Day 12: Data Visualization

Erin Rossiter

24 April, 2023

Announcements

Announcements

– (100% pasted from Day08 slides)
– Final week: project "presentations"
    » i.e., show and tell
    » no slides needed (unless you want)
    » big scraping project? walk us through steps, show final dataset
    » learn a new package? show us how you get it started and one-two cool functions and their outputs
    » everyone has **5 minutes** to share, I will cut you off
    » then ~5 minutes for Q&A
– Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
  - » i.e., show and tell
  - » no slides needed (unless you want)
  - » big scraping project? walk us through steps, show final dataset
  - » learn a new package? show us how you get it started and one-two cool functions and their outputs
  - » everyone has **5 minutes** to share, I will cut you off
  - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
    - » i.e., show and tell
    - » no slides needed (unless you want)
    - » big scraping project? walk us through steps, show final dataset
    - » learn a new package? show us how you get it started and one-two cool functions and their outputs
    - » everyone has **5 minutes** to share, I will cut you off
    - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
    - » i.e., show and tell
    - » no slides needed (unless you want)
    - » big scraping project? walk us through steps, show final dataset
    - » learn a new package? show us how you get it started and one-two cool functions and their outputs
    - » everyone has **5 minutes** to share, I will cut you off
    - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
    - » i.e., show and tell
    - » no slides needed (unless you want)
    - » big scraping project? walk us through steps, show final dataset
    - » learn a new package? show us how you get it started and one-two cool functions and their outputs
    - » everyone has **5 minutes** to share, I will cut you off
    - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
    » i.e., show and tell
    » no slides needed (unless you want)
    » big scraping project? walk us through steps, show final dataset
    » learn a new package? show us how you get it started and one-two cool functions and their outputs
    » everyone has **5 minutes** to share, I will cut you off
    » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
  - » i.e., show and tell
  - » no slides needed (unless you want)
  - » big scraping project? walk us through steps, show final dataset
  - » learn a new package? show us how you get it started and one-two cool functions and their outputs
  - » everyone has **5 minutes** to share, I will cut you off
  - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
  - » i.e., show and tell
  - » no slides needed (unless you want)
  - » big scraping project? walk us through steps, show final dataset
  - » learn a new package? show us how you get it started and one-two cool functions and their outputs
  - » everyone has **5 minutes** to share, I will cut you off
  - » then ~5 minutes for Q&A
- Please take the CIF!

Announcements

- (100% pasted from Day08 slides)
- Final week: project "presentations"
  - » i.e., show and tell
  - » no slides needed (unless you want)
  - » big scraping project? walk us through steps, show final dataset
  - » learn a new package? show us how you get it started and one-two cool functions and their outputs
  - » everyone has **5 minutes** to share, I will cut you off
  - » then ~5 minutes for Q&A
- Please take the CIF!

Random presentation order

```
set.seed(1)
sample(1:14)
```

```
## [1]  9  4  7  1  2 12  3  6 10  8  5 11 13 14
```

- Rasheed
- Nico
- Max
- Ali
- Emily
- Jing
- McKenzie
- Ben
- Prithvi
- Abigail
- Adriana
- Binh
- Jody
- Emma

Today

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Today

- Data viz
    - » Base R
    - » ggplot
    - » Shiny
    - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
    - » Save 30-45 minutes for this

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Today

- Data viz
  - » Base R
  - » ggplot
  - » Shiny
  - » (again, not extending to theory of how to visualize results)
- Debrief/feedback
  - » Save 30-45 minutes for this

Data Viz

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

- Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
- Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
- Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

– Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
– Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
– Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

- Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
- Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
- Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

- Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
- Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
- Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

– Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
– Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
– Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Base R plotting and ggplot are both used for creating visualizations in R, but they differ in the way they approach plotting and the syntax used.

Similarities:

– Both can be used for lots of plot types, like scatter plots, line charts, bar charts, histograms. . .
– Both allow you to customize plot aesthetics like colors, shapes, sizes. . .
– Both take data types like data frames, matrices, vectors. . .

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- Default settings
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- Learning curve
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
    - » Base R–plots each element one by one (you say exactly *what* and in what order)
    - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
    - » Base R–limited default settings for aesthetics
    - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
    - » base R–easy to get started
    - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
        - • but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
    - » Base R–plots each element one by one (you say exactly *what* and in what order)
    - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
    - » Base R–limited default settings for aesthetics
    - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
    - » base R–easy to get started
    - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
        - but, easier once you know it (less lines of code)!

Base R vs. ggplot

Differences:

- **Syntax**
  - » Base R–plots each element one by one (you say exactly *what* and in what order)
  - » ggplot–layers different elements (more abstract, you say *how* not exactly *what*)
- **Default settings**
  - » Base R–limited default settings for aesthetics
  - » ggplot–wide range of default settings that are already high quality/professional.
- **Learning curve**
  - » base R–easy to get started
  - » ggplot–steeper learning curve, more complex syntax, lots of companion packages to learn
    - • but, easier once you know it (less lines of code)!

Resources

Cheatsheet

R Graph Gallery

Base R Cheatsheet

RShiny

What is it?

- A toolkit to write interactive web applications using R code
- Let's check out this gallery
    - » Tip: if you ever see an academic have a shiny app with data you're interested in, just email them
- Note: dashboards are beyond the scope of today

What is it?

- A toolkit to write interactive web applications using R code
- Let's check out this gallery
  - » Tip: if you ever see an academic have a shiny app with data you're interested in, just email them
- Note: dashboards are beyond the scope of today

What is it?

- A toolkit to write interactive web applications using R code
- Let's check out this gallery
  - » Tip: if you ever see an academic have a shiny app with data you're interested in, just email them
- Note: dashboards are beyond the scope of today

What is it?

- A toolkit to write interactive web applications using R code
- Let's check out this gallery
    » Tip: if you ever see an academic have a shiny app with data you're interested in, just email them
- Note: dashboards are beyond the scope of today

How to share your app

- Host app on shinyapps.io or another hosting service
  - » Book for how to post app on shinyapps.io
- Then, you can embed on your own website

How to share your app

- Host app on shinyapps.io or another hosting service
  - » Book for how to post app on shinyapps.io
- Then, you can embed on your own website

How to share your app

- Host app on shinyapps.io or another hosting service
  - » Book for how to post app on shinyapps.io
- Then, you can embed on your own website

Big Picture

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- – In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- – My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- – Build a community

Also:

- – R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - • building blocks of R
        - • reading others' code, even if its not how you code
        - • (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
  - » collection
  - » organization
  - » cleaning
  - » wrangling
  - » sharing
  - » literacy/fluency
    - building blocks of R
    - reading others' code, even if its not how you code
    - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Big Picture

My goals in the course:

- In terms of research, my main goal was to makes your lives easier and faster, and to open up avenues you might not have pursued:
    - » collection
    - » organization
    - » cleaning
    - » wrangling
    - » sharing
    - » literacy/fluency
        - building blocks of R
        - reading others' code, even if its not how you code
        - (base R vs dplyr or ggplot2)
- My goal was also to build a very solid foundation so you are confident you can teach yourself the newest, coolest thing when no longer a class devoted to it
- Build a community

Also:

- R is a marketable skills in and of itself, esp in combination with statistics skills

Debrief

Debrief

- Open to *any* thoughts
  - » What went well, what didn't
  - » Most useful, least useful
  - » What you'd change, what you'd keep the same. . .

1. What are the most challenging aspects of learning R?
2. How has your understanding of R as a tool, what you're capable of with it, etc. improved through this class?
3. Which R programming concepts did you find most useful and applicable to your research/work/future plans?
4. How would you rate the pace and difficulty level of the class?
5. Were there any areas/topics that you would have liked to spend more/less time on?
6. What suggestions do you have for improving the class for future students?
7. Were the homeworks, and labs helpful for your learning? What about the project?