Project

Vicki M. Zhang

January 29, 2020

```
knitr::opts_chunk$set(echo = F, message = F, warning = F,
                      fig.width = 6, fig.height = 6)

library(rethinking)
```

```
## Loading required package: rstan
```

```
## Warning: package 'rstan' was built under R version 3.5.3
```

```
## Loading required package: StanHeaders
```

```
## Warning: package 'StanHeaders' was built under R version 3.5.3
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
## rstan (Version 2.19.2, GitRev: 2e1f913d3ca3)
```

```
## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```
## For improved execution time, we recommend calling
## Sys.setenv(LOCAL_CPPFLAGS = '-march=native')
## although this causes Stan to throw an error on a few processors.
```

```
## Loading required package: parallel
```

```
## Loading required package: dagitty
```

```
## Warning: package 'dagitty' was built under R version 3.5.3
```

```
## rethinking (Version 1.93)
```

```
##
## Attaching package: 'rethinking'
```

```
## The following object is masked from 'package:stats':
##
##     rstudent
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.5.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

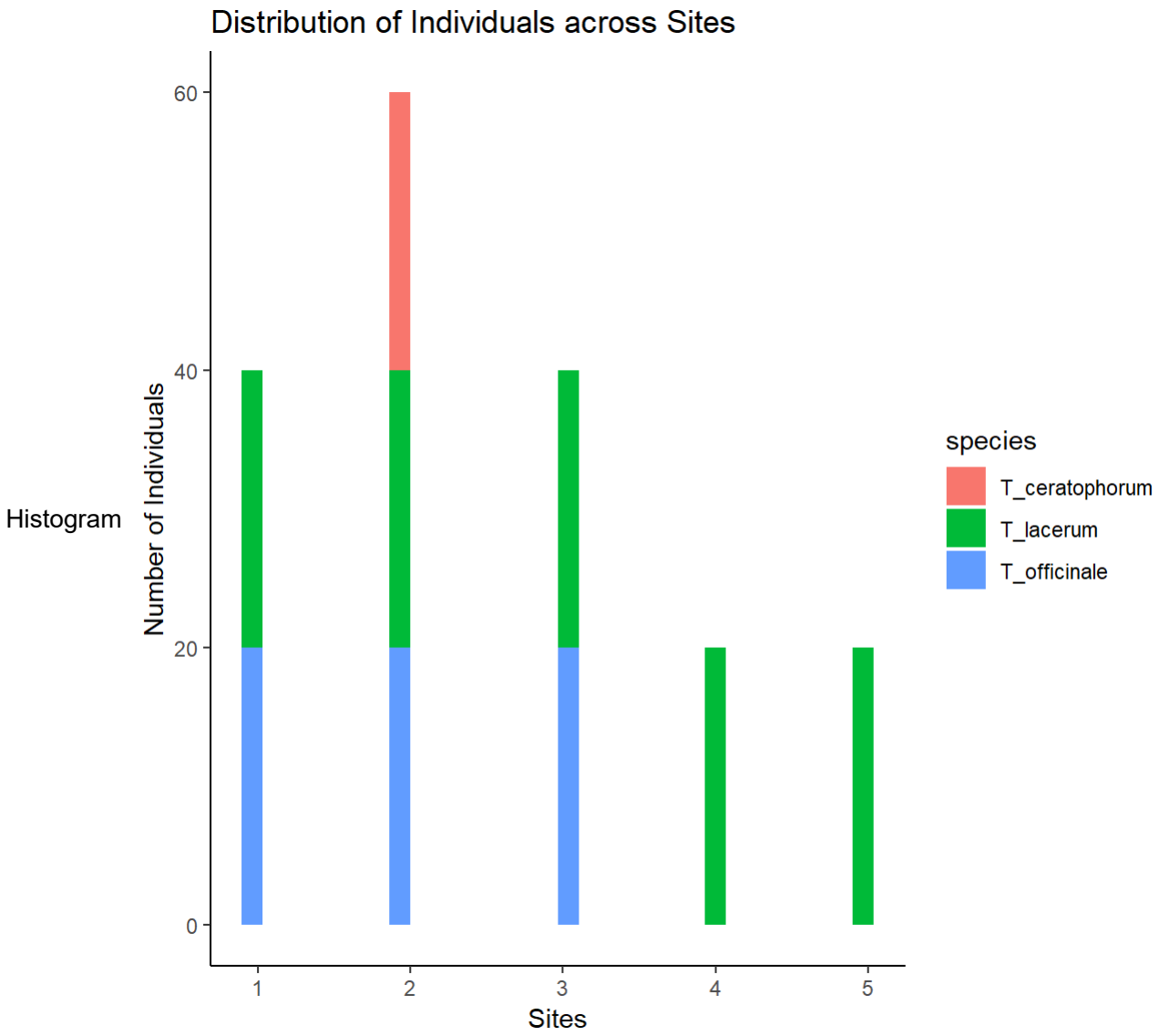
```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Visualizing all data

```
## 'data.frame':    180 obs. of  8 variables:
##  $ plant  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ date   : Factor w/ 3 levels "27-Aug-18","28-Aug-18",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ site   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ town   : Factor w/ 2 levels "N","Y": 2 2 2 2 2 2 2 2 2 2 ...
##  $ species: Factor w/ 3 levels "T_ceratophorum",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ leaf1  : int  10 1 0 0 0 0 0 0 0 0 ...
##  $ leaf2  : int  0 0 0 0 3 0 0 0 0 0 ...
##  $ leaf3  : int  2 1 0 0 20 0 0 0 0 0 ...
```

Description:

- 180 rows of data
- columns:
    1. “plant” = plant ID (1 to 180)
    2. “date” = date of data collection
    3. “site” = site of data collection (1-5)
    4. “town” = binary Y/N of whether plants were found in Churchill townsite
    5. “species” = three *Taraxacum* species (*T. officinale* is invasive, *T. lacerum* and *T. ceratophorum* are native)
    6. “leaf1”, “leaf2”, “leaf3” = measures of herbivory of 3 oldest leaves (i.e., largest, closest to ground)



Species-Level Differences of Herbivory

Number of individuals per site

```
## # A tibble: 9 x 4
## # Groups:   site, town [5]
##    site town  species            n
##   <int> <fct> <fct>          <int>
## 1     1 Y     T_lacerum         20
## 2     1 Y     T_officinale      20
## 3     2 Y     T_ceratophorum    20
## 4     2 Y     T_lacerum         20
## 5     2 Y     T_officinale      20
## 6     3 Y     T_lacerum         20
## 7     3 Y     T_officinale      20
## 8     4 N     T_lacerum         20
## 9     5 N     T_lacerum         20
```

Notes:

- sites 1-3 were in town, sites 4-5 were not in town
- data for all three species were only collected in site 2
    - only data for *T. ceratophorum* in site 2
- sites 1 and 3 have data for only *T. lacerum* and *T. officinale*
- sites 4 and 5 (i.e., out of town sites) have data only for *T. lacerum*

Number of individuals per species

```
## # A tibble: 3 x 2
##   species            n
##   <fct>          <int>
## 1 T_ceratophorum    20
## 2 T_lacerum        100
## 3 T_officinale      60
```

N =

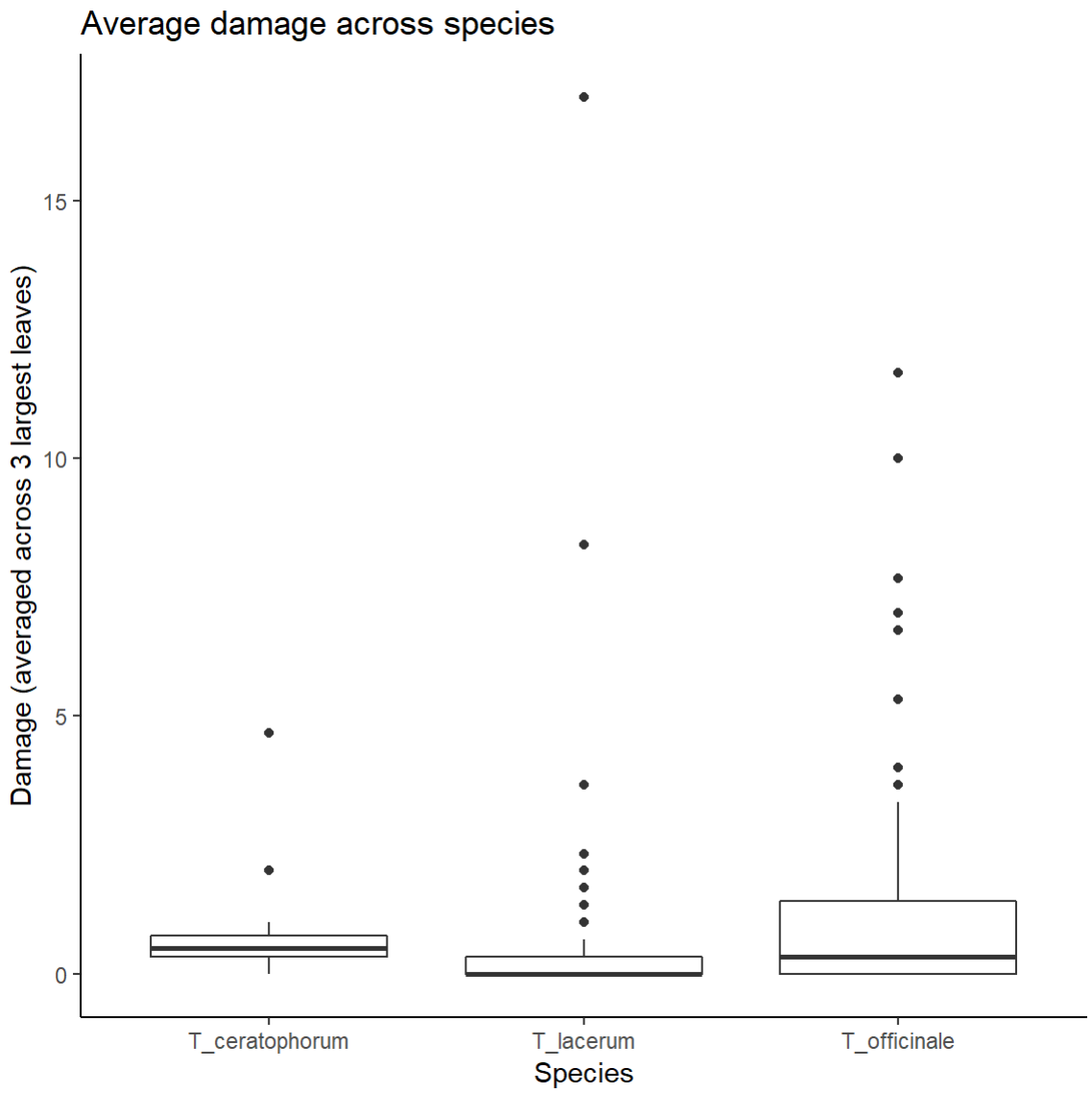
- *T. ceratophorum* = 20
- *T. lacerum* = 100
- *T. officinale* = 60

Averaged damage of all plants in three species

```
## # A tibble: 3 x 5
##   species        leaf1 leaf2 leaf3 average
##   <fct>          <dbl> <dbl> <dbl>   <dbl>
## 1 T_ceratophorum  0.8  1.05   0.45   0.767
## 2 T_lacerum       0.24 0.79   0.71   0.580
## 3 T_officinale    0.5  0.833  3.15   1.49
```

Table:

- “leaf1”, “leaf2” and “leaf3” columns are averaged damage percentages for each leaf
- “average” is the average of columns “leaf1”, “leaf2” and “leaf3”



Notes:

- *T. officinale* (the invasive species) has the greatest herbivory damage
- no general pattern over which leaf is most damaged by herbivory
    - dandelion leaves are in a floret near the base of the plant
    - since all the leaves are close to each other, the lack of difference makes biological sense

Number of plants per species with no damage

```
## # A tibble: 3 x 2
##   species            n
##   <fct>          <int>
## 1 T_ceratophorum     3
## 2 T_lacerum         61
## 3 T_officinale      24
```

Notes:

- *T. ceratophorum* = 20
    - 3 individuals with no damage
- *T. lacerum* = 100

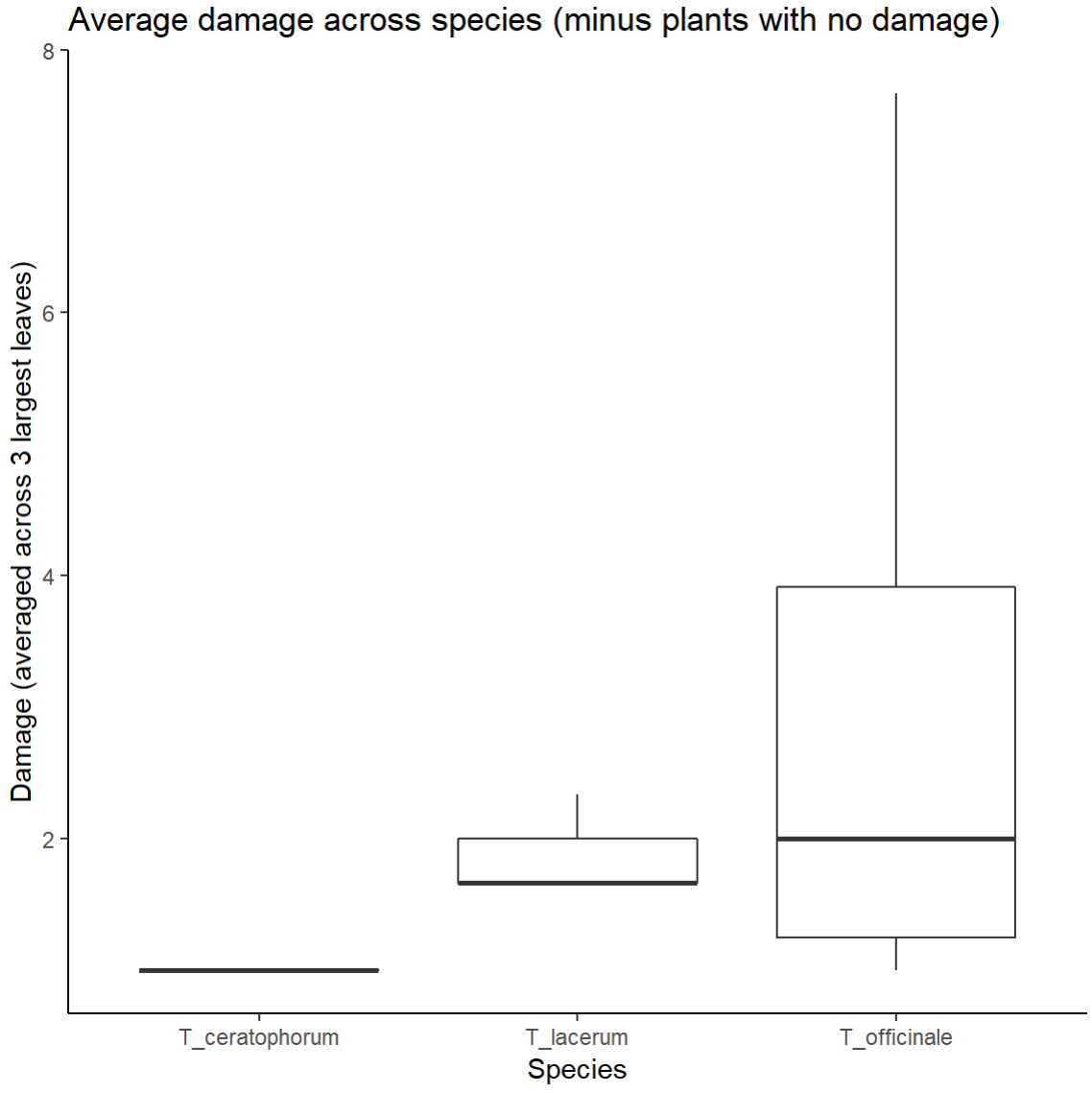
    - 61 individuals with no damage
- *T. officinale* = 60
    - 24 individuals with no damage

Averaged damage of all plants minus plants with no damage

```
## # A tibble: 3 x 5
##   species        leaf1 leaf2 leaf3 average
##   <fct>          <dbl> <dbl> <dbl>   <dbl>
## 1 T_ceratophorum     1  1     1       1
## 2 T_lacerum          3  1.67  1       1.89
## 3 T_officinale       2  5.75  1.75    3.17
```

Table:

- “leaf1”, “leaf2” and “leaf3” columns are averaged damage percentages for each leaf
- “average” is the average of columns “leaf1”, “leaf2” and “leaf3”
- this only includes plants with damage (filtered out plants with no damage)



Notes:

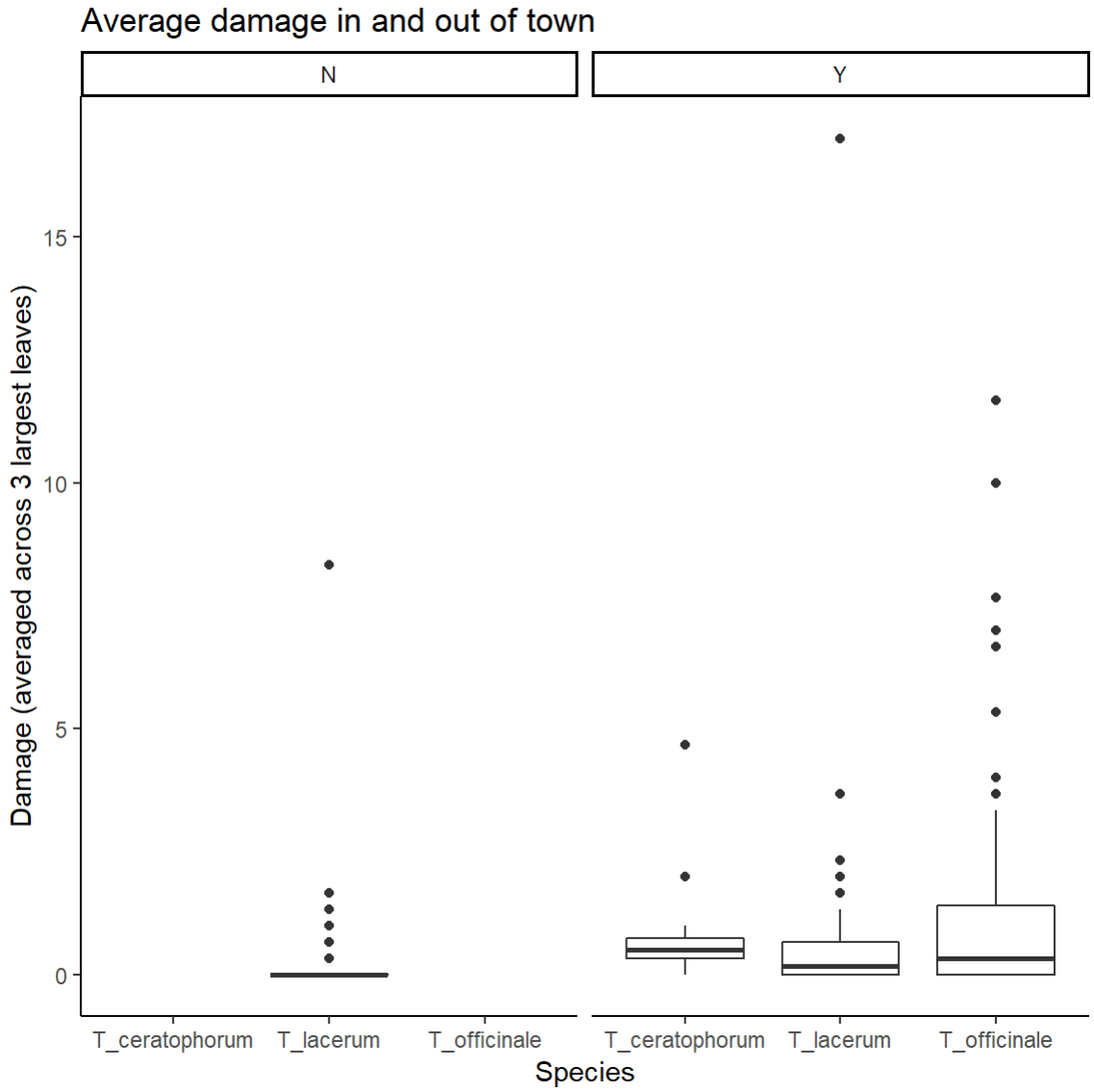
- *T. officinale* (the invasive species) again has the greatest herbivory damage

- no general pattern over which leaf is most damaged by herbivory

Differences in Herbivory between sites

Average damage in and out of town

```
## # A tibble: 4 x 6
## # Groups:   town [2]
##   town  species        leaf1 leaf2 leaf3 average
##   <fct> <fct>          <dbl> <dbl> <dbl>   <dbl>
## 1 N     T_lacerum      0.075 0.225 0.775   0.358
## 2 Y     T_ceratophorum 0.8   1.05  0.45    0.767
## 3 Y     T_lacerum      0.35  1.17  0.667   0.728
## 4 Y     T_officinale   0.5   0.833 3.15    1.49
```

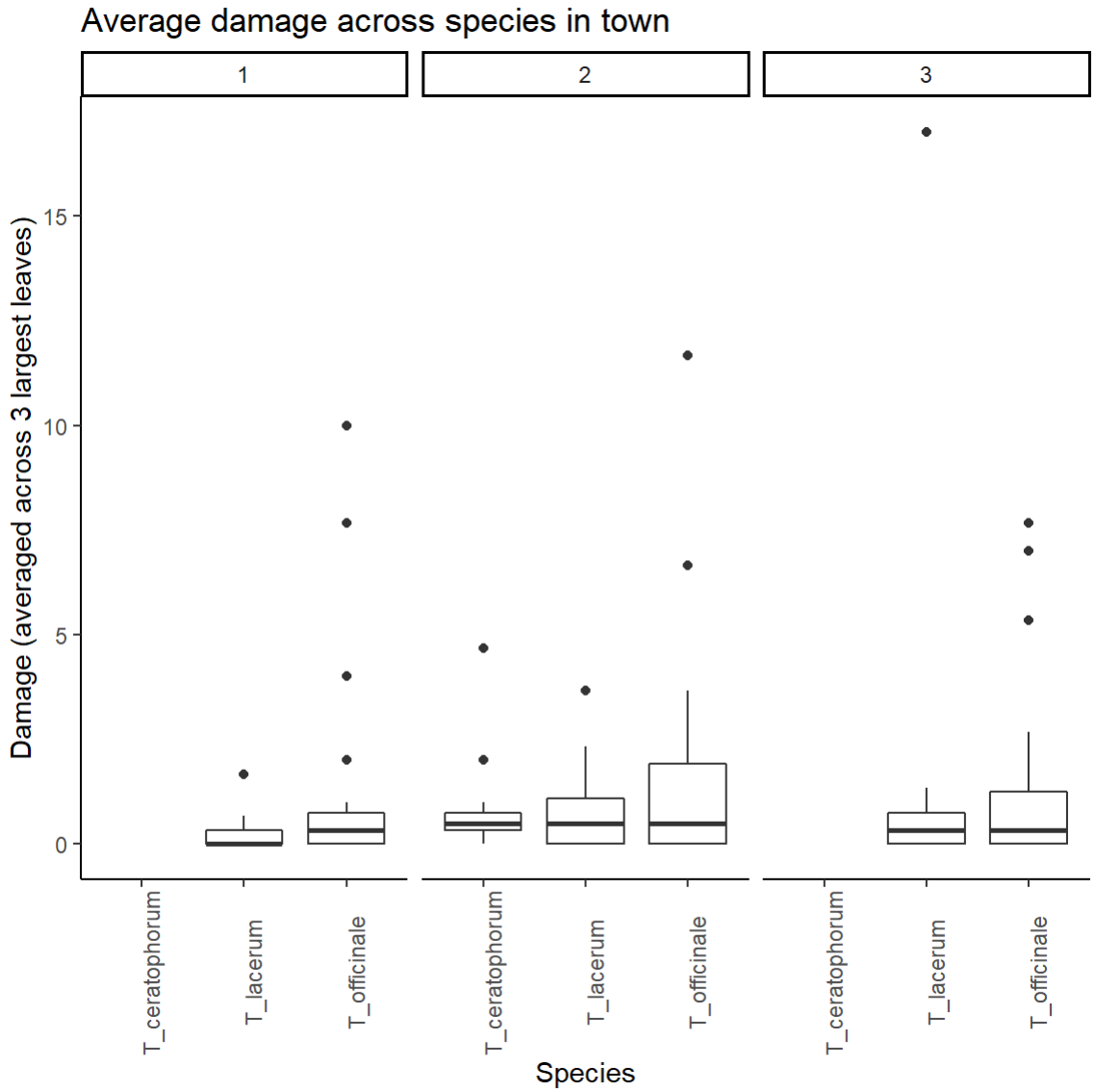


Notes:

- *T. lacerum* was the only species that had data both in and out of town
- more damage in town

Average damage across species in town

```
## # A tibble: 3 x 6
## # Groups:   town [1]
##   town  species        leaf1 leaf2 leaf3 average
##   <fct> <fct>          <dbl> <dbl> <dbl>   <dbl>
## 1 Y     T_ceratophorum  0.8  1.05  0.45    0.767
## 2 Y     T_lacerum       0.35 1.17  0.667   0.728
## 3 Y     T_officinale    0.5  0.833 3.15    1.49
```



Notes:

- *T. officinale* (invader) had most herbivory damage
- more damage in town

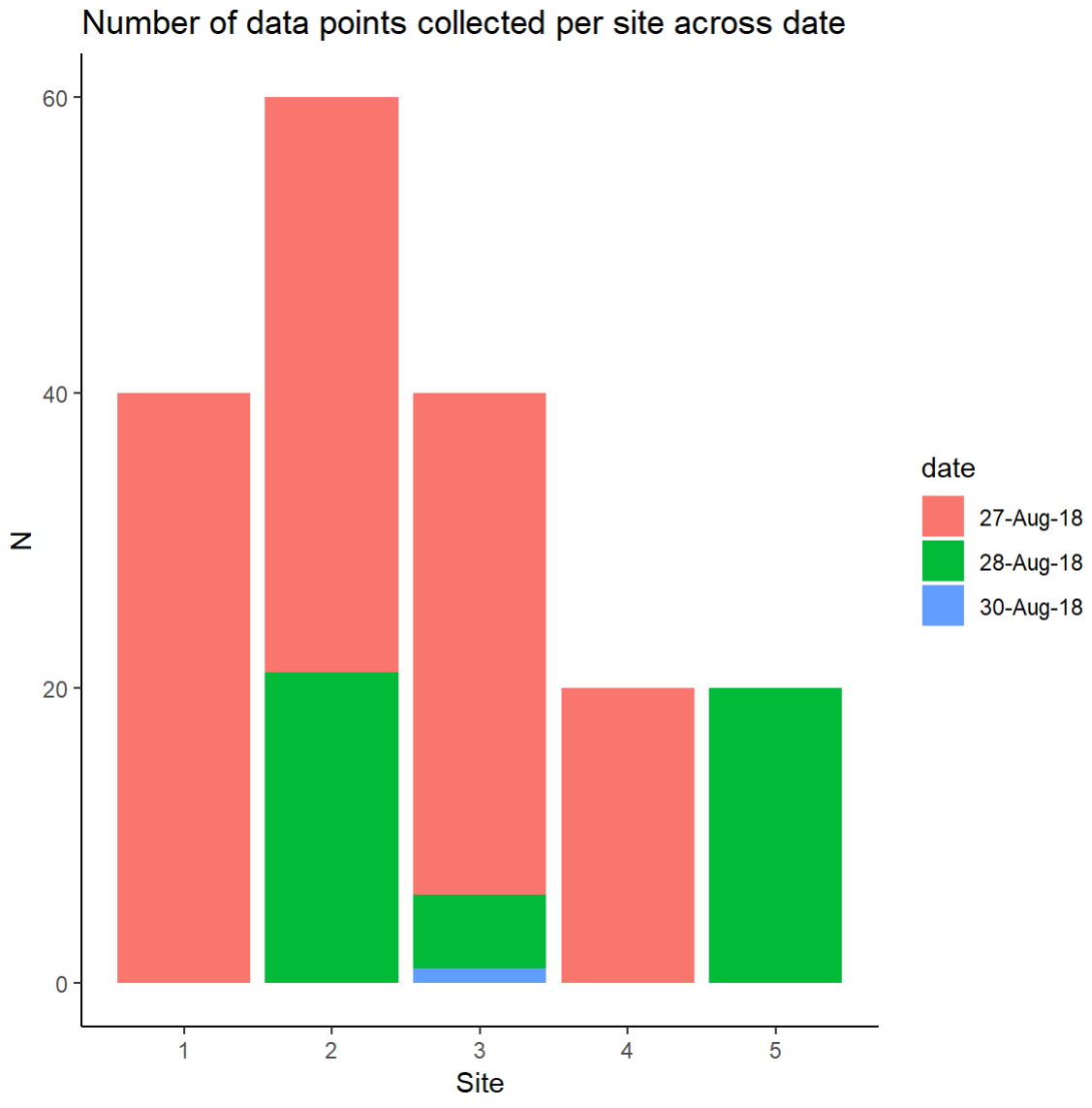
Effects of Date of Data Collection

Date of data collection across sites

```
## # A tibble: 8 x 3
## # Groups:   date [3]
##   date       site     n
##   <fct>     <int> <int>
## 1 27-Aug-18     1    40
## 2 27-Aug-18     2    39
## 3 27-Aug-18     3    34
## 4 27-Aug-18     4    20
## 5 28-Aug-18     2    21
## 6 28-Aug-18     3     5
## 7 28-Aug-18     5    20
## 8 30-Aug-18     3     1
```

Notes

- most of the data (N = 133) collected on August 27th
    - data collected from sites 1-4
- some data collected Aug 28 (N = 46) and Aug 30 (N = 1)
    - data collected on Aug 28th from sites 2, 3, 5
    - one data point collected Aug 30th from site 3

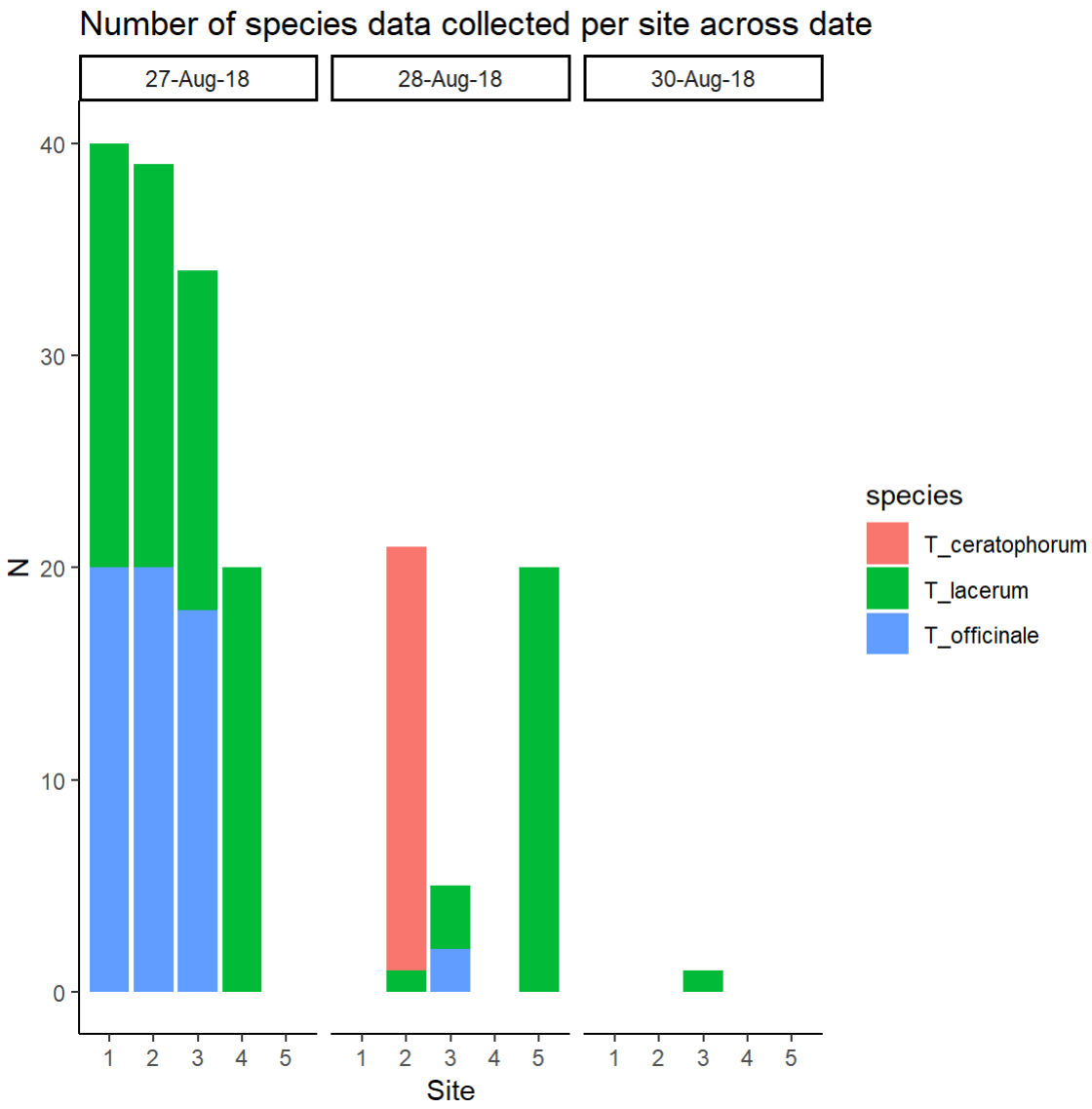


Date of data collection across species

```
## # A tibble: 6 x 3
## # Groups:   date [3]
##   date      species            n
##   <fct>     <fct>          <int>
## 1 27-Aug-18 T_lacerum         75
## 2 27-Aug-18 T_officinale      58
## 3 28-Aug-18 T_ceratophorum    20
## 4 28-Aug-18 T_lacerum         24
## 5 28-Aug-18 T_officinale       2
## 6 30-Aug-18 T_lacerum          1
```

Notes

- Aug 27th data (N = 133)
  - only *T. lacerum* and *T. officinale* data collected
- Aug 28th data (N = 46)
  - data from all three species collected
  - N (*T. ceratophorum*) = 20
  - N (*T. officinale*, the invader) = 2
  - N (*T. lacerum*) = 24
- Aug 29th only has one data point for *T. lacerum*



General thoughts

1. The variables that are affecting the system here are:

- date (Aug 27, 28, 30)
- site (1-5)
- in or out of town (Y/N)

2. I am looking at herbivory differences across the three Taraxacum species

- not sure if I should use all three herbivory measures, choose one, or average the three measurements into a new measure

3. I don't have data about the specific dates of data collection (like weather conditions, temperature, etc.). I might be able to get that data and can add it to the model. It may be important, or I can just generalize these differences into the "date" variable