



Shiny

Web application framework for R

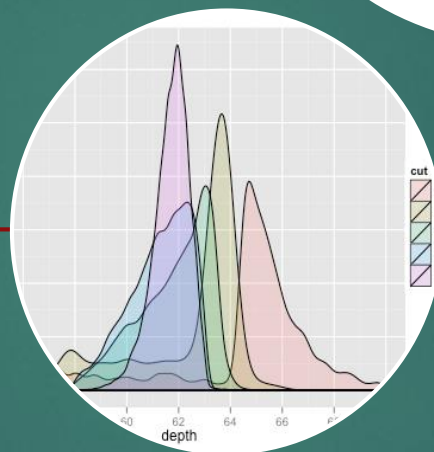
What's in the R ?



$$p(x) = \frac{e^{-m} \times m^x}{x!}$$

Statistical computing :

Modelling,
Classification,
Statistical tests...



Graphs :

Basic plot,
lattice,
ggplot ...

CRAN

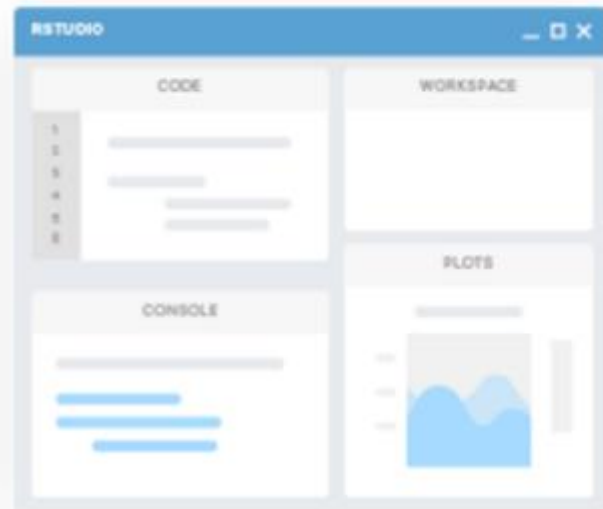
Comprehensive

R

Archive

Network

Some R tools ...



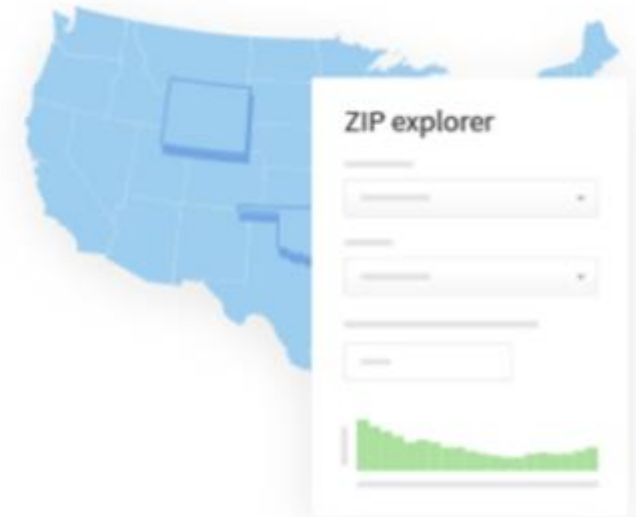
RStudio

RStudio makes R easier to use. It includes a code editor, debugging & visualization tools.



R Packages

Our developers create popular packages to expand the features of R. Includes ggplot2, dplyr, R Markdown & more.



Shiny

Shiny helps you make interactive web applications for visualizing data. Bring R data analysis to life.

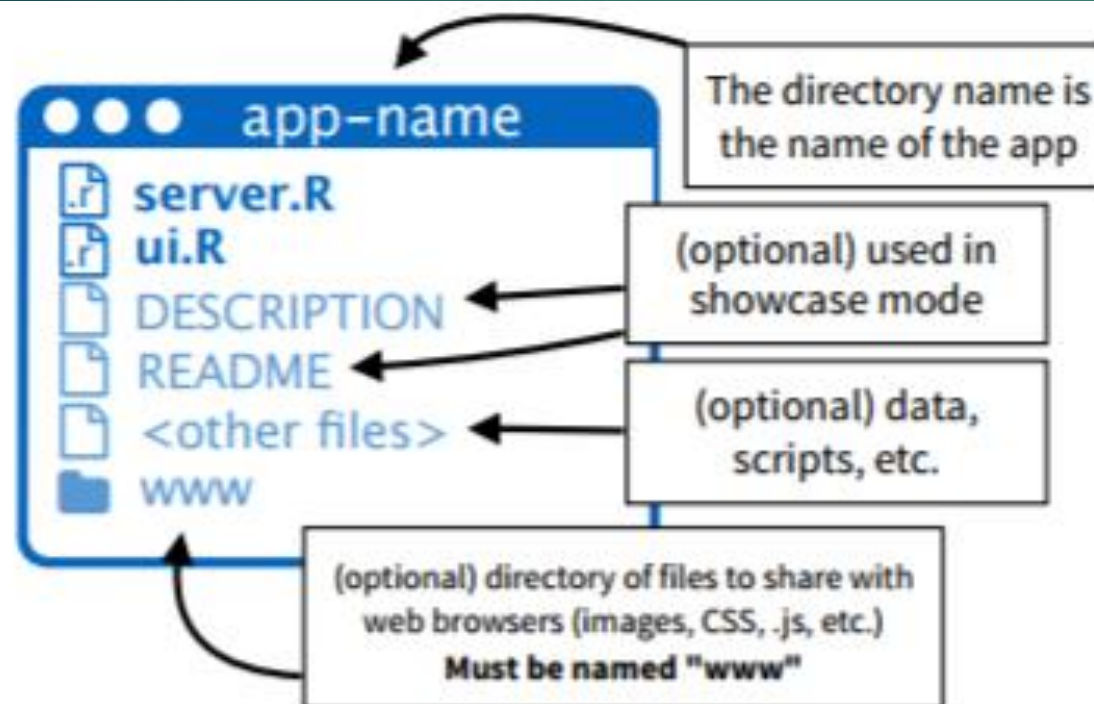
Shiny

- ▶ Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R.
- ▶ Shiny helps you turn your analyses into interactive web applications without requiring HTML, CSS, or JavaScript knowledge.

Shiny's structure

1. Structure

Each app is a directory that contains a server.R file and usually a ui.R file (+optional extra files)



« The form and substance »

2. server.R

The server script contains the instructions that your computer needs to build your app.

3. ui.R

The user-interface (ui) script controls the layout and appearance of your app.

2. server.R

A set of instructions that build the R components of your app

A – the minimum necessary code
shinyServer function + input + output

B – Save R component to share in your UI as
output\$<Component_Name>
Must be unique!

C – Create output with render function
(according to their type, here text and plot)

D – Give to render function the R code the server need to build the component
(here subset of the dataset mtcars)

E – Refer to widget values with
input\$<Widget_Name>

server.R

```
# load libraries, scripts, data

A shinyServer(function(input, output) {

  # make user specific variables

  B output$text <- renderText({
    input$title
  })

  C output$plot <- renderPlot({
    D x <- mtcars[ , input$x]
    E y <- mtcars[ , input$y]
    plot(x, y, pch = 16)
  })

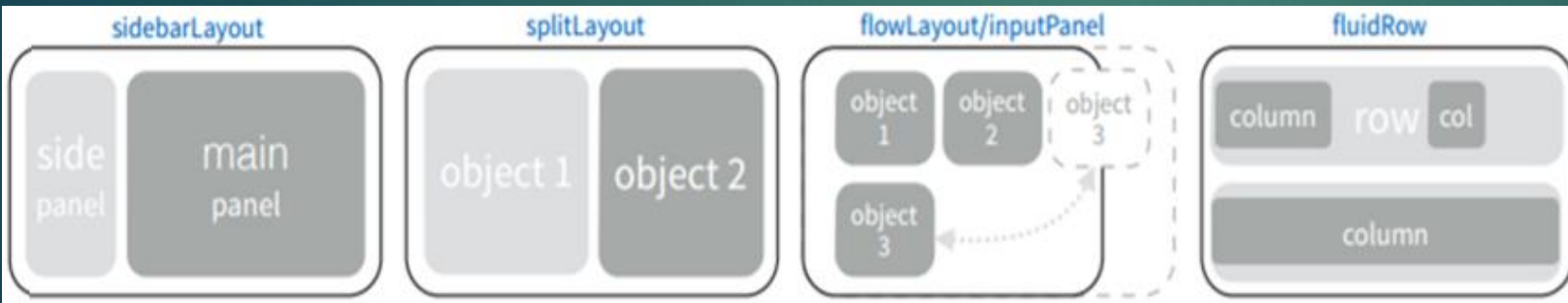
})
```

3. ui.R

A description of your app's User Interface (UI),
The web page that displays your app.

A – the minimum necessary code
shinyUI and **fluidPage** functions

B – Build a layout for your UI



C – Create widgets
Must be unique also!

D – Place R components (the output objects that you
defined in server.R



Output\$plot <- renderPlot({ ... }) ↔ plotOutput("plot")

ui.R

```
A shinyUI(fluidPage(  
  titlePanel("mtcars data"),  
  B sidebarLayout(  
    sidebarPanel(  
        
      textInput("title", "Plot title:",  
        value = "x v y"),  
        
      C selectInput("x", "Choose an x var:",  
        choices = names(mtcars),  
        selected = "disp"),  
        
      selectInput("y", "Choose a y var:",  
        choices = names(mtcars),  
        selected = "mpg")  
    ),  
    mainPanel(  
      D h3(textOutput("text")),  
        plotOutput("plot")  
    )  
  )  
))
```

4. Run your app

runApp – run from local files

runGitHub – run from files hosted on www.GitHub.com

runGist – run from files saved as a gist gist.github.com

runURL – run from files saved at any URL

5. Share your app

Launch your app as a live web page that users can visit online

ShinyApps.io

Host your apps on RStudio's server. Free and paid options
www.shinyapps.io

Shiny Server

Build your own linux server to host apps. Free and open source.
shiny.rstudio.com/deploy

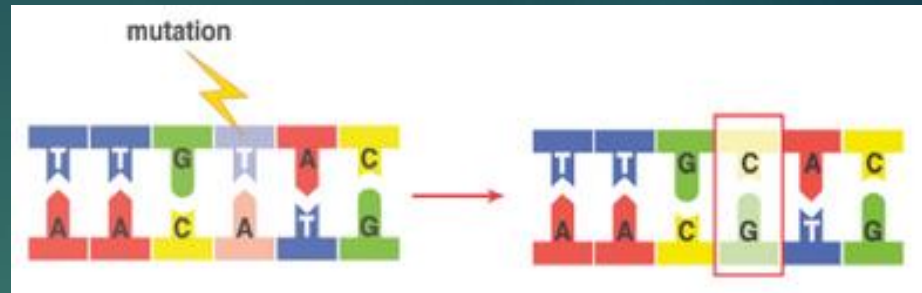
Shiny Server Pro

Build a commercial server with authentication, resource management, and more.
shiny.rstudio.com/deploy

Feedback

Context

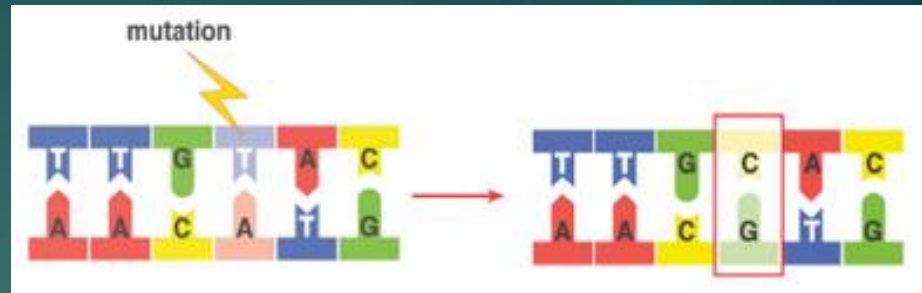
Project : Find a link between sequencing data (more than 2.000 genes, that is to say 800.000 mutations) and phenotypes (disease as diabetes), on 10.000 patients. It is call a case/control study because we are looking for genetic variants (i.e. mutations) associated with cases (diabetic patients) but not with controls (healthy patients). Lots of different strategies are possible to tackled to explore this huge database ... So lots of results will have to be visualized, and only a few may be conclusive !



Feedback

Context

Project : Find a link between sequencing data (more than 2.000 genes, that is to say 800.000 mutations) and phenotypes (disease as diabetes), on 10.000 patients. It is call a case/control study because we are looking for genetic variants (i.e. mutations) associated with cases (diabetic patients) but not with controls (healthy patients). Lots of different strategies are possible to tackled to explore this huge database ... So lots of results will have to be visualized, and only a few may be conclusive !



Aim of the App

- Launch the analysis pipeline automatically and in a reproducible way (like automatic pilot)
- Make the decisions and the selection of the parameters easier for the user (choose a strategy)
- Produce transitional and exportable results (view the tables and graphs..)
- Export a report automatically filled (if you have made an interesting discovery, save it in a Word report and share it!)

✉ Mathilde Boissel
(Biostatistician)

🌐 DATA and Parameters

📄 Methodology

📊 Analysis

📄 Save the results

Data selection

Focus on a gene

ABCDE

Choose a phenotype to analyse

CC_Diabete

Is the variable binary ?

- ☒ YES
- ☐ NO (it is a quantitative trait)

Choose a group of covariables

At least one must be checked

- ☒ SEX
- ☐ AGE
- ☐ BMI

Threshold for rare snps

By default, threshold is set to 1%



For the analysis of rare snps, which cluster(s) do you want?

- ☒ Make a unique cluster (at the gene scale)
- ☐ Discriminate snps on their consequences

Exclude synonymous snps from analyses ?

- ☐ YES
- ☒ NO

Mathilde Boissel
(Biostatistician)

DATA and Parameters

Methodology

Analysis

- » Data summary
- » Rare Analysis
- » Frequent Analysis

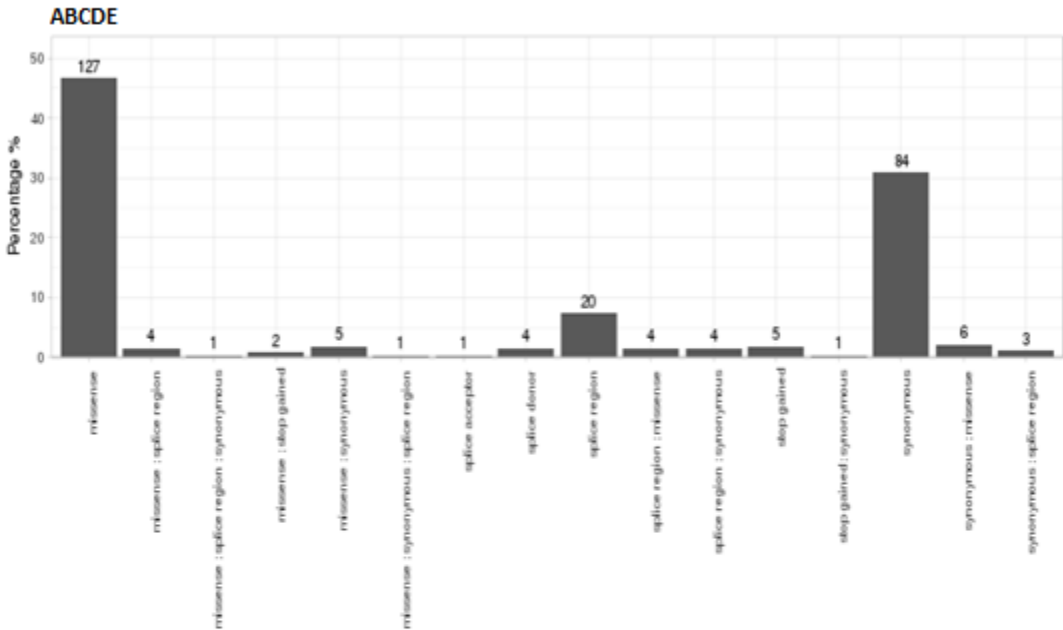
Save the results

Phenotypes available (in the whole project)

Trait	Controls	Cases
CC_Obesity	N or Mean (sd)	N or Mean (sd)
N	1,068	2,439
AGE	46 (10.8)	49.1 (13.1)
BMI	22.3 (1.83)	37.1 (7.36)
SEX	M:976 / F:1,463	M:376 / F:692
FG	5.48 (0.831)	5.15 (0.503)

Trait	Controls	Cases
CC_Diabetes	N or Mean (sd)	N or Mean (sd)
N	3,220	2,025
AGE	51 (10.6)	58.7 (11.5)
BMI	28.5 (5.89)	25.9 (5.78)
SEX	M:1,346 / F:1,874	M:1,161 / F:864
FG	5.11 (0.423)	8.06 (2.1)

Distribution of SNPs annotated



Gene description

Show 10 entries

Search:

	SNP	CHR	POS	Ref/Alt alleles	Missing genotype	Genotype 0	Genotype 1	Genotype 2	MAF	Type	AA change
1	11_174145_C.T	11	174145	C/T	59	9333	1	0	0.0000535676023141204	missense	V1578I ; V1577I
2	11_174145_G.A	11	174145	G/A	59	9319	15	0	0.000803514034711806	synonymous	F/F

Model

Association between CC_Diabetes and SNPs is tested and adjusted for SEX, AGE. All variants are analysed as a single cluster (at the gene scale e.g. ABCDE) without subcluster based on mutation's type. Note: Synonymous variants have been kept for the analyses.

Parameters from the rare analysis

Clusters	Pi hat	CI 2.5%	CI 97.5%	OR
Gene	0.56492	0.37244	0.75954	1.75931

Stat MiST Analysis

Score π	Score π P value	Score τ	Score τ P value	Overall P value
34.10133	5.231541e-9	117.71963	0.000031	5.03930 e-12

Phenotype

Trait	Controls	Cases
CC_Diabetes	N or Mean (sd)	N or Mean (sd)
N	3081	1956
SEX	M:1295 / F:1786	M:1120 / F:836
AGE	51.1 (10.6)	58.7 (11.5)

Genotype

After quality control (removing variants with missing genotypes or variance null), 175 variants were analysed.

Model

Association between CC_Diabetes and SNPs is tested and adjusted for SEX, AGE. All variants are analysed as a single cluster (at the gene scale e.g. ABCDE) without subcluster based on mutation's type. Note: Synonymous variants have been kept for the analyses.

Parameters from the rare analysis

Clusters	Pi hat	CI 2.5%	CI 97.5%	OR
Gene	0.56492	0.37244	0.75954	1.75931

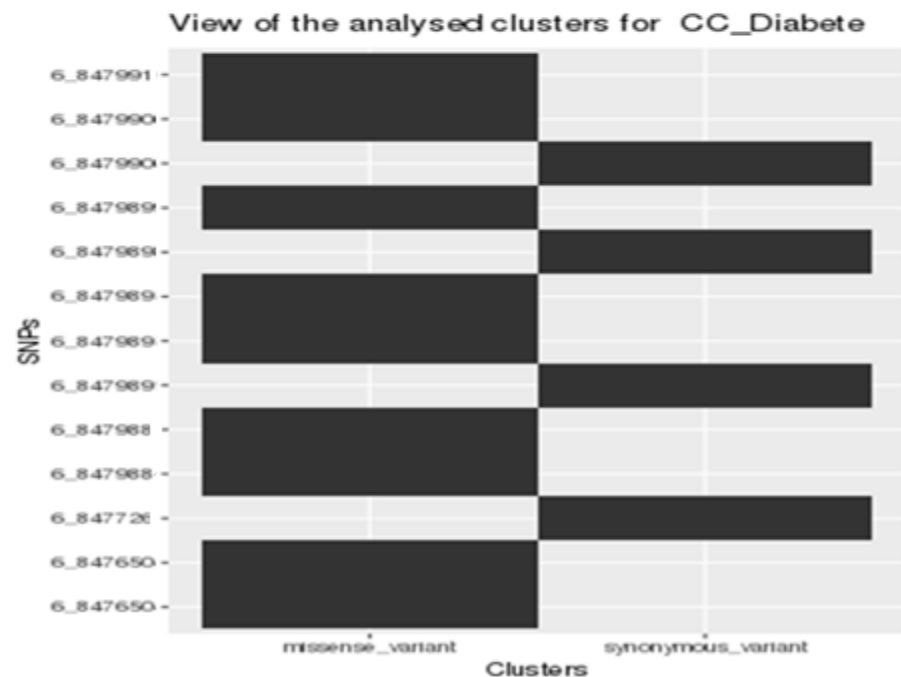
Stat MiST Analysis

Score π	Score π P value	Score τ	Score τ P value	Overall P value
34.10133	5.231541e-9	117.71963	0.000031	5.03930 e-12

Phenotype

Trait	Controls	Cases
CC_Diabetes	N or Mean (sd)	N or Mean (sd)
N	3081	1956
SEX	M:1295 / F:1786	M:1120 / F:836
AGE	51.1 (10.6)	58.7 (11.5)

Genotype



Parameters from the frequent analysis

Y name	Covar names	SNP	Beta	CI 2.5%	CI 97.5%	SE	P value	OR
CC_Diabetes	SEX + AGE	11_174145_C.T	-0.0757431481983427	-0.247507883731516	0.0941796334093889	0.087126249416729	0.384656078445749	0.927054291546812
CC_Diabetes	SEX + AGE	11_174184_C.A	0.0133154230922582	-0.0735530462201233	0.100333584648954	0.044352787402314	0.764012286731366	1.01340446812347
CC_Diabetes	SEX + AGE	11_174192_C.T	0.129193085014022	0.0367989387917354	0.221549542023665	0.0471227942150148	0.00611358794726859	1.13790981628825
CC_Diabetes	SEX + AGE	11_174281_G.A	-0.397523644292917	-0.756697985584093	-0.0501108609142992	0.179891942697669	0.0271196519747857	0.671982053919233
CC_Diabetes	SEX + AGE	11_174342_G.A	0.0902745443592244	-0.0535072777360506	0.233206022207621	0.073116507174846	0.216954432809784	1.09447472432306
CC_Diabetes	SEX + AGE	11_174368_G.A	-0.170052972188977	-0.407326514193013	0.0626965695079752	0.119806408186241	0.155783024859879	0.843620127007949
CC_Diabetes	SEX + AGE	11_174487_G.A	-0.0165083458786036	-0.100805519053416	0.0676987318050064	0.0429801389077352	0.700909943970258	0.983627170123463
CC_Diabetes	SEX + AGE	11_174499_C.T	-0.0593012307104572	-0.181951404498309	0.0625431057894172	0.0623550186882068	0.3415912857241	0.942422839693558
CC_Diabetes	SEX + AGE	11_174524_G.A	-0.0317652112162269	-0.117279068580874	0.0536582051590679	0.0436005979091227	0.466277034647059	0.968734003258607
CC_Diabetes	SEX + AGE	11_174917_G.A	0.128677394977473	-0.0955348207180027	0.350818021582389	0.113773018556639	0.258054680832863	1.13732315881321

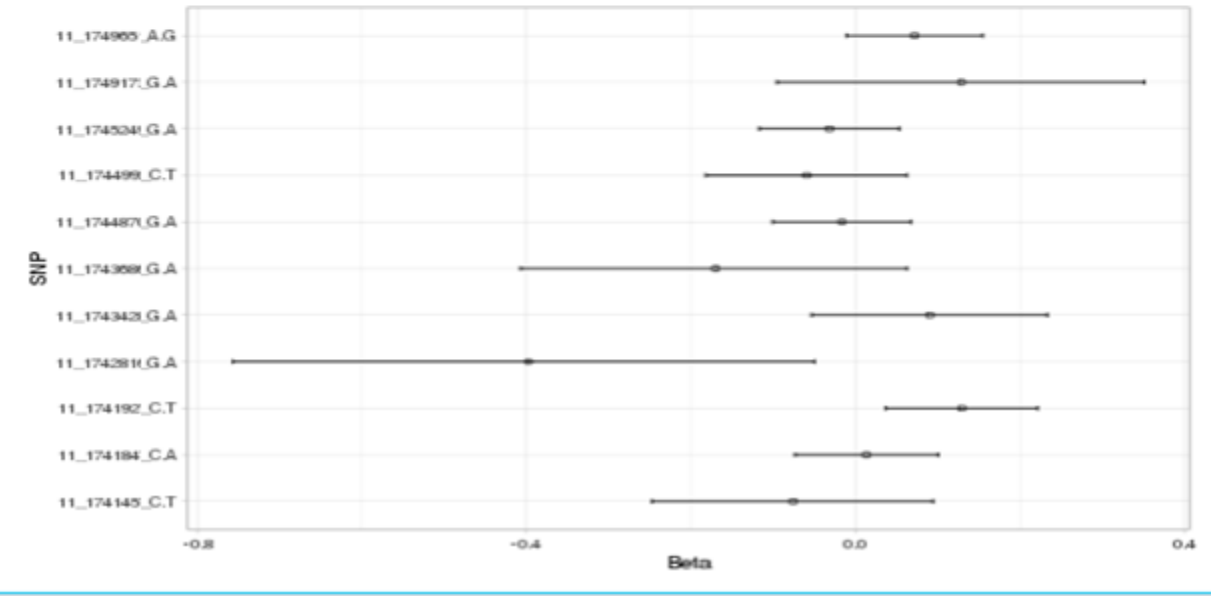
Phenotype

Show 10 entries

Search:

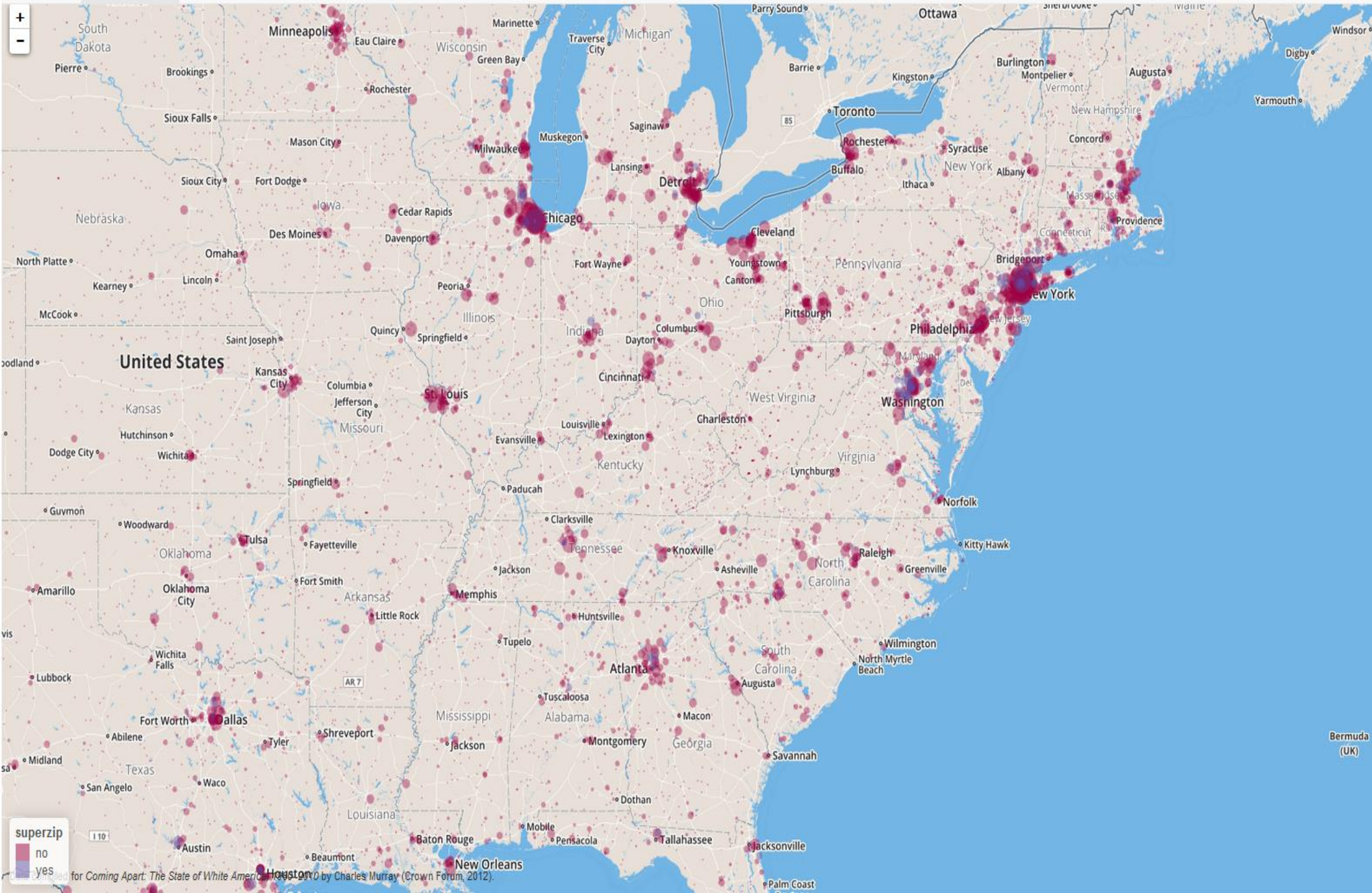
	Distribution.per.SNPs	CC_Diabetes = 0	CC_Diabetes = 1
1	11_174145_C.T	3198	2018
2	11_174184_C.A	3198	2018
3	11_174192_C.T	3197	2018
4	11_174281_G.A	3197	2018
5	11_174342_G.A	3198	2018
6	11_174368_G.A	3196	2016
7	11_174487_G.A	3184	2007
8	11_174499_C.T	3177	2006
9	11_174524_G.A	3197	2018

Estimations and Confidence intervals



Other inspiring apps in shiny's gallery ...

<https://shiny.rstudio.com/gallery/>



ZIP explorer

Color

Is SuperZIP?

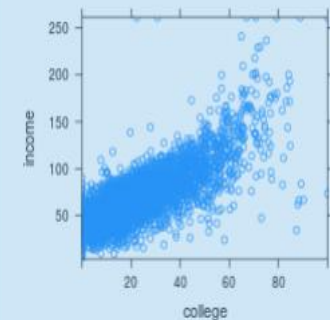
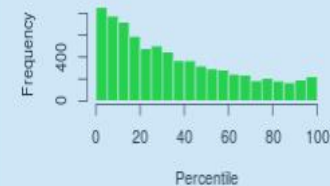
Size

Population

SuperZIP threshold (top n percentile)

5

SuperZIP score (visible zips)

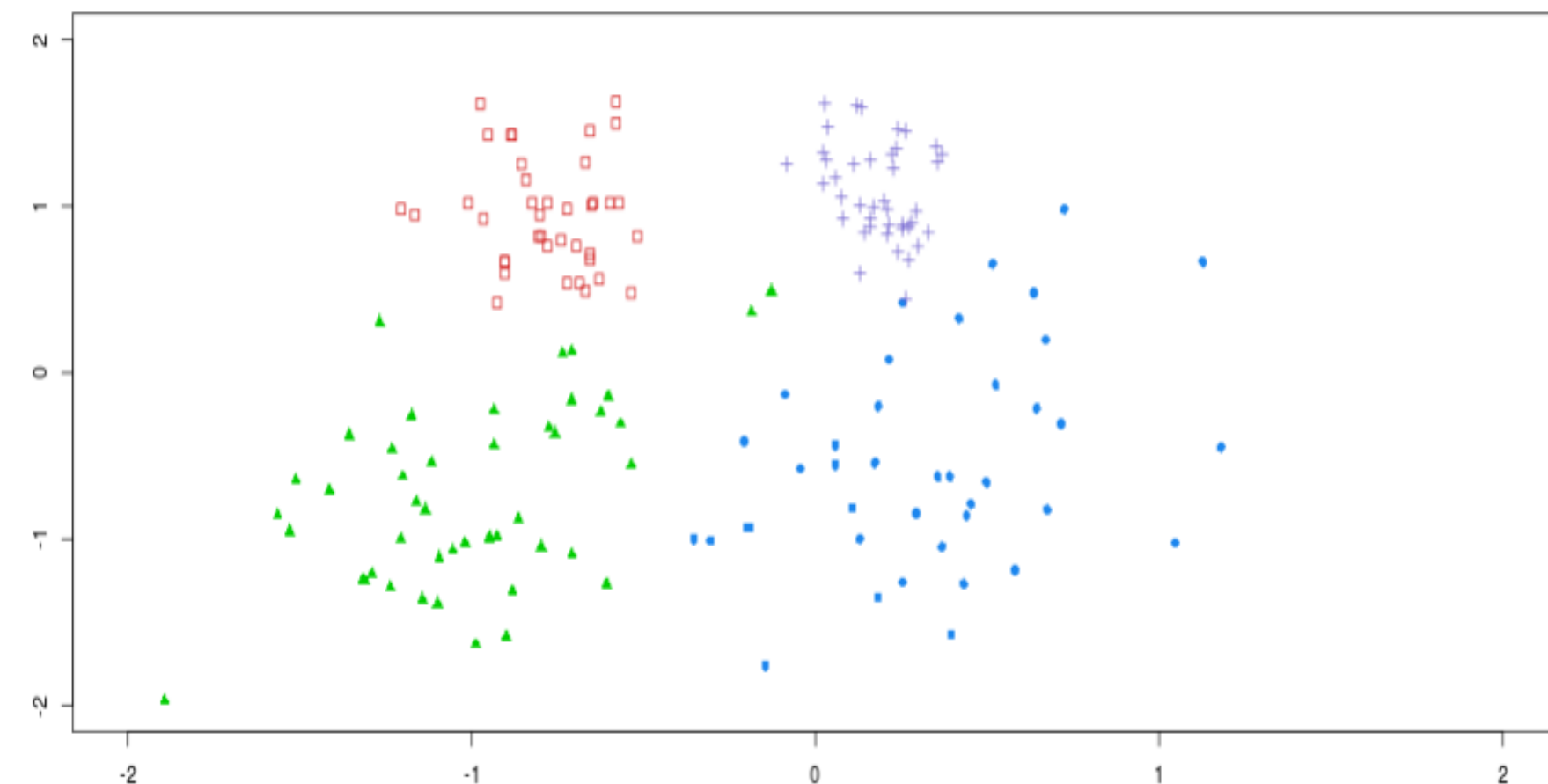


superzip

no
yes

For Coming Apart: The State of White America by Charles Murray (Crown Forum, 2012).

Dynamic Clustering in Shiny



Points:

178

Clear Points

server.R

ui.R

show below

```
library(shiny)
library(mclust)

function(input, output, session) {

  # Create a spot where we can store additional
  # reactive values for this session
  val <- reactiveValues(x=NULL, y=NULL)

  # Listen for clicks
  observe({
    # Initially will be empty
    if (is.null(input$clusterClick)){
      return()
    }

    isolate({
      val$x <- c(val$x, input$clusterClick$x)
      val$y <- c(val$y, input$clusterClick$y)
    })
  })

  # Count the number of points
  output$numPoints <- renderText({
    length(val$x)
  })

  # Clear the points on button click
  observe({
    if (input$clear > 0){
```

Thank you !

And thanks to...

Mickaël Canouil,
Amélie Bonnefond,
Stefan Gaget,
Cécile Lecoeur,
All the team's members who have
worked on this NGS project from UMR8199



Links :

<https://cran.r-project.org>

<https://www.rstudio.com>

<http://shiny.rstudio.com/>

<http://shiny.rstudio.com/gallery/>

