A PETER WALL INTERNATIONAL RESEARCH ROUNDTABLE

Building a Bio-analytical Theory for the Analysis of Marine Mammal Movements

May 4-8, 2015

Using the Grammar of Graphics and Interactivity to explore Biologging Data in R

Vincenzo Coia
Department of Statistics
The University of British Columbia

Tiphaine Jeanniard Du Dot
Department of Zoology
The University of British Columbia

May 6, 2015

Follow Along

- See this demo’s GitHub repository to follow along with this demonstration:

https://github.com/vincenzocoia/sea-mammal-gg

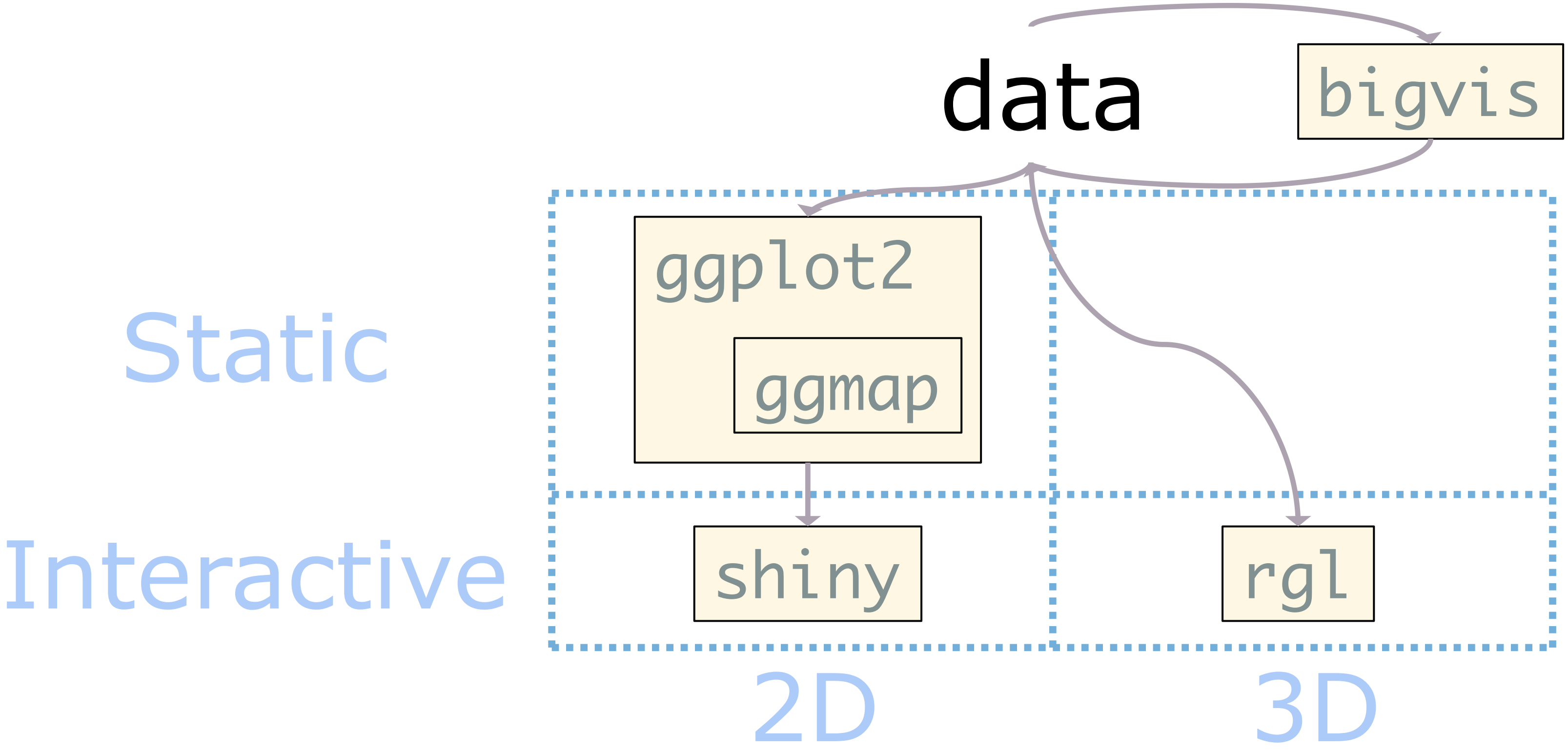
Goals

All simultaneously.

1. Conceptual:
   - How to display data w/ lots of variables (Grammar + Interaction)
2. Software demo in R
   - (Won’t focus on syntax much)
3. Exploratory Analysis of biologging data

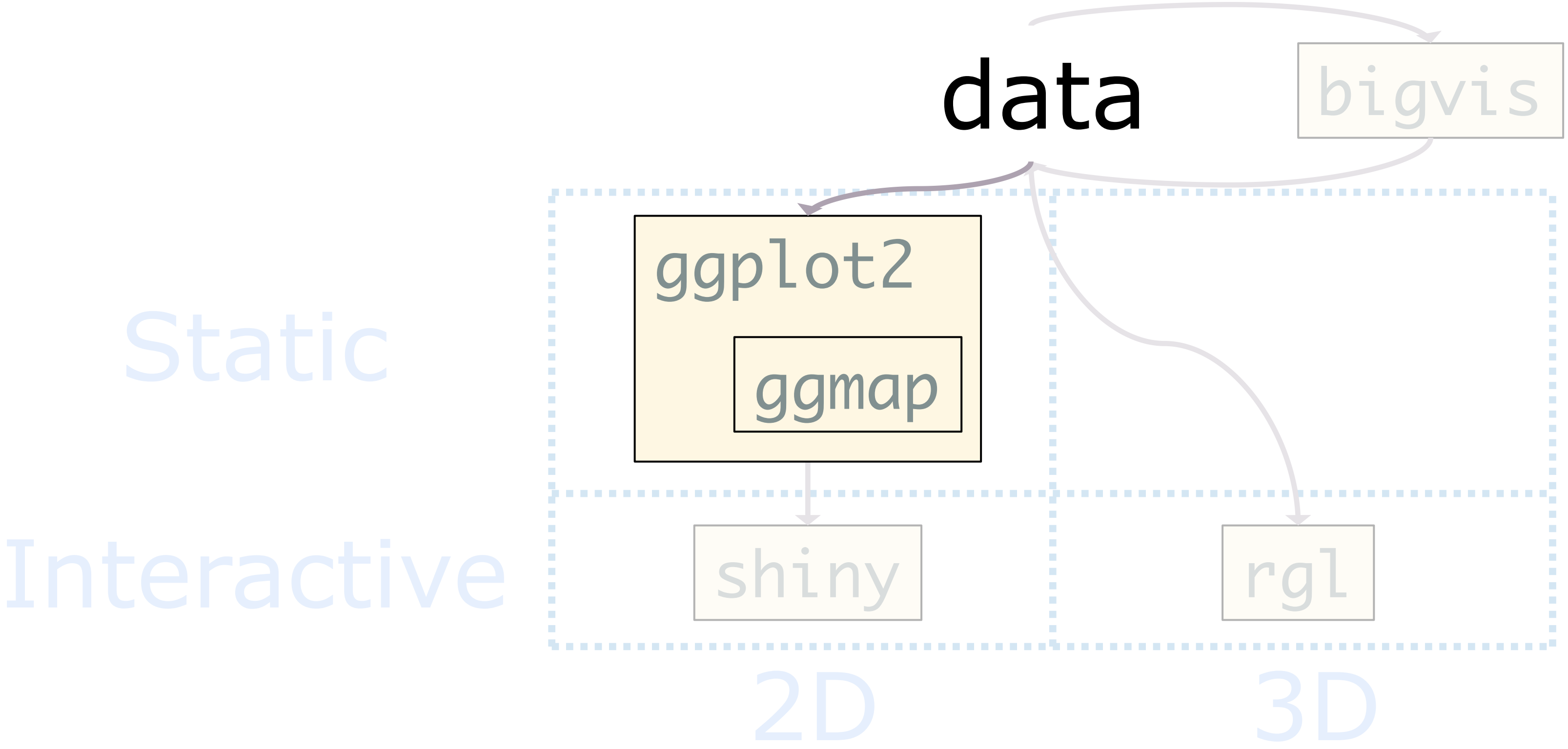
Visualization Map

- Some (of many) R packages for making visuals.



Use the grammar of graphics to see where the seal went.

EXERCISE 1: SEAL'S PATH



Wilkinson, Leland. (2005) “The Grammar of Graphics”. Springer.

Why Grammar?

- Provides a framework for constructing statistical graphics
  - … ***any*** statistical graphic??

> …the grammar effectively defines the parameter space of statistical graphics.
>
> — Hadley Wickham

- There are **6 components***

Wickham, Hadley (2010). “ggplot2: Elegant Graphics for Data Analysis”. Springer

*Or at least I believe there are. They’re not crystal-clear in the book ^.

1. Aesthetics

- A mapping from *variables* to things we can *perceive*.

↔ ↕

```
x = longitude
y = latitude
```

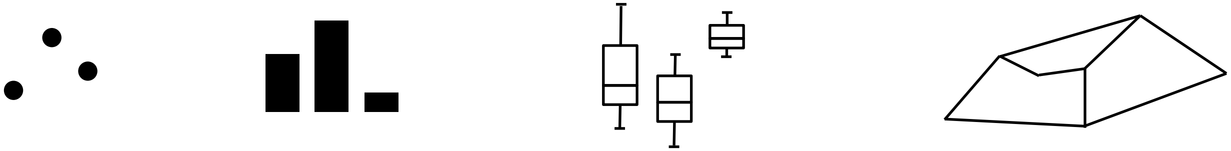
- Other examples:
  - Colour, shape, size

2. Geometric Objects

- The type of object plotted.

```
geom = ‘path’
```

- Other examples (there’s lots!):
  - Points, Bars, Boxplots, Polygons,…



We’ll stop at this component for now.

http://developer.r-project.org/Logo/

ggmap

- Use the grammar on top of raster maps. (see tutorial)
- Workflow:

1. Get a map:

```
my_map <- get_map(...)
```

From:

Google Maps, OpenStreetMap, Stamen Maps, Cloudmade Maps

Apparently has 1000s of map styles!

ggmap

2. Plot it.

```
ggmap(my_map)
```

3. Add ggplot lines

Your would-be first line in making a ggplot2 plot

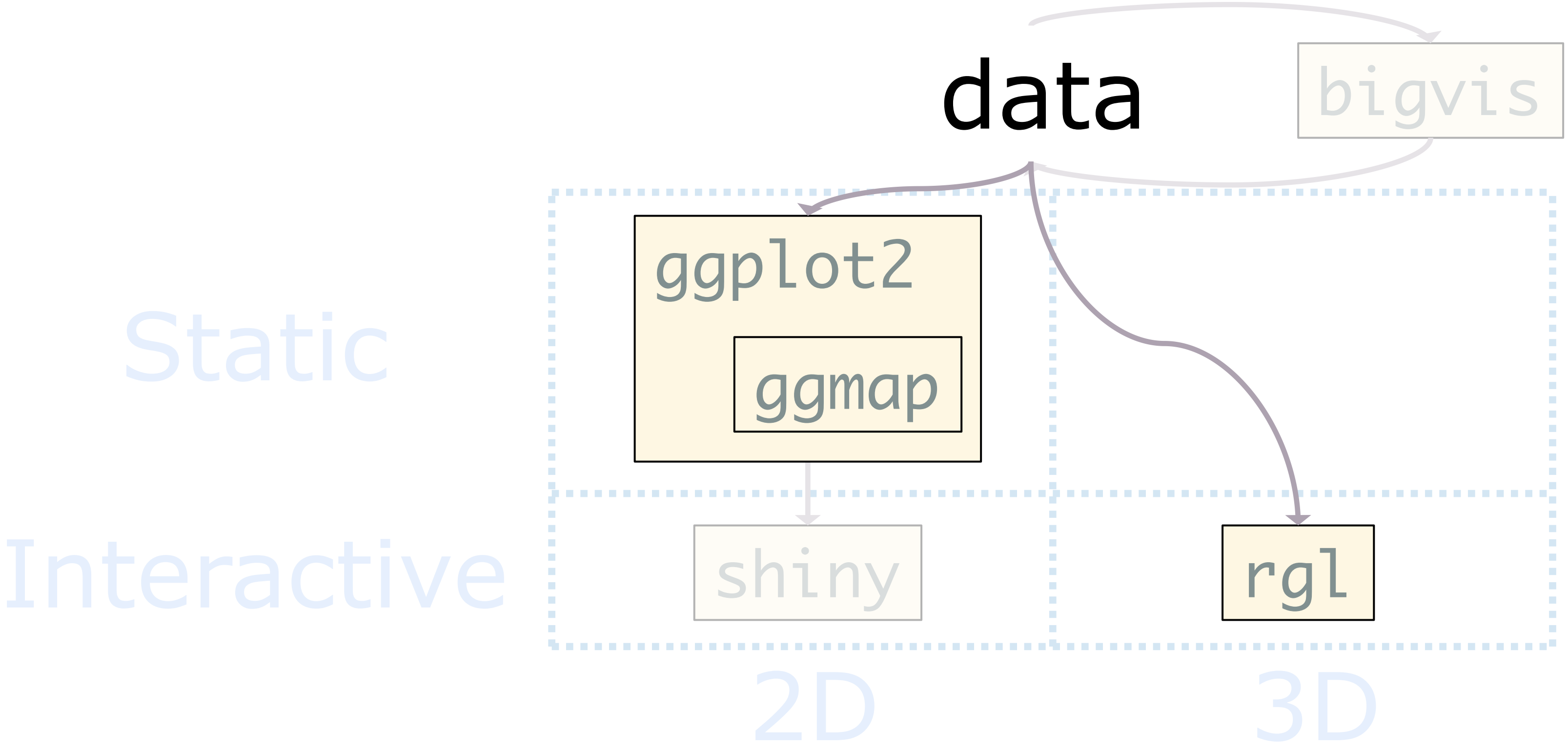
```
ggmap(my_map, base_layer = ggplot(...)) +
   #--ggplot lines--#
```

Add ggplot lines as usual

http://developer.r-project.org/Logo/

- View magnetometer/accelerometer data in 3D.
- View in 2D with ggplot2.

EXERCISE 2: MAGNETOMETER



rgl

- An interactive 3D data plotter.

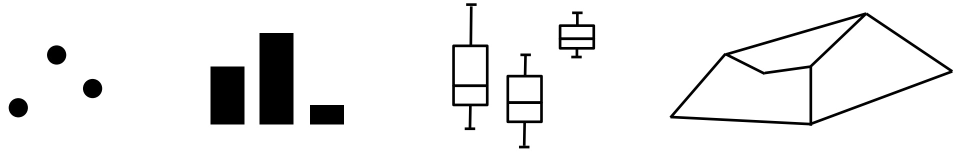
Can rotate and zoom interactively

- Does not (?) use the grammar.
- Let's take a look at the magnetometer data.

```
plot3d(x, y, z)
```

http://developer.r-project.org/Logo/

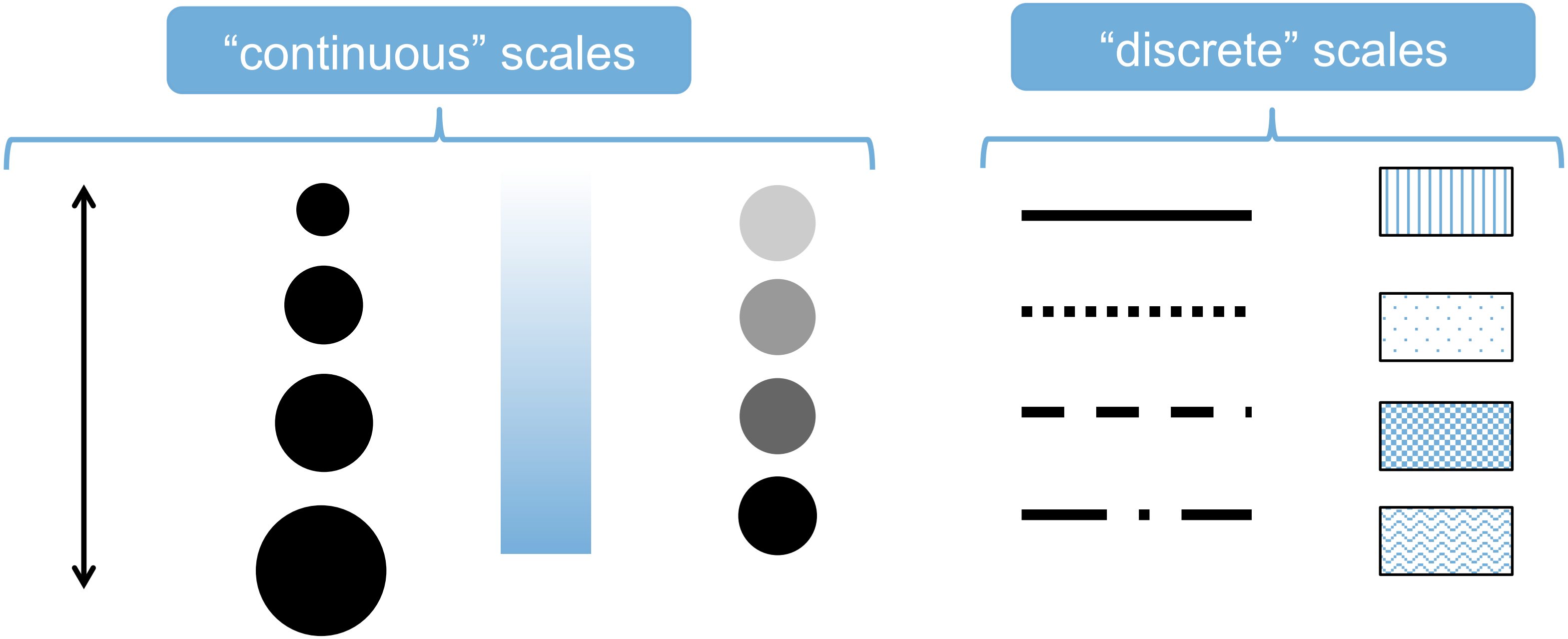
Back to ggplot2

- We can display the data in 2D.
- Need to know the other grammar components.
- Recall:
  1. Aes.: ⟷ ↕ Colour sHaPe siZe
  2. Geom.: 

http://developer.r-project.org/Logo/

3. Scale

- …are **levels** of an aesthetic.
- Use *legends* and *axes* to communicate the scale.



4. Statistical Transformations

- Plots a statistical transformation of the data.

```
Data → mean regression
```

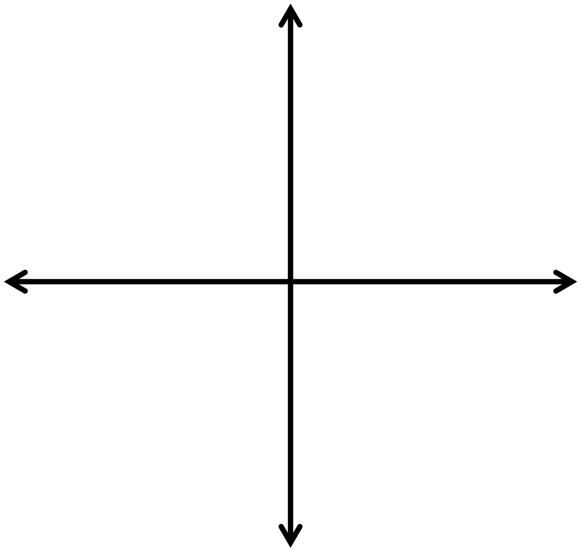
```
Data → count
```

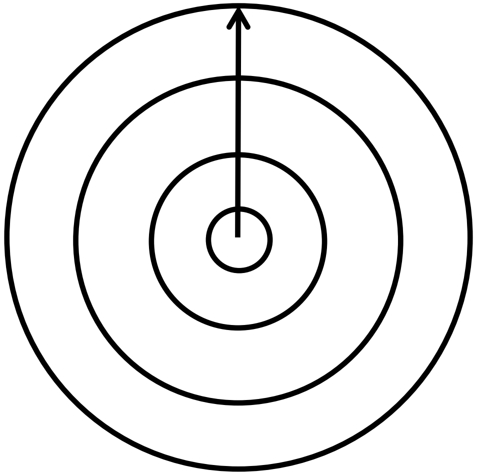
```
Data → quantiles
```

- Typically a default is paired with geom.

5. Coordinates

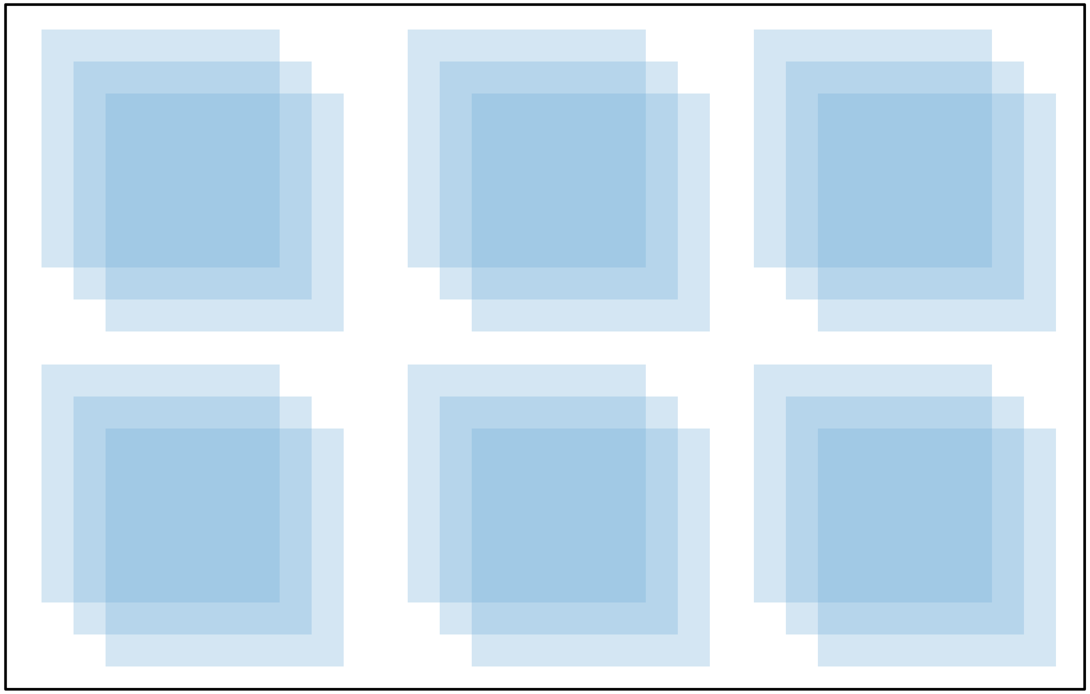
- How to **position** objects on the plot.





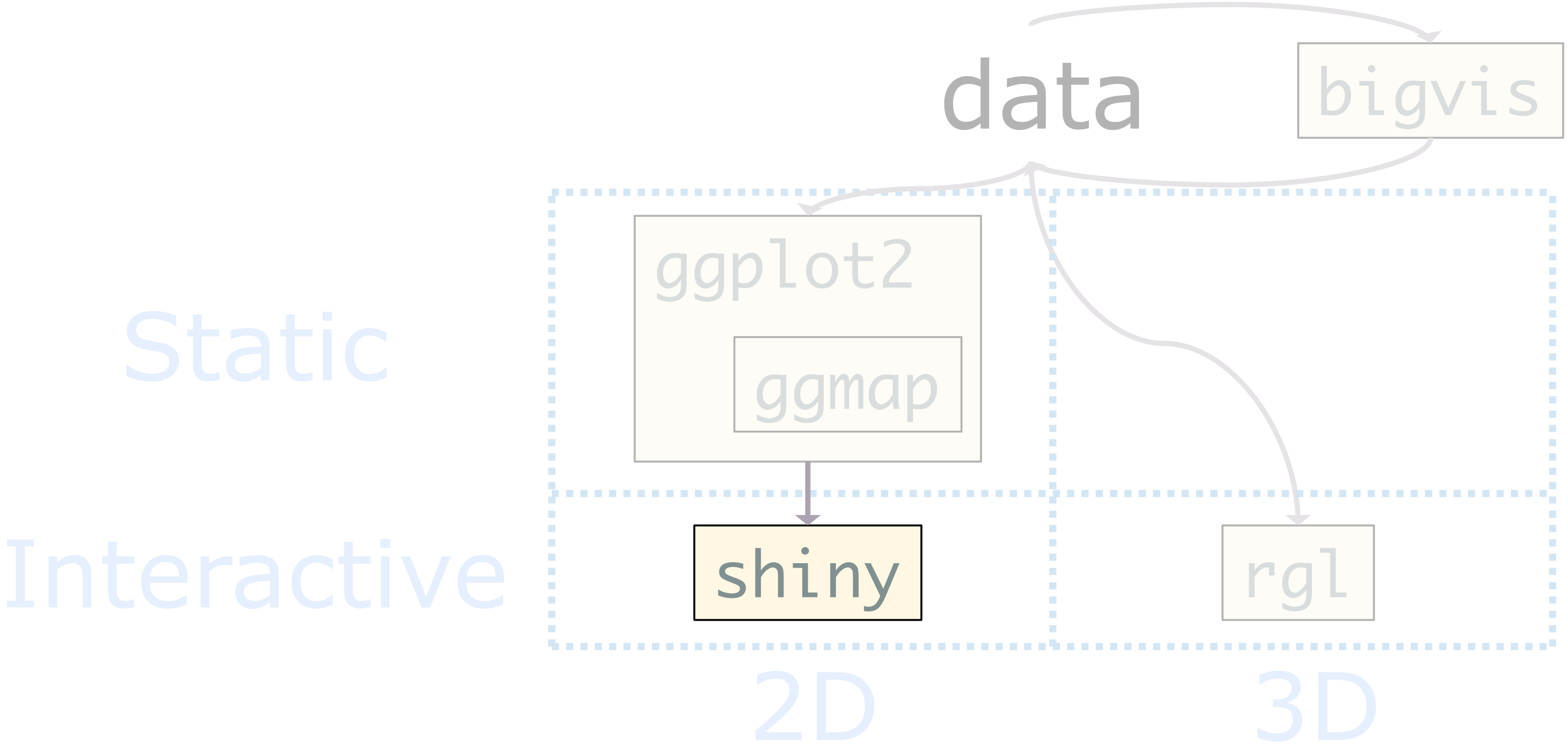
6. Layers

- Combine components 1-5 to get a plot "**layer**".
- Put layers in ≤3D array



http://developer.r-project.org/Logo/

EXERCISE 3: CONNECT AND BRING TO LIFE

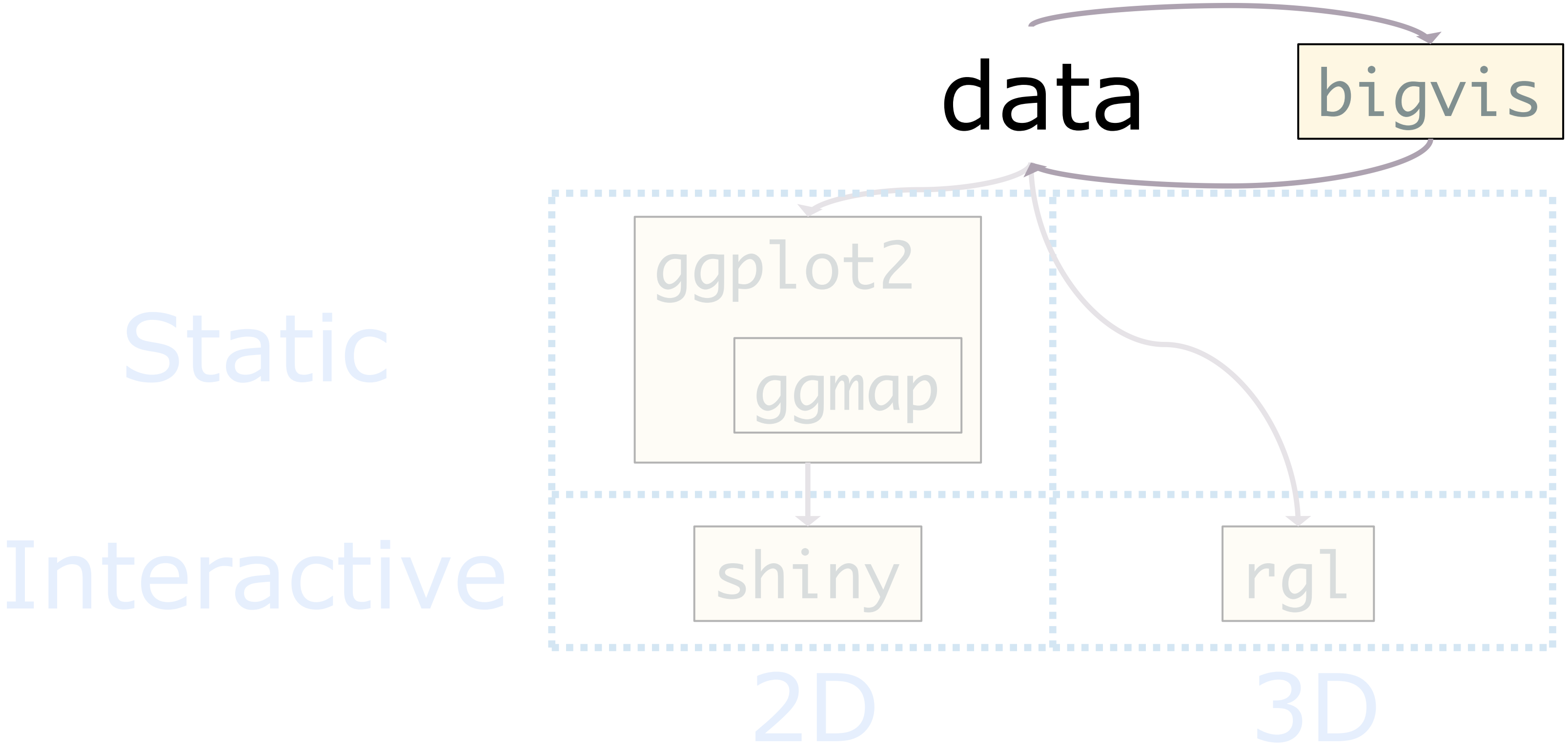


shiny

- Make **apps** anyone can use
- I list it as 2D, but it might accommodate 3D (rgl) too.
- To make, code two .R files:
  - *User interface*
  - *Server* (the "workhorse")

http://developer.r-project.org/Logo/

BIGVIS



data

bigvis

ggplot2

ggmap

shiny

rgl

Static

Interactive

2D

3D

What to do about overplotting?

- We already saw two:

1. Subset the data
2. Use alpha-transparency

- Alternative:

3. Use *bins*: ‘bigvis’ can help!

bigvis

1. **Bin** the data

```
bin(x)
```

2. **Summarize** data in each bin
   - Count; Mean; s.d.; etc.

```
condense(<<binned object(s)>>)
```

   - Limited to a 2D grid!

bigvis

- Ready to plot!
- But, plot may be:
  - …noisy, so **smooth** the bin data

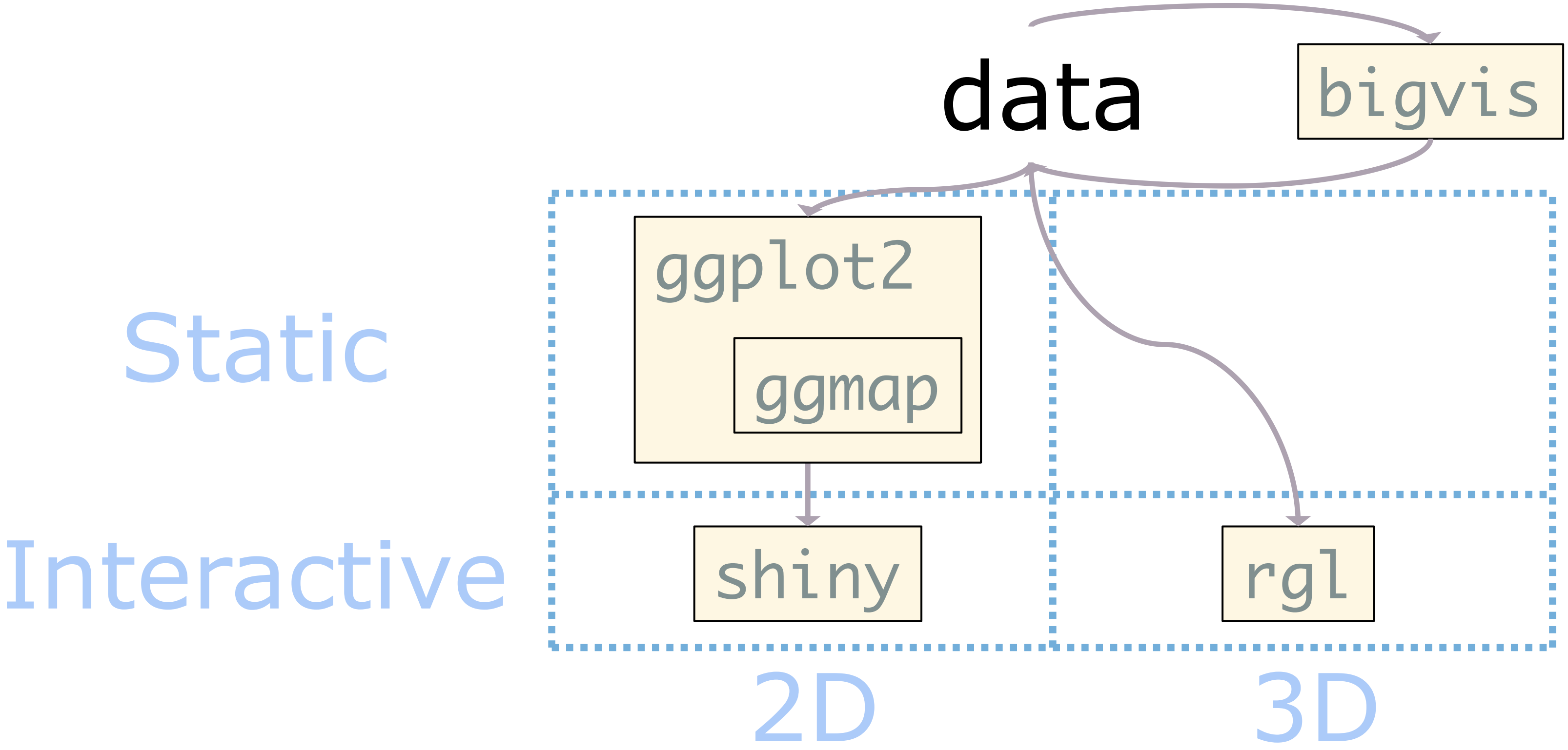
    ```
    smooth()
    ```

  - …imbalanced, so **remove** outliers

    ```
    peel()
    ```

- *Now* plot.

CONCLUSIONS



data

bigvis

Static

Interactive

ggplot2

ggmap

shiny

rgl

2D

3D

Gallery

- gg* plots from James Cheshire
  - Bike Routes + Pollution Map
  - Journeys to Work Map
  - General ggplot2
- rgl:
  - Matt Leonawicz’ Great Circles.
  - See demo for more plots
- shiny’s official gallery.

Summary

- Visualize high-D data with:
  - grammar components (="plotting parameters")
  - exploration through interactivity
- R puts you in control
  - but is not without limitations!
- Did we learn something about our friend the seal?

Useful framework to promote creativity!