Section 3 Clusters

2022-11-04

```r
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(matrixStats)
```

```
##
## Attaching package: 'matrixStats'
##
## The following object is masked from 'package:dplyr':
##
##      count
```

```r
results <- data.frame(location = c("Negative Control", "Off-Campus Fountain", "On-Campus Fountain", "Of

results$total_cfu_mean <- rowMeans(results[,2:4], na.rm = TRUE)
results <- results %>%
  mutate(total_cfu_sd = rowSds(as.matrix(.[c("total_cfu_count_rep1", "total_cfu_count_rep2", "total_cfu
  mutate(total_cfu_se = total_cfu_sd/sqrt(3))

results <- results %>%
  mutate(campus = case_when(
    str_detect(location, "^On-Campus") ~ "On-Campus",
    str_detect(location, "^Off-Campus") ~ "Off-Campus",
    str_detect(location, "Control$") ~ "Control")) %>% mutate(source = case_when(
    str_detect(location, "Fountain") ~ "Fountain",
    str_detect(location, "Sink") ~ "Sink",
    str_detect(location, "Restaurant") ~ "Restaurant",
    str_detect(location, "Control$") ~ "Control"
  ))
```

```r
t.score <- qt(0.025, df = 2, lower.tail = F)

results %>%
  ggplot() +
  geom_bar(aes(x = location, y = total_cfu_mean), stat = "identity", fill = "blue") +
  geom_point(aes(x = location, y = total_cfu_count_rep1)) +
  geom_point(aes(x = location, y = total_cfu_count_rep2)) +
  geom_point(aes(x = location, y = total_cfu_count_rep3)) +
  geom_errorbar(aes(x= location, ymin = total_cfu_mean - (t.score * total_cfu_se), ymax = total_cfu_mean
```

```
theme(axis.text.x = element_text(angle = 35, hjust = 1)) +
labs(title = "Total Colony Forming Units Across Water Sources", x = "Location", y = "Total CFU Count/
```

Total Colony Forming Units Across Water Sources
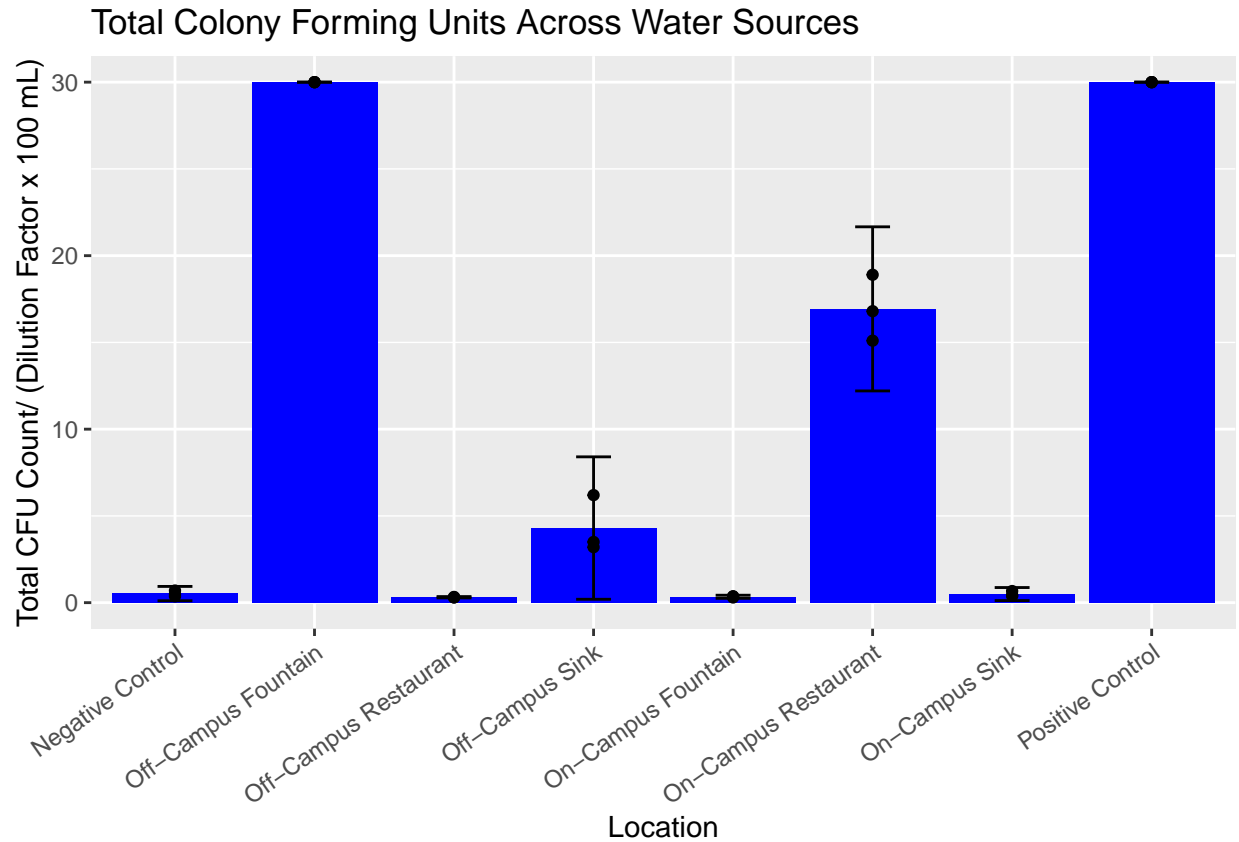

```
results %>%
  filter(campus != "Control") %>%
  group_by(campus) %>%
  summarise(campus_mean_cfu = mean(total_cfu_mean)) %>%
  ggplot(aes(x = campus, y = campus_mean_cfu, fill = source)) +
  geom_bar(position = "stack", stat="identity", fill = "blue") +
  theme(axis.title.x = element_blank()) +
  labs(title = "Off-Campus vs On-Campus Mean Total Colony Forming Units", y = "Total CFU Count/ (Diluti
```

Off–Campus vs On–Campus Mean Total Colony Forming Units
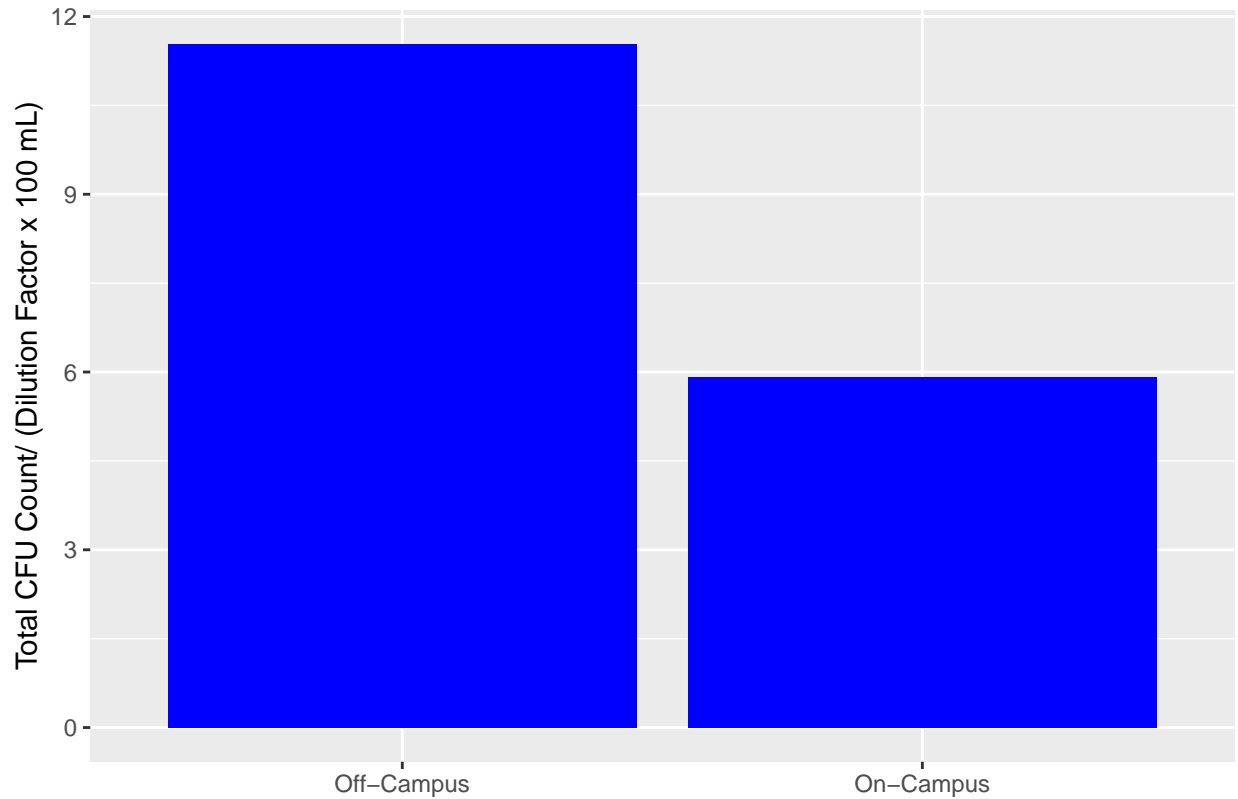

```
results %>%
  ungroup()
```

```
##               location total_cfu_count_rep1 total_cfu_count_rep2
## 1      Negative Control                 0.37                 0.50
## 2    Off-Campus Fountain               30.00                30.00
## 3     On-Campus Fountain                0.35                 0.37
## 4        Off-Campus Sink                3.50                 3.20
## 5         On-Campus Sink                0.39                 0.43
## 6 Off-Campus Restaurant                 0.30                 0.33
## 7  On-Campus Restaurant                16.80                18.90
## 8      Positive Control                30.00                30.00
##   total_cfu_count_rep3 total_cfu_mean total_cfu_sd total_cfu_se       campus
## 1                 0.70      0.5233333   0.16623277  0.095974534      Control
## 2                30.00     30.0000000   0.00000000  0.000000000   Off-Campus
## 3                 0.30      0.3400000   0.03605551  0.020816660    On-Campus
## 4                 6.20      4.3000000   1.65227116  0.953939201   Off-Campus
## 5                 0.67      0.4966667   0.15143756  0.087432514    On-Campus
## 6                 0.31      0.3133333   0.01527525  0.008819171   Off-Campus
## 7                15.10     16.9333333   1.90350554  1.098989435    On-Campus
## 8                30.00     30.0000000   0.00000000  0.000000000      Control
##        source
## 1     Control
## 2    Fountain
## 3    Fountain
## 4        Sink
```

```
## 5        Sink
## 6 Restaurant
## 7 Restaurant
## 8     Control
```

```
results %>%
  filter(campus != "Control") %>%
  ggplot(aes(x= source, y=total_cfu_mean, fill = campus)) +
  geom_bar(position = "stack", stat = "identity") +
  labs(title = "Source Types and Campus Colony Forming Units", x = "Source Type", y = "Total CFU Count/
```