

*powered by*

*R programming*

# Introducción general al uso de R - Clase 01 -

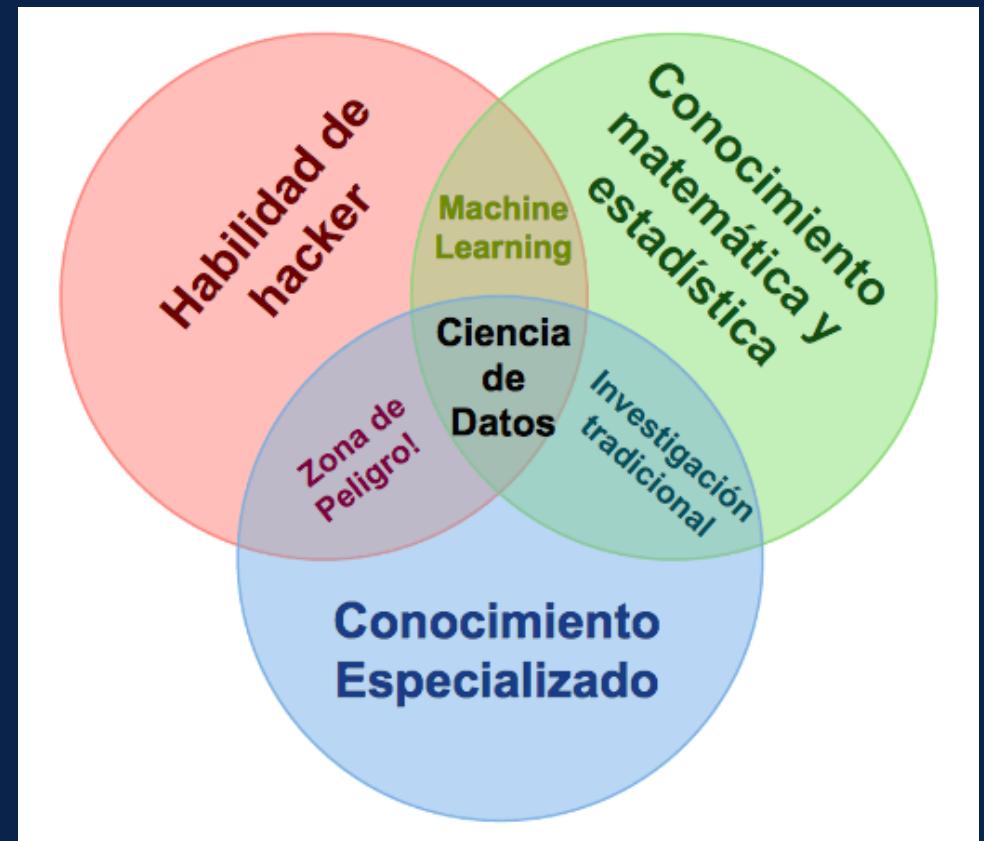
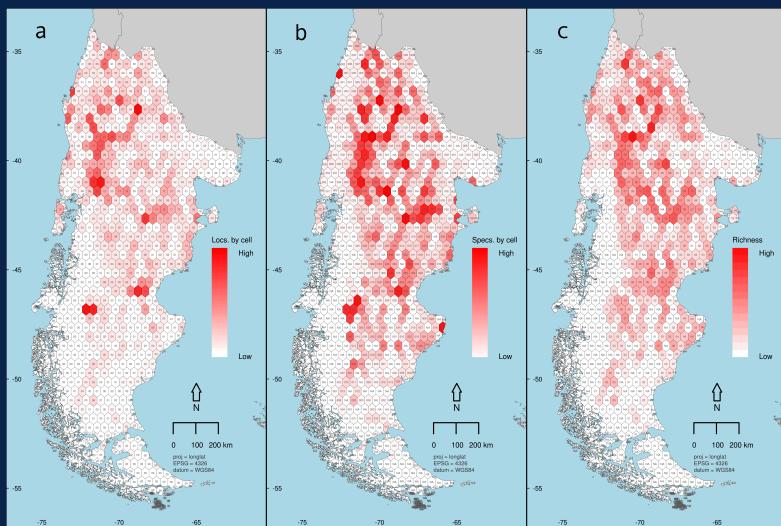
[minolicnp@gmail.com](mailto:minolicnp@gmail.com)

**Ignacio Minoli** *pwp*

Observatorio de Biodiversidad del Bosque Atlántico (OBBA)  
Instituto de Biología Subtropical (IBS) - CONICET - UNaM.



# Presentación



# Formación y Experiencia

- Tecnologías: **Software Libre | Código Abierto | Linux | Programación en R | Programación en Bash-Shell | Programación en Python | Programación en HTML | SublimeText | Rstudio | Shiny Environment | Maxent | Marxan**
- Especialidades: **Biogeografía | Bases de Datos | Programación en R | Cartografía | Mapas | Cambio Climático | Modelado de Nicho Ecológico | GIS | Análisis Espaciales**
- Áreas de desarrollo e investigación: **Manipulación y curación de grandes bases de datos. Alcance internacional, nacional y regional. Biogeografía, Ecoinformática y Sistemas de Información Geográfica.**
- Áreas grales: **Programación, Estadística y Cartografía.**



# Proyectos actuales



# Historia personal con R

## Recomendaciones...

- ¿Me conviene utilizar R u otro lenguaje de programación?
- Cursos de programación en R **versus** cursos de Estadística.
- Uso y/o lectura diaria - feeds.
- Autodidacta. Enfoque: preguntas - hipótesis - problemas puntuales.
- Biblioteca personal de scripts.

```
1 herramienta_principal <- "R is awesome and diverse"
2 #
3 scripts_propios <- c("Lecturas diarias", "Prueba error")
4 #
5 independencia_desarrollo <- (filosofia = "Software.Aberto_Gratis_Multiplataforma")
```

# Posibles usos

- Análisis estadísticos

```
> LinearModel <- lm(dist ~ speed, data=cars)
> print(LinearModel)

Call:
lm(formula = dist ~ speed, data = cars)

Coefficients:
(Intercept)      speed
-17.579         3.932

> |
```

```
> summary(LinearModel)

Call:
lm(formula = dist ~ speed, data = cars)

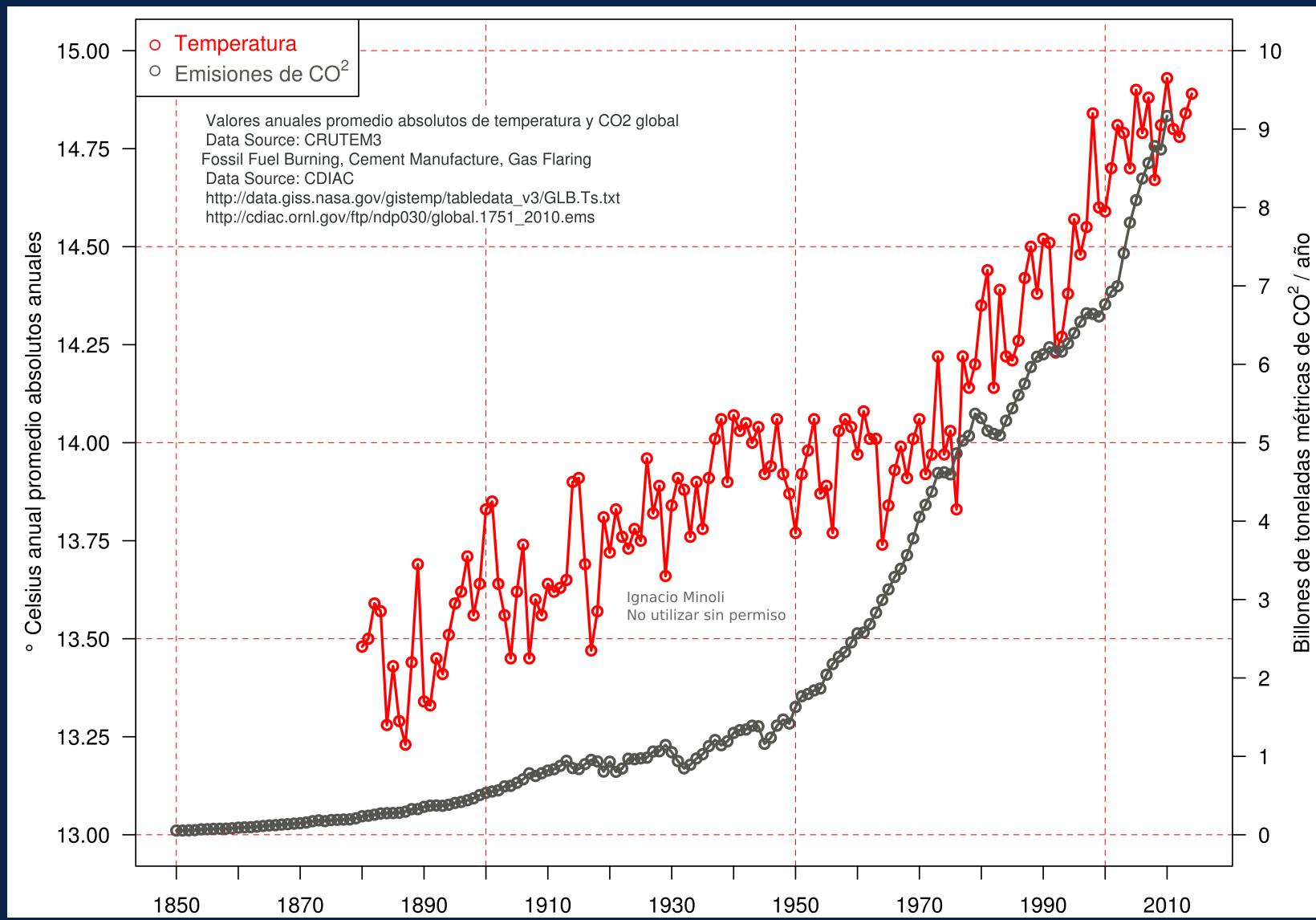
Residuals:
    Min      1Q  Median      3Q     Max 
-29.069 -9.525 -2.272  9.215 43.201 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -17.5791   6.7584 -2.601  0.0123 *  
speed        3.9324   0.4155  9.464 1.49e-12 *** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 15.38 on 48 degrees of freedom
Multiple R-squared:  0.6511,    Adjusted R-squared:  0.6438 
F-statistic: 89.57 on 1 and 48 DF,  p-value: 1.49e-12
> |
```

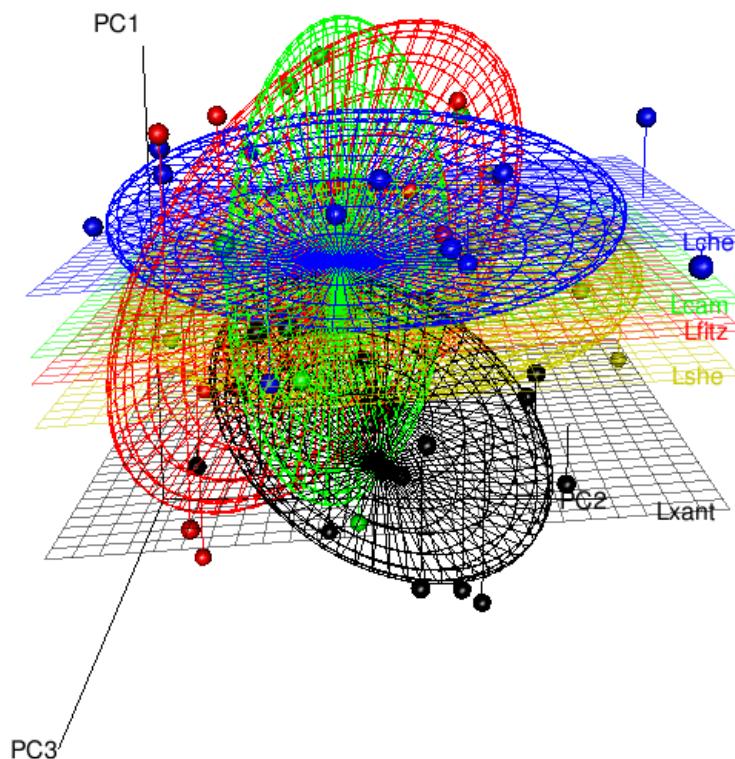
- Descriptiva / Inferencial (Paramétrica No Paramétrica). Uni, Multivariada, Bayesiana, etc.

# Posibles usos

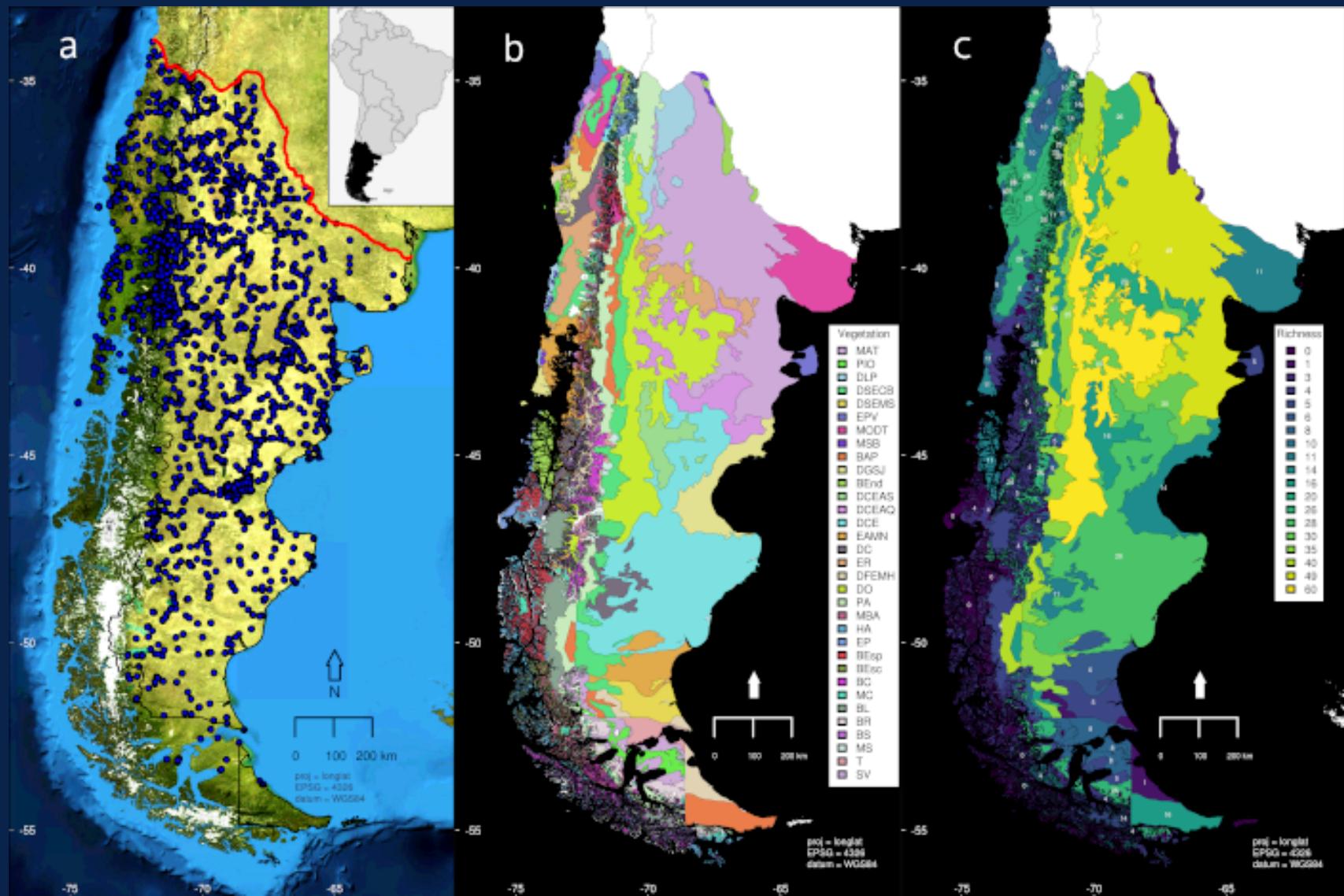


# Posibles usos

- Plots



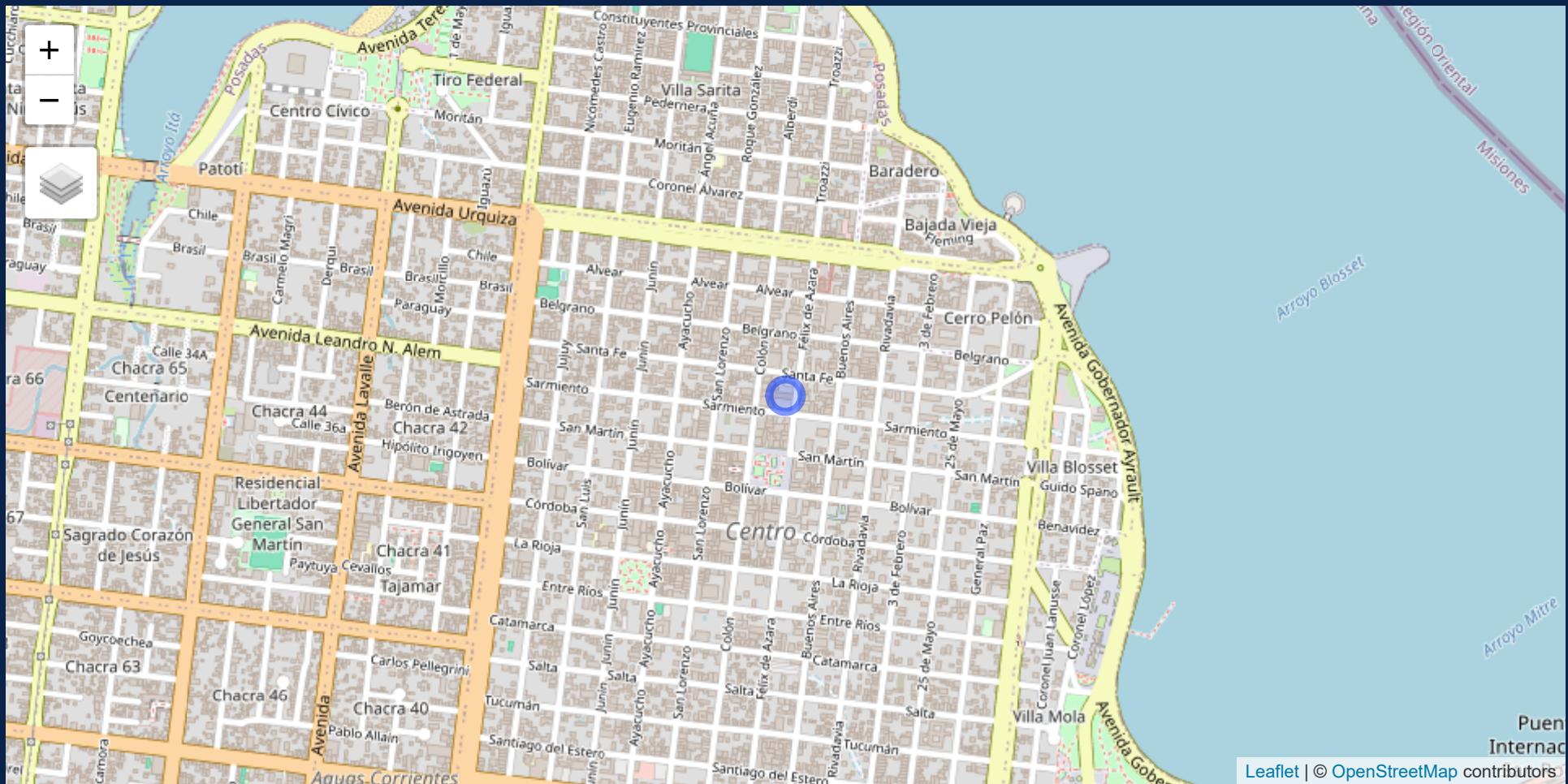
# Posibles usos



- Mapas

# Posibles usos

- Mapas interactivos



# Posibles usos

- Tablas

2019 Women's World Cup Predictions												
TEAM	GROUP	Team Rating		Chance of Finishing Group Stage In ...			Knockout Stage Chances					
		SPI	OFF.	1ST PLACE	2ND PLACE	3RD PLACE	MAKE ROUND OF 16	MAKE QTR-FINALS	MAKE SEMIFINALS	MAKE FINAL	WIN WORLD CUP	
USA 6 pts.	F	98.3	5.5	0.6	83%	17%	-	✓	78%	47%	35%	24%
France 6 pts.	A	96.3	4.3	0.5	>99%	<1%	<1%	✓	78%	42%	30%	19%
Germany 6 pts.	B	93.8	4.0	0.7	98%	2%	-	✓	89%	48%	28%	12%
Canada 6 pts.	E	93.5	3.7	0.6	39%	61%	-	✓	59%	36%	20%	9%
England 6 pts.	D	91.9	3.5	0.6	71%	29%	-	✓	69%	43%	16%	8%
Netherlands 6 pts.	E	92.7	3.9	0.7	61%	39%	-	✓	59%	37%	19%	8%
Australia 3 pts.	C	92.8	4.2	0.9	13%	54%	34%	>99%	54%	26%	10%	5%

# Posibles usos

- Tablas interactivas

Num	Cat	Lon	Lat	Clase	Orden	Genero
1		-54.4	-26.2		Entomobryomorpha	Lepidocyrtus
2		-54.4	-26.2		Sympyleona	Collophora
3		-54.4	-26.2		Sympyleona	Collophora
4		-54.4	-26.1		Sympyleona	Neosminthurus
5		-54.4	-26.1		Entomobryomorpha	Folsomina
6		-54.4	-26.1		Entomobryomorpha	Proisotoma
7		-54.4	-26.1		Poduromorpha	Neotropiella
8		-54.4	-26.1		Entomobryomorpha	Isotoma

[CSV](#)[Excel](#)[Previous](#)

1

2

3

4

5

...

327

[Next](#)

# Posibles usos

- Apps web interactivas para webs de proyectos, disponibilizar resultados de papers, exposiciones en congresos, etc:

Ejemplos -> <https://fauna-atropellada.org.ar/filtro-por-especies/>

- Filtros reactivos - interactivos para una rápida vizualización de datos tabulados, búsqueda de información en big data, presentaciones de powerpoint, etc.

App

# Estructura del Curso

- Carga teórica por las mañanas.
- Prácticas con scripts-data reproducibles por las tardes.
- Puesta en común teórica-práctica al final de las tardes.
- Ejercicio final del curso aplicando el contenido del mismo.
- Los pdf de las teorías y prácticas serán entregados.



# ¿Qué es el lenguaje R?

- R es un entorno y lenguaje de programación con un enfoque -> análisis estadístico + gráficos.
- Fue creado en 1993 por Ross Ihaka y Robert Gentleman. Fines 70' lenguaje S.
- El Comprehensive R Archive Network (CRAN) se creó en 1997. Alojamiento de: archivos ejecutables, código, documentación y librerías o paquetes.
- Actualmente desarrollado por el R Development Core Team.
- Sistema GNU y se distribuye bajo la licencia GNU GPL: Win, Mac, Linux.
- 2008 -> UTF-8. Idiomas; 2010 -> Soporte Win 64 bits
- Líneas de comandos (CLIs).
- Múltiples graphical user interface (GUIs).
- Múltiples Integrated development environment (IDEs).



# Características principales

- Open-source. Soporte multiplataforma. Compatible con otros lenguajes de programación.
- Programación orientada a objetos. Permite que los usuarios lo extiendan definiendo sus propias funciones.
- Enormes capacidades gráficas. Manejo y Almacenaje de datos.
- Comunidad de ayuda - soporte muy grande y diversa. Usuarios estadísticos: nuevas ideas y metodologías.
- Amplia oferta de paquetes - funciones. Realizar cálculos estadísticos complejos.

# Características principales

- Paralelizar o trabajar en clusters. Operaciones aritméticas vectoriales.
- Interactúa con bases de datos: ROracle, Open Database Connectivity Protocol, RMySQL, etc.
- Amplia variedad de datos. Genera reportes en cualquier formato.
- Data Wrangling, Web Scraping, Machine Learning, Información de APIs, servers, etc.

# Características principales

# Características principales

# Características principales



# Características principales



# Características principales



# Características principales

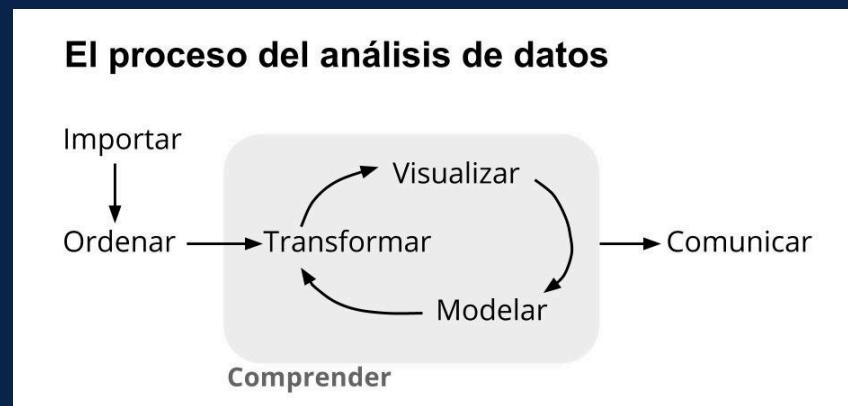
# Importancia de la programación

- Nada de lo expuesto anteriormente sería posible sin la programación...



# Alcances de la programación

- Si, por supuesto que los costos-beneficios valen la pena y la curva de aprendizaje es diferente en lenguajes - personas.
- Poder de creación gratuito, innovador, colaborativo, escalable, reproducible, transparente, continuo e ilimitado...



# QR - presentación

**PDF Presentación: Clase 01 - Teoría**



# Fin de la presentación

**¡¡ Muchas gracias por su atención!!**



¡¡¡ Muchas gracias !!!

¿ Preguntas ?