

BIO 423 - Lab 2

Benjamin Blonder

Spring 2019

Learning outcomes

- Visualize data using the `ggplot2` package
- Aggregate and manipulate data frame subsets
- Estimate basic demographic parameters from census data
- Explore hypotheses for the drivers of demography

Annual plant population dynamics

We will explore long-term plant demography data collected for annual plants in the Sonoran Desert by D. L. Venable (University of Arizona). These plants germinate from seed after winter rains, then quickly grow, potentially set seed, and then die back within a few short months. For several decades, Venable has collected records of germination, survivorship, and fecundity of multiple species co-occurring in the same set of long-term plots.

A total of 72 plots are located at 725 m elevation on Tumamoc Hill near Tucson, Arizona. These data represent averages across all plots, summarized for each year. Plots are 0.1 m² in area.

Raw data come from <http://www.eebweb.arizona.edu/faculty/venable/LTREB/metadata.htm> (see PDF file in this folder for details).

The data contain columns for species (`species`), year (`year`), number of seedlings that germinated per plot per year (`germ`), number of seedlings that survived until reproductive age (`survive`), and per-capita seed production of reproductive-age plants (`bx`). Species are coded as the first two letters of the genus, then first two letters of the specific epithet (e.g. `erci` for *Erodium cicutarium*).

Loading in the data

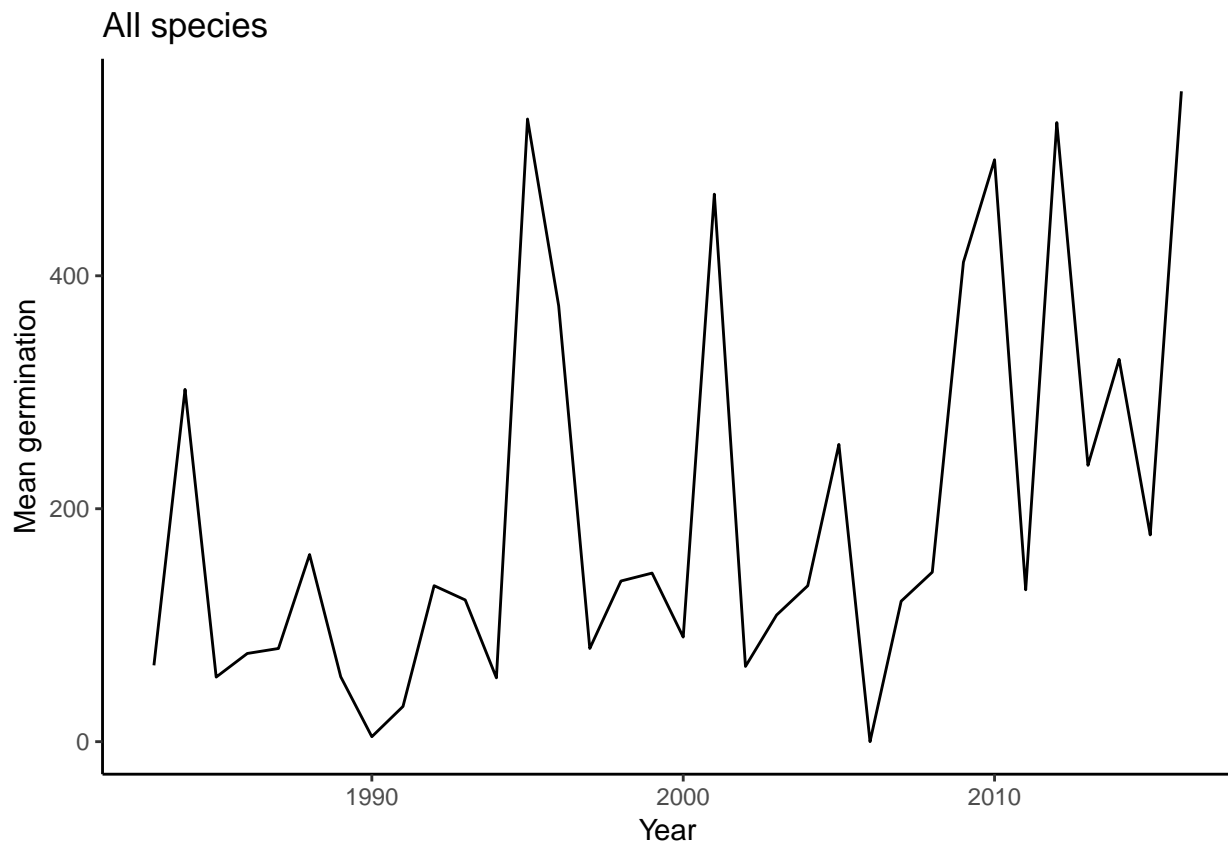
```
# load the ggplot2 package (if the below line does not run,  
# you need to install the package via Tools > Install Packages in Rstudio)  
library(ggplot2)  
  
# read in the data - we need the na.strings argument because  
# missing values are coded as dots in the CSV file  
data_venable = read.csv("Venable LTREB - SpeciesByYear.csv", na.strings=".")  
# select appropriate columns  
data_venable = data_venable[,c("year", "species", "germ", "survive", "bx")]  
# inspect the data  
str(data_venable)  
  
## 'data.frame':   510 obs. of  5 variables:  
##  $ year   : int   1983 1983 1983 1983 1983 1983 1983 1983 1983 1983 ...  
##  $ species: Factor w/ 15 levels "dapu","drcu",...: 1 2 3 4 5 6 7 8 9 10 ...  
##  $ germ   : int   21 0 50 49 14 0 98 0 91 2 ...  
##  $ survive: int   14 0 27 41 5 0 87 0 15 0 ...
```

```
## $ bx      : num  4 NA 10.4 71.2 23 ...

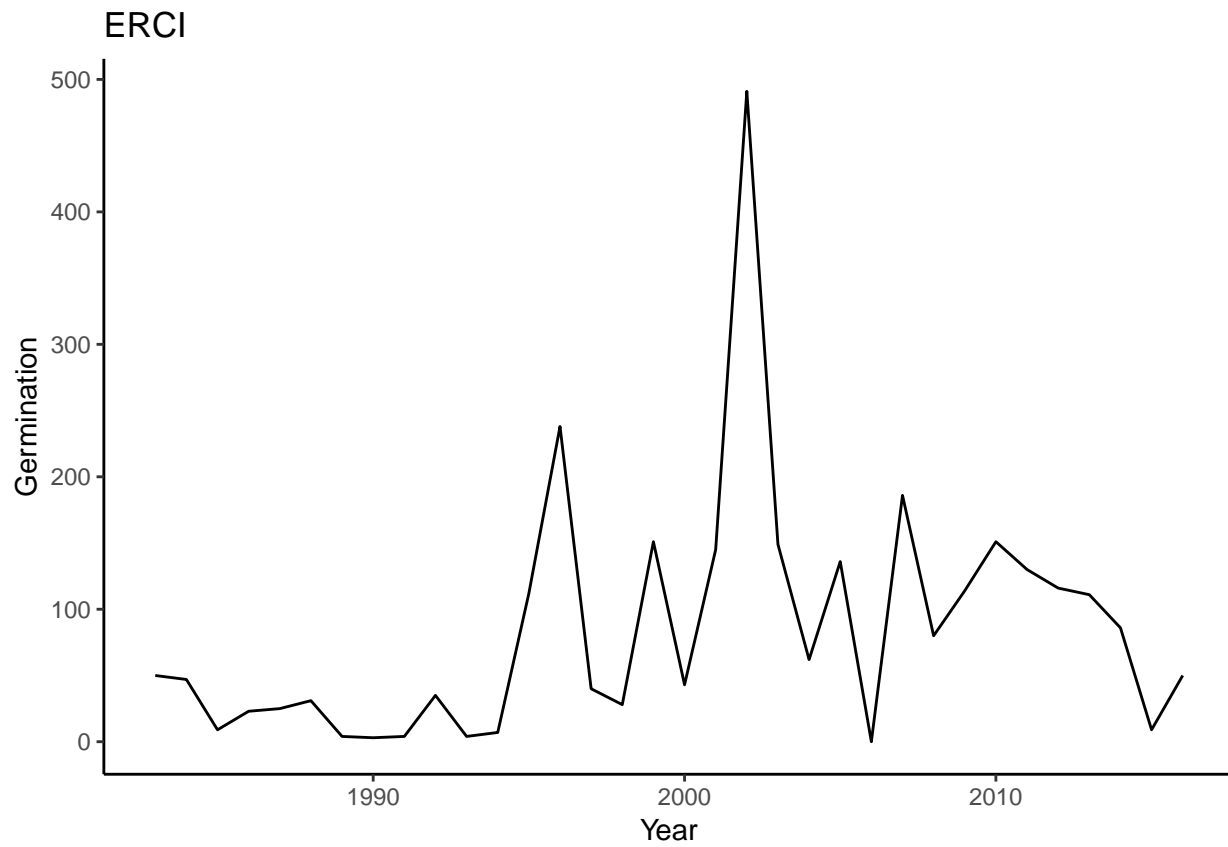
# calculate germination per year across all species using the aggregate function
germ_by_year = aggregate(germ~year,data=data_venable,mean)
# inspect the output
str(germ_by_year)

## 'data.frame':  34 obs. of  2 variables:
## $ year: int  1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 ...
## $ germ: num  65.5 302.5 55.5 75.7 80 ...

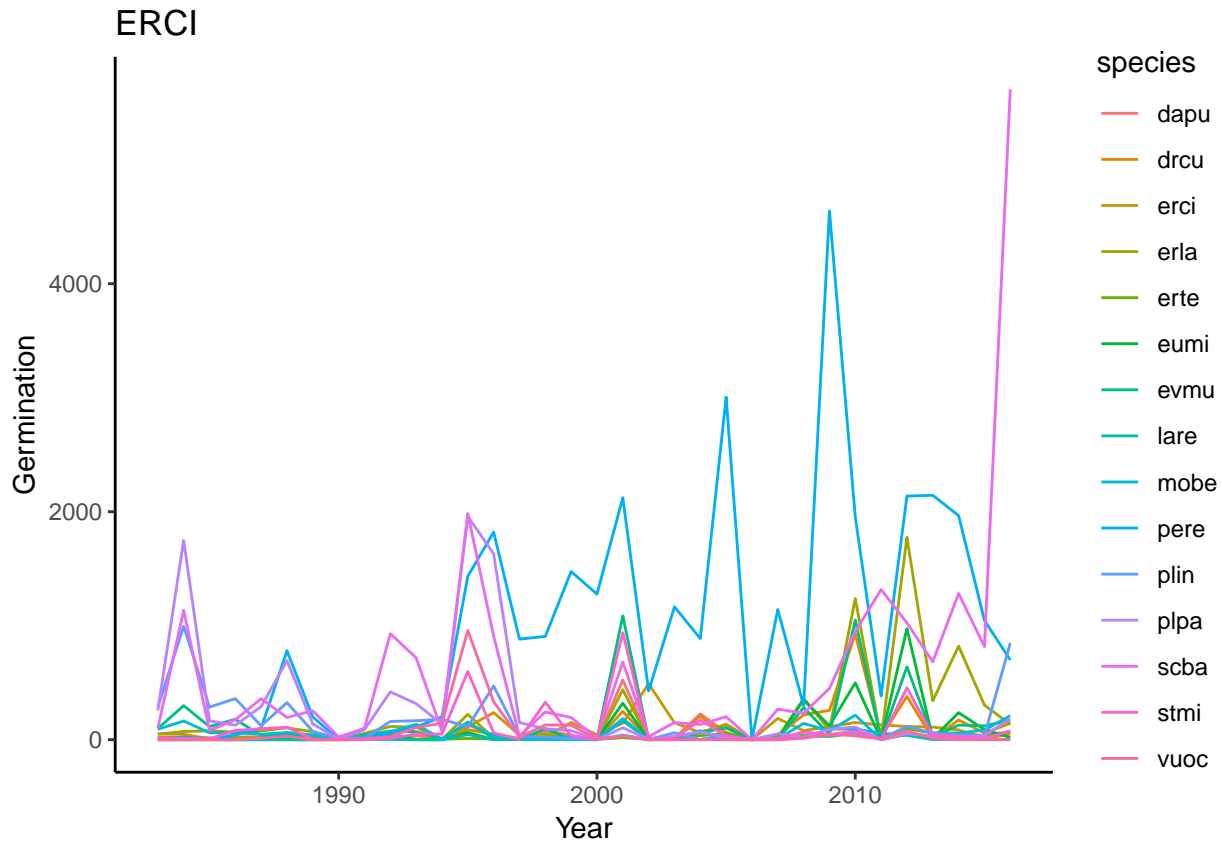
# plot with ggplot (this is a new function)
ggplot(data=germ_by_year,aes(x=year,y=germ)) +
  geom_line() + theme_classic() + ggtitle("All species") +
  xlab("Year") + ylab("Mean germination")
```



```
# subset the data for a single species
data_venable_erci = data_venable[data_venable$species=="erci",]
# print data and then plot
ggplot(data=data_venable_erci,aes(x=year,y=germ)) +
  geom_line() + theme_classic() + ggtitle("ERCI") +
  xlab("Year") + ylab("Germination")
```



```
# plot multiple species  
ggplot(data=data_venable,aes(x=year,y=germ,col=species)) +  
  geom_line() + theme_classic() + ggtitle("ERCI") +  
  xlab("Year") + ylab("Germination")
```



Questions

1. What year has the highest number of adult survivals for species *Schismus barbatus* (**scba**)? How about *Plantago patagonica* (**plpa**)?
2. Several species appear to have correlated increases and decreases in germination and survival. Propose two hypotheses for why these species have similar population dynamics.
3. Calculate the survivorship fraction (**lx**) for each species and year. Remember that survivorship is equal to the number of survivors divided by the number of germinants. You will need to make a new column in the data frame (hint: `data_venable$lx = ...`). Plot survivorship over time for *Pectocarya recurvata* (**pere**).
4. Calculate the reproductive value (**lxbx**) for each species and year. Remember that mean reproductive value is the product of survivorship and per capita seed production of reproductive-age plants. Plot reproductive value over time for *Pectocarya recurvata* (**pere**). Propose two hypotheses for why reproductive value may change over time.
5. Which two species have the highest mean reproductive value across all years? Are either of these species non-native invaders? (Hint - you can look up species names in **Venable LTREB - Species_list_2016_2.csv** and then check Google for native ranges of these species)

Optional question for graduate students

Consider loading in the National Weather Service monthly precipitation data for the Tucson airport (**nws_twc_monthly_precip.csv**). To what extent does winter precipitation predict the demographic responses of species? You will need to convert the T values ('trace') to 0s, then join the two datasets together (hint, use **join** in the **plyr** package), then run a linear mixed model.

What to hand in

- A single Word Document including:
 - written answers (1-2 sentences) and figures for each question above
 - A copy of your R script (the contents of your `.R` file pasted into the Word document)
 - Author contribution statement