

REX: Gall Plots

Moriah Young, Kara Dobson

July 28, 2021

COLLABORATORS: Phoebe Zarnetske, Mark Hammond, Emily Parker, Kristin Wolford DATA INPUT: Cleaned gall csv from the shared Google drive
DATA OUTPUT: Plots visualizing gall data - Rmd is located in the "R/T7_warmx_plant_traits" folder in the rex-analysis Git repository PROJECT: REX

```
# Clear all existing data  
rm(list=ls())
```

```
# Load packages  
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3      v purrr  0.3.4  
## v tibble  3.0.6      v dplyr  1.0.4  
## v tidyr   1.1.2      v stringr 1.4.0  
## v readr   1.4.0      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

```
# Set working directory from .Renvirom  
dir <- Sys.getenv("ANALYSIS_DIR")  
list.files(dir)
```

```
## [1] "animal"           "GHG"  
## [3] "Irrigation"       "lookup_tables"  
## [5] "microbes"         "REX_Data_Management.gdoc"  
## [7] "REX_Data_Template.gsheel" "soil"  
## [9] "T7_ANPP"          "T7_plant_comp"  
## [11] "T7_plant_phenology" "T7_warmx_insect"  
## [13] "T7_warmx_plant_traits" "T7_warmx_VOC"  
## [15] "weather"
```

```
# Read in data  
galls <- read.csv(file.path(dir, "T7_warmx_plant_traits/L1/T7_warmx_galls_L1.csv"))
```

```
# making it so that plant_height, gall_diameter, and gall_height are on the same row for the same plant  
# so each unique plant will have just one row ompared to two in the raw data.
```

```

#galls2 <- galls %>%
#   select(-date) %>%
#   group_by(rep, footprint, treatment, plant_num, gall_present) %>%
#   summarise(across(everything(), na.omit)) %>%
#   distinct()

galls3 <- galls %>% dplyr::select(rep, footprint, treatment, plant_num, gall_present, plant_height) %>%
  na.omit() # creating a data frame with just the plant_height variable (not gall_height or gall_diameter)
# NAs

galls4 <- galls %>% dplyr::select(-plant_height) %>%
  na.omit() %>% # get rid of plant_height variable and NAs
  full_join(galls3) # joining the two data sets together

```

```
## Joining, by = c("rep", "footprint", "treatment", "plant_num", "gall_present")
```

```

# The cleaned dataset has data on the same plant that was collected on different dates. So some plants have multiple
# observations.

galls5 <- galls4 %>% dplyr::select(-date) # get rid of date - we don't need this and some have NAs for date

# create a dataframe that only contains galled plants
gall_only <- galls5[!(galls5$gall_present == "no_gall"),]

# Emily only wants to look at the effects of warming on goldenrod & galls so get rid of "drought" and "warmed_drought"
# gall_warmed <- galls2[!(galls2$treatment == "drought" | galls2$treatment == "warmed_drought"),]

```

Looking at Plant Height - Galls vs No Galls

```

# Take subplot average of plant height
plant_height <- galls5 %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(plant_height = mean(plant_height, na.rm = TRUE))

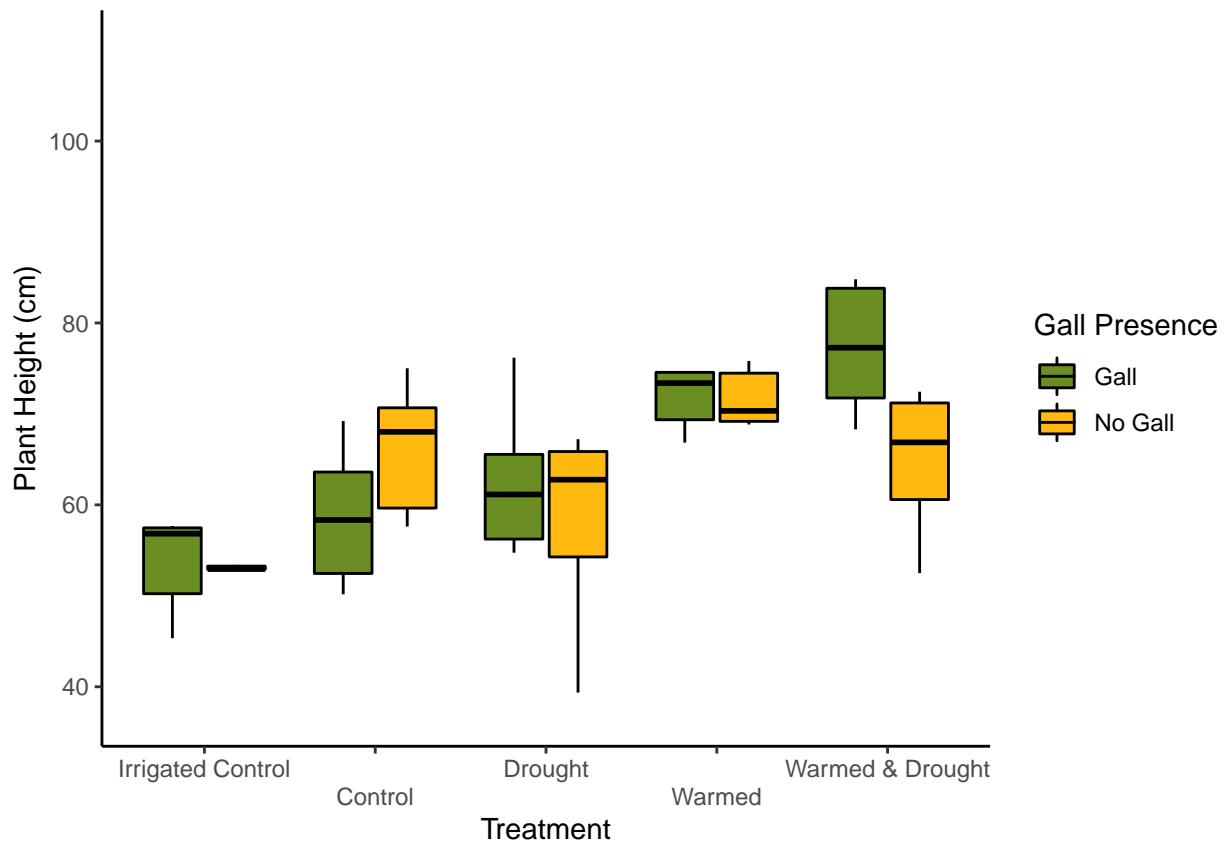
```

'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the 'group_by()' function.

```

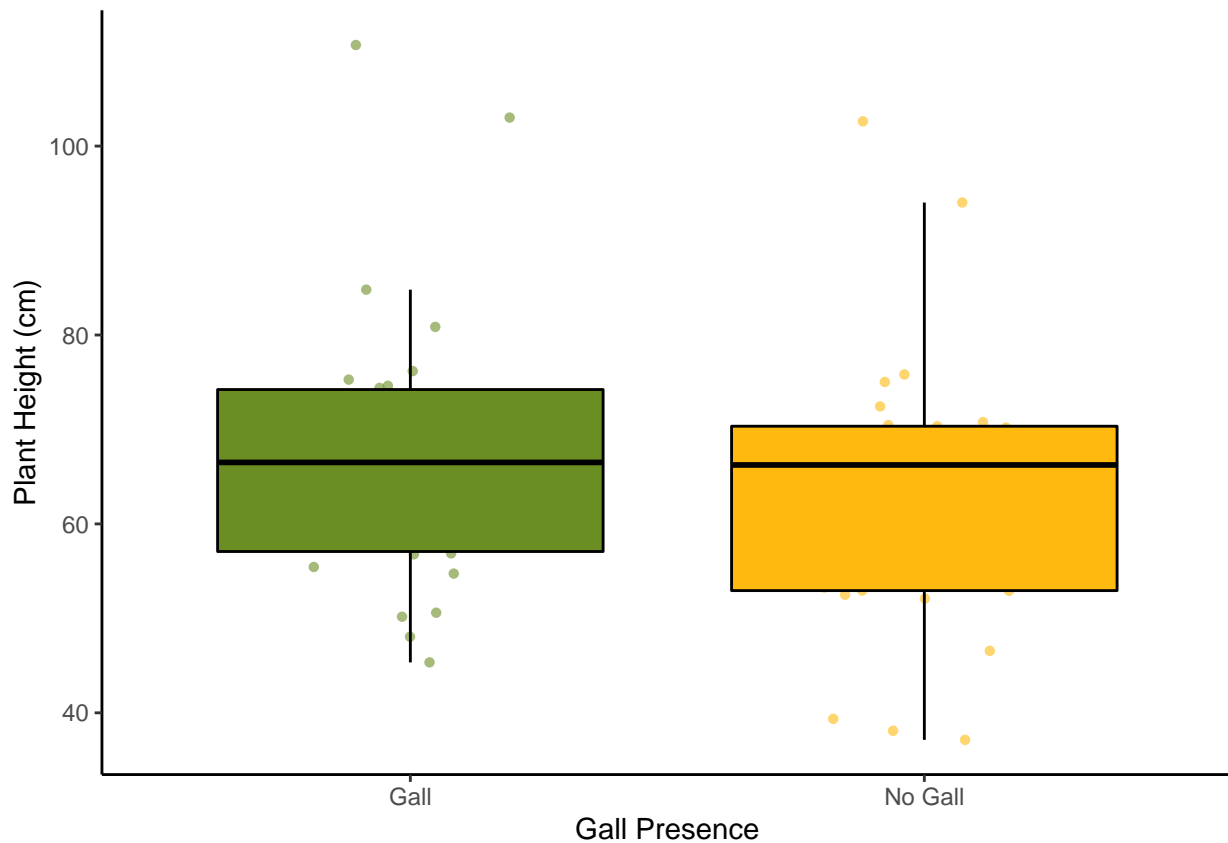
# Boxplot
ggplot(plant_height, aes(x = treatment, y = plant_height, fill = gall_present)) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Treatment", y = "Plant Height (cm)", fill = "Gall Presence") +
  scale_fill_manual(values = c("olivedrab", "darkgoldenrod1"), labels = c("Gall", "No Gall")) +
  scale_x_discrete(limits = c("irr_control", "ambient", "drought", "warmed", "warmed_drought"),
    labels=c("ambient" = "Control",
             "drought" = "Drought",
             "irr_control" = "Irrigated Control",
             "warmed" = "Warmed",
             "warmed_drought" = "Warmed & Drought"),
    guide = guide_axis(n.dodge=2)) +
  #geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present))
  theme_classic()

```



```
# Gall average plot
ggplot(plant_height, aes(x = gall_present, y = plant_height, fill = gall_present)) +
  geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present)) +
  scale_color_manual(values = c("gall" = "olivedrab", "no_gall" = "darkgoldenrod1")) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Gall Presence", y = "Plant Height (cm)") +
  scale_fill_manual(values = c("olivedrab", "darkgoldenrod1")) +
  scale_x_discrete(labels=c("gall" = "Gall",
                             "no_gall" = "No Gall")) +

  theme_classic() +
  theme(legend.position = "none")
```

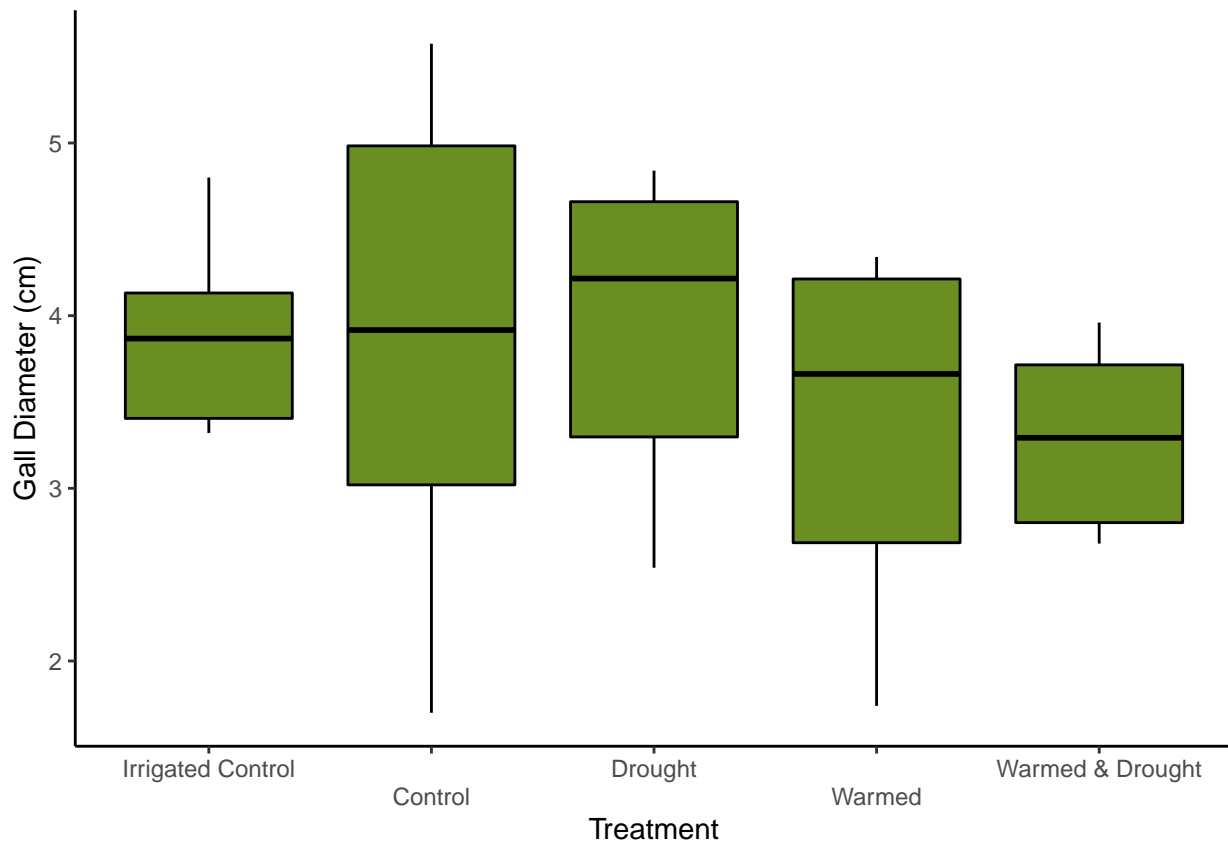


Looking at Gall Diameter

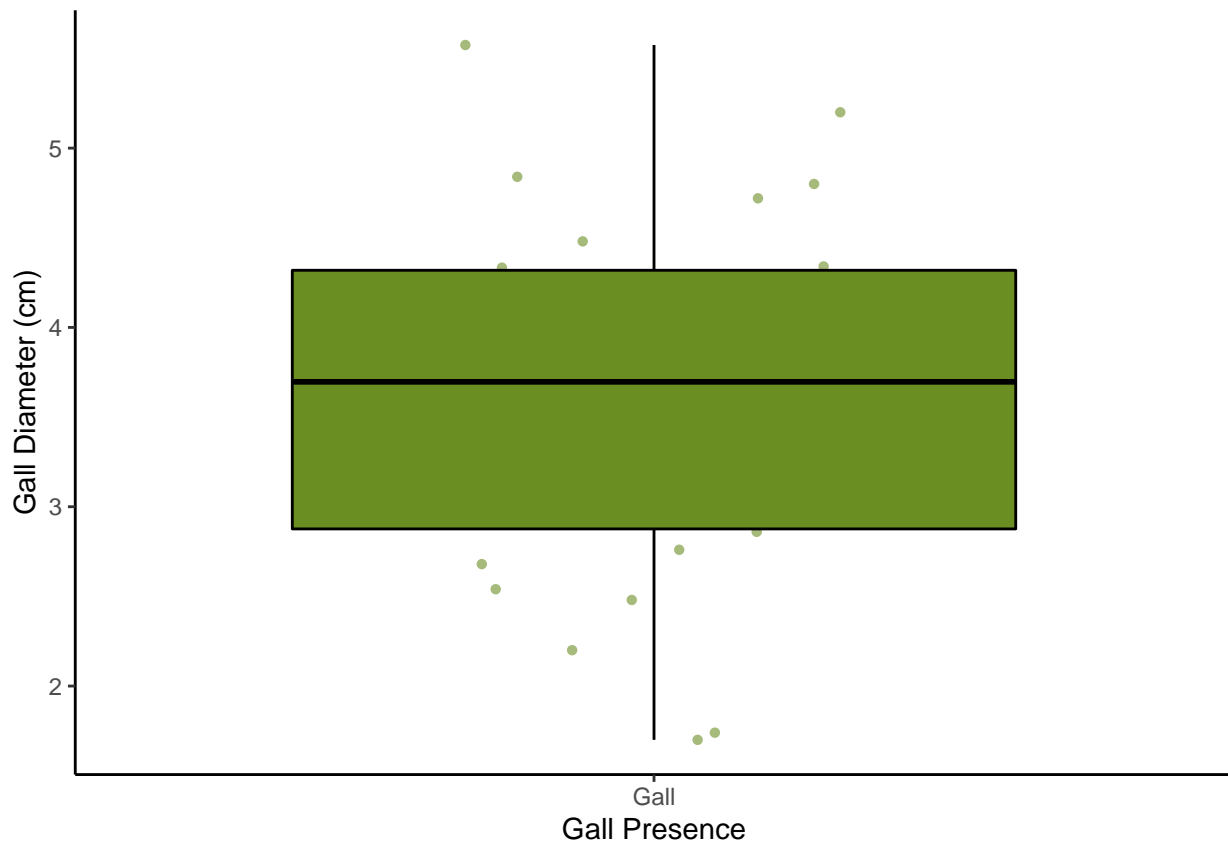
```
# Take subplot average of gall diameter
gall_diameter <- gall_only %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(gall_diameter = mean(gall_diameter, na.rm = TRUE))
```

'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the '.groups' argument.

```
# Boxplot
ggplot(gall_diameter, aes(x = treatment, y = gall_diameter, fill = gall_present)) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Treatment", y = "Gall Diameter (cm)", fill = "Gall Presence") +
  scale_fill_manual(values = c("olivedrab"), labels = c("Gall")) +
  scale_x_discrete(limits = c("irr_control", "ambient", "drought", "warmed", "warmed_drought"),
    labels=c("ambient" = "Control",
             "drought" = "Drought",
             "irr_control" = "Irrigated Control",
             "warmed" = "Warmed",
             "warmed_drought" = "Warmed & Drought"),
    guide = guide_axis(n.dodge=2)) +
  #geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present))
  theme_classic() +
  theme(legend.position = "none")
```



```
# Gall average plot
# don't need this one
ggplot(gall_diameter, aes(x = gall_present, y = gall_diameter, fill = gall_present)) +
  geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present)) +
  scale_color_manual(values = c("gall" = "olivedrab", "no_gall" = "darkgoldenrod1")) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Gall Presence", y = "Gall Diameter (cm)") +
  scale_fill_manual(values = c("olivedrab", "darkgoldenrod1")) +
  scale_x_discrete(labels=c("gall" = "Gall")) +
  theme_classic() +
  theme(legend.position = "none")
```

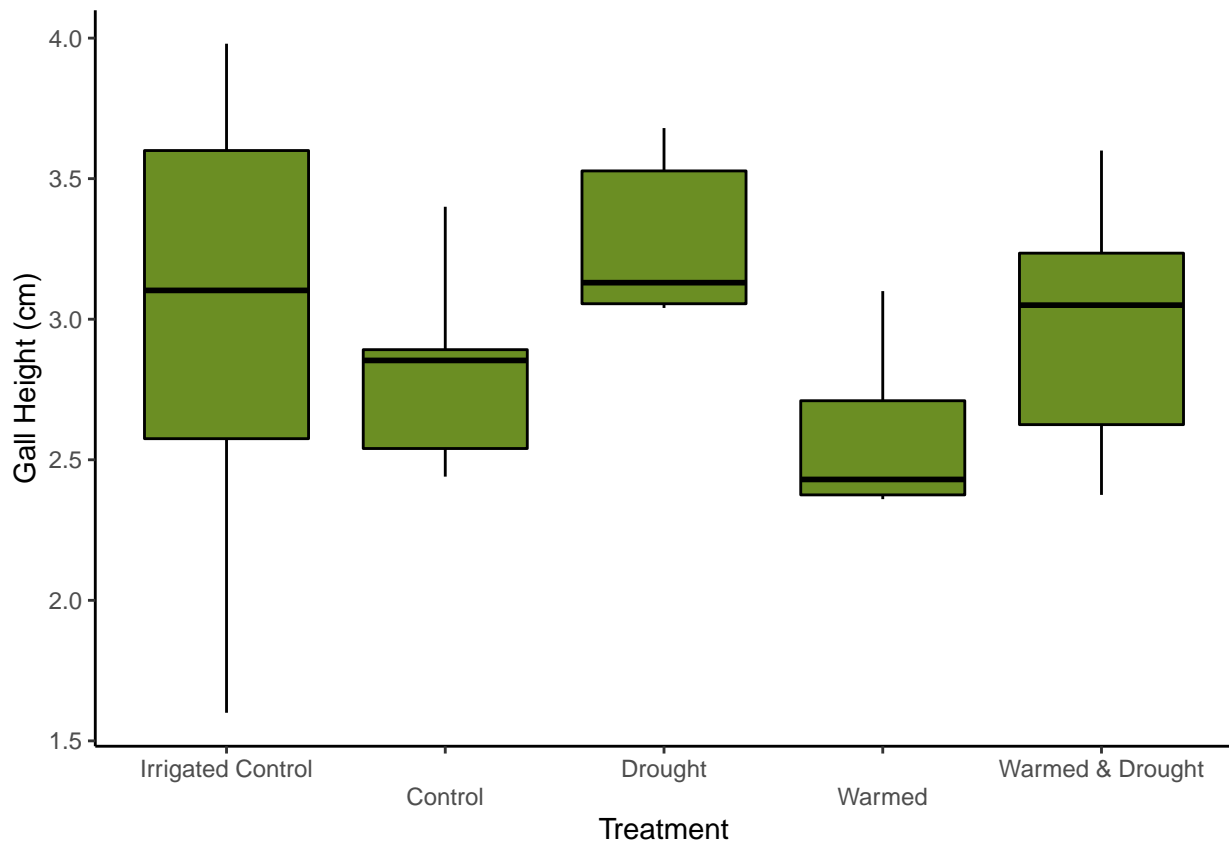


Looking at Gall Height

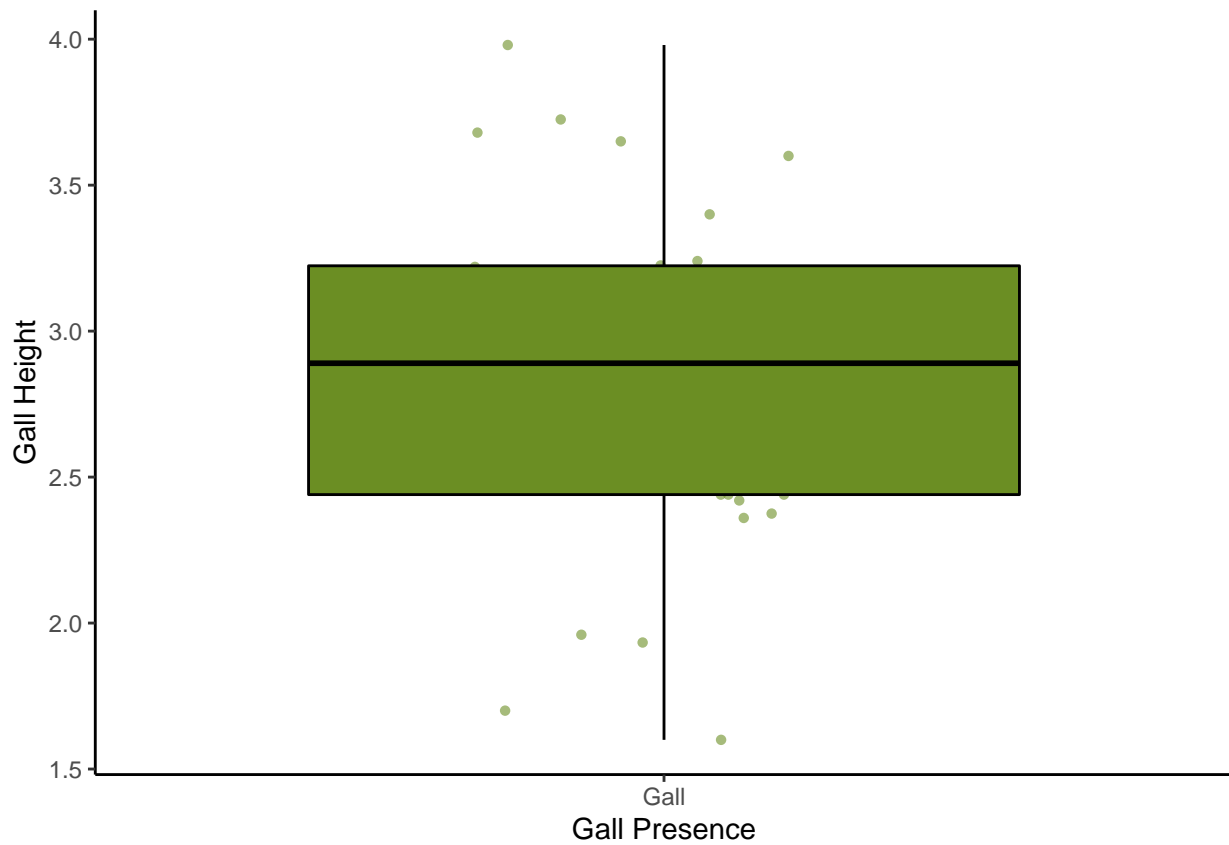
```
# Take subplot average of gall height
gall_height <- gall_only %>%
  group_by(rep, footprint, treatment, gall_present) %>%
  summarize(gall_height= mean(gall_height, na.rm = TRUE))
```

'summarise()' has grouped output by 'rep', 'footprint', 'treatment'. You can override using the '.groups' argument.

```
# Boxplot of gall height by treatment
ggplot(gall_height, aes(x = treatment, y = gall_height, fill = gall_present)) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Treatment", y = "Gall Height (cm)", fill = "Gall Presence") +
  scale_fill_manual(values = c("olivedrab"), labels = c("Gall")) +
  scale_x_discrete(limits = c("irr_control", "ambient", "drought", "warmed", "warmed_drought"),
    labels=c("ambient" = "Control",
             "drought" = "Drought",
             "irr_control" = "Irrigated Control",
             "warmed" = "Warmed",
             "warmed_drought" = "Warmed & Drought"),
    guide = guide_axis(n.dodge=2)) +
  #geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present))
  theme_classic() +
  theme(legend.position = "none")
```

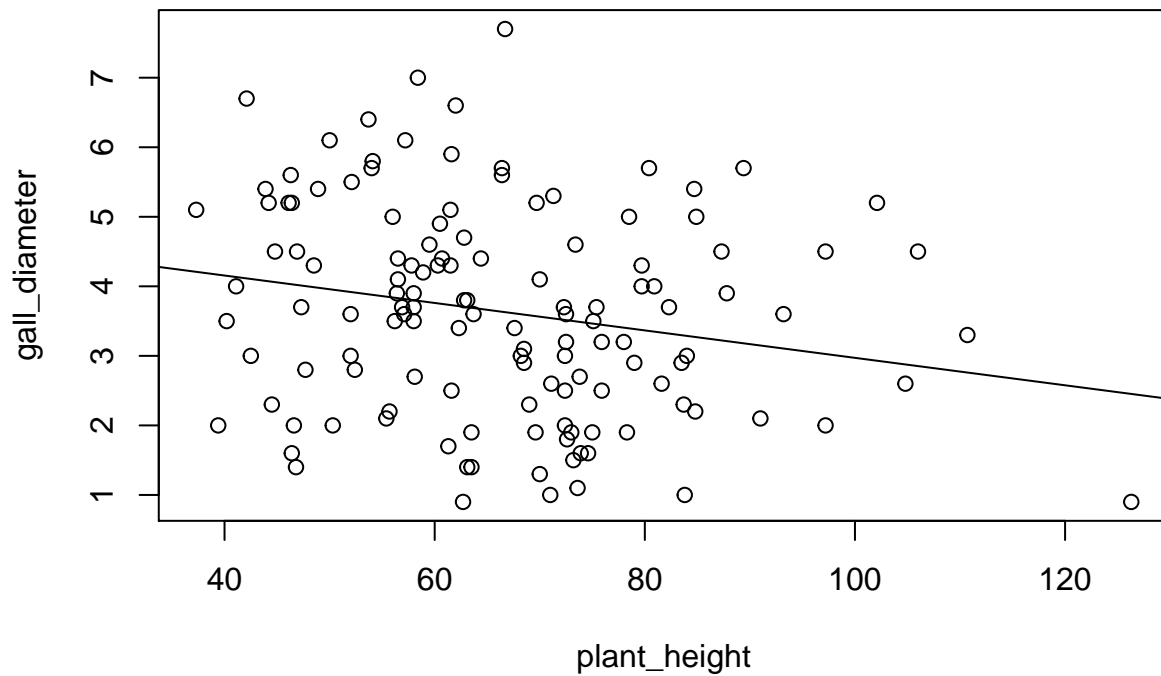


```
# Gall average plot
# don't need this one
ggplot(gall_height, aes(x = gall_present, y = gall_height, fill = gall_present)) +
  geom_jitter(shape=16, position=position_jitterdodge(), alpha = 0.6, aes(colour = gall_present)) +
  scale_color_manual(values = c("gall" = "olivedrab")) +
  geom_boxplot(color = "black", outlier.shape = NA) +
  labs(x = "Gall Presence", y = "Gall Height") +
  scale_fill_manual(values = c("olivedrab", "darkgoldenrod1")) +
  scale_x_discrete(labels=c("gall" = "Gall")) +
  theme_classic() +
  theme(legend.position = "none")
```



Height x Gall diameter or gall height

```
lm_gall <- lm(gall_diameter ~ plant_height, data = gall_only)
plot(gall_diameter ~ plant_height, data = gall_only) +
abline(lm_gall)
```




```
## integer(0)
```

```
summary(lm_gall)
```

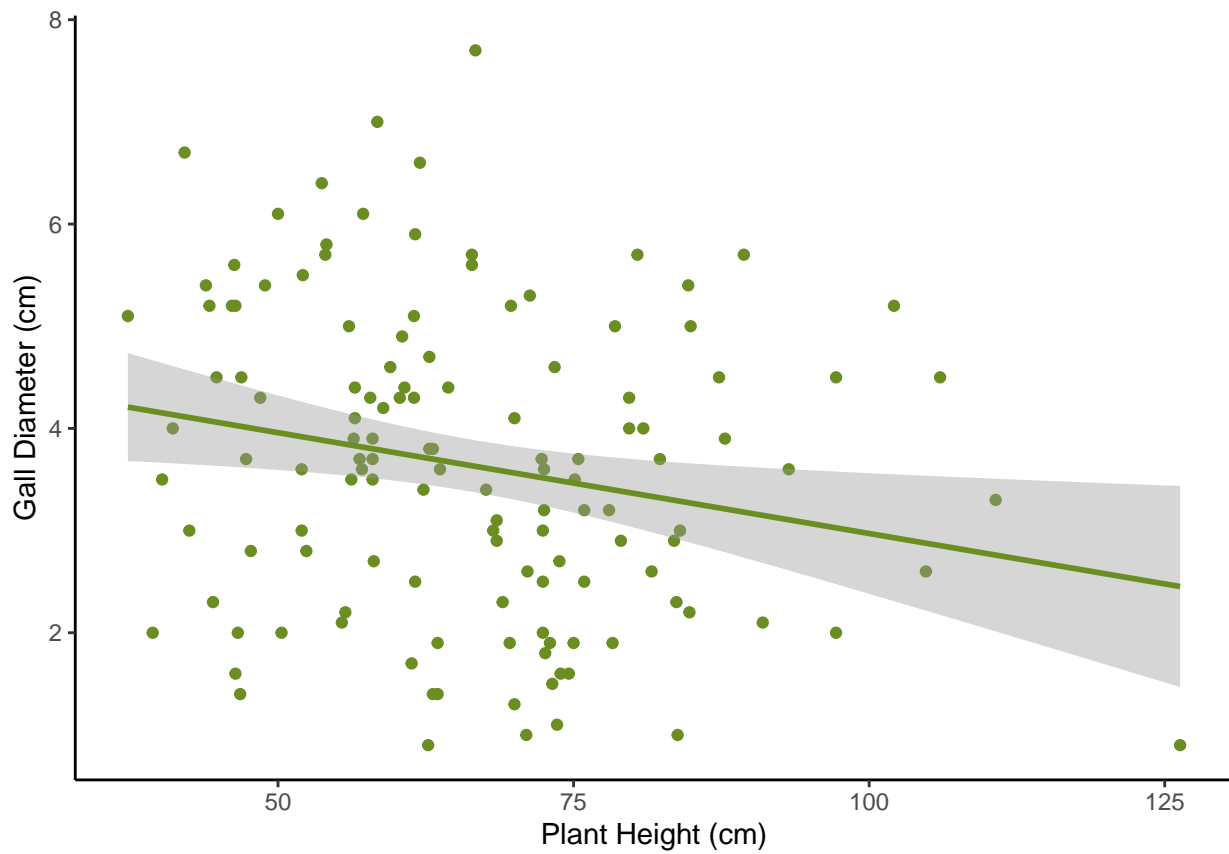
```
##
## Call:
## lm(formula = gall_diameter ~ plant_height, data = gall_only)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8076 -1.0916 -0.0941  1.1214  4.0714
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.944937    0.548298   9.019 2.37e-15 ***
## plant_height -0.019735    0.008022  -2.460  0.0152 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.469 on 128 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.04515,    Adjusted R-squared:  0.03769
## F-statistic: 6.052 on 1 and 128 DF,  p-value: 0.01522
```

```
# ggplot version
ggplot(gall_only, aes(plant_height, gall_diameter)) +
  geom_point(color='olivedrab') +
  geom_smooth(method='lm', color="olivedrab") +
  labs(x = "Plant Height (cm)", y = "Gall Diameter (cm)") +
  theme_classic()
```

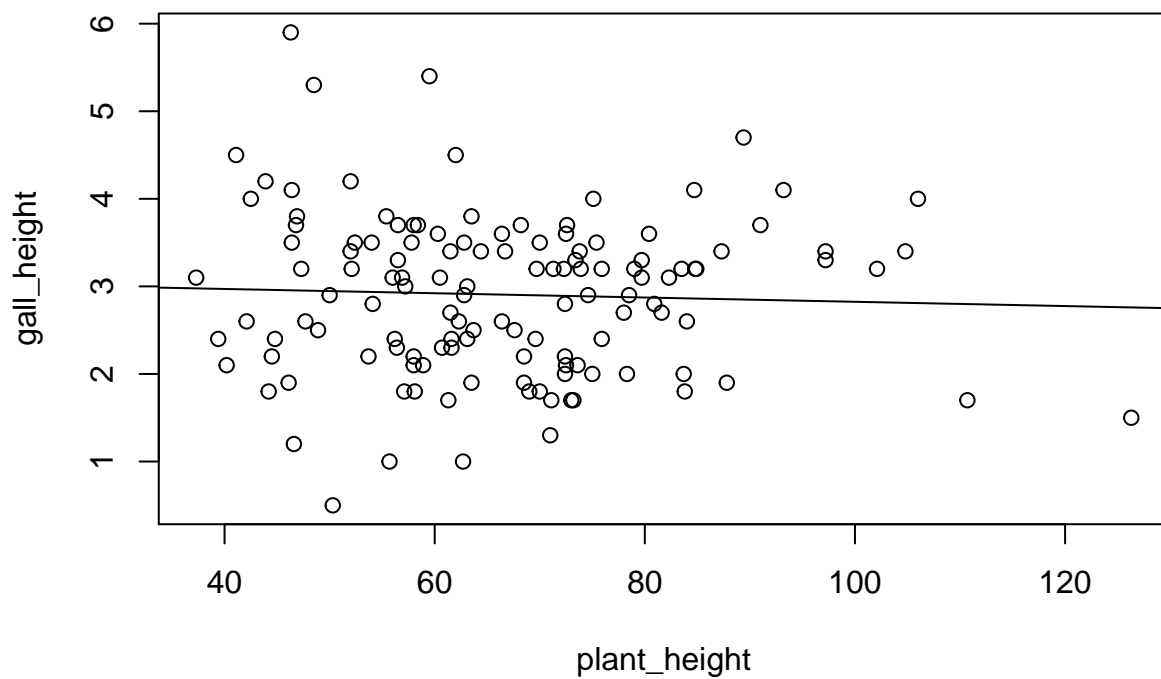
```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



```
lm_gall2 <- lm(gall_height ~ plant_height, data = gall_only)
plot(gall_height ~ plant_height, data = gall_only) +
  abline(lm_gall2)
```



```
## integer(0)
```

```
summary(lm_gall12)
```

```
##
## Call:
## lm(formula = gall_height ~ plant_height, data = gall_only)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4456 -0.7204  0.1042  0.5718  2.9447
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.068449   0.342059   8.971 3.1e-15 ***
## plant_height -0.002443   0.005005  -0.488  0.626
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9165 on 128 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.001858, Adjusted R-squared: -0.00594
## F-statistic: 0.2383 on 1 and 128 DF, p-value: 0.6263
```

```
# ggplot version
ggplot(gall_only, aes(plant_height, gall_height)) +
  geom_point(color='olivedrab') +
  geom_smooth(method='lm', color="olivedrab") +
  labs(x = "Plant Height (cm)", y = "Gall Height (cm)") +
  theme_classic()
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```

