



David Garcia Almenzar

Licenciado en Matemáticas, Master de Estadística Aplicada con R Software

GRUPO CAJAMAR • Universidad de Almería

 <https://www.linkedin.com/in/david-garcia-almenzar-159499129>

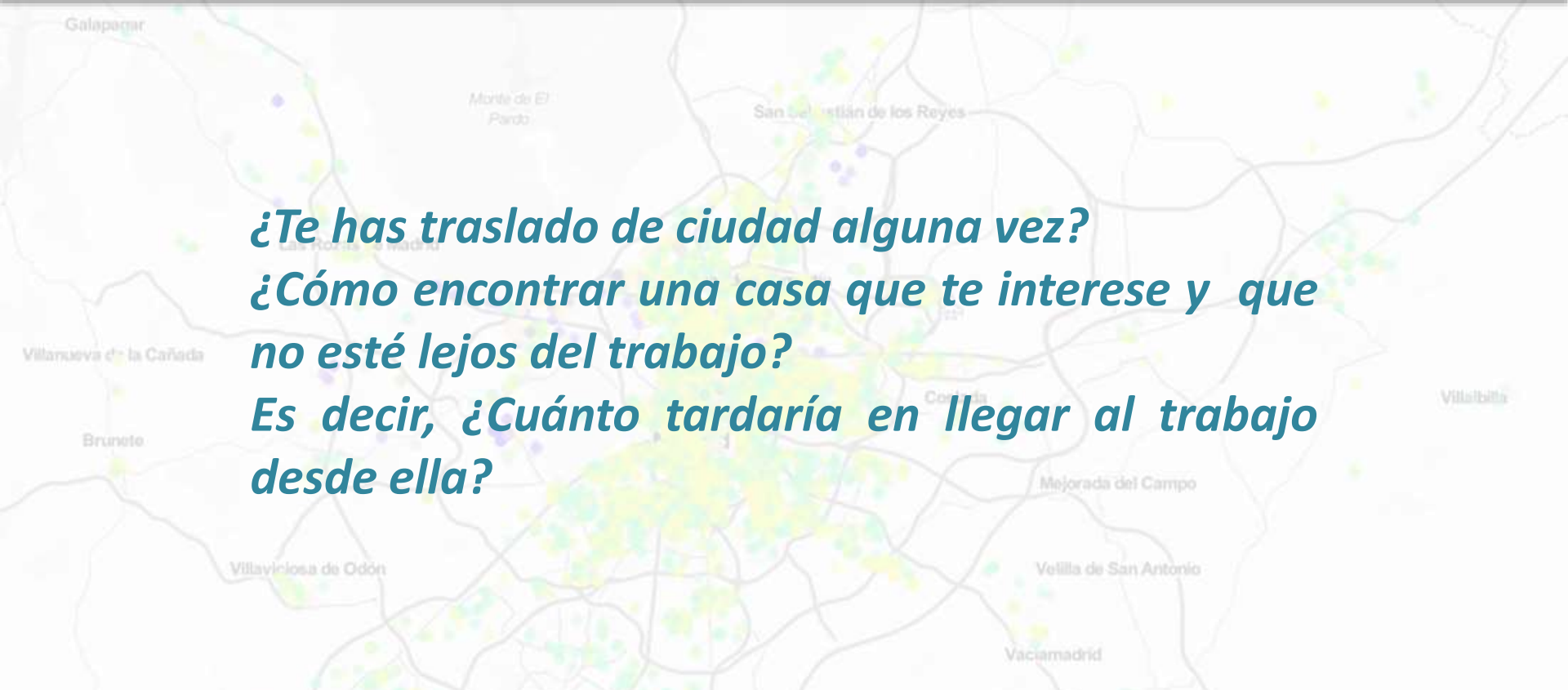
 @dgalmenzar



IoT con R: Alquilar una casa

IX JORNADAS DE USUARIOS DE R

Granada, 16 y 17 de noviembre de 2017



*¿Te has trasladado de ciudad alguna vez?
¿Cómo encontrar una casa que te interese y que
no esté lejos del trabajo?
Es decir, ¿Cuánto tardaría en llegar al trabajo
desde ella?*

IoT con R: Alquilar una casa



IoT con R: Alquilar una casa

Galapagar



Villavieja



Villavieja

Vaciamadrid



IoT con R: Alquilar una casa

Viviendas de interés

Importe: 512 - 1.000 (0,000 - 10.000)

Núm habitaciones: 2 - 3 (0 - 14)

Tiempo máximo de desplazamiento (min): 15 (0 - 120)

¿Como vas a desplazarte al trabajo?

- WALK
- PUBLIC
- CAR

Dirección	Importe *	Num. Hab *	Metros	Distancia (m)	Tiempo (min) *	Medio Trans
4524 calle de Lazaga, 21, Cuatro-Casillejos, Madrid	650	2	46	1.26	14.87	WALK
8596 Carranza, Trafalgar, Madrid	675	2	40	3.46	11.82	CAR
385 calle de María Teresa Sáenz de Heredia, 9, Ventas, Madrid	700	3	60	5.63	14.53	CAR
1271 plaza de redondela, 13, Pilar, Madrid	700	2	70	4.82	12.32	CAR
15873 calle de Ramón Gómez de la Serna, 121, Peñagrande, Madrid	700	2	50	5.57	12.17	CAR
15296 Nuevos Ministerios-Ríos Rosas, Madrid	700	2	45	0.72	9.42	WALK
12370 Cartagena, 110, Prosperidad, Madrid	720	2	37	2.91	8.22	CAR
15609 Ambroz, Madrid	730	2	61	9.34	14.52	CAR
3623 calle de Raimundo Fernández Vialverde, Cuatro Caminos, Madrid	730	2	36	0.7	8.53	WALK
1076 san marcos, Chueca-Justicia, Madrid	750	2	40	4.23	14.97	CAR

Showing 1 to 10 of 427 entries

Mapa: ES: 1553 Alquiler de 1000 euros/mes. Num. hab: 2

Nota: El desarrollo está realizado para las viviendas en alquiler de la comunidad de Madrid disponibles en el portal de Idealista con datos de julio respecto a una dirección determinada (Paseo de la Castellana 87).

Librerías de extracción

library(rvest) # para realizar scraping

- **read_html()** # para leer el html de una url
- **html_nodes()** # permite usar el selector de css en el html
- **html_text()** # transforma el html en texto

library(placement) # para trabajar con la api de google maps

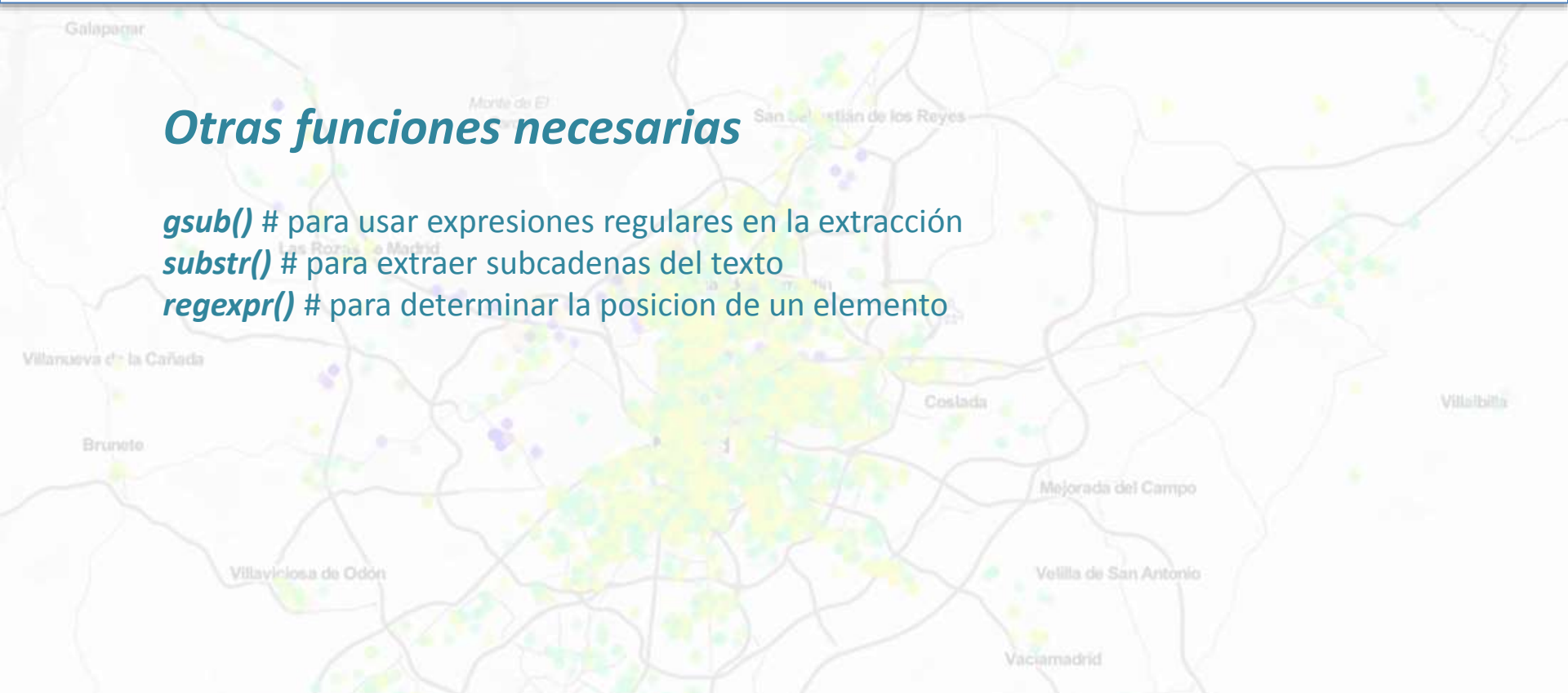
- **geocode_url()** # para obtener las coordenadas de una dirección
- **drive_time()** # para obtener el tiempo y la distancia entre dos direcciones según el medio de transporte

Otras funciones necesarias

gsub() # para usar expresiones regulares en la extracción

substr() # para extraer subcadenas del texto

regexr() # para determinar la posición de un elemento



IoT con R: Alquilar una casa

Limpieza del Dataset

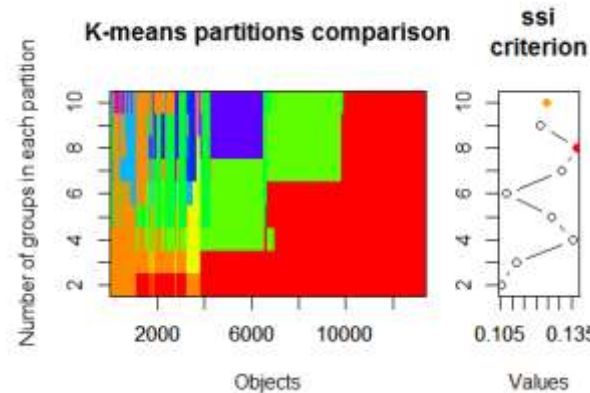
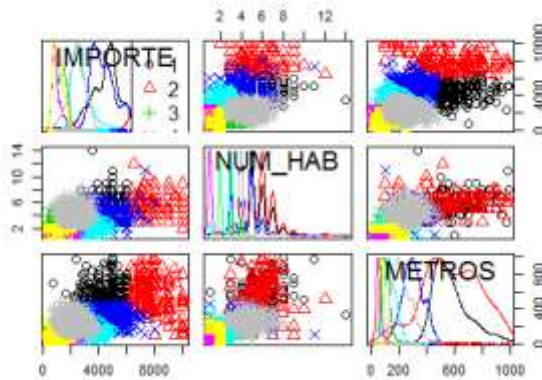
Una vez realizada la extracción disponemos de un dataset con 16.101 registros totales

Motivo	Num reg eliminados	Registros
Importe elevado (sup los 10.000€)	48	16.053
Tamaño erroneo (hab. de menos 5m)	9	16.044
No se ha podido obtener todos los datos	2.369	13.675
El geoposicionamiento es erroneo	242	13.433



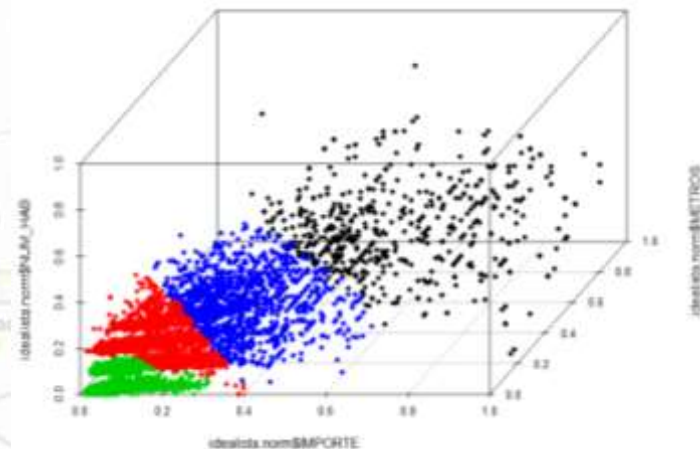
Análisis del Dataset

Con los datos depurados realizamos un análisis de grupos a través de un K-means sobre las variables, importe, tamaño y número de habitaciones.



IoT con R: Alquilar una casa

VAR	CLASS	DEF CLASS	Min.	1st Qu.:	Median :	Mean :	3rd Qu.:	Max.
IMPORTE	1	UNIFAMILIAR	290	707	890	964	1150	3030
	2	FAMILIAR	230	950	1475	1496	1900	4000
	3	FAMILIAR +	700	2600	3100	3238	3700	6495
	4	ALTO STANDING	2300	4500	5500	5880	7310	10000
NUM HAB	1	UNIFAMILIAR	1	1	1	1,43	2	2
	2	FAMILIAR	1	3	3	3	3	6
	3	FAMILIAR +	1	3	4	4	5	8
	4	ALTO STANDING	1	5	5	5,61	6	14
M2	1	UNIFAMILIAR	12	48	60	61,9	75	250
	2	FAMILIAR	15	90	110	113,9	130	300
	3	FAMILIAR +	50	187	231	243	300	700
	4	ALTO STANDING	100	430	530	549	650	990



Librerías de visualización

library(shiny) # para crear una web interactiva

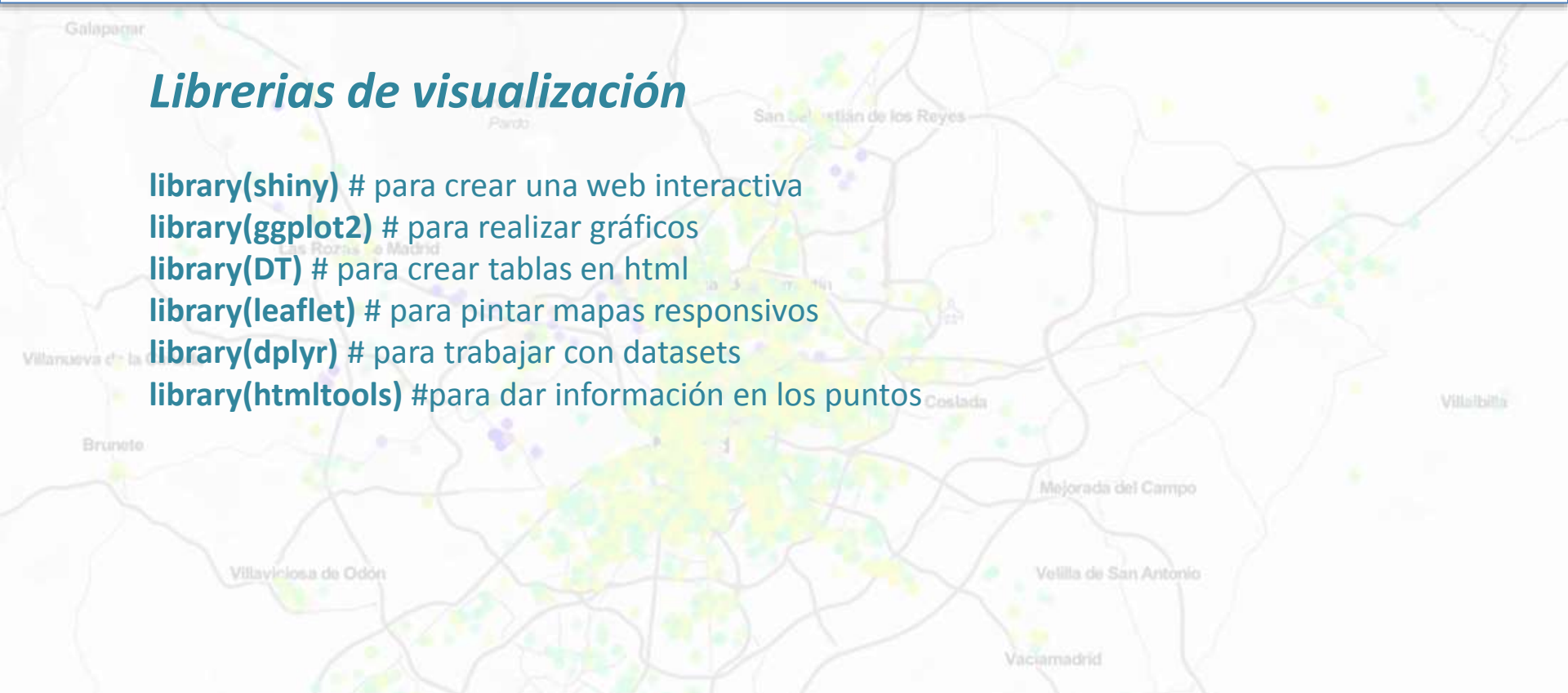
library(ggplot2) # para realizar gráficos

library(DT) # para crear tablas en html

library(leaflet) # para pintar mapas responsivos

library(dplyr) # para trabajar con datasets

library(htmltools) # para dar información en los puntos



IoT con R: Alquilar una casa

Viviendas de interés

Importe: 500 — 1.000 (range up to 10.000)

Núm habitaciones: 3 (range up to 14)

Tiempo máximo de desplazamiento (min): 15 (range up to 120)

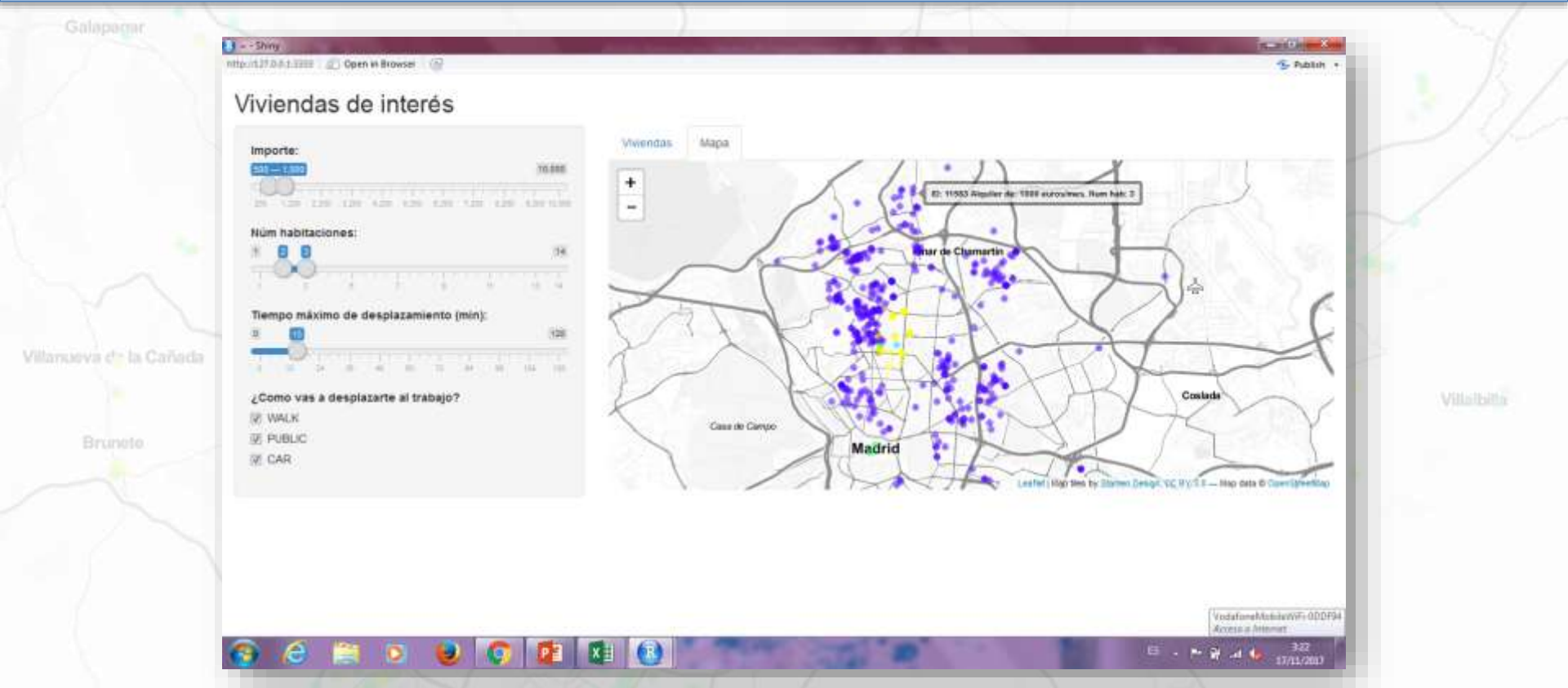
¿Como vas a desplazarte al trabajo?

- WALK
- PUBLIC
- CAR

Showing 1 to 10 of 427 entries

Dirección	Importe *	Num. Hab *	Metros	Distancia (m)	Tiempo (min) *	Medio Trans
4524 calle de Lazaga, 21, Cuatro-Castillejos, Madrid	650	2	46	1.26	14.87	WALK
6566 Carranza, Trafalgar, Madrid	675	2	40	3.46	11.92	CAR
365 calle de María Teresa Sáenz de Heredia, 9, Ventas, Madrid	700	3	60	5.63	14.53	CAR
1271 plaza de redondela, 13, Pinar, Madrid	700	2	70	4.82	12.32	CAR
15873 calle de Ramón Gómez de la Serna, 121, Peñagrande, Madrid	700	2	50	5.57	12.17	CAR
15296 Nuevos Ministerios-Ríos Rosas, Madrid	700	2	45	0.72	9.42	WALK
12379 Cartagena, 110, Prosperidad, Madrid	720	2	37	2.91	8.22	CAR
15609 Ambroz, Madrid	730	2	61	9.34	14.52	CAR
3623 calle de Raimundo Fernández Villaverde, Cuatro Caminos, Madrid	730	2	36	0.7	8.53	WALK
1076 san marcos, Chueca-Justicia, Madrid	750	2	40	4.23	14.97	CAR

IoT con R: Alquilar una casa



Para realizar este proyecto me he apoyado en:

<https://www.r-bloggers.com/placement-an-r-package-to-access-the-google-maps-api/>

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

<https://rstudio.github.io/leaflet/>

<https://gis.stackexchange.com/questions/133625/checking-if-points-fall-within-polygon-shapefile>

Thank You

cajamardatalab@bcc.es

