Data visualization with ggplot

Based on R-ecology lesson - Data carpentry

Marco Chiapello, PhD

March 29, 2017

Importance of data visualization

Ascomb's quartet

Use the file example "scraping.R" in Resource directory

WHAT IS GGPLOT2

- ggplot2 is a plotting system for R, based on the **grammar of graphics**
- ggplot2 makes **simple to create complex plots** from data in a dataframe
- help creating **publication quality plots** with a **minimal amount of settings** and tweaking
- ggplot graphics are built step by step by **adding new elements**

Plotting with ggplot2

To build a ggplot we need to:

- **bind the plot** to a specific data frame using the data argument

```
surveys_complete <- read.csv("surveys_complete.csv")
ggplot(data = surveys_complete)
```

Plotting with ggplot2

To build a ggplot we need to:

- **bind the plot** to a specific data frame using the data argument

```
surveys_complete <- read.csv("surveys_complete.csv")
ggplot(data = surveys_complete)
```

- define **aesthetics** (aes), by selecting the variables to be plotted

```
ggplot(data = surveys_complete,
       aes(x = weight, y = hindfoot_length))
```

Plotting with ggplot2

- add **geoms** – graphical representation of the data in the plot

```
ggplot(data = surveys_complete,
    aes(x = weight, y = hindfoot_length)) +
  geom_point()
```

Plotting with ggplot2

- add **geoms** – graphical representation of the data in the plot

```
ggplot(data = surveys_complete,
    aes(x = weight, y = hindfoot_length)) +
  geom_point()
```

- The **+** in the ggplot2 package is particularly useful because it allows you to modify existing ggplot objects

PLOTTING WITH GGPLOT2

- Set up plot "templates" and conveniently explore different types of plots

```
# Create
surveys_plot <- ggplot(data = surveys_complete,
                       aes(x = weight, y = hindfoot_length))

# Draw the plot
surveys_plot + geom_point()
surveys_plot + geom_line()
surveys_plot + geom_count()
surveys_plot + geom_rug()
```