Introduction

Background

Every year since 1995, citizen scientists gather for a weekend in September at the floodplain (the pools area) of the Mana Pools National Park and perform transects, with the goal of counting the number of animals of specific species in this pools area of the park. Counts are done in the mornings and the afternoons of the Saturday and Sunday.

Area

The Mana Pools National Park is located in Northern Zimbabwe against the Zambezi river. The park itself is 219600ha and is home to a large variety of game animals, including the big 5. A 4500ha area of the park is made up of a floodplain (the pools area) which retains water better than the rest of the park and generally has more lush vegetation.

Data available

The data available is present in 3 datasets The first contains the counts of animal species found as a result of the aforementioned transects. Species counted are: elephant, buffalo, eland, zebra, waterbuck, kudu, impala, nyala, bushbuck, wild dog, lion, leopard, hyena, jackal, warthog and baboon. Counts are separated according to day and time of day. The second and third datasets are for climatic and vegetation conditions for the 219600ha park and the 4500ha pools area, respectively. The frequency of this climatic and vegetation data is monthly.

study objectives

The primary object of this study is to display my proficiency with wrangling and presenting data using R, thus any correlations that are determined are for interests sake alone. That being said, I plan to study the correlations between game animals present and the amount of precipitation which fell during the previous wet season in the park and pools area, and the correlations between game animals present and various park and pools climatic and vegetation and factors present in the dry season leading up to each game count. Considering that the pools area retains water better and produces better quality vegetation than the rest of the park, and that animals will migrate locally to find better feeding and drinking resources, I expect that one will see that animal counts will vary according to yearly climatic and vegetation scores (increasing with better food and water resources at the pools and poorer food and water resources in the rest of the park and vice-versa).

Table importing and cleaning

Installing the tidyverse

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

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

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

Importing the game count df (gc), the park climatic data (pad) df, and the pools climatic data (roi) df.

```
gc <- read_csv('C:/Users/chris/Desktop/Coding/R_work/Projects/Game_counts/game_counts.csv')
```

```
## Rows: 80 Columns: 19

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (2): Day, Time
## dbl (17): Year, Elephant, Buffalo, Eland, Zebra, Waterbuck, Kudu, Impala, Ny...

##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
pad <- read_csv('C:/Users/chris/Desktop/Coding/R_work/Projects/Game_counts/PARK_additional_data.csv')
```

```
## Rows: 265 Columns: 18

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (17): aet, def, pdsi, pet, pr, ro, soil, srad, swe, tmmn, tmmx, vap, vp...
## date  (1): date

##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
roi <- read_csv('C:/Users/chris/Desktop/Coding/R_work/Projects/Game_counts/ROI_additional_data.csv')
```

```
## Rows: 265 Columns: 18

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (17): aet, def, pdsi, pet, pr, ro, soil, srad, swe, tmmn, tmmx, vap, vp...
## date  (1): date
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
gc
```

```
## # A tibble: 80 x 19
##     Year Day     Time  Elephant Buffalo Eland Zebra Waterbuck  Kudu Impala Nyala
##    <dbl> <chr>   <chr>    <dbl>   <dbl> <dbl> <dbl>     <dbl> <dbl>  <dbl> <dbl>
##  1  1995 Saturd~ am         155     381   179   460       234    66   3005     0
##  2  1995 Saturd~ pm         211     634   209   416       248   105   2345     0
##  3  1995 Sunday  am         159     465   136   426       238    62   2376     0
##  4  1995 Sunday  pm         173     304   171   581       242   140   2131     0
##  5  1996 Saturd~ am         210     376   173   273       350    54   4099     0
##  6  1996 Saturd~ pm         251     330   167   259       305    51   3443     0
##  7  1996 Sunday  am         163     431   254   240       404    23   3686     0
##  8  1996 Sunday  pm         327     443   233   231       336    47   2893     0
##  9  1997 Saturd~ am          83     516   260   226       393    55   3480     0
## 10  1997 Saturd~ pm         220     639   176   241       354    70   3455     0
## # ... with 70 more rows, and 8 more variables: Bushbuck <dbl>, Wild dog <dbl>,
## #   Lion <dbl>, Leopard <dbl>, Hyena <dbl>, Jackal <dbl>, Warthog <dbl>,
## #   Baboon <dbl>
```

Fortunately gc was small enough to take a quick scan through it. I noticed that in one year the Saturday and Sunday were recorded in shorthand ('Sat' and 'Sun'), so I converted them to Saturday and Sunday.

```
gc <- gc %>% mutate(Day = recode(Day, 'Sun'= 'Sunday', 'Sat'='Saturday'))
gc
```

```
## # A tibble: 80 x 19
##     Year Day     Time  Elephant Buffalo Eland Zebra Waterbuck  Kudu Impala Nyala
##    <dbl> <chr>   <chr>    <dbl>   <dbl> <dbl> <dbl>     <dbl> <dbl>  <dbl> <dbl>
##  1  1995 Saturd~ am         155     381   179   460       234    66   3005     0
##  2  1995 Saturd~ pm         211     634   209   416       248   105   2345     0
##  3  1995 Sunday  am         159     465   136   426       238    62   2376     0
##  4  1995 Sunday  pm         173     304   171   581       242   140   2131     0
##  5  1996 Saturd~ am         210     376   173   273       350    54   4099     0
##  6  1996 Saturd~ pm         251     330   167   259       305    51   3443     0
##  7  1996 Sunday  am         163     431   254   240       404    23   3686     0
##  8  1996 Sunday  pm         327     443   233   231       336    47   2893     0
##  9  1997 Saturd~ am          83     516   260   226       393    55   3480     0
## 10  1997 Saturd~ pm         220     639   176   241       354    70   3455     0
## # ... with 70 more rows, and 8 more variables: Bushbuck <dbl>, Wild dog <dbl>,
## #   Lion <dbl>, Leopard <dbl>, Hyena <dbl>, Jackal <dbl>, Warthog <dbl>,
## #   Baboon <dbl>
```

As I'm interested in the total counts per year, I could collapse the counts for Saturdays and Sundays, and am and pm into single counts representing each year, this df I called gc_yearly. I then removed all carnivores, warthog and baboon as I'm interested in animals that will locally migrate to and from the pools area, and carnivores, baboons and warthogs generally have well defined territories and so won't migrate locally. I then multiplied the number of animals of a species by the species LSU (large stock unit). This gives a more realistic relative pressure of each species on the environment (a factor to consider when adding species to make additional categories, such as browsers). I then added columns which summate the grazers, browsers and angostic herbivores, as these herbivore types use different vegetation types (grass, trees and shrubs, or both, respectively).

```
gc_yearly <- gc %>% group_by(Year) %>% summarize(Elephant=sum(Elephant, na.rm=TRUE), Buffalo=sum(Buffal

gc_yearly <- gc_yearly %>% select(Year, Elephant, Buffalo, Eland, Zebra, Waterbuck, Kudu, Nyala, Bushbu

gc_yearly <- gc_yearly %>% mutate(Elephant=Elephant*10, Buffalo=Buffalo*1.07, Eland=Eland*1.08, Zebra=Z

gc_yearly <- gc_yearly %>% mutate(Grazers = Buffalo + Zebra, Browsers = Kudu + Bushbuck, Agnostic = Ele

gc_yearly
```

```
## # A tibble: 20 x 13
##     Year Elephant Buffalo Eland Zebra Waterbuck  Kudu Nyala Bushbuck Impala
##    <dbl>    <dbl>   <dbl> <dbl> <dbl>     <dbl> <dbl> <dbl>    <dbl>  <dbl>
##  1  1995     6980   1909.  751. 1243.       481  201.   0       1.3   1873.
##  2  1996     9510   1691.  893.  662.       698.  94.5  0       0.91  2683.
##  3  1997     5620   2181.  989.  599.       692  111.   0       0.65  2555.
##  4  1998     5870   3797.  977.  896.       630  119.   0       0.26  3124.
##  5  1999     8220   2223. 1161   967.       717   68.6  0       0.39  2611.
##  6  2000     7700   5231. 1216. 1063.       863  197.   0       0.13  2690.
##  7  2001     6490   2985. 1112.  756.       834   57.2  0       0.26  2387.
##  8  2002     6390   1509. 1352.  707.       792  163.   0       0.52  3419.
##  9  2003     5140   2472. 1147.  591.       852   71.8  0       0.26  2396.
## 10  2004     5370   4014. 1064.  535.       863  145.   0.23    0     2481.
## 11  2005     7930   1559. 1264.  785.       934  270    0.23    0.78  3326.
## 12  2006     9660   2391. 1053   658.       642  126.   0       0.13  3082.
## 13  2007     9930   2400. 1356.  960.       648  172.   0.69    0.13  3933.
## 14  2008    10950   3122. 1232.  698.      1120  153.   0       0.26  3460.
## 15  2009    10870    681.  956.  323.       804.  94.5  0       0.78  3375.
## 16  2010    15300   2227.  927.  327.       840. 217.   0.23    0.52  3594.
## 17  2011    10690   1727.  807.  586.       718. 244.   0       0.13  4344.
## 18  2012     9900    745.  816.  500.       694. 216    2.07    3.38  4418.
## 19  2013     9470    622.  671.  521.       574  186.   1.38    0.39  4777.
## 20  2014    13050   1655.  566.  530.       522  205.   0       0.26  4422.
## # ... with 3 more variables: Grazers <dbl>, Browsers <dbl>, Agnostic <dbl>
```

I then looked at the pad df.

```
pad
```

```
## # A tibble: 265 x 18
##    date          aet     def  pdsi   pet       pr    ro  soil  srad   swe  tmmn
##    <date>      <dbl>   <dbl> <dbl> <dbl>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1 1994-01-01 1267.      0   -172. 1267. 230.     11.5  1180. 2017.     0  210.
##  2 1994-02-01 1162.      4.02 -223. 1166. 130.      6.57 1255. 2122.     0  201.
##  3 1994-03-01  706.   862.   -330. 1569.  27.2     1.32  807. 2700.     0  190.
##  4 1994-04-01  428.   963.   -399. 1390.  23.1     1      598. 2355.     0  173.
##  5 1994-05-01  119.  1087.   -458. 1206.   0        0     480. 2153.     0  139.
##  6 1994-06-01   77.6  903.   -471.  981.   0        0     402. 1903.     0  112.
##  7 1994-07-01   55.3 1054.   -468. 1109.   0        0     347. 1994.     0  117.
```

```
## 8 1994-08-01   41.5 1339.   -447. 1381.   0        0      306. 2176.     0  138.
## 9 1994-09-01   34.4 1811.   -420. 1846.   0.0519   0      273. 2604.     0  181.
## 10 1994-10-01  284.  1707.   -326. 1990.  27.1      1.19  247. 2454.     0  212.
## # ... with 255 more rows, and 7 more variables: tmmx <dbl>, vap <dbl>,
## #   vpd <dbl>, vs <dbl>, NDVI <dbl>, EVI <dbl>, water <dbl>
```

I renamed the columns to their full names and dropped the columns which are irrelevant to the study. I then used the date column to create a Year column and a Month column. As the game counts occurred in the beginning of September every year, I added two columns which gave an altered Year and Month value such that the Year_altered and Month_altered columns ran from September to August. This would make it easier to group climatic data for the periods before game counts were done.

```
pad <- pad %>% rename(Actual_evapotranspiration=aet,
                      Climate_water_deficit=def,
                      Palmer_drought_severity_index=pdsi,
                      Predicted_evapotranspiration=pet,
                      Precipitation=pr,
                      Runoff=ro,
                      Soil_moisture=soil,
                      Surface_radiation=srad,
                      Snow_water_Equivalent=swe,
                      Min_temp=tmmn,
                      Max_temp=tmmx,
                      Vapor_pressure=vap,
                      Vapor_pressure_deficit=vpd,
                      Wind_speed=vs,
                      Normalized_diff_veg_index=NDVI,
                      Enhanced_veg_index=EVI,
                      Water=water)
pad <- pad %>% select(c(1,4,6,8,16,17,18))
pad <- pad %>% mutate(Year=year(date),Month=month(date))
pad <- pad %>% select(c(2,3,4,5,6,7,8,9))
pad <- pad %>% mutate(Year_altered=lead(Year, n=4))
pad <- pad %>% mutate(Month_altered=lead(Month,n=4))
pad
```

```
## # A tibble: 265 x 10
##    Palmer_drought_severity_index Precipitation Soil_moisture Normalized_diff_ve~
##                            <dbl>         <dbl>         <dbl>               <dbl>
##  1                         -172.      230.           1180.                     0
##  2                         -223.      130.           1255.                     0
##  3                         -330.       27.2           807.                     0
##  4                         -399.       23.1           598.                     0
##  5                         -458.        0             480.                     0
##  6                         -471.        0             402.                     0
##  7                         -468.        0             347.                     0
##  8                         -447.        0             306.                     0
##  9                         -420.        0.0519        273.                     0
## 10                         -326.       27.1           247.                     0
## # ... with 255 more rows, and 6 more variables: Enhanced_veg_index <dbl>,
## #   Water <dbl>, Year <dbl>, Month <dbl>, Year_altered <dbl>,
## #   Month_altered <dbl>
```

I then applied to roi what I had applied to pad.

```r
roi <- roi %>% rename(Actual_evapotranspiration=aet,
                      Climate_water_deficit=def,
                      Palmer_drought_severity_index=pdsi,
                      Predicted_evapotranspiration=pet,
                      Precipitation=pr,
                      Runoff=ro,
                      Soil_moisture=soil,
                      Surface_radiation=srad,
                      Snow_water_Equivalent=swe,
                      Min_temp=tmmn,
                      Max_temp=tmmx,
                      Vapor_pressure=vap,
                      Vapor_pressure_deficit=vpd,
                      Wind_speed=vs,
                      Normalized_diff_veg_index=NDVI,
                      Enhanced_veg_index=EVI,
                      Water=water)
roi <- roi %>% select(c(1,4,6,8,16,17,18))
roi <- roi %>% mutate(Year=year(date),Month=month(date))
roi <- roi %>% select(c(2,3,4,5,6,7,8,9))
roi <- roi %>% mutate(Year_altered=lead(Year, n=4))
roi <- roi %>% mutate(Month_altered=lead(Month,n=4))
roi
```

```
## # A tibble: 265 x 10
##    Palmer_drought_severity_index Precipitation Soil_moisture Normalized_diff_ve~
##                            <dbl>         <dbl>         <dbl>               <dbl>
##  1                         -176.         225.          1147.                   0
##  2                         -220.         125.          1151.                   0
##  3                         -328.          20.1          799.                   0
##  4                         -397.          20.2          620.                   0
##  5                         -447.           0            509.                   0
##  6                         -457.           0            433.                   0
##  7                         -454.           0            377.                   0
##  8                         -437.           0            335.                   0
##  9                         -411.           0            301.                   0
## 10                         -322.          25.8          274.                   0
## # ... with 255 more rows, and 6 more variables: Enhanced_veg_index <dbl>,
## #   Water <dbl>, Year <dbl>, Month <dbl>, Year_altered <dbl>,
## #   Month_altered <dbl>
```

I then found the total precipitation for the prior wet season (Months_altered 2 to 9, annually) and the Palmer_drought_severity_index, Soil_moisture, Enhanced_veg_index and Water for the 3 months prior to the game count (Months_altered 10 to 12, annually). These new dfs were named pad_wet_precip and pad_dry_clim, respectively for the park data and roi_wet_precip and roi_dry_clim, respectively for the pools data. To these dfs I added a column to represent the area the df was referring to to (i.e. pools or park).

```r
pad_wet_precip <- pad %>% filter(Month_altered >= 2 & Month_altered <= 9) %>% group_by(Year_altered) %>%
pad_dry_clim <- pad %>% filter(Month_altered==10|Month_altered==11|Month_altered==12) %>% group_by(Year_

roi_wet_precip <- roi %>% filter(Month_altered >= 2 & Month_altered <= 9) %>% group_by(Year_altered) %>%
roi_dry_clim <- roi %>% filter(Month_altered==10|Month_altered==11|Month_altered==12) %>% group_by(Year_
```

```
pad_wet_precip
```

```
## # A tibble: 22 x 3
##     Year Precipitation Area
##    <dbl>         <dbl> <chr>
##  1  1994          411. park
##  2  1995          512. park
##  3  1996          683. park
##  4  1997         1036. park
##  5  1998          678. park
##  6  1999          950. park
##  7  2000          824. park
##  8  2001         1028. park
##  9  2002          651. park
## 10  2003          912. park
## # ... with 12 more rows
```

```
pad_dry_clim
```

```
## # A tibble: 22 x 6
##     Year Palmer_drought_severity_in~ Soil_moisture Enhanced_veg_ind~ Water Area
##    <dbl>                       <dbl>         <dbl>             <dbl> <dbl> <chr>
##  1  1994                      -1386.         1055.                0  3.02  park
##  2  1995                      -2338.          514.                0  2.02  park
##  3  1996                      -1322.         1361.                0  2.01  park
##  4  1997                       1424.         2886.                0  2.11  park
##  5  1998                       -911.         1261.                0  2.02  park
##  6  1999                        703.         2167.                0  1.01  park
##  7  2000                        879.         2461.             6662. 0.846 park
##  8  2001                       1489.         2556.             7333. 2.02  park
##  9  2002                      -1106.         1055.             6086. 1.01  park
## 10  2003                        703.         2345.             7121. 1.01  park
## # ... with 12 more rows
```

```
roi_wet_precip
```

```
## # A tibble: 22 x 3
##     Year Precipitation Area
##    <dbl>         <dbl> <chr>
##  1  1994          390. pools
##  2  1995          515. pools
##  3  1996          661. pools
##  4  1997          976. pools
##  5  1998          637. pools
##  6  1999          894. pools
##  7  2000          781. pools
##  8  2001          990. pools
##  9  2002          620. pools
## 10  2003          899. pools
## # ... with 12 more rows
```

```r
roi_dry_clim
```

```
## # A tibble: 22 x 6
##     Year Palmer_drought_severity_in~ Soil_moisture Enhanced_veg_ind~ Water Area
##    <dbl>                       <dbl>         <dbl>             <dbl> <dbl> <chr>
##  1  1994                      -1348.         1146.                 0  3.33 pools
##  2  1995                      -2185.          485.                 0  2.23 pools
##  3  1996                      -1323.         1331.                 0  2.20 pools
##  4  1997                       1472.         3435.                 0  2.21 pools
##  5  1998                       -937.         1329.                 0  2.21 pools
##  6  1999                        895.         2761.                 0  1.11 pools
##  7  2000                        633.         2285.              5011.  1.02 pools
##  8  2001                       1632.         3261.              5826.  2.27 pools
##  9  2002                      -1104.         1111.              4941.  1.13 pools
## 10  2003                        812.         2561.              5715.  1.15 pools
## # ... with 12 more rows
```

I then merged pad__wet__precip and roi__wet__precip, and pad__dry__clim and roi__dry__clim. These dfs were called wet__precip and dry__clim, respectively.

```r
wet_precip <- pad_wet_precip %>% full_join(roi_wet_precip) %>% arrange(Year)
```

```
## Joining, by = c("Year", "Precipitation", "Area")
```

```r
wet_precip
```

```
## # A tibble: 44 x 3
##     Year Precipitation Area
##    <dbl>         <dbl> <chr>
##  1  1994          411. park
##  2  1994          390. pools
##  3  1995          512. park
##  4  1995          515. pools
##  5  1996          683. park
##  6  1996          661. pools
##  7  1997         1036. park
##  8  1997          976. pools
##  9  1998          678. park
## 10  1998          637. pools
## # ... with 34 more rows
```

```r
dry_clim <- pad_dry_clim %>% full_join(roi_dry_clim) %>% arrange(Year)
```

```
## Joining, by = c("Year", "Palmer_drought_severity_index", "Soil_moisture", "Enhanced_veg_index", "Wate
```

```r
dry_clim
```

```
## # A tibble: 44 x 6
##     Year Palmer_drought_severity_in~ Soil_moisture Enhanced_veg_ind~ Water Area
##    <dbl>                       <dbl>         <dbl>             <dbl> <dbl> <chr>
```

```
## 1 1994 -1386. 1055. 0 3.02 park
## 2 1994 -1348. 1146. 0 3.33 pools
## 3 1995 -2338. 514. 0 2.02 park
## 4 1995 -2185. 485. 0 2.23 pools
## 5 1996 -1322. 1361. 0 2.01 park
## 6 1996 -1323. 1331. 0 2.20 pools
## 7 1997 1424. 2886. 0 2.11 park
## 8 1997 1472. 3435. 0 2.21 pools
## 9 1998 -911. 1261. 0 2.02 park
## 10 1998 -937. 1329. 0 2.21 pools
## # ... with 34 more rows
```

I then merged gc__yearly and wet__precip to make gc__precip and gc__yearly and dry__clim to make gc__clim

```
gc_precip <- gc_yearly %>% full_join(wet_precip)
```

```
## Joining, by = "Year"
```

```
gc_clim <- gc_yearly %>% full_join(dry_clim)
```

```
## Joining, by = "Year"
```

```
gc_precip
```

```
## # A tibble: 44 x 15
##     Year Elephant Buffalo Eland Zebra Waterbuck  Kudu Nyala Bushbuck Impala
##    <dbl>    <dbl>   <dbl> <dbl> <dbl>     <dbl> <dbl> <dbl>    <dbl>  <dbl>
##  1  1995     6980   1909.  751. 1243.       481  201.      0     1.3   1873.
##  2  1995     6980   1909.  751. 1243.       481  201.      0     1.3   1873.
##  3  1996     9510   1691.  893.  662.      698.  94.5     0     0.91  2683.
##  4  1996     9510   1691.  893.  662.      698.  94.5     0     0.91  2683.
##  5  1997     5620   2181.  989.  599.       692  111.      0     0.65  2555.
##  6  1997     5620   2181.  989.  599.       692  111.      0     0.65  2555.
##  7  1998     5870   3797.  977.  896.       630  119.      0     0.26  3124.
##  8  1998     5870   3797.  977.  896.       630  119.      0     0.26  3124.
##  9  1999     8220   2223. 1161   967.       717   68.6     0     0.39  2611.
## 10  1999     8220   2223. 1161   967.       717   68.6     0     0.39  2611.
## # ... with 34 more rows, and 5 more variables: Grazers <dbl>, Browsers <dbl>,
## #   Agnostic <dbl>, Precipitation <dbl>, Area <chr>
```

```
gc_clim
```

```
## # A tibble: 44 x 18
##     Year Elephant Buffalo Eland Zebra Waterbuck  Kudu Nyala Bushbuck Impala
##    <dbl>    <dbl>   <dbl> <dbl> <dbl>     <dbl> <dbl> <dbl>    <dbl>  <dbl>
##  1  1995     6980   1909.  751. 1243.       481  201.      0     1.3   1873.
##  2  1995     6980   1909.  751. 1243.       481  201.      0     1.3   1873.
##  3  1996     9510   1691.  893.  662.      698.  94.5     0     0.91  2683.
##  4  1996     9510   1691.  893.  662.      698.  94.5     0     0.91  2683.
##  5  1997     5620   2181.  989.  599.       692  111.      0     0.65  2555.
##  6  1997     5620   2181.  989.  599.       692  111.      0     0.65  2555.
```

```
##  7  1998     5870   3797.  977.  896.      630  119.      0     0.26  3124.
##  8  1998     5870   3797.  977.  896.      630  119.      0     0.26  3124.
##  9  1999     8220   2223. 1161   967.      717   68.6     0     0.39  2611.
## 10  1999     8220   2223. 1161   967.      717   68.6     0     0.39  2611.
## # ... with 34 more rows, and 8 more variables: Grazers <dbl>, Browsers <dbl>,
## #   Agnostic <dbl>, Palmer_drought_severity_index <dbl>, Soil_moisture <dbl>,
## #   Enhanced_veg_index <dbl>, Water <dbl>, Area <chr>
```

Incase the y axis scales needed for graphs resulted in difficulty seeing the trends in some species/groups while encompassing other species/groups. I also created a table with species/group column normalized as a percentage of the maximum value for the species/group seeing in a single year.

```
gc_max <- gc_yearly %>% summarize(across(2:13, max)) %>% mutate(Year=100) %>% relocate(Year)
gc_max <- gc_max %>% slice(rep(1:n(), each=20))
gc_yearly_norm <- gc_yearly/gc_max*100

gc_norm_precip <- gc_yearly_norm %>% full_join(wet_precip)
```

```
## Joining, by = "Year"
```

```
gc_norm_clim <- gc_yearly_norm %>% full_join(dry_clim)
```

```
## Joining, by = "Year"
```

Data analysis

Correlation between precipitation and animal counts

I graphed park and pools precipitations against LSU values for each species, and park and pools precipitations against LSU values for each herbivore type.

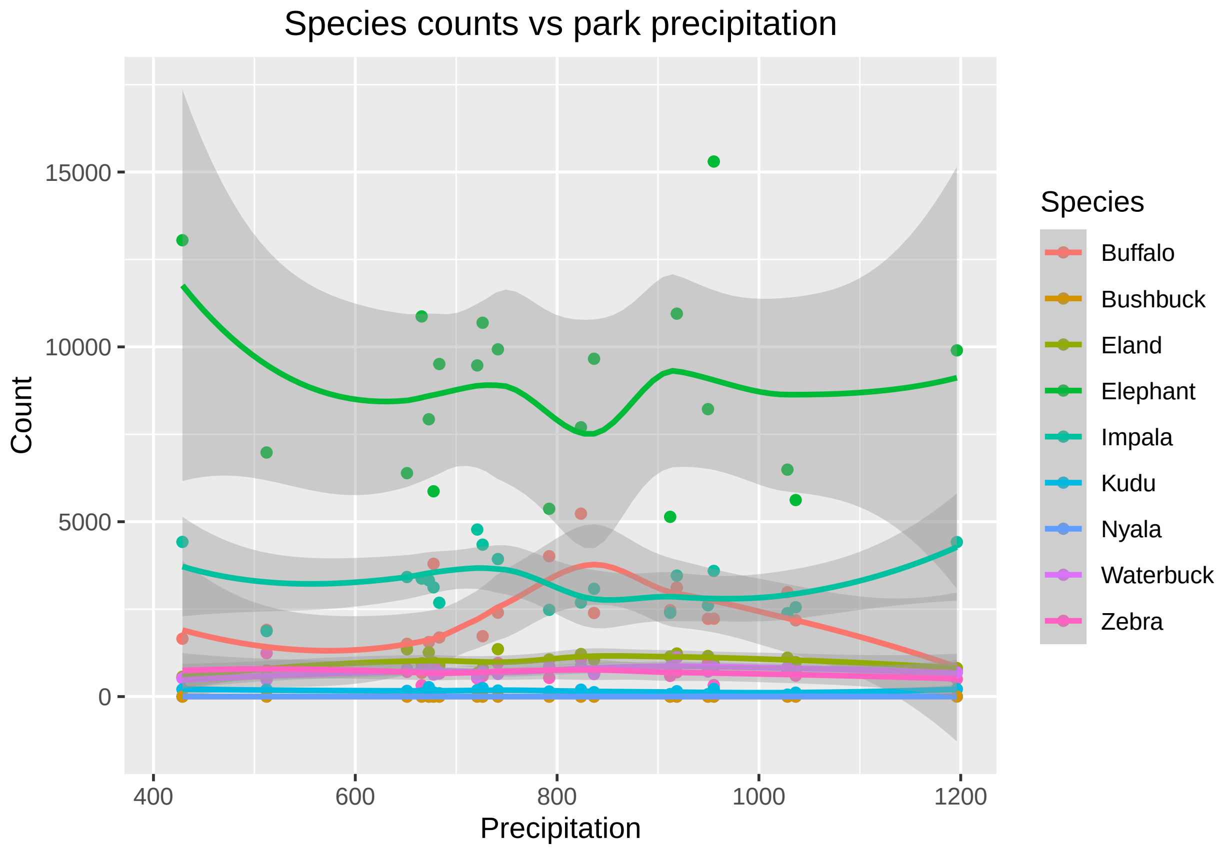
```
gc_precip_for_spp_plot <- gc_precip %>% gather(c(2,3,4,5,6,7,8,9,10), key='Species', value='Count')

ggplot(data=filter(gc_precip_for_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Precipitation, y=Co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 18 rows containing non-finite values (stat_smooth).

## Warning: Removed 18 rows containing missing values (geom_point).
```

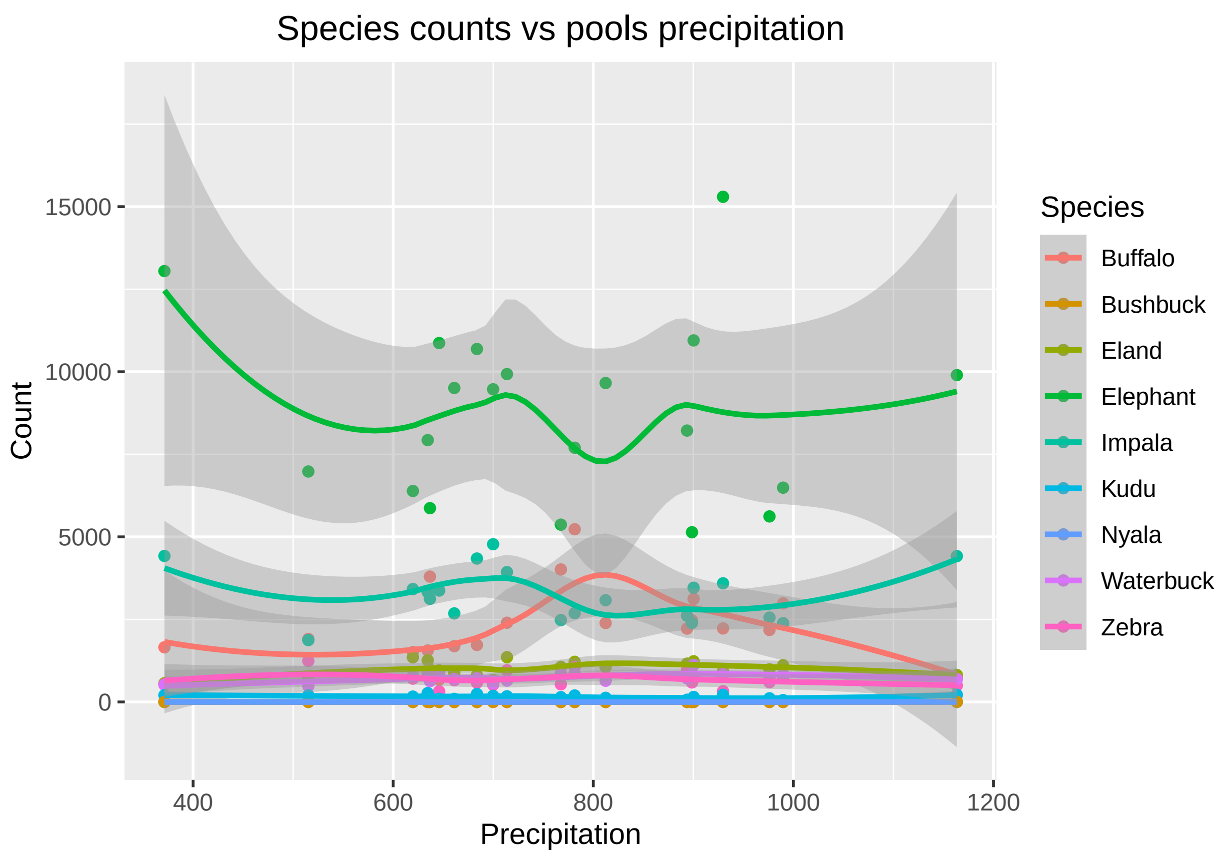



```
ggplot(data=filter(gc_precip_for_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Precipitation, y=(
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 18 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 18 rows containing missing values (geom_point).
```



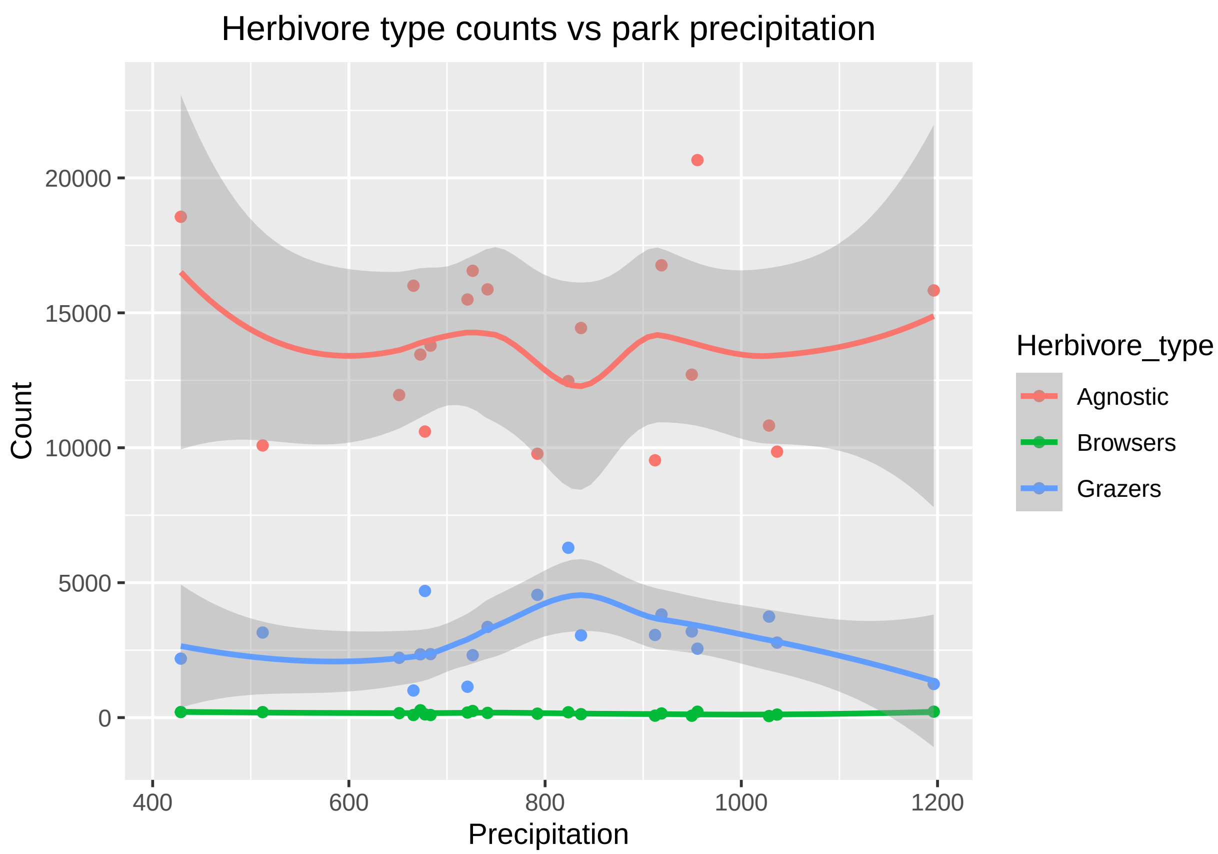

```
gc_precip_for_type_plot <- gc_precip %>% gather(c(11,12,13), key='Herbivore_type', value='Count')

ggplot(data=filter(gc_precip_for_type_plot, Area=="park")) + geom_point(mapping=aes(x=Precipitation, y=C
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 6 rows containing non-finite values (stat_smooth).

## Warning: Removed 6 rows containing missing values (geom_point).
```

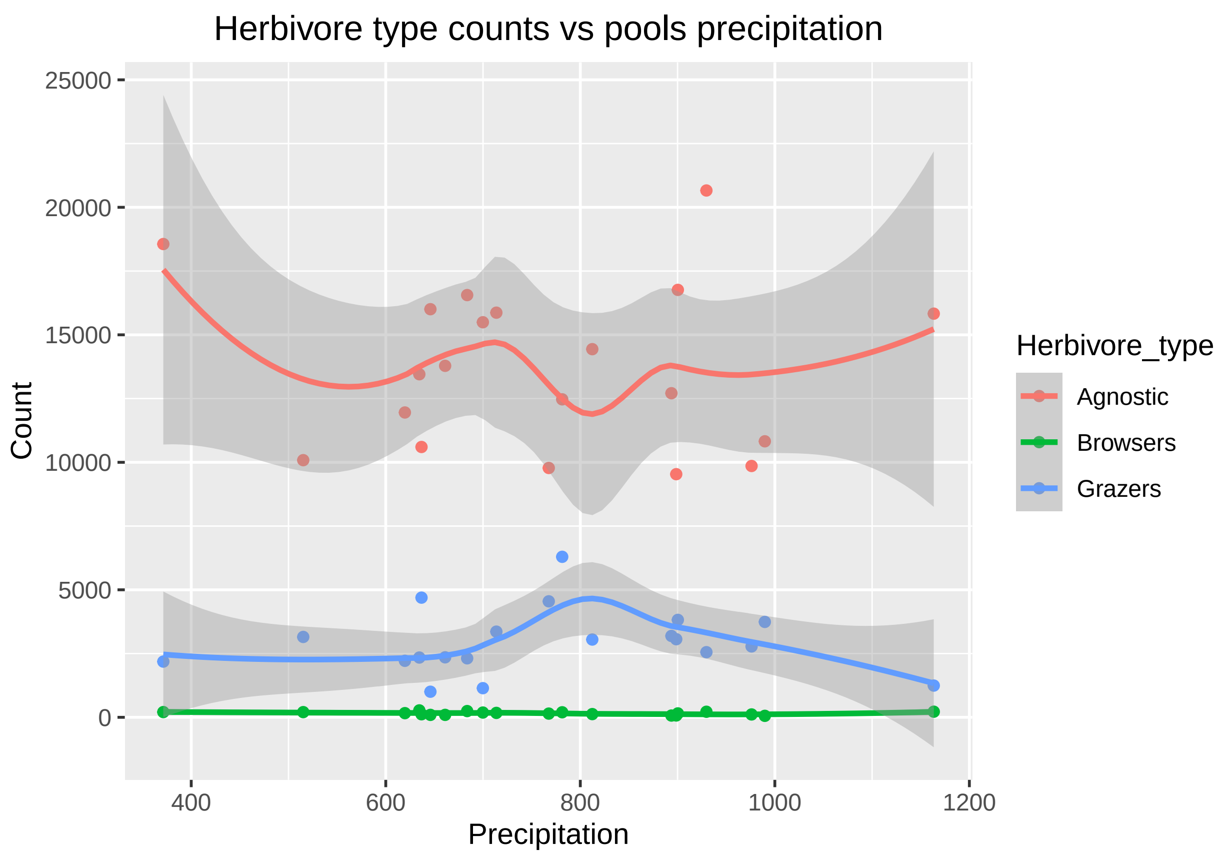



```
ggplot(data=filter(gc_precip_for_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Precipitation, y=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



I then graphed the park and pools precipitations against the normalized LSU values for each species, and park and pools precipitations against the normalized LSU values for each herbivore type.

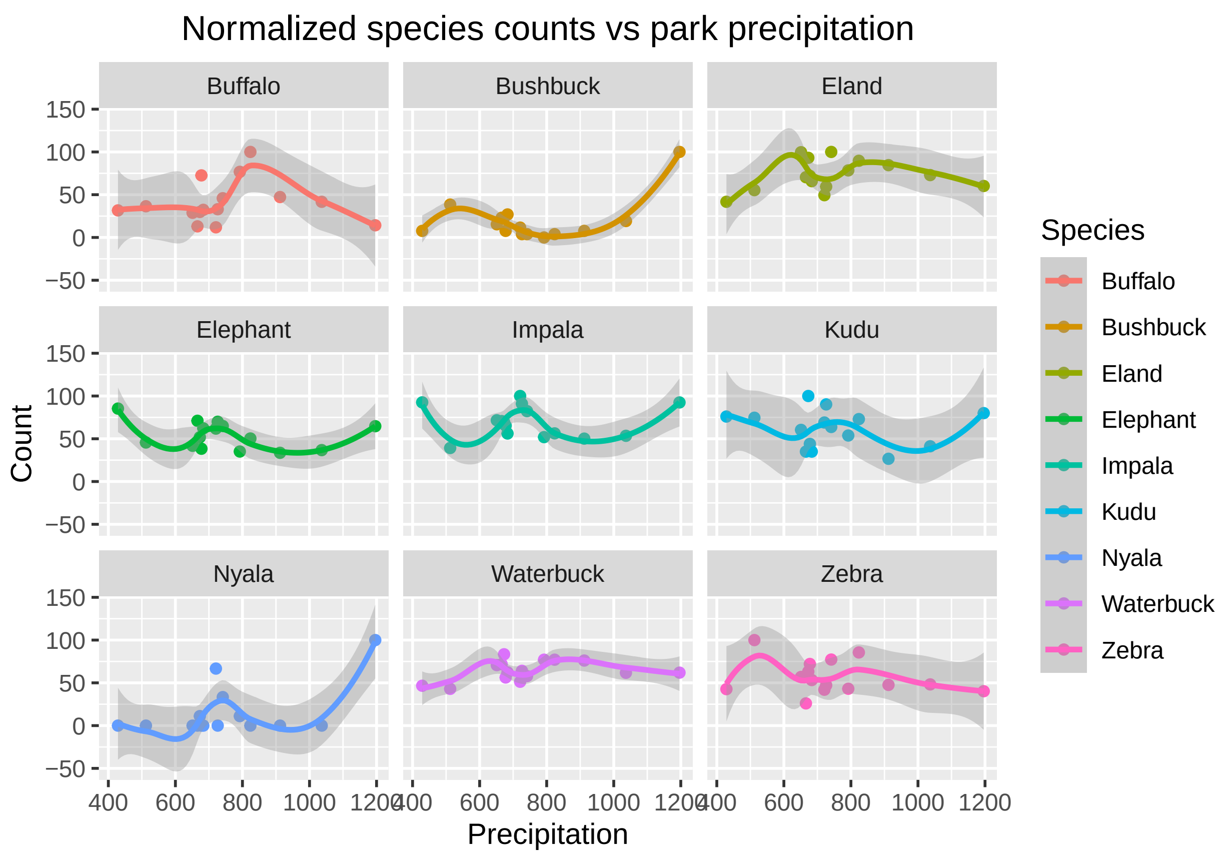
```
gc_precip_norm_spp_plot <- gc_norm_precip %>% gather(c(2,3,4,5,6,7,8,9,10), key='Species', value='Count

ggplot(data=filter(gc_precip_norm_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Precipitation, y=(
```

```
## ‘geom_smooth()‘ using method = ’loess’ and formula ’y ~ x’

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

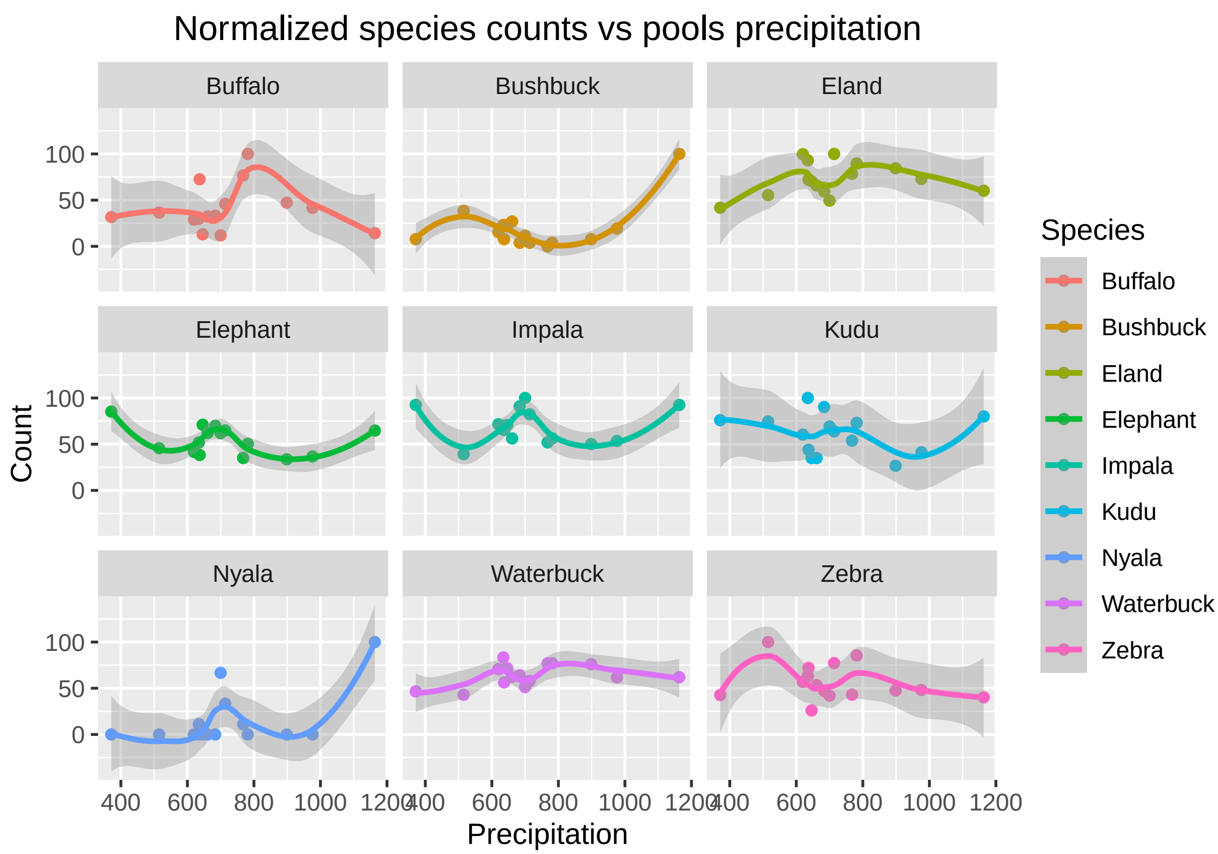



```
ggplot(data=filter(gc_precip_norm_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Precipitation, y=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```



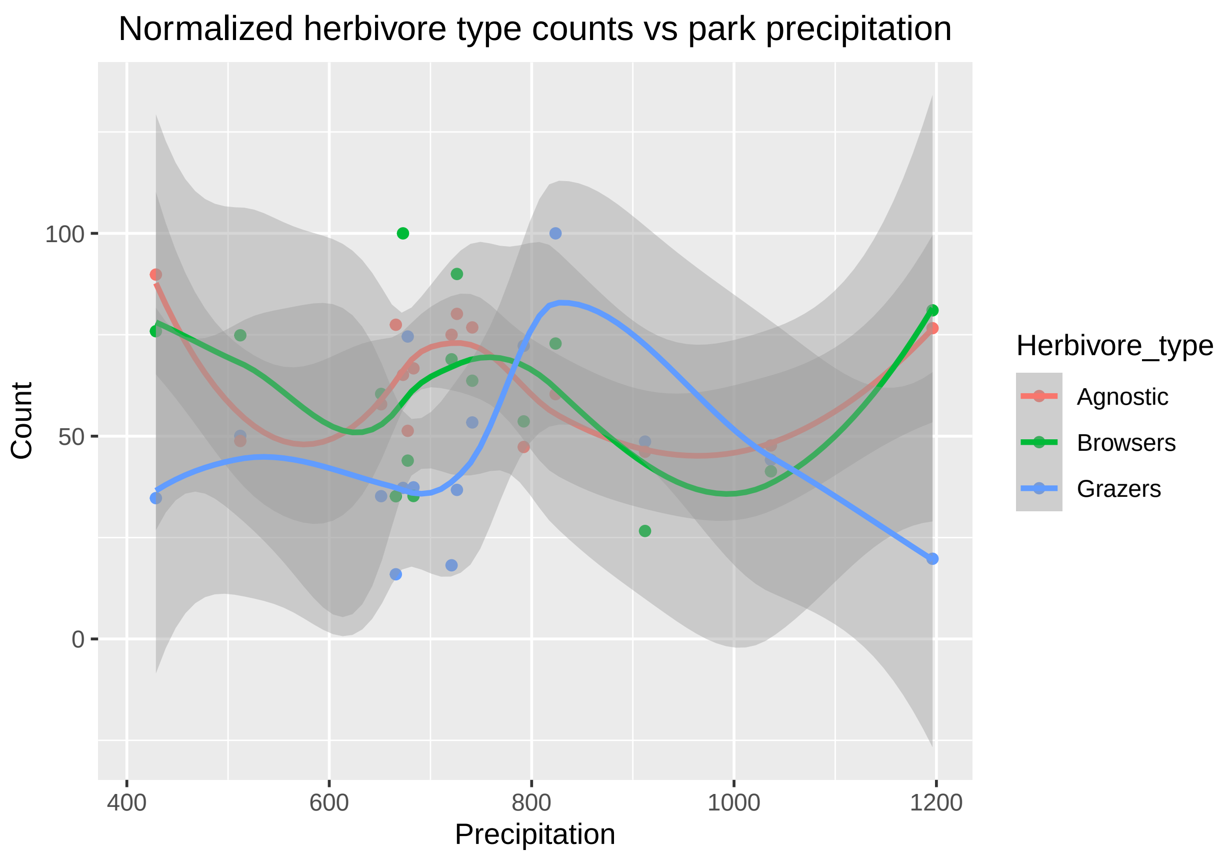

```
gc_precip_norm_type_plot <- gc_norm_precip %>% gather(c(11,12,13), key='Herbivore_type', value='Count')

ggplot(data=filter(gc_precip_norm_type_plot, Area=="park")) + geom_point(mapping=aes(x=Precipitation, y=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).

## Warning: Removed 21 rows containing missing values (geom_point).
```

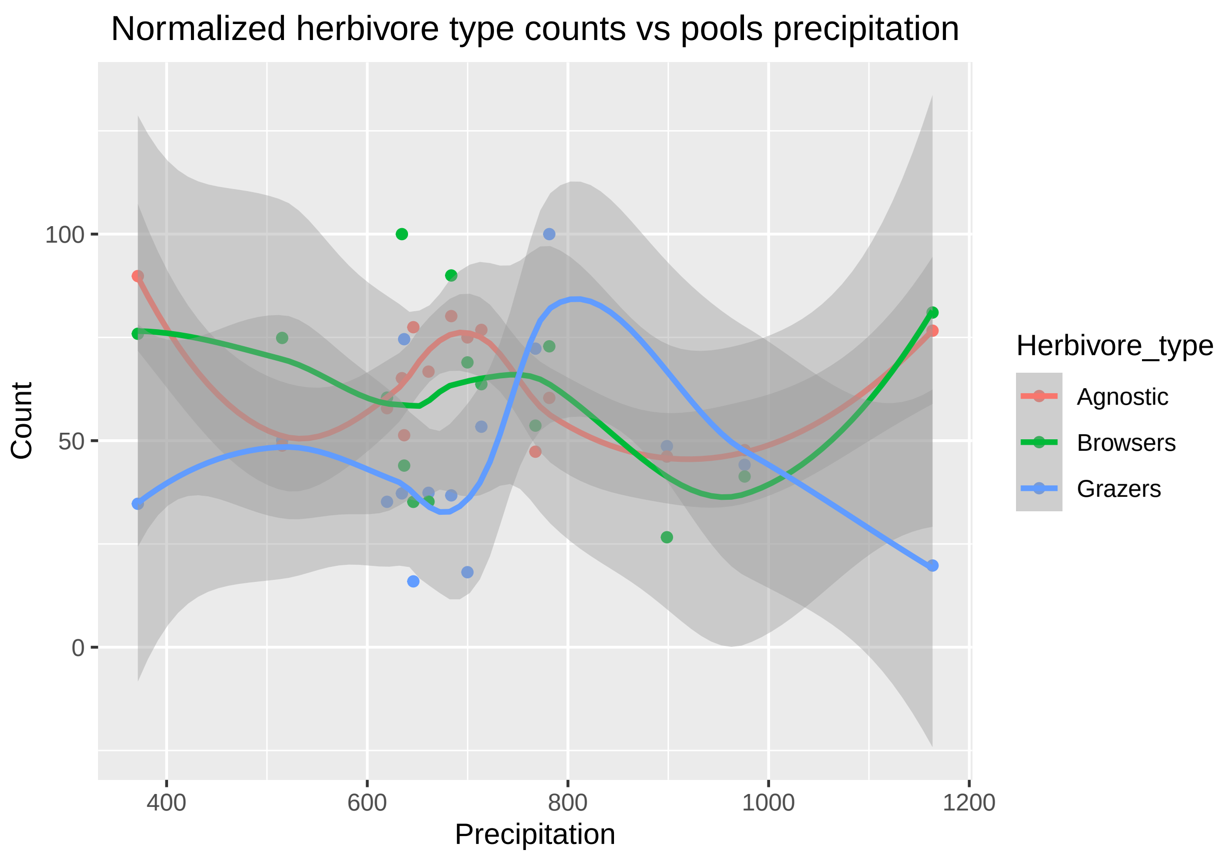



```
ggplot(data=filter(gc_precip_norm_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Precipitation, y
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).

## Warning: Removed 21 rows containing missing values (geom_point).
```



Correlation between soil moisture, drought index, vegetation, and water, and animal counts

I then graphed park and pools dry season climatic factors against LSU values for each species, and park and pools climatic factors against LSU values for each herbivore type.

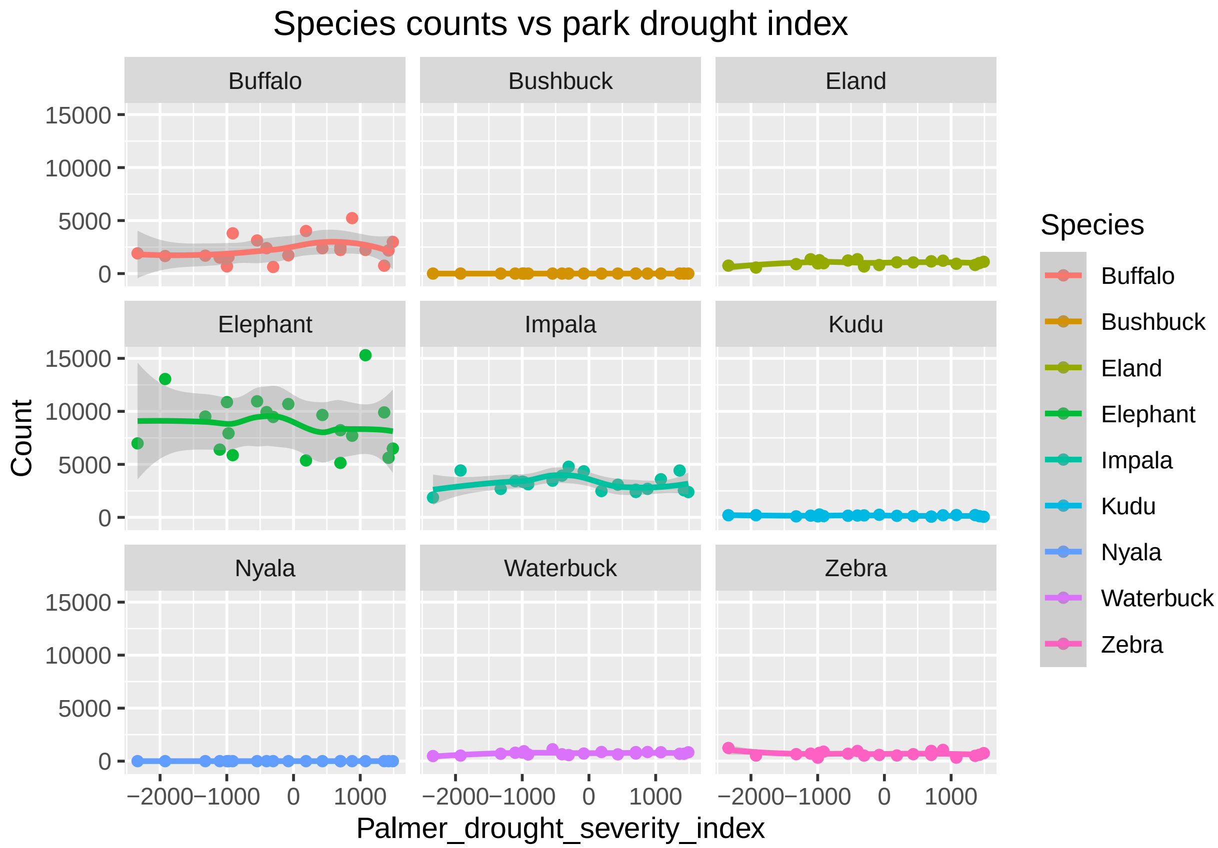
```
gc_clim_for_spp_plot <- gc_clim %>% gather(c(2,3,4,5,6,7,8,9,10), key='Species', value='Count')

ggplot(data=filter(gc_clim_for_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Palmer_drought_severi
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 18 rows containing non-finite values (stat_smooth).

## Warning: Removed 18 rows containing missing values (geom_point).
```

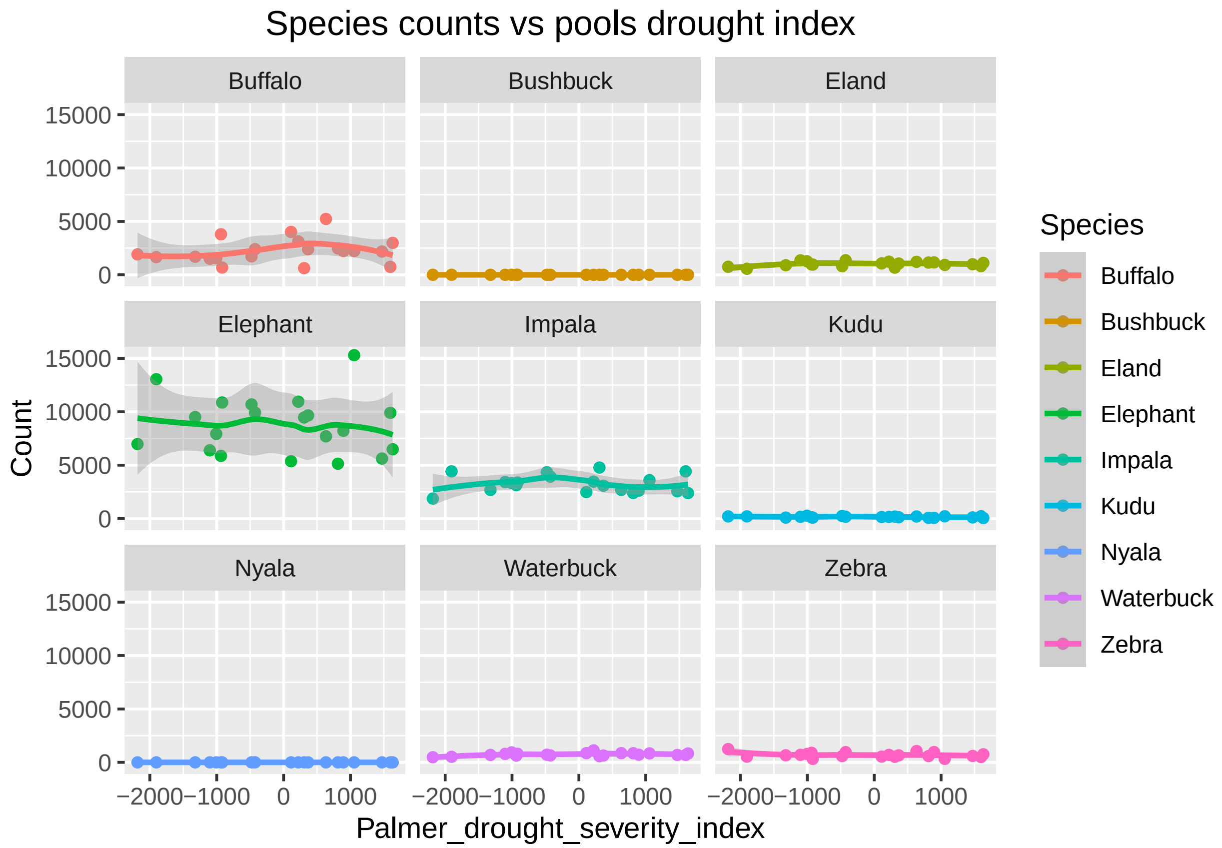



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 18 rows containing non-finite values (stat_smooth).

## Warning: Removed 18 rows containing missing values (geom_point).
```

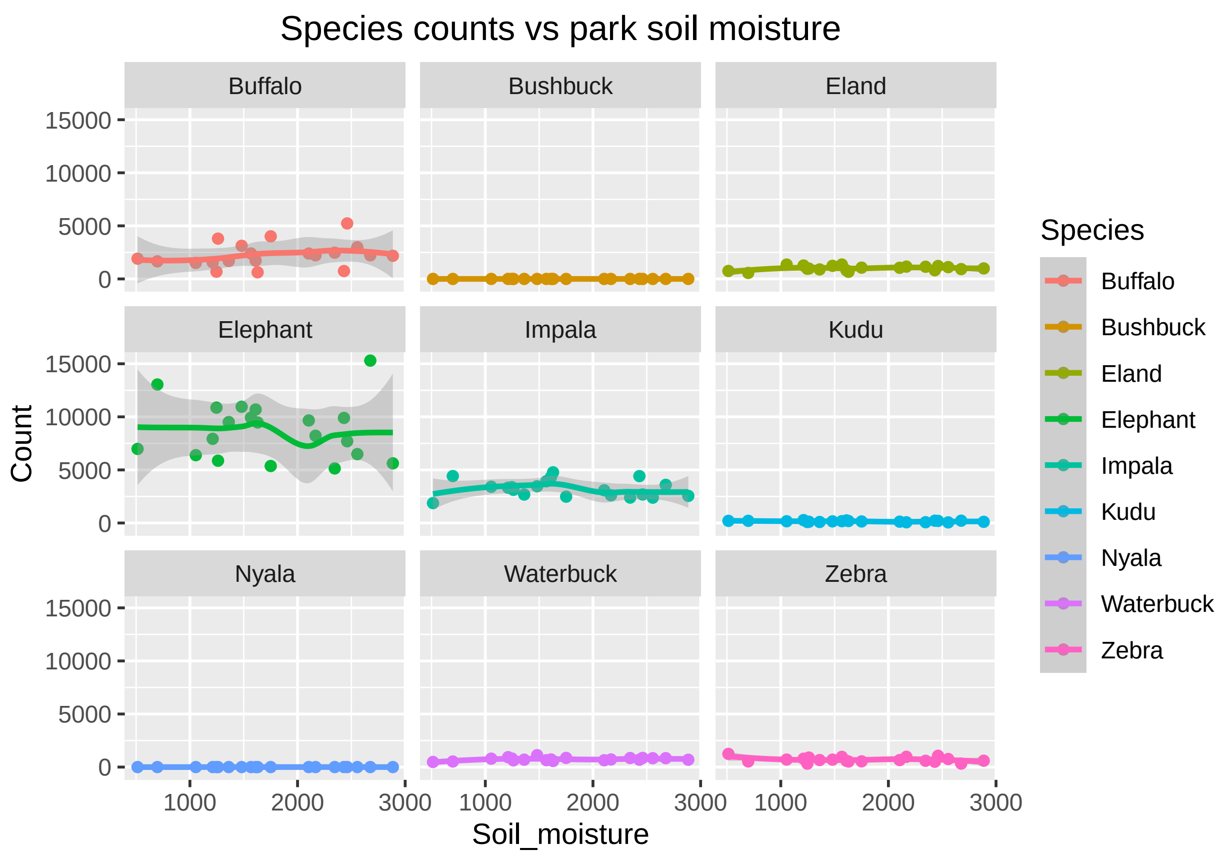



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Soil_moisture, y=Cou
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 18 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 18 rows containing missing values (geom_point).
```

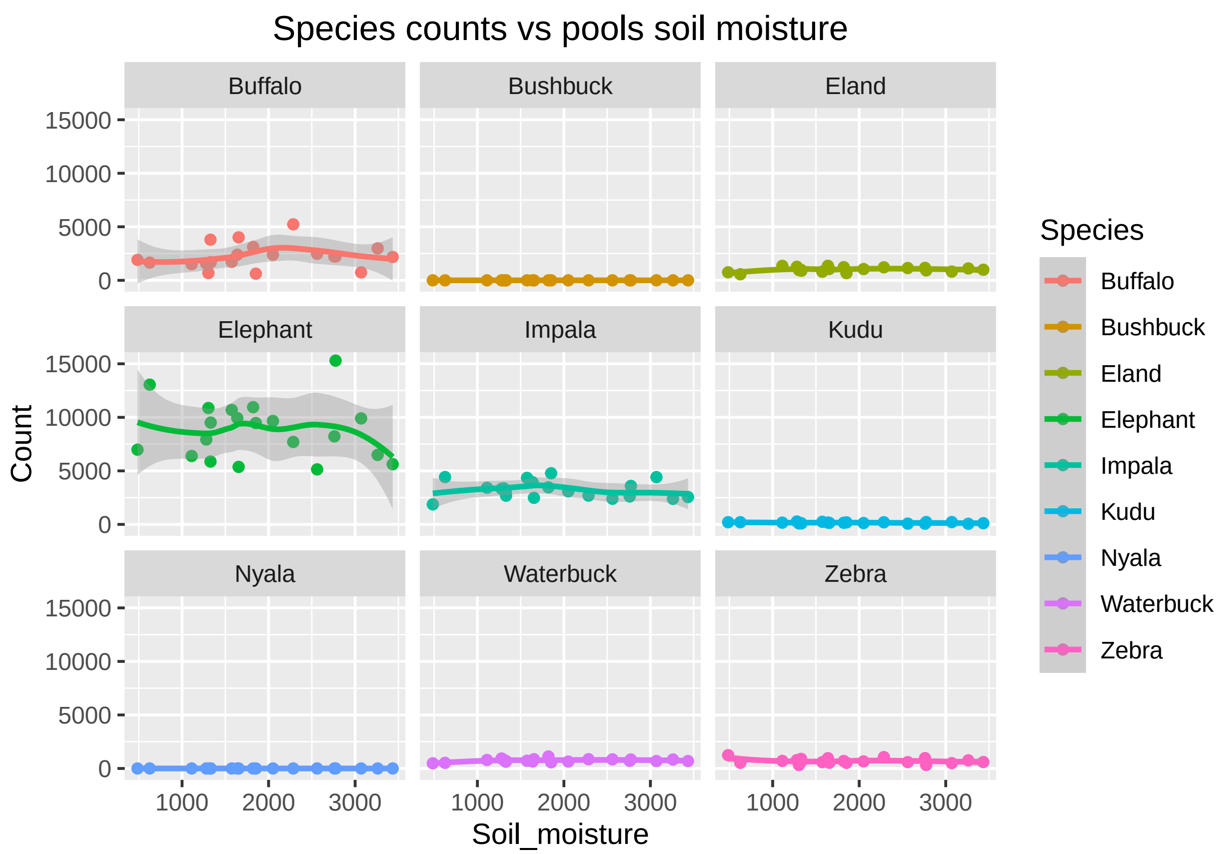



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Soil_moisture, y=Cou
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 18 rows containing non-finite values (stat_smooth).

## Warning: Removed 18 rows containing missing values (geom_point).
```

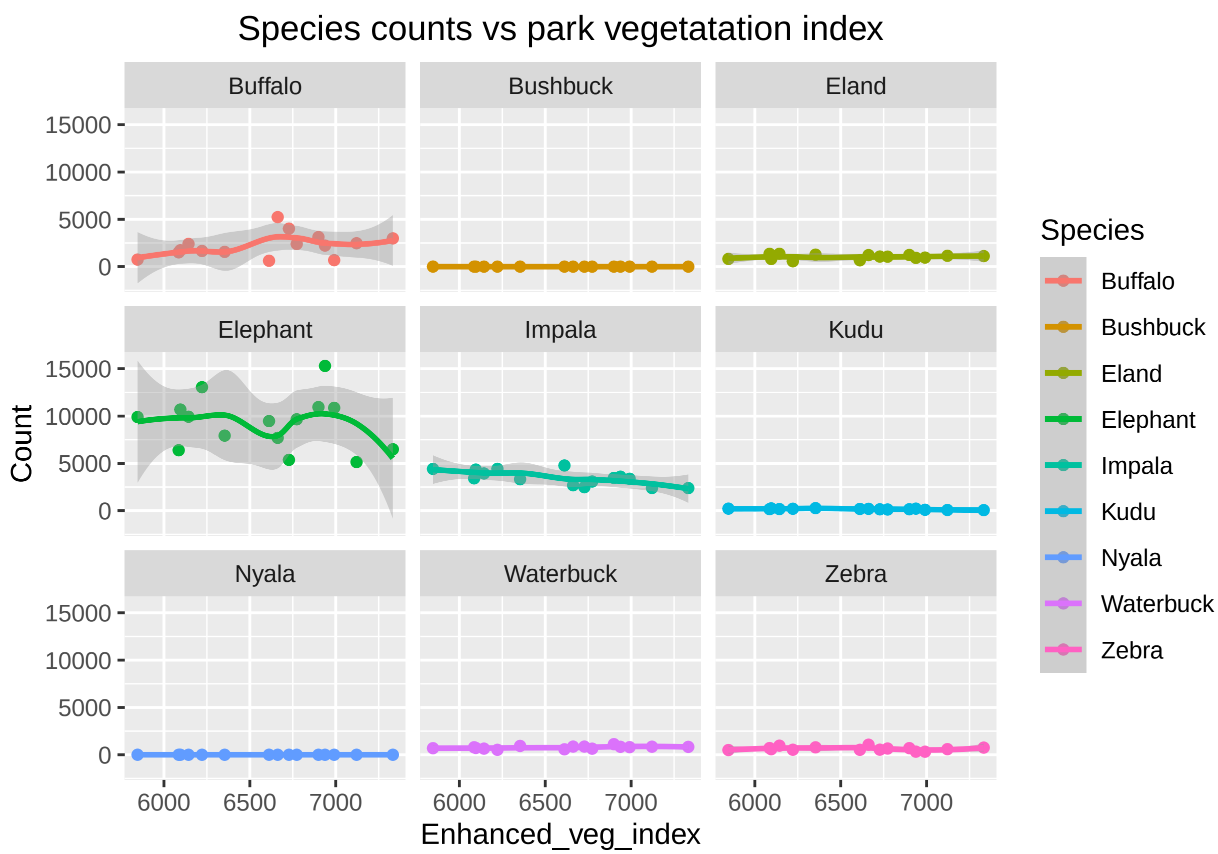



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="park" & Enhanced_veg_index > 0)) + geom_point(mapping=ae
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 9 rows containing non-finite values (stat_smooth).

## Warning: Removed 9 rows containing missing values (geom_point).
```

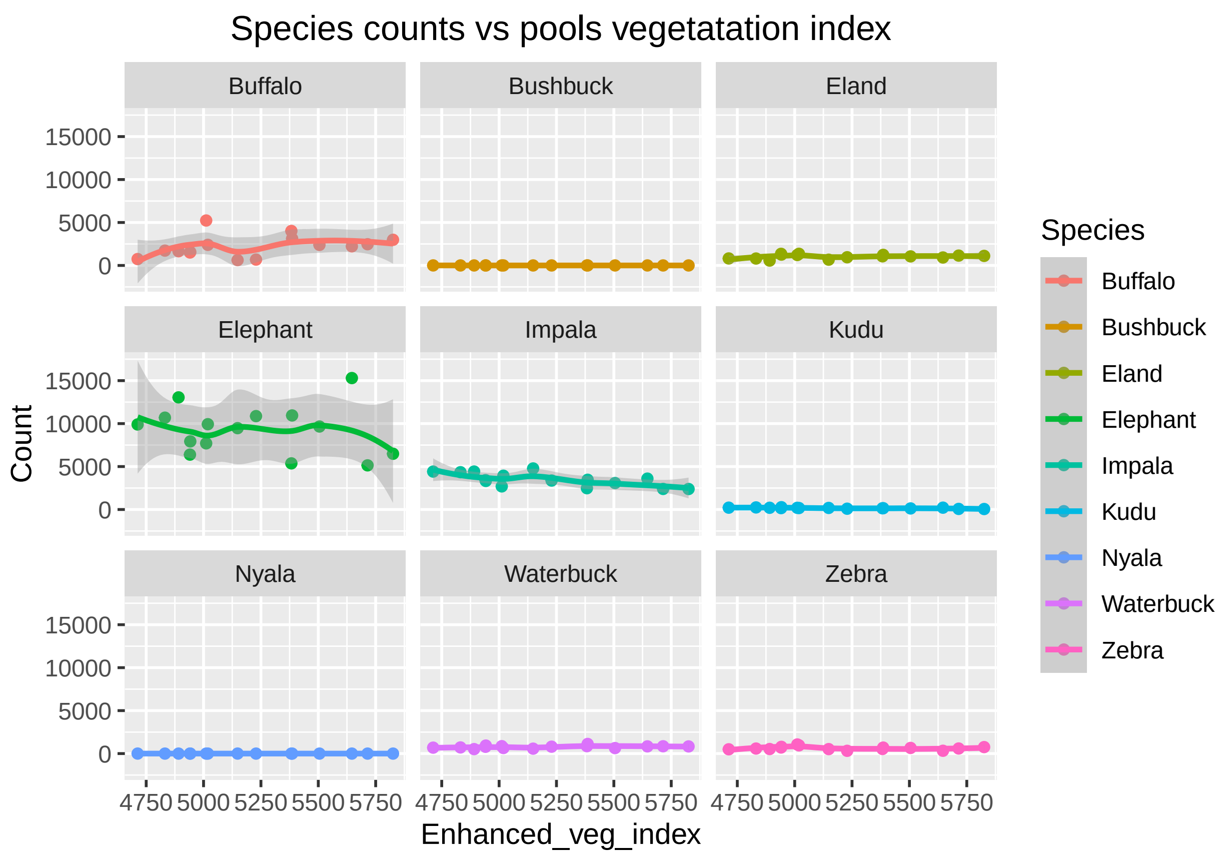



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="pools" & Enhanced_veg_index > 0)) + geom_point(mapping=a
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 9 rows containing non-finite values (stat_smooth).

## Warning: Removed 9 rows containing missing values (geom_point).
```

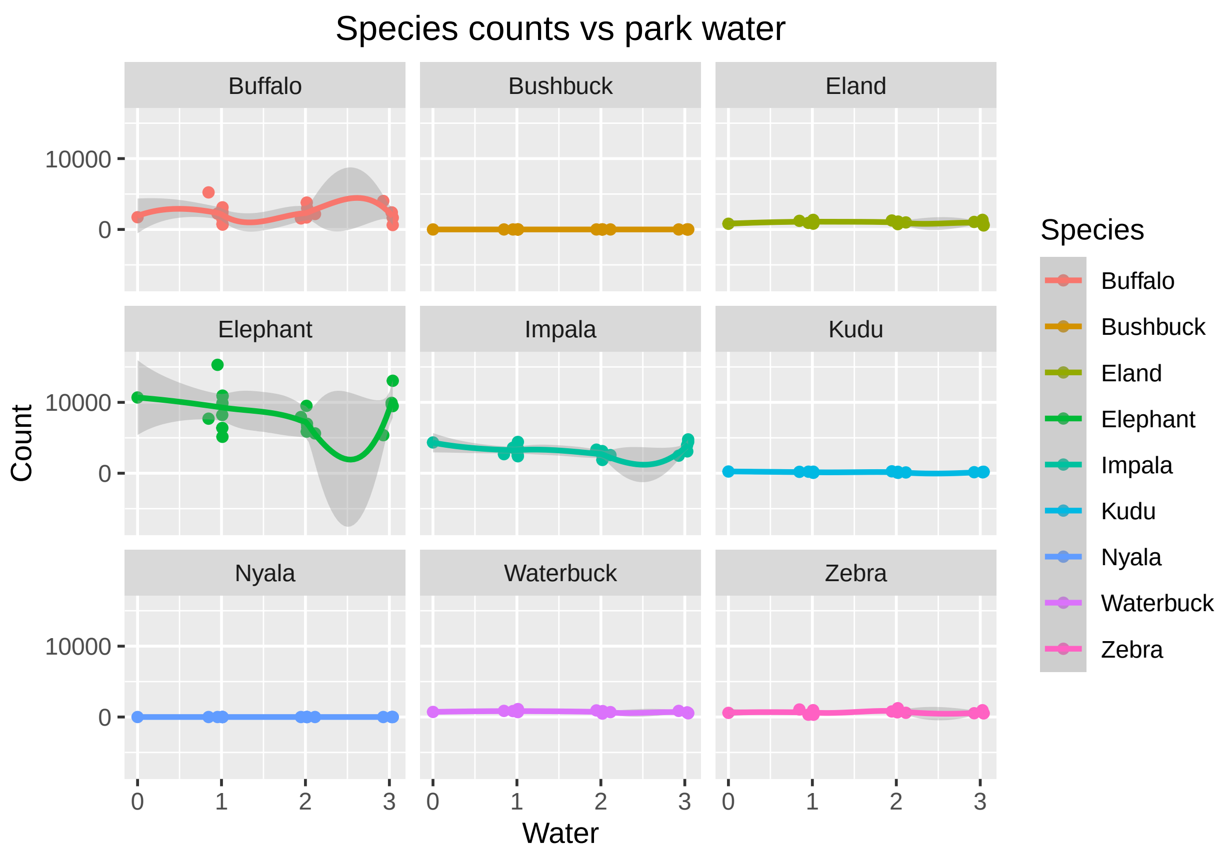



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Water, y=Count, colo
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 18 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 18 rows containing missing values (geom_point).
```

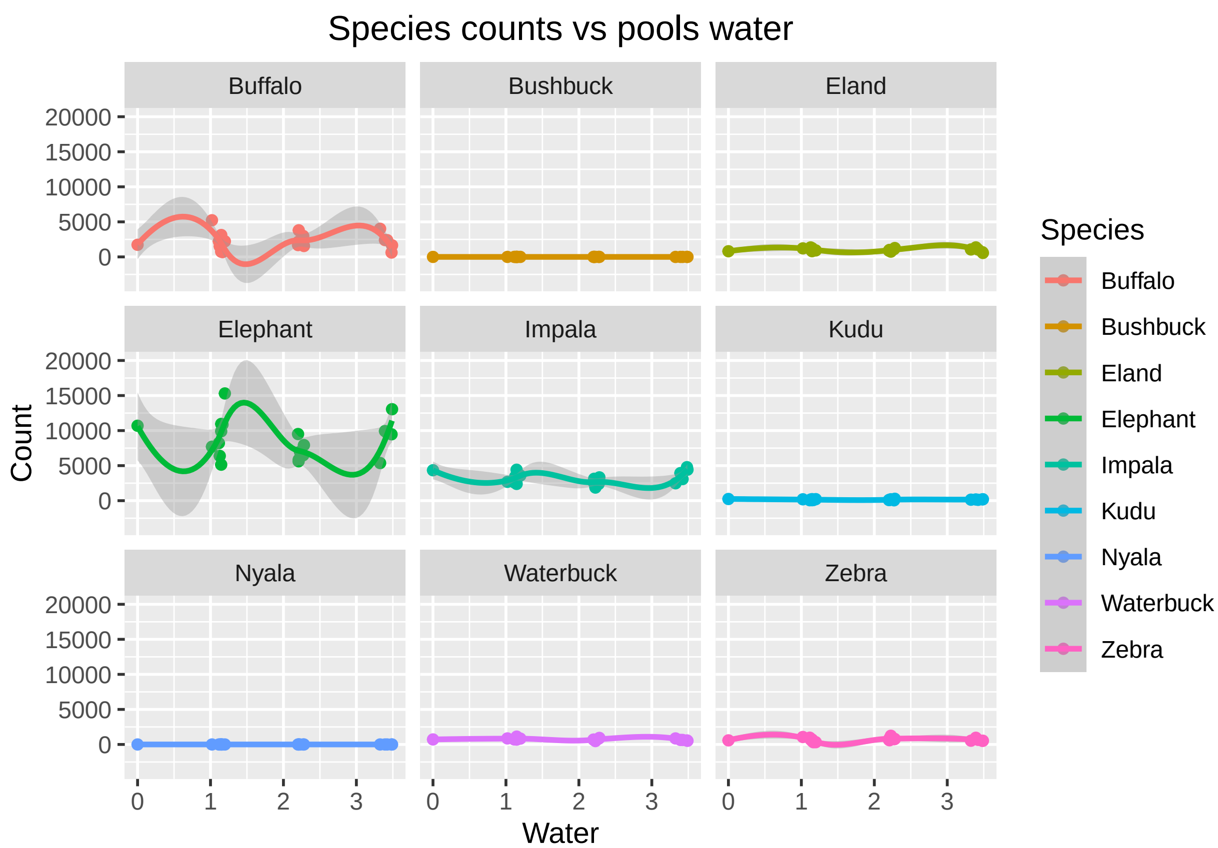



```
ggplot(data=filter(gc_clim_for_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Water, y=Count, colo
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 18 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 18 rows containing missing values (geom_point).
```



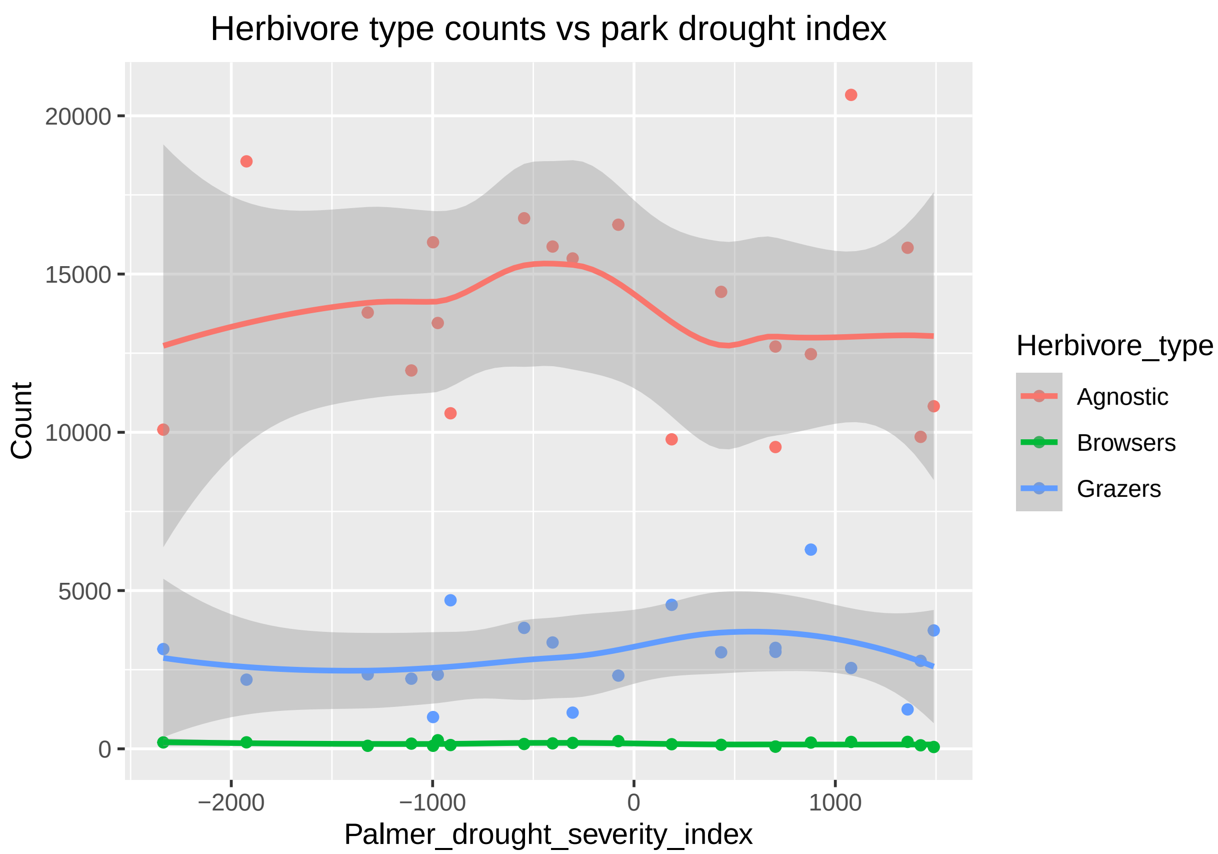

```
gc_clim_for_type_plot <- gc_clim %>% gather(c(11,12,13), key='Herbivore_type', value='Count')

ggplot(data=filter(gc_clim_for_type_plot, Area=="park")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 6 rows containing non-finite values (stat_smooth).

## Warning: Removed 6 rows containing missing values (geom_point).
```

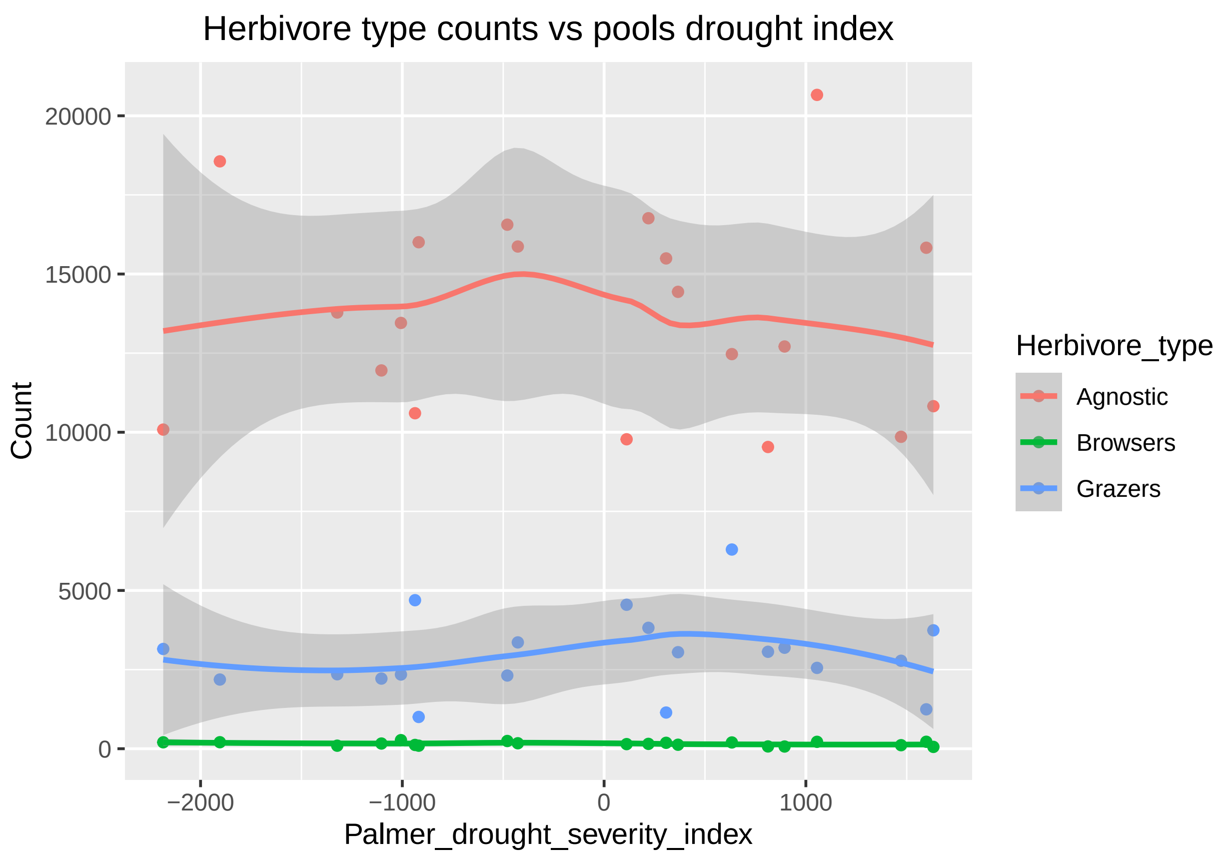



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 6 rows containing non-finite values (stat_smooth).

## Warning: Removed 6 rows containing missing values (geom_point).
```

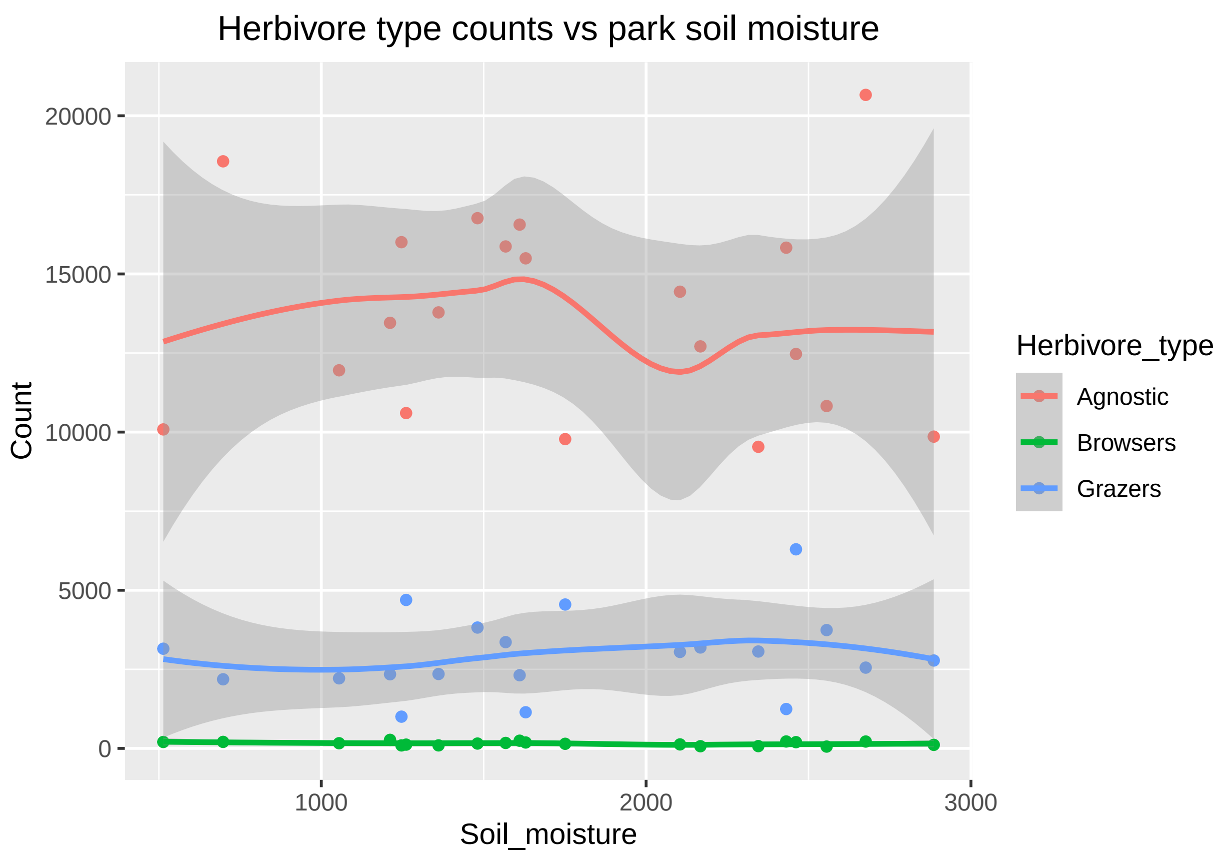



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="park")) + geom_point(mapping=aes(x=Soil_moisture, y=Cou
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```

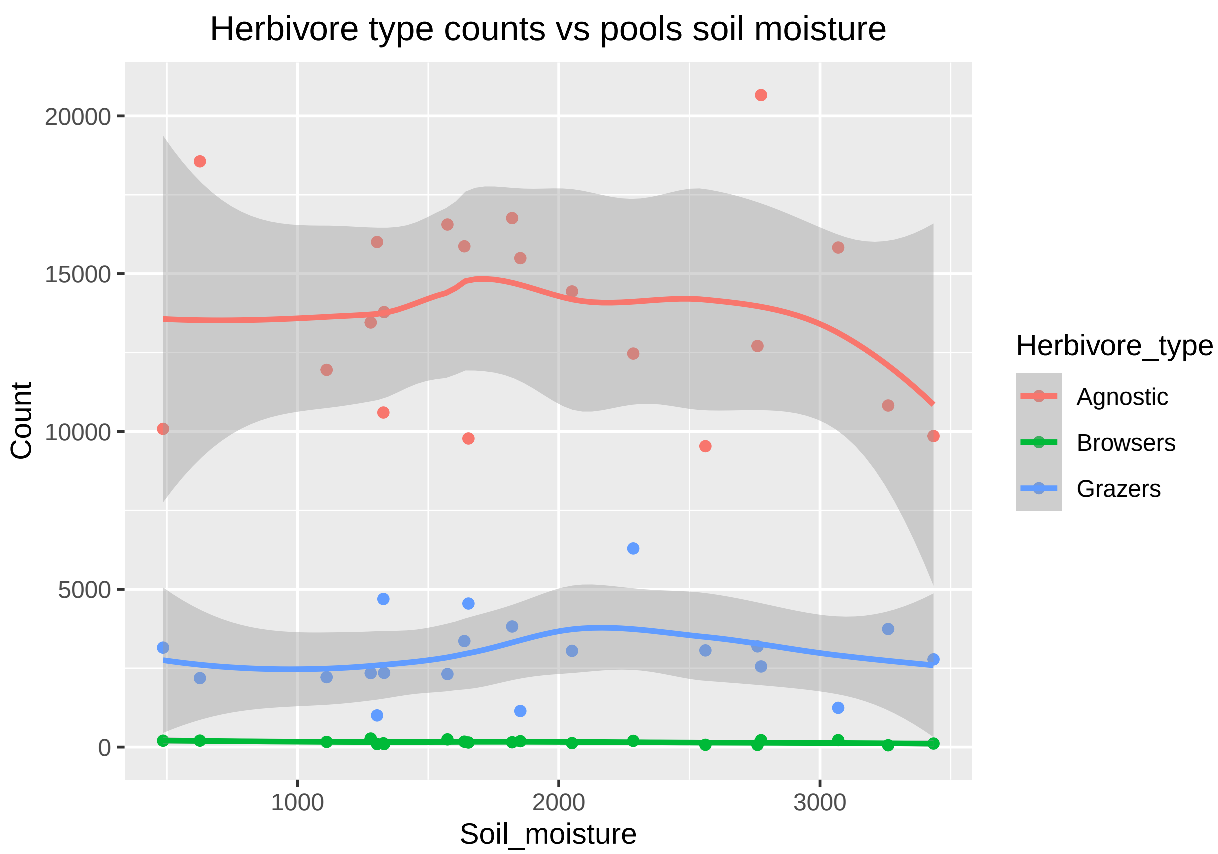



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Soil_moisture, y=Co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```

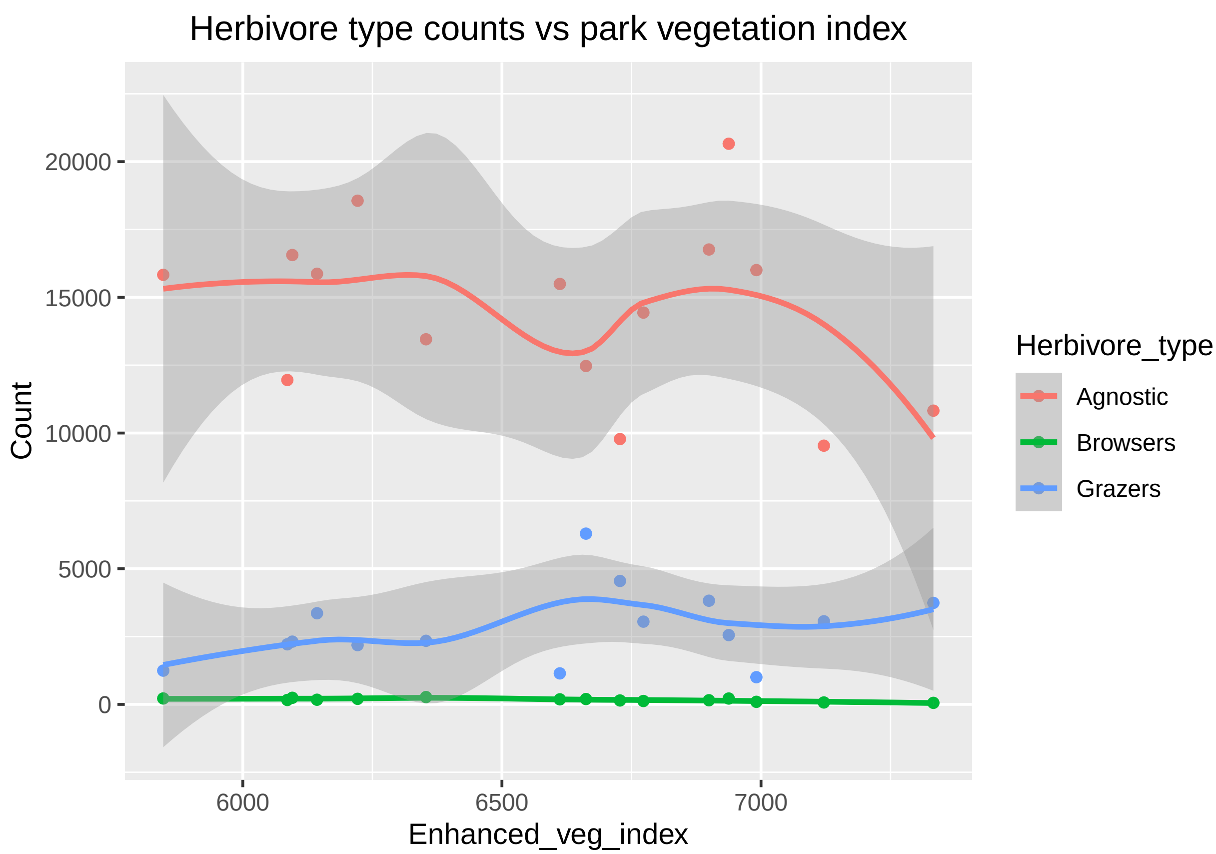



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="park" & Enhanced_veg_index > 0)) + geom_point(mapping=a
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 3 rows containing non-finite values (stat_smooth).

## Warning: Removed 3 rows containing missing values (geom_point).
```

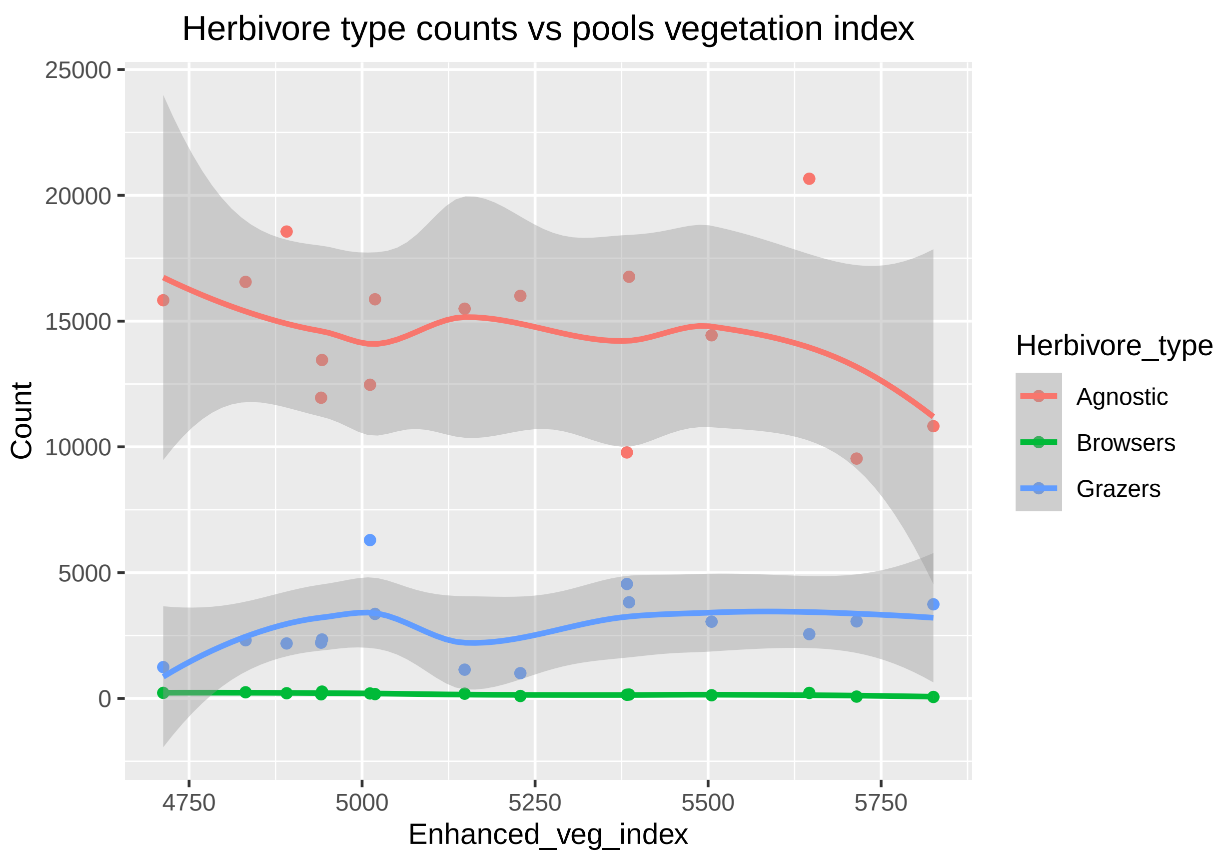



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="pools" & Enhanced_veg_index > 0)) + geom_point(mapping=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
```

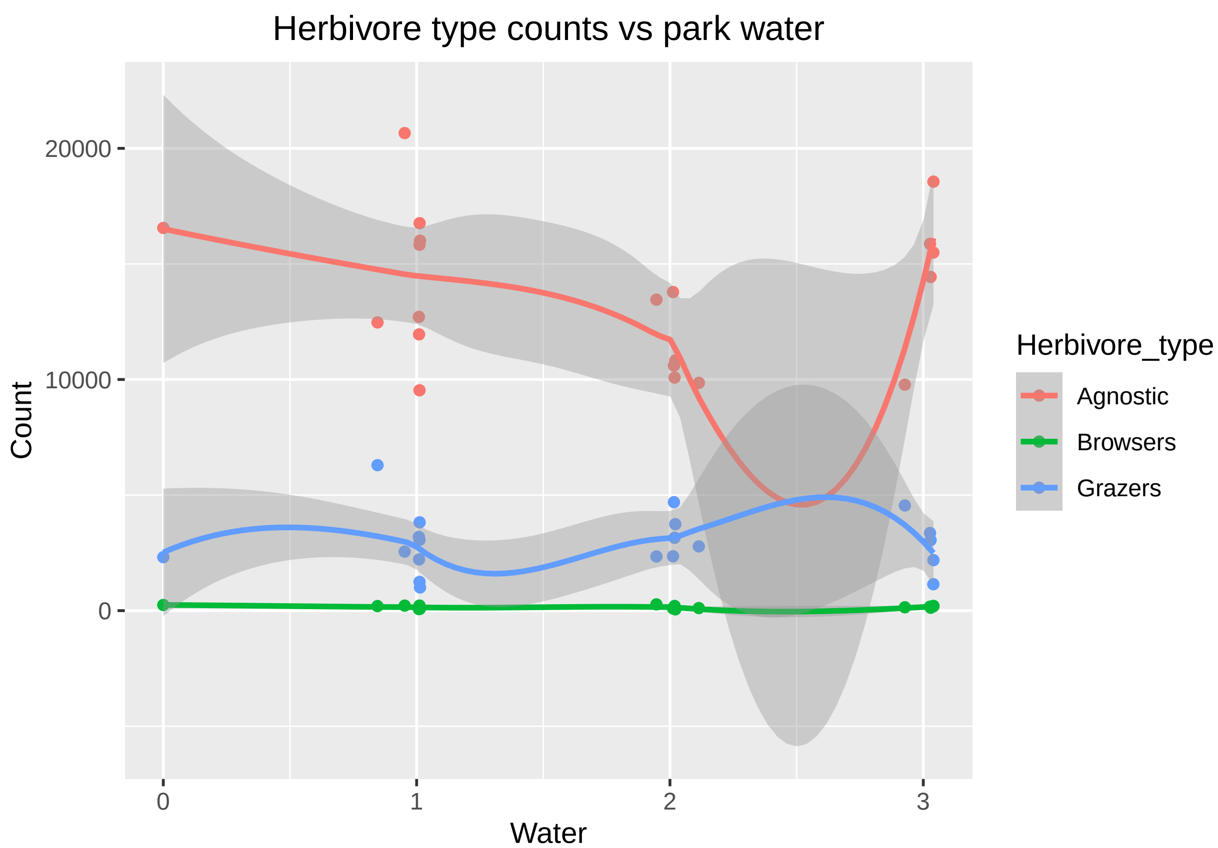



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="park")) + geom_point(mapping=aes(x=Water, y=Count, col
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 6 rows containing non-finite values (stat_smooth).

## Warning: Removed 6 rows containing missing values (geom_point).
```

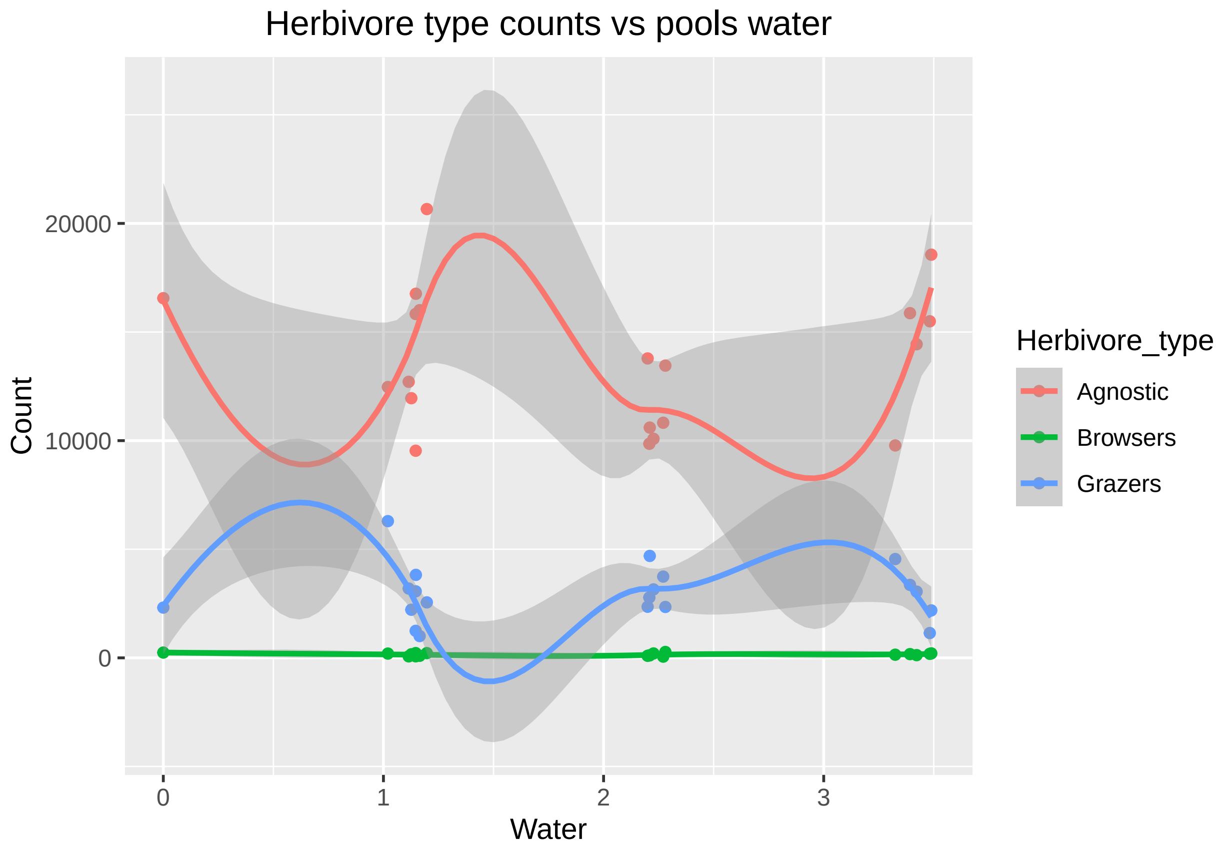



```
ggplot(data=filter(gc_clim_for_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Water, y=Count, co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



I then graphed park and pools dry season climatic factors against the normalized LSU values for each species, and park and pools climatic factors against the normalized LSU values for each herbivore type.

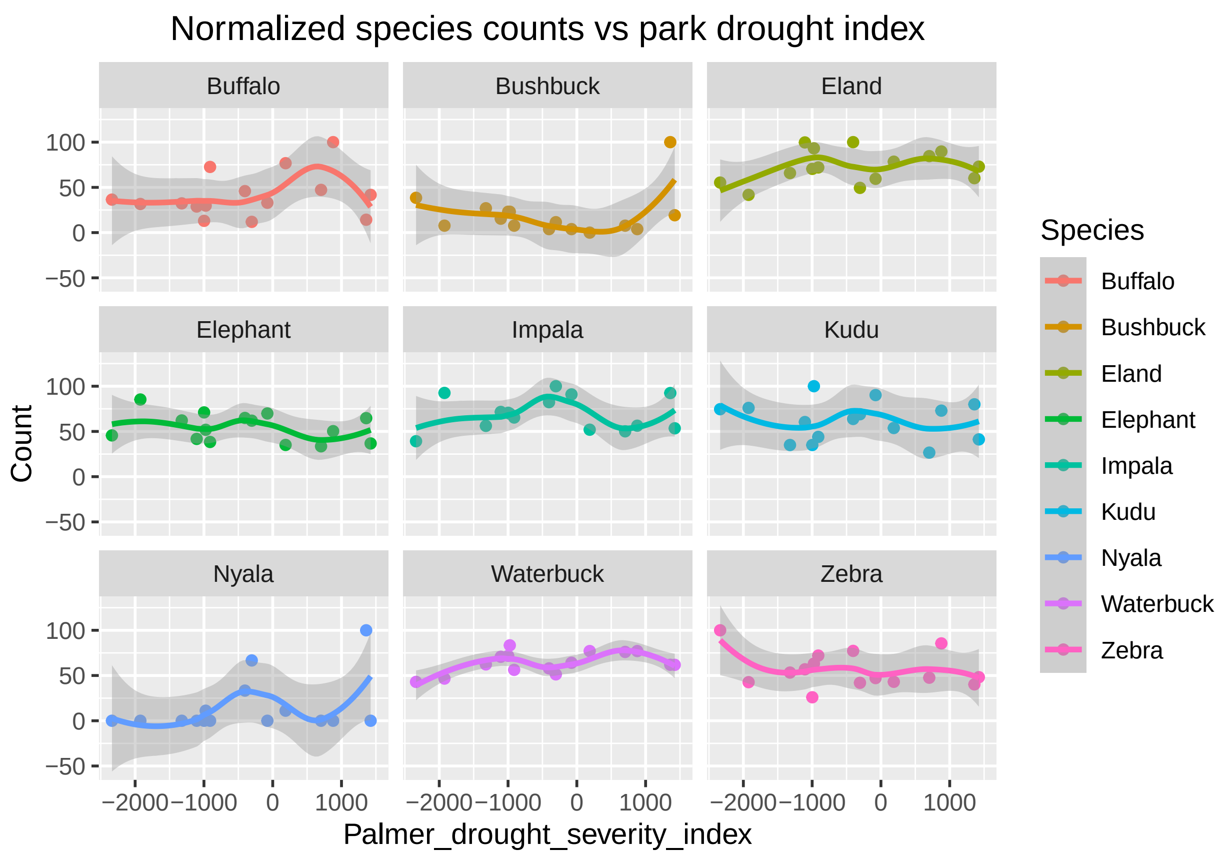
```
gc_clim_norm_spp_plot <- gc_norm_clim %>% gather(c(2,3,4,5,6,7,8,9,10), key='Species', value='Count')

ggplot(data=filter(gc_clim_norm_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

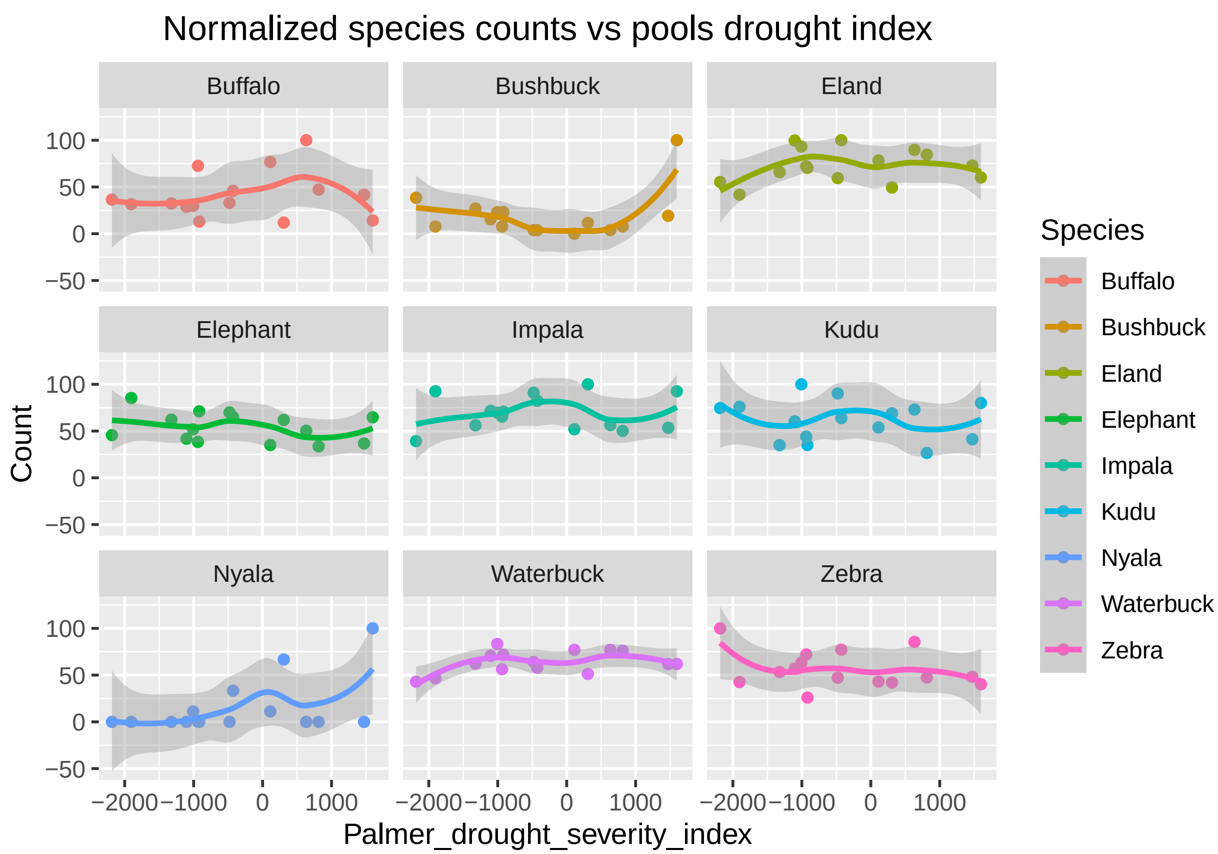



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

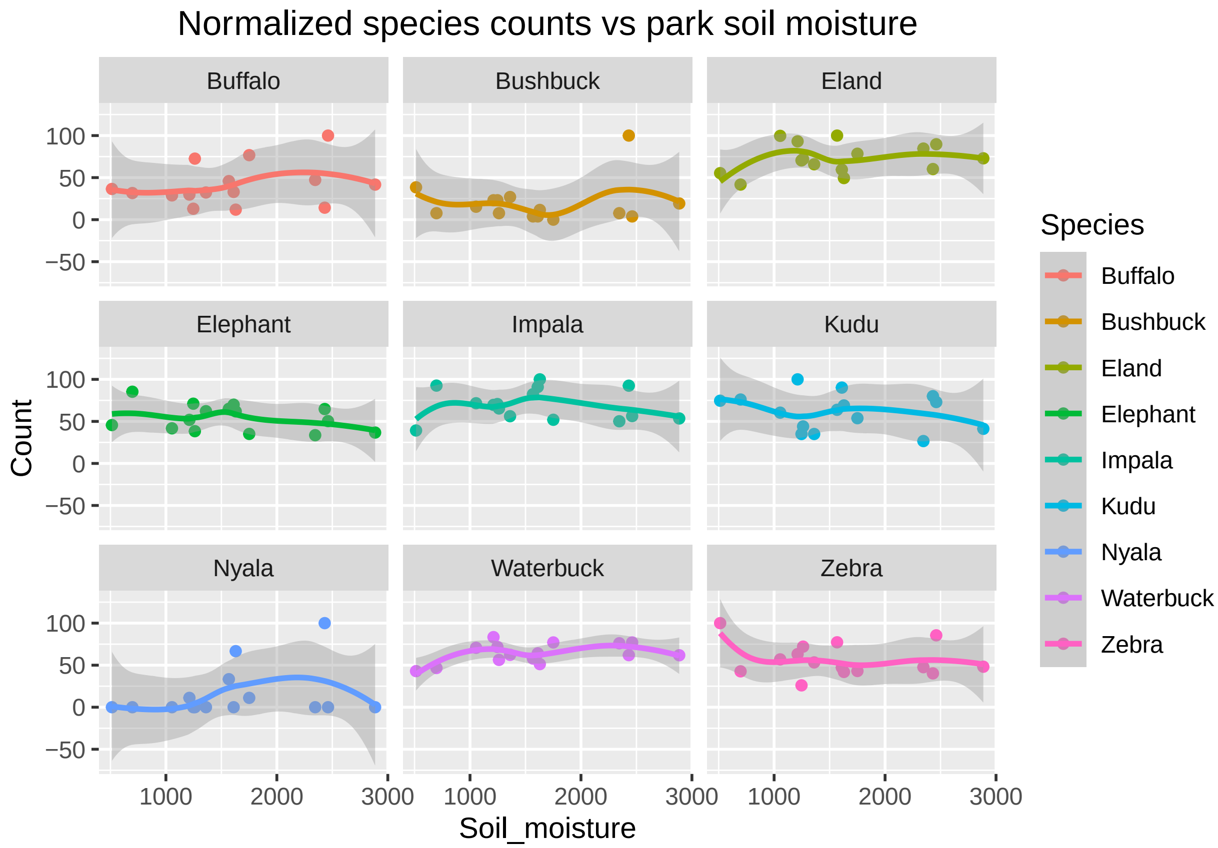



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Soil_moisture, y=Cou
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

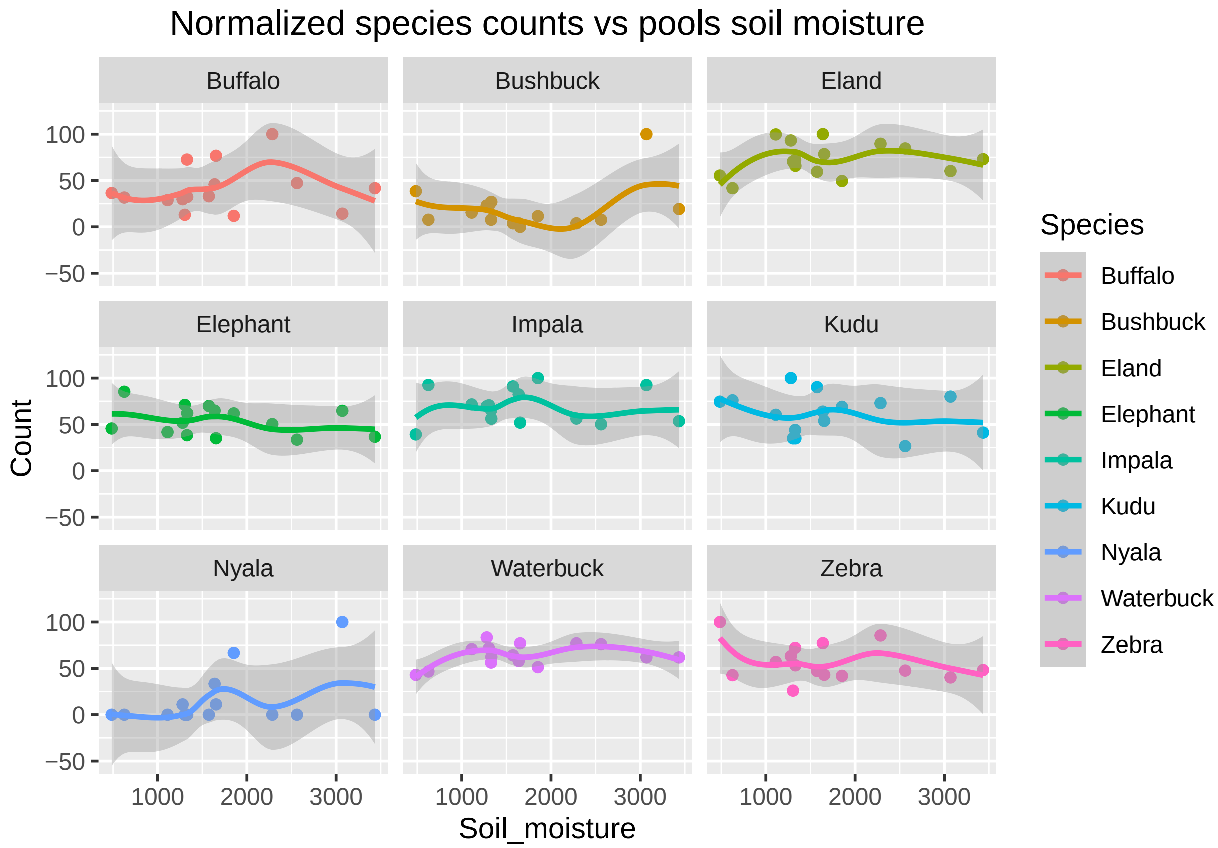



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Soil_moisture, y=Co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

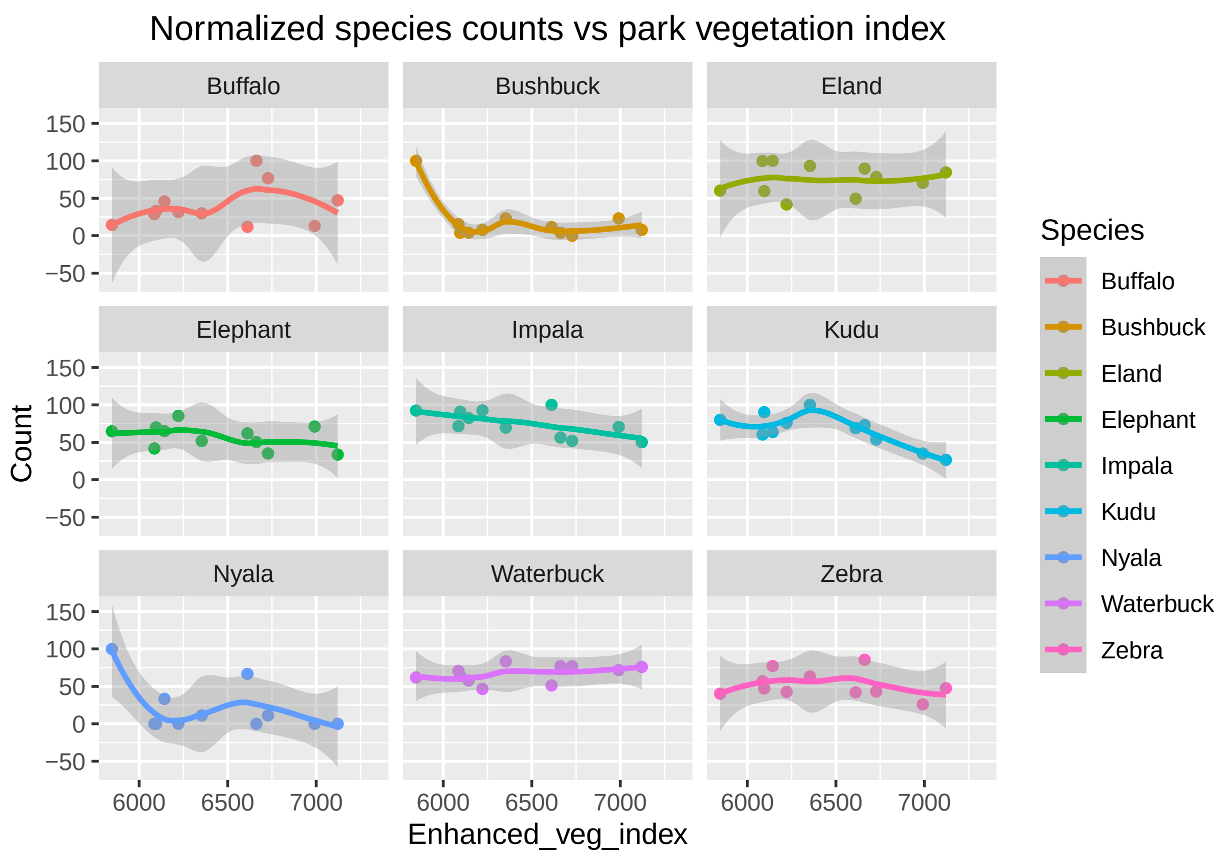



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="park" & Enhanced_veg_index > 0)) + geom_point(mapping=a
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 45 rows containing non-finite values (stat_smooth).

## Warning: Removed 45 rows containing missing values (geom_point).
```

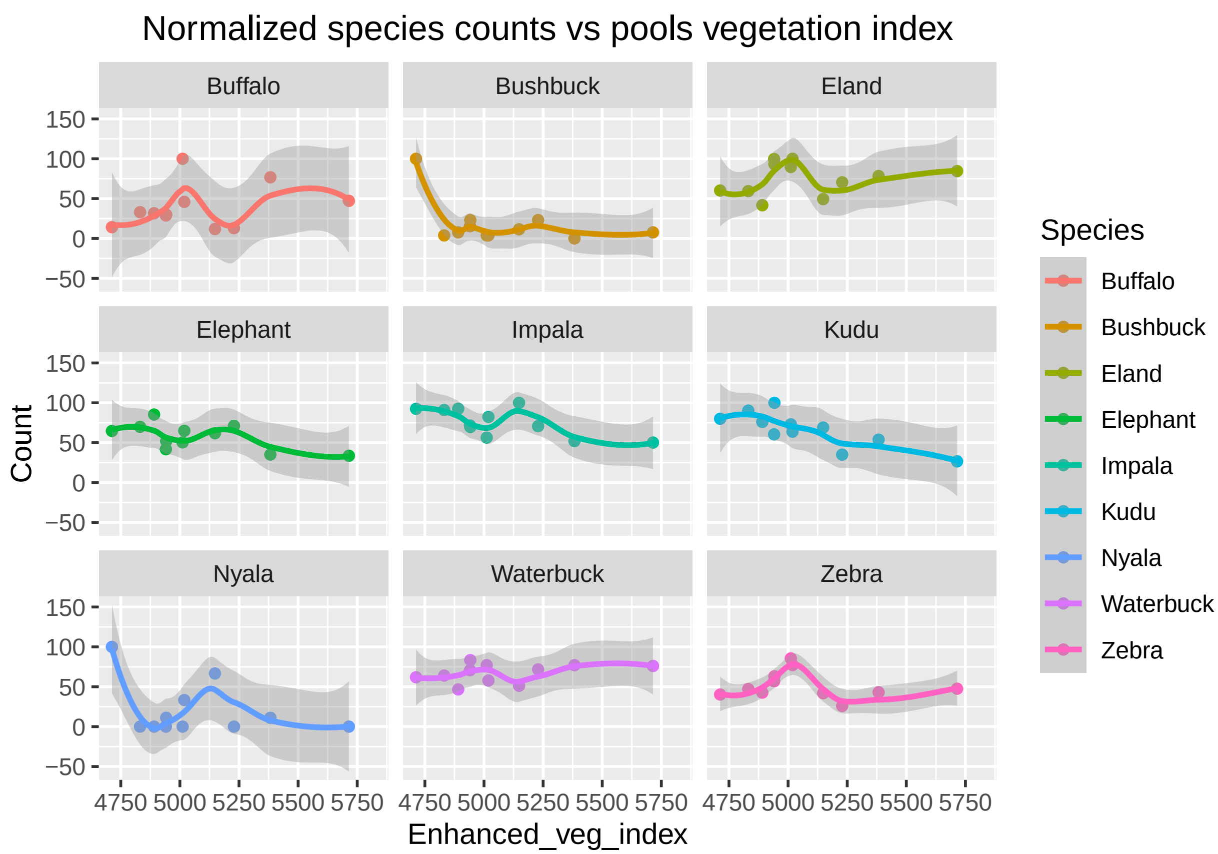



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="pools" & Enhanced_veg_index > 0)) + geom_point(mapping=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 45 rows containing non-finite values (stat_smooth).

## Warning: Removed 45 rows containing missing values (geom_point).
```

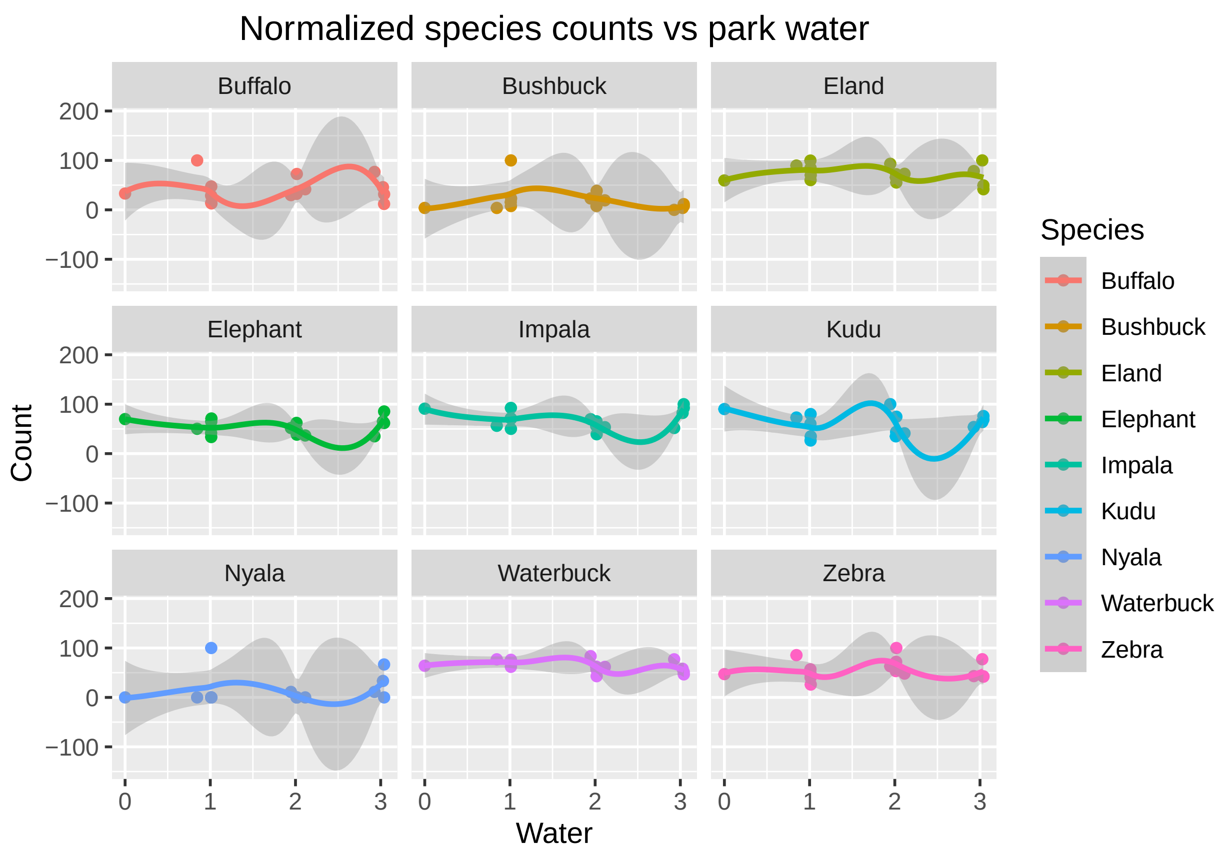



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="park")) + geom_point(mapping=aes(x=Water, y=Count, col
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 63 rows containing non-finite values (stat_smooth).

## Warning: Removed 63 rows containing missing values (geom_point).
```

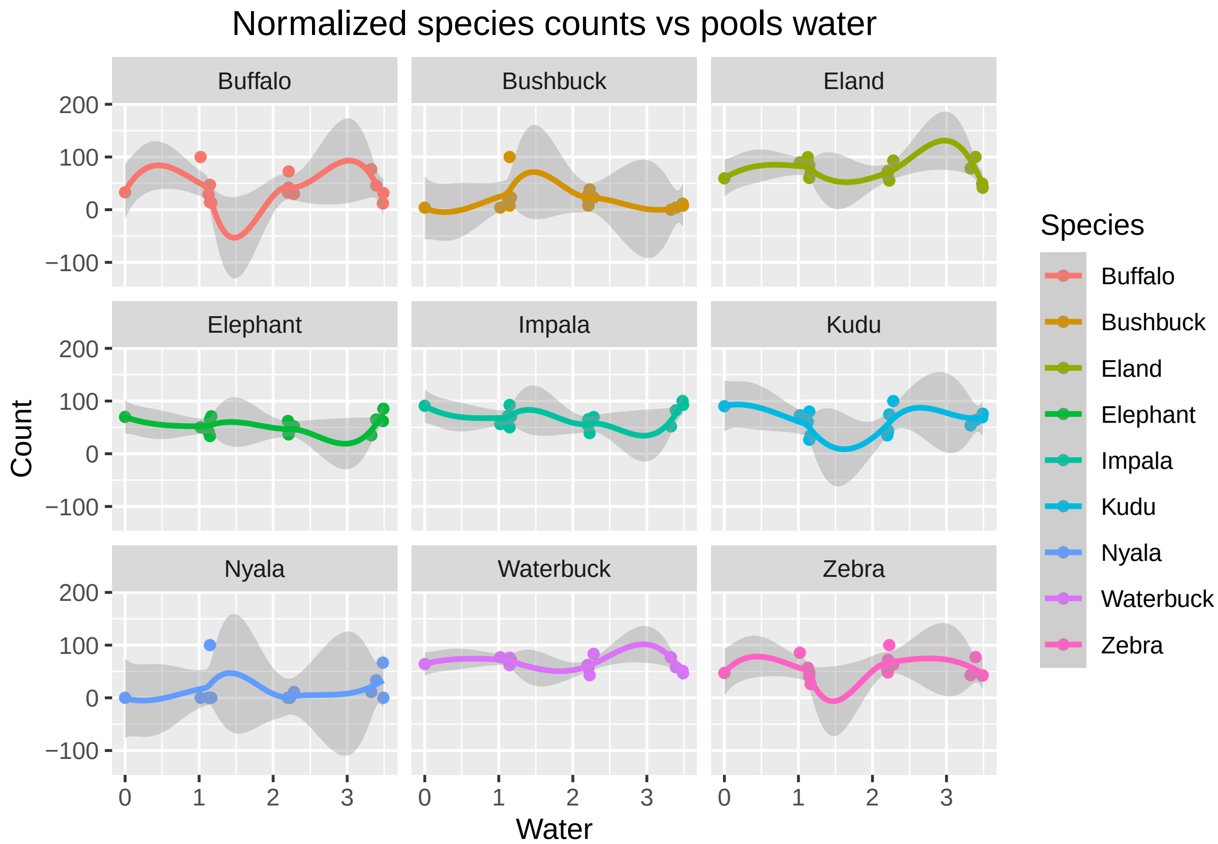



```
ggplot(data=filter(gc_clim_norm_spp_plot, Area=="pools")) + geom_point(mapping=aes(x=Water, y=Count, co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 63 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 63 rows containing missing values (geom_point).
```



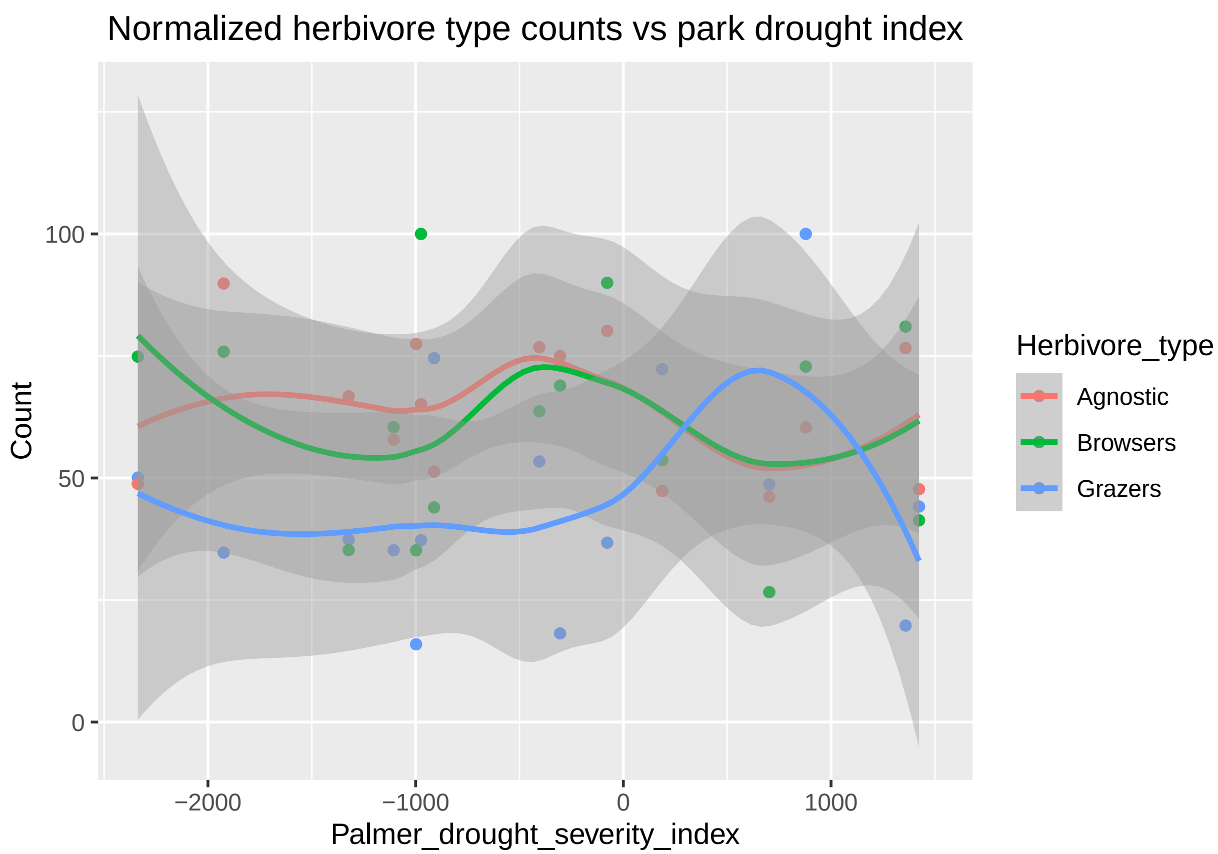

```
gc_clim_norm_type_plot <- gc_norm_clim %>% gather(c(11,12,13), key='Herbivore_type', value='Count')

ggplot(data=filter(gc_clim_norm_type_plot, Area=="park")) + geom_point(mapping=aes(x=Palmer_drought_seve
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).

## Warning: Removed 21 rows containing missing values (geom_point).
```

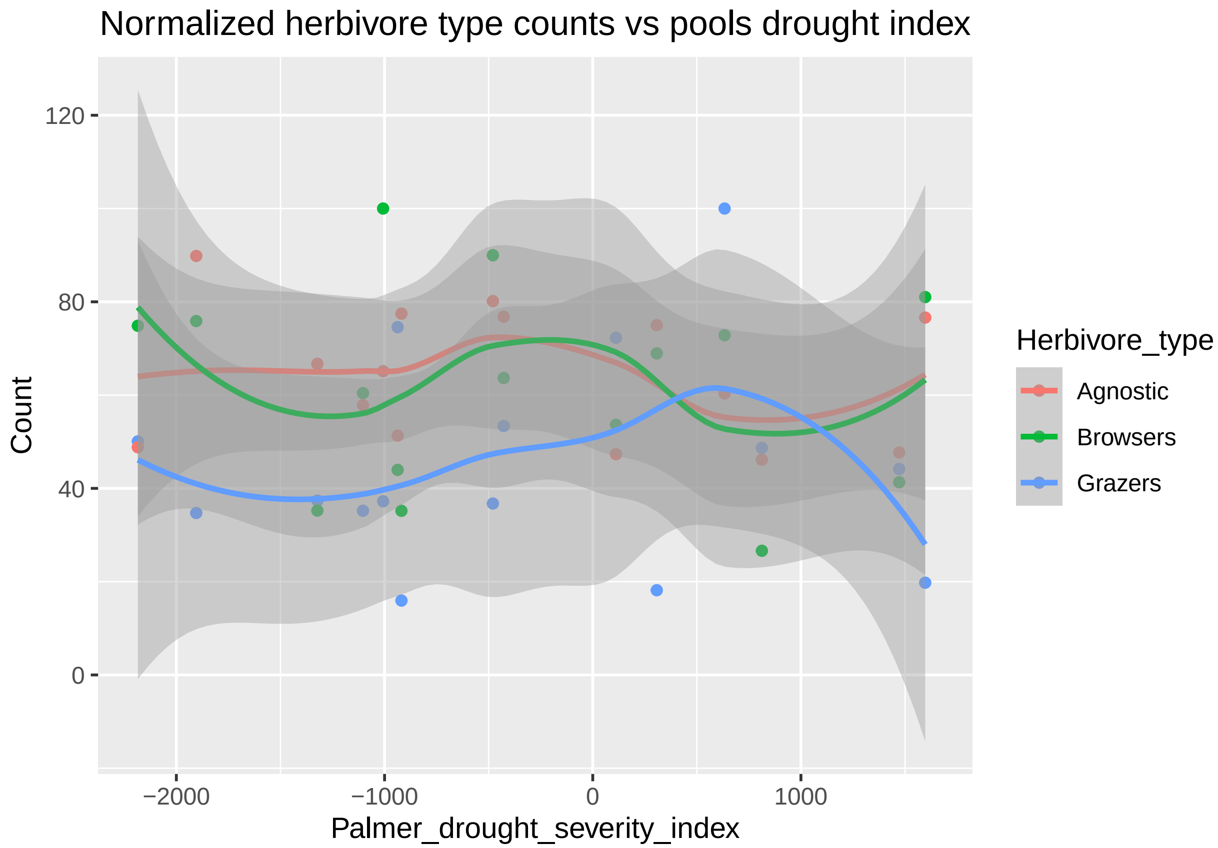



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Palmer_drought_sev
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 21 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 21 rows containing missing values (geom_point).
```

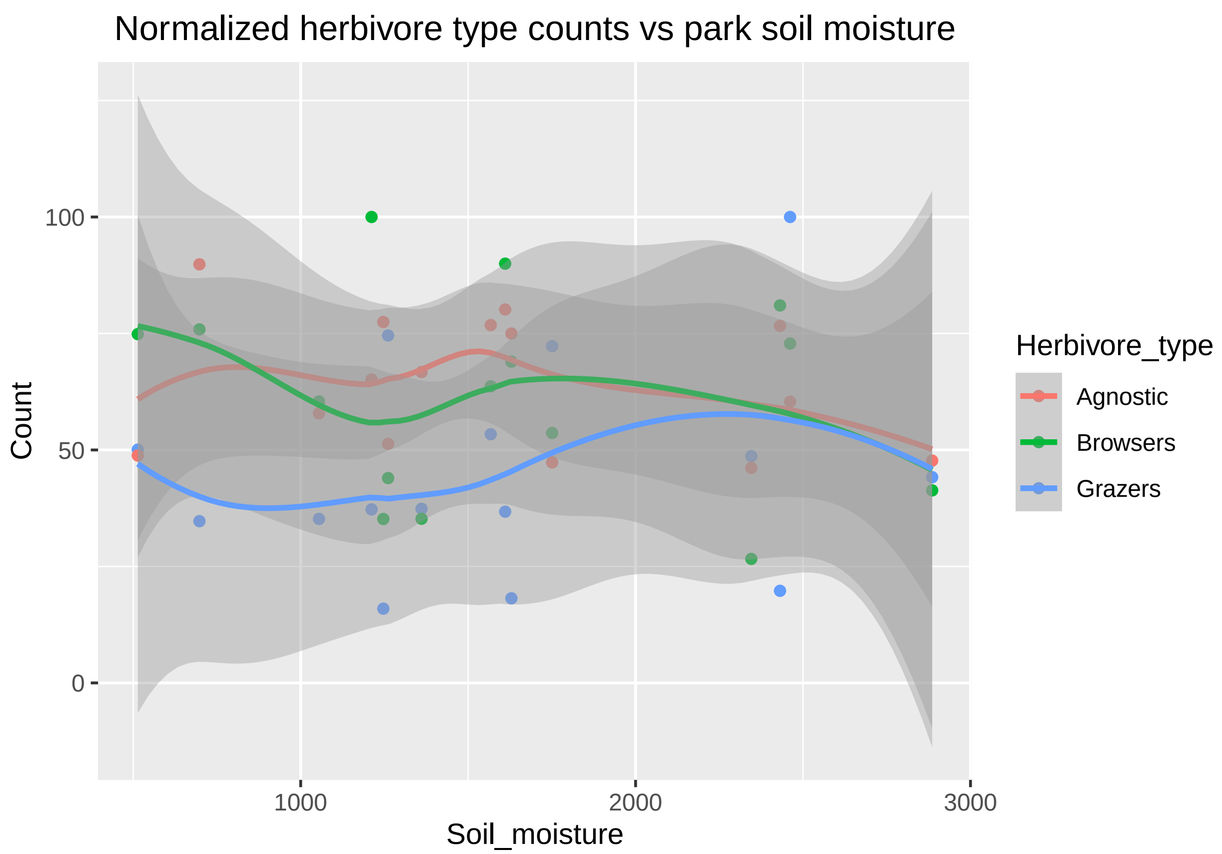



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="park")) + geom_point(mapping=aes(x=Soil_moisture, y=Co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 21 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 21 rows containing missing values (geom_point).
```

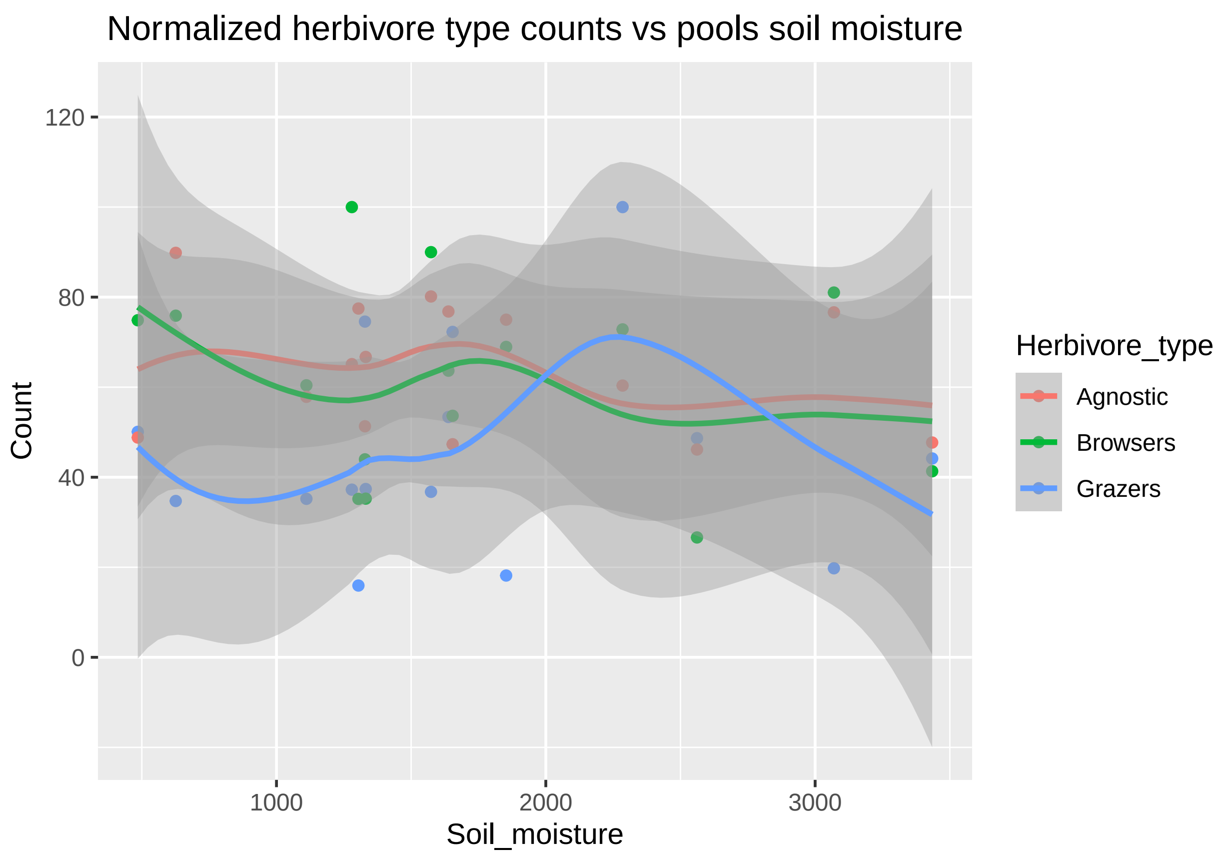



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Soil_moisture, y=C
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).

## Warning: Removed 21 rows containing missing values (geom_point).
```

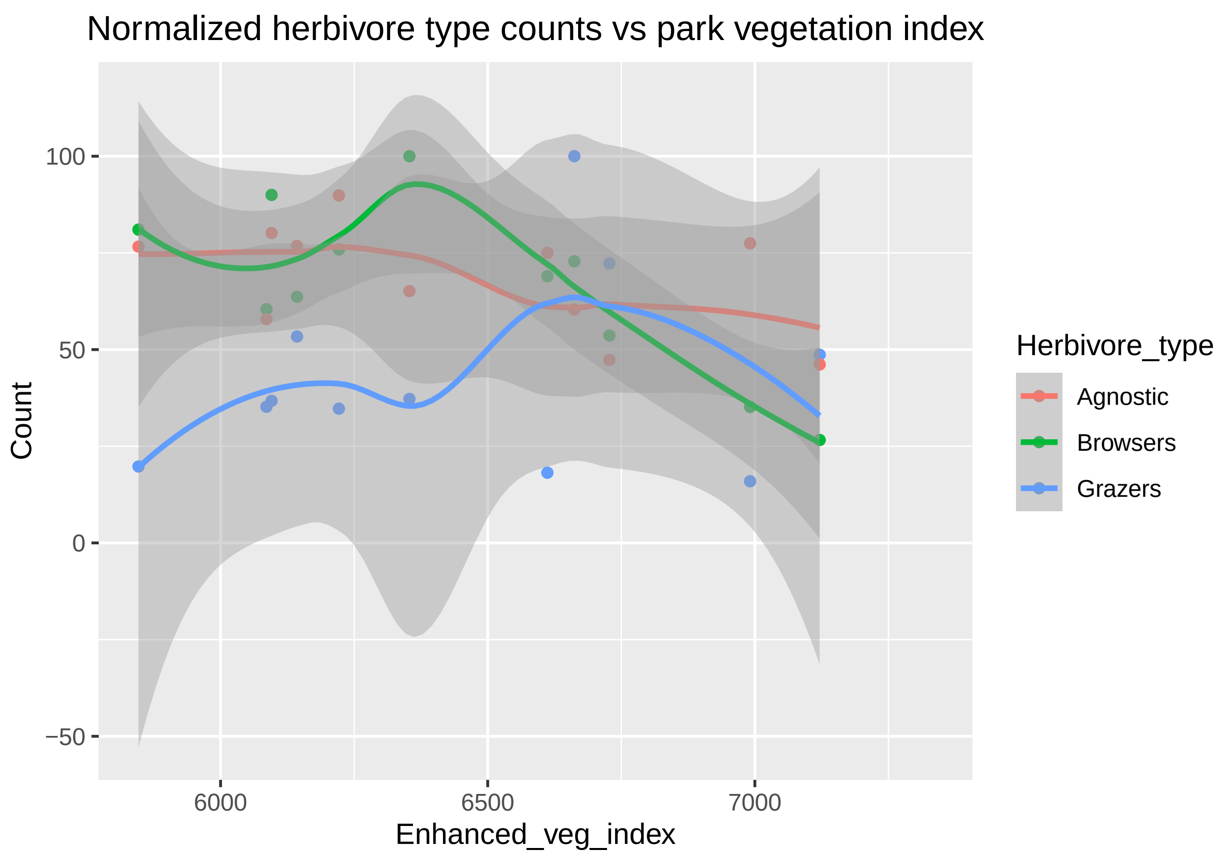



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="park" & Enhanced_veg_index > 0)) + geom_point(mapping=
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 15 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 15 rows containing missing values (geom_point).
```

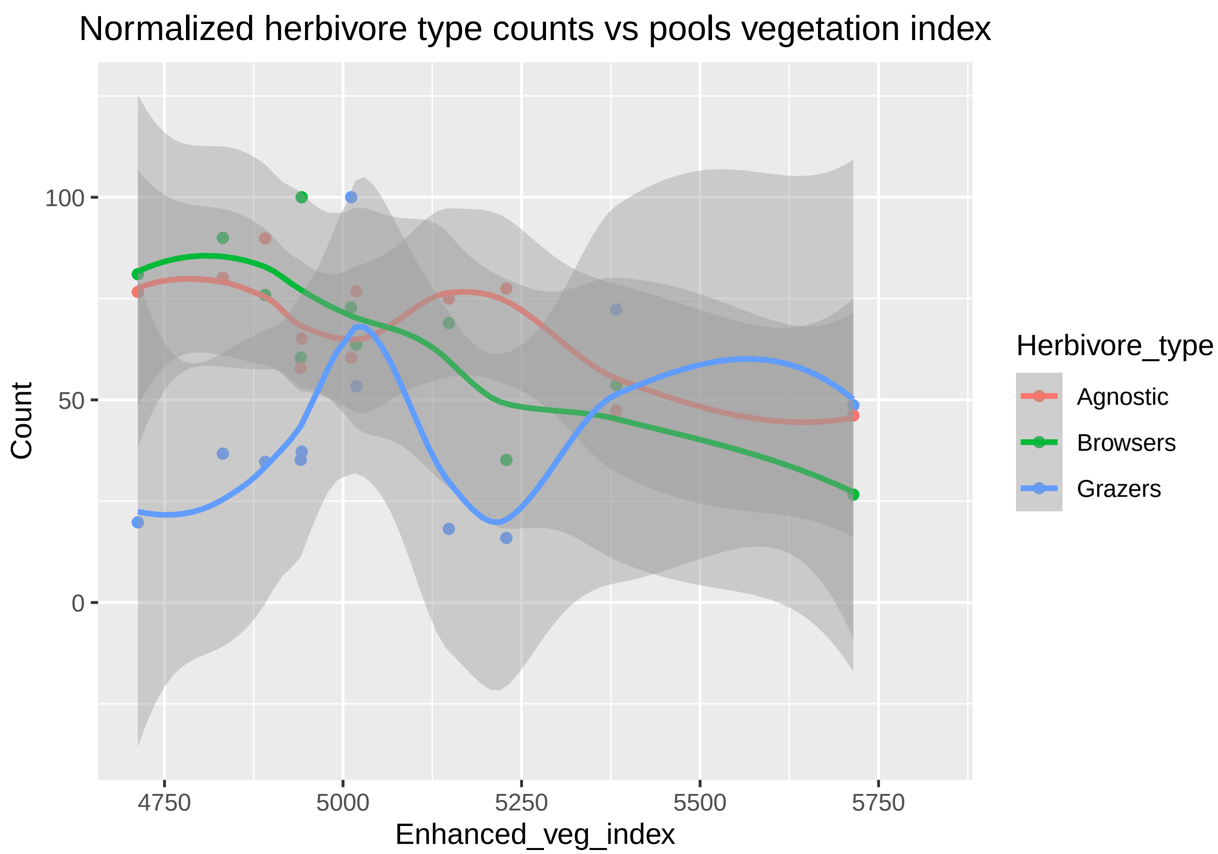



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="pools" & Enhanced_veg_index > 0)) + geom_point(mappin
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 15 rows containing non-finite values (stat_smooth).

## Warning: Removed 15 rows containing missing values (geom_point).
```

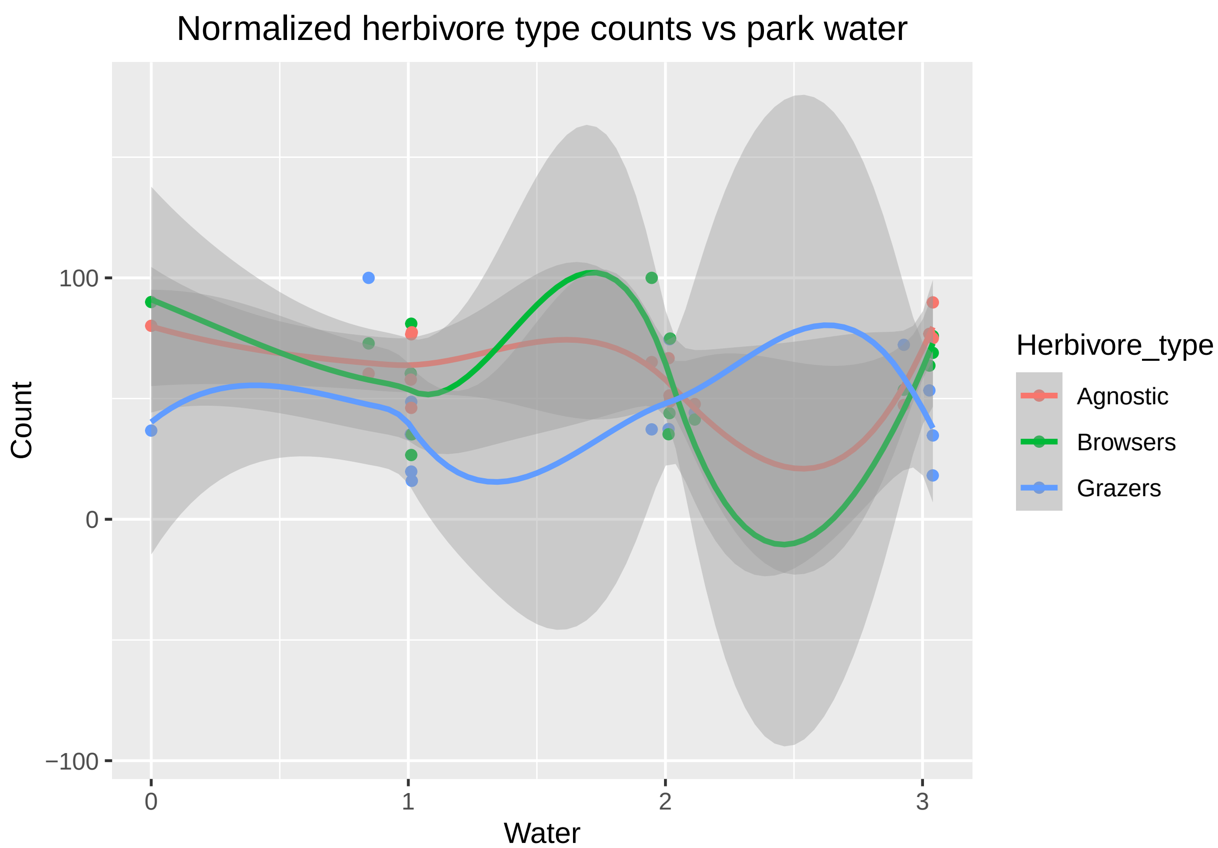



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="park")) + geom_point(mapping=aes(x=Water, y=Count, co
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 21 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 21 rows containing missing values (geom_point).
```

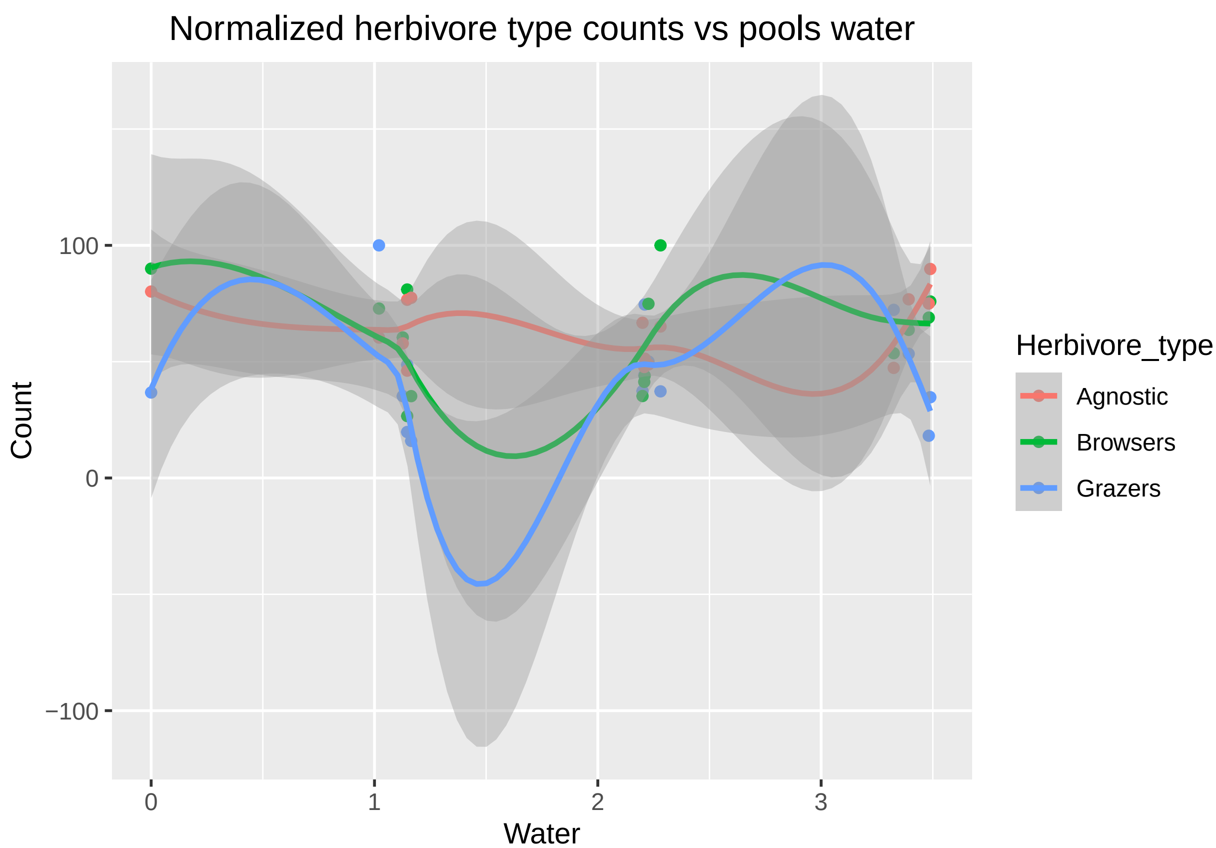



```
ggplot(data=filter(gc_clim_norm_type_plot, Area=="pools")) + geom_point(mapping=aes(x=Water, y=Count, c
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).

## Warning: Removed 21 rows containing missing values (geom_point).
```



Discussion

Previous wet season precipitation

The amount of precipitation during the previous wet season did not appear to be strongly correlated with game counts (individual species or herbivore type). This may be due to the large time period between the previous wet season and the game count, variation in climatic factors (such as temperature, wind and solar radiation) during this interim period may have confounded any relationship.

Soil moisture, drought index, vegetation and water

There did not appear to be any easily discernible correlation between drought index and water, and game counts. Soil moisture may have and a slight negative correlation with agnostic herbivores and browsers, but not for grazers. This correlation was present for both park and pools soil moisture. There was a strongly negative correlation between the enhanced vegetation index (pool and park) and browsers, and a less strong negative correlation between enhanced vegetation index (pool and park) and agnostic herbivores. This appears to indicate that extent of local migrations is more influenced by presence/absence of vegetation than by the presence/absence of water. Please note that although the extent of browsers local migration is more varied, this does not mean that browsers are more likely to migrate depending on water or vegetation. It may be the case that grazers always migrate depending on the season, despite what the climatic or vegetation conditions are, thus resulting in more uniform game counts from year to year. This study only looks at the variation of migratory patterns, not whether species/herbivore groups are more likely to exhibit season local migrations.