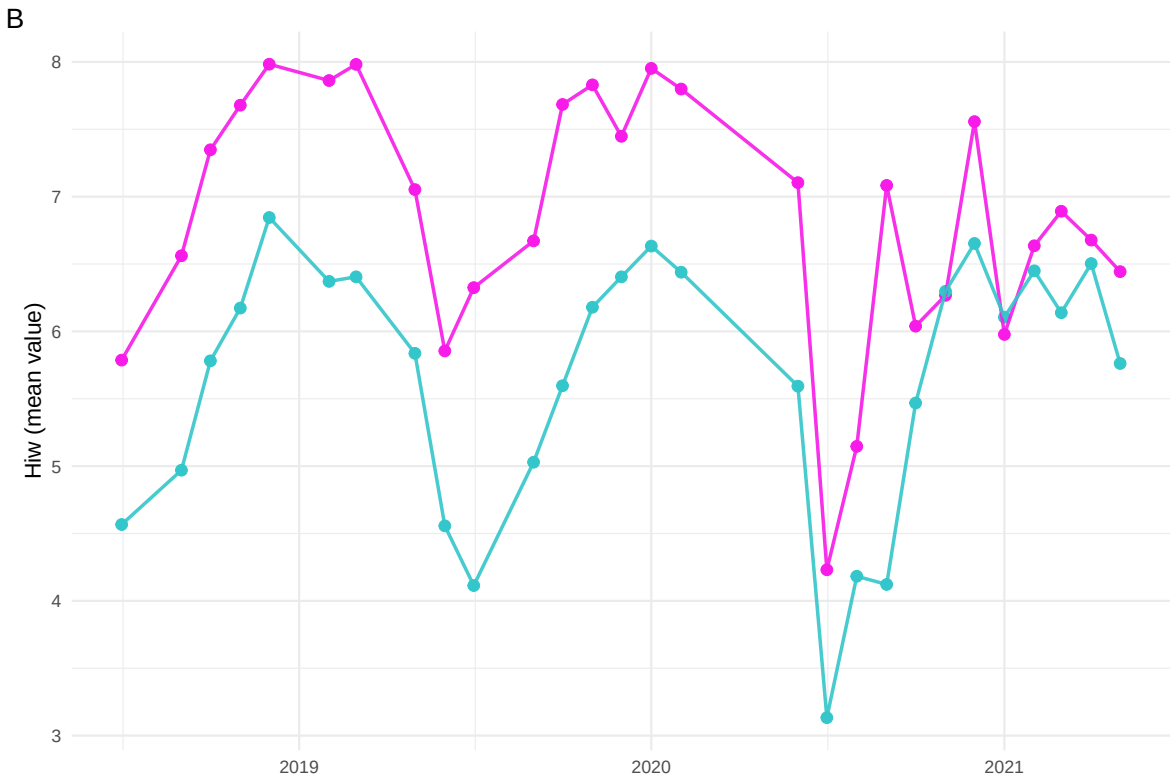
```
# glimpse(pc)
pc.twb <- pc %>%
    group_by(date,sesso) %>%
    summarise(TwB=mean(TwB,na.rm=T), .groups = 'drop')
```

```
pc.twb %>%
    ggplot(aes(date, TwB, col=sesso))+
    geom_point(size=4) +
    geom_line(alpha=0.9, lwd=1.3) +
    labs(y='Tw/B (mean value)', x='', tag = 'A')+
    theme(legend.position = 'none')
```




```
pc.hiw <- pc %>%
    group_by(date,sesso) %>%
    summarise(Hiw=mean(Hiw,na.rm=T), .groups = 'drop')
```

```
pc.hiw %>%
    ggplot(aes(date, Hiw, col=sesso))+
    geom_point(size=4) +
    geom_line(alpha=0.9, lwd=1.3) +
    labs(y='Hiw (mean value)', x='', tag = 'B')+
    theme(legend.position = 'none')
```

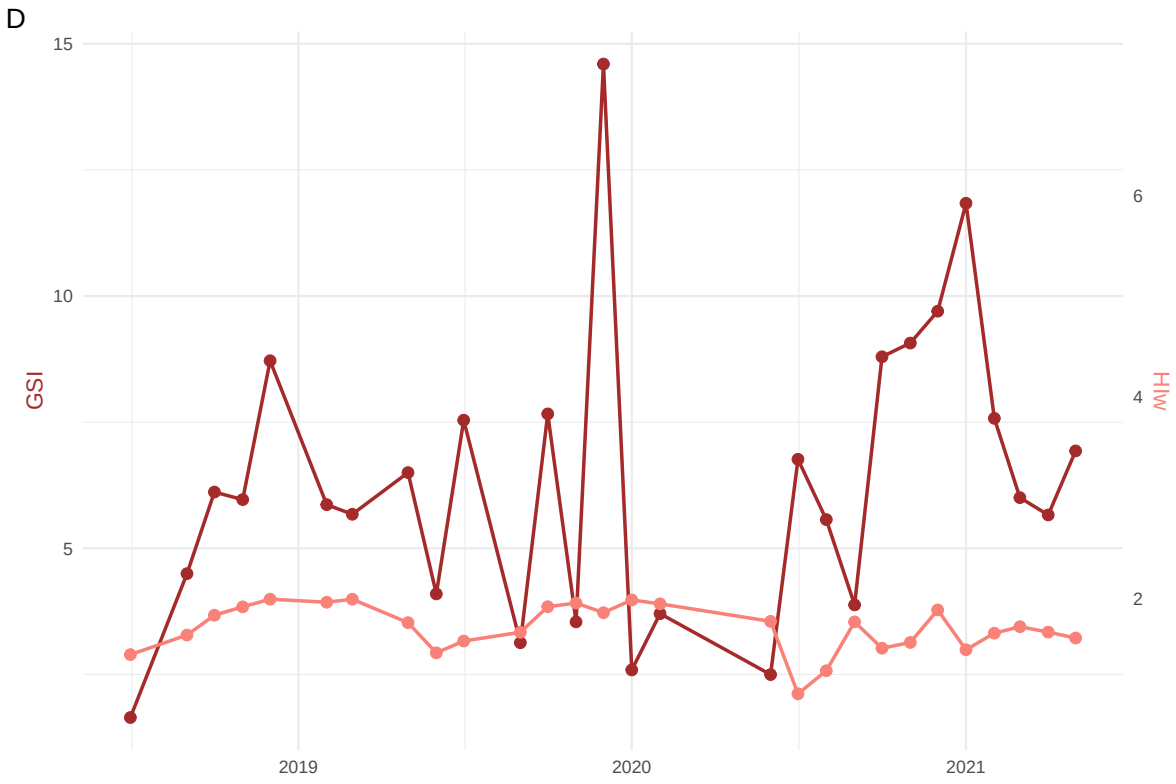


GSI

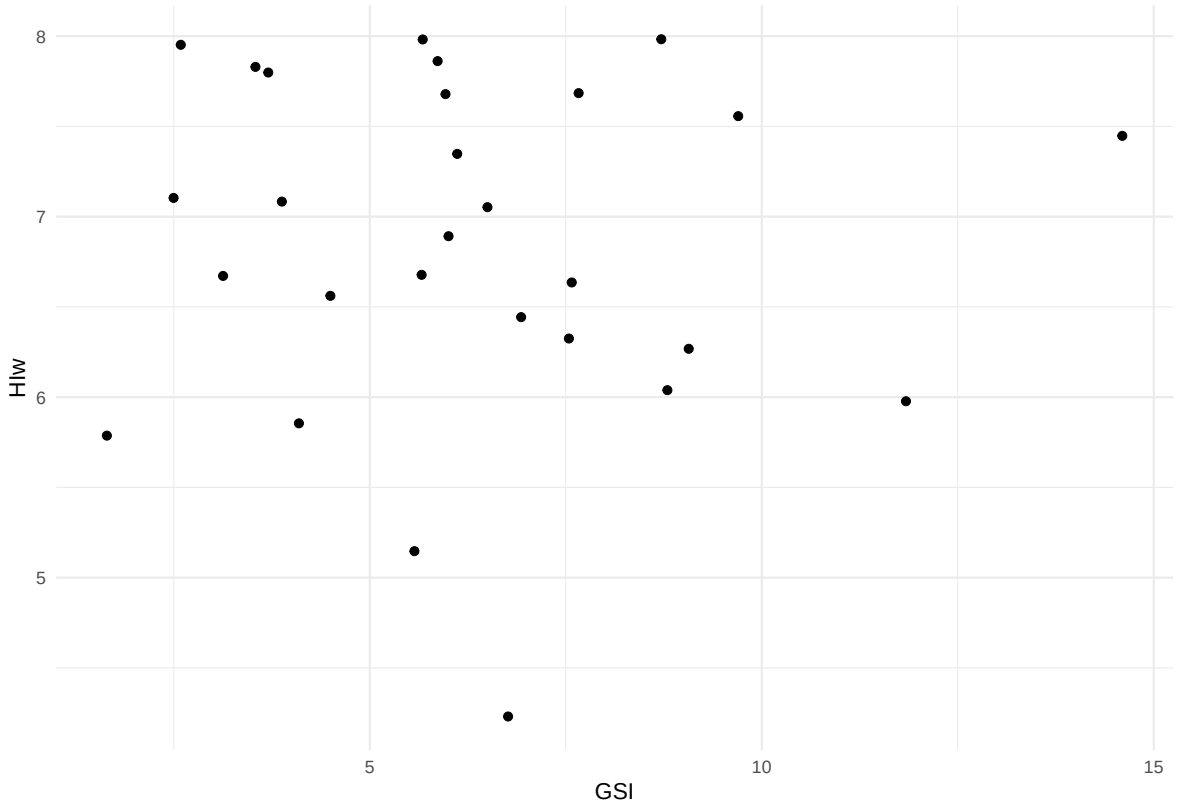
```
pc.ind <- pc %>%
    filter(sesso=='F') %>%
    mutate(GSI=peso.gonadi/(peso.tot-peso.gonadi)*100) %>%
    group_by(date) %>%
    summarise(GSI=mean(GSI, na.rm = T),
              HIw=mean(Hiw, na.rm = T))
```

```
pc.ind %>%
    ggplot(aes(date,GSI))+
    geom_point(col='brown',size=4)+
    geom_line(col='brown',lwd=1.3)+
    geom_point(aes(date,HIw/2),col='#F98178',size=4)+
    geom_line(aes(date,HIw/2),col='#F98178',lwd=1.3)+
    scale_y_continuous(sec.axis = sec_axis(~. /2, name = "HIw")) +
    labs(x='', tag = 'D') +
```

```
    theme(
        axis.title.y.left = element_text(color = "brown"),
        axis.title.y.right = element_text(color = "#F98178")
    )
```

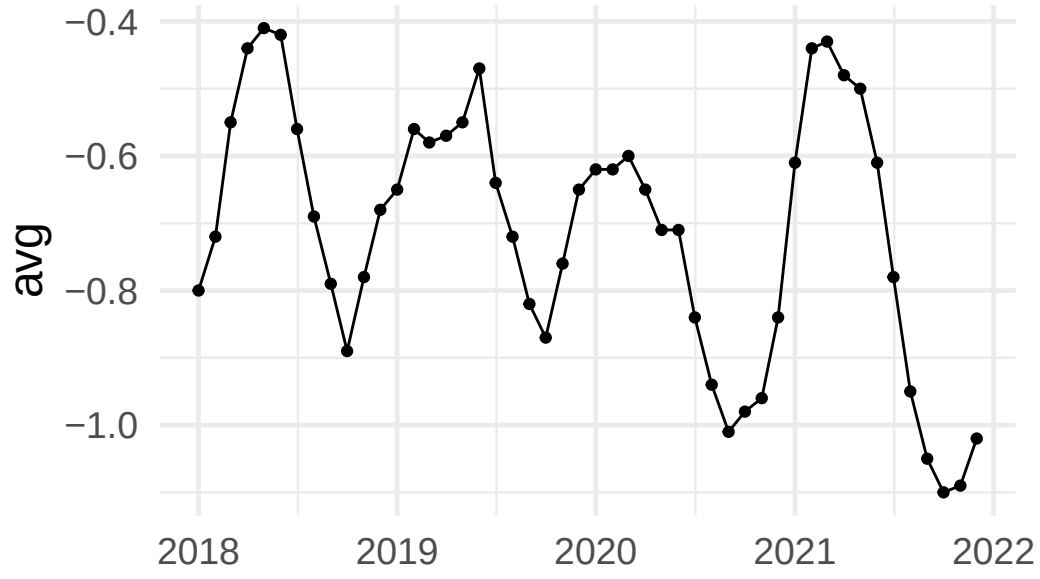



```
pc.ind %>%
    ggplot(aes(GSI,HIw))+
    geom_point(size=3)
```



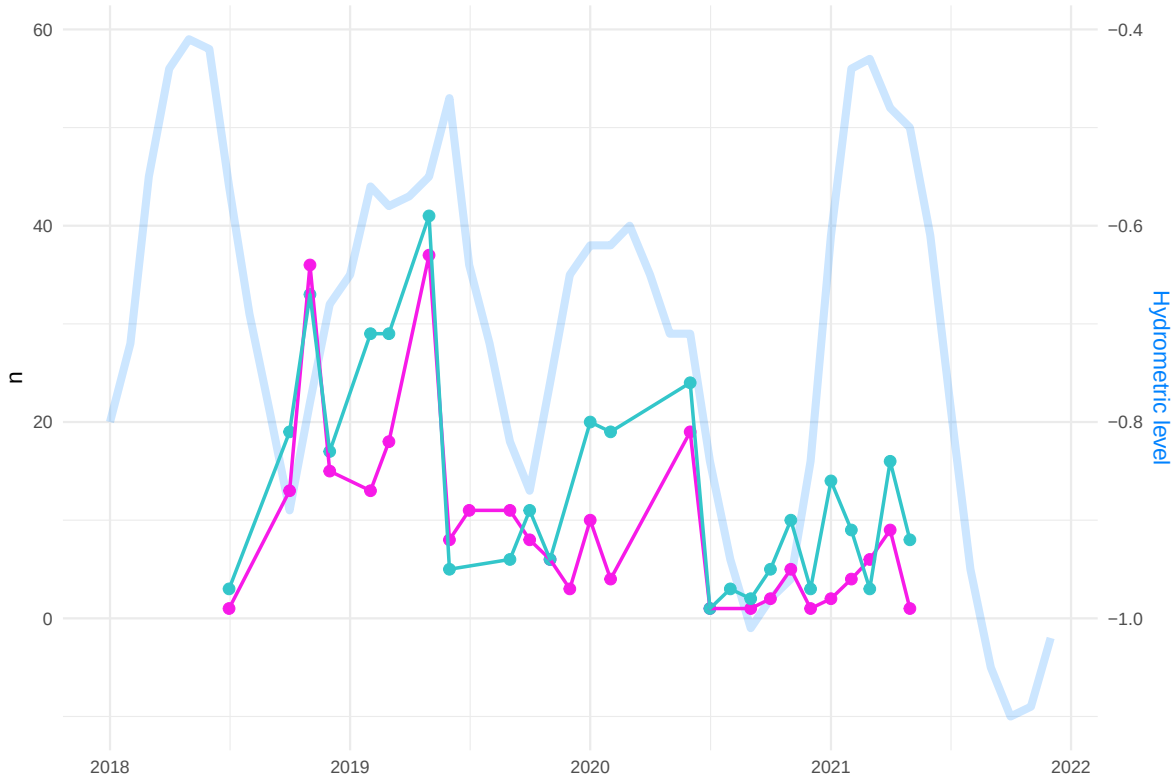
Confronto con livelli idrometrici

```
livello <- read.table(here('input/idrometro/idrometroSSavino.txt'),
                      header = T)
liv <- livello %>%
    pivot_longer(cols = X01:X12, values_to = 'avg', names_to = 'month')
liv$month <- rep(1:12,4)
liv <- liv %>%
    mutate(date=as_date(paste(year,month,'01',sep='-')))

liv %>%
    ggplot(aes(date,avg))+
    geom_point()+
    geom_line()+
    labs(x='')
```



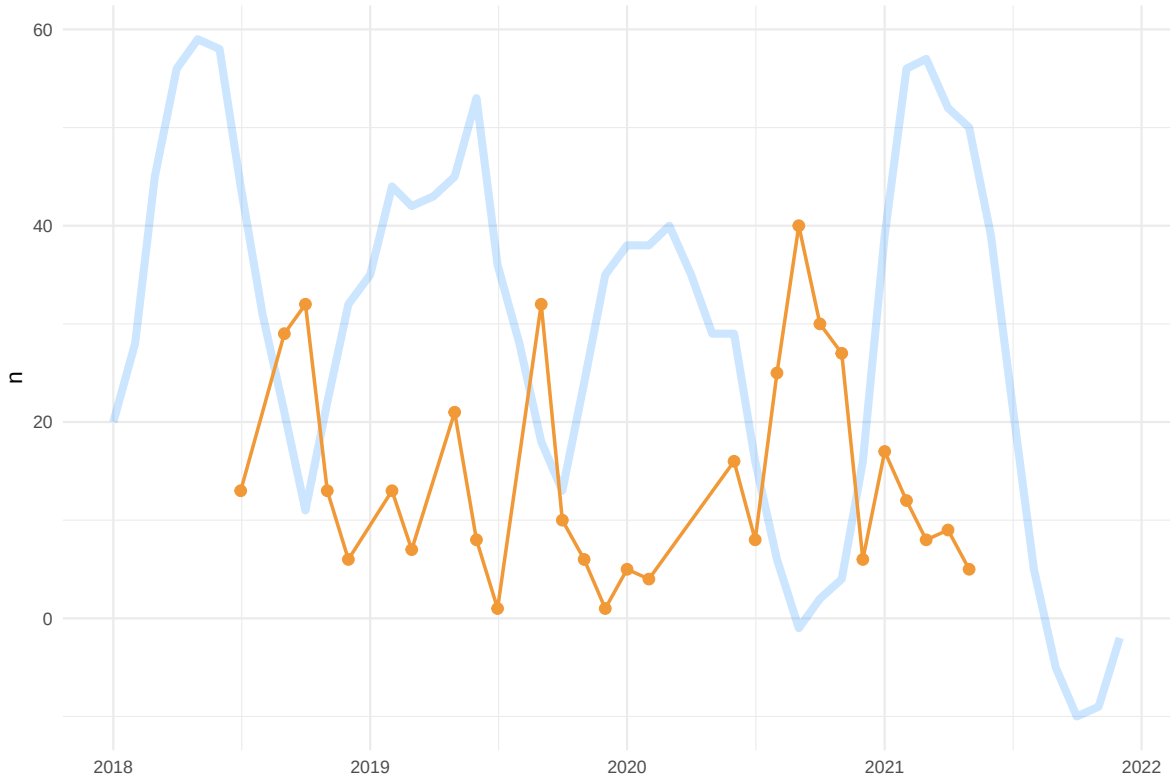
mute

```
# head(mute)

mute %>%
    ggplot(aes(date,n))+
    #geom_point(data=liv,aes(date,(avg+1)*100), col='#0082FF')+
    geom_line(data=liv,aes(date,(avg+1)*100),
              col='#0082FF', alpha=0.2, lwd=3)+
    geom_point(aes(col=sesso), size=4)+
    geom_line(aes(col=sesso), lwd=1.3)+
    scale_y_continuous(sec.axis = sec_axis(~. /100-1,
                                           name = "Hydrometric level")) +
    labs(x='') +
    theme(
        legend.position = 'none',
        axis.title.y.right = element_text(color = '#0082FF')
    )
```



Femmine mature

```
# head(pc.f)

pc.f %>%
    filter(maturità.gonadi=='MATURE') %>%
    count(date,maturità.gonadi) %>%
    ggplot(aes(date,n))+
    geom_line(data=liv,aes(date,(avg+1)*100),
              col='#0082FF', alpha=0.2, lwd=3)+
    geom_point(col='#F19837', size=4)+
    geom_line(col='#F19837', lwd=1.3)+
    labs(x='')
```


60
40
20
0
n
2018
2019
2020
2021
2022