

# Taller de introducción a R

Darwin Del Castillo, MD

# About me

- ▶ Soy médico egresado de la UNMSM (Lima, Perú)
- ▶ Estudiante de la Maestría en Salud Pública en University of Washington (Seattle, WA).
- ▶ Anteriormente fui asistente de investigación en CRONICAS, Centro de Excelencia en Enfermedades Crónicas (UPCH).
- ▶ Actualmente trabajando como asistente de investigación en el Departamento de Salud Global de UW y en el Equipo de Informática Clínica del Institute for Health Metrics and Evaluation (IHME).
- ▶ Trabajo en la intersección de epidemiología y ciencias de datos en salud. Mi interés en investigación es **el origen temprano de las enfermedades no transmisibles** y en **ENT en edades tempranas**.



# Contenido del curso

Seis sesiones con los siguientes temas:

- ▶ Introducción a R y RStudio (IDE)
- ▶ Limpieza de datos en R/RStudio
- ▶ Estadística descriptiva: Tablas de resumen y gráficos básicos
- ▶ Estadística inferencial: Pruebas de hipótesis bivariadas
- ▶ Modelo de regresión lineal simple: supuestos, ajuste y diagnóstico
- ▶ Extras: regresión múltiple y otros modelos, introducción a version control con GitHub

# Objetivos del curso

- ▶ Conocer las bases del lenguaje de programación R y su IDE RStudio
- ▶ Realizar limpieza de datos en R
- ▶ Identificar pruebas de hipótesis y gráficos adecuados útiles para tus trabajos de investigación (**tesis**)
- ▶ Conocer conceptos básicos de version control con GitHub
- ▶ Reconocer cuando pedir ayuda y qué pedir exactamente

# ¿Qué es R?

- ▶ R es un lenguaje de programación y un entorno de desarrollo para análisis estadístico y gráfico.
- ▶ Utilizado principalmente en biomedicina por sus capacidades estadísticas y su habilidad para generar gráficos de alta calidad.
- ▶ R base es útil, el poder del código base de R se amplifica cuando añadimos paquetes (explicaremos esto más adelante)


# ¿Por qué R? — Comunidad

r-bloggers.com

Menu ▾

## R-bloggers

R news and tutorials contributed by hundreds of R bloggers



### How Are P-values Distributed Under The Null?

I sometimes use this fun interview question for aspiring data scientists: How are p-values distributed assuming the null hypothesis is true? I've heard a lot of reasonable answers, including: It should be centered towards large values it should have almost zero mass below ... [Continue reading →](#)

January 22, 2025

### rOpenSci 2024 Highlights

In 2024, we remained committed to open science and open source software. We continued to build a welcoming and inclusive community, through innovation and collaboration. Here are the highlights of the year: R-Univers R-Univers, rOpenSci's platform for finding and publishing R ... [Continue reading →](#)

January 21, 2025

stackoverflow.com

Products

2

## [r]

Ask Question

R is a free, open-source programming language and software environment for statistical computing, bioinformatics, information graphics, and general computing.

Watching ▾

Go to Wiki

510,387 questions

Newest

Active

More ▾

Filter

0 votes

0 answers

11 views

### R Trouble With Paste0 in lm

I am trying to use paste0 so that I can run a series of linear models (with a quadratic term), one for each independent variable. Unfortunately, when I...

R r lm

JohnS 21 asked 34 mins ago

0 votes

0 answers

5 views

### Custom sort order on dynamic legend items for ggpat-tern and ggplot items

I'm trying to sort legend entries across both ggplot and ggpattern items. The map will not always include all the layers, so the legend should be d...

R r sorting ggplot2 legend ggpattern

mmooore 185 asked 34 mins ago

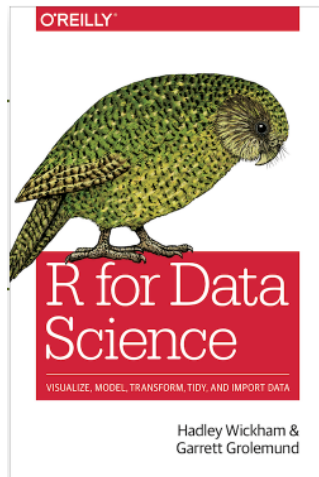
0 votes

0 answers

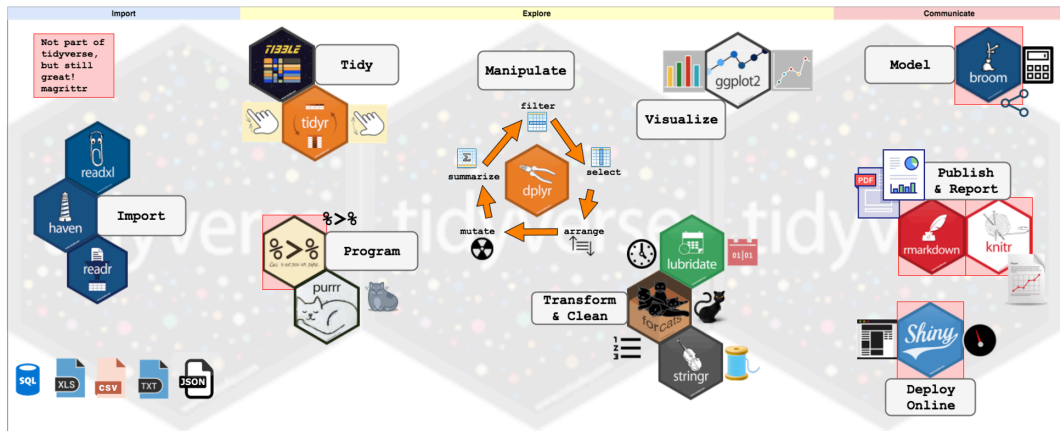
11 views

### How to use column names in a list of values in a loop?

I am trying to run a process that will can that saves specific lag lengths of different variables as values to be used later. The code below is what I am...




# ¿Por qué R? — Flexibilidad









# ¿Por qué R? — Gráficos, documentos, presentaciones, páginas web

Darwin Del Castillo

About CV Publications Teaching Index



**Darwin Del Castillo**



Welcome to my new site!

I am a Peruvian physician-scientist, epidemiologist, data analyst, musician, gamer, and a Master in Public Health candidate at University of Washington, Seattle in the [Department of Global Health](#). My research focus is to understand the link between adverse experiences during pregnancy, the perinatal period, and infancy, and the development of cardiometabolic diseases during youth and young adulthood.



## ¿Por qué R? — Lo más importante

