# Interactive data visualization

Lecture 15

Louis SIRUGUE

01/2022

## Last time we saw

## **Shapefiles and rasters**

#### Two main types of geolocalized datasets:

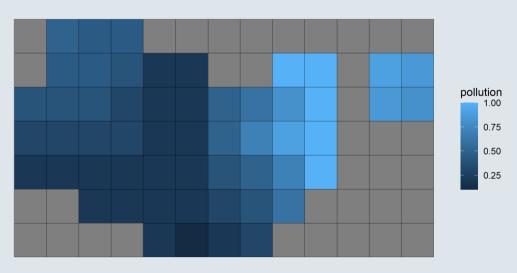
## **Shapefiles**

- One row per entity/one column per variable
- A geometry variable with the coordinates of the points/polylines/polygons

# Median income — High — Low

#### **Rasters**

- Works like a picture, with cells like pixels
- And each cell can take a given value, e.g. pollution observed from satellites



## Last time we saw

## **Coordinates Reference Systems**

- A Coordinate Reference System (CRS) is a model of the Earth in which each location is coded using degrees
  - It allows to project the surface of the globe on a plane
  - But there is a **tradeoff** between preserving:

**Shape** (like the Mercator projection)



**Scale** (like the Equal-Area Cylindrical projection)



- Most projections are somewhere in between
- For France: Lambert 93 projection (EPSG:2154)

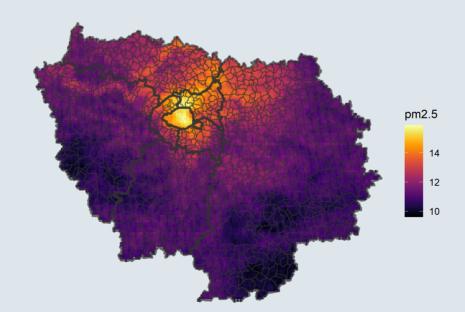
→ First thing to do: **reprojection** 

## Last time we saw

## Operations on geolocalized data

#### **Zonal statistics**

- Computing statistics on areas delimited by a shapefile from values of a raster
  - Project shapefile and raster the same way
  - Compute the mean/max/... of cell values



#### **Centroids**

- The centroid is the arithmetic mean position of all the points in the polygon
  - To compute distances between polygons
  - A centroid is not always within its polygon

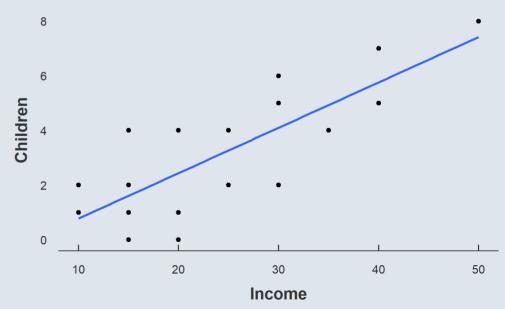


- We've seen in previous lectures that when regressing y on x:
  - Controlling for z allows to net out the relationship between x and y from how they both relate to z
  - Interacting x with z allows to estimate how the relationship between x and y varies with z
- Given what I've seen in the homeworks it seems unclear for many of you

#### → So let's get back to it with some visualization

```
library(tidyverse)
data <- read.csv("household_data.csv")
head(data, 7) # fake data</pre>
```

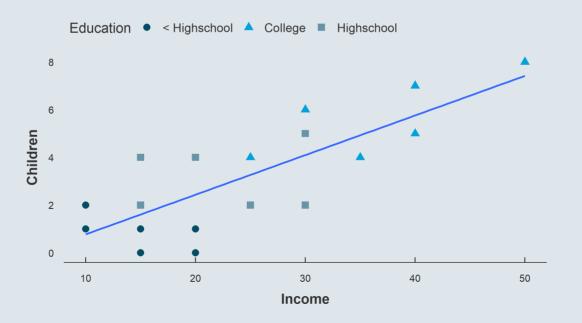
```
Income Children
                         Education
##
## 1
                    1 < Highschool
         20
## 2
                    1 < Highschool
         10
                    2 < Highschool
## 3
         10
                    0 < Highschool
## 4
         15
## 5
                    1 < Highschool
         15
                    0 < Highschool
## 6
         20
                        Highschool
## 7
         15
```



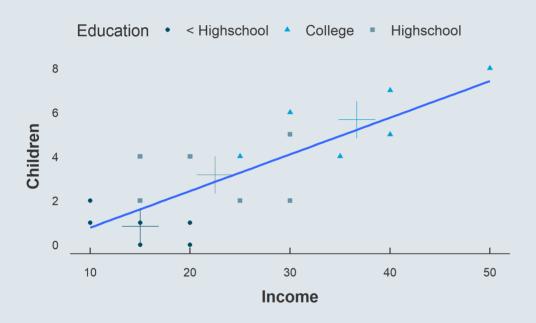
• There's a clear positive relationship

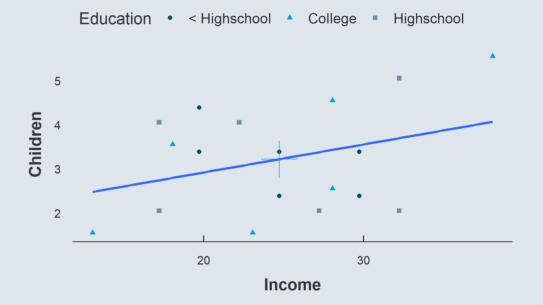
```
## Estimate Pr(>|t|)
## (Intercept) -0.885 0.319
## Income 0.166 0.000
```

- But what if this relationship was driven by a third variable?
- Maybe it's just that more educated parents tend to earn more and to have more children



• **Controlling** for education does the same to the slope **as recentering** the graph with respect to education :





- The crosses are located at the average x and y values for each education group
  - Controlling for education shifts x and y by group such that crosses superimpose

• Here when we **do not control** for education:

$$Children_i = \alpha + \beta Income_i + \varepsilon_i$$

- We estimate the overall relationship (here, significantly positive)
- But when we **control** for education:

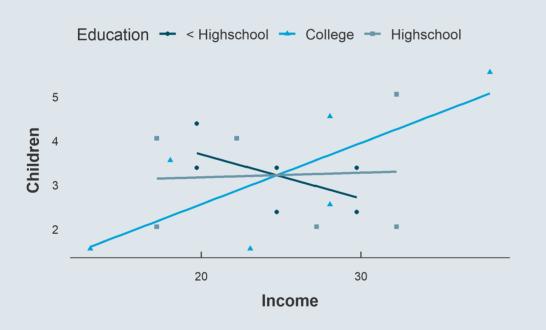
$$Children_i = \alpha + \beta Income_i + \gamma_1 1\{Education_i = ext{Highschool}\} + \gamma_2 1\{Education_i = ext{College}\} + arepsilon_i$$

- We estimate the relationship net of the effect of education (here, not significant)
- **Interacting** the two variables is going one step further:

$$Children_i = lpha + eta Income_i + \gamma_1 1\{Education_i = ext{Highschool}\} + \gamma_2 1\{Education_i = ext{College}\} + \delta_1 Income_i imes 1\{Education_i = ext{Highschool}\} + \delta_2 Income_i imes 1\{Education_i = ext{College}\} + arepsilon_i$$

- It is not simply taking into account the fact that education may plays a role
- o It estimates by how much the relationship between x and y varies according to z

• Interacting income with education provides one slope per education group:



##		Estimate	Pr(> t )
##	(Intercept)	2.333	0.225
##	Income	-0.100	0.411
##	EducationCollege	-1.768	0.553
##	EducationHighschool	0.596	0.819
##	<pre>Income:EducationCollege</pre>	0.239	0.095
##	<pre>Income:EducationHighschool</pre>	0.111	0.445

- The principle is the same when the third variable is continuous:
  - o Controlling nets out the slope from how the third variable enters the relationship
  - o Interacting gives by how much the slope changes on expectation when the third variable increases by 1
  - And we can control for/interact with multiple third variables

## Today: Interactive data visualization

#### 1. Introduction to shiny apps

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

#### 2. Our first shiny app

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 3. More advanced tools

- 3.1. Input randomization
- 3.2. HTML formatting

## 4. Wrap up!

# Today: Interactive data visualization

## 1. Introduction to shiny apps

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

#### 1.1. General structure

- **Shiny** is an R package that makes it easy to build **interactive web apps** straight from R:
  - Shiny app to find the colleges that fit your criteria
  - Shiny app to visualize data on movies
  - The online quizzes of this course
- To make a Shiny app you should create an R script and name it app.R
- You shiny app should contain two components
  - 1. The **user interface:** What is displayed on the screen, what the user can interact with
  - 2. The **server:** Where the calculations are made to display the interactive components accordingly

```
library(shiny)

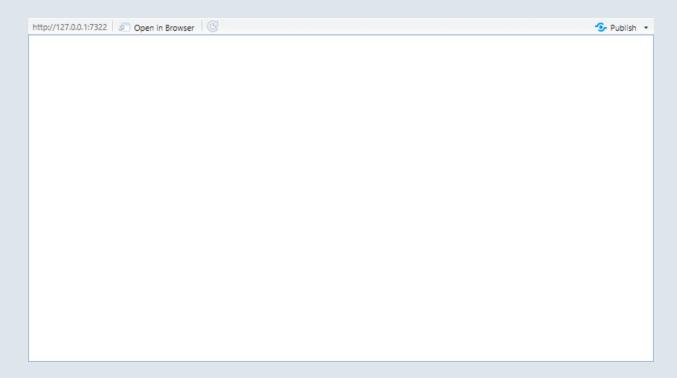
ui <- fluidPage()
server <- function(input, output) {}

shinyApp(ui = ui, server = server)</pre>
```

→ R will detect that you are creating a shiny app and you will have access to the RunApp · button

#### 1.1. General structure

• You can already click on it to view you blank app:



→ We have to program what we want to appear in the user interface

#### 1.2. User interface

- There are many different types of **input widgets** to place in the UI:
  - **numericInput():** Write a numeric input
  - **textInput():** Write a character input
  - checkbox[Group]Input(): Box[es] to tick
  - radioButtons(): One item to tick
  - **selectInput():** Select an item from a dropdown list
  - 0 ...
- These functions take the following **arguments**:
  - **inputId:** The identifier of the input for use in the server function
  - **label:** The title of the input widget that will appear in the UI
  - **choices:** The list of input options for multiple choices inputs
  - **selected:** Which option is selected by default when multiple choices
  - **value:** What is filled by default in the text/numeric input boxes

→ Let's try out a few of them

#### 1.2. User interface

Write a number o

Write text ...

☐ Tick the box

Boxes to check

- $\Box$  A
- ✓ B

Select option

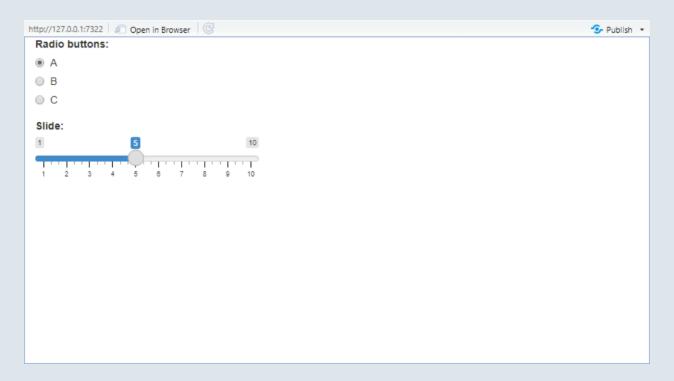
A 🕶

#### 1.2. User interface

- In the UI these elements should be separated with a commas
  - Let's do a user interface with radio buttons and a slider:

#### 1.2. User interface

Now we have some inputs but nothing happens when we use them



→ We should make an output that will react to these inputs

#### 1.3. Server

- We can add a reactive table in the server function:
  - Put a standard tibble() in a **reactive({})** environment
  - Fill this table with the inputs that should be called by their id as **input\$inputId**
  - Assign it to an output through renderTable({})

- Every time an input will change, **reactive({})** function will notice it
  - And the updated table will be stored into the output named "table" through renderTable({})

#### 1.3. Server

- But for this reactive table to appear on the app, we should put it in the UI
  - Use tableOutput("output\_label") to render a reactive table
  - Don't forget the comma!

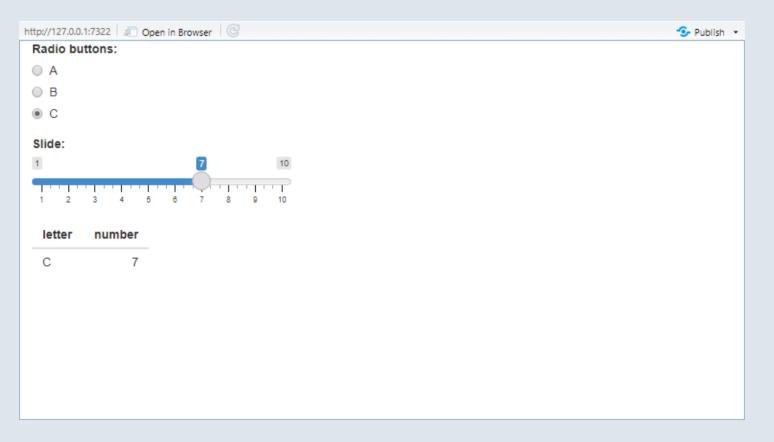
- We now have all the components of a shiny app:
  - Some input widgets in the UI
  - Reactive functions in the server
  - The processed output in the UI

#### 1.3. Server

```
library(shiny)
library(tidvverse)
ui <- fluidPage(</pre>
  radioButtons(inputId = "radio", label = "Radio buttons:",
               choices = c("A", "B", "C"), selected = "A"),
  sliderInput(inputId = "slider", label = "Slide:",
              min = 1, max = 10, step = 1, value = 5),
  tableOutput("table")
server <- function(input, output) {</pre>
  reactive tibble <- reactive({tibble(letter = input$radio,
                                        number = input$slider)})
  output$table <- renderTable({reactive tibble()})</pre>
shinyApp(ui = ui, server = server)
```

#### 1.3. Server

• We created an interactive table:



#### 1.3. Server

- As you might have guessed, tableOutput() and renderTable() only work for tables
  - There are specific UI and server function for each type of interactive element

### Main interactive elements in Shiny:

Interactive element	Server render function	UI output function	
Table	renderTable()	tableOutput()	
Plot	renderPlot()	plotOutput()	
Console output	renderPrint()	verbatimTextOutput()	
Text	renderText()	textOutput()	
UI element	renderUI()	uiOutput()	

## 1.4. Layout

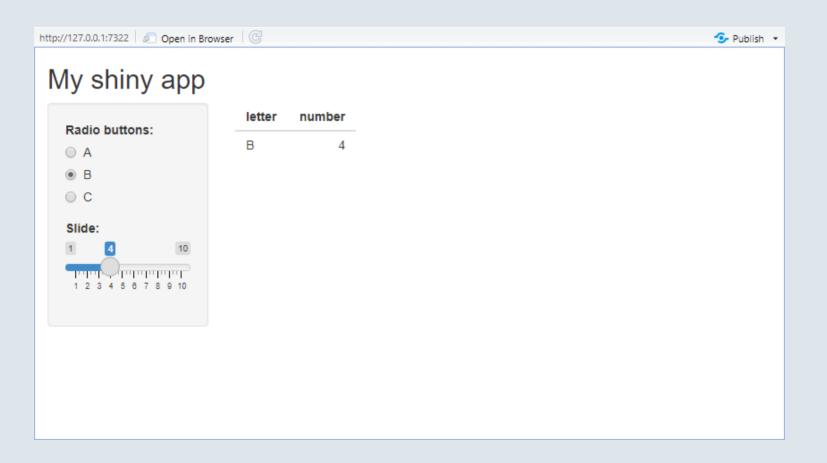
- Right now the presentation is not very convenient
  - Everything is stacked at the left of the page
- The sidebarLayout() allows to display:
  - A control panel on the left with the inputs
  - A main panel on the right with the reactive outputs

## 1.4. Layout

• We can also add a title to our app using the titlePanel() function:

```
ui <- fluidPage(</pre>
  titlePanel("My shiny app"),
  sidebarLayout(
    sidebarPanel(width = 3,
      radioButtons(inputId = "radio", label = "Radio buttons:",
                   choices = c("A", "B", "C"), selected = "A"),
      sliderInput(inputId = "slider", label = "Slide:",
                  min = 1, max = 10, step = 1, value = 5)
    ),
    mainPanel(width = 9, tableOutput("table"))
```

## 1.4. Layout



## Overview

## 1. Introduction to shiny apps ✓

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

#### 2. Our first shiny app

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 3. More advanced tools

- 3.1. Input randomization
- 3.2. HTML formatting

## 4. Wrap up!

## Overview

## 1. Introduction to shiny apps ✓

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

## 2. Our first shiny app

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 2.1. Import data on shiny

- We now know everything we need to build our first app
  - Let's make an app that allows to visualize the relationship between department-level characteristics
  - o dep\_data.csv contains department characteristics at the department-year level from 2012 to 2017

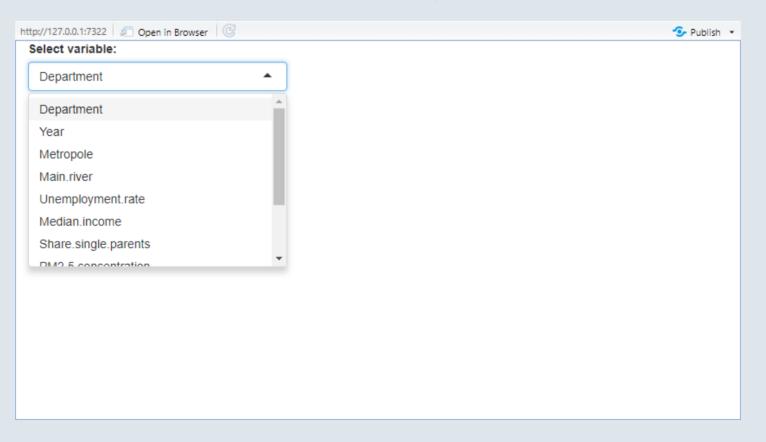
```
dep_data <- as_tibble(read.csv("dep_data.csv"))</pre>
head(dep_data)
## # A tibble: 6 x 11
     Department Year Metropole Main.river Unemployment.rate Median.income
##
     <chr>
                                                                        <dbl>
##
                <int>
                           <int>
                                      <int>
                                                         <fdb>>
## 1 Ain
                                                          6.85
                 2012
                                                                       21122
## 2 Ain
                 2013
                                                          7.22
                                                                       21490.
## 3 Ain
                 2014
                                                                       21700.
                                                          7.12
## 4 Ain
                 2015
                                                          7.38
                                                                       22020.
## 5 Ain
                 2016
                                                          7.35
                                                                       22272
## 6 Ain
                 2017
                               0
                                                          6.78
                                                                       22640
## # ... with 5 more variables: Share.single.parents <dbl>,
      PM2.5.concentration <dbl>, Population <int>, Log.population <dbl>,
## #
       Log.median.income <dbl>
## #
```

#### 2.1. Import data on shiny

- Importing data in a Shiny app is no different than usual
  - We can simply assign it to an object before specifying the UI and the server functions
  - Every object stored at the beginning of the script is accessible by the app

## 2.1. Import data on shiny

• We can now access the variable of the dataset from the dropdown list:



## 2.1. Import data on shiny

- Let's make a control panel containing:
  - Two dropdown lists for the x and y variables
  - A slider for the year of observation
- But using sliderInput() would display years with a comma, e.g., 2,012
  - To display years conveniently we should use sliderTextInput() from shinyWidgets

```
library(shiny)
library(tidyverse)
library(shinyWidgets)
```

• And we do not want the Department and Year variables to appear in our lists

```
dep_data <- as_tibble(read.csv("dep_data.csv"))
depvars <- names(dep_data)[!names(dep_data) %in% c("Department", "Year")]</pre>
```

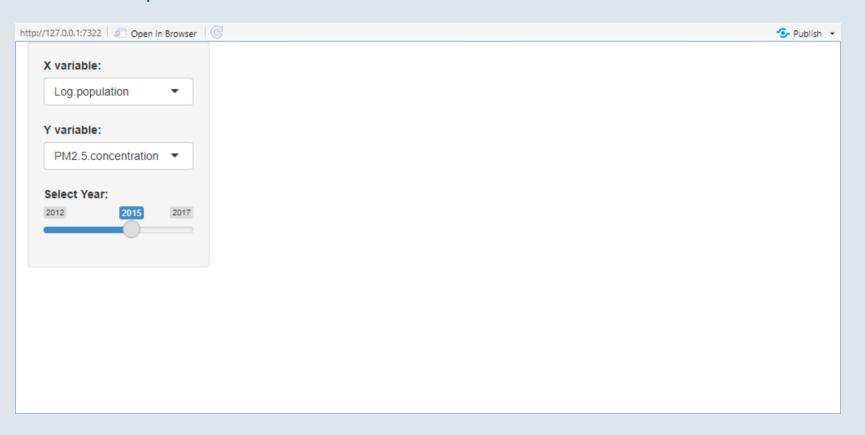
#### 2.1. Import data on shiny

• The desired UI writes as follows

```
ui <- fluidPage(</pre>
  sidebarLayout(
    sidebarPanel(width = 3,
      selectInput(inputId = "xvar", label = "X variable:",
                  choices = depvars, selected = "Log.population"),
      selectInput(inputId = "yvar", label = "Y variable:",
                  choices = depvars, selected = "PM2.5.concentration"),
      sliderTextInput(inputId = "year", label = "Select Year:",
                      choices = 2012:2017, selected = 2015)
    mainPanel(width = 9)
server <- function(input, output) {}</pre>
shinyApp(ui = ui, server = server)
```

## 2.1. Import data on shiny

• We now have a control panel that is linked to our data:

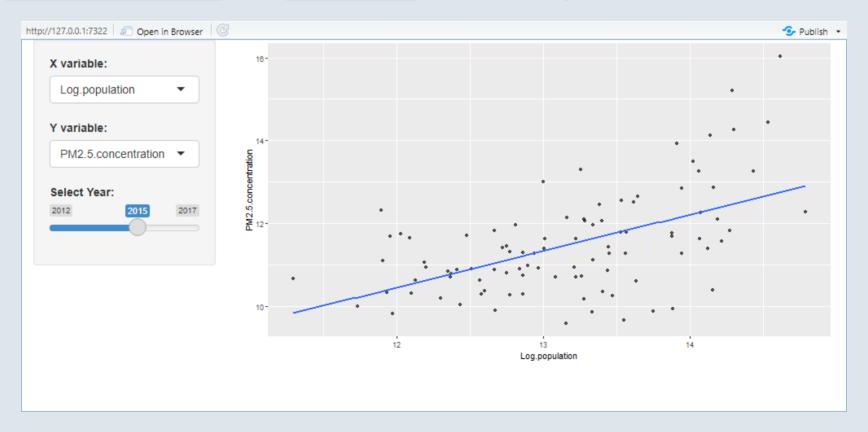


#### 2.2. Interactive plot

- To do an interactive plot that reacts to these inputs we should:
  - 1. Program the plot in a reactive({}) environment for the inputs to be updated
  - 2. Put the reactive plot in a render function to render the updated plot
  - 3. Include the resulting output in the user interface
- Because in aes() input\$xvar and input\$yvar should be treated as variable names instead of strings, they should be put in the get() function

## 2.2. Interactive plot

• Adding plotOutput("plot") in the mainPanel() ui function we get:



#### 2.2. Interactive plot

- But we can make the graph even more interactive using the **plotly** package
  - It allows to have information on a data point in a tooltip on hover
- In the reactive({}) environment we should:

- But we can make the graph even more interactive using the **plotly** package
  - It allows to have information on a data point in a tooltip on hover
- In the reactive({}) environment we should:
  - 1. Put the ggplot in the ggplotly() function

```
ggplotly(
    ggplot(dep_data %>% filter(Year == input$year),
        aes(x = get(input$xvar), y = get(input$yvar))) +
        geom_point(alpha = .6) +

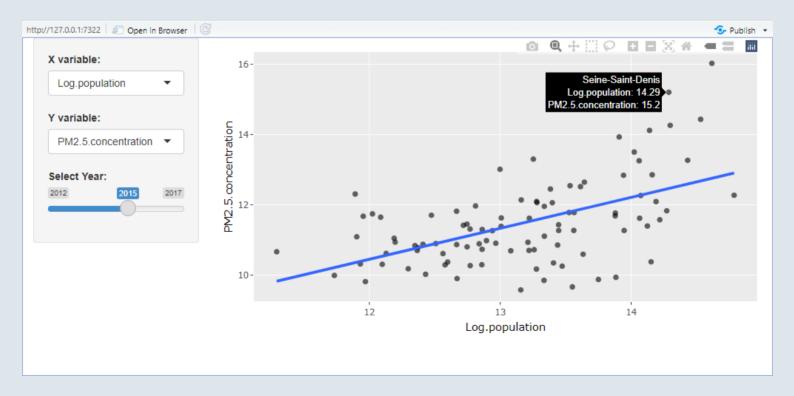
#

geom_smooth(method = "lm", se = F) + xlab(input$xvar) + ylab(input$yvar)
#
```

- But we can make the graph even more interactive using the **plotly** package
  - It allows to have information on a data point in a tooltip on hover
- In the reactive({}) environment we should:
  - 1. Put the ggplot in the ggplotly() function
  - 2. Format the tooltip as the 'text' axis in aes()

- But we can make the graph even more interactive using the **plotly** package
  - It allows to have information on a data point in a tooltip on hover
- In the reactive({}) environment we should:
  - 1. Put the ggplot in the ggplotly() function
  - 2. Format the tooltip as the 'text' axis in aes()
  - 3. Assign the text axis to the tooltip argument of ggplotly()

- We also have to:
  - Replace the server renderPlot() function by renderPlotly()
  - Replace the UI plotOutput() function by plotlyOutput()



#### 2.3. Interactive regression results

- We can also include an interactive regression table
  - We should put the **stargazer()** function in a reactive({}) environment:

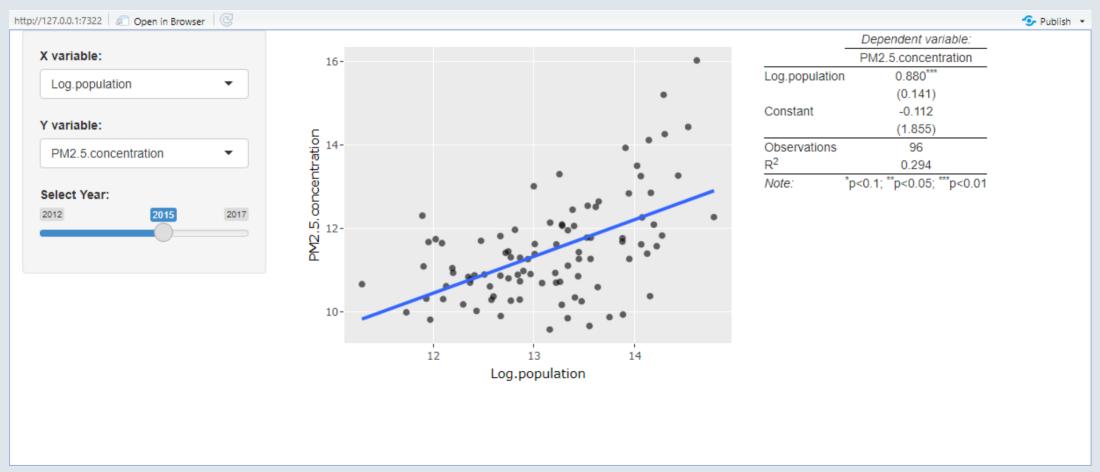
• Then use the render function dedicated to console output: **renderPrint()**:

```
output$reg_table <- renderPrint({reg_table()})</pre>
```

• And use the **htmlOuput()** UI output function in a column layout to put it side to side with the plot:

```
mainPanel(width = 9, column(7, plotlyOutput("plot")), column(5, htmlOutput("reg_table")))
```

## 2.3. Interactive regression results



## Overview

### 1. Introduction to shiny apps ✓

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

### 2. Our first shiny app ✓

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 3. More advanced tools

- 3.1. Input randomization
- 3.2. HTML formatting

## Overview

### 1. Introduction to shiny apps ✓

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

### 2. Our first shiny app ✓

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 3. More advanced tools

- 3.1. Input randomization
- 3.2. HTML formatting

#### 3.1. Input randomization

- → A nice feature would be a button to **randomly select** the x and y **variables** and the year
  - Adding a button is easy, we can simply add an actionButton() widget in the control panel:

- But what should happen in the server when clicking on it is different from what we're used to:
  - We do not want a reactive output to place in the UI
  - We want the selected inputs to change
    - → This is why we're gonna use observeEvent() instead of reactive()

#### 3.1. Input randomization

- The arguments of **observeEvent()** are:
  - The **id** of the input that should trigger the actions
  - The **actions** to take when the input is triggered

```
observeEvent(input$random, {
    # SELECT RANDOM INPUTS
})
```

- The action we want is to change the status of the input widgets
  - This can be done with functions of the form update[SelectInput/SliderTextInput/...]()
  - The first argument should be "session"
  - And the following arguments are those of the widget that we can update

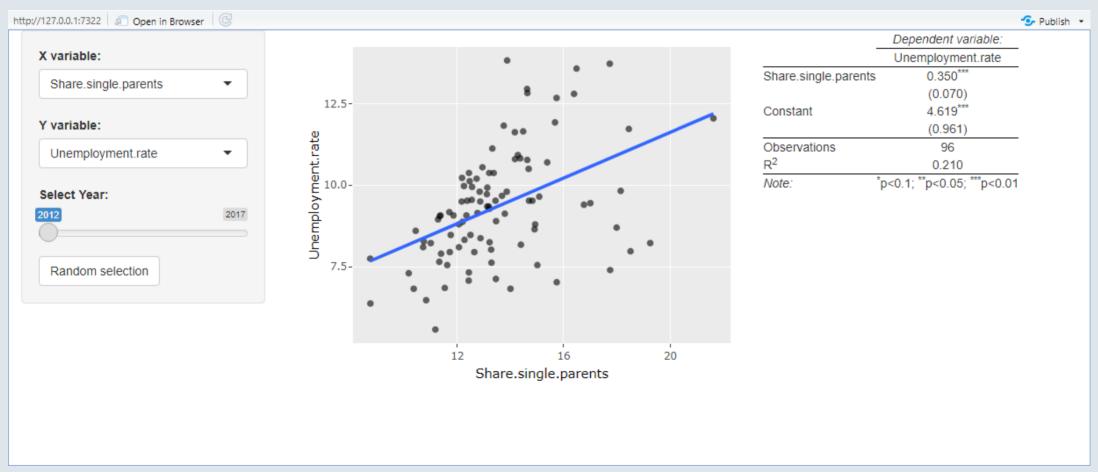
sample (depvars, 1) picks 1 variable name randomly from the vector depvars

### 3.1. Input randomization

• For it to work, the session argument should also be added to the server() function

```
server <- function(input, output, session) {</pre>
  . . .
 observeEvent(input$random, {
   updateSelectInput(session, inputId = "xvar", label = "X variable:",
                      choices = depvars, selected = sample(depvars, 1))
   updateSelectInput(session, inputId = "yvar", label = "Y variable:",
                      choices = depvars, selected = sample(depvars, 1))
   updateSliderTextInput(session, inputId = "year", label = "Select Year:",
                          choices = 2012:2017, selected = sample(2012:2017, 1))
 })
```

### 3.1. Input randomization



#### 3.2. HTML formatting

- We can apply the final touch to our application by including some **html formatting** 
  - To include html in the UI we can simply use the **HTML()** function

#### Some html tags

Text

HTML("<b>Text</b>")

Text

Text

HTML("<i>Text

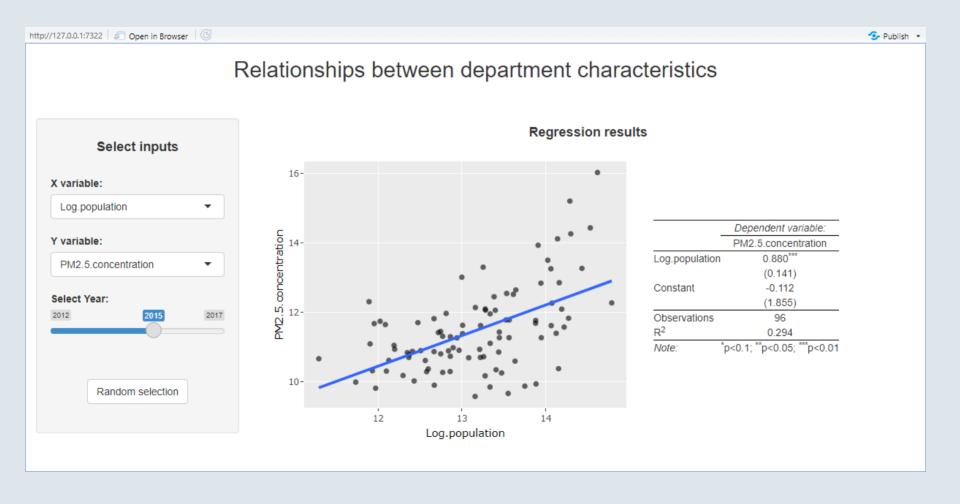
HTML("<center>Text

Text

### 3.2. HTML formatting

```
ui <- fluidPage(</pre>
 HTML("<center><h2>Relationships between department characteristics</h2></center><br>>"),
 sidebarLayout(
   sidebarPanel(width = 3,
     HTML("<center><h4><b>Select inputs</b></h4></center><br>"),
     selectInput(...), selectInput(...), sliderTextInput(...),
     HTML("<br><<br><"), actionButton(...), HTML("</center><br>")
   mainPanel(width = 9,
            HTML("<center><h4><b>Regression results</b></h4></center>"),
            column(width = 7, plotlyOutput("plot")),
```

#### 3.2. HTML formatting



## Overview

### 1. Introduction to shiny apps ✓

- 1.1. General structure
- 1.2. User interface
- 1.3. Server
- 1.4. Layout

### 2. Our first shiny app ✓

- 2.1. Import data in Shiny
- 2.2. Interactive plot
- 2.3. Interactive regression results

#### 3. More advanced tools ✓

- 3.1. Input randomization
- 3.2. HTML formatting

#### **General structure**

### A shiny app is composed of

#### A **user interface** function

• It is what is displayed to the user

0

0

#### A **server** function

It is what should be computed in the background

0

0

#### **General structure**

### A shiny app is composed of

#### A **user interface** function

• It is what is displayed to the user, including:

Input widgets

0

#### A **server** function

• It is what should be computed in the background

0

0

#### **General structure**

### A shiny app is composed of

#### A **user interface** function

- It is what is displayed to the user, including:
  - Input widgets

2

#### A **server** function

- It is what should be computed in the background:
  - Update inputs with reactive({})

0

#### **General structure**

### A shiny app is composed of

#### A **user interface** function

- It is what is displayed to the user, including:
  - Input widgets

0

#### A **server** function

- It is what should be computed in the background:
  - Update inputs with reactive({})
  - Render output with render[Table/Plot/...]()

#### **General structure**

### A shiny app is composed of

#### A **user interface** function

- It is what is displayed to the user, including:
  - Input widgets
  - Reactive outputs

#### A **server** function

- It is what should be computed in the background:
  - Update inputs with reactive({})
  - Render output with render[Table/Plot/...]()

```
library(shiny)
#
ui <- fluidPage(</pre>
server <- function(input, output) {</pre>
shinyApp(ui = ui, server = server)
```

```
library(shiny)
data(iris)
ui <- fluidPage(</pre>
  selectInput(inputId = "x", label = "Select X variable:",
              choices = names(iris), selected = names(iris)[1]),
  selectInput(inputId = "y", label = "Select Y variable:",
              choices = names(iris), selected = names(iris)[2]),
server <- function(input, output) {</pre>
shinyApp(ui = ui, server = server)
```

```
library(shiny)
library(tidyverse)
data(iris)
ui <- fluidPage(</pre>
  selectInput(inputId = "x", label = "Select X variable:",
              choices = names(iris), selected = names(iris)[1]),
  selectInput(inputId = "y", label = "Select Y variable:",
              choices = names(iris), selected = names(iris)[2]),
server <- function(input, output) {</pre>
  reactive_plot <- reactive({</pre>
    ggplot(iris, aes(x = get(input$x), y = get(input$y))) +
      geom point() + xlab(input$x) + ylab(input$y)
 })
shinyApp(ui = ui, server = server)
```

```
library(shiny)
library(tidyverse)
data(iris)
ui <- fluidPage(</pre>
  selectInput(inputId = "x", label = "Select X variable:",
              choices = names(iris), selected = names(iris)[1]),
  selectInput(inputId = "y", label = "Select Y variable:",
              choices = names(iris), selected = names(iris)[2]),
server <- function(input, output) {</pre>
  reactive_plot <- reactive({</pre>
    ggplot(iris, aes(x = get(input$x), y = get(input$y))) +
      geom point() + xlab(input$x) + ylab(input$y)
 })
  output$plot <- renderPlot({reactive_plot()})</pre>
shinyApp(ui = ui, server = server)
```

```
Select X variable:
library(shiny)
                                                                          Sepal.Length
library(tidyverse)
data(iris)
                                                                         Select Y variable:
                                                                          Petal.Length
ui <- fluidPage(</pre>
  selectInput(inputId = "x", label = "Select X variable:",
               choices = names(iris), selected = names(iris)[1]),
  selectInput(inputId = "y", label = "Select Y variable:",
               choices = names(iris), selected = names(iris)[2]),
  plotOutput("plot")
server <- function(input, output) {</pre>
  reactive_plot <- reactive({</pre>
    ggplot(iris, aes(x = get(input$x), y = get(input$y))) +
      geom point() + xlab(input$x) + ylab(input$y)
  })
  output$plot <- renderPlot({reactive_plot()})</pre>
shinyApp(ui = ui, server = server)
```