



*¿Por qué usar R y no otro?
Perspectiva desde la opinión
de un usuario asiduo (versión 2024)*

Dr. Luis Alan Navarro Navarro

**Investigador por México-El Colegio de Sonora
Centro de Estudios en Gobierno y Asuntos
Públicos (CEGAP)**

alanphd.com

¿Qué es "R"?

- *R es un sistema para análisis estadísticos y gráficos creado por Ross Ihaka y Robert Gentle-man 1996 (de ahí en nombre de "R").*
- *R se distribuye gratuitamente bajo los términos de la GNU General Public License; su desarrollo y distribución son llevados a cabo por varios estadísticos conocidos como el Grupo Nuclear de Desarrollo de R (Paradis 2002).*

Ventajas de R

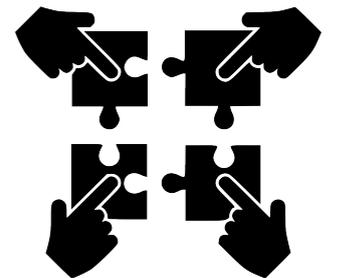
- *Es gratuito y está disponible en Internet para todos. Así como su entorno de desarrollo integrado (IDE) Rstudio.*

- *Corre prácticamente en todos los sistemas operativos.* 

Referencia [1]

- *Existen muchos tutoriales, foros de discusión y ayuda que proveen siempre fórmulas y código que resuelve una gran cantidad de problemas y retos. Todos creados libremente por una extensa comunidad de usuarios y desarrolladores.*

- *Favorece el trabajo colaborativo. Es fácil compartir el código pues es solo texto.*



Referencia [2]

- *Repetibilidad: Reproducir rutinas, volver a correrlas con datos nuevos. A la vez, otros pueden replicar tu análisis y revisarlo.*



Referencia [3]

Ventajas de R (continuación)

- *Existe una amplia y diversa red de colaboradores que contribuyen a mejorar el programa que está en constante evolución.*
- *Estimula el pensamiento crítico y sistémico en la solución de problemas.*
- *La Programación Orientada a Objetos es una herramienta que brinda muchas ventajas, comparada con los programas de "correr la rutina e imprimir en papel (pantalla)".*
- *Posee un amplio espectro de formas de presentar los resultados. PDF, HTML, SVG, etc. Así como múltiples extensiones de archivo.*

Ventajas de R (continuación)

- *Hoy en día puedes echar mano de la Inteligencia Artificial (Ai) para generar código.*
- *Por ejemplo: https://youtu.be/P-72GeozxDA?si=k_nx0He3i-BNJF9I*
- *Es importante tener un buen conocimiento de R al auxiliarse de la Ai. Te ahorra tiempo, pero al final es el humano el que juzga si el análisis esta correcto.*



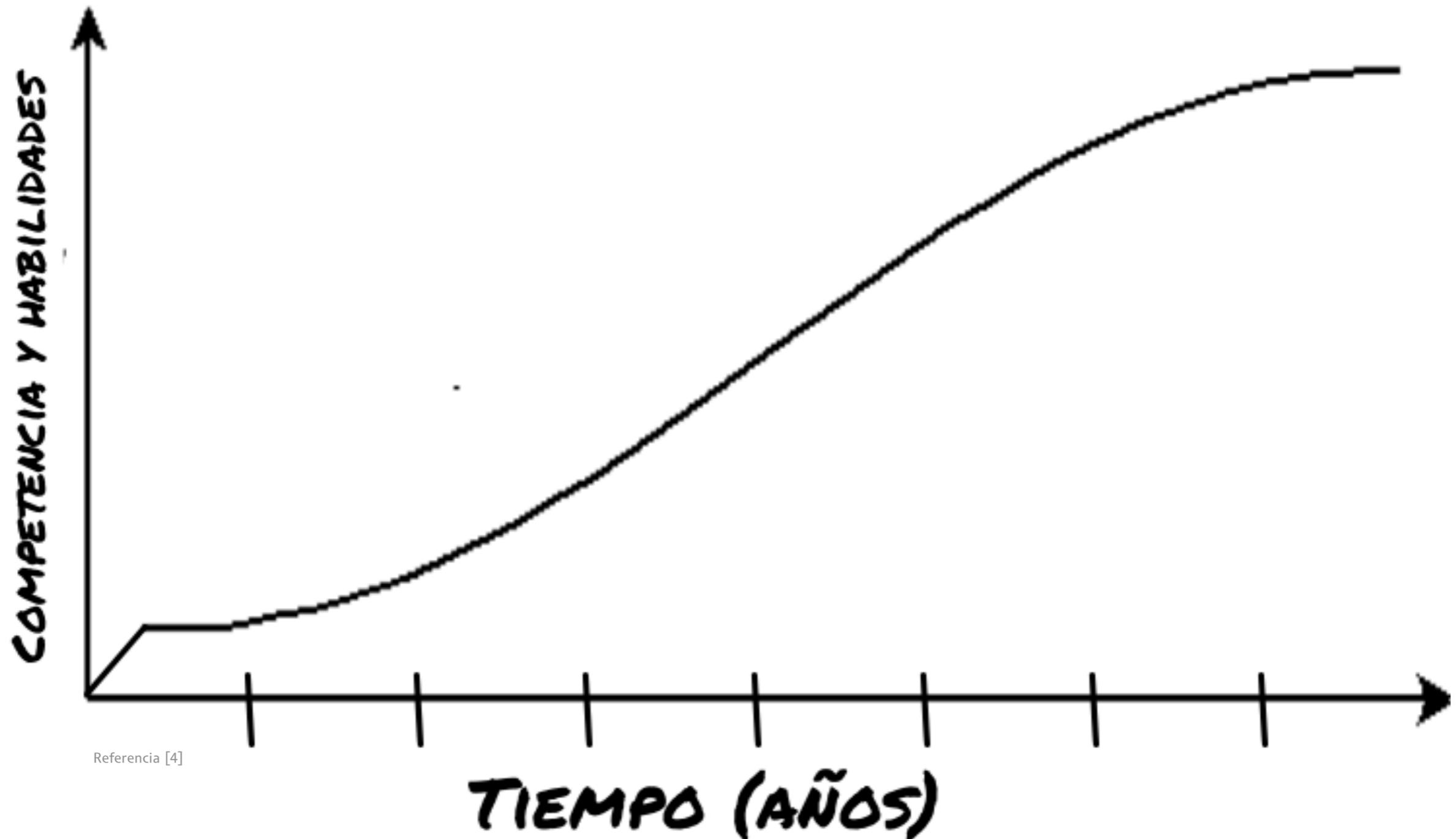
Usando ChatGPT para crear código de R

22 views • 1 month ago

Desventajas

- *Se requiere de un buen nivel de inglés técnico, ya que la gran mayoría de foros, tutoriales y publicaciones están en este idioma.*
- *La curva de aprendizaje es lenta. Por otra parte, el programa está en constante evolución. Algunos paquetes dejan de actualizarse pero surgen otros con mayor capacidad y rapidez analítica. Hay que estar actualizándose constantemente.*
- *El usuario debe de entender muy bien la herramienta estadística y metodológica que pretende aplicar en R.*
- *El escribir "código" asusta e intimida a muchos principiantes sin experiencia en programación.*

R tiene una curva de aprendizaje "lenta"



Desventajas (continuación)

- Hay que ser obstinado e insistente en la búsqueda de una solución a un modelo o algoritmo que se está codificando en R; a base de "prueba y error" y de buscar en todos los recursos disponibles en Internet, pensando siempre que "alguien" es probable que ya haya tenido el mismo problema antes que tú.
- En ocasiones en poco lento. Especialmente en rutinas complejas y con base de datos grandes. Es común dejar el análisis corriendo por horas.
- El usuario a veces tarda en hacer la abstracción necesaria para entender la Programación Orientada a Objetos. Por ejemplo, una gráfica de caja que es a la vez "imagen" pero puede declararse como "objeto", lo mismo una salida de un análisis de regresión, etc.

Desventajas (continuación)

- *"Otro de los inconvenientes es que los códigos son extremadamente sensibles a los errores tipográficos, por lo que una simple coma, fuera de lugar interrumpe el proceso de análisis generando mensajes de error" (Celis de la Rosa y Labrada, 2014).*

- *Como muchos libros de "campismo" lo sugieren muchos se pierden por qué dejan de creer/confiar en su brújula, como lo sugiere este segmento de Kreps (2016):*

which case we should believe it without question, for if we go contrary to the teachings of the instrument we will find that 99 times out of 100 the compass is right and we are dead wrong. One of the greatest mistakes a man can make when he gets turned around in the big timber is to doubt his compass, but many people will take a chance on their very unreliable instinct rather than to trust a perfectly trustworthy instrument which was brought into the woods to serve them on just such occasions. But one need



- *Lo mismo pasa con R. Los estudiantes "dudan" del programa y dicen "no esta funcionando", cuando lo que en realidad tienen es un "error de dedo".*

Los paquetes



Los paquetes ("packages")

- *Son paquetes, librerías, extensiones, programas, etc., que contienen funciones y bases de datos especializadas para un tipo de análisis.*
- ~~*Cuando abres R se carga una funcionalidad básica (alrededor de 1,300 funciones).*~~
- *Una de las cosas que hacen a R excepcional es la posibilidad de extender su funcionalidad instalando y cargando nuevos paquetes o programas.*
- *Estos paquetes están disponibles para los usuarios de R en varios repositorios.*

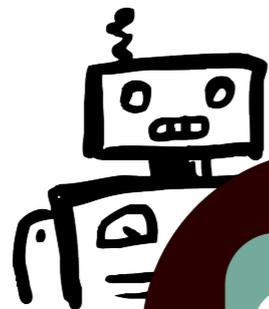
Los paquetes ("packages")



How many libraries are installed by default in R when you open it?



R, by default, comes with 30 packages pre-installed. These are the base and recommended packages that ship with R. Here is a list of these default packages:

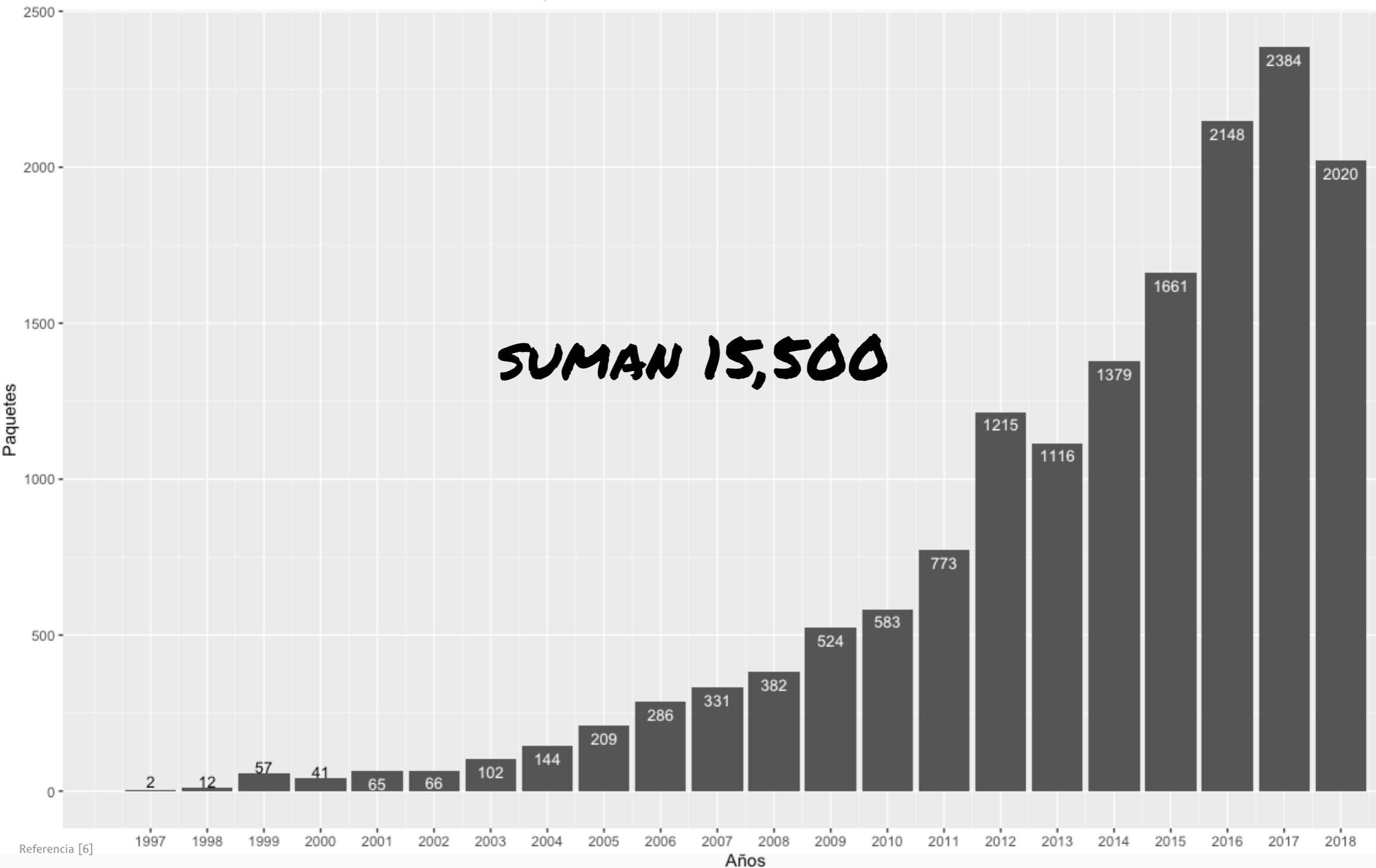


R has a substantial number of functions installed by default. A typical installation of R includes around 1500 functions across its base and recommended packages. These functions cover a wide range of tasks, from basic arithmetic and data manipulation to more complex statistical modeling and graphical functions.

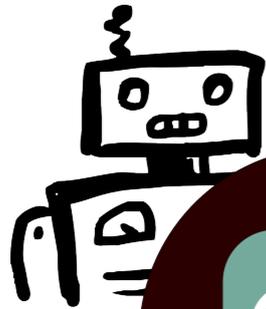
Los paquetes ("packages") (continuación)

- *No hay restricciones para que cualquier persona agregue un paquete nuevo. Un paquete es una forma en la que puedes compilar tu trabajo, análisis (funciones) y bases de datos.*
- *Instalar un paquete no hace que éste esté listo para ser usado, hay que cargar el paquete a la sesión de trabajo.*
- *Los paquetes no ocurren en el vacío ni son independientes uno de los otros. Es decir, éstos usan funciones de otros paquetes.*
- *A veces hay problemas con la versión de R; o con la versión de los paquetes de los que un paquete depende. El uso de, por ejemplo RStudio Cloud te permite cambiar "on the fly" de versión de R.*

Evolución de los paquetes (fecha liberación)



Los paquetes ("packages")



What is the current number of packages total that R has?



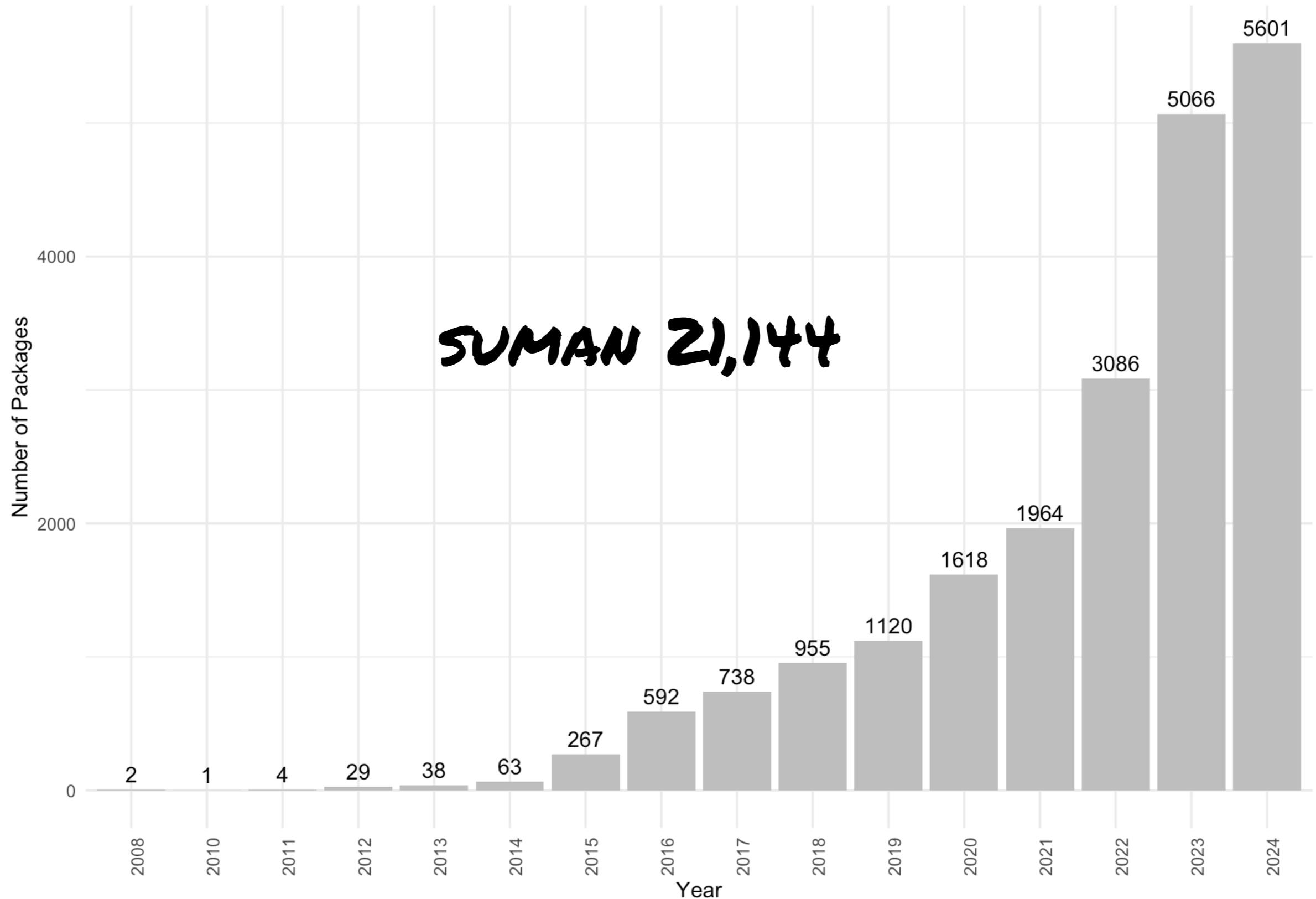
Talked to gpts.webpilot.ai

As of the latest data on August 2, 2024, there are a total of 20,105 packages available on CRAN .
This number is continuously increasing as new packages are added regularly.

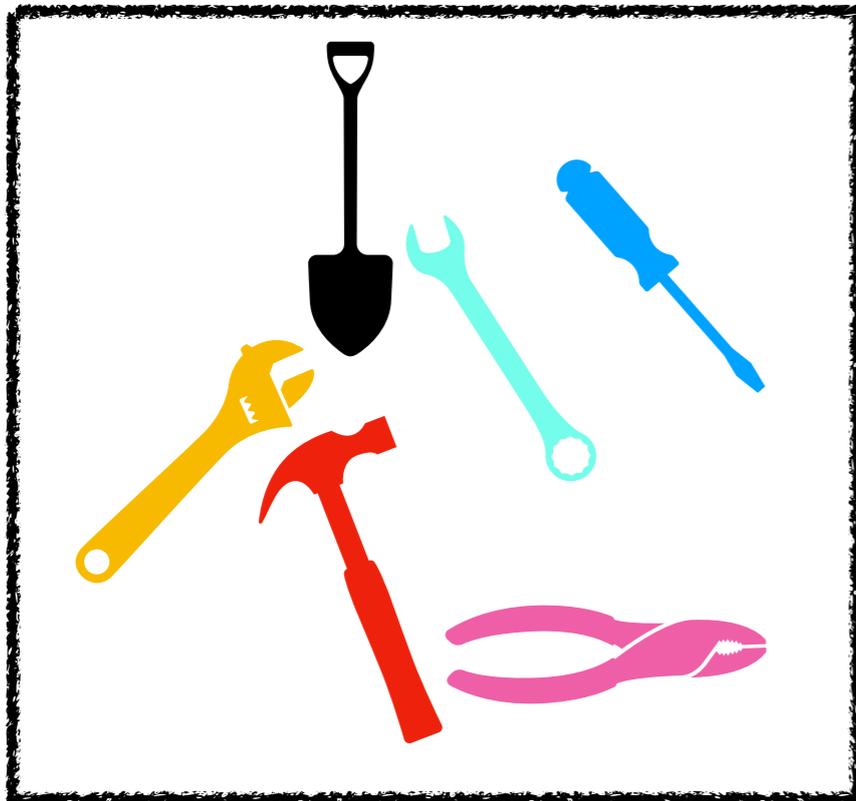


Evolución de los paquetes (fecha liberación)

Number of R Packages Released by Year (1999-2024)



¿Cómo saber qué paquete usar?



UN PAQUETE CONSISTE DE
HERRAMIENTAS.

¿QUÉ ANÁLISIS QUIERES HACER?

¿CUÁL ES TU BASE DE DATOS?

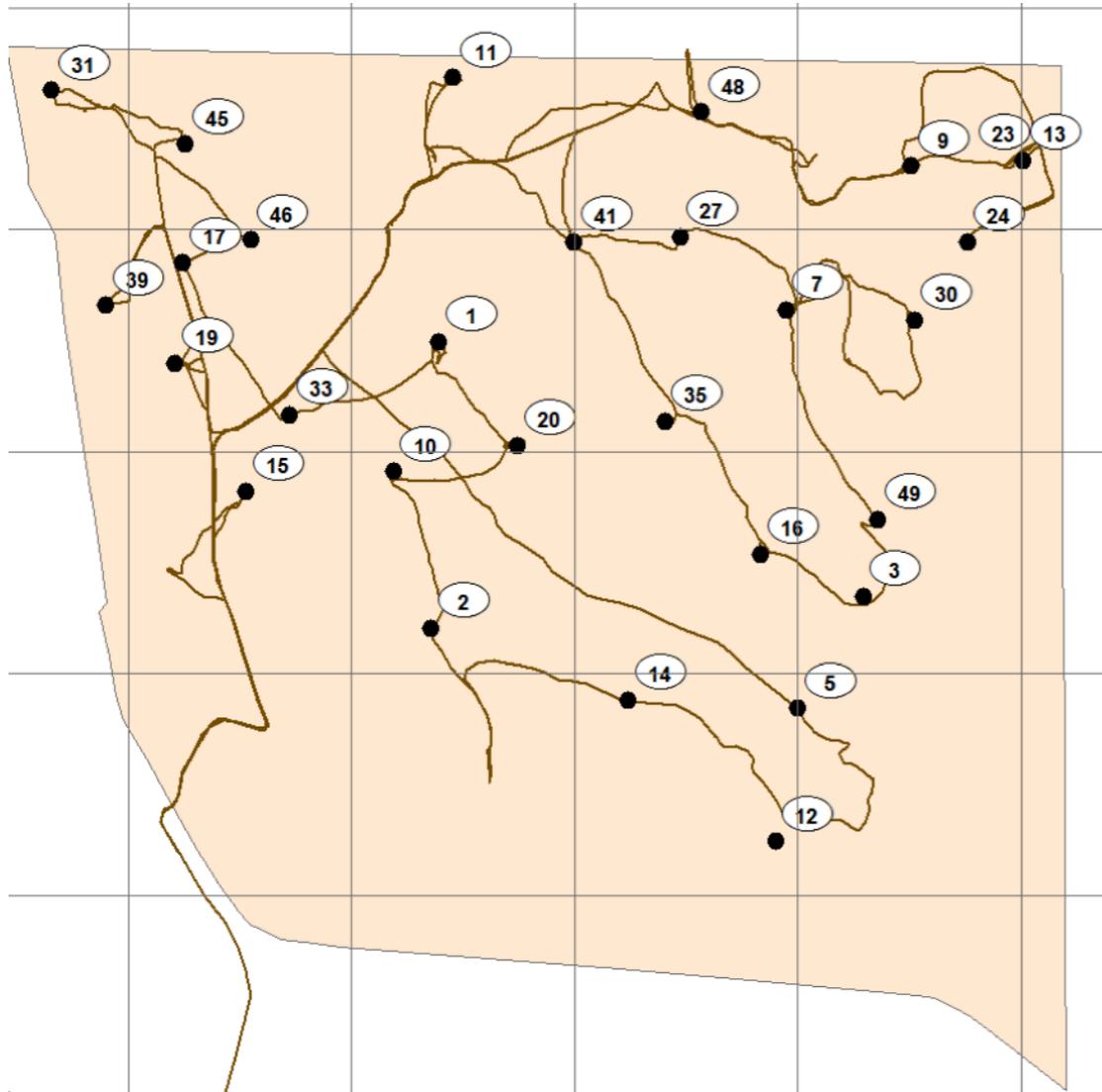
**PIENSA, ES MUY SEGURO QUE
ALGUIEN YA TUVO ESE PROBLEMA**

ANTES QUE TU, CON SUERTE LO

RESOLVIÓ Y COMPARTIÓ SU

EXPERIENCIA EN INTERNET.

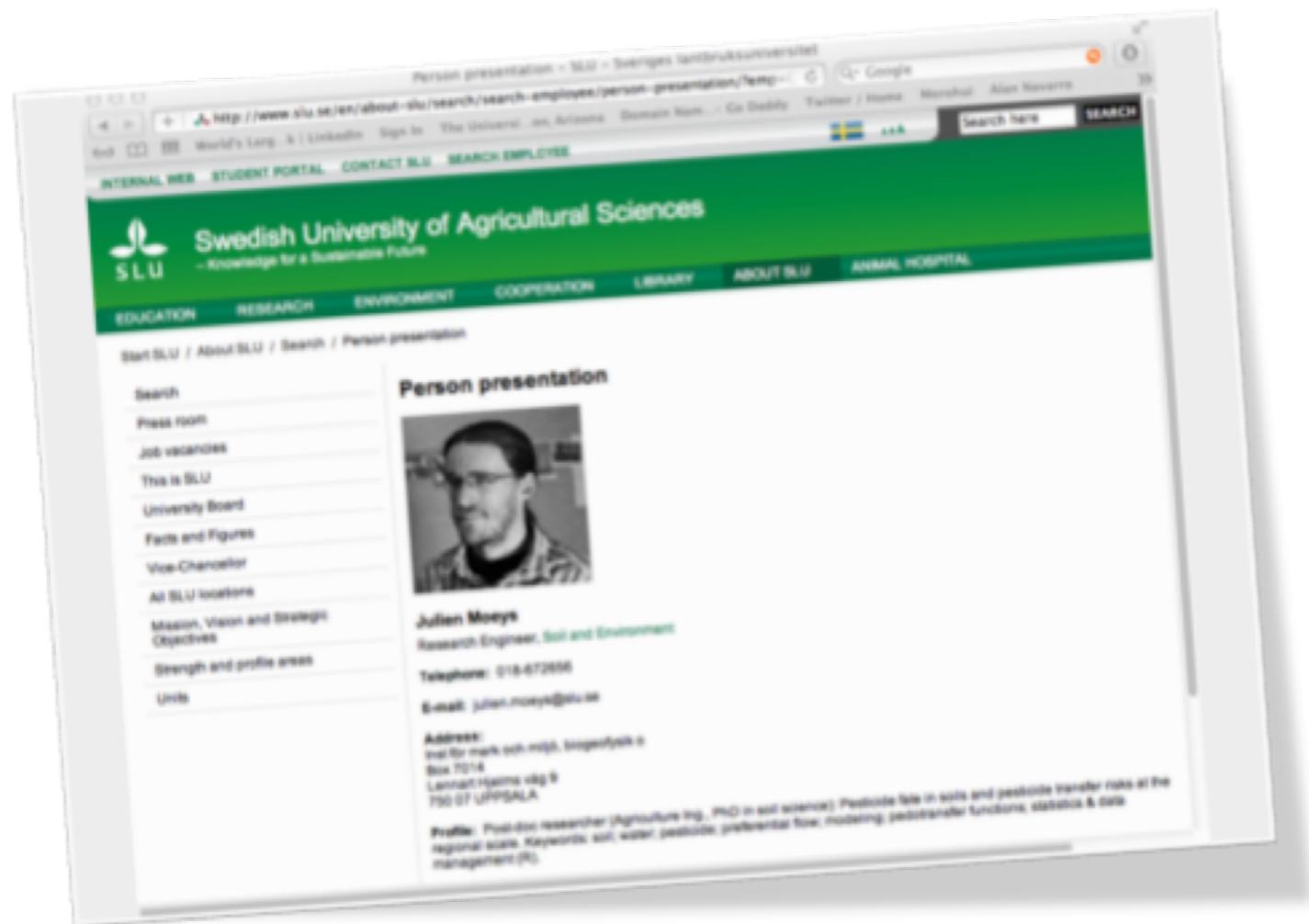
Ejemplo 1. Muestras de textura del suelo



- El polígono representa un predio.
- Las líneas son el "track" del GPS.
- Los números en círculo, puntos de muestreo.
- Se tomaron muestras de suelo.
- Se determinó, en el laboratorio del Departamento de Agricultura y Ganadería de la Universidad de Sonora, las tres clases texturales: arena, limo y arcilla.
- El equipo de trabajo buscamos una forma práctica de presentar la información de texturas.

Ejemplo 1. Muestras de textura del suelo

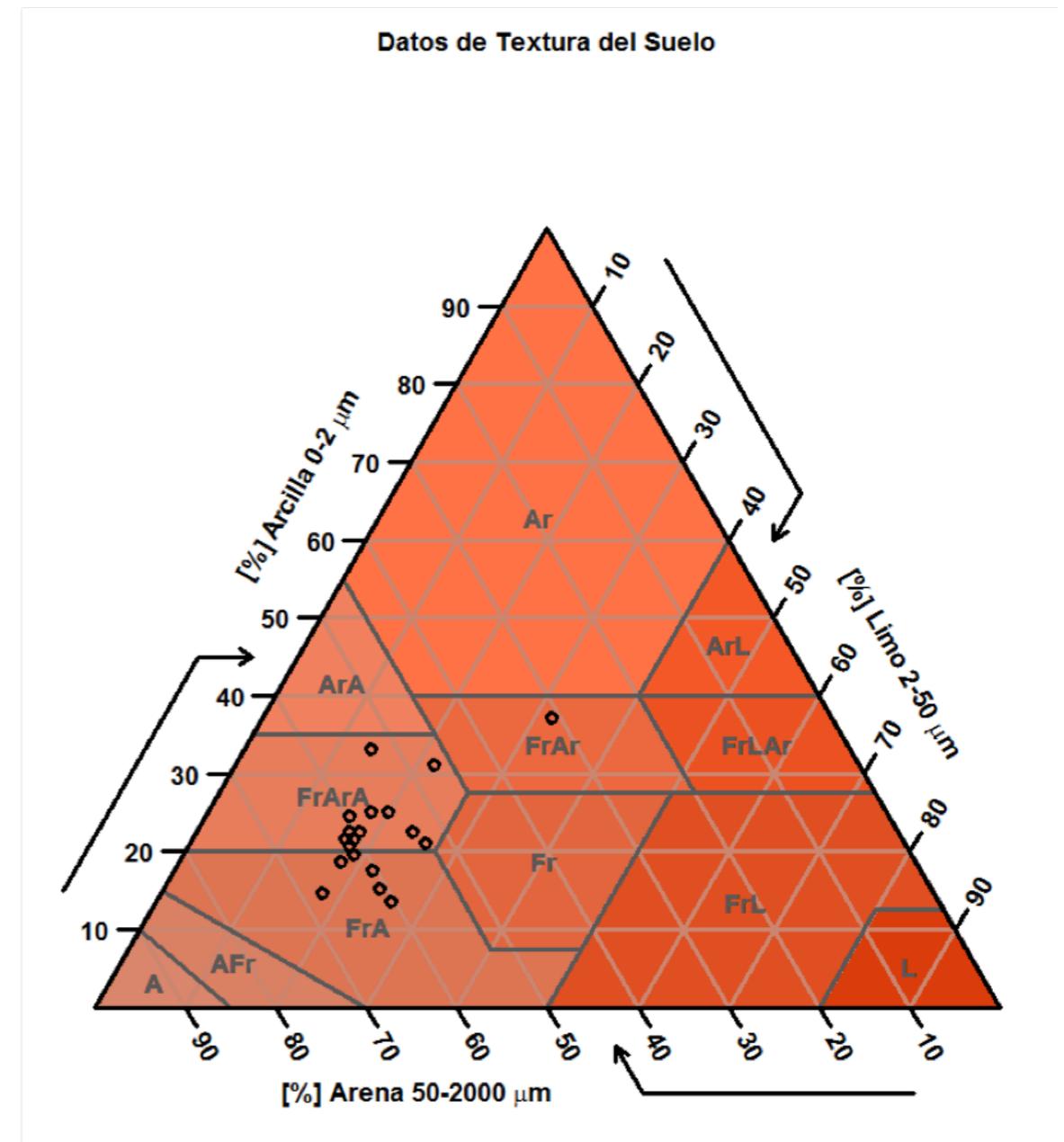
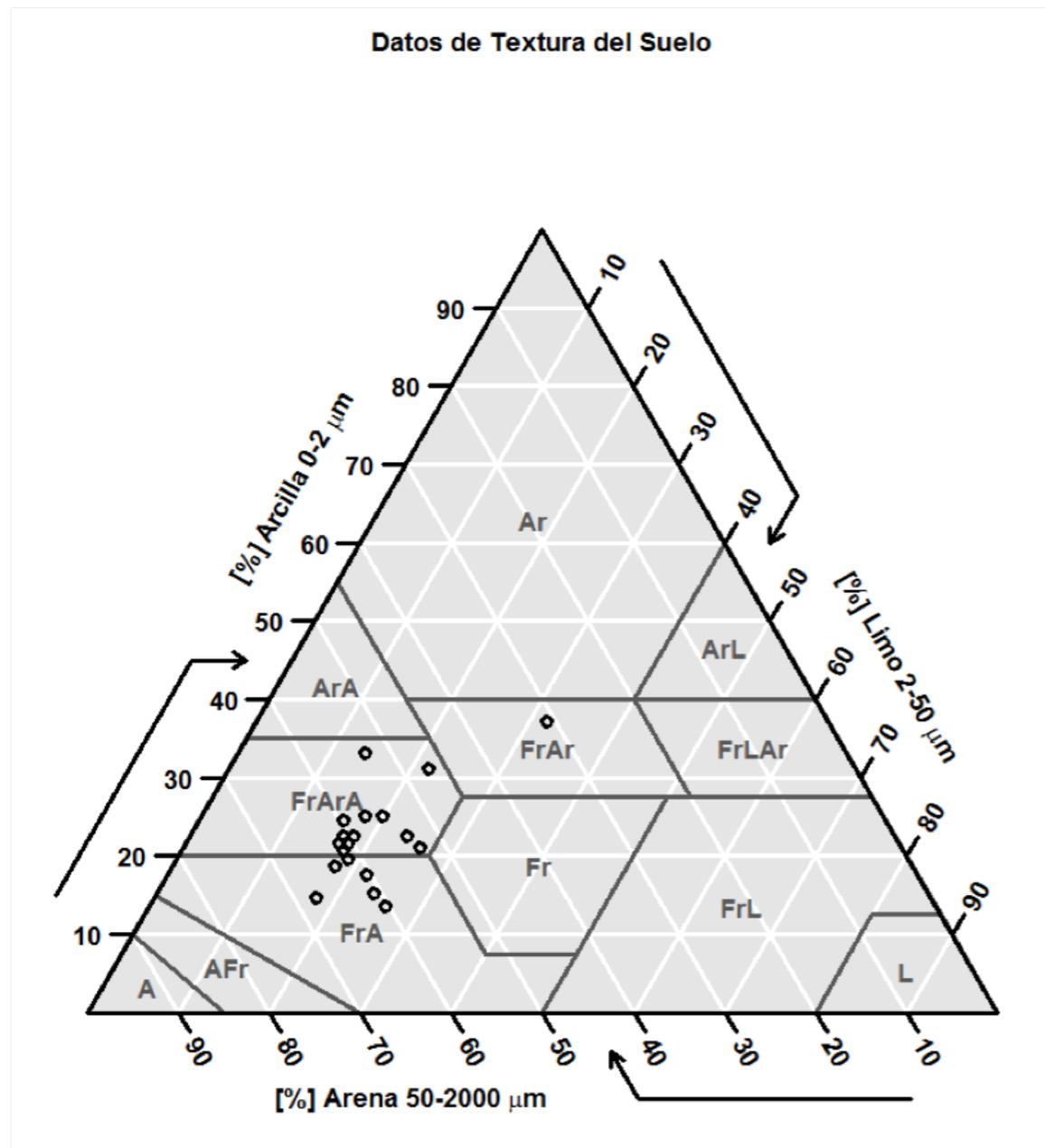
(Continuación)



- Bueno, alguien antes que nosotros, ya tuvo ese problema y lo resolvió usando R.
- El programa se llama "soiltexture"
- El desarrollador es el Dr. Julien Moeys <https://www.slu.se/en/cv/julien-moeys/>
- Y gracias a su altruismo lo puso a disposición de todos nosotros.

Ejemplo 1. Muestras de textura del suelo

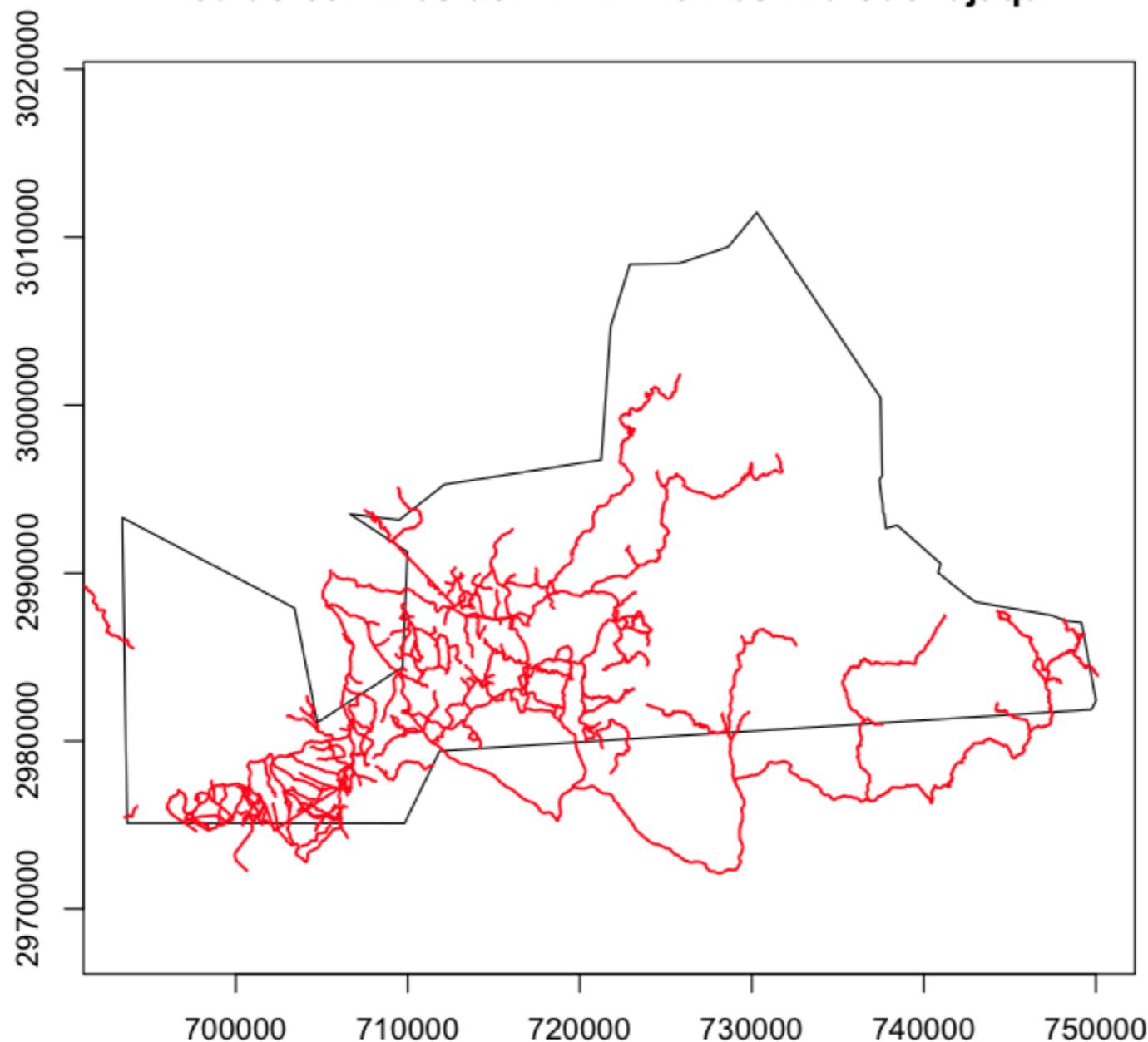
(Continuación)



- Los puntos muestran gráficamente los grupos de textura de suelo predominantes. "FrArA" que se lee "franco arenoso arcilloso".

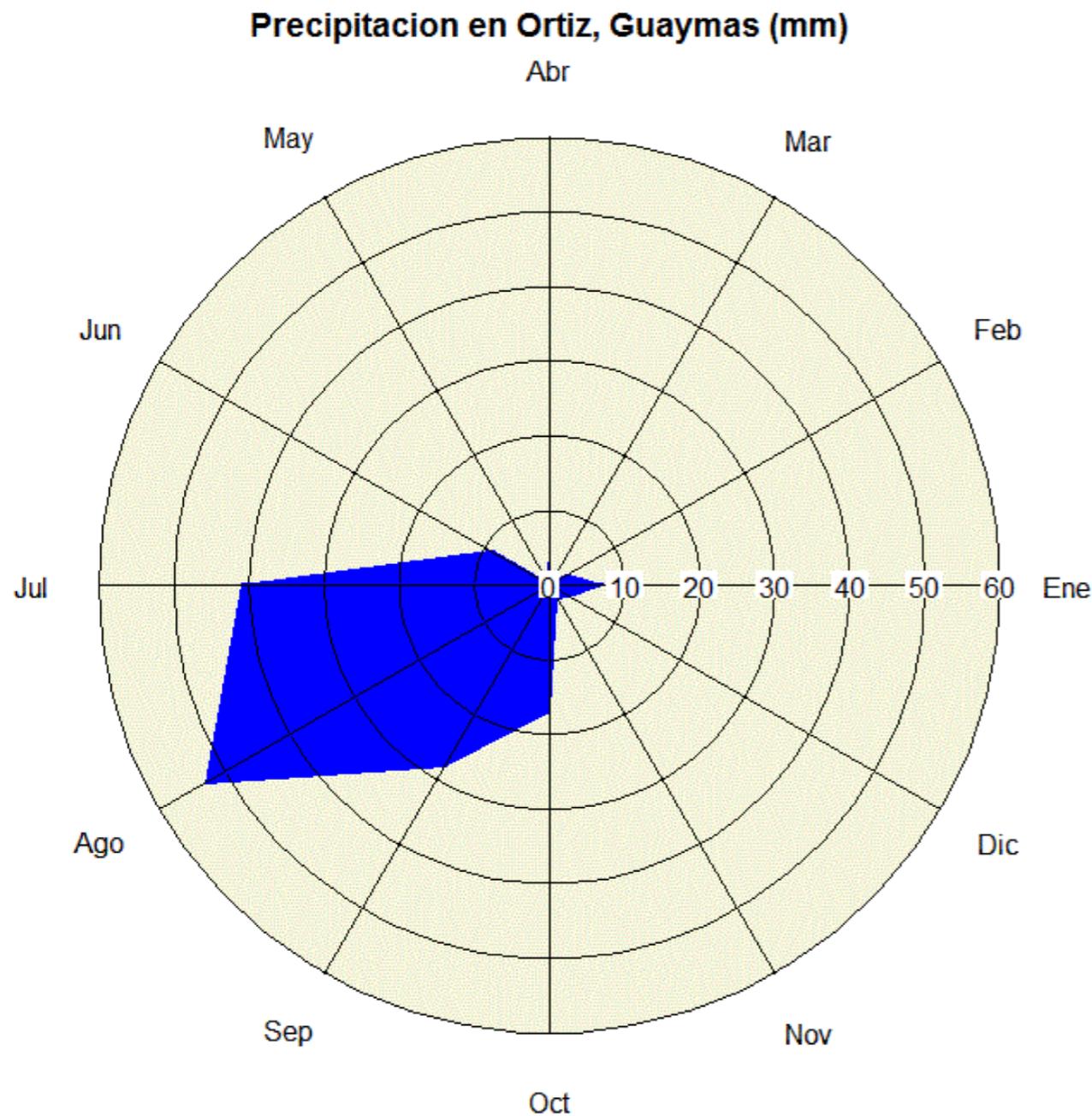
Ejemplo 2. La red de caminos en un área natural protegida en Álamos, Sonora

Red de Caminos del APFF Alamos Rio Cuchujaqui



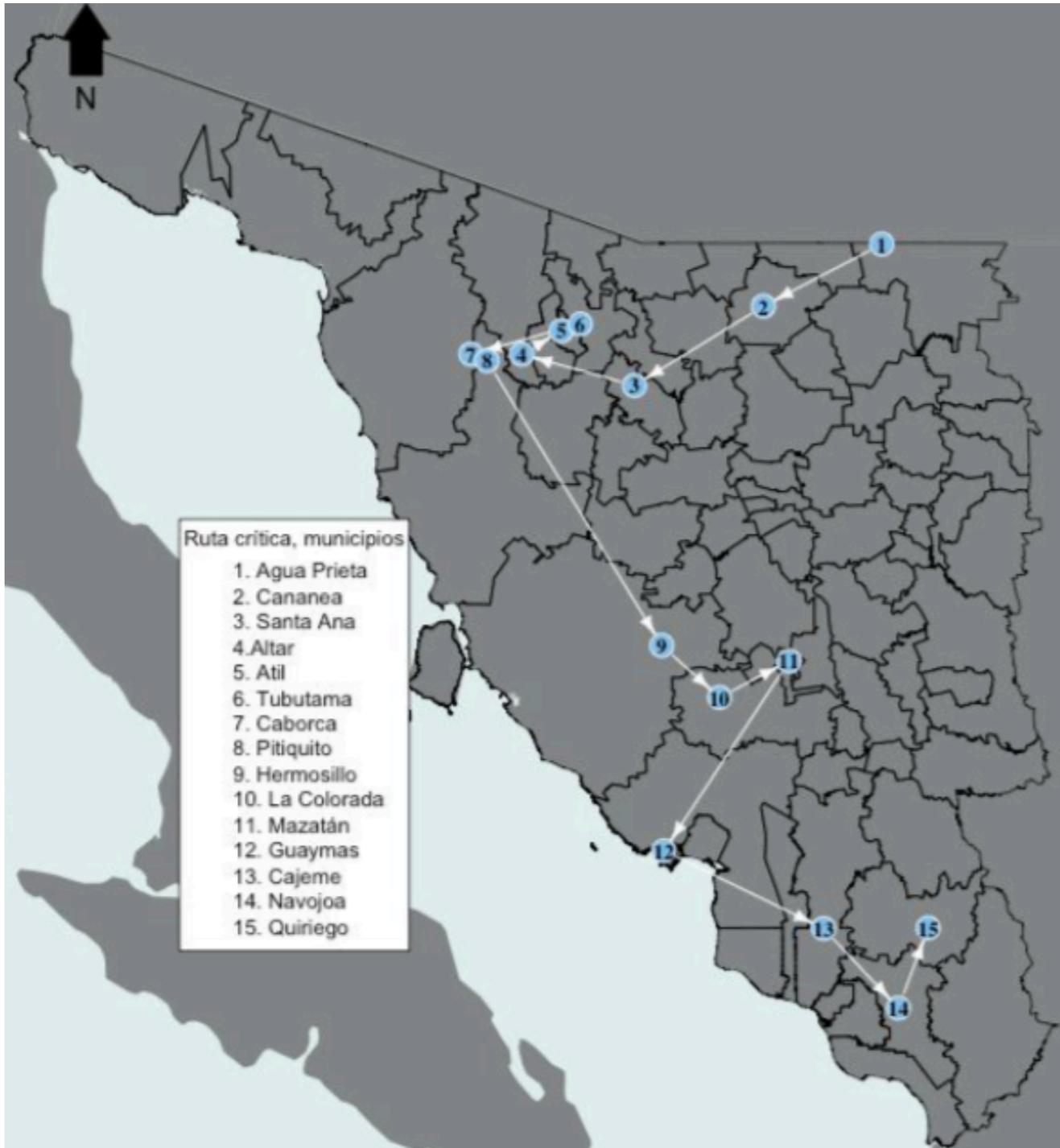
- *La línea negra es el límite del polígono del Área Natural de Protección de Flora y Fauna (APFF) "Álamos Río Cuchujaqui"*
- *En rojo está la red de caminos.*
- *Usé el paquete "mapprocs" para elaborar una gráfica con estos dos elementos.*

Ejemplo 3. ¿Cómo se distribuye la lluvia?



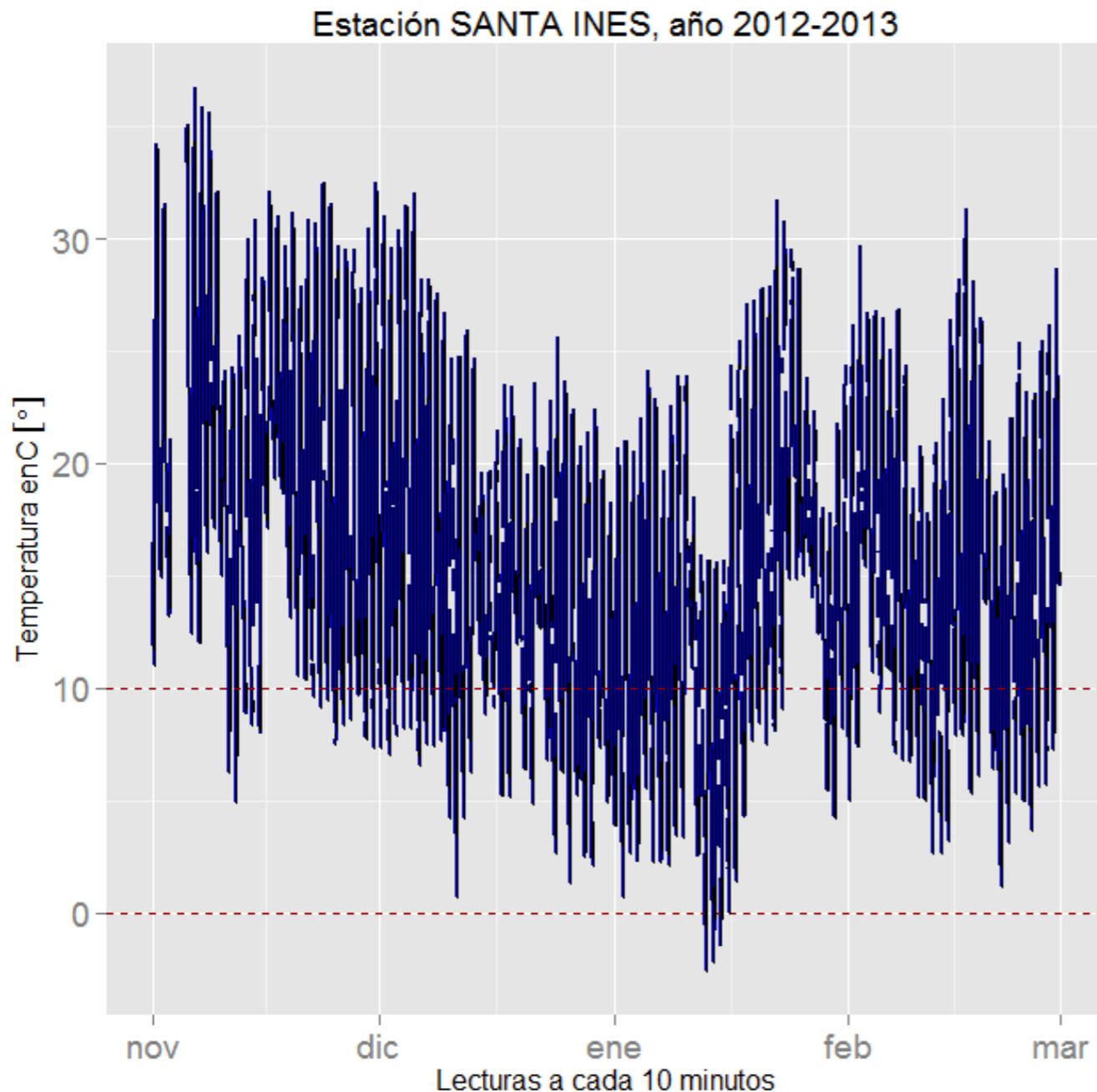
- *Buscaba una forma rápida y visual de mostrar al lector cómo se distribuía la lluvia en los meses del año y el volumen de lluvia recibido.*
- *Lugar Ortiz, Guaymas, Sonora.*
- *Utilice el paquete "plotrix".*

Ejemplo 4. Se pueden combinar redes con mapas



- *R* posee mucha capacidad analítica para el análisis de redes (sociales).
- Se hace una red, se georeferencian los nodos, se grafican sobre un mapa.
- Utilice los paquetes "GISTools", "igraph", "maptools".

Ejemplo 5. Es muy eficiente creando series de tiempo



- *La gráfica muestra las lecturas a cada 10 minutos de temperatura en una estación automatizada ubicada en la Costa de Hermosillo.*
- *Se utilizó el paquete "zoo".*

Múltiples bases de datos y disciplinas



Puedes importar datos de muchas fuentes



Referencia [10]



Referencia [7]



Referencia [9]



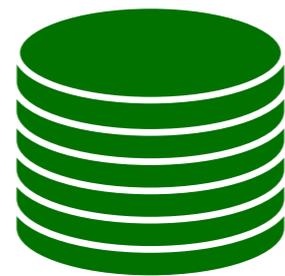
Referencia [11]



Referencia [8]



Referencia [12]



Referencia [13]

Conclusiones

- *R es un programa que demanda mucho tiempo, pero invertirle vale la pena.*
- *Las posibilidades son infinitas y sigue creciendo, cada vez se relaciona con otras plataformas como Google Earth, Netlogo, ArcGis, Python, etc.*
- *Atiende problemática específica. La mayoría de los científicos proveen el artículo científico y las bases de datos con las que puedes repetir el análisis.*
- *Usar R estimula la creatividad y la capacidad analítica del usuario.*
- *R está en constante mejora y evolución. Cada vez posee más herramientas de análisis, es más rápido y se integra a otros programas. También mejora los medios visuales (gráficas, mapas, etc.) para representar datos.*

Referencias y fuentes de imágenes

Celis de la Rosa Alfredo de Jesús y Vanessa Labrada Martagón (2014) *Bioestadística (3era Edición)* Editorial El Manual Moderno S.A. de C.V. México, D.F.
Ihaka R. & Gentleman R. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 299–314
Kreps, E. (2016) Woodcraft. Read Books Ltd.
Paradis, E. (2002) R para principiantes. Traducido por: Jorge A. Ahumada. Documento disponible on-line: http://cran.r-project.org/doc/contrib/rdebuts_es.pdf
R Development Core Team (2011). *R: A language and environment fo statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>

Referencia [1] Sistemas operativos <http://communitysnowobs.org/tutorials/mac-win-linux/>

Referencia [2] Icono colaboración <https://www.onlinewebfonts.com/icon/461234>

Referencia [3] Icono copiar y pegar <https://icon-library.net/icon/copy-icon-png-1.html>

Referencia [4] Imagen de curva tomada y adaptada de: <http://malaher.org/2007/03/pet-peeve-learning-curve-misuse/>

Referencia [5] Dependencia entre paquetes. <https://www.datacamp.com/community/tutorials/r-packages-guide>

Referencia [6] Datos para la gráfica tomados de: <https://gist.github.com/daroczig/3cf06d6db4be2bbe3368>

Referencia [7] Stata logotipo de: <https://www.statanordic.com/ebook-details.html>

Referencia [8] SPSS logotipo de: <https://www.monash.edu/research-portal/vlab/applications/spss>

Referencia [9] SAS logotipo de: <https://tigerware.lsu.edu/Software/183/SAS#>

Referencia [10] Excel logotipo de: <https://www.cleverducks.com/2018/01/04/moving-your-office-opening-a-new-office-your-it-service-provider-is-a-key-partner-in-the-process/excel-icon/>

Referencia [11] HTML icono de: https://www.flaticon.com/free-icon/html-file_81970

Referencia [12] SHP icono de: <http://geohipster.com/2017/10/09/shapefile-to-geohipster-80-of-successful-gis-work-is-having-a-good-folder-structure/>

Referencia [13] KML icono de: <https://www.redgps.com/blog-noticias/exportacion-de-recorridos-en-kml-55>

Fonts: "PlaytimeWithHotToddlersOblique", "RoughTypewriter", "Cisalpin LT Std", "Permanent Marker"