Data Wrangling and Data Analysis
Data Visualization (Creating Dashboards)

Hakim Qahtan

Department of Information and Computing Sciences

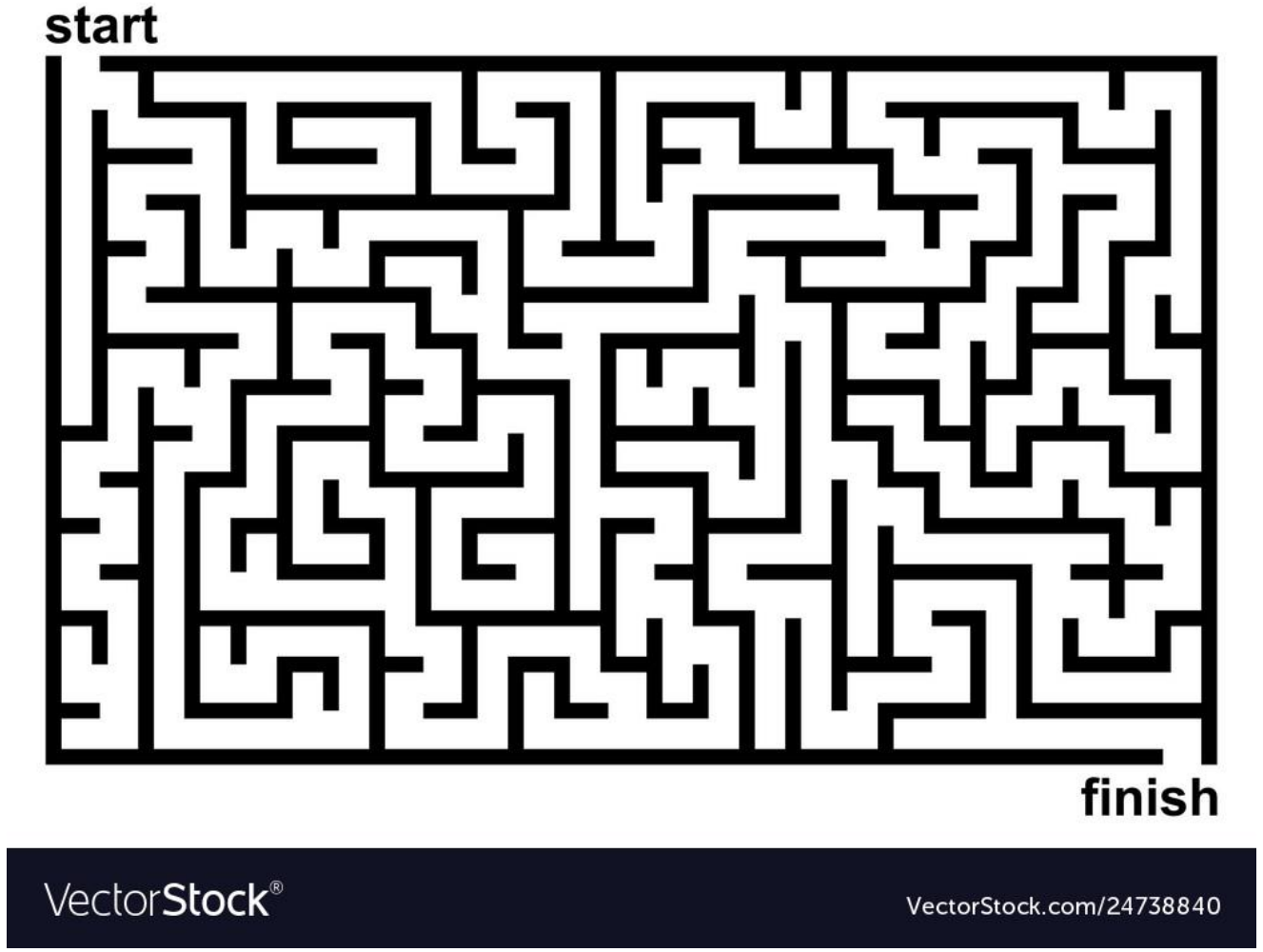
Utrecht University

Topics for Today

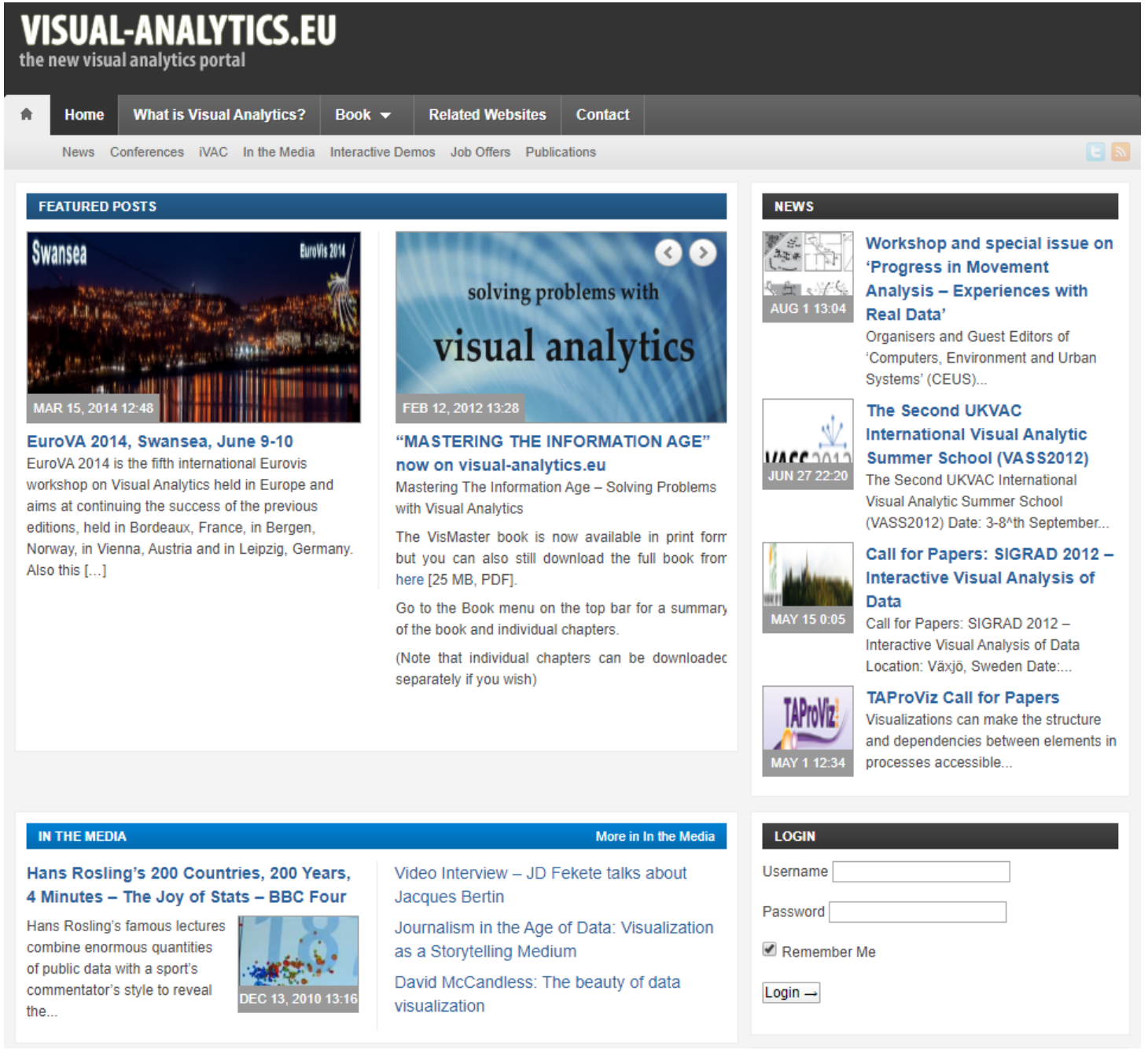
- Introduction
- Mistakes in Data Visualization
- Examples of Visualization Tools for Data Analytics
- Different Ways to Represent Data
- Designing a Dashboard

Introduction

- Danger of getting lost in data, which may be:
  - Irrelevant to the current task in hand
  - Processed in an inappropriate way
  - Presented in an inappropriate way



Introduction (Cont.)

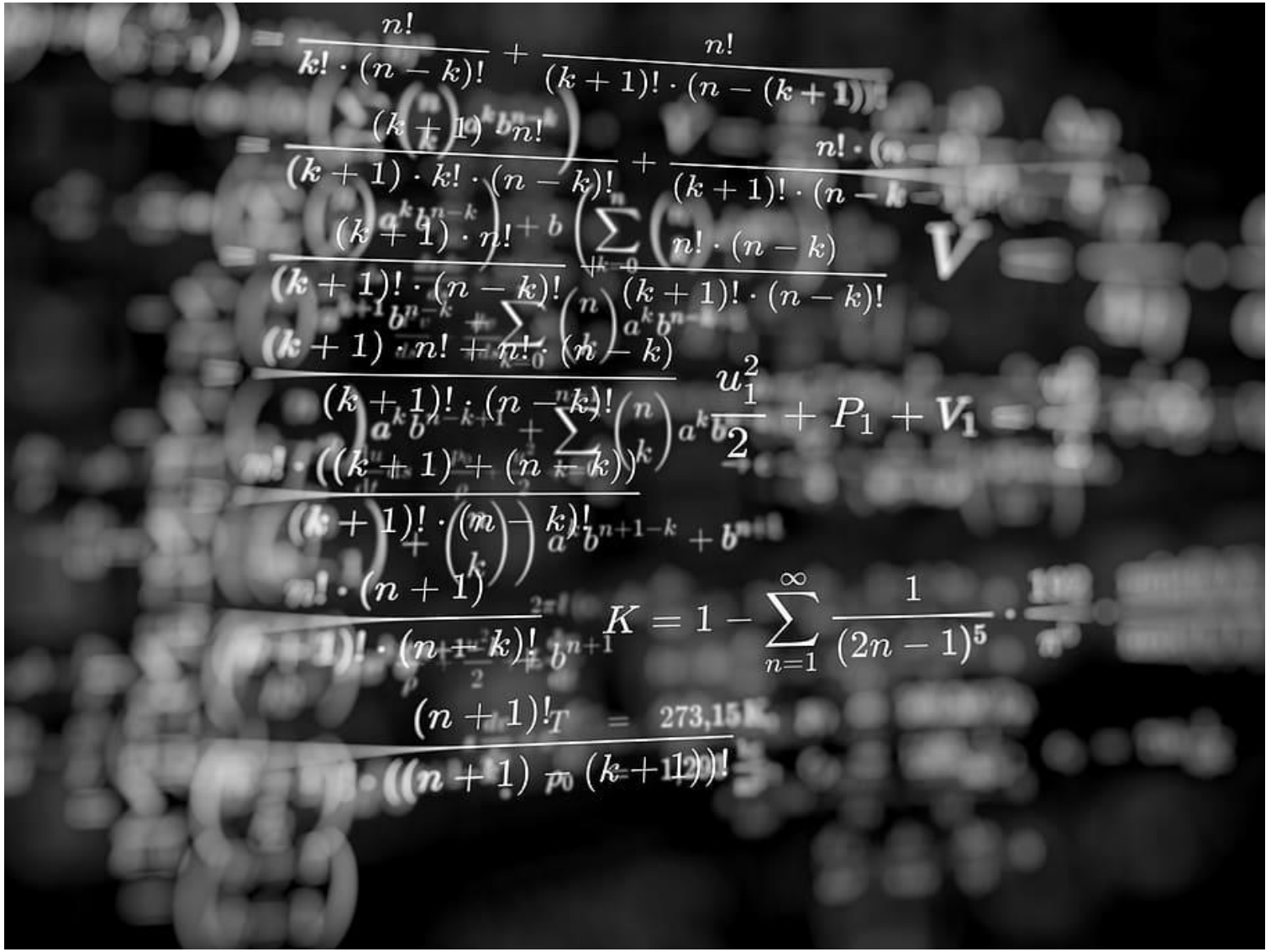


Introduction (Cont.)

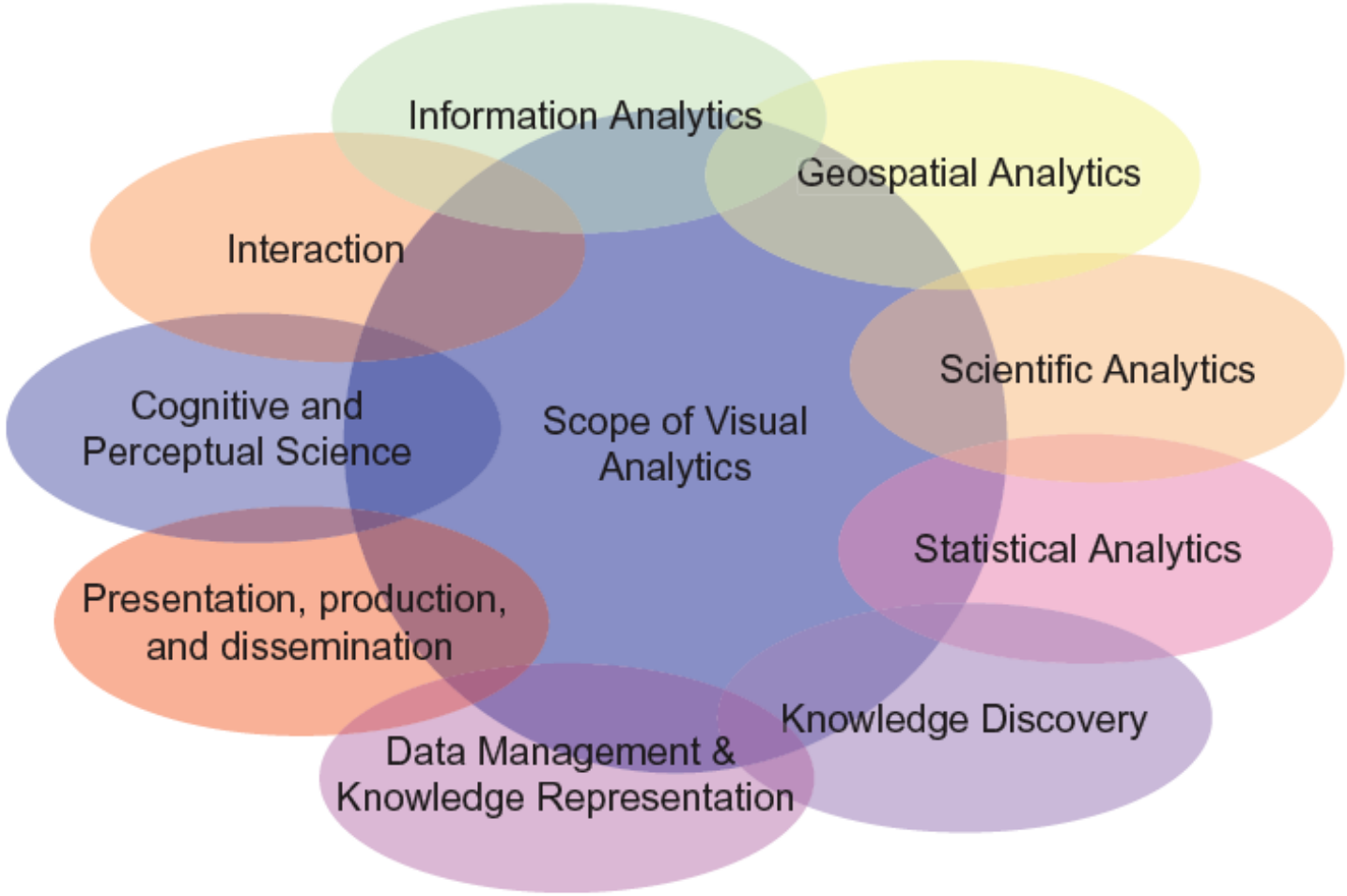
- Good graphics
  - Point relationships, trends or patterns
  - Explore data to infer new things
  - To make something easy to understand
  - To observe a reality from different viewpoints
  - To achieve an idea to be memorized

Introduction (Cont.)

- Data visualization is a way of expressing
  - Like math, music, drawing or writing
- So, it has some rules to respect



Introduction (Cont.)



Keim, D.A., Mansmann, F., Schneidewind, J. and Ziegler, H., Challenges in Visual Data Analysis, Tenth International Conference on Information Visualisation (IV'06), pp. 9-16, 2006.

Mistakes in Data Visualization

Mistakes in visualization

If it's an image that displays and explains information quickly and clearly, it's an infographic. But we've collected some that are head-craning, eye-squinting, eyebrow-raising nightmares that leave you more confused than before you clicked 'next'. The result is an exciting gallery of infographics that tell you nothing.

If you manage to understand all 16 of the infographics in this gallery, write to data@theguardian.com and we'll pick one lucky person to win our incredible prize (the prize is a ruler and a ball-point pen).

Mona Chalabi

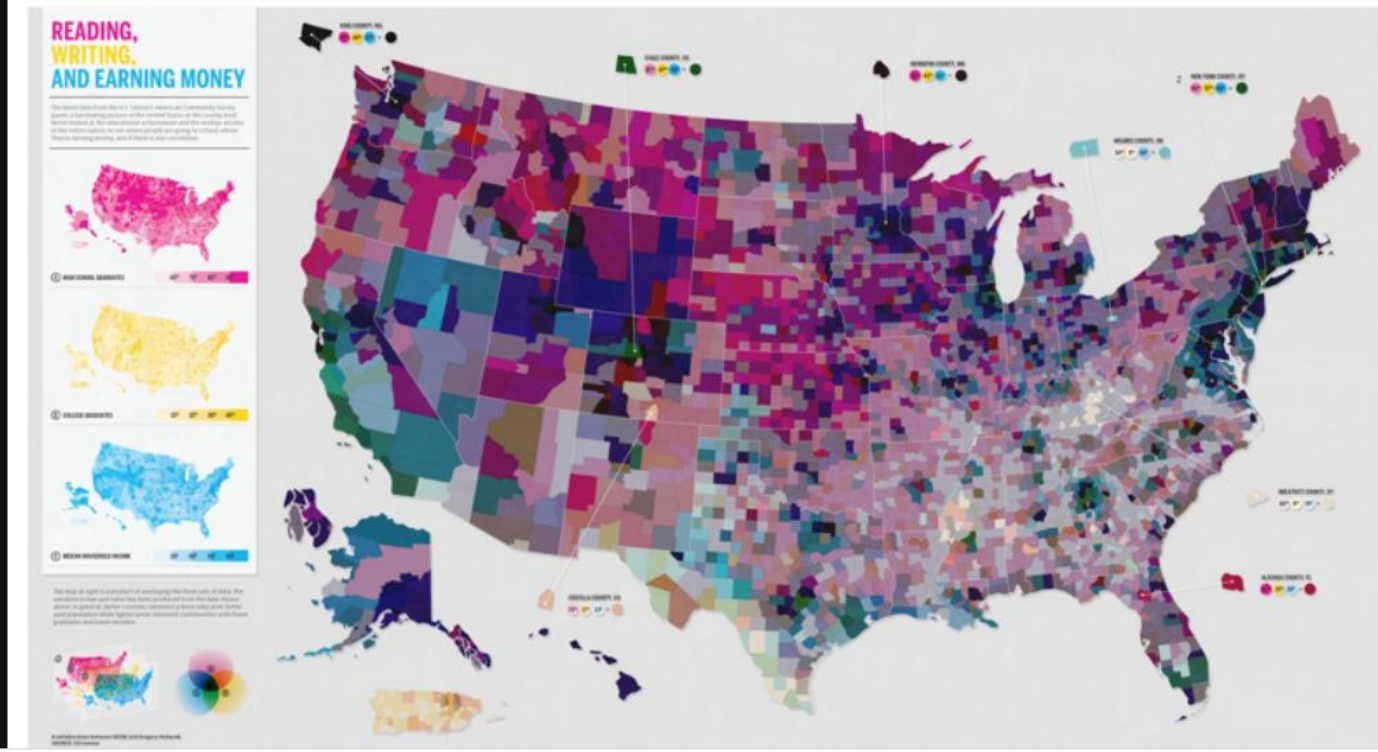
@MonaChalabi Email

Thu 1 Aug 2013 12.55 BST

939 131

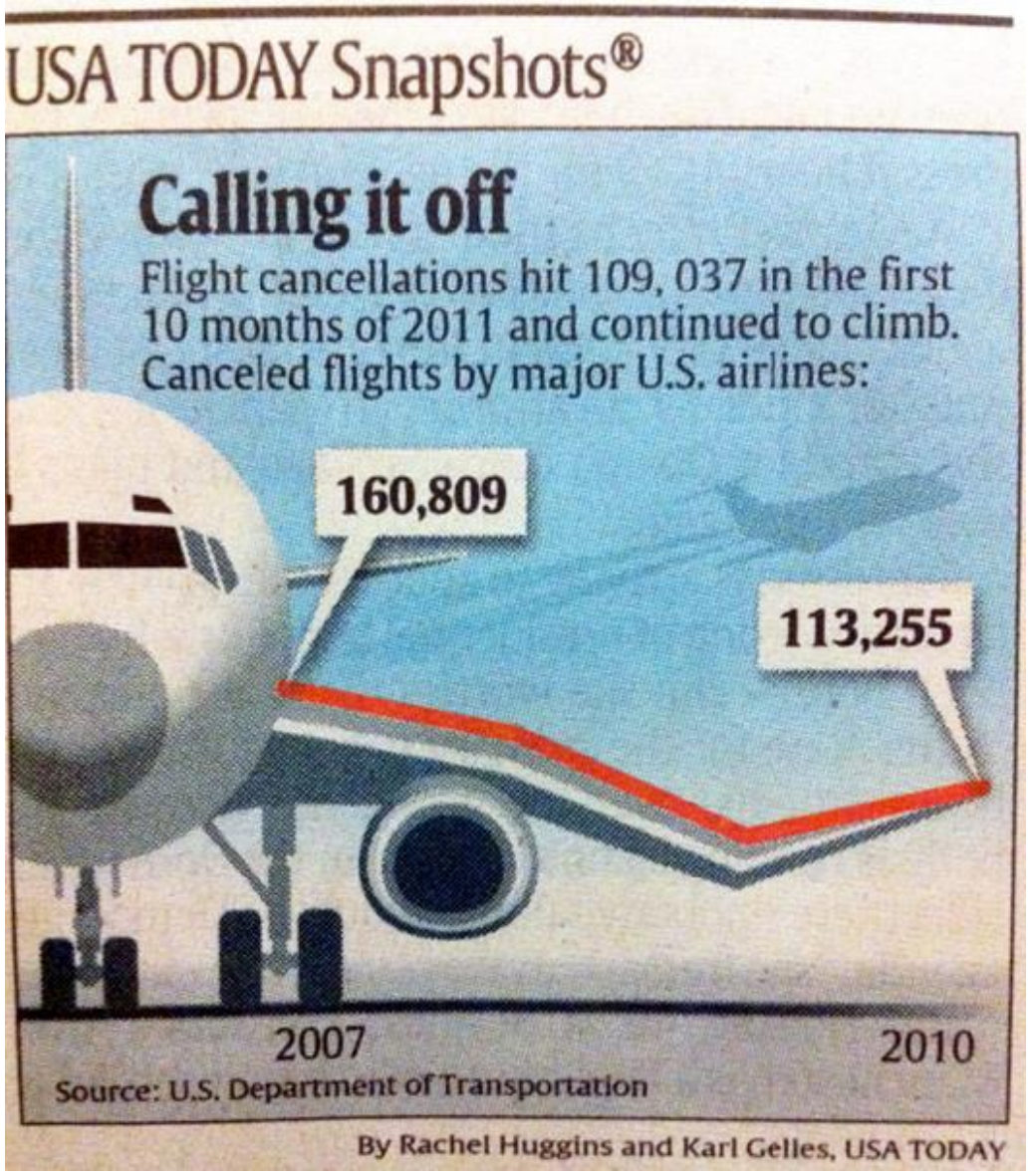
It's quite ironic that you need a map to navigate this map of the US. The key in the bottom-left corner explains that the colours are based on three maps overlayed. Oh! Right! We get it now! (We don't really, it's just embarrassing to admit sometimes).

Photograph: GOOD/Gregory Hubacek



https://www.theguardian.com/news/datablog/gallery/2013/aug/01/16-useless-infographics

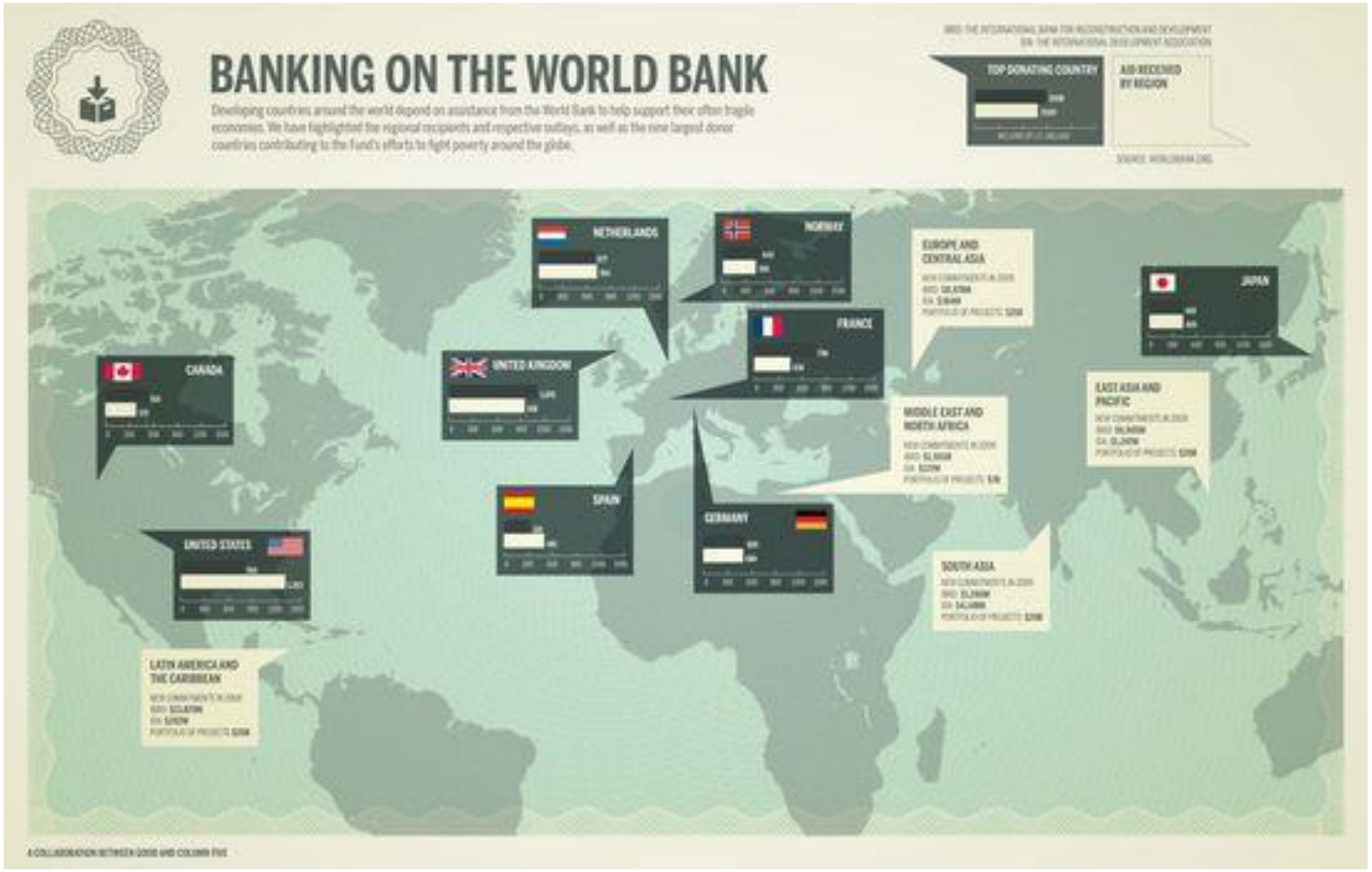
Mistakes in Data Visualization (Cont.)



- Problems?

Mistakes in Data Visualization (Cont.)

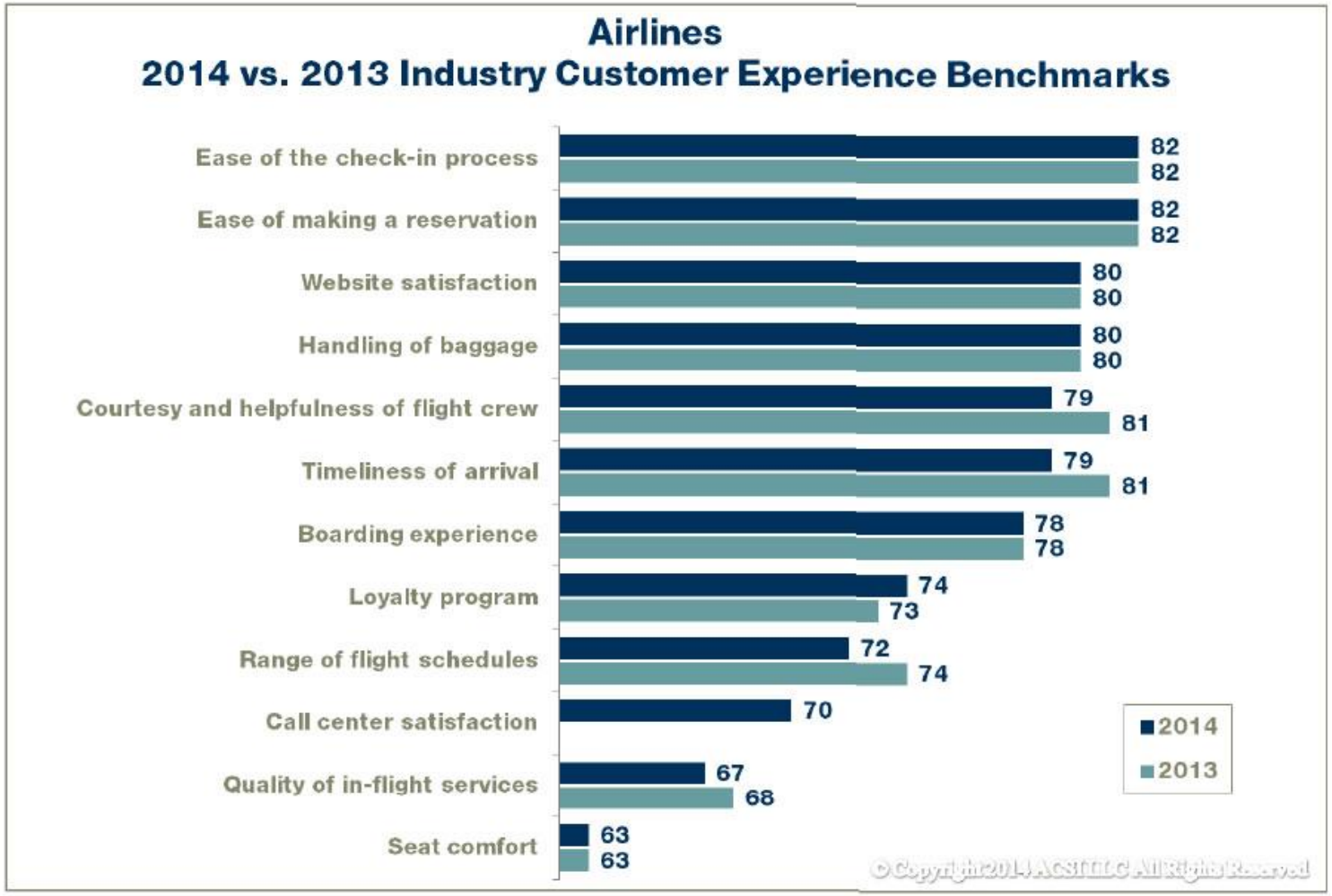
- Problems?
- Try to identify:
  1. The biggest donor in 2008
  2. The smallest donor in 2009
  3. The variation between 200 and 2009
  4. Which region received the biggest amount of money



https://blogs.elpais.com/.a/6a00d8341bfb1653ef0153903125d9970b-550wi

Mistakes in Data Visualization (Cont.)

- Problems?
  - Bar values should start at zero



https://www.qualitydigest.com/IQedit/Images/Articles_and_Columns/2014/April_2014/ACSI/airlines-lg.jpg

Tools for Data Analytics

Requirements for DA Visualization Tools

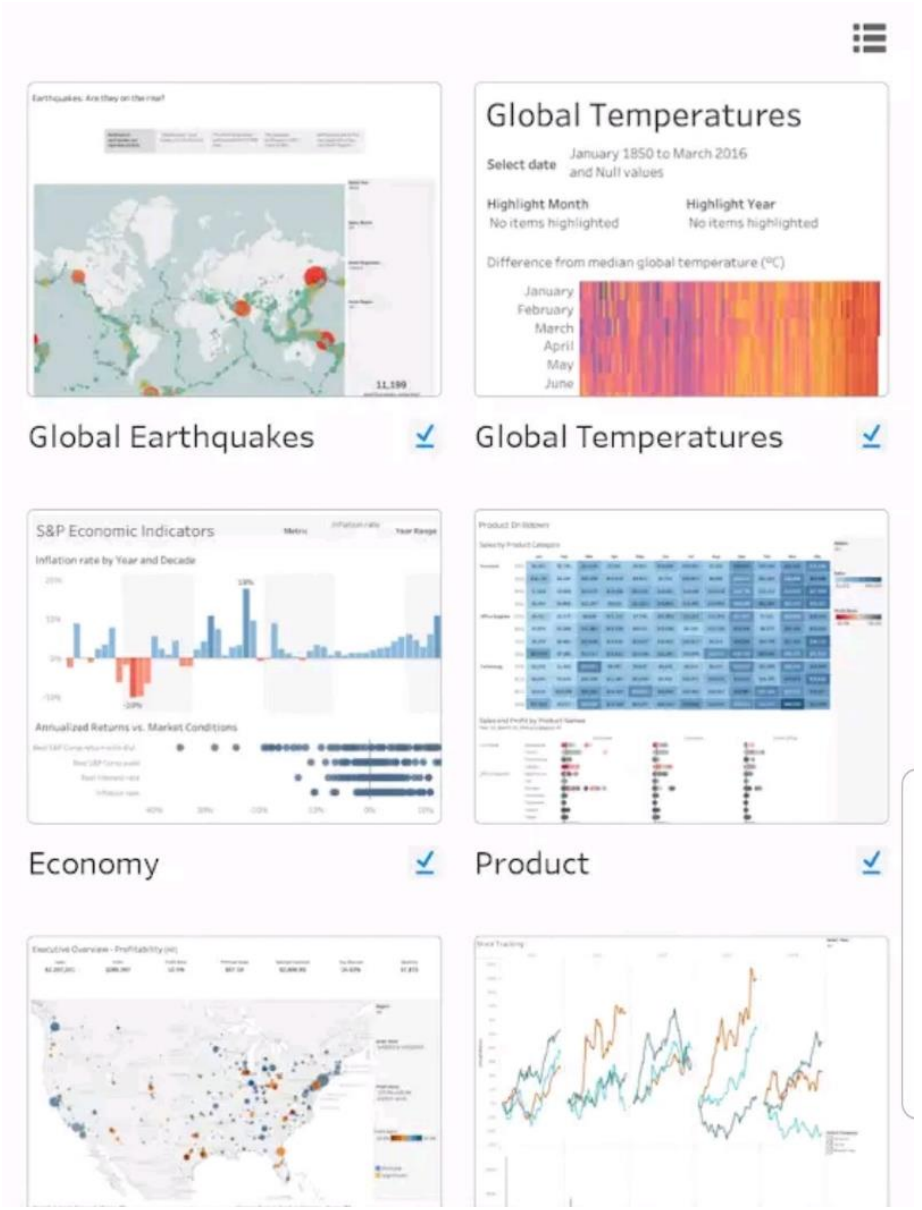
- Ability to choose different visuals and graphs
  - Choose between pie charts, bar charts, and line graphs to express the data better
- Accurate trend tracking capability
  - Keep track of the interesting trends in your data
  - Ex.: you can see which products are selling each month, and which demographic is buying
- High level of security
  - Without proper security, you could risk losing information that could be crucial to your enterprise
- Simplified and self-explainable software interfacing

Tableau

- Connect to and aggregate data.
  - Flat files, databases
  - Perform joining and clustering
- Create insightful analysis.
  - Explore data, create good looking charts and dashboards online
  - Drag and drop functionality
  - Tell compelling stories with your data
  - Write custom calculations within your visualizations
  - Set up filters

Tableau (Cont.)

- Free for one data source which should be made public.
- Requires purchase of a license for private use
  - 70 USD per month.



D3

Search or jump to... / Pull requests Issues Marketplace Explore

d3 / d3

Used by 82.6k | Watch 4k | Star 90.7k | Fork 21.9k

Code | Issues 3 | Pull requests 1 | Actions | Wiki | Security | Insights

Gallery

Mike Bostock edited this page 26 days ago · 1295 revisions

Wiki ▸ Gallery

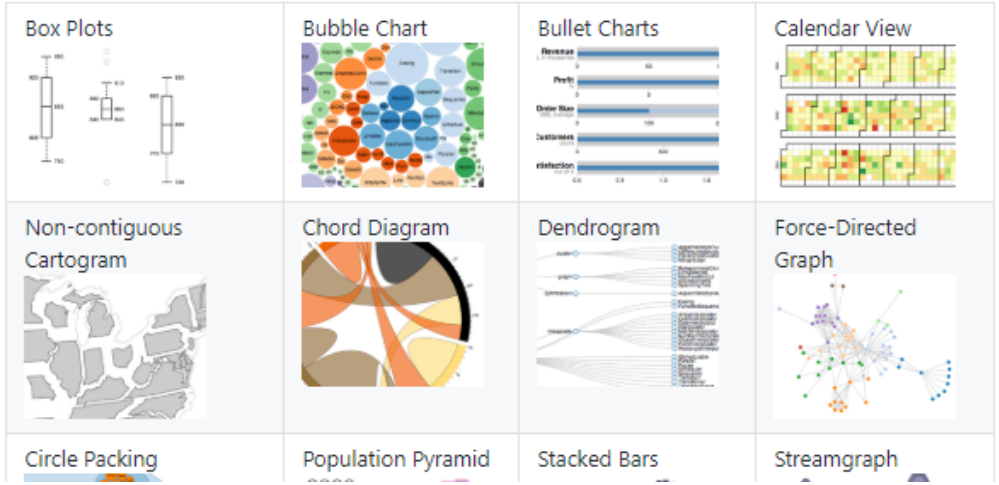
Pages 62

Please see the new D3 Gallery on Observable.

Many examples below are broken or use older versions of D3. We're in the process of cleaning up this gallery, and writing new examples and tutorials. Thanks for your patience!

Welcome to the **D3 gallery**! More examples are available for forking on Observable; see D3's profile and the visualization collection. Please share your work on Observable, or tweet us a link!

Visual Index



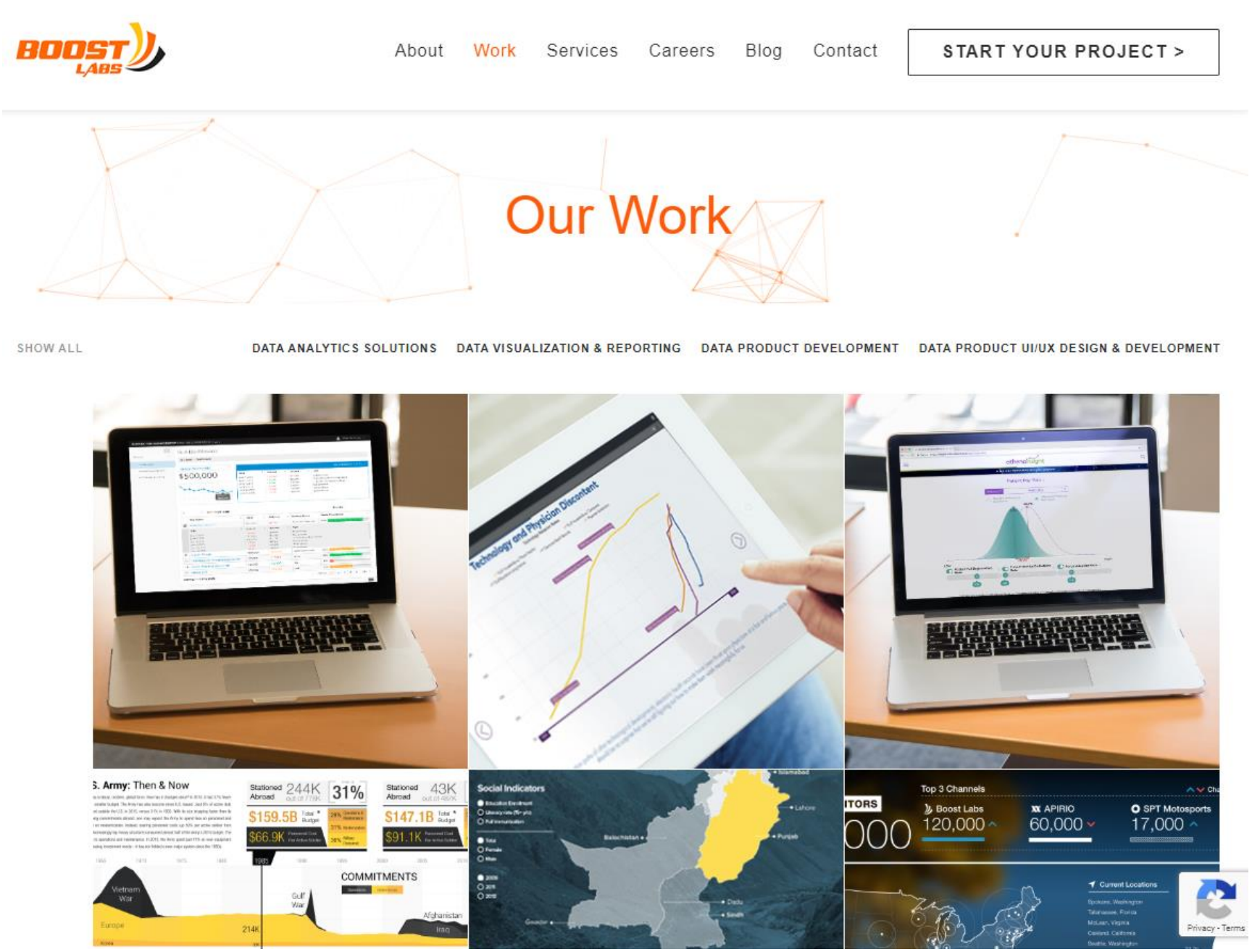
Data-Driven Documents

- Home
- Introduction
- Examples
- Tutorials
- Plugins

Help

- Stack Overflow
- Slack (Invite)
- Google Group
- Gitter

IBM Many Eyes



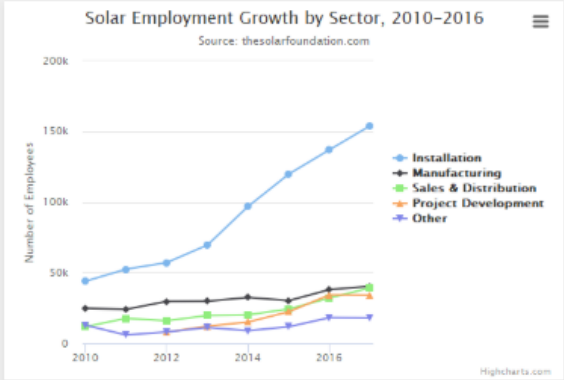
Highcharts

HIGHCHARTS PRODUCTS DEMO DOCS SUPPORT BLOG COMMUNITY BUY JOIN OUR TEAM ABOUT US

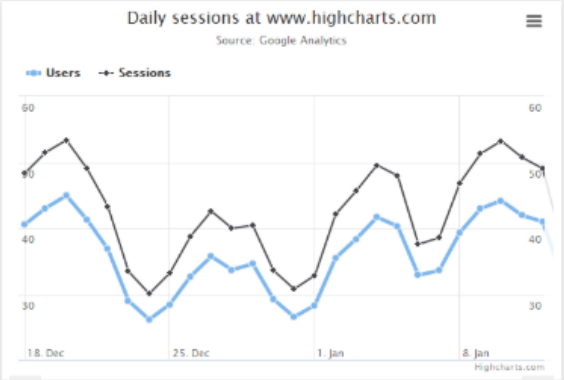
Highcharts Demos

DEFAULT | DARK UNICA | SAND SIGNIKA | GRID LIGHT

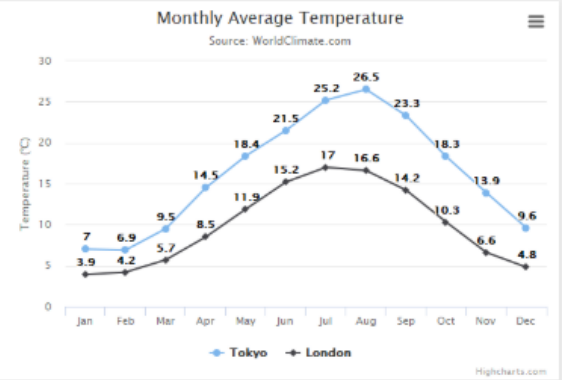
Line charts



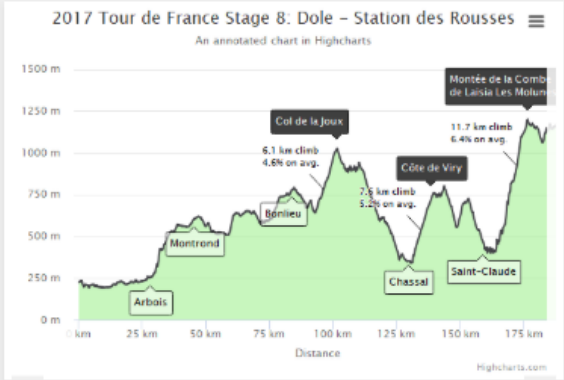
Basic line



Ajax loaded data, clickable points

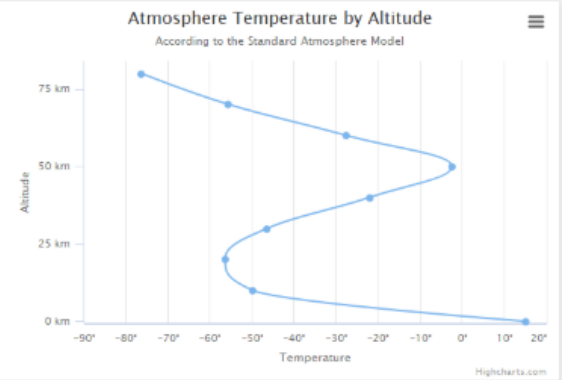


With data labels



With annotations

Time series, zoomable



Spline with inverted axes

Google Charts

Google Charts

Search

Language

Home | Guides | Reference | Support | Send fee

Overview

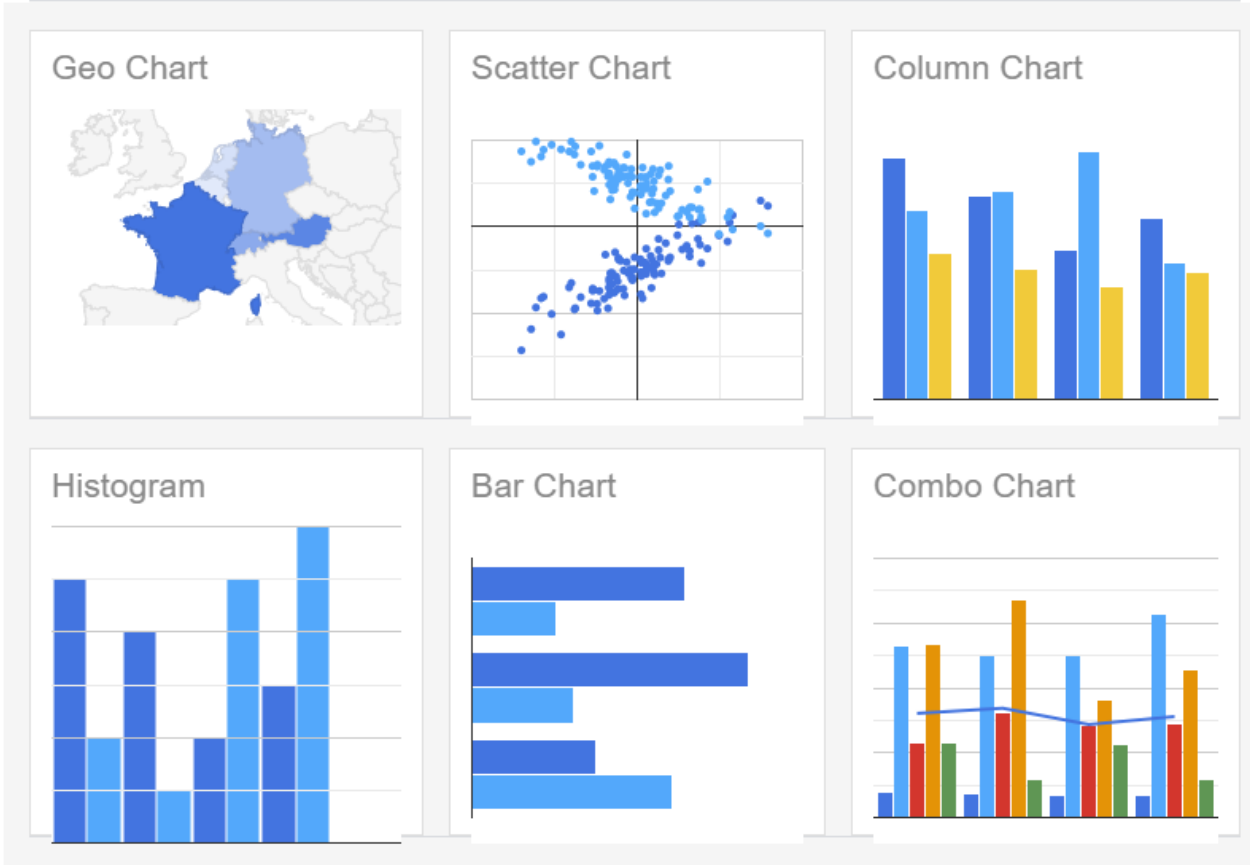
Hello, Charts!
Quickstart
Load the Charts Library
Prepare the Data
Customize the Chart
Draw the Chart
Draw Multiple Charts

Chart Types
Chart Gallery
Annotation Charts
Area Charts
Bar Charts
Bubble Charts
Calendar Charts
Candlestick Charts
Column Charts
Combo Charts
Diff Charts
Donut Charts
Gantt Charts
Gauge Charts
GeoCharts
Histograms
Intervals
Line Charts
Maps
Org Charts
Pie Charts

Chart Gallery

Our gallery provides a variety of charts designed to address your data visualization needs. These charts are based on pure HTML5/SVG technology (adopting VML for old IE versions), so no plugins are required. All of them are interactive, and many are pannable and zoomable. Adding these charts to your page can be done in a few simple steps.

Some additional community-contributed charts can be found on the Additional Charts page.



NodeXL

CodePlex Archive Open Source Project Archive

NodeXL

download archive

NodeXL: Network Overview, Discovery and Exploration for Excel

NodeXL is an Excel 2007/2010/2013 template for viewing network graphs, with a set of .NET Framework 3.5 class libraries that can be used to add network graphs

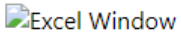
home issues discussions

NODEXL

A project from the socialmedia RESEARCH FOUNDATION

Donate to the Social Media Research Foundation's NodeXL Development Fund

NodeXL is the MSPaint of Networks.



NodeXL Basic is a free, open-source template for Microsoft® Excel® 2007, 2010, 2013 and 2016 that makes it easy to explore network graphs. With NodeXL, you can enter a network edge list in a worksheet, click a button and see your graph, all in the familiar environment of the Excel window.

NodeXL Pro offers additional features that extend NodeXL Basic, providing easy access to social media network data streams, advanced network metrics, and text and sentiment analysis, and powerful report generation. NodeXL Pro can create insights into social media streams with just a few clicks.

NodeXL Pro released on October 12, 2015.

NodeXL Graph Gallery

Visit the NodeXL Graph Gallery to see the wide variety of graphs that have been created by the NodeXL community.

What is Power BI?

Power BI is a business analytics solution that lets you visualize your data and share insights across your organization, or embed them in your app or website. Connect to hundreds of data sources and bring your data to life with live dashboards and reports.

WATCH OVERVIEW > WATCH DEMO >

Microsoft Power BI

- Unlimited access to on-site and in-cloud data to give a centralized data access hub
- Comes with a robust tutorial system to help in learning how to use it properly
- Allows to import and publish data online for your collaborators
- Free and 10$/month for pro-package

More DA Visualization Tools to Consider

- IBM Watson Analytics
- Infogram
- ChartBlocks
- Datawrapper
- Sisense
- Looker
- Periscope Data
- Zoho Analytics
- Domo
- Qlikview
- Klipfolio
- SAP Analytics Cloud
- Kibana
- Plotly
- Chartio
- …

Choose How to Represent Your Data

Numbers

https://www.paulkrekorian.org/los_angeles_tourism_on_the_rise

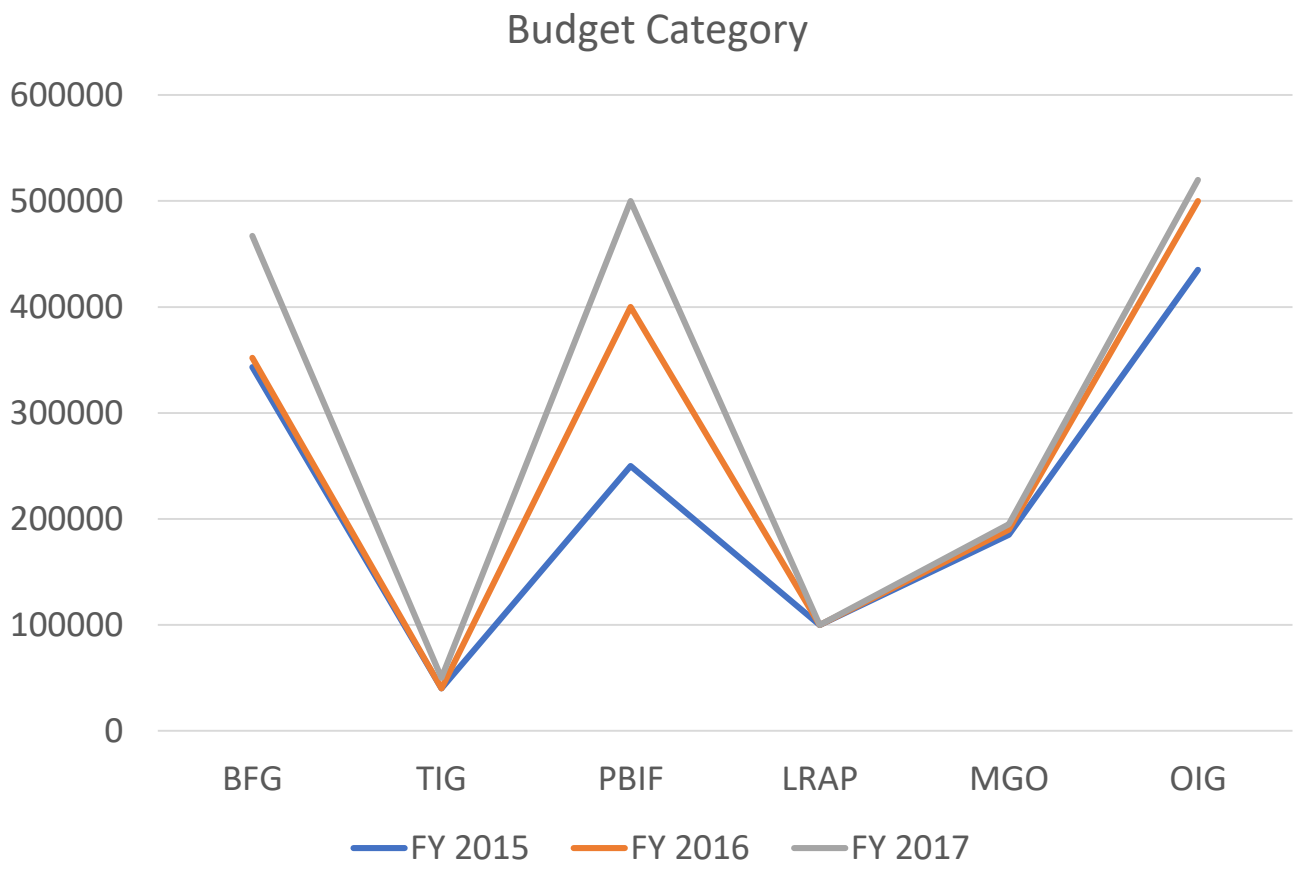
Tables

- Could be helpful when talking with sophisticated audience
- Also, when is the data is reasonably small

Budget Category	FY 2015	FY 2016	FY 2017
BFG	343150	352000	467000
TIG	40000	40000	50000
PBIF	250000	400000	500000
LRAP	100000	100000	100000
MGO	185000	190000	195000
OIG	435000	500000	520000

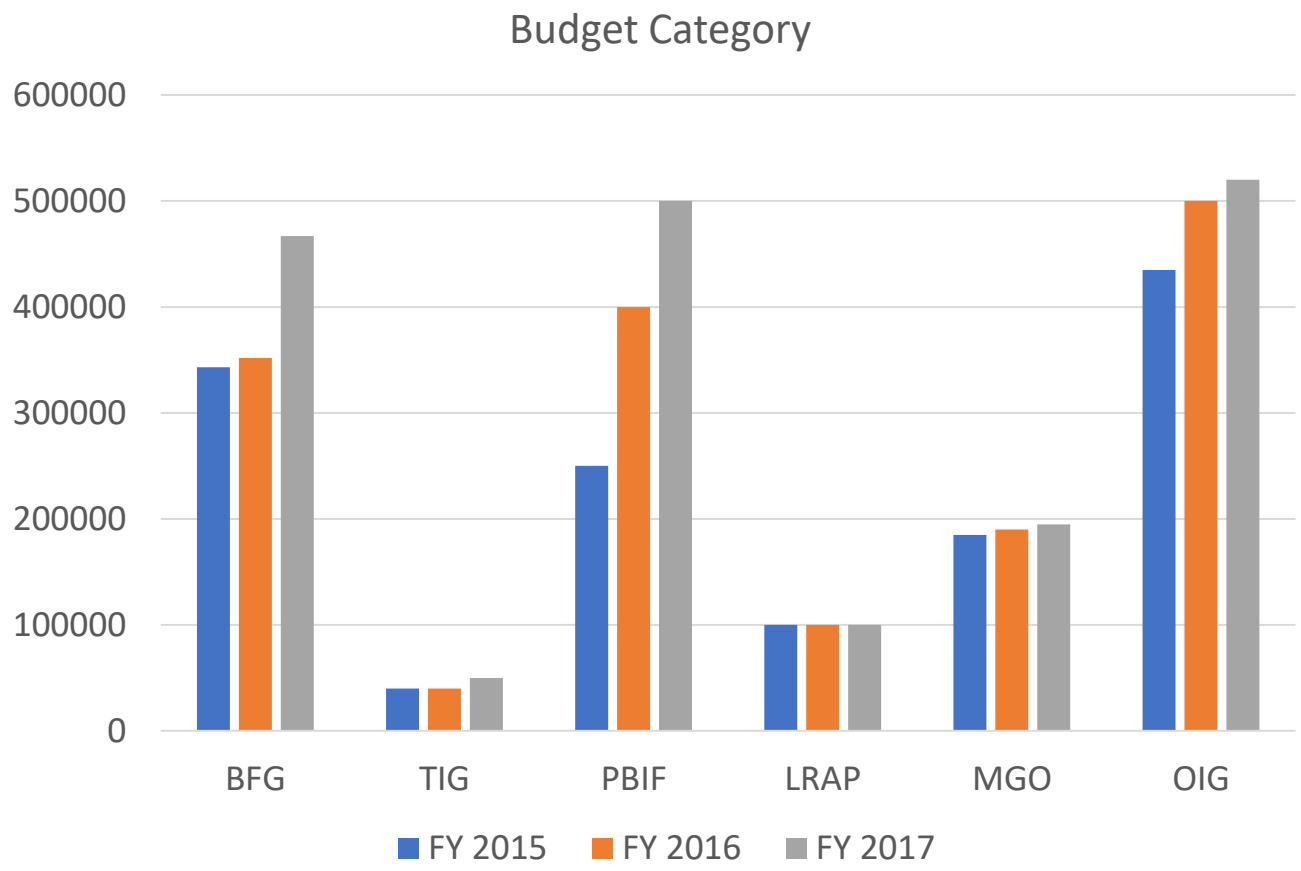
Lines

- Likely, best way to show trends in the data



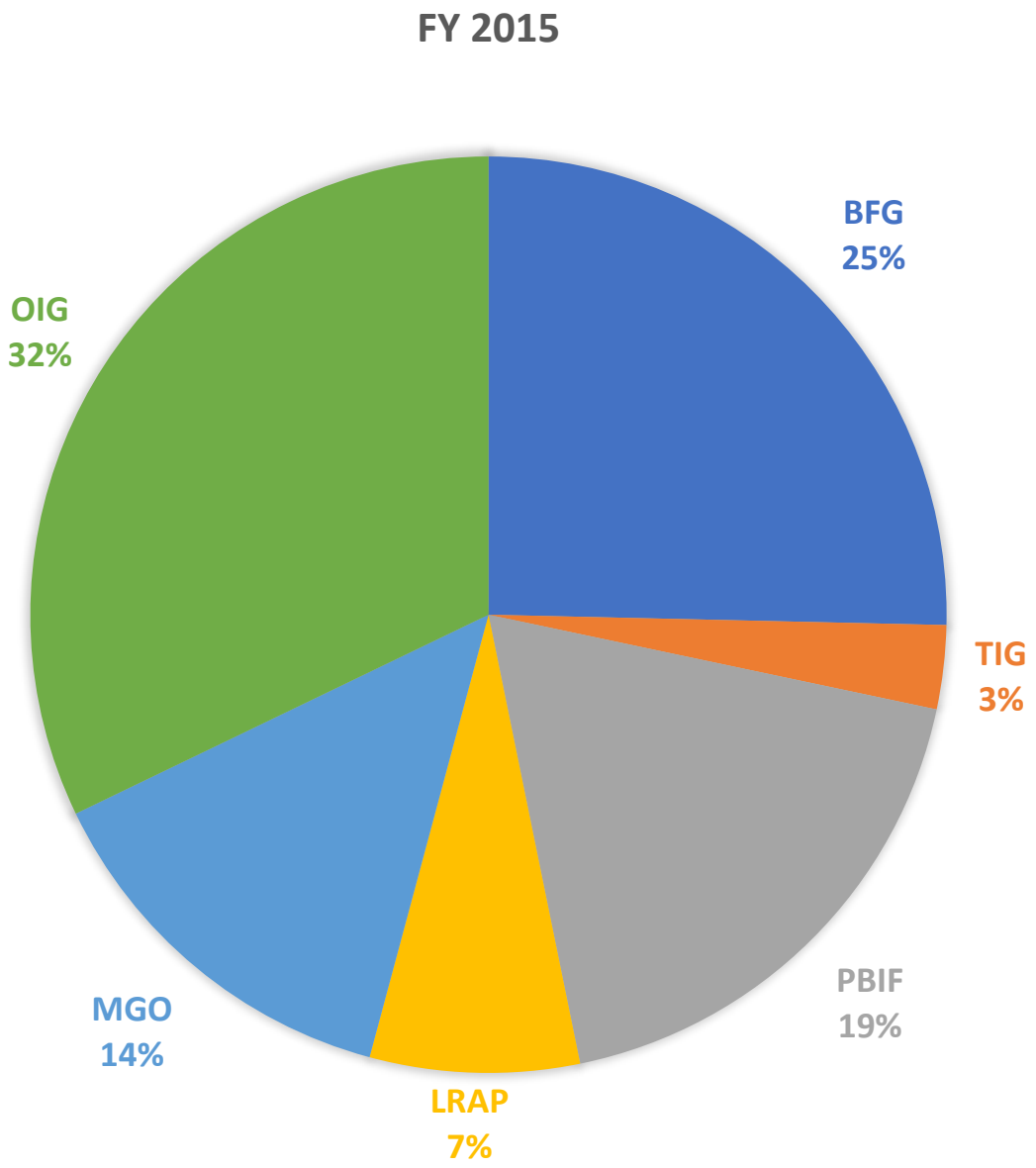
Bars

- Most popular way of presenting data
- Shows insights of the data that is not suitable for line charts



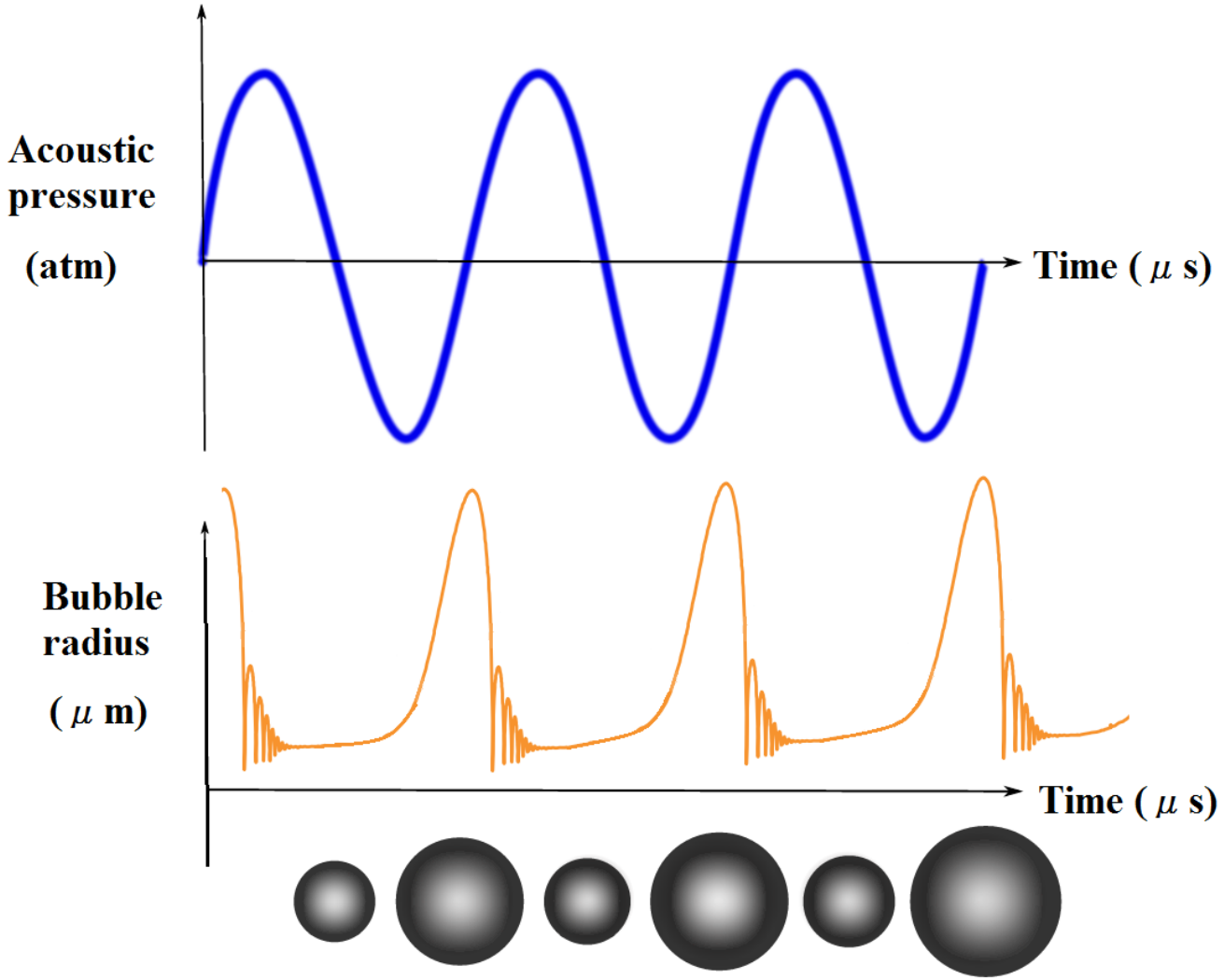
Pie Charts

- Nothing pie charts can do better than bar charts
- Could be reasonable to highlight two or three percentages among the rest



Bubble Plot

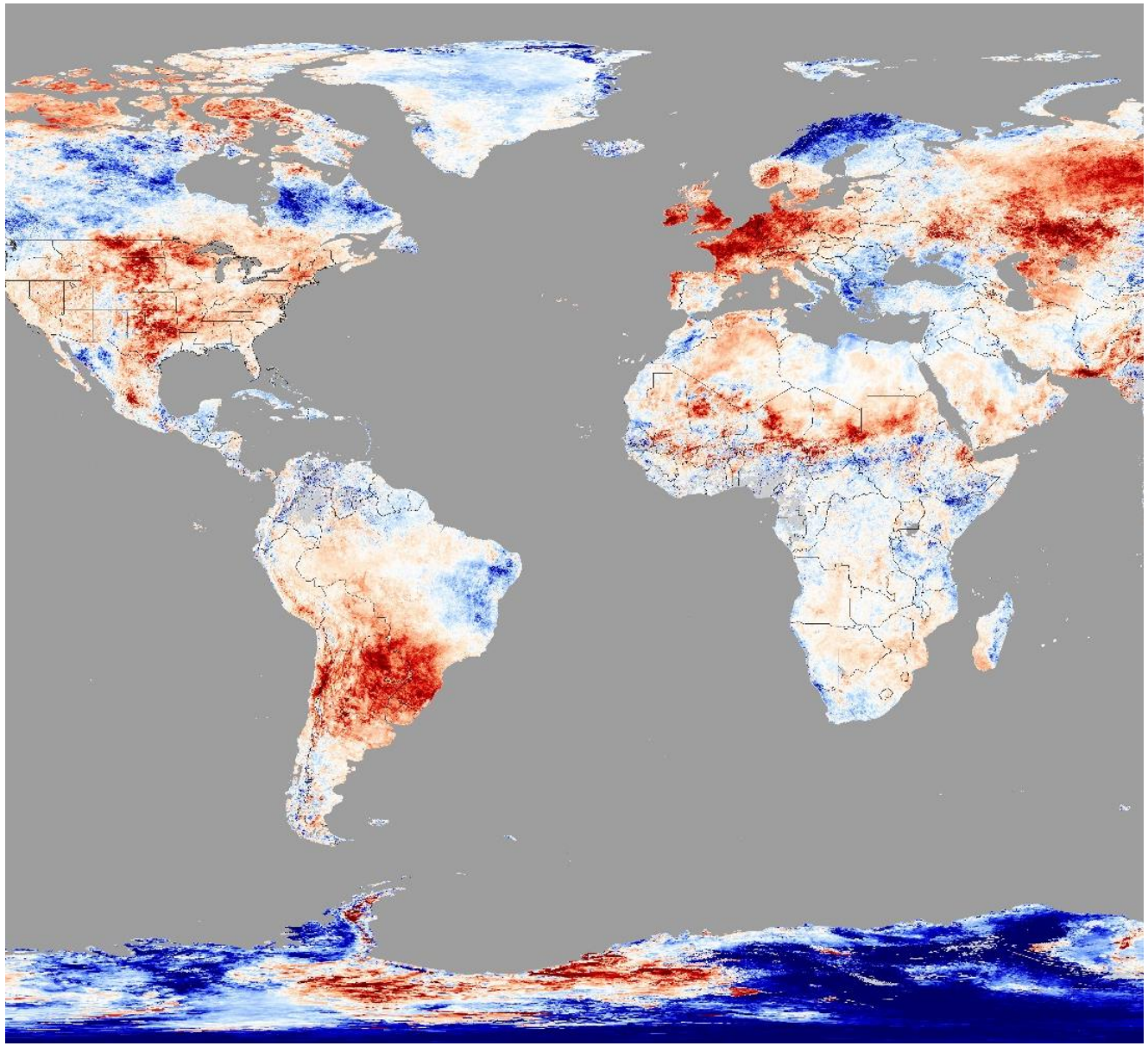
- Scatter (or bubble) plots can show the trends of a lot of different data points



https://commons.wikimedia.org/wiki/File:Acoustic_pressure_vs_bubble_radius.png

Map Plot

- Maps can be a powerful way to represent data geographically
- In the figure, land surface temperature anomaly (-10 :10^{o}C)



https://eoimages.gsfc.nasa.gov/images/imagerecords/6000/6803/globallstanom_tmo_2006193_lrg.jpg

Designing a Dashboard

Dashboards

Fundamentals

Perception
Vision
Color
Principles

Techniques

Representation
Presentation
Interaction

Application

Dashboard
Visual
Analytics

Dashboards (Cont.)

"Most information dashboards that are used in business today fall far short of their potential"

Stephen Few

Dashboards – Definition

"A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance"

Few2007

Dashboards – Characteristics

- Visual displays
- Display information needed to achieve specific objectives
- Fits on a single computer screen
- Are used to monitor information at a glance
- Have small, concise, clear, intuitive display mechanisms
- Are customized

Dashboards – Common Mistakes

- Supplying inadequate context for the data
  - Fail to provide adequate context to make the measures meaningful
- Displaying excessive detail or precision
  - Show unnecessary detail
- Choosing a deficient measure
  - Use of measures that fail to directly express the intended message
- Choosing inappropriate display media
  - Common problem with pie charts
- Introducing meaningless variety
  - Exhibit unnecessary variety of display media

Dashboards – Common Mistakes (Cont.)

- Using poorly designed display media
  - A legend was used to label and assign values to the slices of the pie. This forces our eyes to bounce back and forth between the graph and the legend to glean meaning, which is a waste of time and effort when the slices could have been labeled directly.
  - The order of the slices and the corresponding labels appears random. Ordering them by size would have provided useful information that could have been assimilated instantly.
  - The bright colors of the pie slices produce sensory overkill. Bright colors ought to be reserved for specific data that should stand out from the rest.
- Encoding quantitative data inaccurately

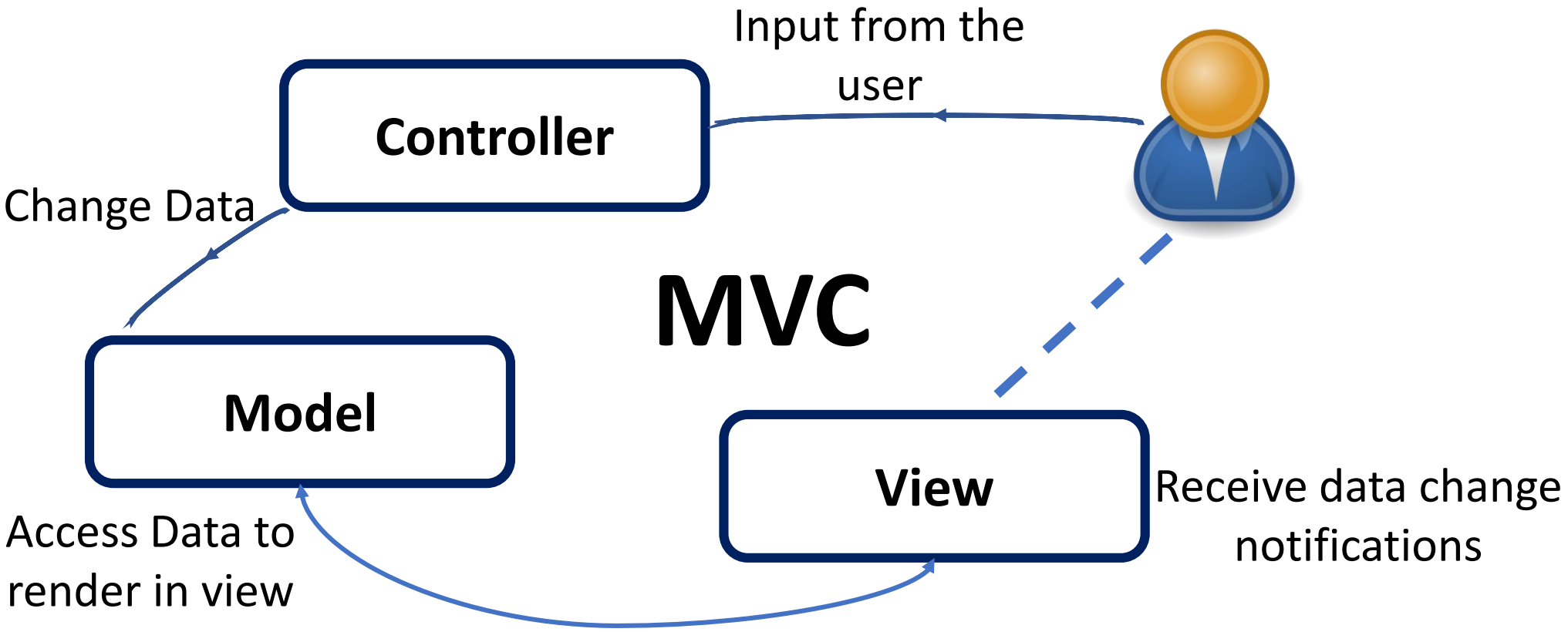
Dashboards – Common Mistakes (Cont.)

- Arranging the data poorly
  - The most important data ought to be prominent
  - Data that require immediate attention ought to stand out
  - Data that should be compared ought to be arranged and visually designed to encourage comparisons
- Highlighting important data ineffectively or not at all
  - Fail to differentiate data by its importance
    - Giving relatively equal prominence to everything on the screen
- Cluttering the display with useless decoration
  - Try to look something that is not
  - It results in useless and distracting decoration

Dashboards – Common Mistakes (Cont.)

- Misusing or overusing color
  - Too much color undermines its power
- Designing an unattractive visual display
  - The fundamental challenge of dashboard design is to effectively display a great deal of often disparate data in a small amount of space

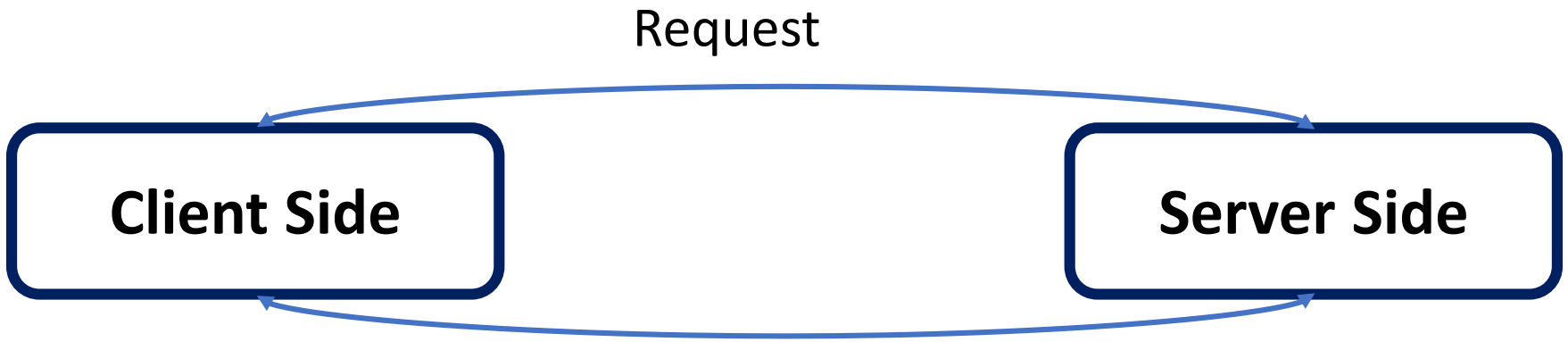
Model-View-Control



"The controller is essential and explicit: you have to specify what to do when you receive user requests and what resources you are going to mobilize to carry out the necessary tasks outlined in the model"

(Ribeiro 2016)

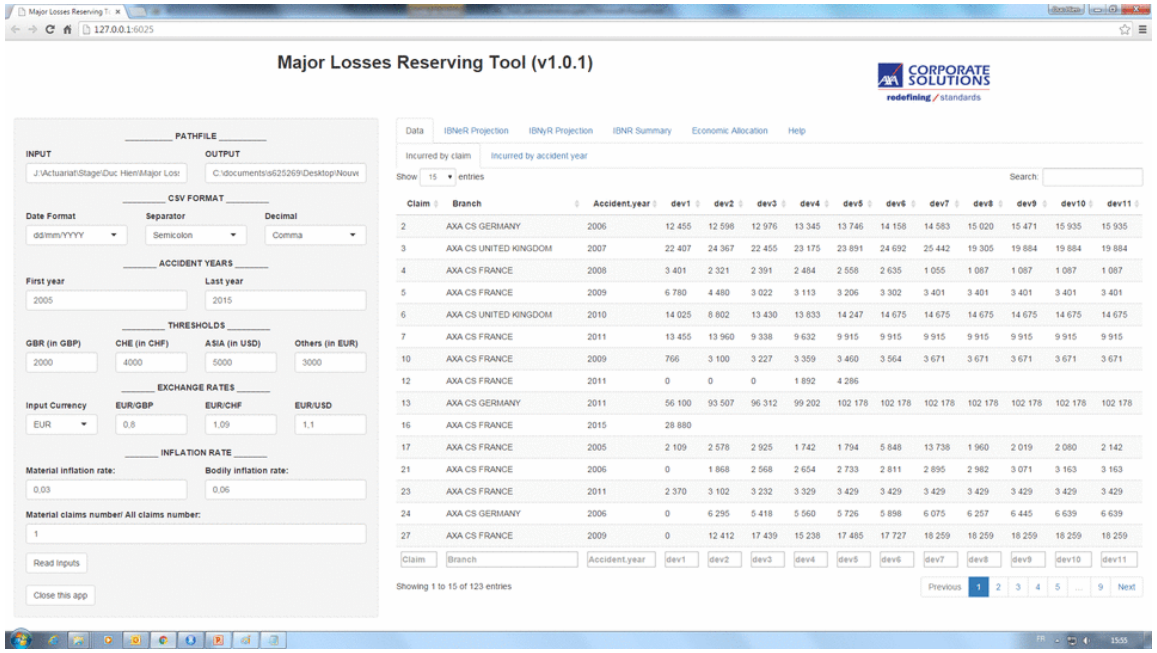
Reactive Web Framework

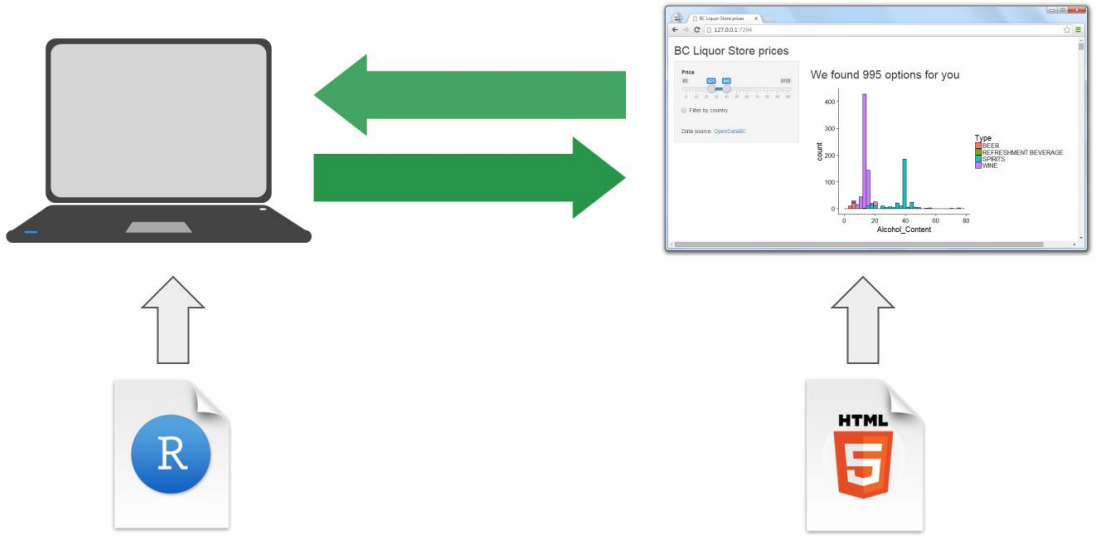


"Reactive Systems are highly responsive, giving users effective interactive feedback"

https://www.reactivemanifesto.org
http://littleactuary.github.io/blog/Web-application-framework-with-Shiny/

Reactive Web Framework

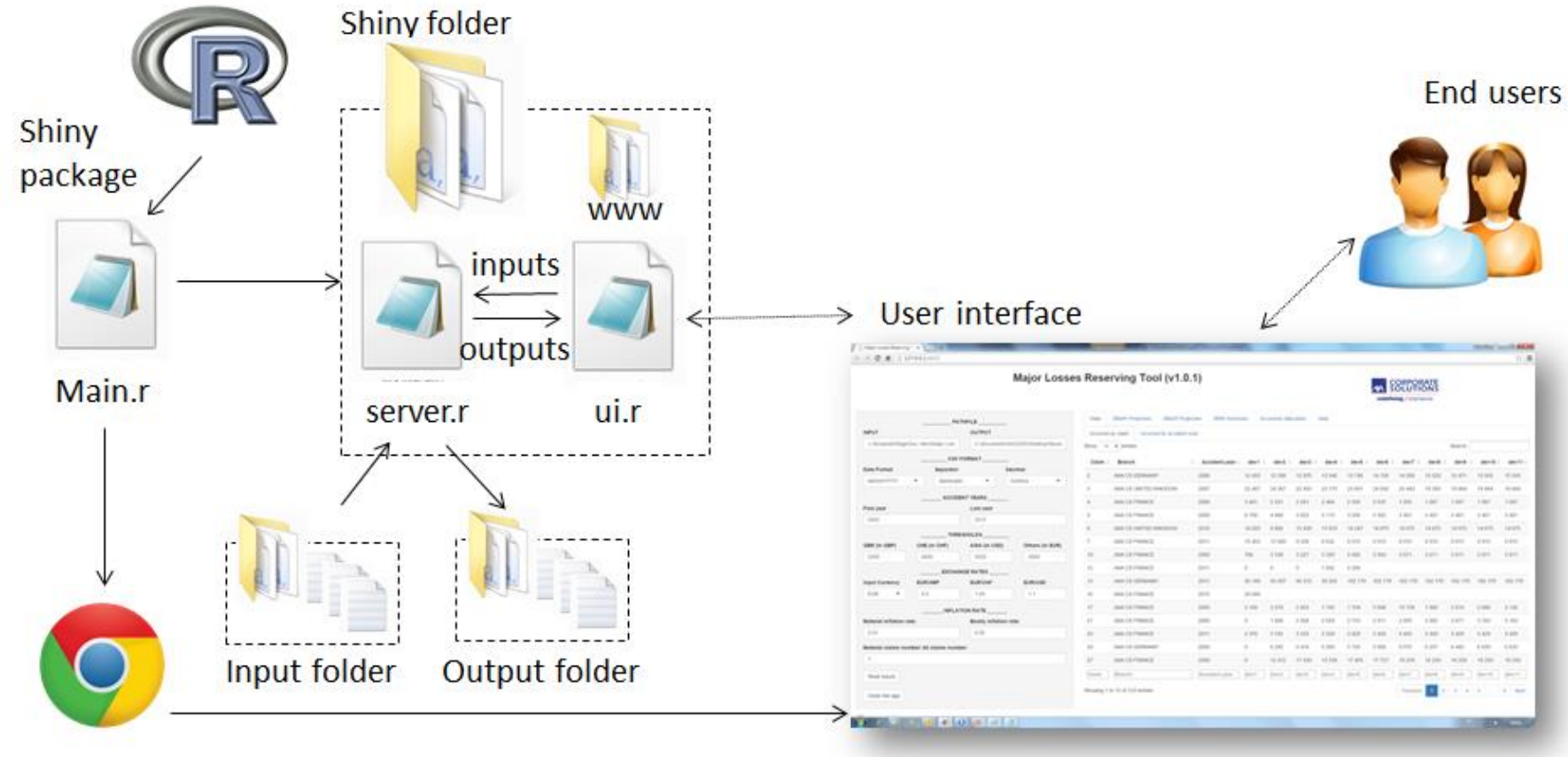




R-Shiny

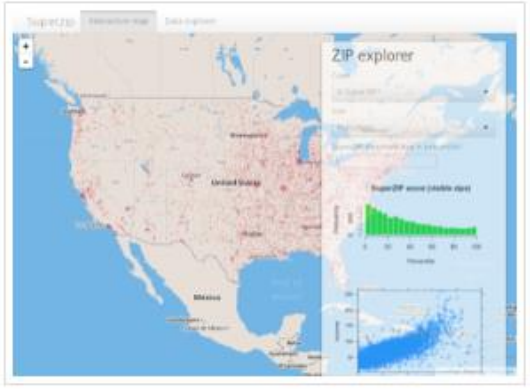
- What is Shiny?
  - R package that allows programmers to:
    - Quickly build a user interface to explore and work with data.
    - Publish their charts and graphs online.
- Why Shiny?
  - Interactive display and manipulation of data
  - Interactive communication sessions between the user's browser and a server without having to poll the server for a reply
  - Easy to develop and share with clients and project teams
  - Entirely extensible - custom input/output
  - Open source library

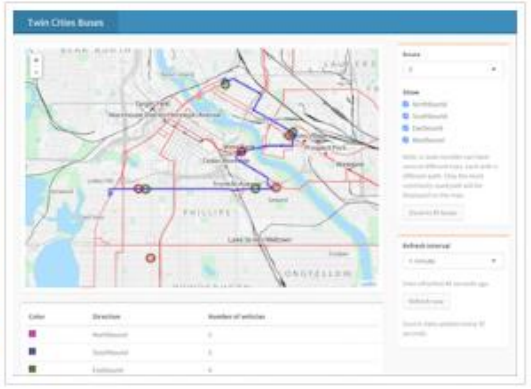
R-Shiny (Cont.)

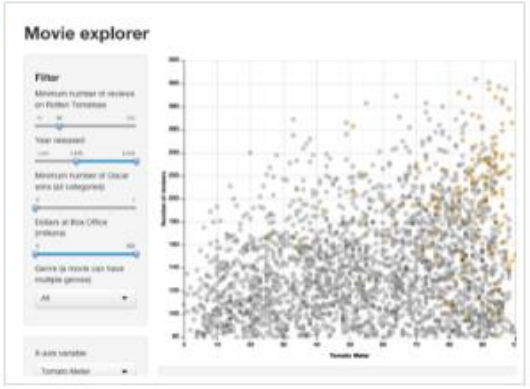

Shiny folder
Shiny package
www
inputs
outputs
Main.r
server.r
ui.r
End users
User interface
Input folder
Output folder
Major Losses Reserving Tool (v1.0.1)

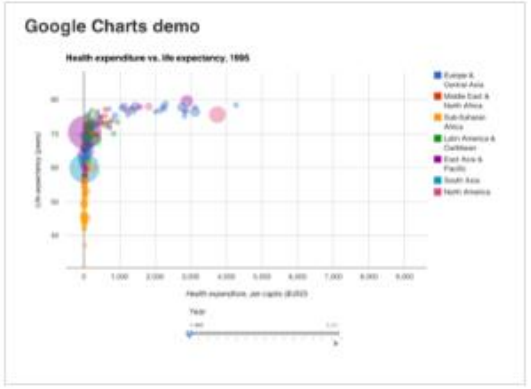
R-Shiny (Cont.)

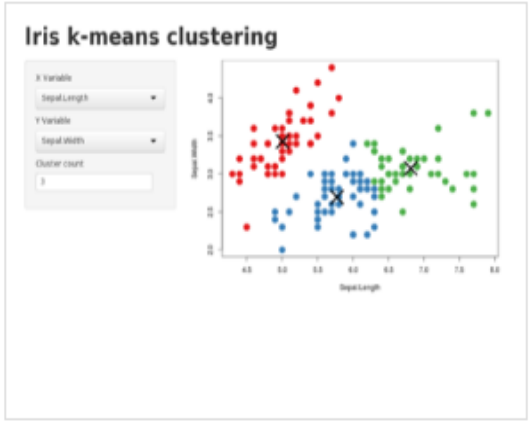
- Wide variety of templates that you can get inspired by

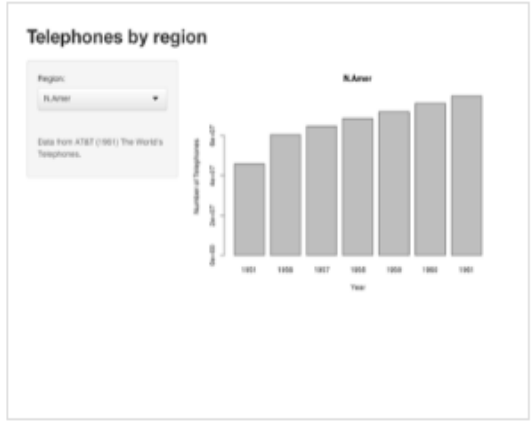


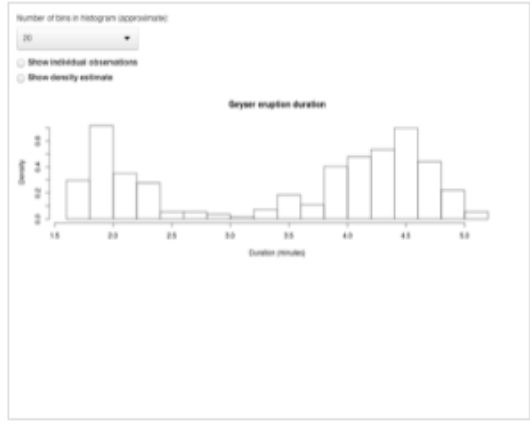


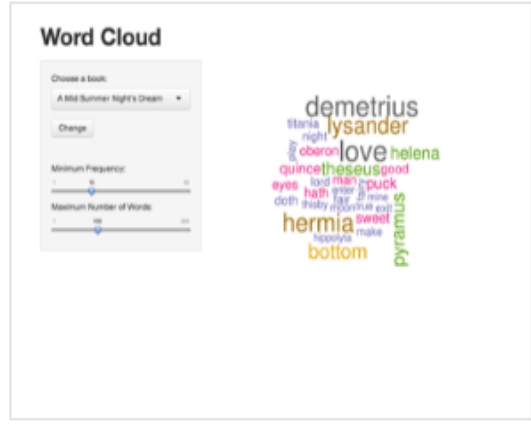












R-Shiny (Cont.)

- Shiny application consists of two main components
  - The user interface (frontend) component
    - Allows the user to interact with the application
  - The server (backend) component:
    - Listens to the requests of the user and respond accordingly based on predefined set of functionalities
- A nice tutorial can be found at

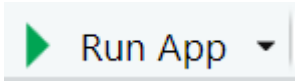
  https://shiny.rstudio.com/tutorial/

R-Shiny (Cont.)

- Shiny application consists of two main components
  - The user interface (frontend) component
    - Allows the user to interact with the application
  - The server (backend) component:
    - Listens to the requests of the user and respond accordingly based on predefined set of functionalities
- A nice tutorial can be found at

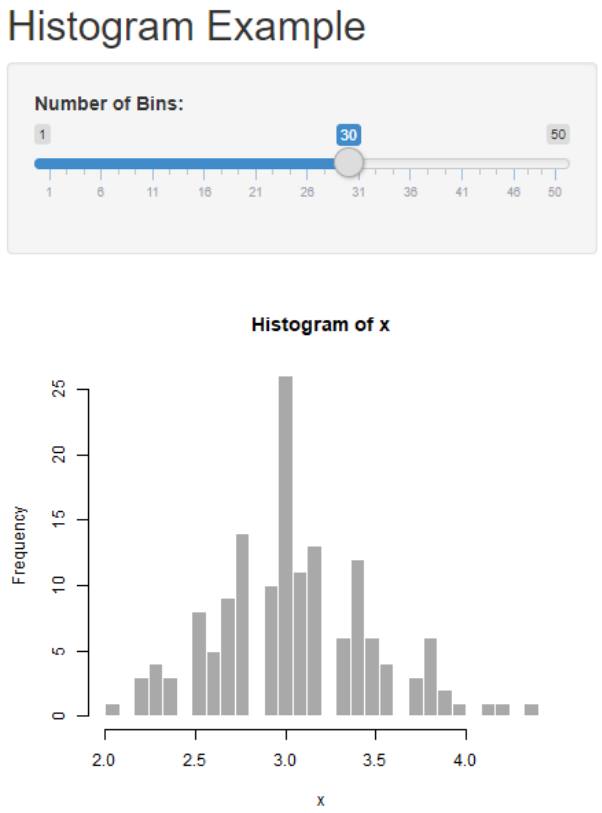
  https://shiny.rstudio.com/tutorial/

Two Ways to Create Shiny Application

1. Use 2 R script files – `ui.R` and `server.R` – Rstudio has made it easy for you to incorporate widgets and charts (ui side) and run algorithms on your data (server.R)
   - Click on 

   or type `runAPP()` in the console.
2. Use `app.R` and store the user interface widgets and charts in one object (let's say *u*) and the algorithms in another object (let's say *s*)
   - Call `shniyApp(ui = u, server = s)`

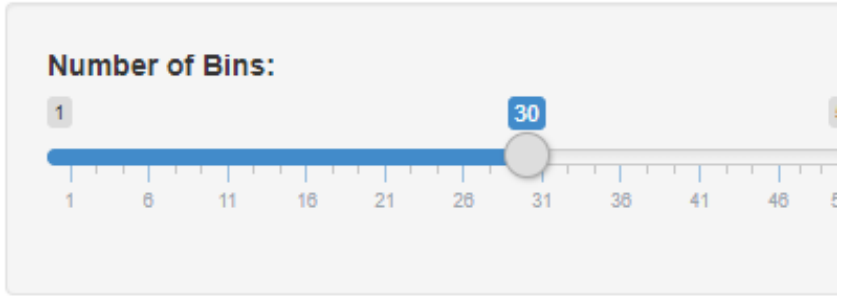
ui.R

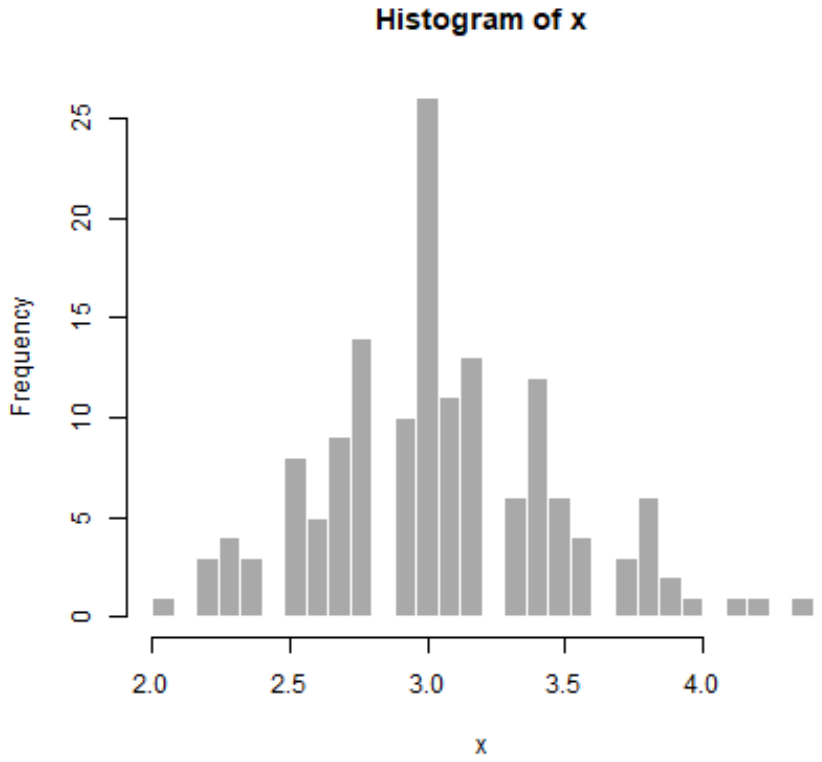


ui.R | Server.R

```
1  library(shinydashboard)
2  fluidPage(
3    # Application Title
4    titlePanel('Histogram Example'),
5    # Sidebar with a slider input for the number of bins
6    sidebarLayout(
7      sidebarPanel(
8        sliderInput('bins', 'Number of Bins:',
9                    min = 1, max = 50, value = 30),
10                 ),
11     mainPanel(
12       plotOutput('histPlot')
13     )
14   )
15 )
```

Histogram Example



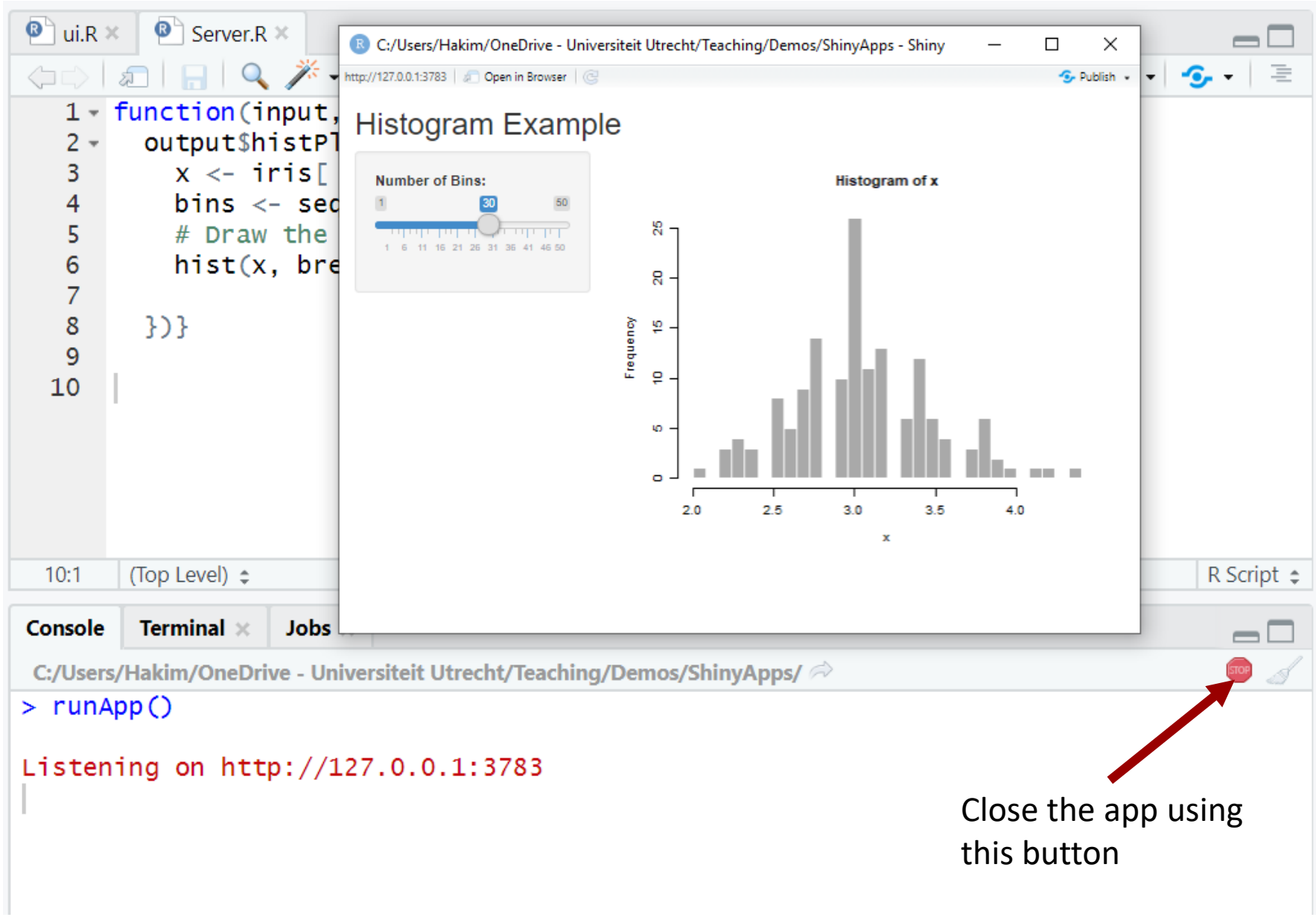


ui.R | Server.R

```
1  function(input, output, session) {
2    output$histPlot <- renderPlot({
3      x <- iris[ , 2]
4      bins <- seq(min(x), max(x), length.out = inp
5      # Draw the histogram using the specified nu
6      hist(x, breaks = bins, col = 'darkgray', bo
7
8    })}
```

server.R

Closing the App



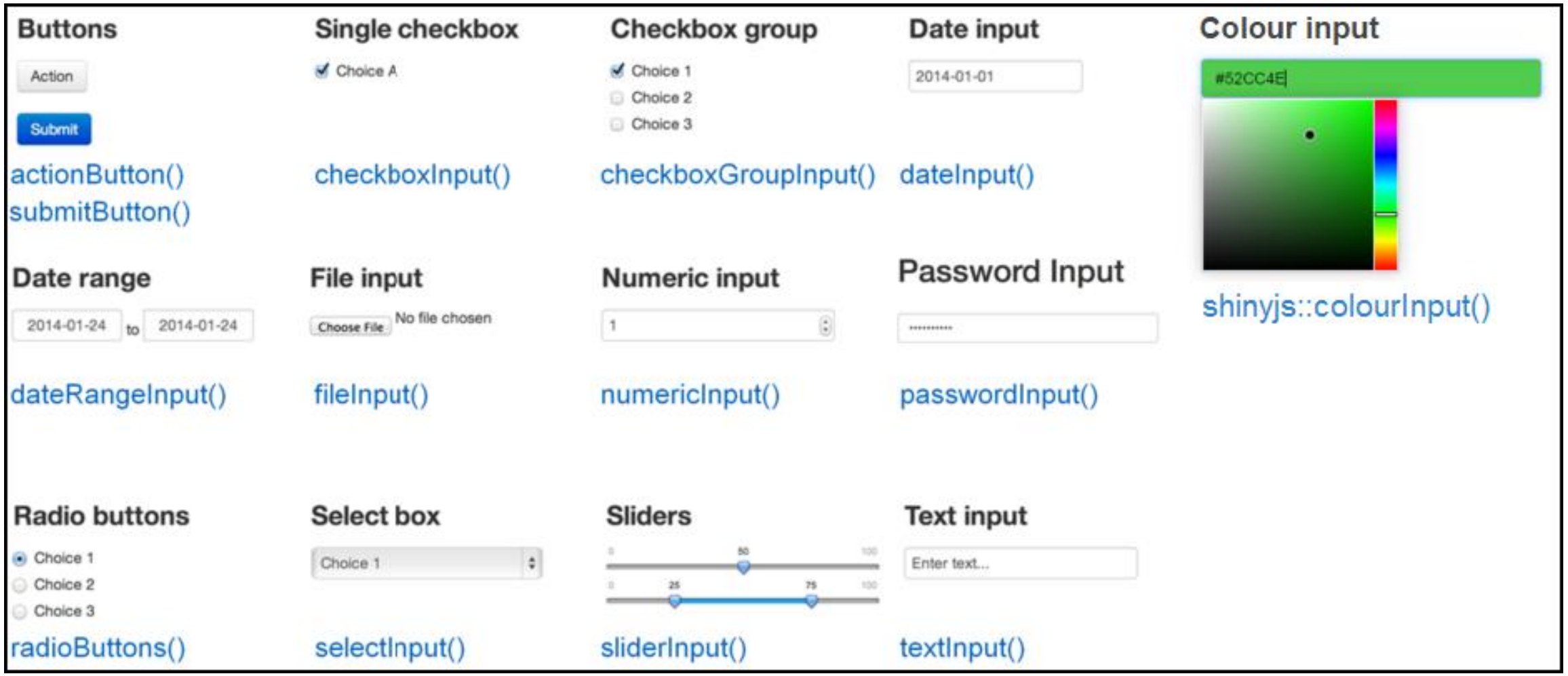
Close the app using this button

Application Template

```
library(shinydashboard)
# or library(shiny) for simple visualizations
u <- dashboardPage(dashboardHeader(),
      dashboardSidebar(), dashboardBody())
# dashboardPage expects three arguments
# You can also use pageWithSidebar() or fluidPage() …
s <- function(input, output){}
shinyApp(ui = u, server = s)
```

What is the output of this app?

R-Shiny – Basic Widgets



Example: numericInput(inputId = "", label = "", min = 0, max = 10, step = 1, value = 5, width = 200)

R-Shiny (Cont.)

- Shiny application consists of two main components
  - The user interface (frontend) component
    - Allows the user to interact with the application
  - The server (backend) component:
    - Listens to the requests of the user and respond accordingly based on predefined set of functionalities
- A nice tutorial can be found at

  https://shiny.rstudio.com/tutorial/

R-Shiny and HTML

- How the parts of a dashboard work together?
- HTML tag functions in Shiny, like div() and p() return objects that can be rendered as HTML

```
# A basic div
div(class = "my-class", "Div content")
## <div class="my-class">Div content</div>

# Nested HTML tags
div(class = "my-class", p("Paragraph text"))
## <div class="my-class">
##   <p>Paragraph text</p>
## </div>
```

R-Shiny and HTML (Cont.)

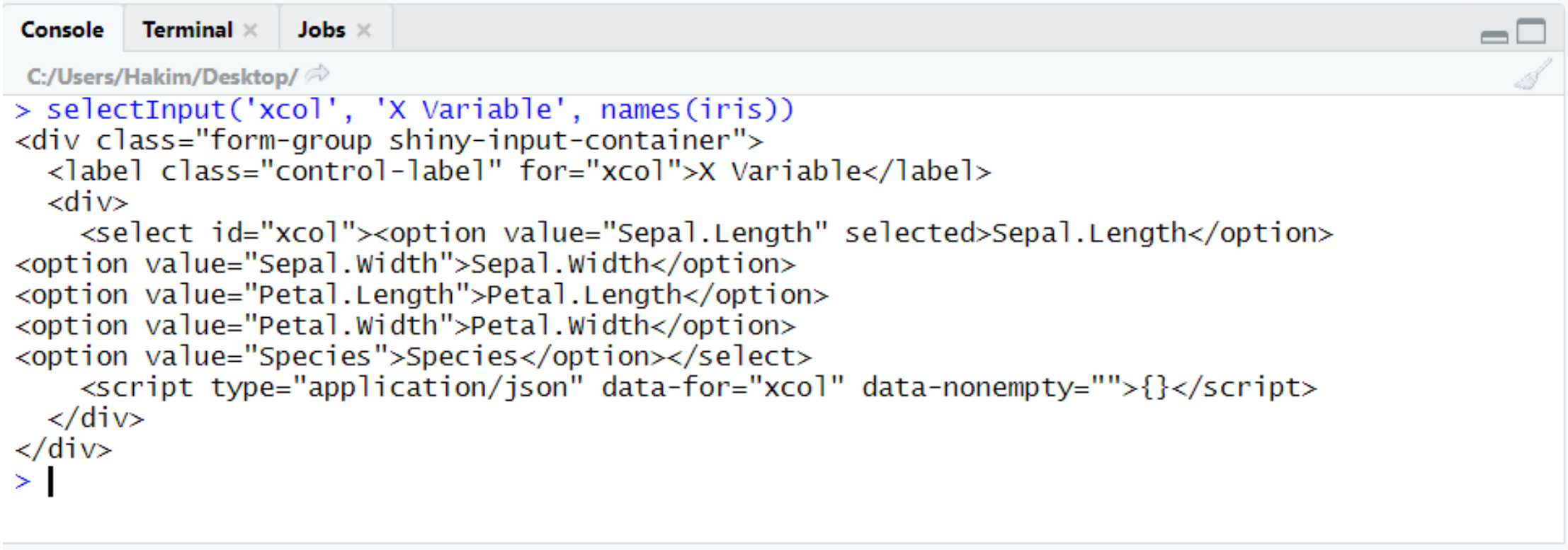
- Some functions in Shiny, returns more complex HTML fragments

```
textInput("Id", "Label")
## <div class="form-group shiny-
input-container">
##   <label for = "Id">
Label</label>
##   <input id="Id" type="text"
class="form-control" value=""/>
## </div>
```

```
sidebarPanel(
  div("First div"),
  div("Second div")
)
## <div class="col-sm-4">
##   <form class="well">
##     <div>First div</div>
##     <div>Second div</div>
##   </form>
## </div>
```

R-Shiny and HTML (Cont.)

- To check the equivalent HTML of shiny functions, run the shiny function in the r-console



Demo

- In the demo, you will learn how to:
  - Create simple interactive plot using Shiny
  - Select parameters using selectInput() and numericInput()
  - Select input file and reading them
  - Rend plots
  - Create conditional panels

Reading Material

- Useful resources:
  - Shiny official tutorial -http://shiny.rstudio.com/tutorial
  - Cheat sheet - http://shiny.rstudio.com/images/shiny-cheatsheet.pdf
  - Publish your app free - http://www.shinyapps.io
  - Examples - http://www.showmeshiny.com/
  - Tutorial by Dean Attali - http://deanattali.com/blog/building-shiny-apps-tutorial/