

```
library(dplyr)

rladies_global %>%
  filter(city == 'Buenos Aires')
```



| Aceptando la | Incertidumbre



Hola!

Soy Daniela Vazquez

Estoy acá porque quiero compartir mi historia con ustedes :)

Blog: dv.uy | Twitter: [@d4tagirl](https://twitter.com/d4tagirl)



Limitación de responsabilidad

Mis primeras charlas en público!

Mi carrera hasta ahora...



Universidad
(Economía)

Probé:
Matemática,
Diseño Web,
Contabilidad,
...

Trabajé en una
inmobiliaria
Gané una beca de
investigación en
Historia Económica

Me recibí

Equifax
(Analytics
Consultant)

Uruguay y Paraguay,
Modelos de Score a
medida (también algo
de marketing analytics,
collections, ...)

Renuncié!





Y entonces?

Me puse a explorar



Algunas de las cosas que exploré

- Traduje un *comic* que presenta R a estudiantes de secundaria
- Tomé una clase de Machine Learning en la Universidad
- Tomé muchos cursos online de R
- Aprendí cómo resolver un cubo Rubik (no me juzguen...)
- Me postulé al programa "NASA Datanauts" de la NASA
- Empecé mi blog (d4tagirl.com)
- Contacté a la organización R-Ladies Global



Mi vida ahora

Aprendiendo a llevarme bien con la incertidumbre



Algunas cosas que estoy haciendo

- Hablando en público (como ahora!)
- NASA Datanauts Clase de 2017
- Continuando con el blog
- Co-organizando R-Ladies Buenos Aires (Argentina)
- Pensando en fundar R-Ladies Montevideo (Uruguay)
- Buscando un trabajo que me encante como Data Scientist!



El mes pasado...

R-Ladies NY y SF



Lanzamiento de NASA Datanauts DC





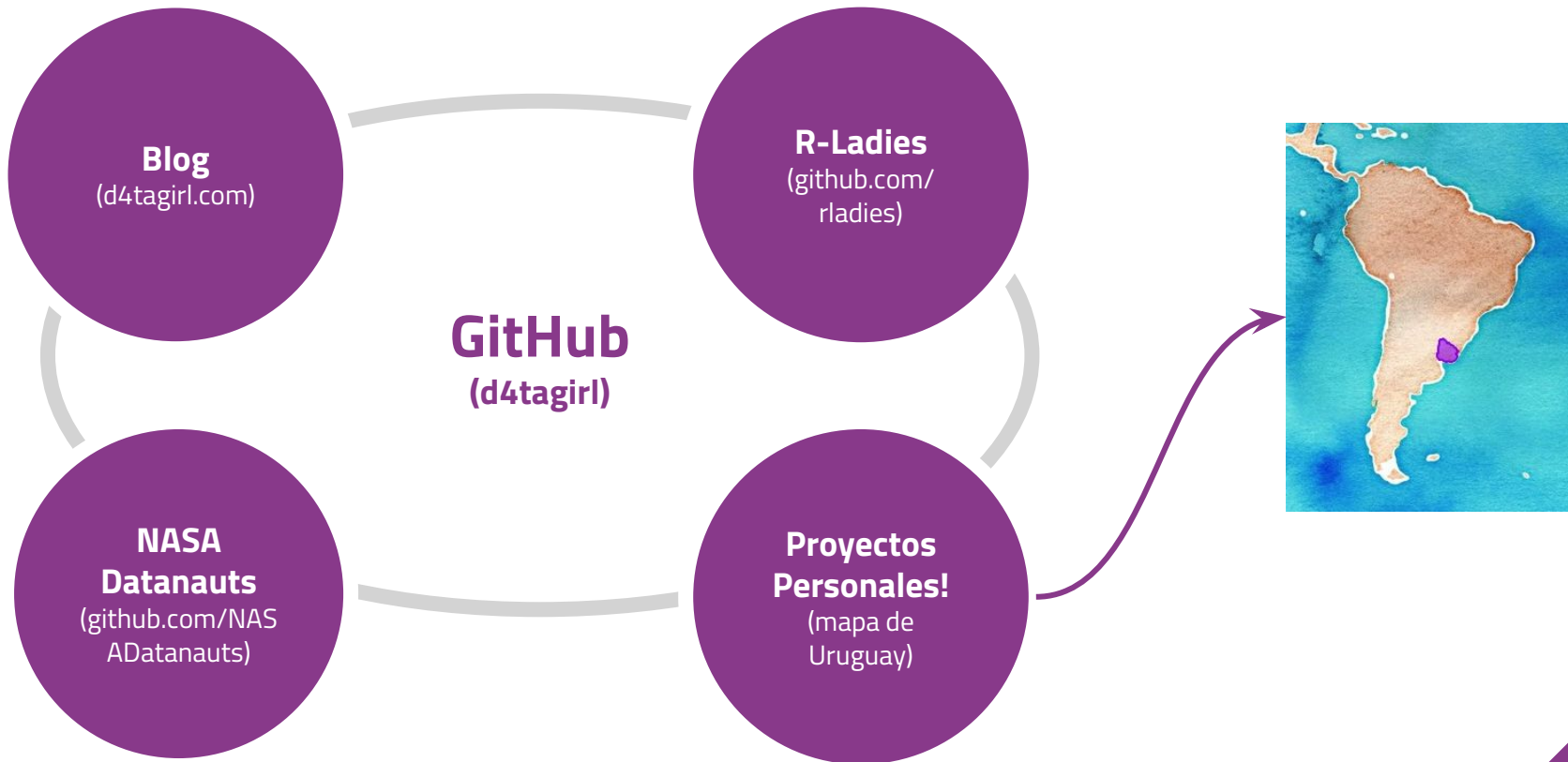
“ *Si trabajo mucho y le hago caso a lo que siento, algo bueno va a suceder!* ”



Cambiando de tema

Algunos aspectos técnicos...

Para qué uso GitHub?



Quién votó por Trump? Lo averiguo usando R!



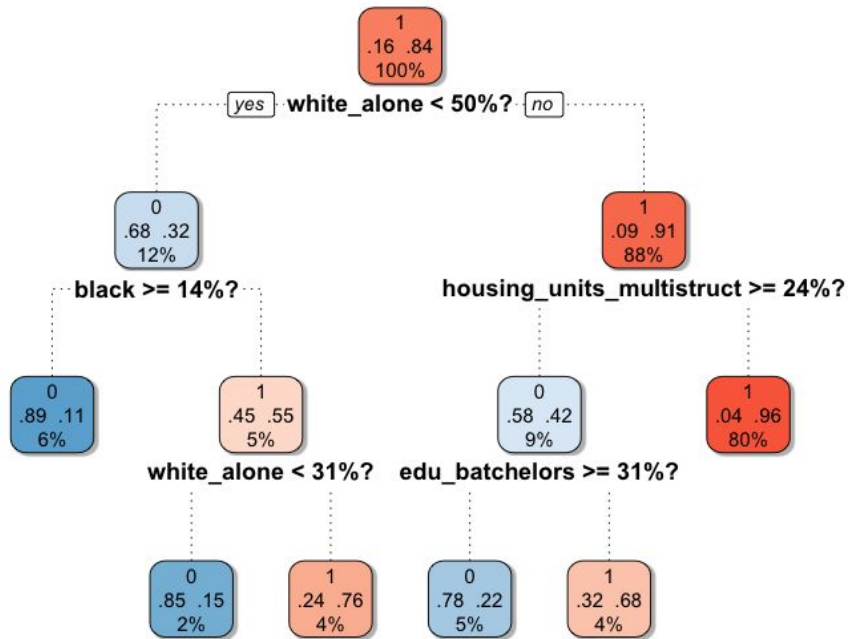
```
set.seed(4444)

trainIndex_2 <- createDataPartition(y = votes$pref_cand_T, p = perc_train,
                                   list = FALSE, times = 1)

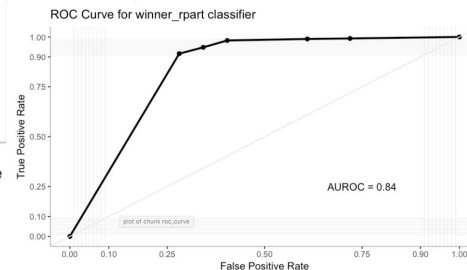
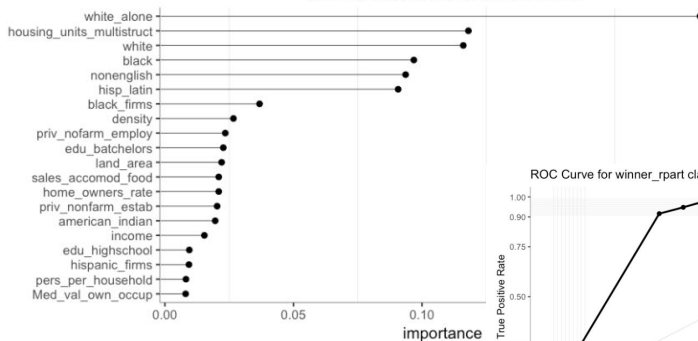
train_2 <- votes %>% subset(ID %in% trainIndex_2)
test_2 <- votes %>% setdiff(train_2)

winner_rpart_2 <- rpart(pref_cand_T ~ .,
                       data = train_2[, -c(1:24, 70)],
                       control = rpart.control(xval = 10, cp = cp))

rpart.plot(winner_rpart_2, main = "Winner candidate in county",
           extra = 104, split.suffix = "%?", branch = 1,
           fallen.leaves = FALSE, box.palette = "BuRd",
           branch.lty = 3, split.cex = 1.2,
           shadow.col = "gray", shadow.offset = 0.2)
```



Importance of county characteristics
in determining the most voted candidate
(20 most relevant, Rpart, scaled to sum 1)





Gracias!!

