

Exploration

Daniel Suh

11/15/2021

```
library(tidyverse)
library(magrittr)
library(lubridate)
library(here)
```

Read in pilot data

```
treatments <- read_csv(here("data/treatments.csv"))
fitness <- read_csv(here("data/fitness.csv"))
mort <- read_csv(here("data/mortality.csv"))
```

View

```
head(treatments)
```

```
## # A tibble: 6 x 7
##   ID      treatment resource temperature birth_date exposure_date isolation_date
##   <chr> <chr>          <dbl>         <dbl> <chr>         <chr>         <chr>
## 1 1A    R1T1            0.1           15 10/19/21    10/23/21    10/25/21
## 2 1B    R2T1            0.5           15 10/19/21    10/23/21    10/25/21
## 3 1C    R3T1            1            15 10/19/21    10/23/21    10/25/21
## 4 2A    R1T2            0.1           20 10/19/21    10/23/21    10/25/21
## 5 2B    R2T2            0.5           20 10/19/21    10/23/21    10/25/21
## 6 2C    R3T2            1            20 10/19/21    10/23/21    10/25/21
```

```
head(mort)
```

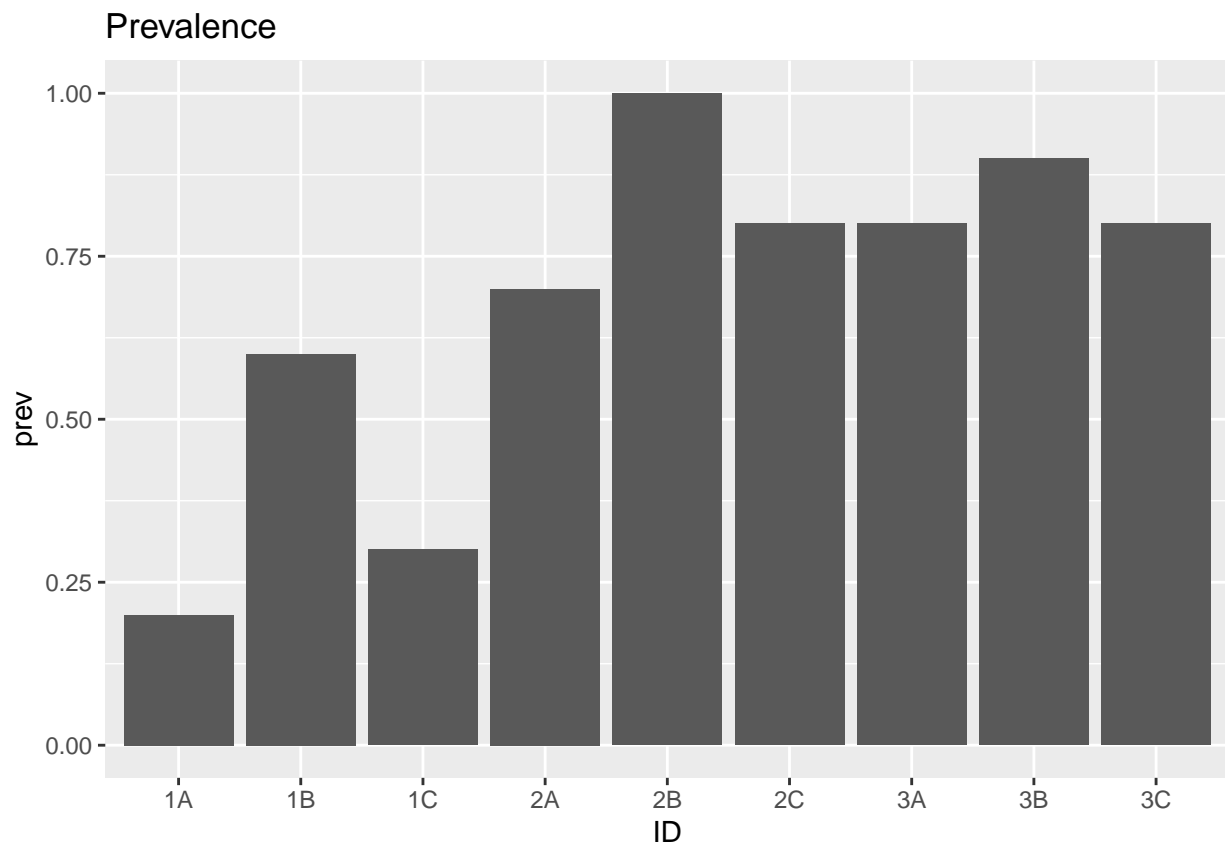
```
## # A tibble: 6 x 6
##   ID      replicate date      length spore_yield infection_status
##   <chr>         <dbl> <chr>      <dbl>         <dbl>         <dbl>
## 1 2C            1 11/6/21     NA           NA             NA
## 2 2C            2 11/4/21     73          8750            1
## 3 2C            3 11/4/21     NA          8438.            1
## 4 2C            4 11/4/21     67          7812.            1
## 5 2C            5 11/8/21     74            0            0
## 6 2C            6 11/4/21     74         15312.            1
```

```
head(fitness)
```

```
## # A tibble: 6 x 4
##   ID      replicate date      fitness
##   <chr>      <dbl> <chr>      <dbl>
## 1 2C          1 10/31/21      0
## 2 2C          2 10/31/21      6
## 3 2C          3 10/31/21      0
## 4 2C          4 10/31/21      1
## 5 2C          5 10/31/21      6
## 6 2C          6 10/31/21      5
```

Summarize prevalence

```
prevalence <- mort %>%
  group_by(ID) %>%
  summarize(n = n(), prev = sum(infection_status,
    na.rm = T)/n()) #NA's removed from numerator but still in denominator
prevalence %>%
  ggplot(., aes(x = ID, y = prev)) + geom_col() +
  labs(title = "Prevalence")
```

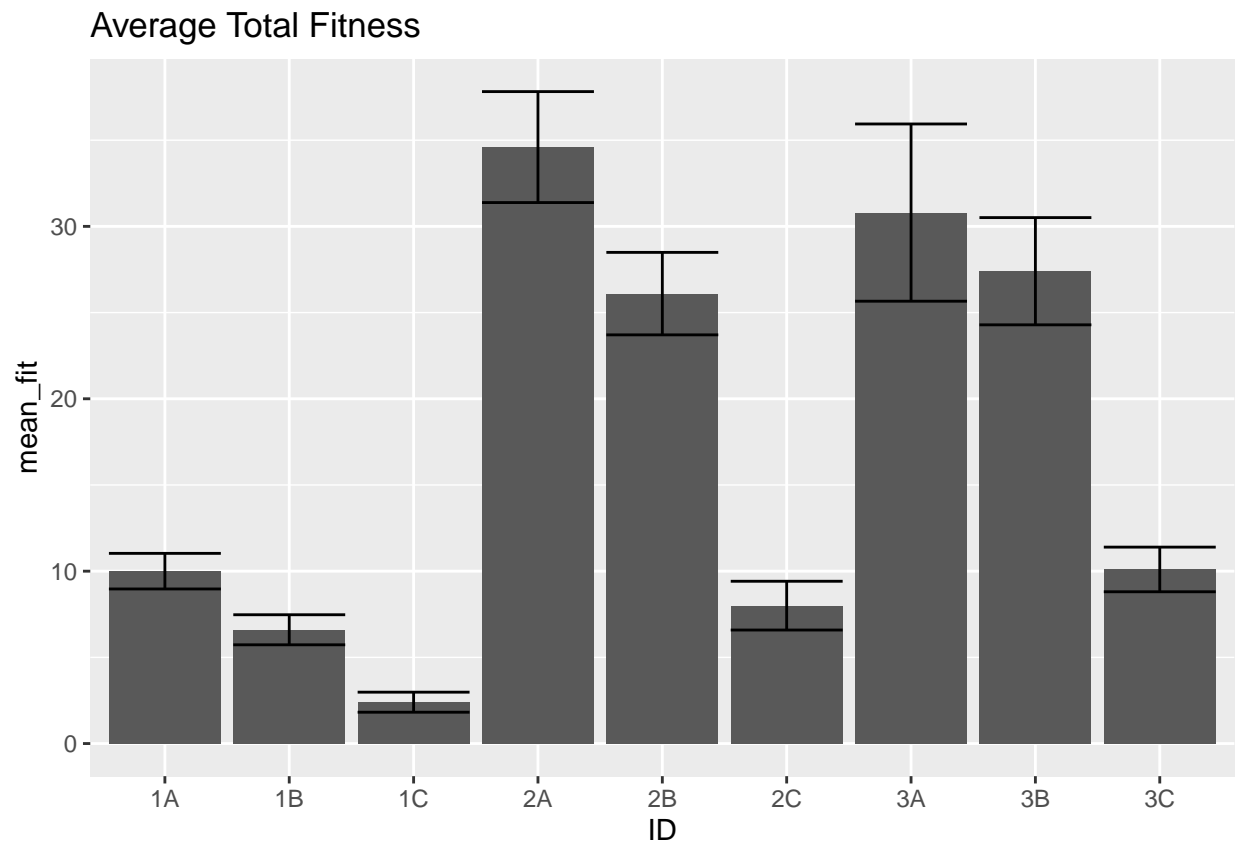


Summarize fitness

```
fit_by_ID_replicate <- fitness %>%
  group_by(ID, replicate) %>%
  summarize(total_fit = sum(fitness, na.rm = T)) # removed NAs
```

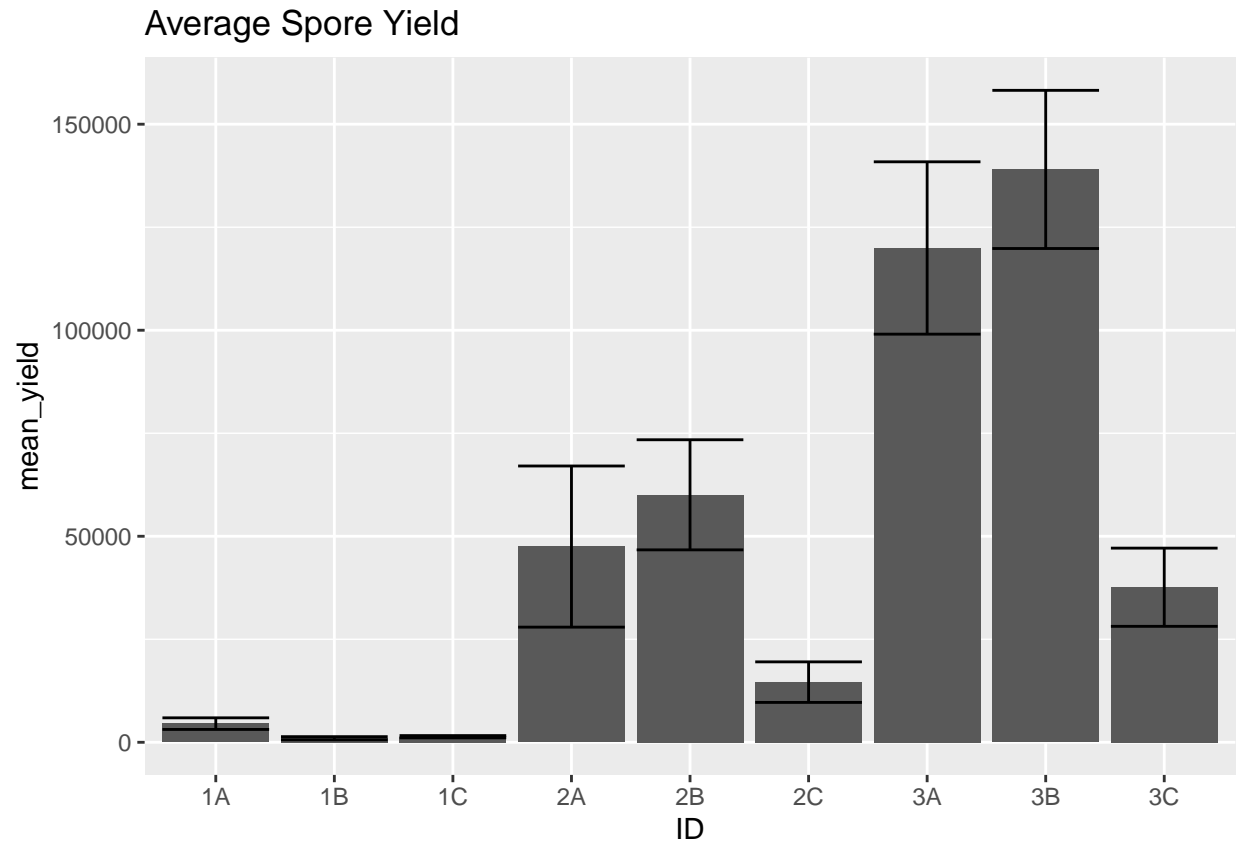
'summarise()' has grouped output by 'ID'. You can override using the '.groups' argument.

```
fit_average <- fit_by_ID_replicate %>%
  group_by(ID) %>%
  summarize(n = n(), mean_fit = mean(total_fit, na.rm = T),
            var = var(total_fit, na.rm = T), se = sqrt(var(total_fit,
            na.rm = T)/n())) # remove NAs
fit_average %>%
  ggplot(., aes(x = ID, y = mean_fit)) + geom_col() +
  geom_errorbar(aes(ymin = mean_fit - se, ymax = mean_fit +
    se)) + labs(title = "Average Total Fitness")
```

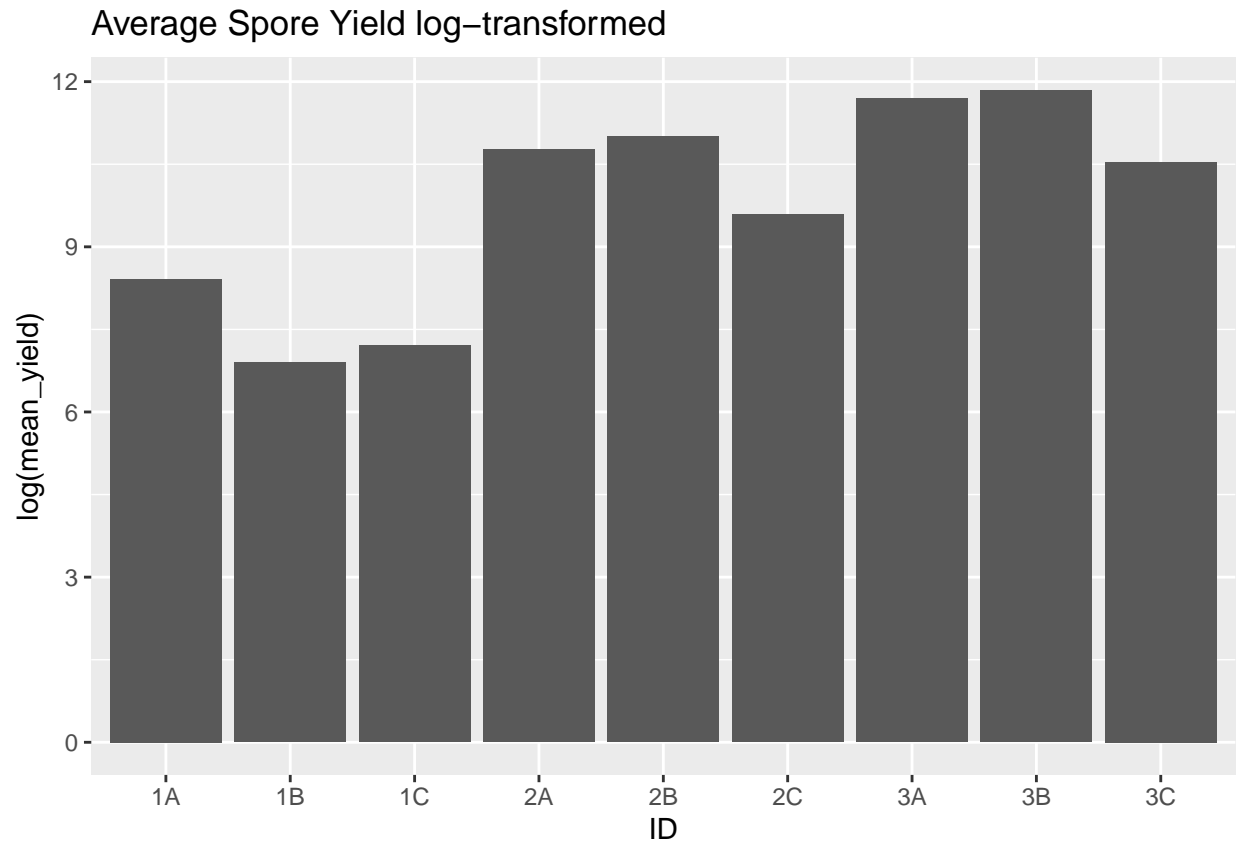


Summarize spore yield

```
spores <- mort %>% filter(spore_yield>0) %>% group_by(ID) %>%
  summarize(mean_yield = mean(spore_yield, na.rm=T),
            # NAs removed but there should be no NAs after the filter
            var = var(spore_yield, na.rm = T),
            se = sqrt(var(spore_yield, na.rm = T)/n()))
spores %>% ggplot(., aes(x=ID, y=mean_yield)) +
  geom_col() +
  geom_errorbar(aes(ymin = mean_yield-se, ymax = mean_yield+se)) +
  labs(title = "Average Spore Yield")
```



```
spores %>% ggplot(., aes(x=ID, y=log(mean_yield))) +  
  geom_col() +  
  labs(title = "Average Spore Yield log-transformed")
```



Summarize lifespan for all

```
lifespan <- left_join(mort, treatments, by = "ID")
lifespan %<>%
  select(ID, replicate, birth_date, date)
mdy(lifespan$birth_date)
```

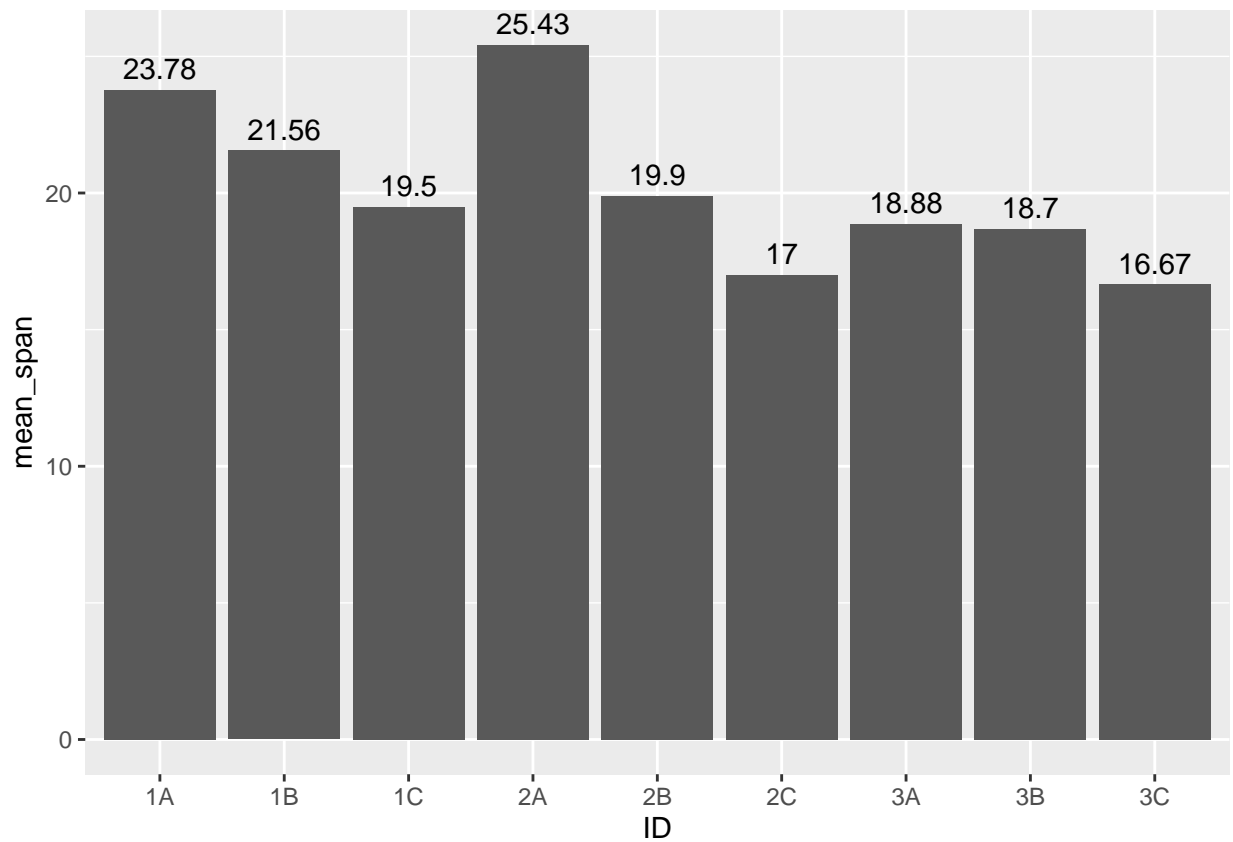
```
## [1] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [13] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [25] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [37] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [49] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [61] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [73] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
## [85] "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19" "2021-10-19"
```

```
lifespan %<>%
  transmute(ID = ID, replicate = replicate, birthday = mdy(birth_date),
            deathday = mdy(date))
lifespan %<>%
  mutate(span = deathday - birthday)

mean_span <- lifespan %>%
  group_by(ID) %>%
  summarize(mean_span = mean(span, na.rm = T))
```

```
mean_span %>%
  ggplot(., aes(x = ID, y = mean_span)) + geom_col() +
  geom_text(aes(label = round(mean_span, digits = 2)),
    vjust = -0.5)
```

Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.
 ## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.



```
# rmarkdown::render(here('scripts/exploration.Rmd'),
# output_format = 'pdf_document', output_file =
# here('pilot.pdf'))
```