Blablacar Case

From data to Business

Deusto
Universidad de Deusto

LIN3S

BlaBlaCar

Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

Ander Fernández

- Business Administration Graduate at Deusto Business School.
- Studied Big Data & Business Intelligence Program (2019-2020)
- I work as Data Scientist at LIN3S.
- Winner of the UniversityHack 2020.
- 1.5 years ago I didn’t know how to code.

anderfernandezj@gmail.com

anderfernandez.com/blog/

Clients at LIN3S

Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

Context

Blablacar is a French company dedicated to facilitating car sharing established in Spain since 2010.

Blablacar wants to explot the data generated with journey offers and demand to improve the business of the company.

For this, Blablacar has provided us with the following information (see next page)

Context - Data

DIA	PAIS	ORIGEN	DESTINO	IMP_KM	ASIENTOS_OFERTADOS	ASIENTOS_CONFIRMADOS	VIAJES_OFERTADOS	VIAJES_CONFIRMADOS	OFERTANTES	OFERTANTES_NUEVOS
01/11/2017	es	A Coruña	Santiago de Compostela	0.055813953	13	1	6	1	5	0
01/11/2017	es	A Guarda	Tui	NA	3	0	1	0	1	0
01/11/2017	es	A Gudiña	Vila Nova de Famalicão	NA	3	0	1	0	1	1
01/11/2017	es	Aguilar de Ca	Madrid	NA	10	0	5	0	5	0
01/11/2017	es	Alatoz	Valencia/València	NA	3	0	1	0	1	0
01/11/2017	es	Albacete	Alhambra	NA	3	0	1	0	1	0
01/11/2017	es	Albacete	Almansa	0.053333333	13	1	5	1	5	0
01/11/2017	es	Albacete	Elda	NA	8	0	3	0	3	0
01/11/2017	es	Albarracín	Elche/Elx	NA	2	0	1	0	1	0
01/11/2017	es	Alcalá de He	Badajoz	NA	3	0	1	0	1	0
01/11/2017	es	Alcorcón	Hellín	NA	2	0	1	0	1	0
01/11/2017	es	Alfaro	Sestao	NA	1	0	1	0	1	0
01/11/2017	es	Almansa	Petrer	NA	1	0	1	0	1	0
01/11/2017	es	Almeida	Valladolid	NA	3	0	1	0	1	0
01/11/2017	es	Almería	Alhama de Murcia	NA	3	0	1	0	1	0
01/11/2017	es	Almería	Granada	0.053814147	68	22	26	13	25	0
01/11/2017	es	Almería	Totana	NA	3	0	1	0	1	0
01/11/2017	es	Alp	Lleida	NA	2	0	1	0	1	0
01/11/2017	es	Alzira	Bicorp	NA	3	0	1	0	1	1
01/11/2017	es	Alzira	Chella	NA	3	0	1	0	1	1
01/11/2017	es	Antequera	Olvera	NA	2	0	1	0	1	0
01/11/2017	es	Aranda de Du	Palencia	NA	3	0	1	0	1	0
01/11/2017	es	Aranda de Du	Santander	NA	0	0	1	0	1	0
01/11/2017	es	Aranjuez	Albacete	NA	4	0	1	0	1	0
01/11/2017	es	Aranjuez	Villena	NA	4	0	1	0	1	0

EDA main results

- There are a lot of route with high offer and low demand.
- Both offer and demand are not equally distributed among all weekdays.
- Trips occurring in a constant bases are more likely to be covered that those than are spontaneous.
- Not all provinces are equal: some have several important cities with large amount of travels while others concentrate most travels in one or two cities.

What information could we add to the one that we already have?

Data enrichment

- **Date-related data:**
  - Basic Data: day of the week, month, week, year...
  - External Data: Hollydays, Holliday eves...

- **Trip Data**:
  - Population: province, Comunidad Autónoma, inhabitants...
  - Geo: Coordinates, geometries
  - Trips: distance, time...

- **New calculated data**
  - Route probability: total, by day, by day type
  - Relative data: %confirmed trips, etc

Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

Know that we know what we have…

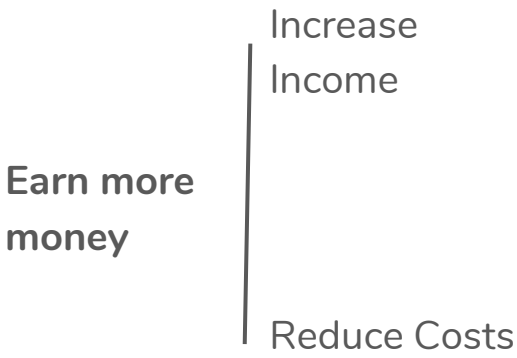
Know that we know what we have...

¿What can we do to improve Blablacar's business with this data?

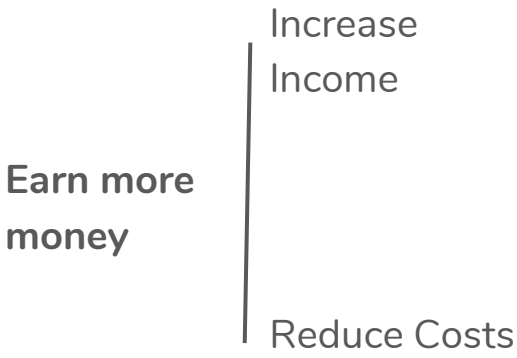
How can we earn more money?

Earn more money

How can we earn more money?

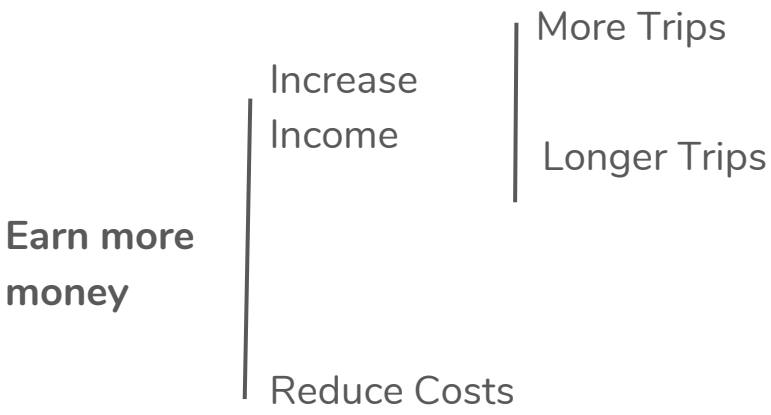


How can we earn more money?



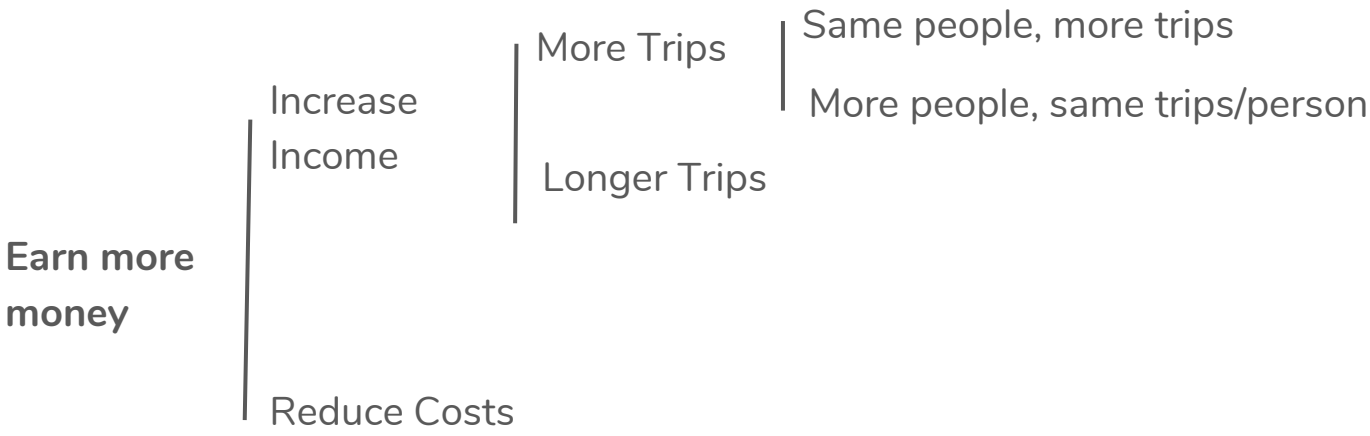
Note: Blablacar’s business model is to charge a 10-15% comission for each trip.

How can we earn more money?



Note: Blablacar's business model is to charge a 10-15% comission for each trip.

How can we earn more money?



Note: Blablacar's business model is to charge a 10-15% comission for each trip.

Why do people don’t use Blablacar?

Why do people don’t use Blablacar?

Safety

There is not the route that I want

I can't plan the trip

I’ve never think about it

Price

I’m a shameful person

All problem has a solution

Problem	Solution
There is not the route that I want	Expand routes by creating connections to be able to go further.
I can't plan the trip	Obtain the probability of a trip on a given day.
I’ve never think about it	Know the penetration of Blablacar by localities to know strong and weak regions.

Our Idea

1. **Ir más lejos**: a new service that (1) predicts the probability of that trip to happen and (2) finds a short and highly probable alternative.

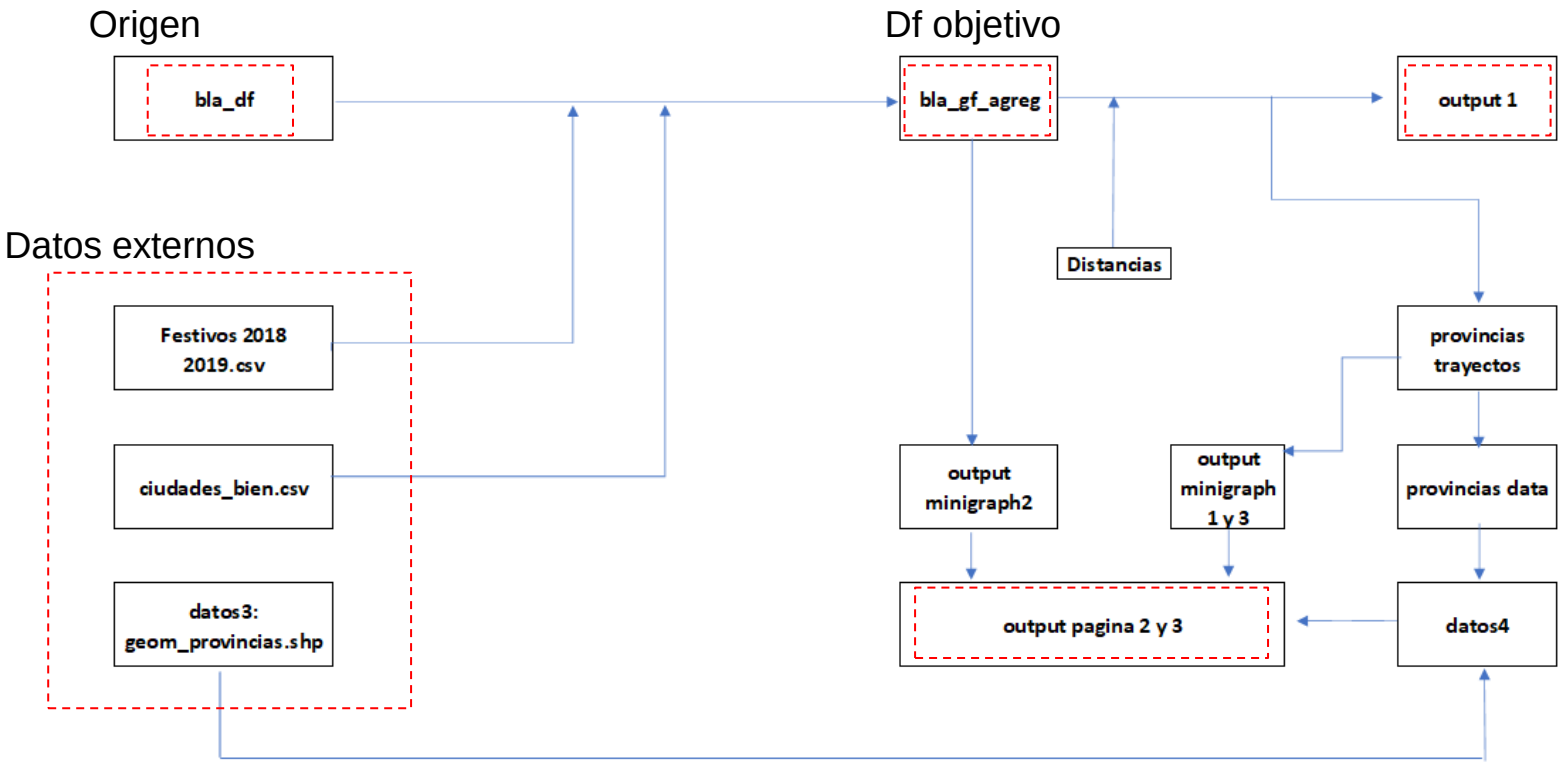
Our Idea

1. **Ir más lejos**: a new service that (1) predicts the probability of that trip to happen and (2) finds a short and highly probable alternative.
2. **Ir a más gente**: a responsive dashboard showing the use of Blablacar in different provinces and the municipalities within this provinces.

Our Idea

1. **Ir más lejos**: a new service that (1) predicts the probability of that trip to happen and (2) finds a short and highly probable alternative.
2. **Ir a más gente**: a responsive dashboard showing the use of Blablacar in different provinces and the municipalities within this provinces.
3. **Cuidando el planeta**: a dashboard showing the saving in the environmental impact thanks to Blablacar in each municipality.

Flujo de datos



Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

What data do we need?

In our case we decided to show:

- **The chosen route**: it has a high visual impact and it is vital to differenciate between the direct route and the alternative route.
- **Provinces**: it offers a high visual impact with a low storage cost. We could show municipality data with normal graphs.

Working with Cartociudad

```
library(caRtociudad)
```

We will use the following functions:

- `cartociudad_geocode`: enables to retrieve geospatial data given a string.
- `cartociudad_get_route`: enables to extract a route given two coordinates.

Visualizing maps with Leaflet

```
library(leaflet)
```

Leaflet allows you to create maps using layers, like ggplot. In this way, we will put the route on the map and include a world map and the names of the cities to see the data well.

In addition, it is a very widespread library that is also used in Python and Javascript. This makes it a library with many interesting options in Shiny.

Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

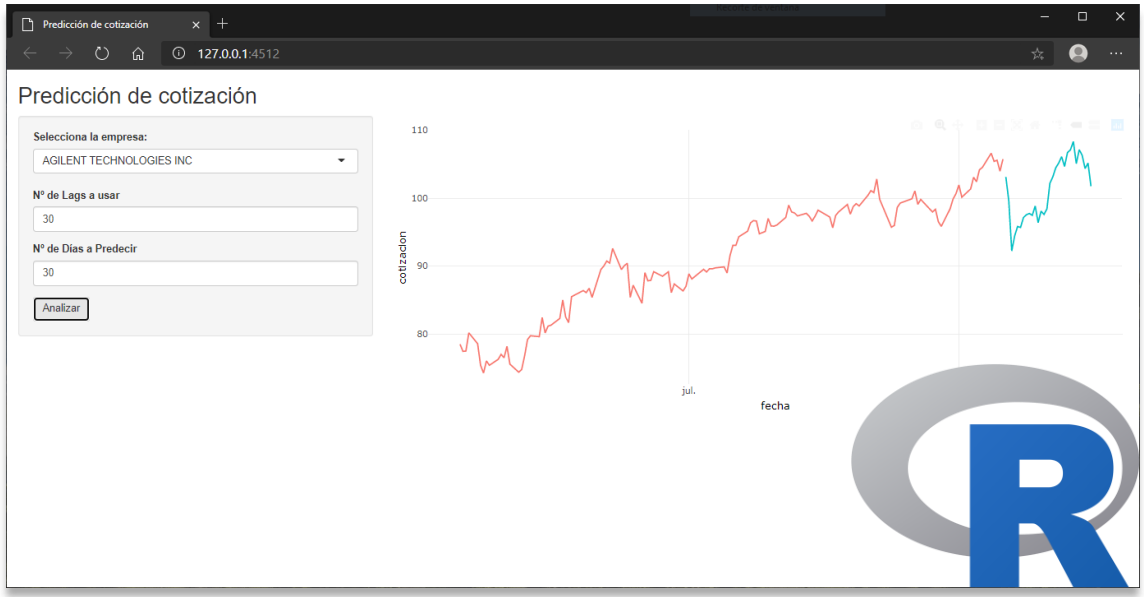
Shiny

Shinyapps if a perfect toot for this case because:

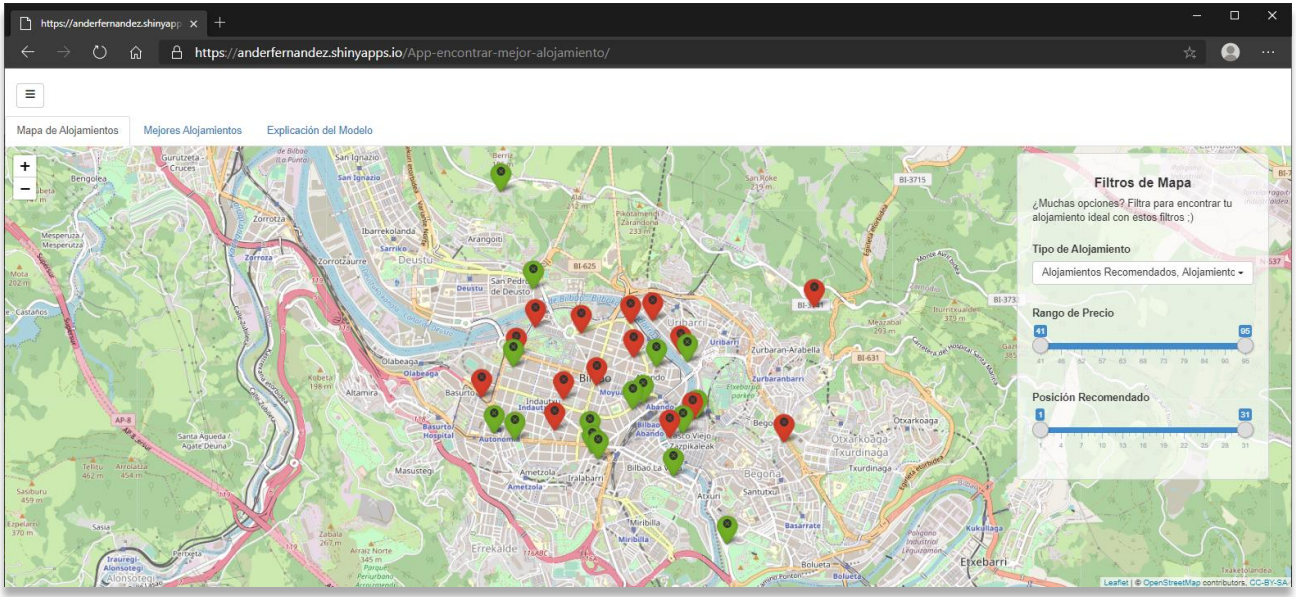
- Enables to show data in a very visual and interactive way. Besides, it enables to include designs and functionalities via HTML, CSS & Javascript.
- Enable to run any algorith or code built in R, opening a lot of possibilities.
- Putting the app in production with Shinyapps.io is easy and free.

Enlace: https://datmen.shinyapps.io/Datmen/

Shiny allows you to create and publish web applications that run R code with which any user can interact



With a single language we can create applications that extract the data through APIs, scraping or access to databases to train and visualize Machine Learning models



App Structure

Initial Part

In this part we load the libraries and data that remain static from the loading of the App.

```
library(shiny)
```

App Structure

Initial Part

In this part we load the libraries and data that remain static from the loading of the App.

```
library(shiny)
```

User Interface

In this part we will create the HTML structure of the page, including: action buttons, sliders, selectors, and everything that the user can enter and that we are going to use on the page.

```
ui <- fluidPage(
  titlePanel(“Hello World!")
)
```

App Structure

Initial Part

In this part we load the libraries and data that remain static from the loading of the App.

```
library(shiny)
```

User Interface

In this part we will create the HTML structure of the page, including: action buttons, sliders, selectors, and everything that the user can enter and that we are going to use on the page.

```
ui <- fluidPage(
  titlePanel("Hello World!")
)
```

Server

This is where we run the code in R and generate dynamic results and charts that will be displayed in the UI section.

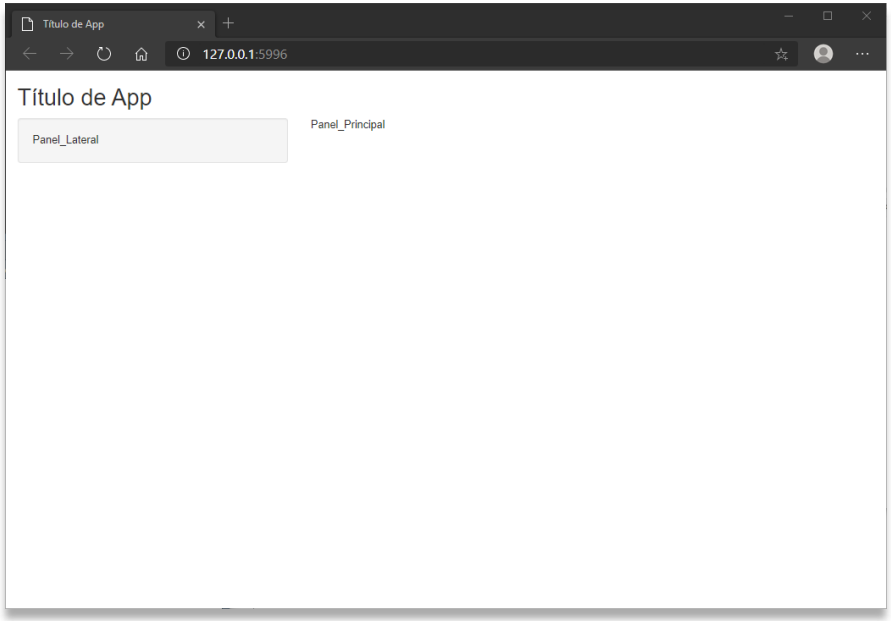
```
server <- function(input, output) {
}
```

Creating the UI

Although we could create the structure we want or even load an HTML document, Shiny offers functions with basic structures.

Example:

```
ui <- fluidPage(
   titlePanel('Título de App'),
   sidebarLayout(
        sidebarPanel('Panel_Lateral'),
        mainPanel('Panel_Principal')
)
)
```

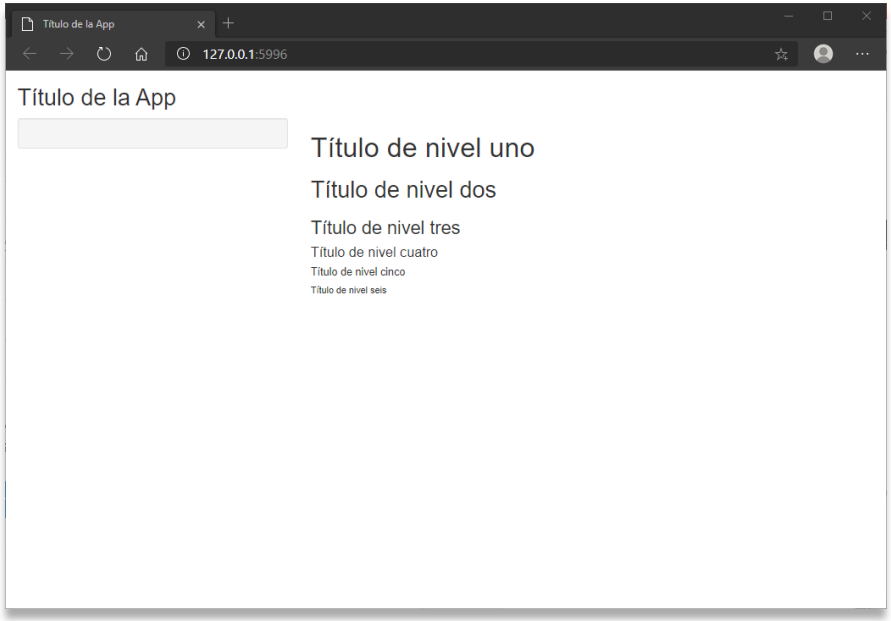


Creating the UI

In addition, in case we do not know about HTML, it offers many functions that allow you to create the UI in HTML5 directly from R.

Example:

```
ui <- fluidPage(
    titlePanel("Título de la App"),
    sidebarLayout(
        sidebarPanel(),
        mainPanel(
     h1("Título de nivel uno"),
     h2("Título de nivel dos"),
     h3("Título de nivel tres"),
     h4("Título de nivel cuatro"),
     h5("Título de nivel cinco"),
     h6("Título de nivel seis")
        )))
```

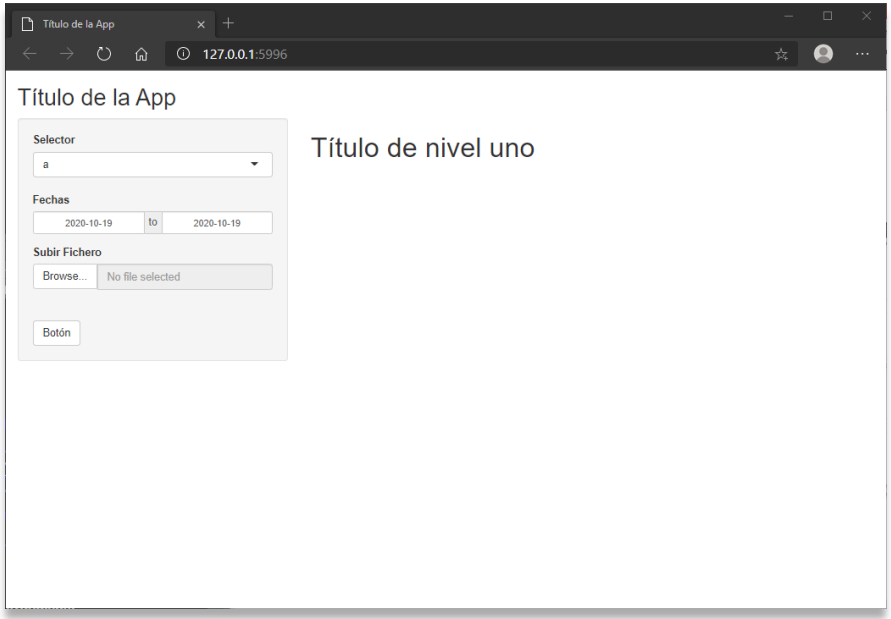


Creating the UI

Likewise, we can also create elements that serve as user inputs, such as: an action button, a selector

Example:

```
ui <- fluidPage(
    titlePanel("Título de la App"),
    sidebarLayout(
        sidebarPanel(
selectInput("a","Selector",letters),
dateRangeInput("b","Fechas"),
fileInput("c","Subir Fichero"),
actionButton("d","Botón")
        ),
        mainPanel(
            h1("Título de nivel uno")
        )
```

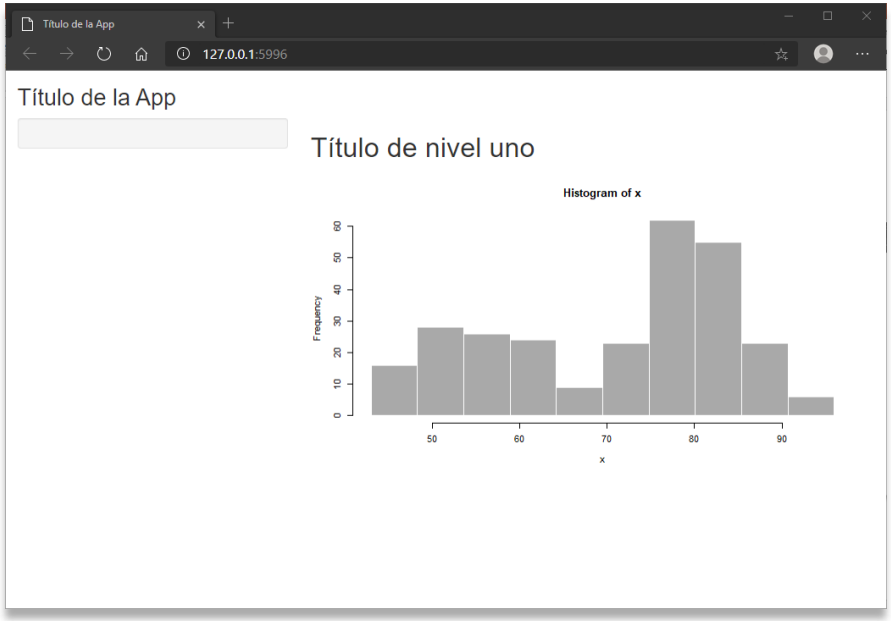


Creating the UI

Finally, we have to indicate where the elements that we create on the server will go. We will do this with the Output functions:

Example:

```
ui <- fluidPage(
    titlePanel("Título de la App"),
    sidebarLayout(
        sidebarPanel(),
        mainPanel(
            h1("Título de nivel uno"),
            plotOutput("distPlot")
        )
    )
)
```



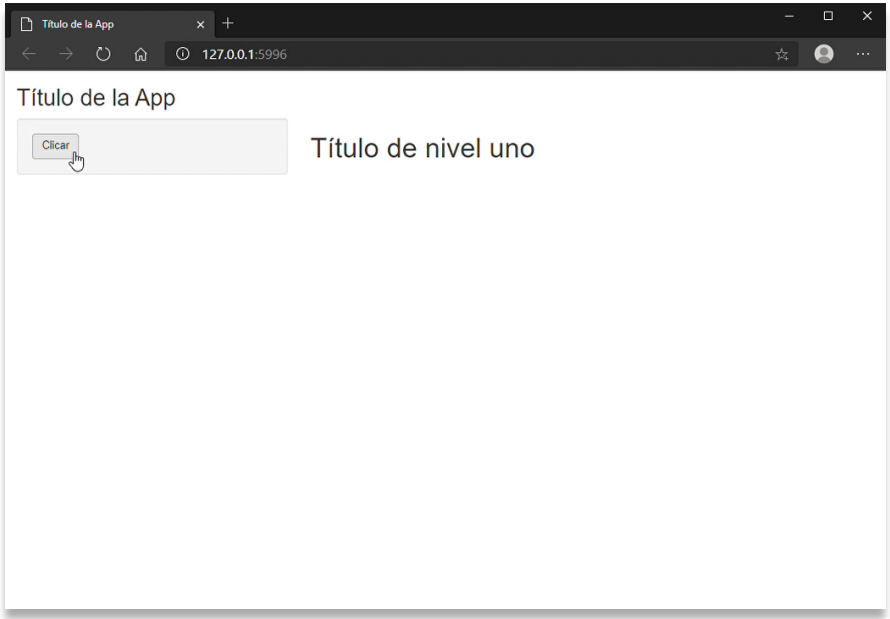
Creating the Server

Reactive elements: They are elements that are updated every time there is any change in the UI or other reactive element or when an element that we indicate changes.

```
datos = eventReactive(input$accion,{
        x = sample(50)
})
```

Render Elements: consists of rendering an element that we will later show in the UI. In the previous case, if when we used plotOutput a graph was displayed, it is because that graph had been rendered on the server.

```
output$texto = renderText(
        datos()
    )
```



Code of the previous example

```
library(shiny)

ui <- fluidPage(
    titlePanel("Título de la App"),
    sidebarLayout(
        sidebarPanel(
            actionButton("accion","Clicar")
        ),
        mainPanel(
            h1("Título de nivel uno"),
            textOutput("texto")
        )
    )
)
```

```
server <- function(input, output) {

    datos = eventReactive(input$accion,{
        x = sample(50)
    })

    output$texto = renderText(
        datos()
    )
shinyApp(ui = ui, server = server)
```

Let’s build an app

Structure of the App

Initial Part

- We install caRtociudad.
- We load the libraries.

UI

- We create a sidebar with a text input and an action button.
- We include the output of the visualization that we will build.

Server

- We extract the geodata of the input we have included.
- We create and render a map with this info.

Structure of the App

Initial Part

- We install caRtociudad.
- We load the libraries.

```
# Install caRtociudad

library("devtools")

devtools::install_github("rOpenSpain/caRtociudad")


# Load libraries

library(caRtociudad)

library(leaflet)

library(sf)

library(geojsonio)

library(googleway)

library(shiny)
```

Structure of the App

UI

- We create a sidebar with a text input and an action button.
- We include the output of the map that we will build.

```
ui <- fluidPage(

    titlePanel("Visualiza tu localidad"),

    # Sidebar with text input and action button

    sidebarLayout(

        sidebarPanel(

            textInput("localidad", "Elige tu localidad:"),

            actionButton("boton",label = "Analizar")

        ),

        # Render de Map

        mainPanel(

           leafletOutput("mapa")

        )

    )

)
```

Structure of the App

Server

- We extract the geodata of the input we have included.
- We create and render a map with this info.

```
server <- function(input, output, session) {

    datos_mapa = eventReactive(input$boton ,{

        ciudad = input$localidad

        map_data = cartociudad_geocode(ciudad)

        plot_poligon = poligono_donostia = st_as_sfc(map_data$geom)

    })


    output$mapa = renderLeaflet({

        datos_mapa = datos_mapa()

        leaflet(datos_mapa) %>%

            addMapPane("background_map", zIndex = 410) %>%

            addMapPane("polygons", zIndex = 420) %>%

            addProviderTiles(providers$Stamen.TonerLite) %>%

            addPolygons(fillColor = "blue", stroke=TRUE, fillOpacity = 0.5)

    })

}
```

Index

1. Presentation
2. Case Context
3. From Data to Business
4. Working with Geo Data
5. Building a Shiny App
6. Taking Blablacar to the next level

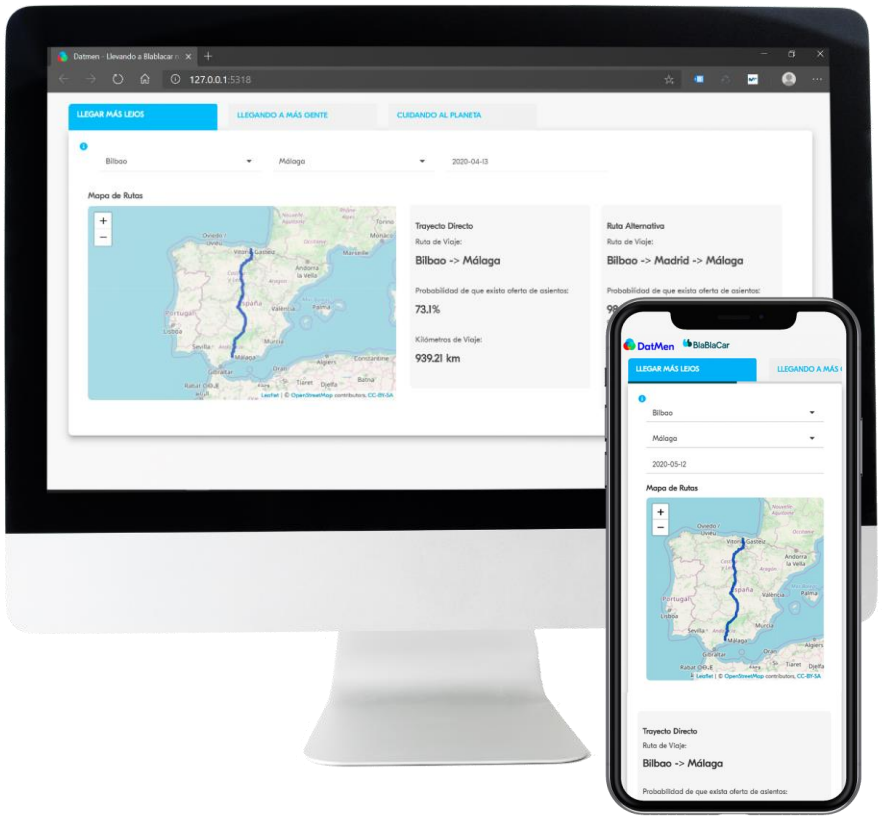
Llegar más lejos,
llegar a más gente,
cuidando el planeta

URL del proyecto
https://datmen.shinyapps.io/Datmen/

Llegar más lejos

- Hay rutas que son poco probables que se den.

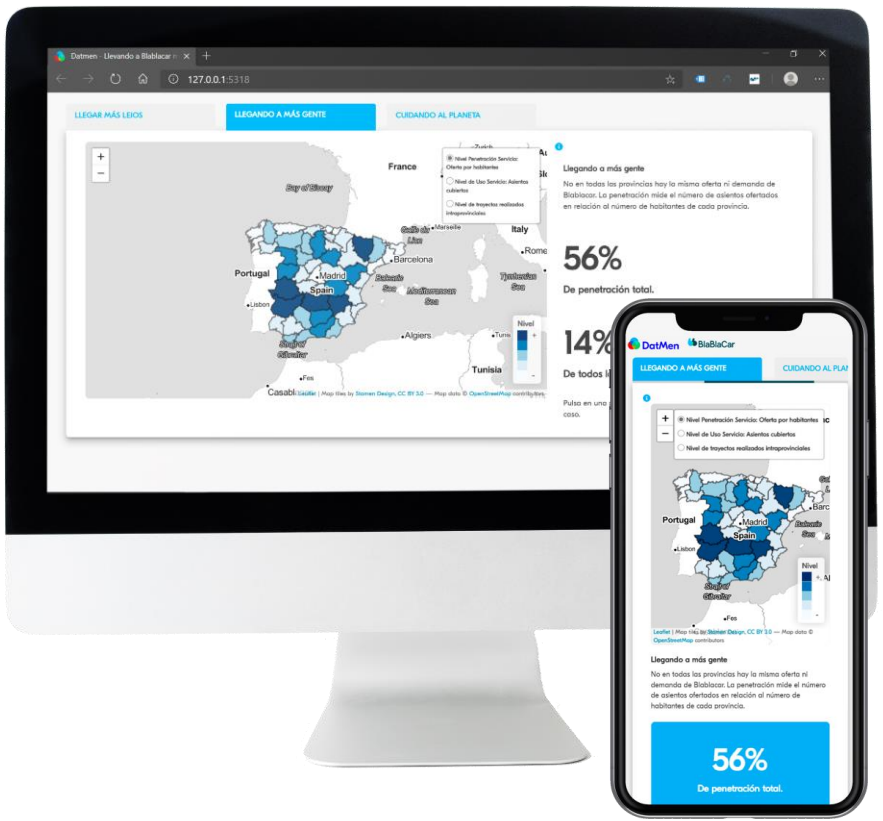
- Vamos a buscar rutas alternativas cortas y probables



Llegar a más gente

..pero, ¿cómo y dónde publicito yo este proyecto?

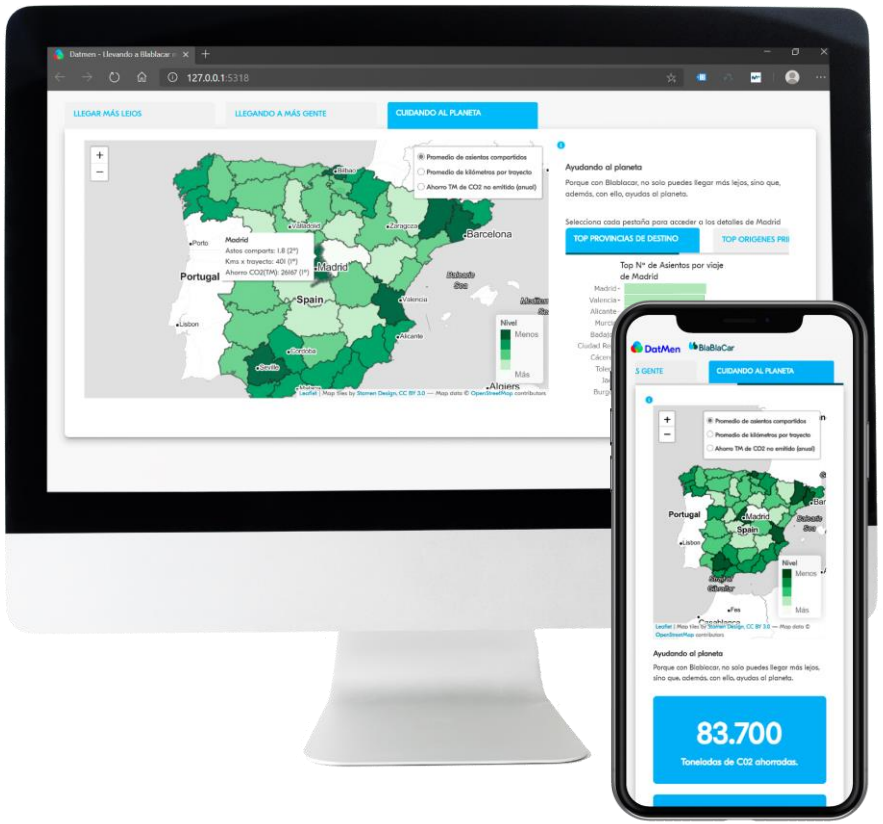
Hemos creado un dashboard para saber la penetración y uso de Blablacar en distintas provincias



Cuidando el planeta

..pero, ¿cómo puedo llegar a esa gente?

Vamos a calcular y promocionar el ahorro ambiental generado para cada localidad



Llegar más lejos, llegar a más gente, cuidando el planeta.

DatMen