

Presentación del curso

Introducción al manejo y análisis de datos biológicos usando R

Dr. Diego Carmona Moreno Bello

(Cátedra Conacyt)
Depto. Ecología Tropical,
UADY

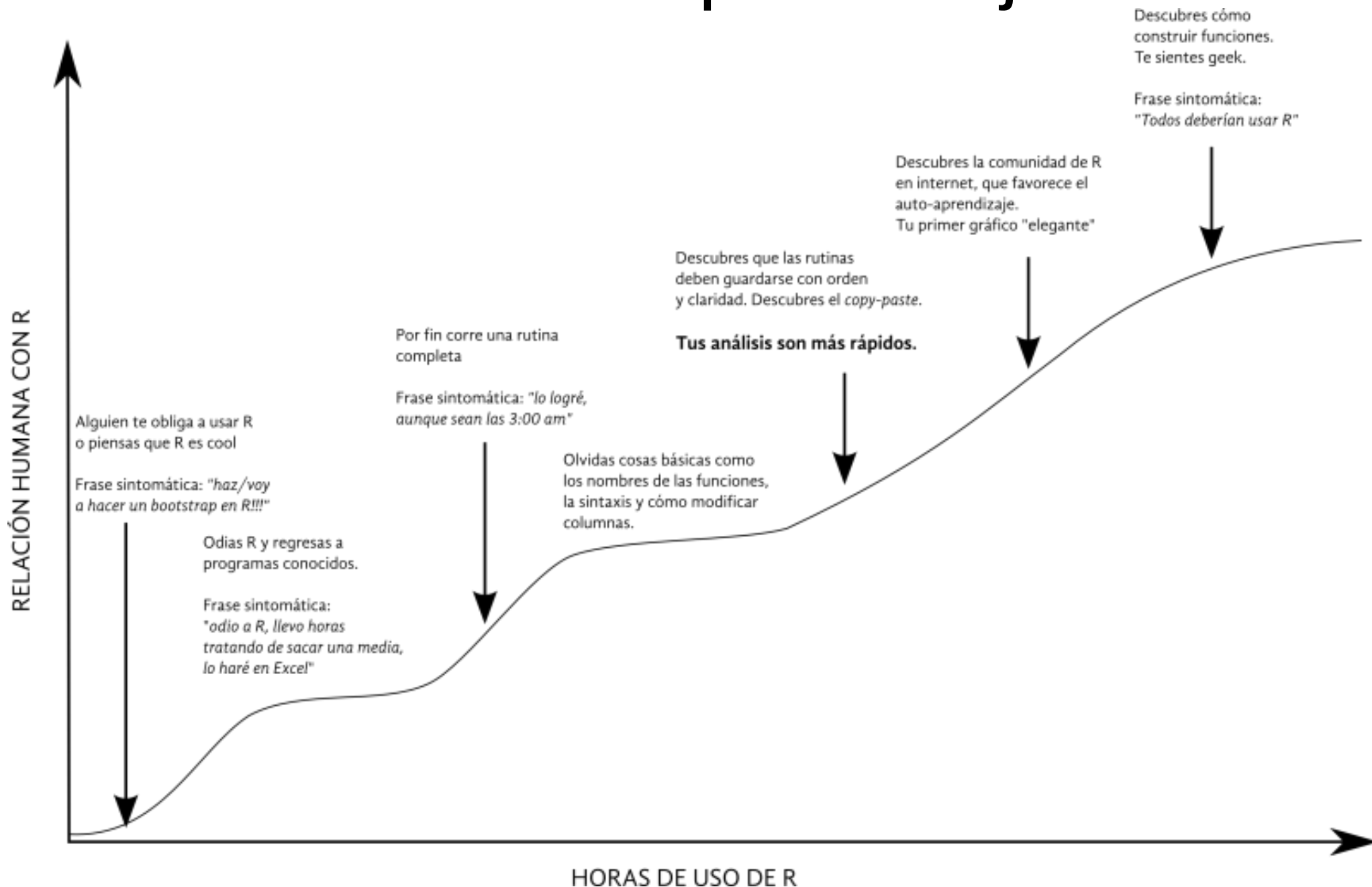
Tesis del curso

- R incrementan nuestra capacidad de aprender, investigar, y crear, favoreciendo nuestra competitividad laboral.
- R se suele enseñar asociado a estadística, cuando es más que un programa de estadística.
- Al inicio, existe un trade-off entre aprender R como lenguaje y estadística

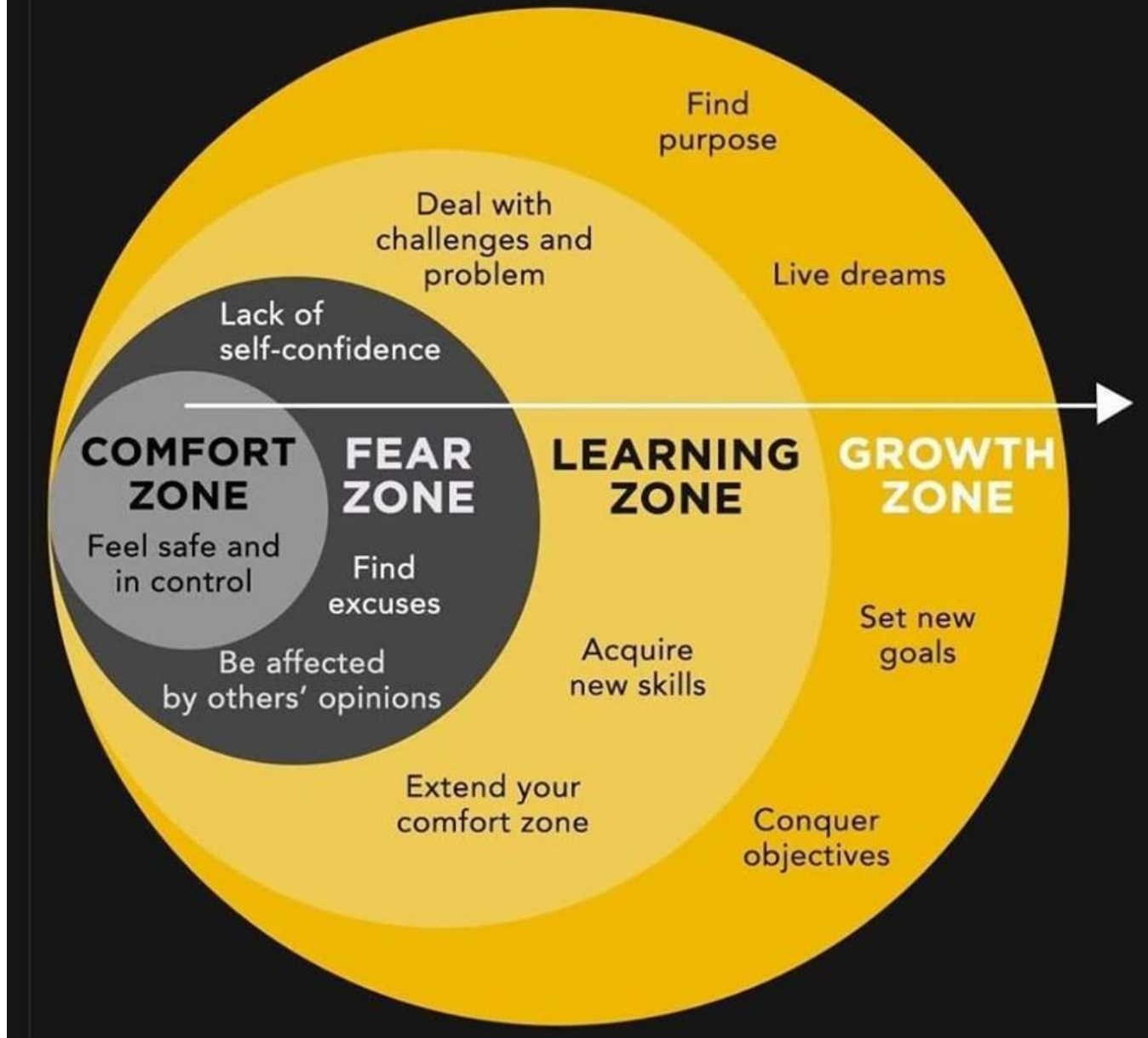
Tesis del curso

- R requiere paciencia y solemos no darnos tiempo ni espacio para aprender este lenguaje.
- La presión académica por analizar datos rápidamente favorece el regresar a herramientas y procedimientos poco eficientes, limitados en reproducibilidad, restrictivos en libertades de usuarios y en capacidad de análisis (ej. excel)
- Resultado: académicos (estudiantes y profesores) que desisten en usar R se limitan y restringen la conformación de comunidades de usuarios que favorecería un incremento en la capacidad de análisis de los individuos y las instituciones.

Curva de aprendizaje



“The Comfort Zone”



El comfort del idiota (Peter Capusotto)

<https://www.youtube.com/watch?v=PdESMwCkn-g>

```

342 } , mc.cores = 4)
343 toc()
344
345 ###----- Model 4 assuming non-independence within studies -----
346 #           but considering the phylogenetic variance component
347 tic()
348 pm1_ho_study<- mclapply(1:3, function(i) {
349   MCMCglmm(Ho ~ 1, random =~ Study+animal,
350     ginverse=list(animal=dat0.all$IA), family = "gaussian",
351     data = dat0.all$database,
352     mev = dat0.all$database$mev, verbose = T, nitt = my_nitt,
353     thin = my_thin, burnin = my_burnin,
354     prior = prior2_strong) } , mc.cores = 4)
355
356
357 ###----- Model 4 assuming non-independence within species but also-----
358 #           considering phylogenetic variance component
359 tic()
360 pm2_ho_specie<- mclapply(1:3, function(i) { MCMCglmm(Ho ~ 1, random =~ Species_ID+animal,
361   ginverse=list(animal=dat0.all$IA), family = "gaussian",
362   data = dat0.all$database,
363   mev = dat0.all$database$mev, verbose = T, nitt = my_nitt,
364   thin = my_thin, burnin = my_burnin,
365   prior = prior2_strong) } , mc.cores = 4)
366
367 toc()
368 ###--- Model 5 including the two random effects and the phylogenetic component ---
369
370 tic()
371 pm3_ho_study_specie<- mclapply(1:3, function(i) { MCMCglmm(Ho ~ 1, random =~ Study+Species_ID+ani
372   ginverse=list(animal=dat0.all$IA), family = "gaussian",
373   data = dat0.all$database,
374   mev = dat0.all$database$mev, verbose = T, nitt = my_nitt,
375

```




```

342 } , mc.cores = 4)
343 toc()
344
345 ###----- Model 4 assuming non-independence within studies -----
346 #           but considering the phylogenetic variance component
347 tic()
348 pm1_ho_study<- mclapply(1:3, function(i) {
349   MCMCglmm(Ho ~ 1, random =~ Study+animal,
350             ginverse=list(animal=dato.all$IA), family = "gaussian",
351             data = dato.all$dabase,
352             mev = dato.all$dabase$mev, verbose = T, pitt = my.pitt
353             thin = my_thin, burnin = 10000, prior = prior2_st)
354 }
355
356
357 ###----- Model 4 assuming non-independence within studies -----
358 #           but considering phylogenetic variance component
359 tic()
360 pm2_ho_specie<- mclapply(1:3, function(i) {
361   ginverse=list(animal=dato.all$IA), family = "gaussian",
362   data = dato.all$dabase,
363   mev = dato.all$dabase$mev, verbose = T, pitt = my.pitt
364   thin = my_thin, burnin = 10000, prior = prior2_st)
365 }
366
367 toc()
368 ###--- Model 5 including the two random effects ---
369 tic()
370 pm3_ho_study_specie<- mclapply(1:3, function(i) {
371   ginverse=list(animal=dato.all$IA), family = "gaussian",
372   data = dato.all$dabase,
373   mev = dato.all$dabase$mev, verbose = T, pitt = my.pitt
374 }
375

```



NERDS VS GEEKS

Objetivo general

1. Establecer las bases fundamentales para lograr un manejo óptimo de R.
2. Sentar las bases necesarias que facilite el aprendizaje de cursos avanzados
3. Reforzar habilidades que permita al alumno ser autodidacta en el uso de este programa.

Índice general

1. Introducción general.
2. Manejo de bases de datos
3. Listas, matrices y arrays
4. Sobre funciones y programación
5. Gráficos
6. Introducción al análisis de datos básico

Evaluación

Estrategia de evaluación	Criterio de evaluación	Ponderación
• Retos	Claridad, estructura, reproducibilidad, consistencia en estilo de escribir código	55%
• Ensayo 1	Claridad, buena argumentación, uso de citas	15%
• Examen a casa (1)		30%



¿Por qué tener un parto natural y no una cesárea?

¿Por qué quieres aprender R?

