Prior to acquiring data, a field survey was conducted to find corn fields that were damaged by elves or other stress factors. We measured and recorded geographic locations of each selected field, and then determined the predominant cause of crop stress in each field by traversing the field on foot and recording whether the field was heavily infested by elves or not. Fields selected for data collection were ones in which elves were the predominant source of stress to the crop, and other fields that exhibited little if any evidence of infestation.

In total, 60 corn fields were selected for data collection, with numbers of fields being approximately equally divided between those infested and fields that were not infested but exhibited evidence of stress caused by other factors.

A set of 22 covariates were computed for each of the 60 fields. Fourteen covariates showed significant differences between elf-stressed fields and other stress (e.g., distance to nearest forest: p-value < 0.0001, f-stat = 843.2, df = 58). Covariates were used to quantify the difference between fields infested with elves and those that were stressed by other factors.

Today

- Make sure we are talking in the same language about: Data
- The confusing lexicon (Part 1)
- Concept and logic maps
- More R fun!

The fine line:

Boring   Challenging   Overwhelming

Housekeeping

- Returning assignments
- E-mails and contact information

Qualitative data

- Relative or perspective-based
  - Fast vs slow
  - Tastes good
    - Does beer taste good?
      Hint: The answer is YES!
- Describing a quality
  - Red, blue, green
  - Old
  - Rough texture

E.g., outside: the trees are still green but the air is crisp and smells of fall

Quantitative data

- Measured
  - 10 km/day vs 3 km/day
  - Emitting light in a range between 620–750 nm (red)
- Quantities / numbers

E.g., outside: ~ 13 hours of daylight, the night time minimum temperature was 55° F, and 4 grass species are seeding.

These are the type of data that we are going to concentrate on in this class.

Qualitative vs. Quantitative

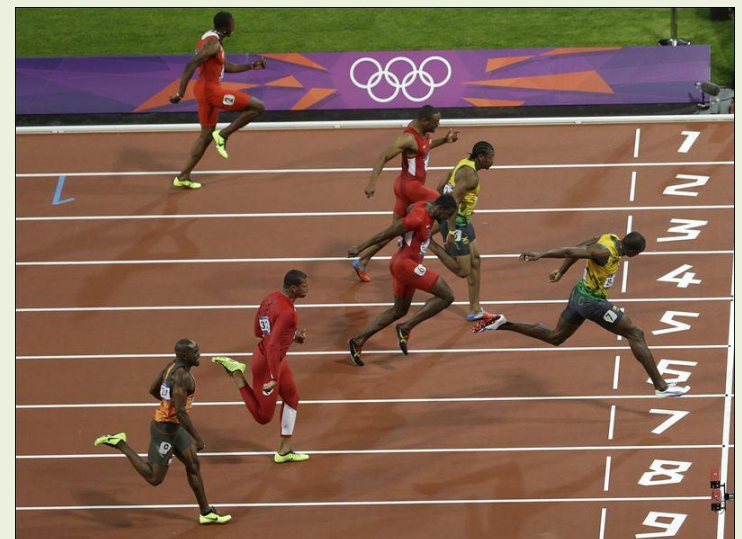It can be dangerous when they are mixed!
Why?

Repeatability.

Fast sprinter vs Slow sprinter

- Hours a week training with a coach
- Caloric intake
- Train on Sundays?
- Train in the afternoon?

- Number of practice starts per week
- Weekly hours in isometric training
- Weekly hours in dynamic training
- Length of hair

Tricky elves!

Prior to acquiring data, a field survey was conducted to find corn fields that were damaged by elves or other stress factors. We measured and recorded geographic locations of each selected field, and then determined the predominant cause of crop stress in each field by traversing the field on foot and recording whether the field was heavily infested by elves or not. Fields selected for data collection were ones in which elves were the predominant source of stress to the crop, and other fields that exhibited little if any evidence of infestation.

In total, 60 corn fields were selected for data collection, with numbers of fields being approximately equally divided between those infested and fields that were not infested but exhibited evidence of stress caused by other factors.

A set of 22 covariates were computed for each of the 60 fields. Fourteen covariates showed significant differences between elf-stressed fields and other stress (e.g., distance to nearest forest: p-value < 0.0001, f-stat = 843.2, df = 58). Covariates were used to quantify the difference between fields infested with elves and those that were stressed by other factors.

Examples:

Alternate plant hosts availability will be difficult to quantify. However, we intend to rate availability based on a 0-2 scale with "0" defined as no alternate plant hosts immediately adjacent to sampled crops, "1" defined as limited alternate host plant availability (such as uncontrolled roadside vegetation), and "2" defined as abundant alternative crop hosts adjacent to the sampled crop (e.g., wheat planted into the corners of a pivot-irrigated corn field).
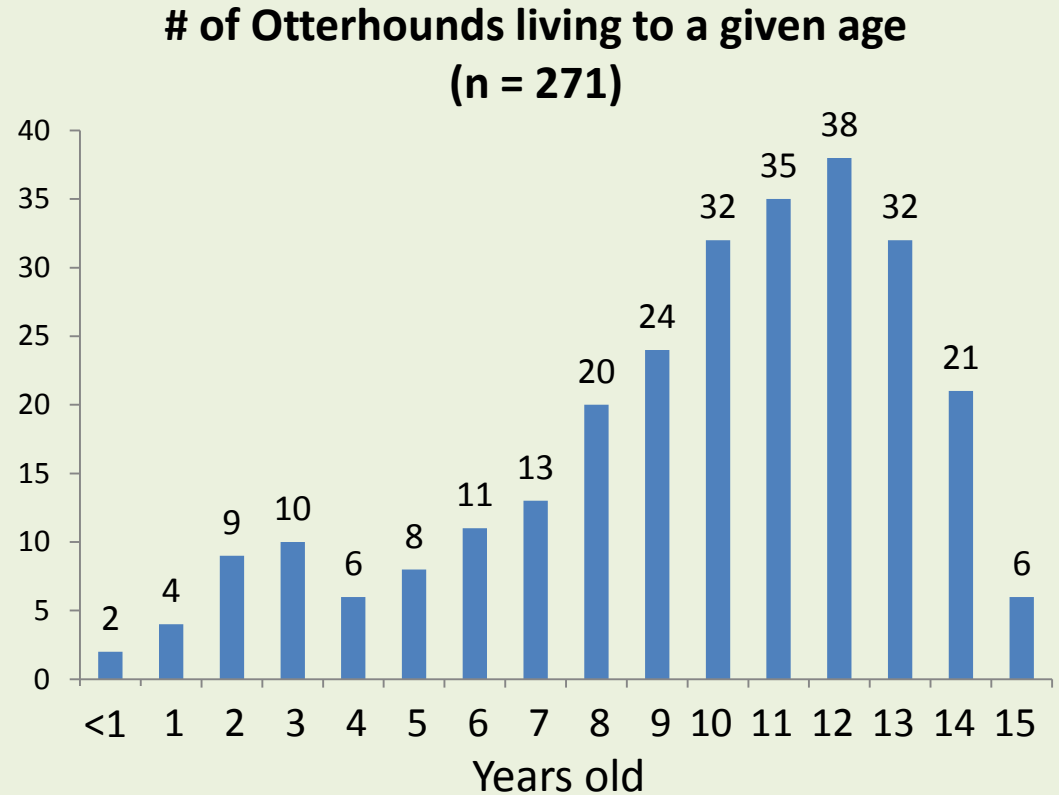
Leaf damage was rated on a scale from 0-10.

-or-

Leaf damage was rated on a scale from 0-10 by measuring the percentage of leaf area affected by hypersensitive cell death.

Quantitative lexicon

Histogram:

- Device for displaying data
- Data are separated into bins

# of Otterhounds living to a given age (n = 271)

Source of confusion of the day (SCD): Average!

Mean	9.63 yrs
Median	10 yrs
Mode	13 yrs

Quantitative lexicon

One of the fun dangers of histograms
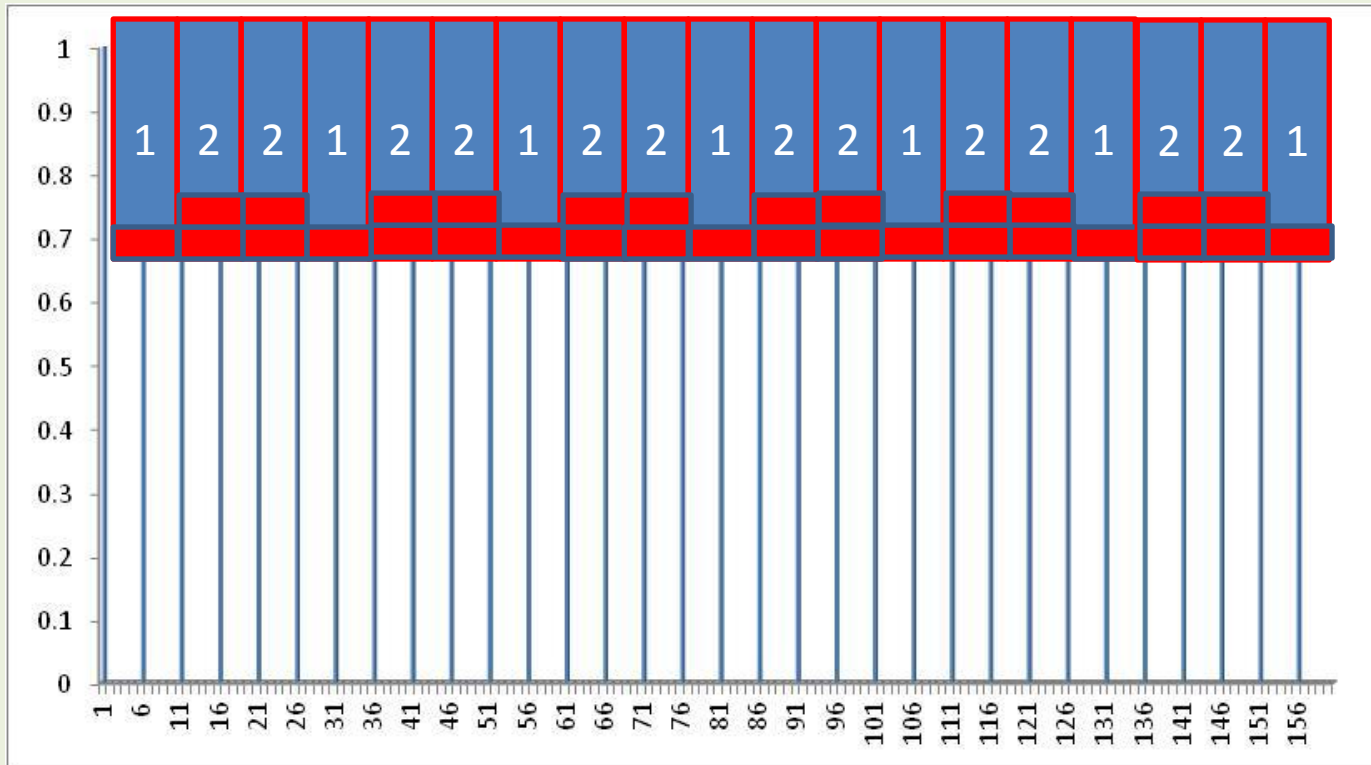
Aliasing!

Frequently observed in image processing
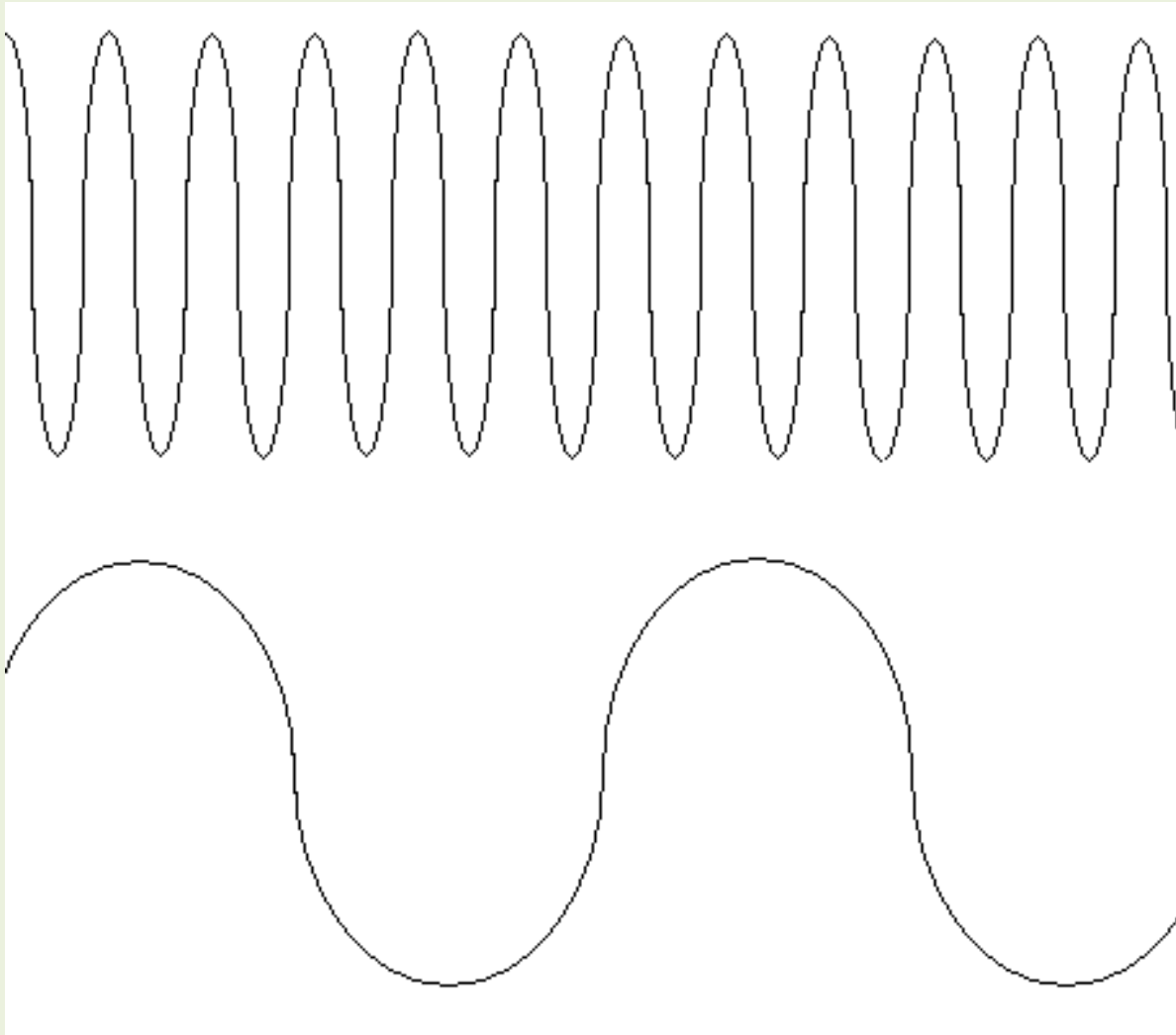
of Otterhounds living to a given age
(n = 271)
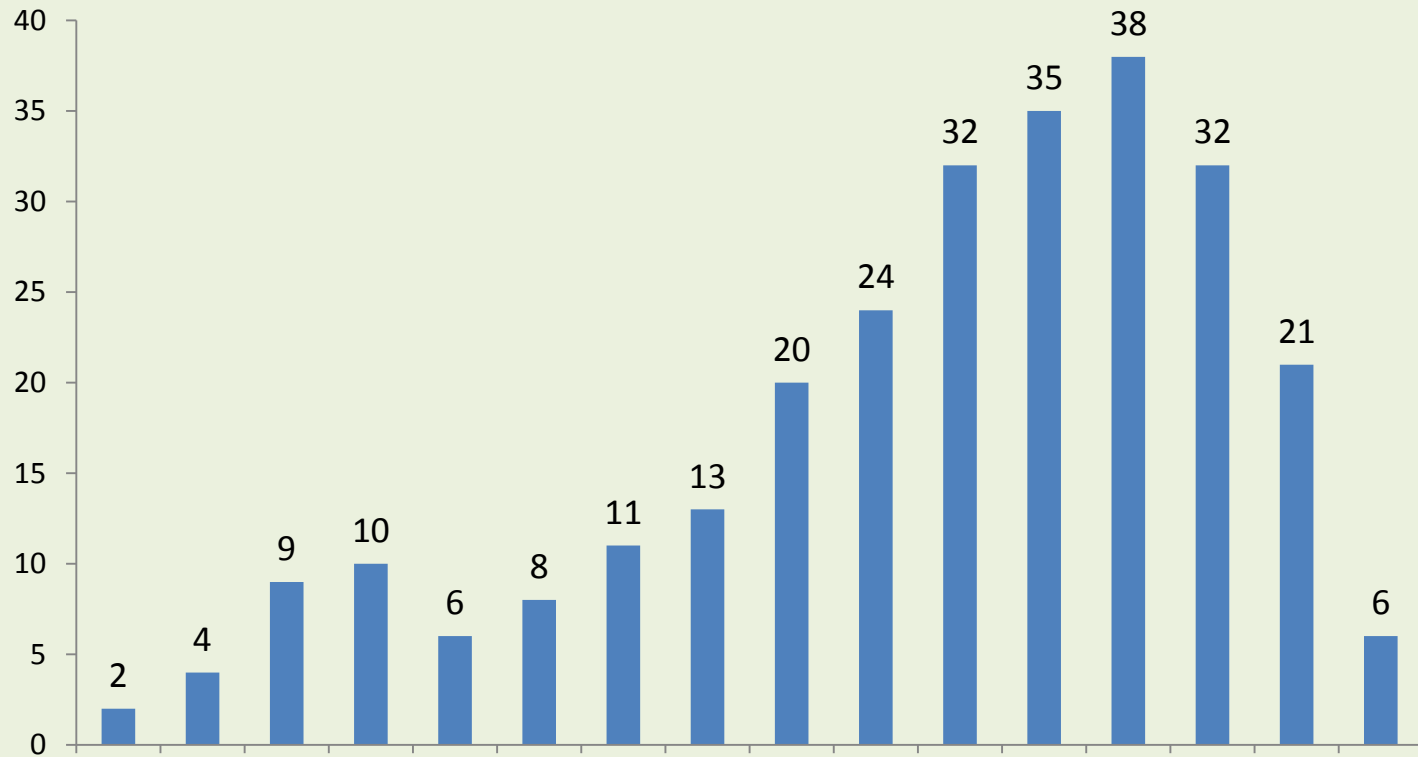
Data separated into age bins / categories

Metadata

- Metadata: data about data
- Meta-analysis: an analysis of meta-data
  - Most frequently a analysis of the results from many studies
- Problem: which studies get published?
  - E.g., the effect of climate change on corn production
    - Do studies that find no effect get published?

Concept models

- Focal element (question, species, concept, ecosystem, etc.)
- Major components or drivers
- What influences those components
- How are these components/factors/drivers connected?

Temperature

Predators and Parasitoids

Heat units

Precipitation

Aestivation period

Environmental factors

Host quality

Planting date

Tillage system

Plant hosts / cropping system

Sunflower stem weevil emergence & ovipositional success

Logic Models

System or situation ➡ **Inputs** ➡ **Results** (Outputs) ➡ **Impacts**
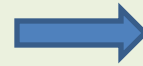
Factors that could influence the system

Ecosystem (Crop)	Ecological factors & management decisions	Change in Sunflower stem weevil emergence & success	Yield Loss
Sunflower stem weevils as a pest of sunflower	Heat units		
	Precipitation		
	Aestivation period		
	Planting date		
	Tillage		
	Predators		

Assignment # 2

- On courses tab:

- http://www.uvm.edu/~scmerril/Courses.html

R fun!

- Questions from last week?
- http://www.uvm.edu/~scmerril/Chapter%202%20Learning%20to%20code%20in%20R.pdf