field_demography.R

mcglinndj

2023-08-14

Author: Dan McGlinn Date: 2023-08-13

```
library(ggplot2)
```

load data

```
dat <- read.csv('../data/canbys_data - field.csv')
```

examine data

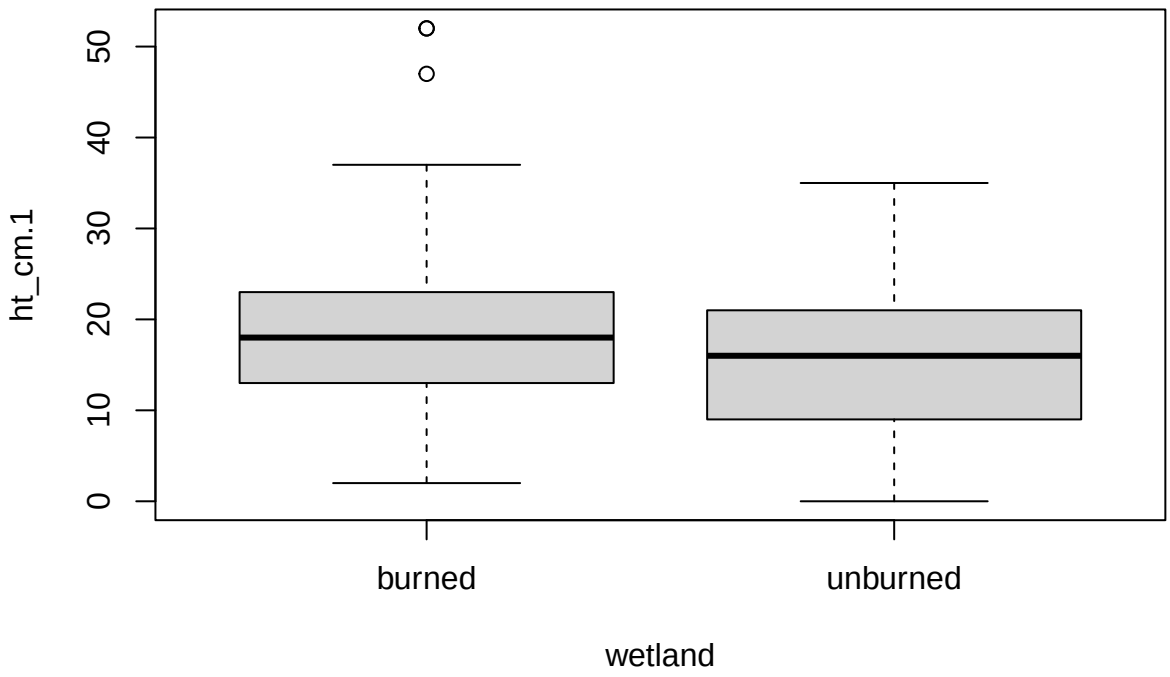
```
head(dat)
```

```
##    wetland loc plant_id status       date ht_cm stem_ct soil.moist
## 1 unburned   0    3-9-H      a 07/10/2023    13       1         NA
## 2 unburned   0    3-8-C      d 07/10/2023     0       0         NA
## 3 unburned   0    1-7-D      a 07/10/2023    15       1         NA
## 4 unburned   1    3-2-B      w 07/10/2023     8       2         NA
## 5 unburned   1    2-7-B      a 07/10/2023    14       1         NA
## 6 unburned   1    2-3-B      a 07/10/2023    32       1         NA
##   water.depth_cm notes     date.1 status.1 ht_cm.1 stem_ct.1 soil.moist.1
## 1              0       2023-07-18        a      14         2           NA
## 2              0       2023-07-18        d       0         0           NA
## 3              0       2023-07-18        a      16         2           NA
## 4              0       2023-07-18        a       9         3           NA
## 5              0       2023-07-18        a      15         2           NA
## 6              0       2023-07-18        a      34         4           NA
##   water.depth_cm.1 notes.1 Third.week..08.09.23. alive.wilt.dead ht_cm.2
## 1                0                            NA              NA      NA
## 2                0                            NA              NA      NA
## 3                0                            NA              NA      NA
## 4                0                            NA              NA      NA
## 5                0                            NA              NA      NA
## 6                0                            NA              NA      NA
##   X..of.stems soil.moist.2 water.depth_cm.2
## 1          NA           NA               NA
## 2          NA           NA               NA
## 3          NA           NA               NA
## 4          NA           NA               NA
## 5          NA           NA               NA
## 6          NA           NA               NA
```

```
dat$status.1 <- as.factor(dat$status.1)
with(dat, table(status.1, wetland))
```

```
##          wetland
## status.1 burned unburned
##                0        2
##        a      77       63
##        d       3        9
##        w       0        7
```

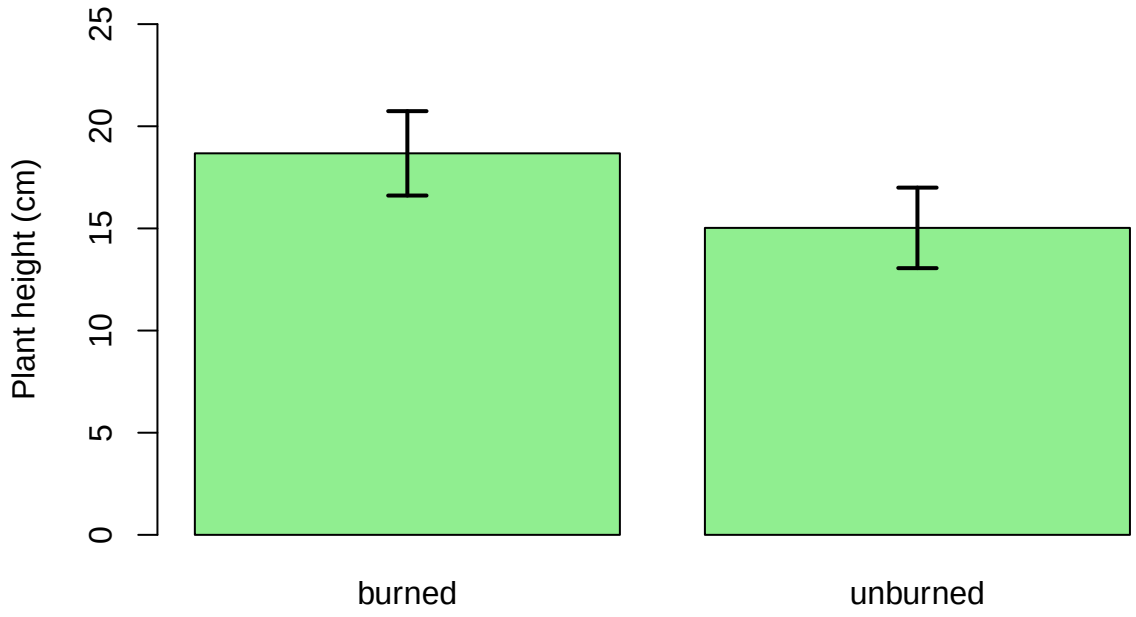
```
boxplot(ht_cm.1 ~ wetland, data = dat)
```



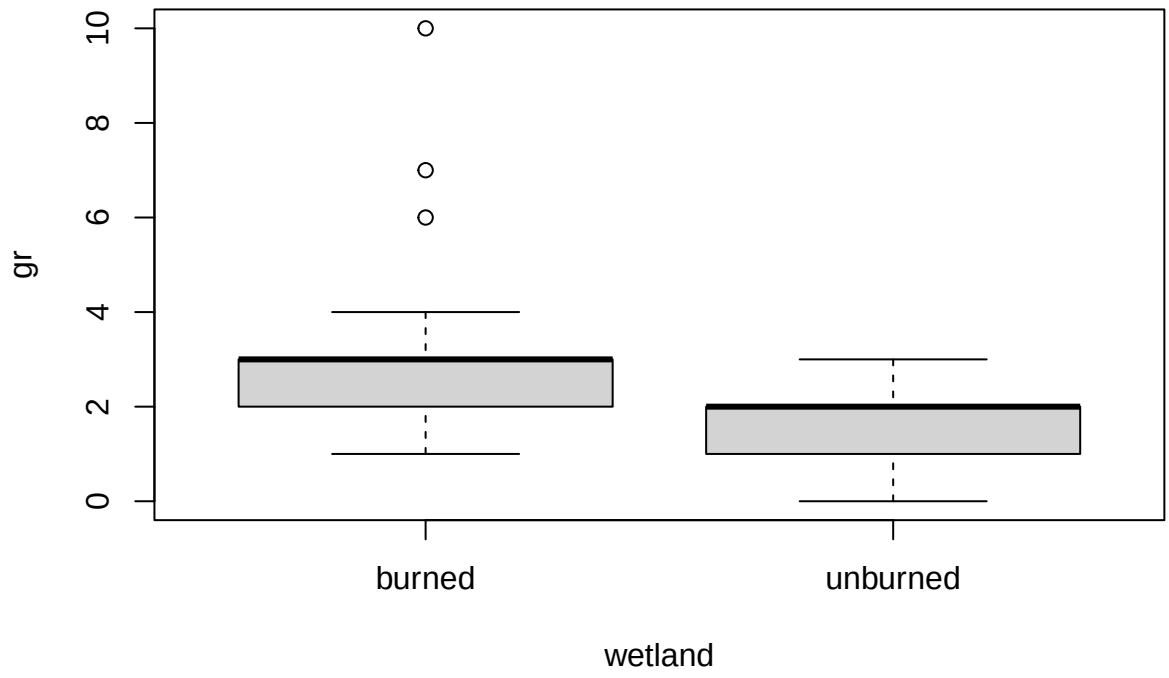

```
t.test(ht_cm.1 ~ wetland, data = dat)
```

```
##
##  Welch Two Sample t-test
##
## data:  ht_cm.1 by wetland
## t = 2.5049, df = 151.67, p-value = 0.01331
## alternative hypothesis: true difference in means between group burned and group unburned is not equal
## 95 percent confidence interval:
##  0.7709552 6.5277461
## sample estimates:
##   mean in group burned mean in group unburned
##               18.67532               15.02597
```

```r
ht_avg <- with(dat, tapply(ht_cm.1, wetland, mean, na.rm = TRUE))
ht_sd <- with(dat, tapply(ht_cm.1, wetland, sd, na.rm = TRUE))
ht_n <- with(dat, tapply(ht_cm.1, wetland, function(x) sum(!is.na(x))))
ht_se <- ht_sd / sqrt(ht_n)

#pdf('./figs/ht_vs_trt.pdf')
ht_plt <- barplot(ht_avg, width = 0.25, ylim = c(0, 25),
                  ylab = 'Plant height (cm)', col = 'lightgreen')
arrows(ht_plt, ht_avg - (ht_se * 1.96), y1 = ht_avg + (ht_se * 1.96),
       angle = 90, code = 3, length = 0.1, lwd=2)
```



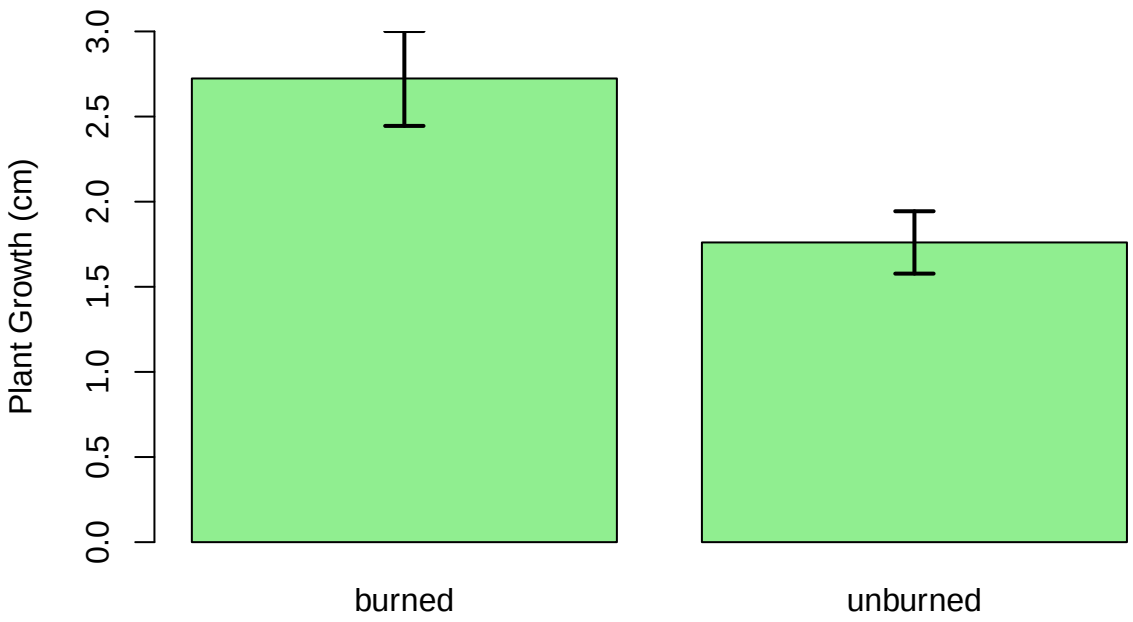

```r
#dev.off()

gr <- dat$ht_cm.1 - dat$ht_cm
# drop plant that lost ht - this appears to be due to damage
gr <- ifelse(gr < 0, NA, gr)

boxplot(gr ~ wetland, data = dat)
```



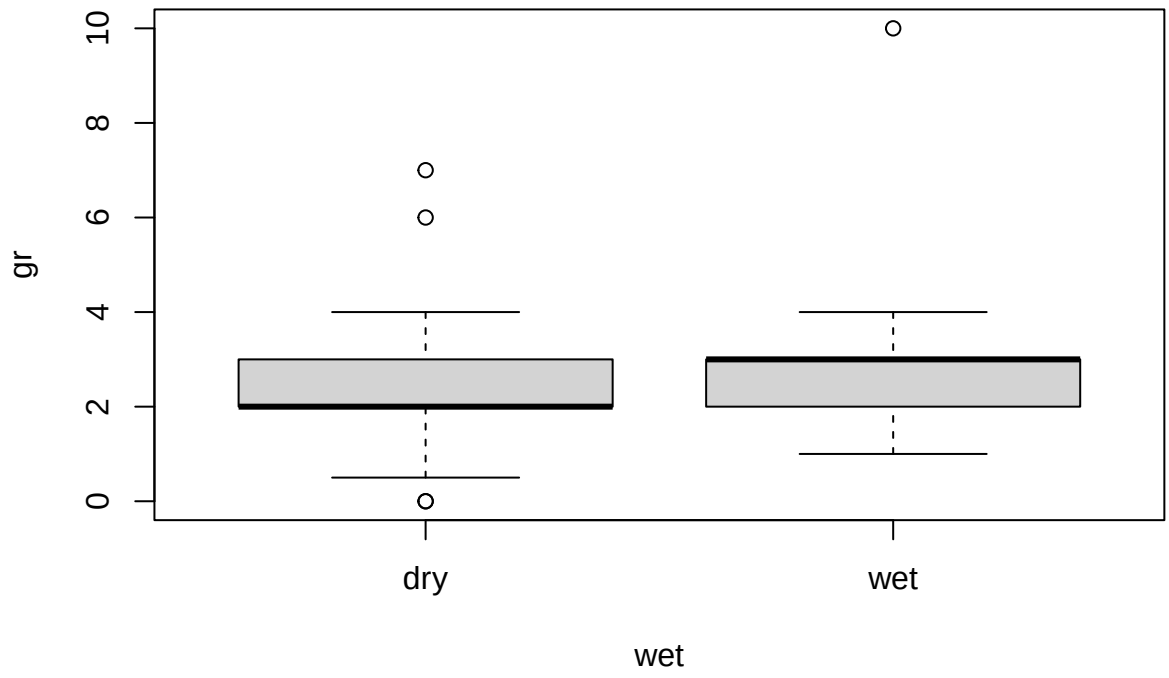

```
t.test(gr ~ wetland, data = dat)
```

```
##
##  Welch Two Sample t-test
##
## data:  gr by wetland
## t = 5.6607, df = 128.2, p-value = 9.407e-08
## alternative hypothesis: true difference in means between group burned and group unburned is not equal
## 95 percent confidence interval:
##  0.6264702 1.2997715
## sample estimates:
##   mean in group burned mean in group unburned
##               2.723684               1.760563
```

```
gr_avg <- with(dat, tapply(gr, wetland, mean, na.rm = TRUE))
gr_sd <- with(dat, tapply(gr, wetland, sd, na.rm = TRUE))
gr_n <- with(dat, tapply(gr, wetland, function(x) sum(!is.na(x))))
gr_se <- gr_sd / sqrt(gr_n)

#pdf('./figs/gr_vs_trt.pdf')
gr_plt <- barplot(gr_avg, width = 0.25, ylim = c(0, 3),
                  ylab = 'Plant Growth (cm)', col = 'lightgreen')
arrows(gr_plt, gr_avg - (gr_se * 1.96),
       y1 = gr_avg + (gr_se * 1.96),
       angle = 90, code = 3, length = 0.1, lwd = 2)
```




```
#dev.off()


wet <- ifelse(dat$water.depth_cm > 0, 'wet', 'dry')
boxplot(gr ~ wet)
```




```
gr_mod <- lm(gr ~ wet + wetland, data = dat)
summary(gr_mod)
```

```
##
## Call:
## lm(formula = gr ~ wet + wetland, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9607 -0.6663  0.0393  0.3337  7.0393
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.5774     0.1502  17.156  < 2e-16 ***
## wetwet            0.3833     0.2392   1.603    0.111
## wetlandunburned  -0.8223     0.1929  -4.263 3.63e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.04 on 144 degrees of freedom
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.1911, Adjusted R-squared:  0.1799
## F-statistic: 17.02 on 2 and 144 DF,  p-value: 2.326e-07
```

```
alive <- ifelse(dat$status.1 == 'd', 0, 1)
alive
```

```
##   [1] 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 0 1 0 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 0 1
##  [75] 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1
## [112] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [149] 1 1 1 1 1 1 1 1 1 1 1 1
```

test if probability of plant being alive is higher in a specific wetland

```
log_mod <- glm(alive ~ wetland, data = dat, family = binomial)
summary(log_mod)
```

```
##
## Call:
## glm(formula = alive ~ wetland, family = binomial, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5626   0.2765   0.2765   0.4854   0.4854
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)       3.2452     0.5885   5.514  3.5e-08 ***
## wetlandunburned  -1.1658     0.6865  -1.698   0.0895 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 85.398  on 160  degrees of freedom
## Residual deviance: 82.097  on 159  degrees of freedom
## AIC: 86.097
##
## Number of Fisher Scoring iterations: 6
```

you can see above not quite statistically significant

```
pseudo_r2 = function(glm_mod) {
  1 -  glm_mod$deviance / glm_mod$null.deviance
}
```

how much variation does the model explain

```
pseudo_r2(log_mod)
```

```
## [1] 0.0386552
```

not much at all only around 4% of the variation what is the probability of survival in each wetland?

```
predict(log_mod,
        newdata = data.frame(wetland = c('unburned', 'burned')),
        type = 'r')
```

```
##         1         2
## 0.8888889 0.9625000
```

```
# another way to calculate the above is as follows
logit2prob <- function(logit){
  odds <- exp(logit)
  prob <- odds / (1 + odds)
  return(prob)
}

logit2prob(coef(log_mod)[1])
```

```
## (Intercept)
##      0.9625
```

```
logit2prob(coef(log_mod)[1] + coef(log_mod)[2])
```

```
## (Intercept)
##   0.8888889
```

95% confidence interval (CI) for raw coefficients

```
confint(log_mod)
```

```
## Waiting for profiling to be done...
```

```
##                     2.5 %     97.5 %
## (Intercept)      2.261698 4.65123168
## wetlandunburned -2.700520 0.08756066
```

burned wetland 95% CI

```
logit2prob(2.2617)
```

```
## [1] 0.905655
```

```
logit2prob(4.6512)
```

```
## [1] 0.9905402
```

unburned wetland 95% CI

```
logit2prob(coef(log_mod)[1] - 2.7005)
```

```
## (Intercept)
##   0.6329035
```

```
logit2prob(coef(log_mod)[1] + 0.08756)
```

```
## (Intercept)
##   0.9655355
```