

Informatics 143

Information Visualization

Lecture 3

*Duplication of course material for any commercial purpose without
the explicit written permission of the professor is prohibited.*

*These course materials are based on books from Claus O. Wilke, Kieran Healy, Edward R. Tufte,
Alberto Cairo, Colin Ware, Tamara Munzner, and others.
Powerpoint theme by Prof. André van der Hoek.*

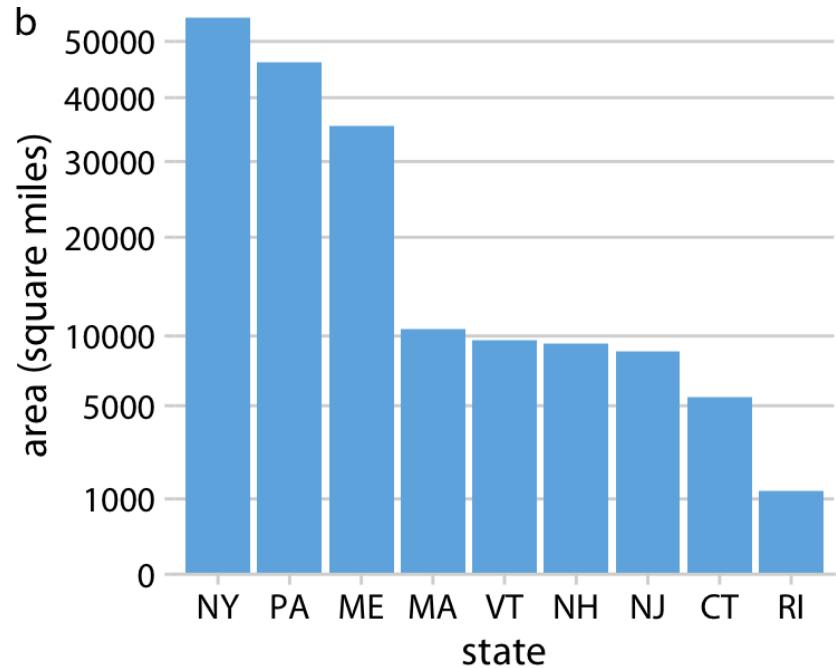
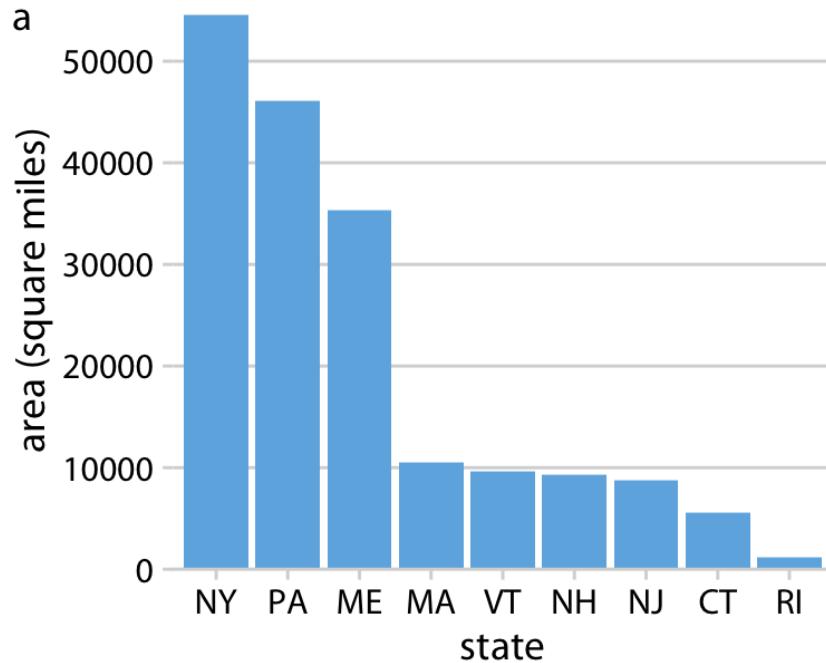
Coordinate systems and axes

- First step in the construction of a visual representation
- **Where the different data values are located?**
- **How they are related to each other?**
 - Horizontally? Vertically? At an angle? Along a circle? On the surface of a sphere? Like a spiral?
 - Linearly? Logarithmically? Discretized? Continuous?

The choice of the position scales and how they are arranged geometrically is what sets a coordinate system in the context of visualization.

A more common solution: square-root scale

- Natural scale for data in squared units



Are there coordinate systems using other axis?



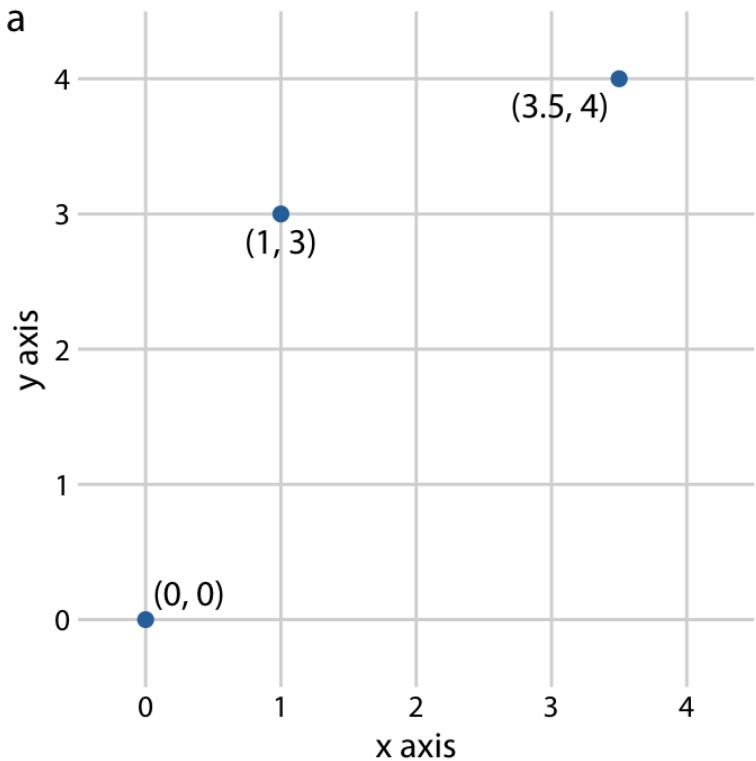
- So far, Cartesian
 - Two data values (p_1, p_2) are mapped into two coordinates (x, y)
 - Axis x and y are positioned at a right angle to each other
 - Mapping $f(p_1, p_2) \rightarrow (x,y)$ can be linear or non-linear (i.e. log, sqrt, ...)

Other coordinates

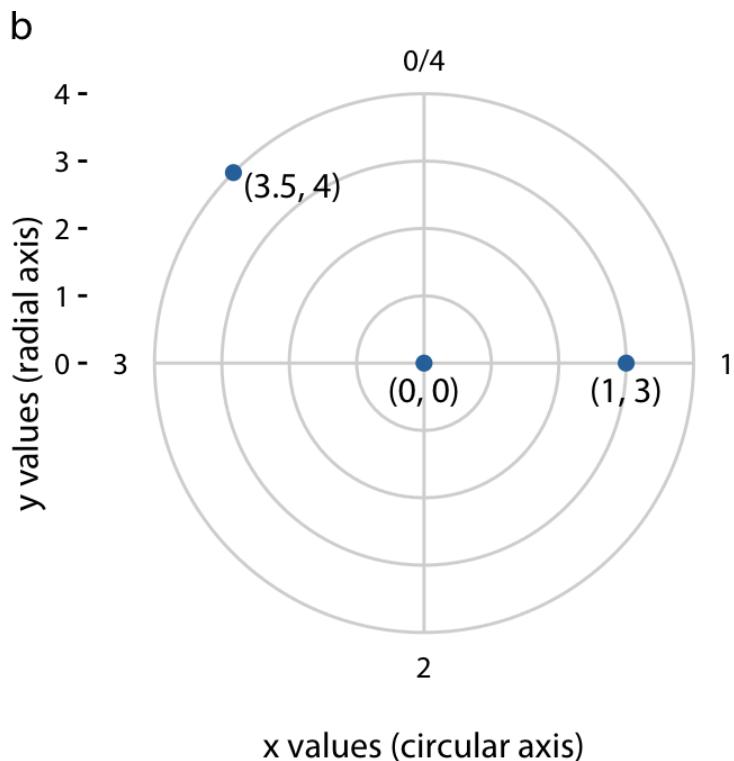
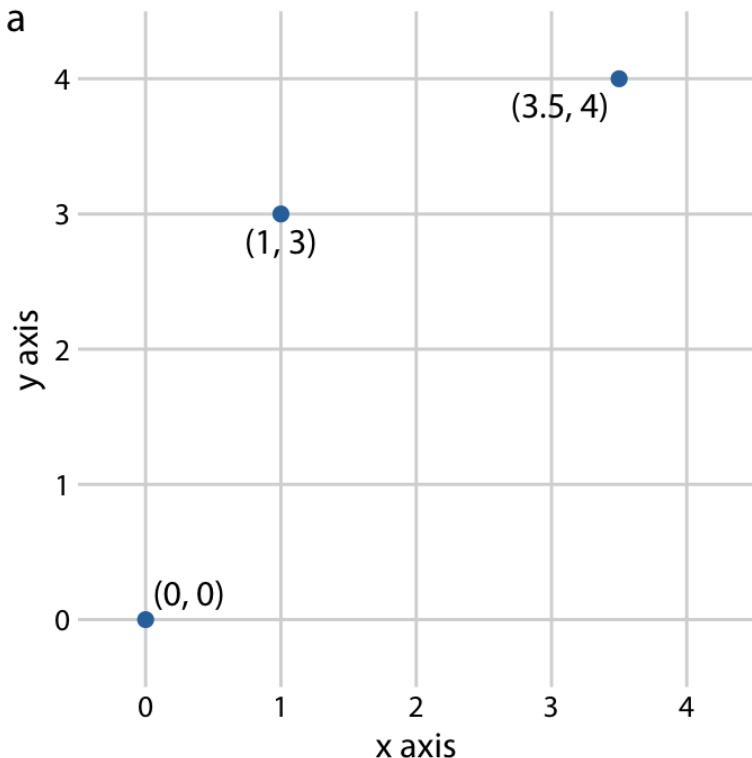


- **Polar coordinates**
 - Two data values (p_1, p_2) are mapped into a pair **(θ, r) angle-distance from the origin**
 - Note for those interested in mathematics : this is closely related to complex numbers

Polar coordinates



Polar coordinates

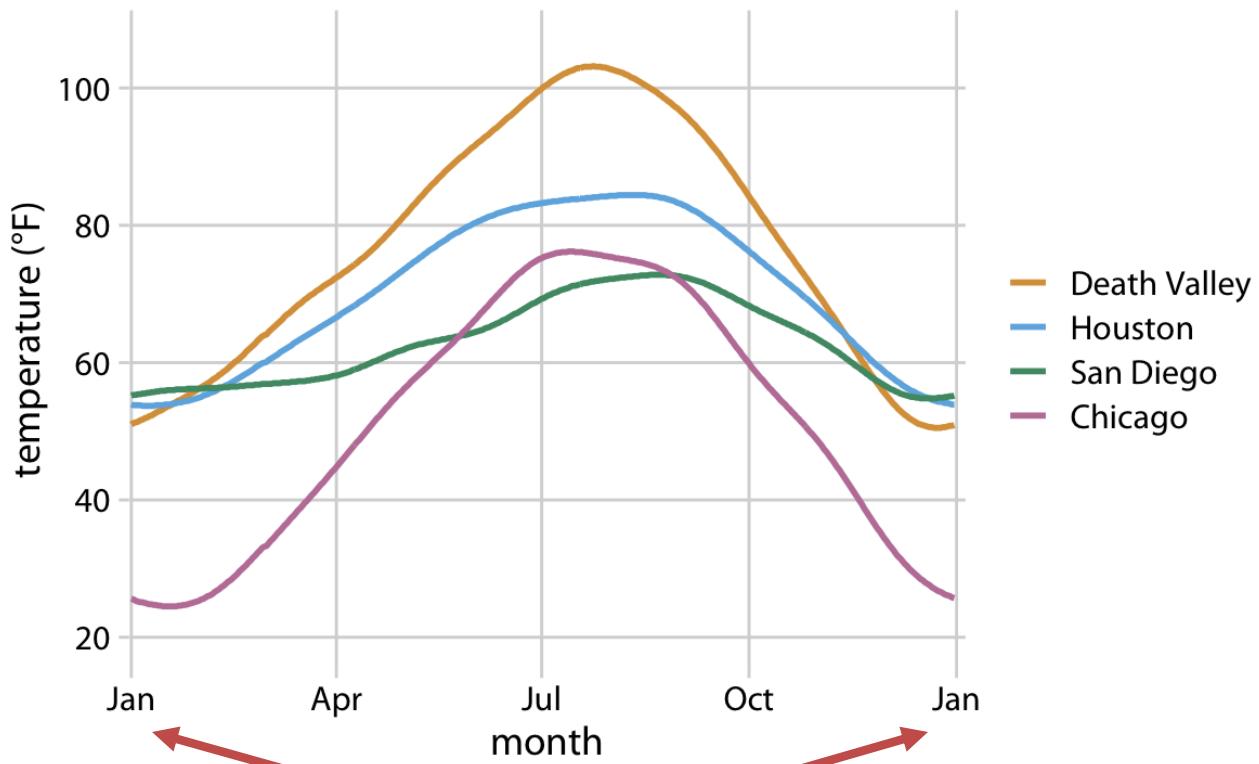


Polar coordinates



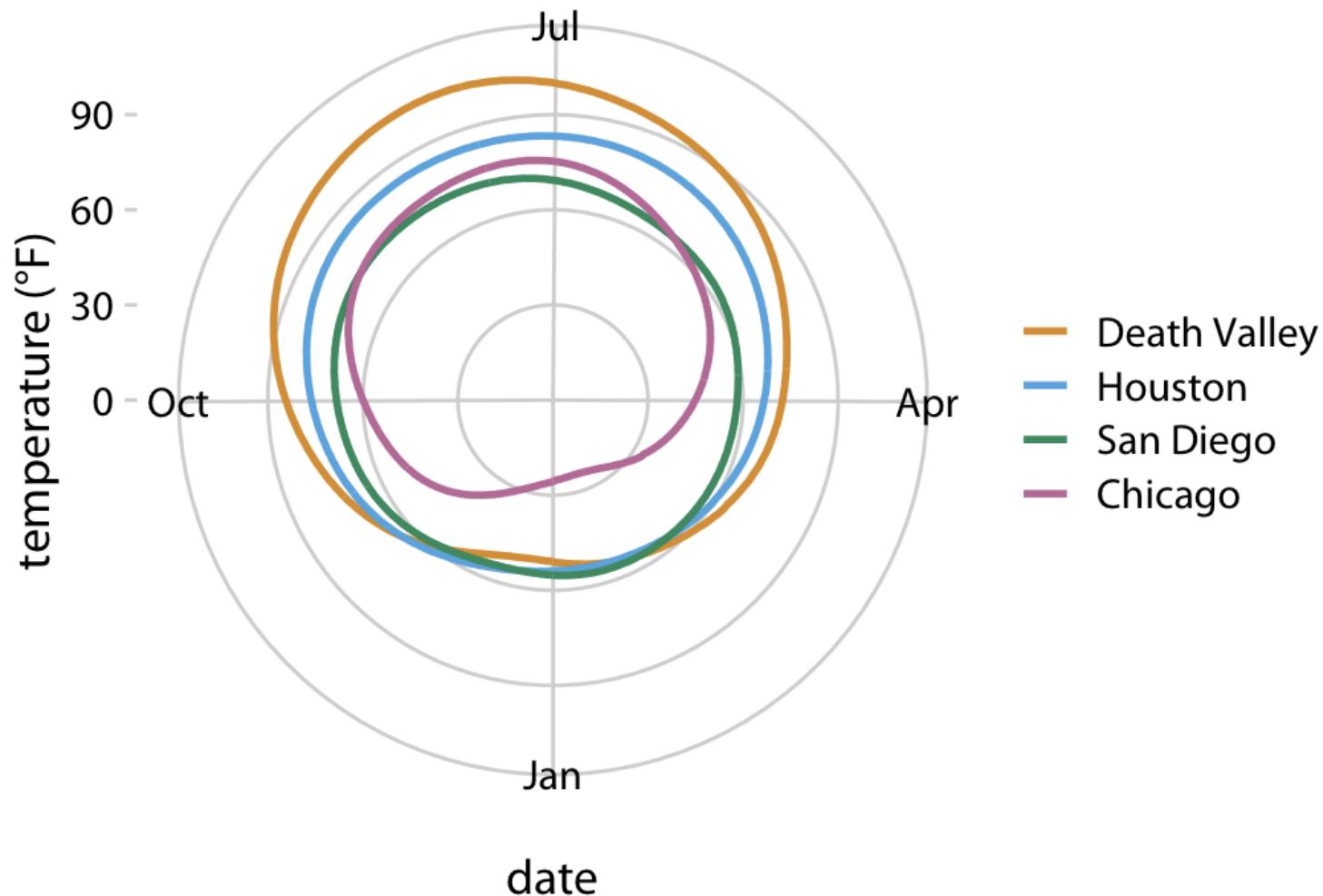
- Polar coordinates
 - Two data values (p_1, p_2) are mapped into a pair (θ, r) angle-distance from the origin
 - Note for those interested in mathematics : this is closely related to complex numbers
 - Good for representing periodic patterns

Polar coordinates



Depending on the *story*, these points should not be apart!

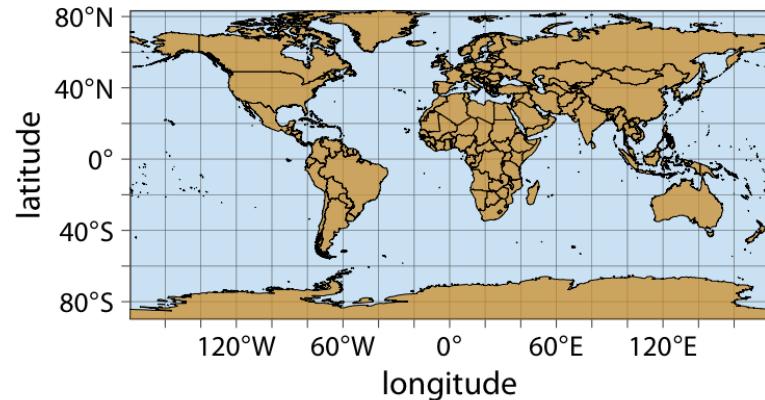
Polar coordinates



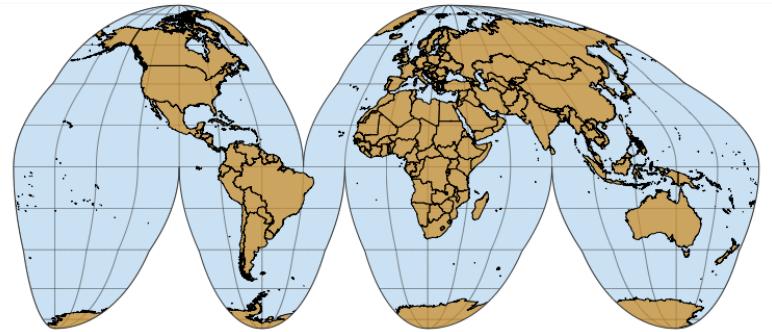
Other coordinates

- Multiple systems are used in geospatial data!

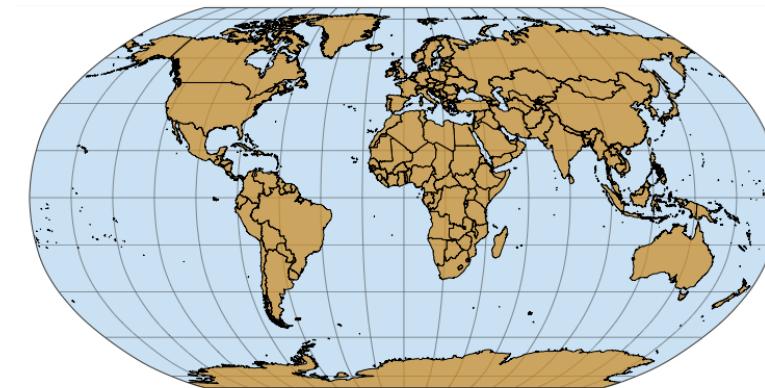
Cartesian longitude and latitude



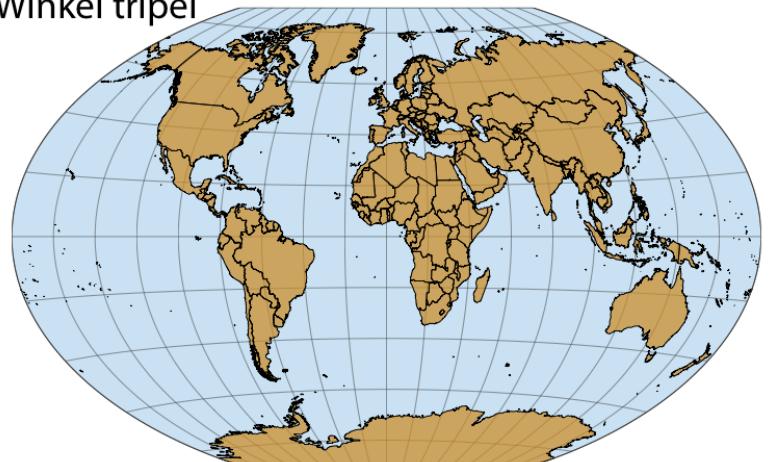
Interrupted Goode homolosine



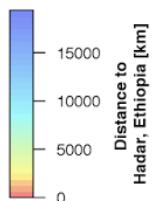
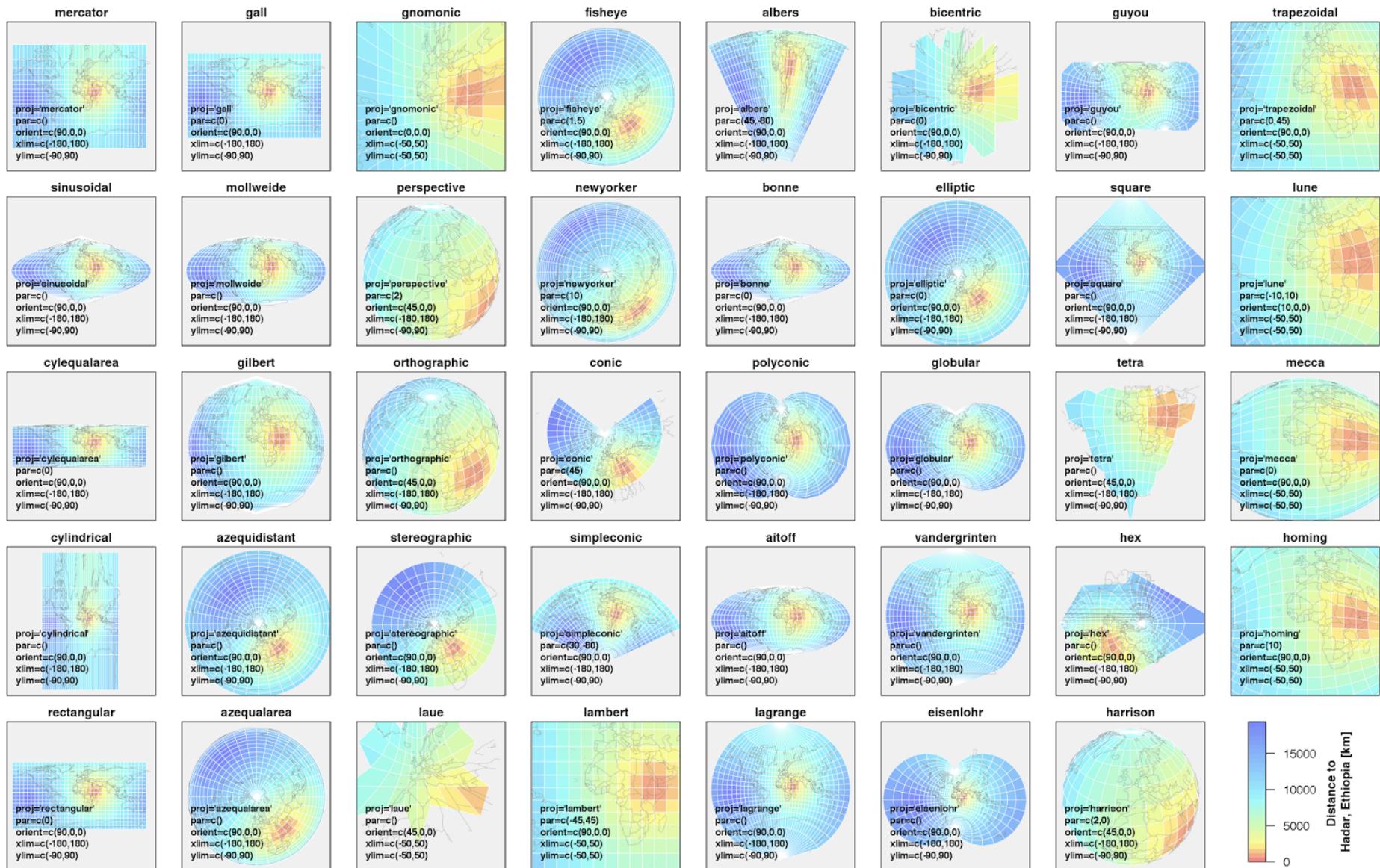
Robinson



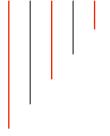
Winkel tripel



Other coordinates



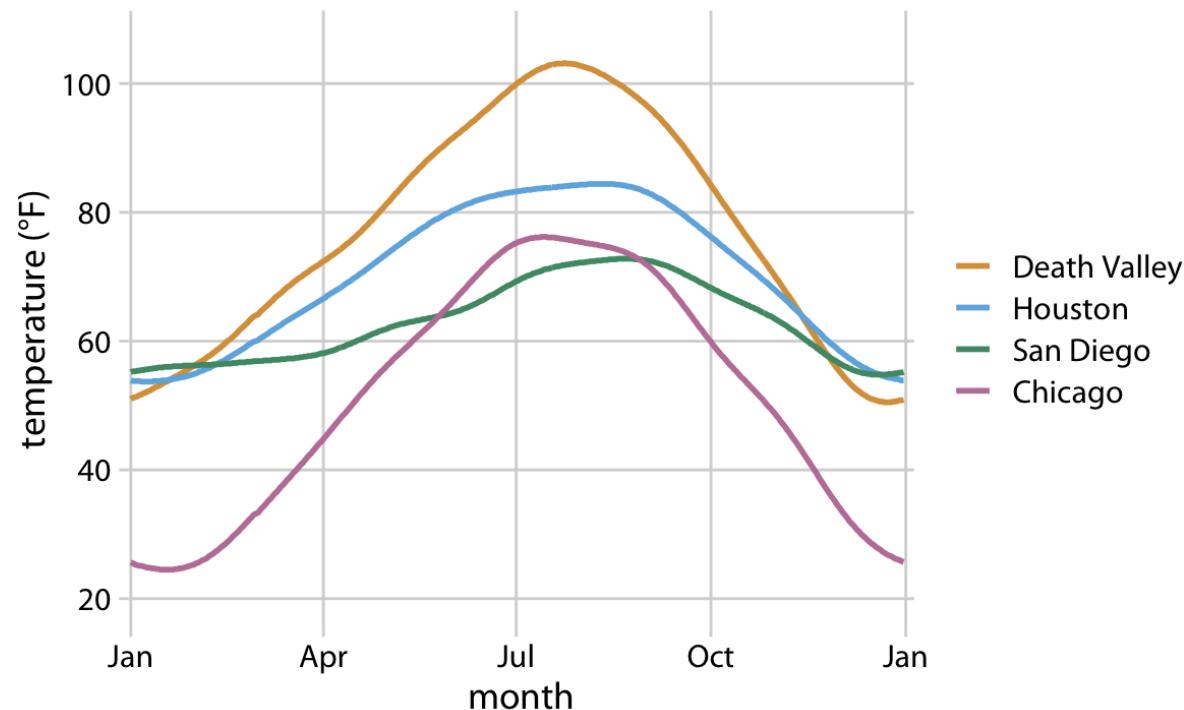
Color



- **Why?**
 - Distinguish data groups
 - Represent data values
 - Highlight important aspects

Color for grouping

- Distinguish discrete items or groups without intrinsic ordering



Color for grouping



- Distinguish discrete items or groups without intrinsic ordering
- **Qualitative color scale**
 - Finite set of colors
 - Clearly distinct from each other
 - No one color should stand out
 - No impression of ordering

Color for grouping

Okabe Ito



Okabe, M., and K. Ito. 2008. "Color Universal Design (CUD): How to Make Figures and Presentations That Are Friendly to Colorblind People." <http://jfly.iam.u-tokyo.ac.jp/color/>.

ColorBrewer Dark2



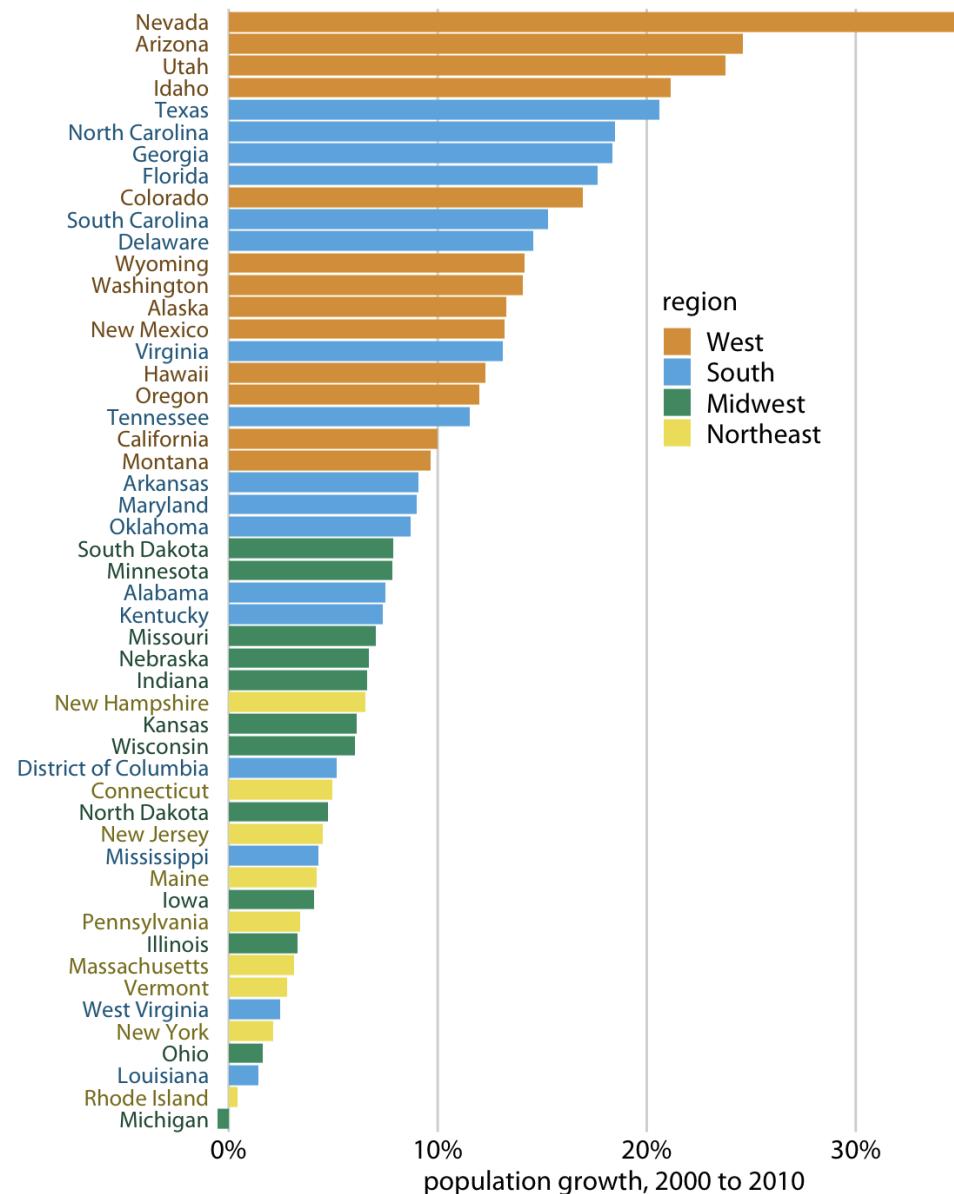
Brewer, Cynthia A. 2017. "ColorBrewer 2.0. Color Advice for Cartography." <http://www.ColorBrewer.org>.

ggplot2 hue



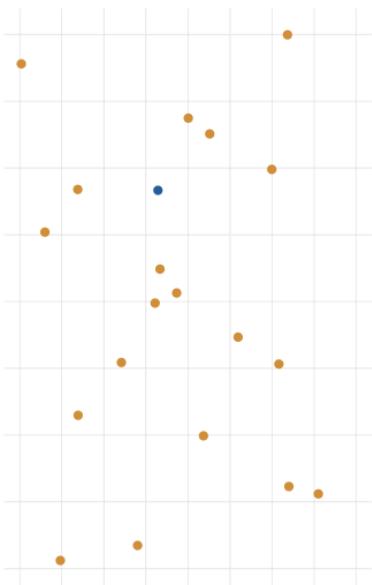
ggplot2

Color for grouping

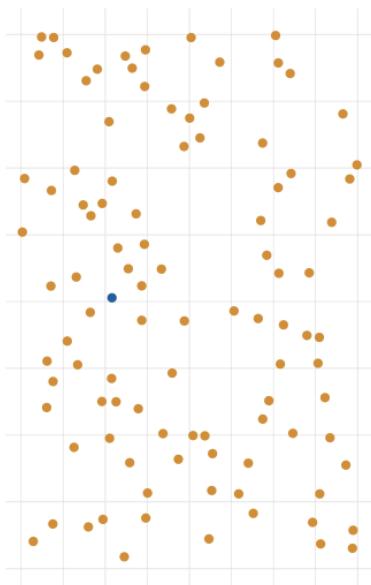


Color for grouping

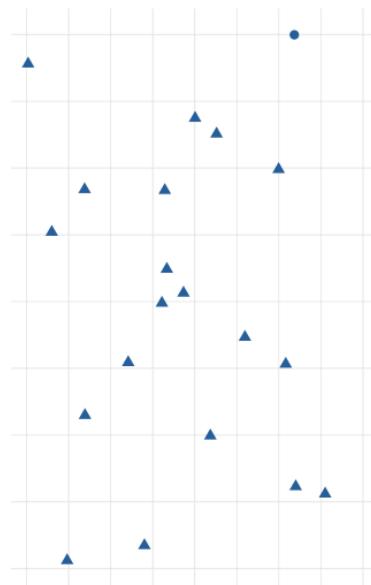
Color Only, N=20



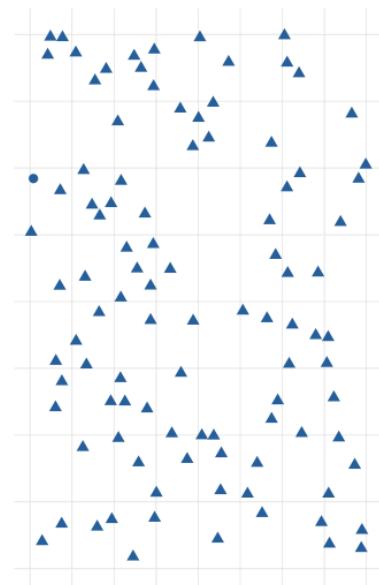
Color Only, N=100



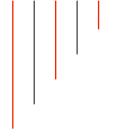
Shape Only, N=20



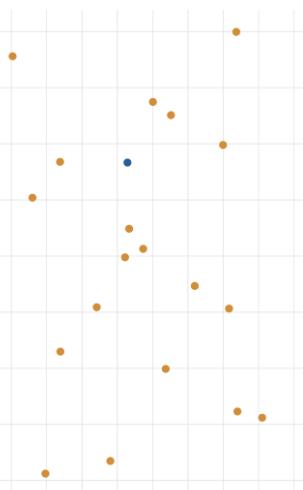
Shape Only, N=100



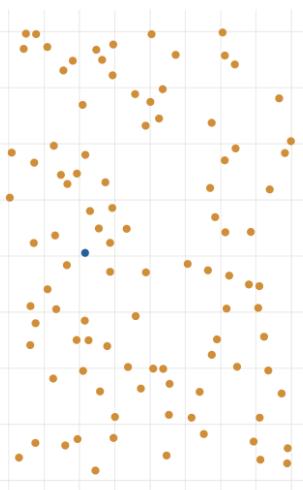
Color for grouping



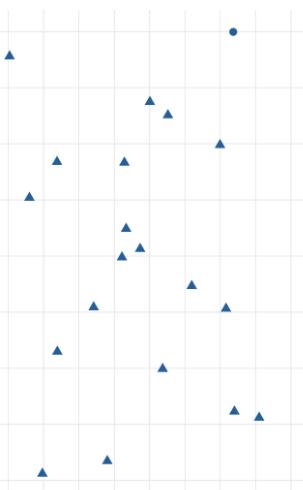
Color Only, N=20



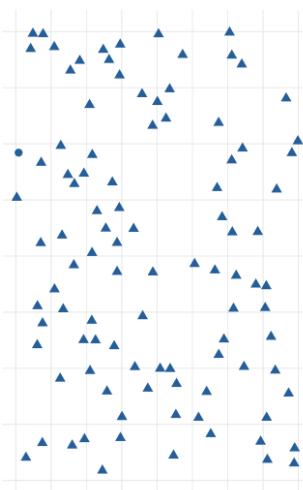
Color Only, N=100



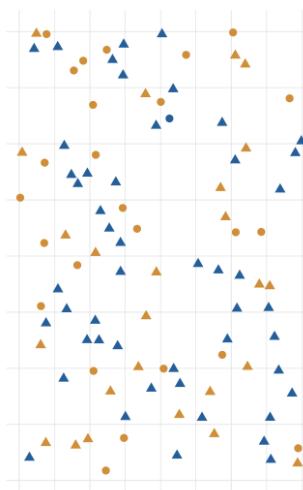
Shape Only, N=20



Shape Only, N=100



Color & Shape, N=100



Color to represent values



- Using color as additional dimension in the visualization
- **Sequential color scale**
 - Usually infinite set of colors
 - Smoothly-varying (i.e. not very distinct from each other if data values are close to each other)
 - Perceptually uniform scale
 - Clear impression of ordering

Color to represent values : sequential



ColorBrewer Blues



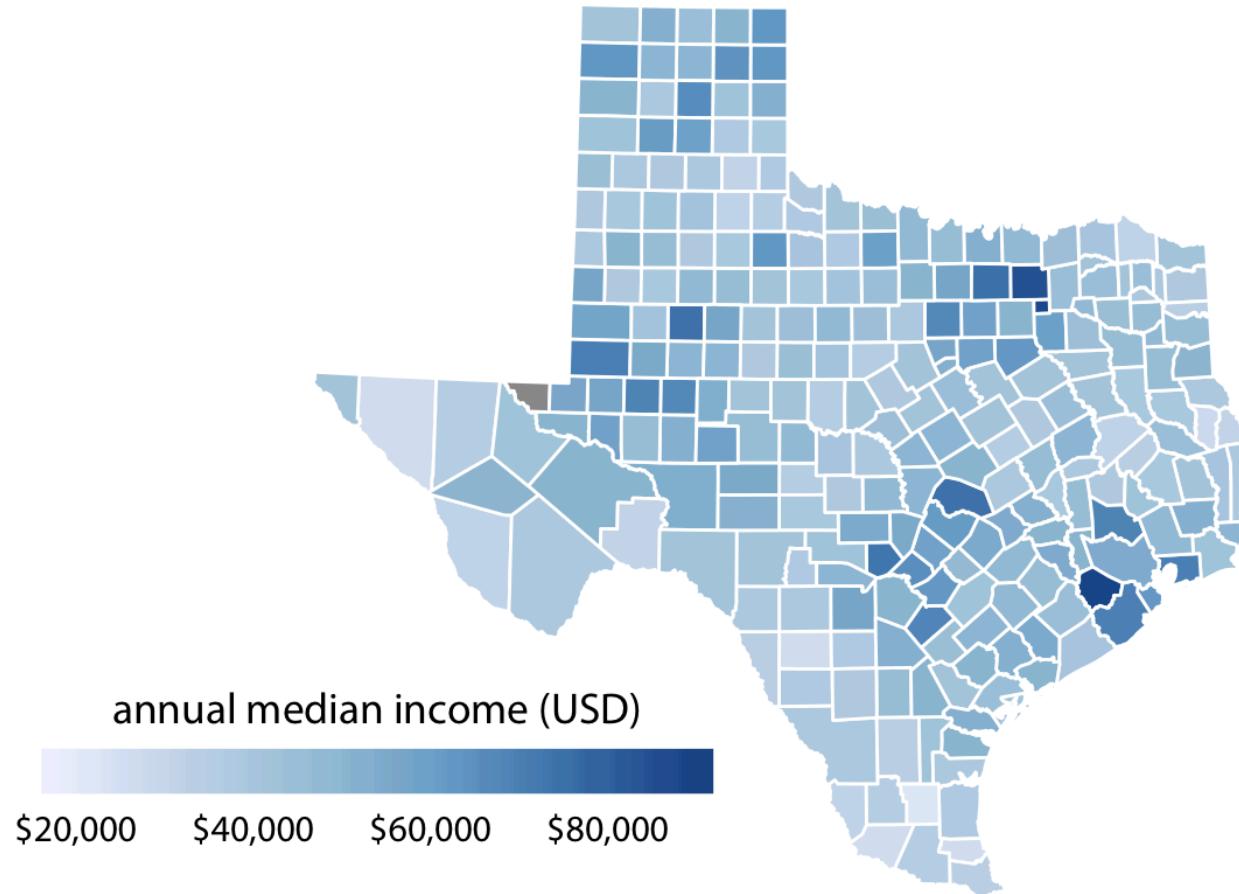
Heat



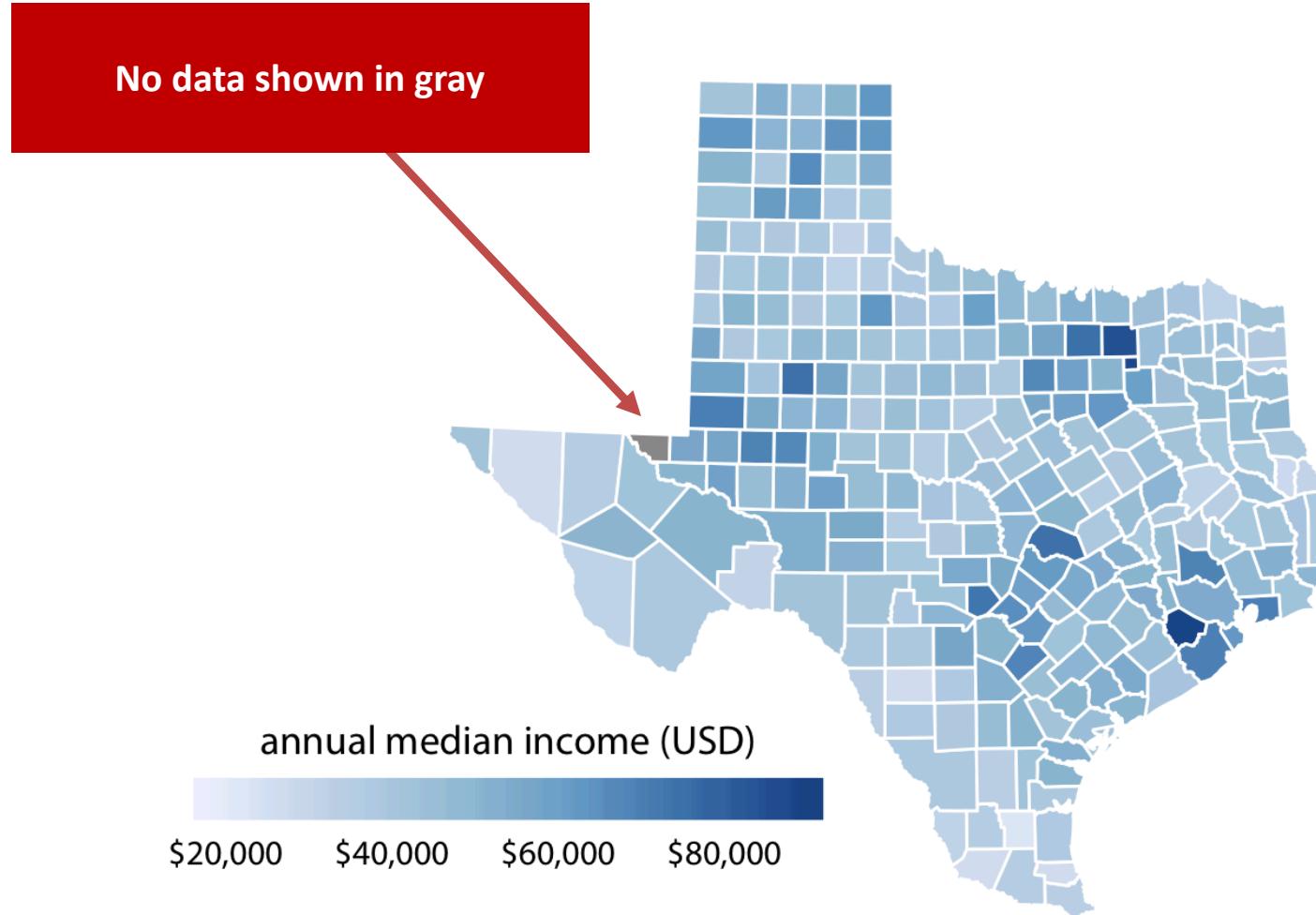
Viridis



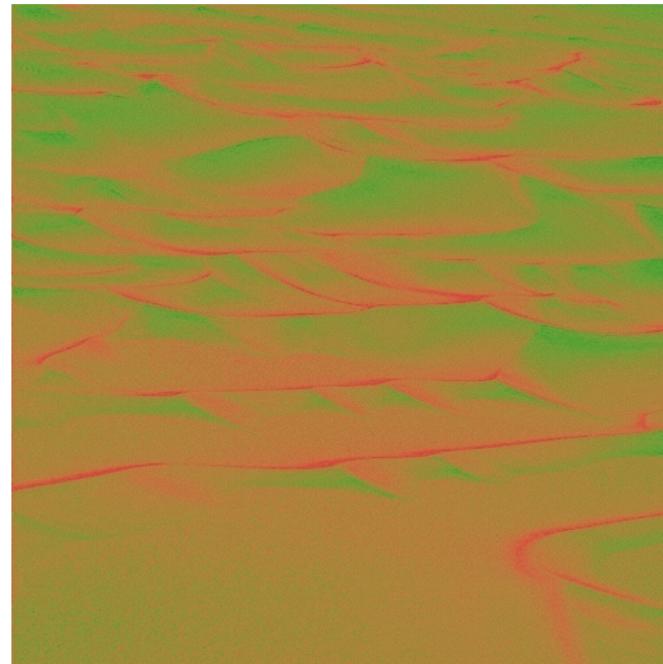
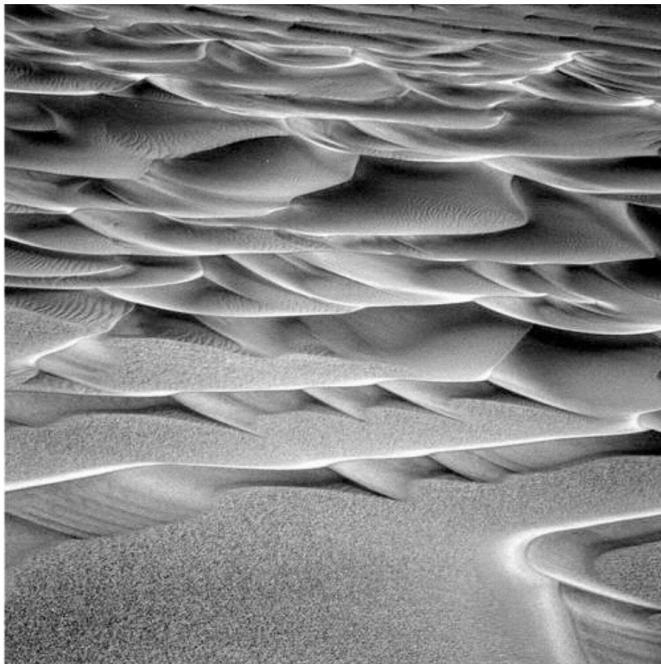
Color to represent values : sequential



Color to represent values : sequential



Color to represent values : sequential



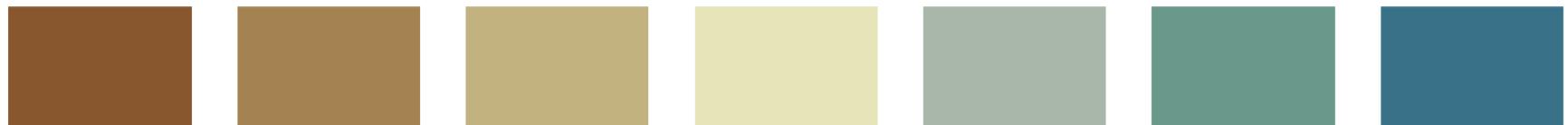
Color to represent values



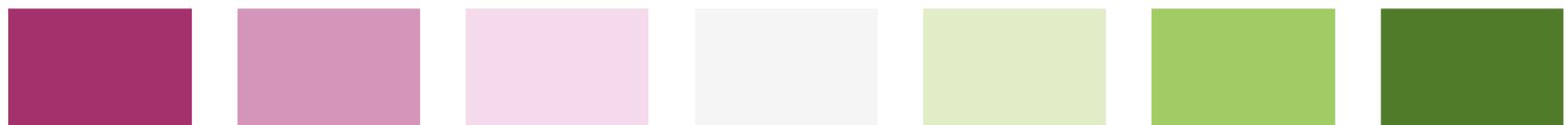
- Using color as additional dimension in the visualization
- **Diverging color scale**
 - Usually infinite set of colors
 - Smoothly-varying (i.e. not very distinct from each other if data values are close to each other)
 - Perceptually uniform scale
 - **Clear impression of ordering but with clear definition for mid-point**
 - Mid-point usually taken as a “weak” part of the scale (i.e. it cannot stand out), so light colors are adopted

Color to represent values : divergent

CARTO Earth



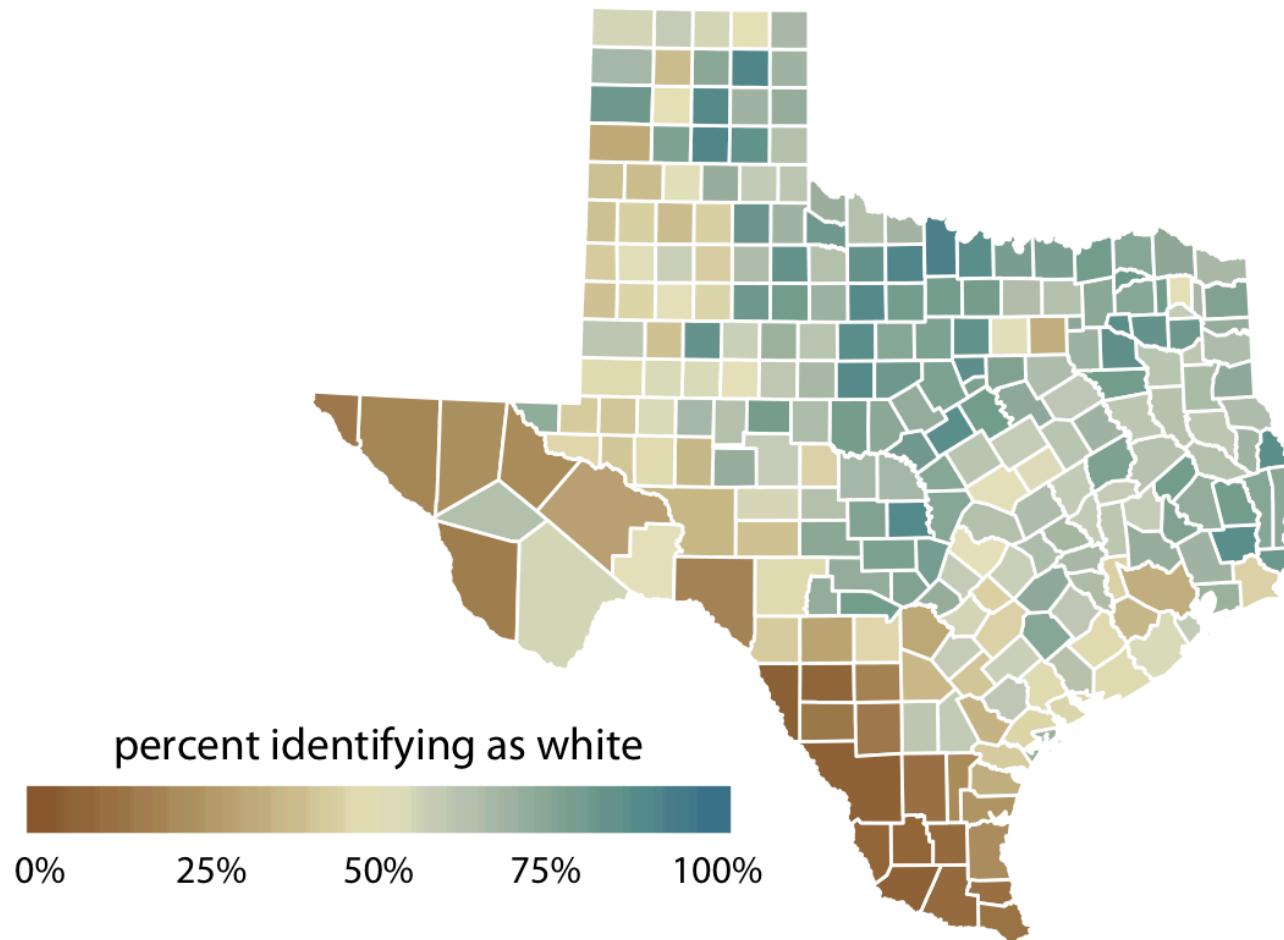
ColorBrewer PiYG



Blue-Red



Color to represent values : divergent



Color to highlight



- Using color to highlight specific elements
- Accent color scale
 - Finite or Infinite set of colors with a **subset that clearly stand out**
 - Can be a modification of qualitative, sequential or divergent scales
 - Example: mixing subdued colors and a matching set of stronger, darker, and/or more saturated colors to create contrast

Color to highlight : accent

Okabe Ito Accent



Grays with accents



ColorBrewer Accent



Color to highlight : accent generation

Lighten / desaturate some colors

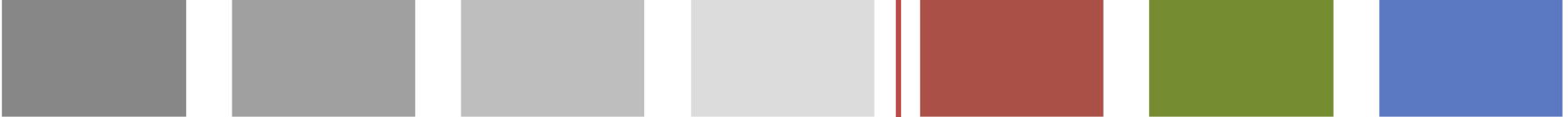
Okabe Ito Accent



Darken others



Color to highlight : accent generation

Gray levels	Colors
Grays with accents	

Color to highlight : accent generation

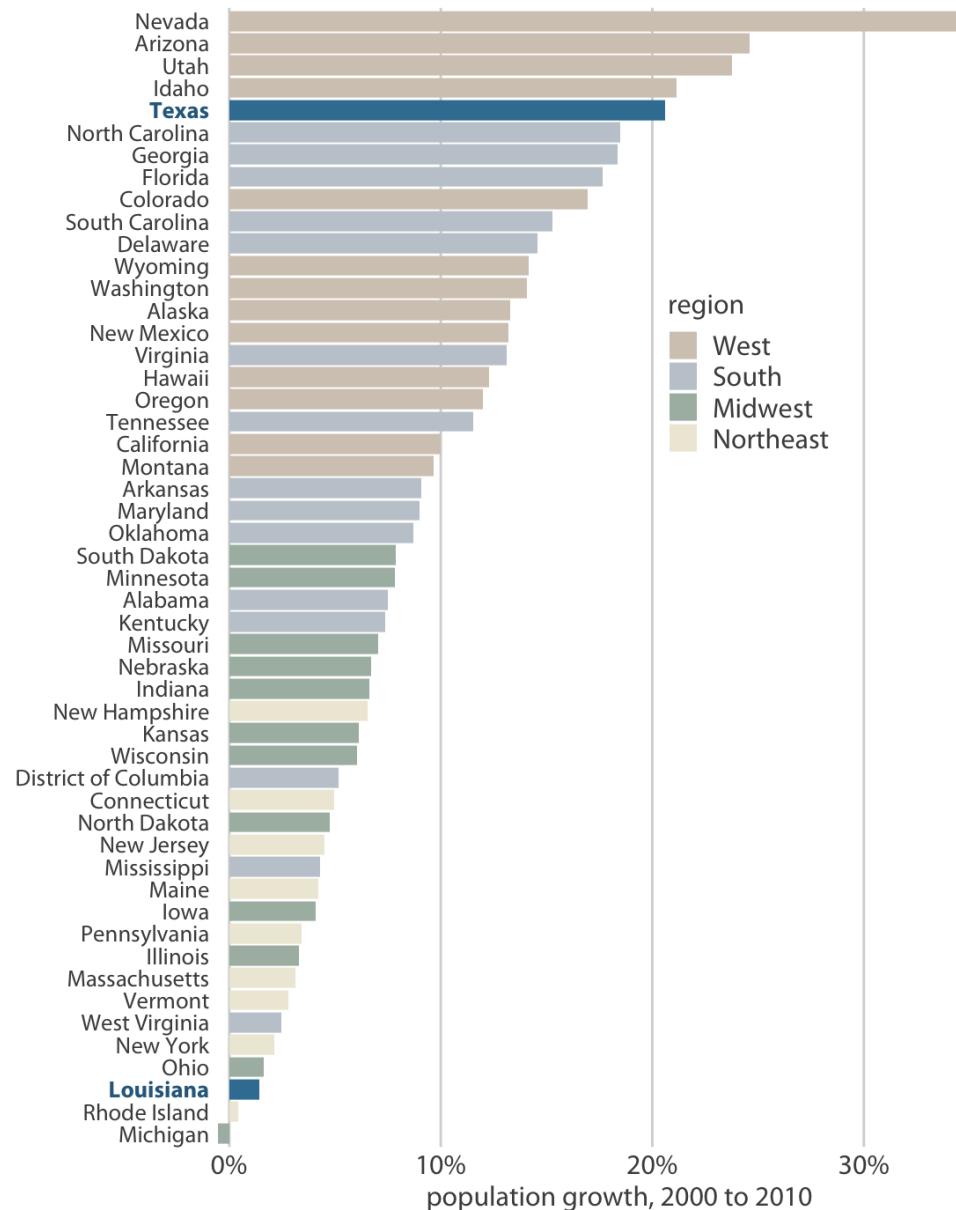


Accent color scale

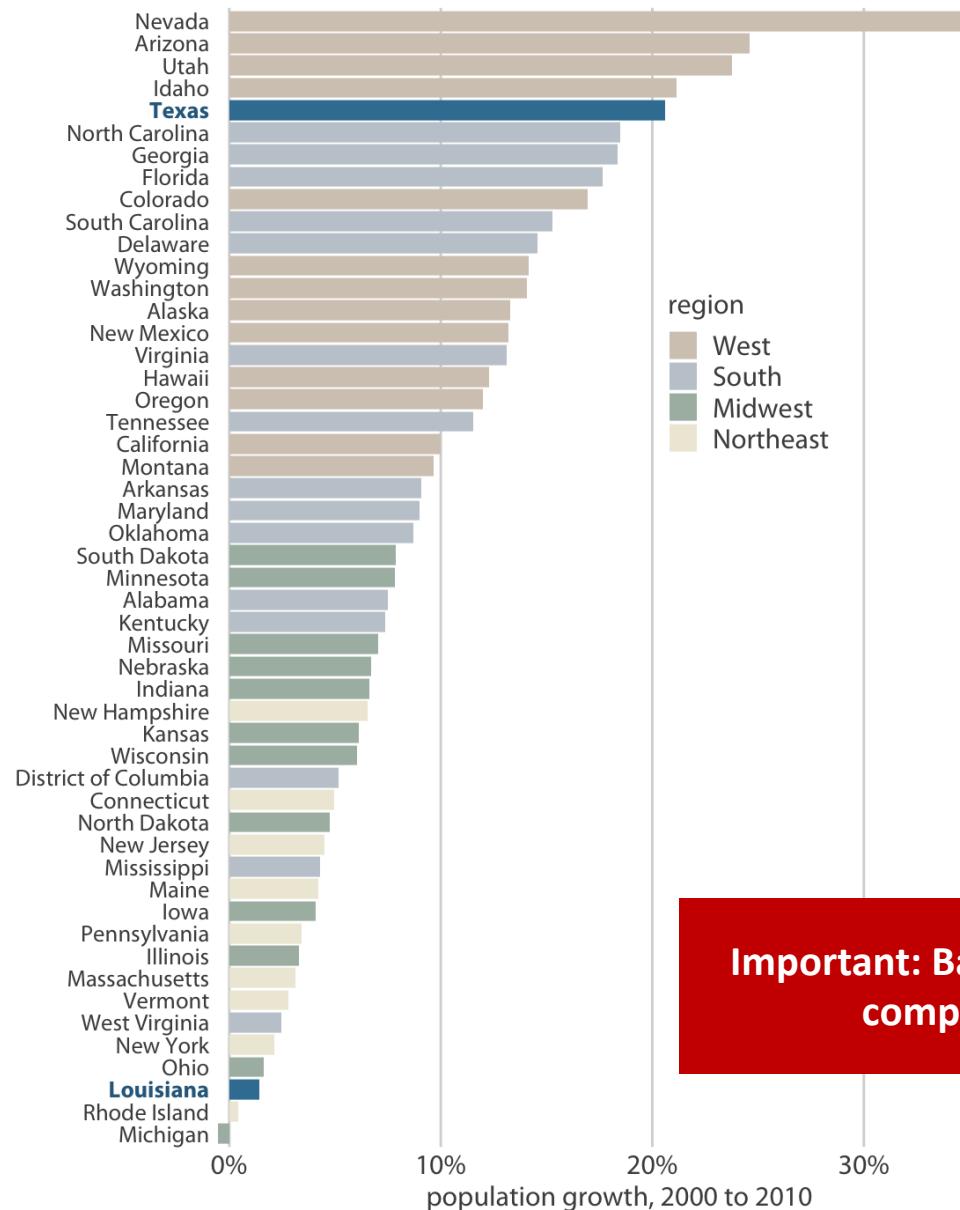
ColorBrewer Accent



Color to highlight

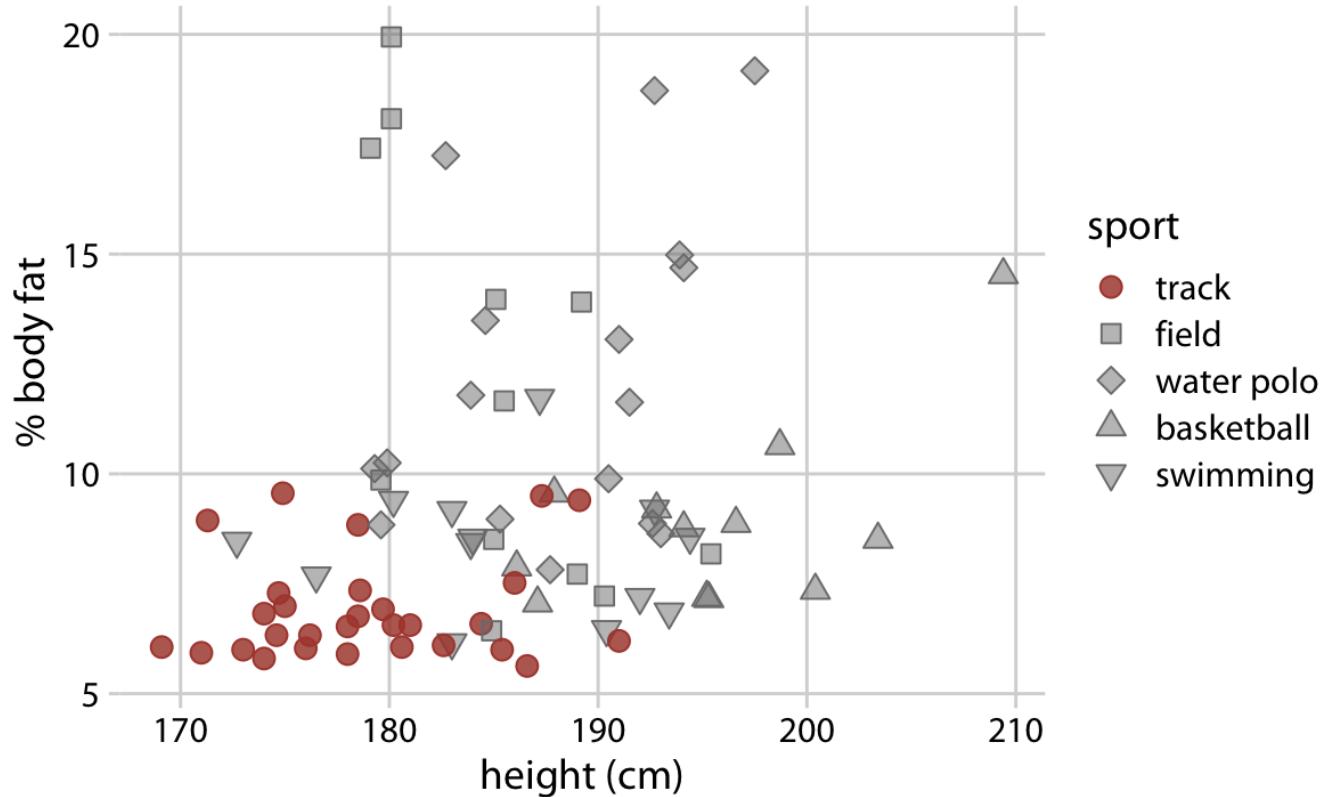
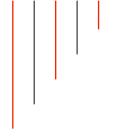


Color to highlight



Important: Baseline colors should not compete for attention.

Color to highlight



Color to highlight



Colorbrewer

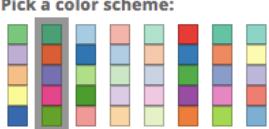
ColorBrewer: Color Advice for Cartography

colorbrewer2.org/#type=qualitative&scheme=Dark2&n=3

Number of data classes: 3

Nature of your data:
 sequential diverging qualitative

Pick a color scheme:



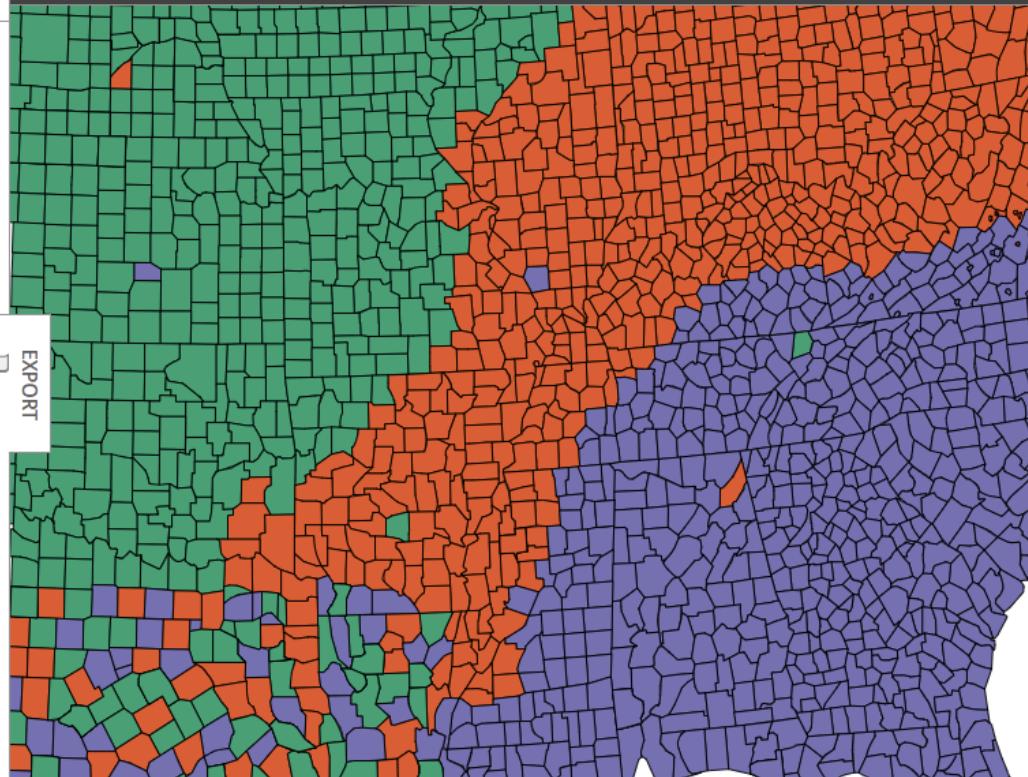
Only show:
 colorblind safe
 print friendly
 photocopy safe

Context:
 roads
 cities
 borders

Background:
 solid color
 terrain

color transparency

3-class Dark2



EXPORT

HEX

#1b9e77
#d95f02
#7570b3

The multiple types of Visualizations



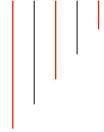
- The same data can be represented with more than one type of visualization

The multiple types of Visualizations



- The same data can be represented with more than one type of visualization
 - Remember: Visualization can be seen as a *map* between *quantifiable features* to *aesthetics*
- Different types of visualizations help you to tell or to discover different stories

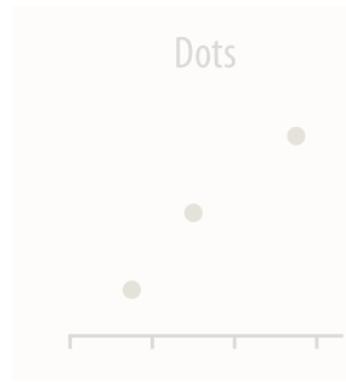
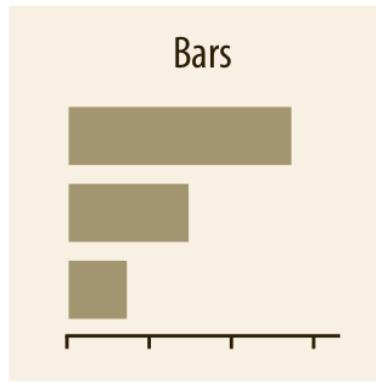
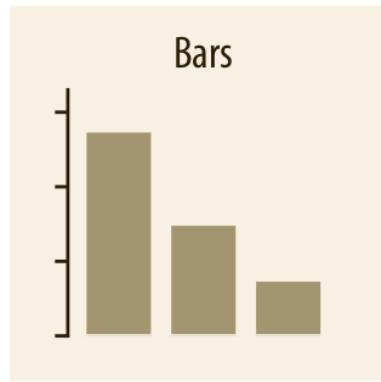
The multiple types of Visualizations



- The same data can be represented with more than one type of visualization
 - Remember: Visualization can be seen as a *map* between *quantifiable features* to *aesthetics*
- Different types of visualizations help you to tell or to discover different stories
 - **Some visualizations can tell misleading stories: you must avoid this!**

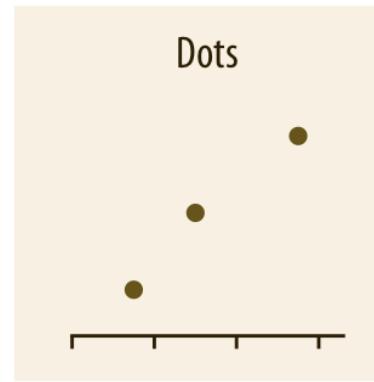
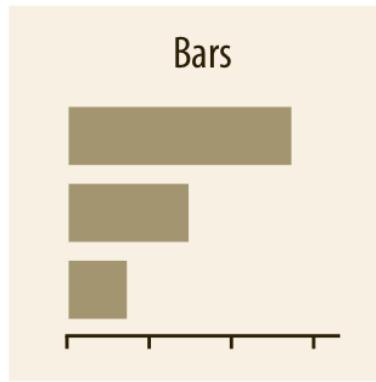
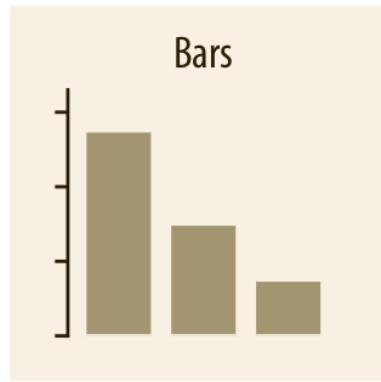
Visualization of Amounts

- Individual group/set



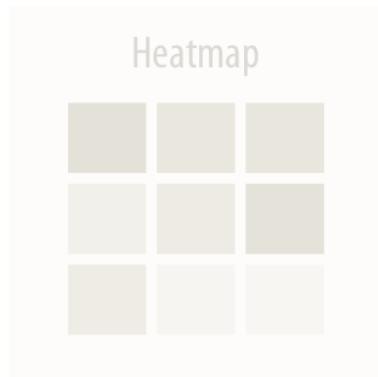
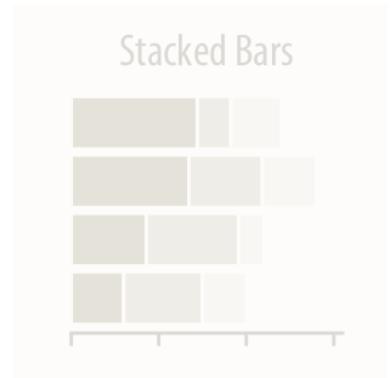
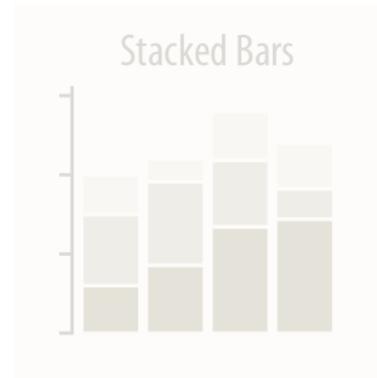
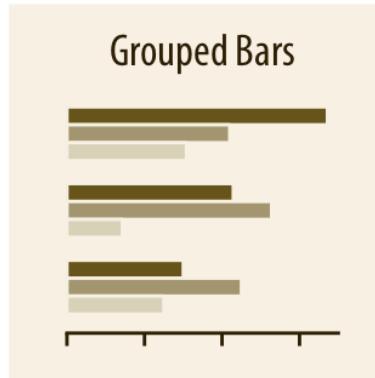
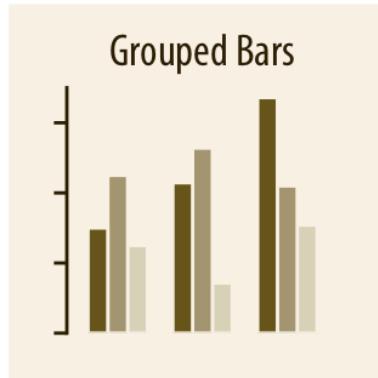
Visualization of Amounts

- Individual group/set



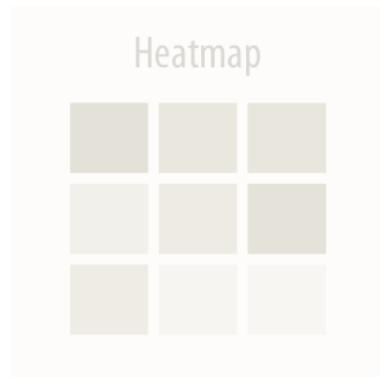
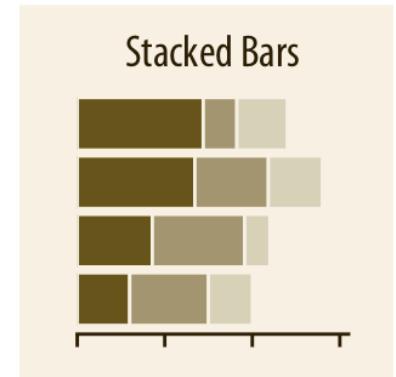
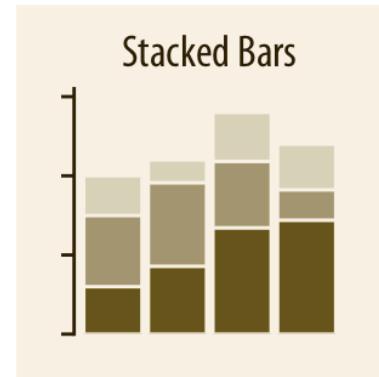
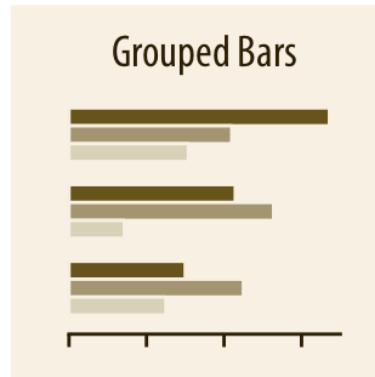
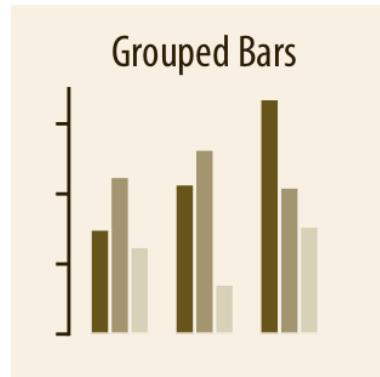
Visualization of Amounts

- Multiple groups/sets



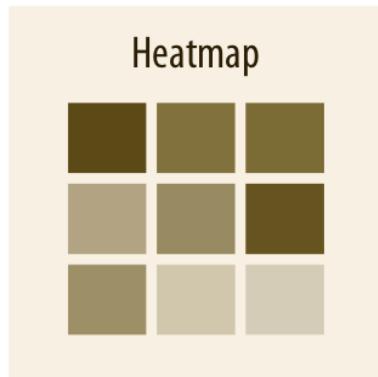
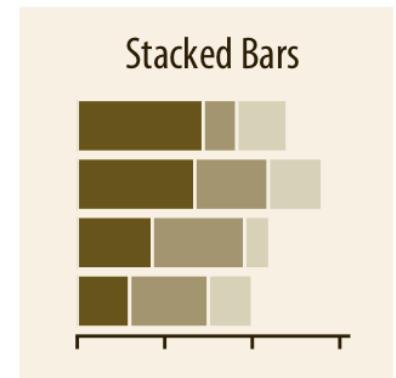
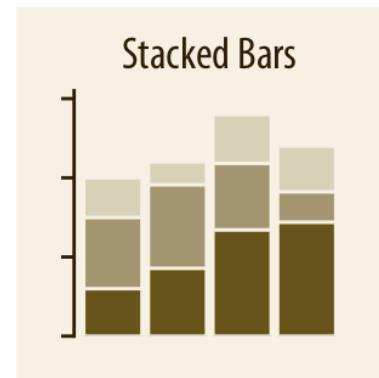
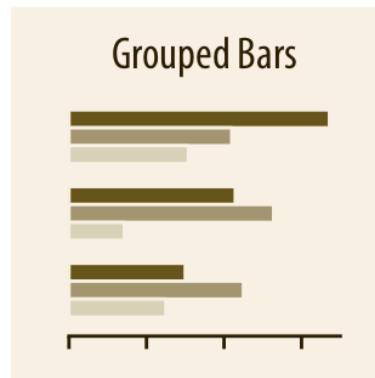
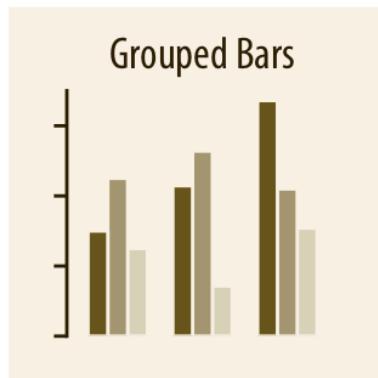
Visualization of Amounts

- Multiple groups/sets



Visualization of Amounts

- Multiple groups/sets

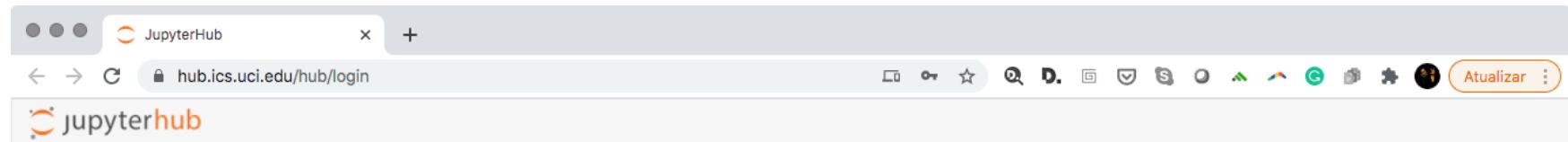


Your first visualization

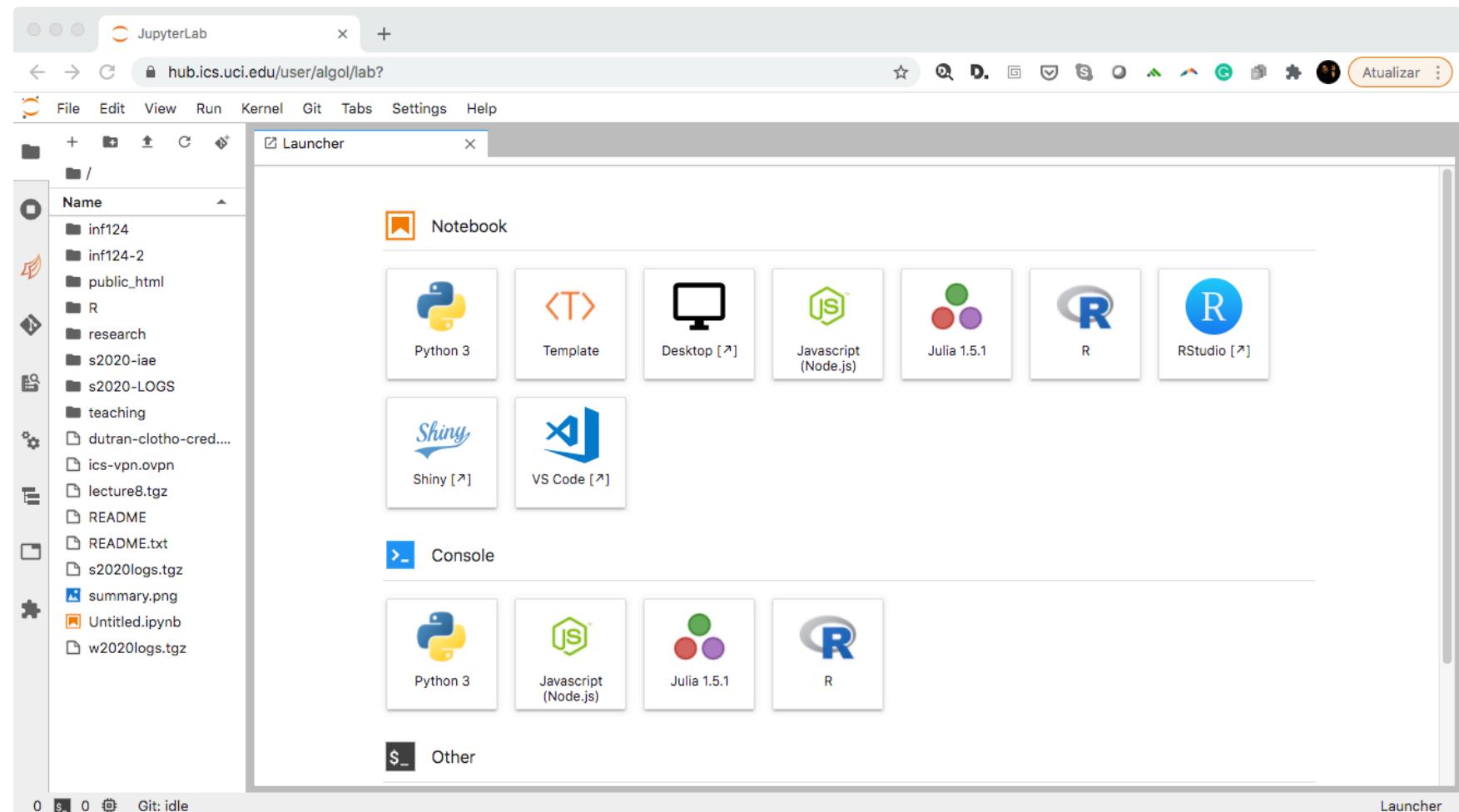


- We will use R in the course
- You can install R and RStudio in your computer, or you can use through the openlab online environment
 - <https://hub.ics.uci.edu/>

Your first visualization



Your first visualization



Your first visualization

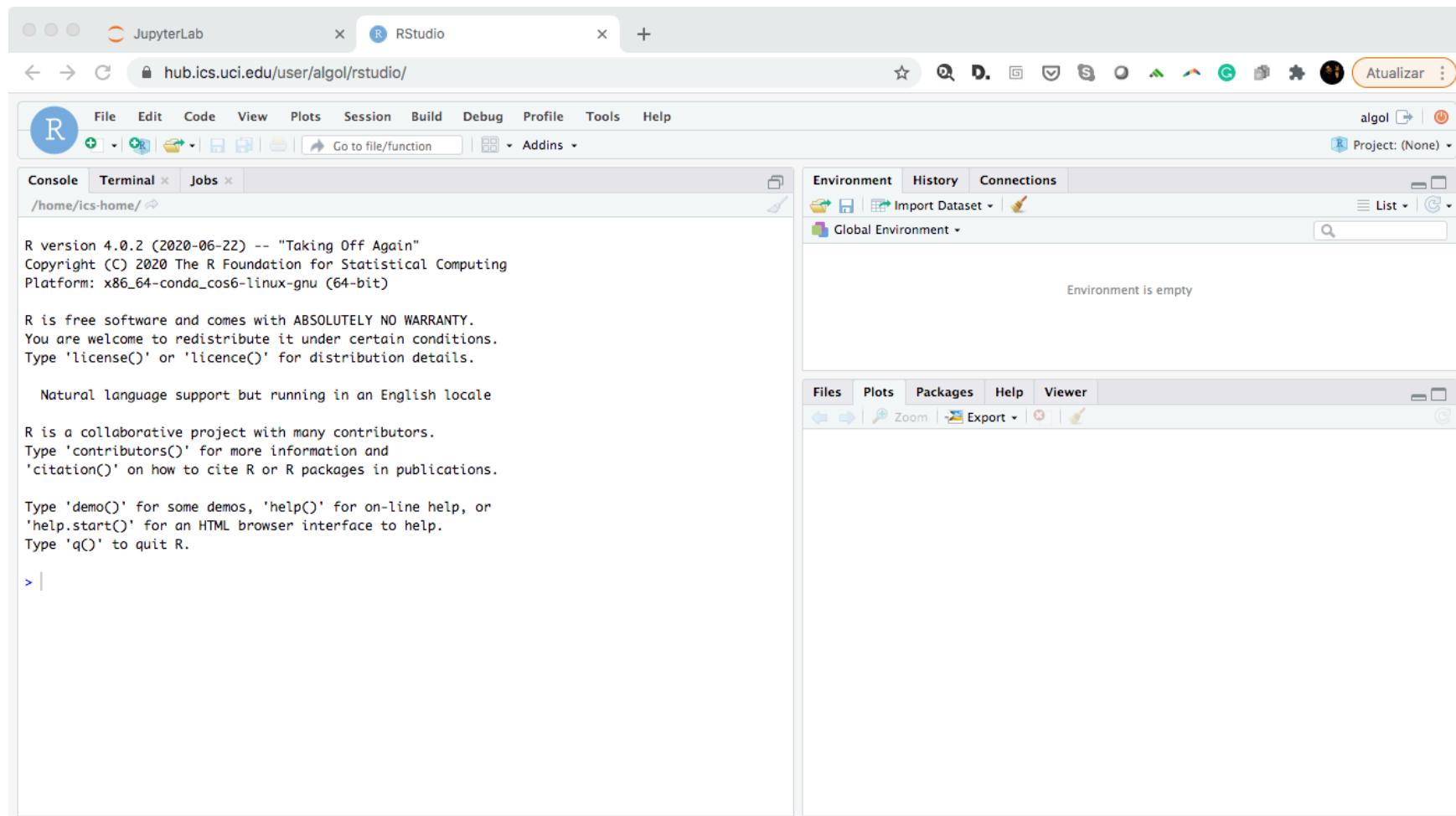
The screenshot shows the RStudio interface running in a web browser window titled "RStudio" at "hub.ics.uci.edu/user/algol/rstudio/". The top navigation bar includes "File", "Edit", "Code", "View", "Plots", "Session", "Build", "Debug", "Profile", "Tools", and "Help". The "Console" tab is selected, displaying the R startup message:

```
R version 4.0.2 (2020-06-22) -- "Taking Off Again"  
Copyright (C) 2020 The R Foundation for Statistical Computing  
Platform: x86_64-conda_cos6-linux-gnu (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.
```

The "Environment" tab in the top right shows an empty global environment. The "Files" tab in the bottom right displays a file browser with the following contents:

Name	Size	Modified
.R	86 B	Dec 17, 2020, 10:38 PM
dutran-clotho-cred.txt	7.5 KB	Dec 16, 2020, 10:06 PM
inf124		
inf124-2		
lecture8.tgz	1.2 KB	Apr 28, 2020, 10:36 PM
public_html		
R		
README	95 B	Jan 7, 2020, 11:20 PM
README.txt	5.9 KB	Jan 7, 2020, 11:20 PM
research		
s2020-lae		
s2020-LOGS		
s2020logs.tgz	31.3 MB	Apr 30, 2020, 6:03 AM

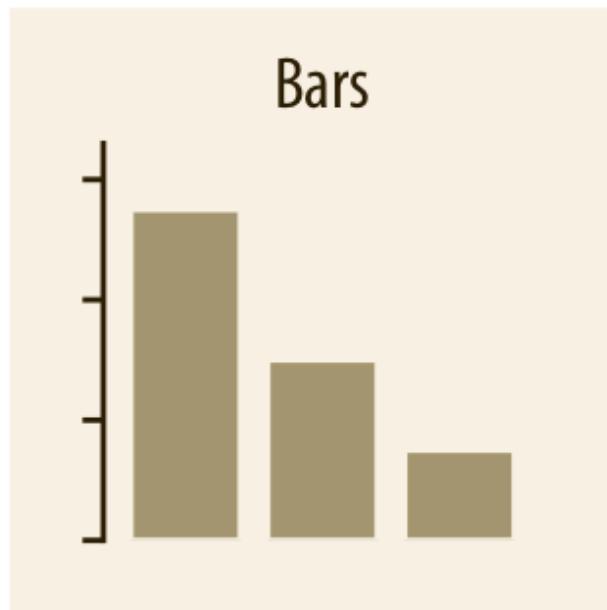
Your first visualization



Your first visualization



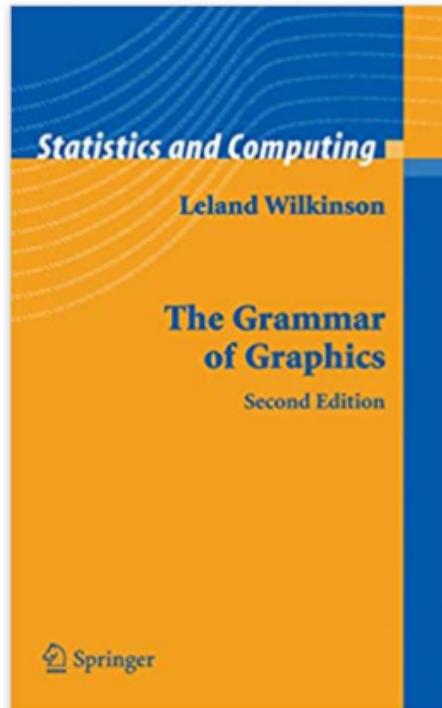
- Possibly the simplest type of visualization



Your first visualization



- We will see how to read data later...
- ggplot2 : R library that introduces a *grammar for graphics*



Your first visualization



- We will see how to read data later...
- ggplot2 : R library that introduces a *grammar for graphics*
- In R, you can import libraries using

`library(...)` or `require(...)`

Your first visualization



- ggplot2 : R library that introduces a *grammar for graphics*
- In R, you can import libraries using
`library(...)` or `require(...)`
- So, we start by importing ggplot2

```
library(ggplot2)
```

Your first visualization



- ggplot2 : R library that introduces a *grammar for graphics*
- So, we start by importing ggplot2

```
library(ggplot2)
```

- Then you can enter some data:

```
myData <- data.frame(name=c("Class A", "Class B", "Class C"),  
                      value=c(8, 10, 15))
```

Your first visualization



- ggplot2 : R library that introduces a *grammar for graphics*
- So, we start by importing ggplot2

```
library(ggplot2)
```

- Then you can enter some data:

```
myData <- data.frame(name=c("Class A", "Class B", "Class C"),  
                      value=c(8, 10, 15))
```

- In a data.frame you can mix data of different data types
- The information can be NA, but the data.frame must be rectangular
(i.e. each row must have the same number of columns)

Your first visualization



- ggplot2 : R library that introduces a *grammar for graphics*
- So, we start by importing ggplot2

```
library(ggplot2)
```

- Then you can enter some data:

```
myData <- data.frame(name=c("Class A", "Class B", "Class C"),  
                      value=c(8, 10, 15))
```

- And create your plot!

```
ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")
```

Your first visualization

The screenshot shows the RStudio interface running in a JupyterLab browser window. The top navigation bar includes tabs for JupyterLab and RStudio, and a URL bar pointing to hub.ics.uci.edu/user/algol/rstudio/. The RStudio menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The R logo icon is in the top-left corner of the main workspace.

Console pane:

```
R version 4.0.2 (2020-06-22) -- "Taking Off Again"  
Copyright (C) 2020 The R Foundation for Statistical Computing  
Platform: x86_64-conda_cos6-linux-gnu (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> library(ggplot2)  
>
```

Environment pane:

Global Environment

Environment is empty

Files pane:

Name	Size	Modified
..		
.R	86 B	Dec 17, 2020, 10:38 PM
dutran-clotho-cred.txt	7.5 KB	Dec 16, 2020, 10:06 PM
inf124		
inf124-2		
lecture8.tgz	1.2 KB	Apr 28, 2020, 10:36 PM
public_html		
R	95 B	Jan 7, 2020, 11:20 PM
README	5.9 KB	Jan 7, 2020, 11:20 PM
README.txt		
research		
s2020-iae		
s2020-locs		

Plots pane:

(No plots are currently displayed)

Your first visualization

The screenshot shows the RStudio interface running in a JupyterLab session. The top navigation bar includes tabs for JupyterLab and RStudio, and a URL bar pointing to hub.ics.uci.edu/user/algol/rstudio/. The main window has several panes:

- Console:** Displays the R startup message, license information, and a command history:

```
R version 4.0.2 (2020-06-22) -- "Taking Off Again"  
Copyright (C) 2020 The R Foundation for Statistical Computing  
Platform: x86_64-conda_cos6-linux-gnu (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.
```

```
> library(ggplot2)
```

```
> myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))
```

```
> |
```
- Global Environment:** Shows a single object named 'myData' with 3 observations and 2 variables.
- Files:** A file browser showing the directory structure of the user's home directory on the server. The contents of the 'ics-home' folder are listed:

Name	Size	Modified
..		
.R	86 B	Dec 17, 2020, 10:38 PM
dutran-clotho-cred.txt	7.5 KB	Dec 16, 2020, 10:06 PM
inf124	1.2 KB	Apr 28, 2020, 10:36 PM
inf124-2		
lecture8.tgz		
public_html		
R	95 B	Jan 7, 2020, 11:20 PM
README	5.9 KB	Jan 7, 2020, 11:20 PM
README.txt		
research		
s2020-iea		
s2020-loca		

Your first visualization

You can inspect your data here

The screenshot shows the RStudio interface running in a browser window. The title bar indicates the tabs are JupyterLab and RStudio, with the URL being hub.ics.uci.edu/user/algol/rstudio/. The RStudio interface includes:

- Console:** Displays the R startup message, license information, and a command history including `library(ggplot2)` and `myData` assignment.
- Environment:** Shows the variable `myData` defined as a data frame with 3 observations and 2 variables.
- Files:** A file browser showing the directory structure under `/home/ics-home/`.

A red arrow points from the text "You can inspect your data here" to the Environment tab in the RStudio interface.

Your first visualization

The screenshot shows the RStudio interface running in a JupyterLab browser window. The top navigation bar includes tabs for 'JupyterLab' and 'RStudio'. The address bar shows the URL 'hub.ics.uci.edu/user/algol/rstudio/'. The RStudio menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The main workspace is divided into several panes:

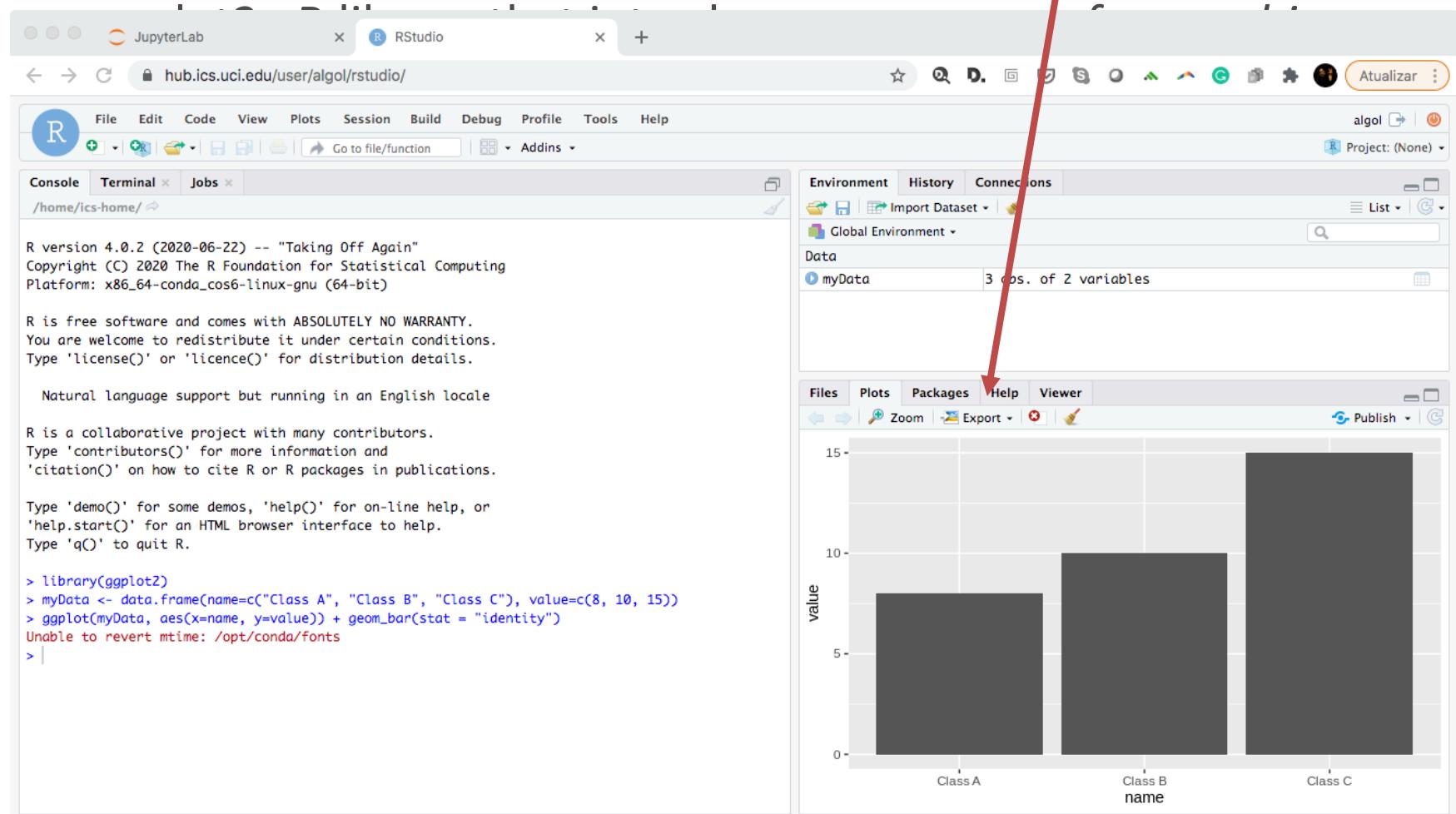
- Console:** Displays the R startup message, license information, and a command history block.
- Environment:** Shows the 'Global Environment' pane with a data frame named 'myData' containing 3 observations and 2 variables.
- Plots:** Displays a bar chart with 'name' on the x-axis ('Class A', 'Class B', 'Class C') and 'value' on the y-axis (0, 5, 10, 15). The bars have values of approximately 8, 10, and 15 respectively.

Code in the Console pane:

```
> library(ggplot2)
> myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))
> ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")
Unable to revert mtime: /opt/conda/fonts
>
```

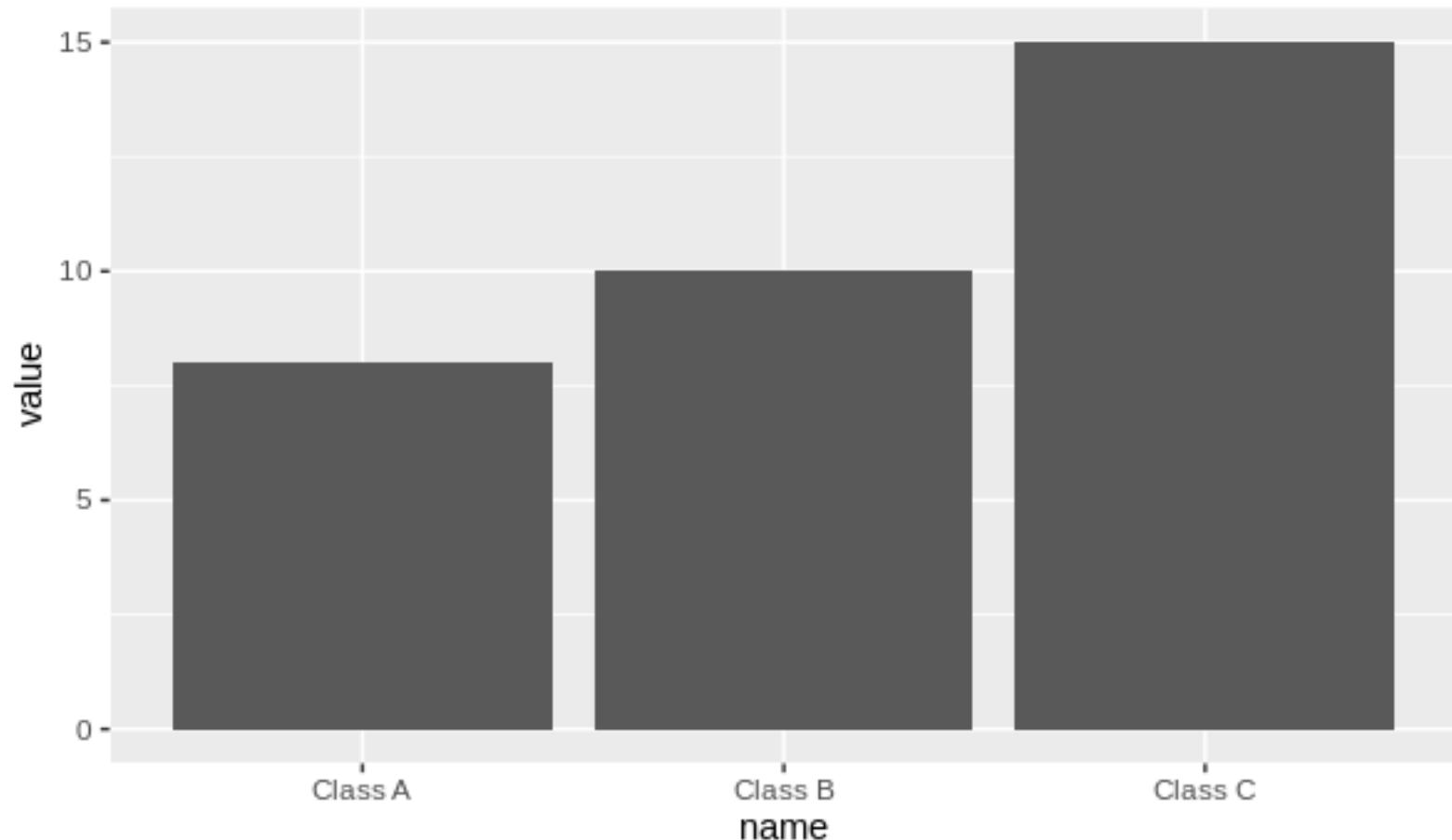
Your first visualization

You can interactively export the plot



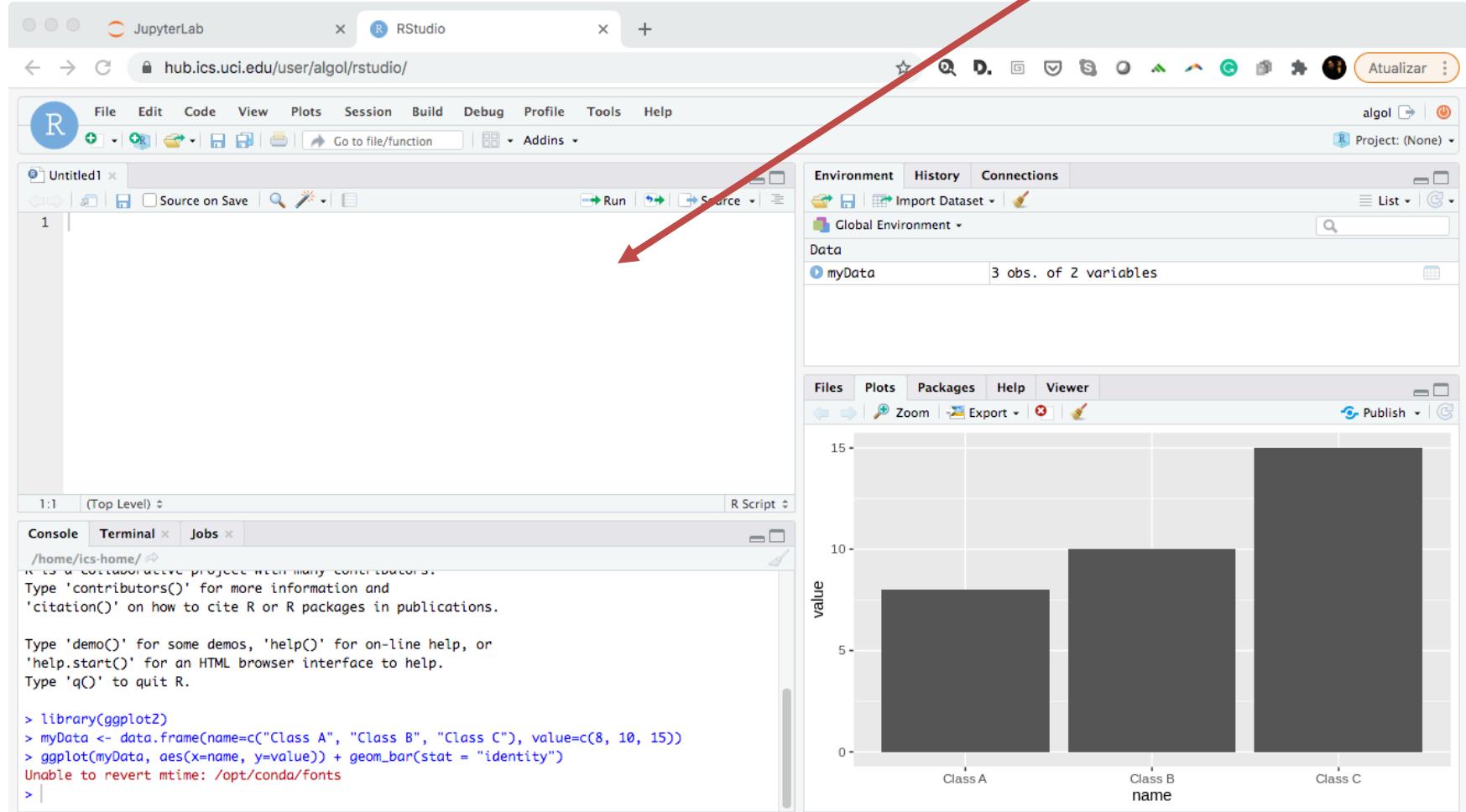
Your first visualization

- Simple but already respectable bar chart



Your first visualization

You can code in here



Your first visualization

You can code in here, save and run

The screenshot shows the RStudio interface running within a JupyterLab environment. On the left, the RStudio interface includes:

- A top bar with tabs for "JupyterLab" and "RStudio".
- A browser header showing "hub.ics.uci.edu/user/algol/rstudio/".
- An "Lecture-3-v1.R" script editor containing R code to create a bar chart.
- A "Console" tab showing the execution of the R code and its output.

The right side of the interface contains the RStudio environment:

- The "Environment" pane shows a global environment with a dataset named "myData" containing 3 observations and 2 variables.
- The "Files" pane shows the directory structure of the user's home folder, including subfolders like "public_html" and "R", and files like "README", "summary.png", and "myFirstBarVisualization.pdf".

Your first visualization

You can direct the plot to a file also

The screenshot shows an RStudio interface with a red arrow pointing from the code editor to the file browser.

Script Editor (Lecture-3-v1.R):

```
1 #  
2 # The first visualization  
3 #  
4 library(ggplot2)  
5 #  
6 # Data to be visualized  
7 myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))  
8 #  
9 # Map the data using a bar chart  
10 # stat = "identity" ensures that the y value is mapped exactly into the visualization  
11 ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")  
12 #  
13 # Save to a file|  
14 ggsave("myFirstBarVisualization.pdf", width = 6, height = 4)  
15
```

Console:

```
/home/ics-home/  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> library(ggplot2)  
> myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))  
> ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")  
Unable to revert mtime: /opt/conda/fonts  
> source('/home/ics-home/inf143/Lecture-3-v1.R')  
> source('/home/ics-home/inf143/Lecture-3-v1.R')  
>
```

File Browser:

Name	Size	Modified
public_html		
R		
README	95 B	Jan 7, 2020, 11:20 PM
README.txt	5.9 KB	Jan 7, 2020, 11:20 PM
research		
s2020-iea		
s2020-LOGS		
s2020logs.tgz	31.3 MB	Apr 30, 2020, 6:03 AM
summary.png	208.8 KB	Apr 9, 2020, 9:08 PM
teaching		
Untitled.ipynb	72 B	Nov 9, 2020, 11:49 PM
w2020logs.tgz	326.5 MB	Apr 30, 2020, 6:02 AM
inf143		
myFirstBarVisualization.pdf	4.4 KB	Jan 12, 2021, 2:57 PM

Your first visualization

You can inspect the plot directly in the browser if you wish to; just click!

The screenshot shows the RStudio interface with the following components:

- Script Editor (Lecture-3-v1.R):** Contains R code for creating a bar chart.
- Environment:** Shows the global environment with a dataset named `myData`.
- File Viewer:** Displays the contents of the `home/ics-home` directory, including files like `public_html`, `R`, and `myFirstBarVisualization.pdf`.
- Console:** Shows the output of running the R script, including the creation of the PDF file.

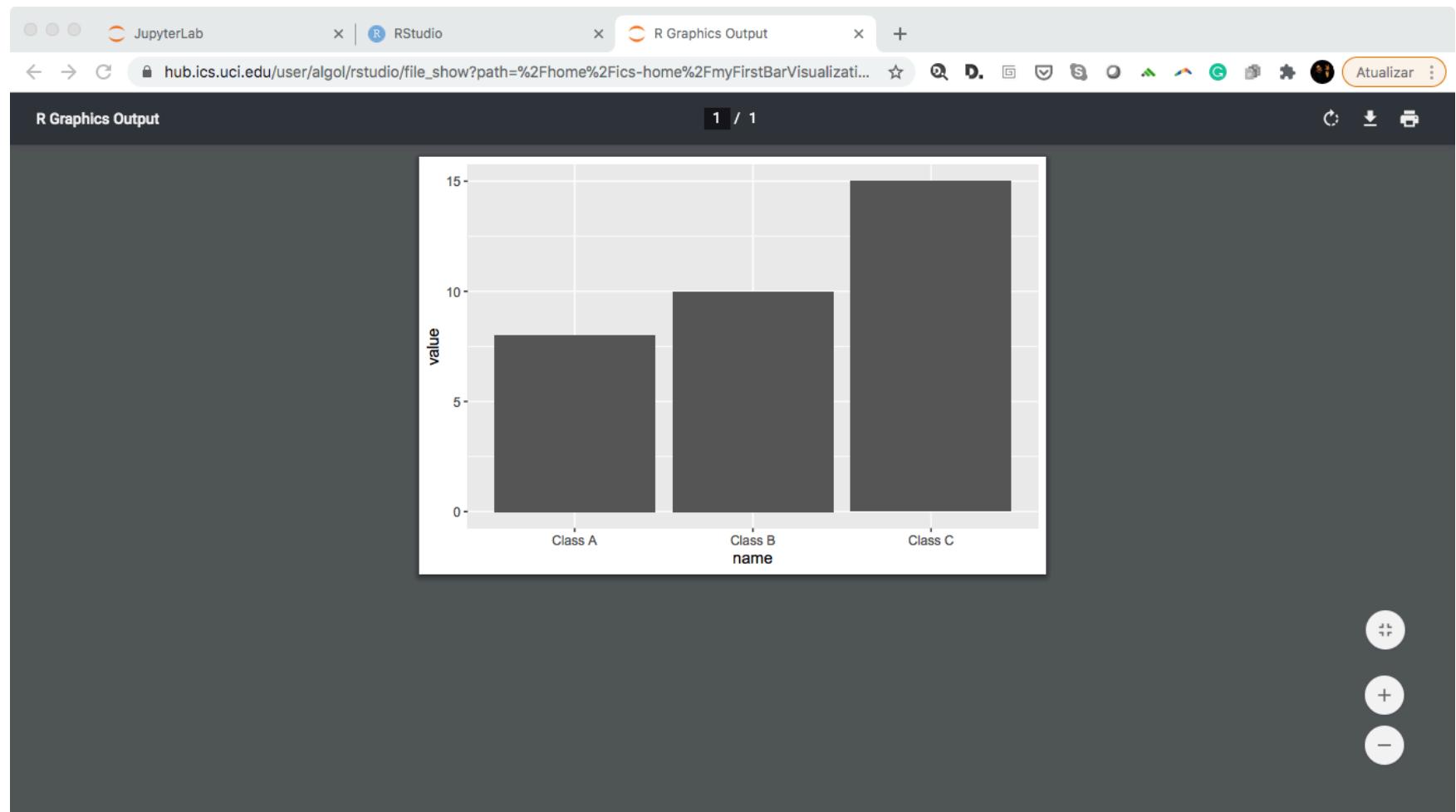
A red arrow points from the text "You can inspect the plot directly in the browser if you wish to; just click!" to the PDF file in the File Viewer pane.

```
1 #  
2 # The first visualization  
3 #  
4 library(ggplot2)  
5 #  
6 # Data to be visualized  
7 myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))  
8 #  
9 # Map the data using a bar chart  
10 # stat = "identity" ensures that the y value is mapped exactly into the visualization  
11 ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")  
12 #  
13 # Save to a file|  
14 ggsave("myFirstBarVisualization.pdf", width = 6, height = 4)  
15
```

```
13:17 (Top Level) R Script  
Console Terminal x Jobs x  
/home/ics-home/  
Execution of 'R' on how to detect or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> library(ggplot2)  
> myData <- data.frame(name=c("Class A", "Class B", "Class C"), value=c(8, 10, 15))  
> ggplot(myData, aes(x=name, y=value)) + geom_bar(stat = "identity")  
Unable to revert mtime: /opt/conda/fonts  
> source('/home/ics-home/inf143/Lecture-3-v1.R')  
> source('/home/ics-home/inf143/Lecture-3-v1.R')  
>
```

Name	Size	Modified
public_html		
R	95 B	Jan 7, 2020, 11:20 PM
README	5.9 KB	Jan 7, 2020, 11:20 PM
README.txt		
research		
s2020-iea		
s2020-LOGS		
s2020logs.tgz	31.3 MB	Apr 30, 2020, 6:03 AM
summary.png	208.8 KB	Apr 9, 2020, 9:08 PM
teaching		
Untitled.ipynb	72 B	Nov 9, 2020, 11:49 PM
w2020logs.tgz	326.5 MB	Apr 30, 2020, 6:02 AM
inf143		
myFirstBarVisualization.pdf	4.4 KB	Jan 12, 2021, 2:57 PM

Your first visualization



Your first visualization

