

# Incorporando R como herramienta de análisis de datos en el curso de Estadística para estudiantes de Computación



M.Sc. Esteban Ballesterero

# Punto de partida

- *Nombre del curso:* Estadística
- *Carrera en la que se imparte:* Ingeniería en Computación y Computadores
- *Características del curso:* Es un curso de estadística inferencial paramétrica clásico, que ofrece al estudiante la posibilidad de aprender técnicas estadísticas empleadas en la toma de decisiones, tales como estimación, pruebas de hipótesis y regresión.
- *Metodología:* en su mayoría las clases son de tipo magistral donde la teoría se desarrolla de manera semiformal. El uso de software queda criterio del docente, pero muchas veces el uso que se le da es más didáctico que como herramienta de análisis de datos.
- *Evaluaciones:* tareas, pruebas cortas y exámenes diseñados por cátedra.

# Punto de partida

**Teorema 16** Considere la población dada por la variable aleatoria  $X$  y el estadístico  $\hat{\Theta}$  asociado al parámetro de población  $\theta$  para muestras de tamaño  $n$  con

$$\text{Var}(\hat{\Theta}) = \bar{\sigma}^2$$

Si  $\hat{\Theta}$  es insesgado ( $E(\hat{\Theta}) = \theta$ ) y su función de distribución es simétrica con respecto a  $\theta$  entonces un intervalo de confianza de  $100(1 - \alpha)\%$  para  $\theta$  tiene extremos

$$\hat{\theta} \pm A_{\alpha/2} \cdot \bar{\sigma}$$

donde:  $\hat{\theta}$  es el valor de  $\hat{\Theta}$  para una muestra de tamaño  $n$  y  $A_{\alpha/2}$  cumple que

$$P\left(\frac{\hat{\Theta} - \theta}{\bar{\sigma}} < A_{\alpha/2}\right) = \frac{\alpha}{2}$$

En este caso,  $\hat{\theta}$  es el centro del intervalo y  $|A_{\alpha/2} \cdot \bar{\sigma}|$  es su radio.

**Prueba.** Considere la variable aleatoria  $A = \frac{\hat{\Theta} - \theta}{\bar{\sigma}}$ , note que  $E(A) = 0$  y como la función de distribución de  $\hat{\Theta}$  es simétrica con respecto a  $\theta$  entonces la función de distribución de  $A$  es

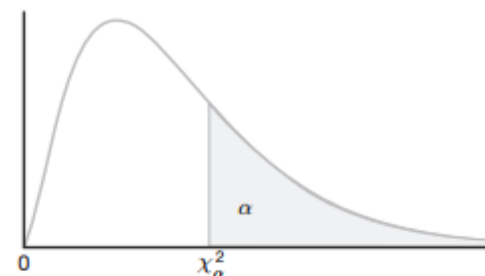
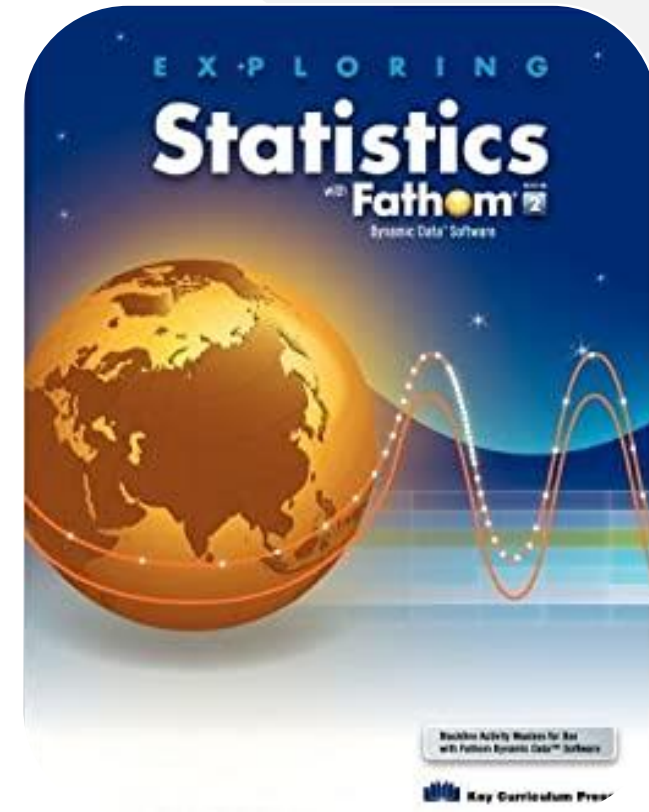


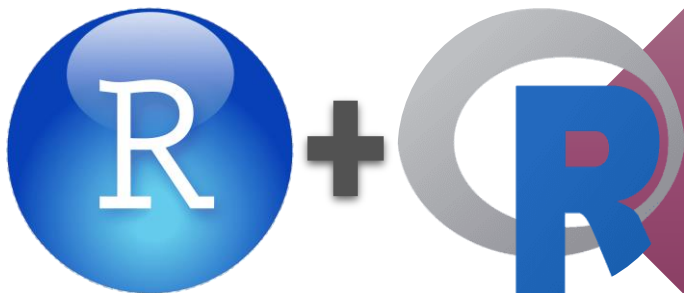
Tabla A.5 Valores críticos de la distribución chi cuadrada

$\nu$	$\alpha$									
	0.995	0.99	0.98	0.975	0.95	0.90	0.80	0.75	0.70	0.50
1	0.00393	0.00439	0.00482	0.00521	0.00564	0.00615	0.00681	0.00738	0.00794	0.00939
2	0.0100	0.0137	0.0175	0.0200	0.0233	0.0270	0.0318	0.0358	0.0394	0.0497
3	0.0717	0.115	0.185	0.216	0.352	0.584	1.005	1.213	1.424	2.366
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	1.923	2.195	3.357
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	2.675	3.000	4.351
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	3.455	3.828	5.348
7	0.989	1.239	1.564	1.690	2.167	2.833	3.822	4.255	4.671	6.346
8	1.344	1.647	2.032	2.180	2.733	3.490	4.594	5.071	5.527	7.344
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	5.899	6.393	8.343
10	2.156	2.558	3.059	3.247	3.940	4.865	6.179	6.737	7.267	9.342
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	7.584	8.148	10.341
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	8.438	9.034	11.340
13	3.565	4.107	4.765	5.009	5.892	7.041	8.634	9.299	9.926	12.340
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	10.165	10.821	13.339
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	11.037	11.721	14.339

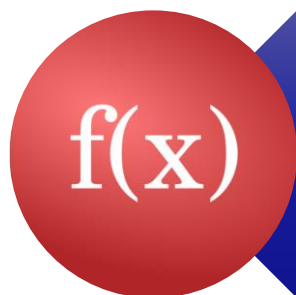
# Punto de partida



# El cambio



Incorporación de **R** como herramienta de análisis de datos a partir de la IDE de **Rstudio**



Usar la App **Probability distributions** para el cálculo de probabilidades



Uso opcional de Excel para crear las bases de datos, con la facilidad de poderlas convertir a formato **.CSV**

# El cambio: metodológico

- En las primeras clases se hicieron algunas actividades de instrumentación para que el estudiante instalara el software requerido, se familiarizara con la plataforma de RStudio y la sintaxis del lenguaje.
- Se abordaban los conceptos teóricos en la clase y luego, se desarrollaban de forma paralela algunos ejemplos de manera tradicional y luego hacía uso de R.
- Los códigos generados para cada temas, se trabajaron en formato Rmarkdown y se exportaban como HTML o PDF para los estudiantes.



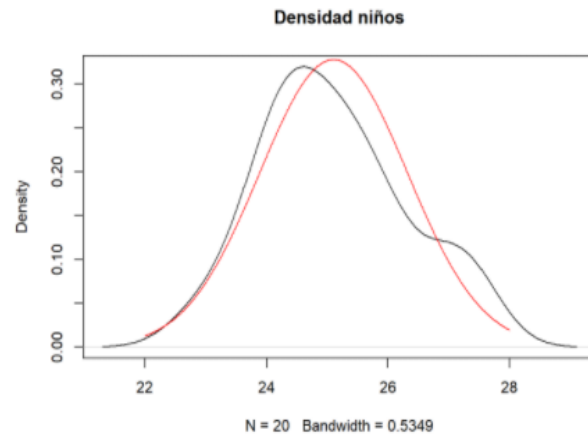
# El cambio: evaluación pruebas

## Examen

Steve Ulate-Jose Okon-Jefferson Torres

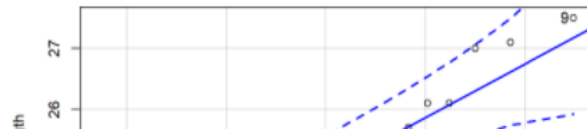
27/6/2020

## Visualización niños

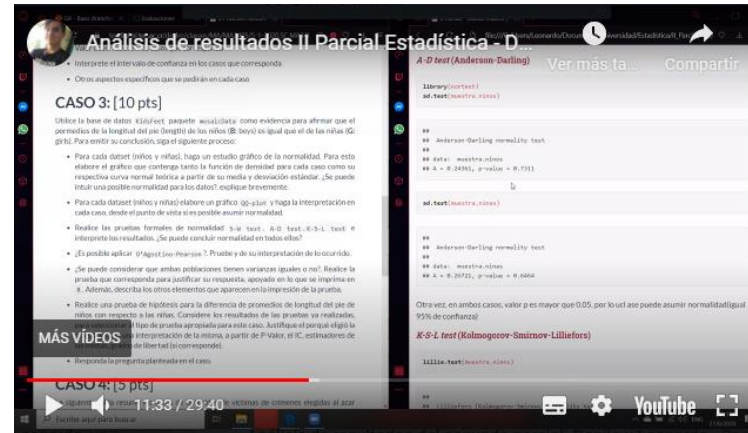


La curva de densidad adopta una forma similar a la campana de Gauss, por lo tanto se puede intuir normalidad

## Usando QQPlot niños



- Los estudiantes realizaron pruebas parciales que debían resolver haciendo uso de R. Para estas pruebas trabajaron tanto individualmente como en parejas y tríos, la solución se enviaba mediante un documento PDF o HTML para su respectiva revisión junto con un vídeo de análisis.
- Por motivo de la pandemia ocasionada por el COVID-19, el curso se tuvo que impartir un 30% presencial y un 70% virtual



# El cambio: evaluación

Pruebas parciales

Entrevista a experto

Taller de estadística descriptiva

Esta se realizó en 3 fases:



# El cambio: evaluación entrevista

- En grupos de 3 personas los estudiantes realizaron una entrevista a profesionales de su campo que trabajan directamente o están vinculados con análisis de datos.
- Se coordinó con el profesional para realizar una videollamada y grabarla, para que posteriormente esta fuese presentada al resto de los compañeros de la clase.
- La entrevista constaba de 3 preguntas relacionadas con la necesidad de que un Ingeniero en Computación conozca de análisis de datos y en particular, de saber procesarlos y analizarlos usando R.
- Entrevistas variadas: sector público, privado, internacionales, con conocimiento en diferentes tipos de datos (espaciales, demográficos, Smart Cities, etc)



# El cambio: evaluación entrevista

## Guía de la entrevista

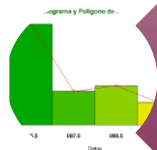
- a. Nuevas tendencias mundiales, **visualizan los datos como el futuro petróleo de la humanidad.** ¿Está de acuerdo con dicho argumento? Refiérase al respecto.
- b. El perfil de profesional en Computación durante muchos años estuvo enfocado en formar desarrolladores, pero el avance de la tecnología hace que el Ingeniero en Sistemas de la actualidad, sea una persona más versátil. **Considerando desde su perspectiva los atributos deseables que deba tener el Ingeniero en Computación del futuro, ¿es el análisis de datos uno de estos?** Explique su respuesta indicando si esta habilidad le daría ventaja o no en el mercado laboral.
- c. En la industria de análisis de datos existen varias herramientas tecnológicas disponibles, algunas de uso libre y otras bajo licencias que requieren de un considerable costo para adquirirlas. Algunos perfiles profesionales que busca el sector empleador, está considerando de manera más frecuente como requisito deseable que la persona a contratar tenga conocimientos avanzados en R, **¿Cuál sería su opinión en relación a que un Ingeniero en Computación conozca de R para analizar datos a nivel avanzado?**
- d. Libremente, si lo desea puede realizar un mensaje de cierre, considerando que este vídeo será visto por un grupo de estudiantes que están en fase de formación para ser Ingenieros en Computación.



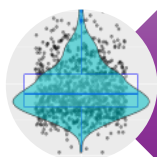
# El cambio: evaluación taller

- En grupos de 3 personas los estudiantes diseñaron una actividad tipo taller sobre un tema de estadística complementario al curso.
- Para el tema asignado se le proporcionaron unos materiales que servirían como base, pero debían investigar y complementarlo
- Cada grupo diseñó una actividad taller para sus compañeros, quiénes debieron trabajar las actividades durante dos semanas
- El actividad taller debía contener el manual en formato HTML generado mediante un documento Rmarkdown con los ejemplos desarrollados, un vídeo explicativo de dicho material y al menos una actividad corta como evaluación para los compañeros de clase.

# El cambio: evaluación taller



Tabulación, graficación y medidas de tendencia central para datos agrupados y no agrupados



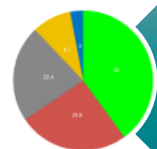
Medidas de Dispersión: Datos agrupados y no agrupados  
Diagramas de caja (Boxplot) y Violin plot



Elaboración de documentos RMarkdown



Gráficos de barras y su uso



Gráficos Lollipop y de pastel



Introducción al text mining: gráficos de palabras (Word Cloud's)


# El cambio: paquetes de R

## R Markdown

Hoja de Referencia

lee mas en [rmarkdown.rstudio.com](https://rmarkdown.rstudio.com)

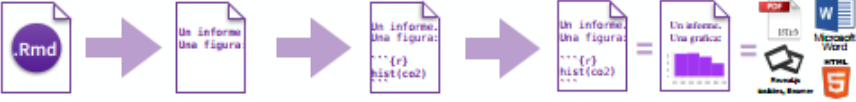
markdown 0.2.50 Actualizado: 8/14



### 1. Flujo de trabajo

R Markdown para escribir informes reproducibles y dinámicos con R. Usalo para incluir código R y resultados en presentaciones, documentos pdf, html, Word y mas. Para crear un informe:

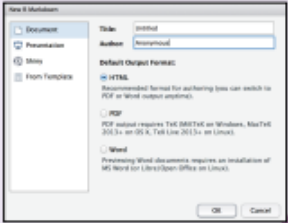
- i. Abre - Abre un archivo con extensión .Rmd.
- ii. Escribe - Escribe contenido con la sintaxis R Markdown.
- iii. Incluye - Incluye código R para crear salidas en el informe.
- iv. Genera - Reemplaza el código R con sus salidas y transforma el informe a una presentación, PDF, HTML o Word.



### 2. Abre archivo


Empieza guardando un archivo de text con la extensión .Rmd o abre RStudio.

- En el menu haz clic en: **File > New File > R Markdown...**
- Una ventana abre donde puedes seleccionar el tipo de formato que quieres tener como resultado de tu archivo .Rmd.
- Selecciona el tipo de formato que quieres como salida con los botones (puedes cambiarlo después)
- Haz clic en OK



### 3. Markdown

Ahora, escribe tu informe en texto plano. usa la sintaxis de markdown para describir el formato de tu texto en el informe final.

sintaxis	resulta en
Texto plano Termina línea con dos espacios para nuevo párrafo. <code>*cursivo*</code> y <code>_cursivo_</code> <code>**negrita**</code> y <code>__negrita__</code> <code>superíndice<sup>24</sup></code> <code>~~tachado~~</code> <code>[etiquetón](www.rstudio.com)</code>	Texto plano Termina línea con dos espacios para nuevo párrafo. cursivo y cursivo negrita y negrita superíndice <sup>24</sup> tachado etiquetón
<code># Encabezado 1</code> <code>## Encabezado 2</code> <code>### Encabezado 3</code> <code>#### Encabezado 4</code> <code>##### Encabezado 5</code> <code>##### Encabezado 6</code>	Encabezado 1 Encabezado 2 Encabezado 3 Encabezado 4 Encabezado 5 Encabezado 6
<code>raya en: --</code> <code>raya en: ---</code> <code>elipsis: ...</code> <code>ecuación en línea: <math>SA = \pi r^2</math></code> <code>imagen: </code> <code>regla horizontal (o nueva diapositiva):</code> <code>---</code>	<code>raya en: --</code> <code>raya en: ---</code> <code>elipsis: ...</code> <code>ecuación en línea: <math>A = \pi r^2</math></code> <code>imagen: </code> <code>regla horizontal (o nueva diapositiva):</code> <code>---</code>
<code>&gt; cita en bloque</code> <code>* lista sin orden</code>	cita en bloque * lista sin orden

### 4. Escoge salida

Escribe un encabezado en YAML que explica que tipo de documento quieres construir con tu archivo R Markdown.

**YAML**

Un encabezado YAML es un conjunto de pares llave:valor al comienzo de tu archivo. Comienza y termina el encabezado con tres rayas (`---`)

```
---
title: "Sin título"
author: "Andrés"
output: html_document
---
```

Este es el comienzo de mi informe. Arriba están los metadatos en un encabezado YAML.

El generador de planillas de RStudio te genera el encabezado YAML

El valor de salida determina que tipo de archivo R construirá con base en tu archivo .Rmd (en Paso 6)

```
library(stats)
library(car)
library(nortest)
library(fBasics)
library(moments)
library(ggplot2)
library(corrplot)
library(corr)
library(ggcorrplot)
library(colorspace)
library(mosaicData)
library(wordcloud)
library(RColorBrewer)
library(randomcoloR)
library(tm)
library(SnowballC)
library(tidyverse)
library(ggalt)
library(datos)
library(nycflights13)
library(dplyr)
library(nycflights13)
```

# El cambio: aprendizaje

- Uso de IDE de Rstudio, aprovechando muchas de sus bondades
- Generar reportes de calidad mediante documentos Rmarkdown exportados en diferentes formatos (uso de diferentes templates para mejorar la presentación, creación de tablas, formato de títulos, viñetas, insertar enlaces web, insertar chunk's de código y propiedades de estos, formato para el texto plano, inserción de texto simbólico usando LaTeX, cargar imágenes)
- Crear proyectos en R
- Bases de datos: cargar, extracción y modificación
- Operaciones con matrices
- Uso de paquetes específicos para diferentes análisis estadísticos estudiados en el curso
- Generación de gráficos de alta calidad
- Operaciones estadísticas básicas en R
- Interpretación de resultados
- Realizar simulaciones

# El cambio: percepción del estudiante

Se aplicó un formulario a los estudiantes para valorar sus apreciaciones sobre los cambios realizados en el curso, de los cuáles 18 de 36 lo respondieron.

Seguidamente algunos de los resultados sobre la utilidad de las herramientas utilizadas en el curso:

	Valoración en escala de 1 a 5 (Deficiente:1, ..., Muy buena:5)				
	1	2	3	4	5
Uso de R para realizar análisis estadístico			2	2	14
Documentos Rmarkdown con los contenidos y ejemplos resueltos en R				6	12



# El cambio: percepción del estudiante

Se les consultó a los estudiantes para cada uno de los recursos utilizados a lo largo del curso, si recomendaban mantenerlo tal cuál, con modificaciones o eliminarlo para una futura impartición de este:

Aspecto a evaluar	Valoración en escala de 1 a 5 (Deficiente:1, ..., Muy buena:5)		
	Eliminar	Mantener con cambios	Mantener de manera similar
Uso de R para realizar análisis estadístico	0	2	16
Documentos Rmarkdown con los contenidos y ejemplos resueltos en R	0	1	17
Actividad taller	3	6	9
Entrevista a experto	2	3	13

# El cambio: percepción del estudiante

Si tiene algún comentario que desea externar en relación al curso, por favor, sírvase escribirlo en el siguiente espacio:

Agradecer al profesor primeramente, ya que de no haber cambiado la modalidad del curso probablemente no fuera tenido experiencia con R a lo largo de toda la carrera. Esto es algo de lo que estoy agradecido ya que se nota que el profesor conoce lo que se está dando en el área laboral y suplir a los estudiantes de herramientas y conocimientos que puedan abrir nuevas puertas para nuevos trabajos es algo que siempre se agradece.

La forma en la cual se abarcó el curso fue muy provechosa y excelente, las presentaciones, y documentos generados de R estuvieron genial. Muy agradecido de la forma en que se dio el curso y por la disposición y amabilidad del profesor.

Excelente el enfoque usando R.

# El cambio: percepción del estudiante

Si tiene algún comentario que desea externar en relación al curso, por favor, sírvase escribirlo en el siguiente espacio:

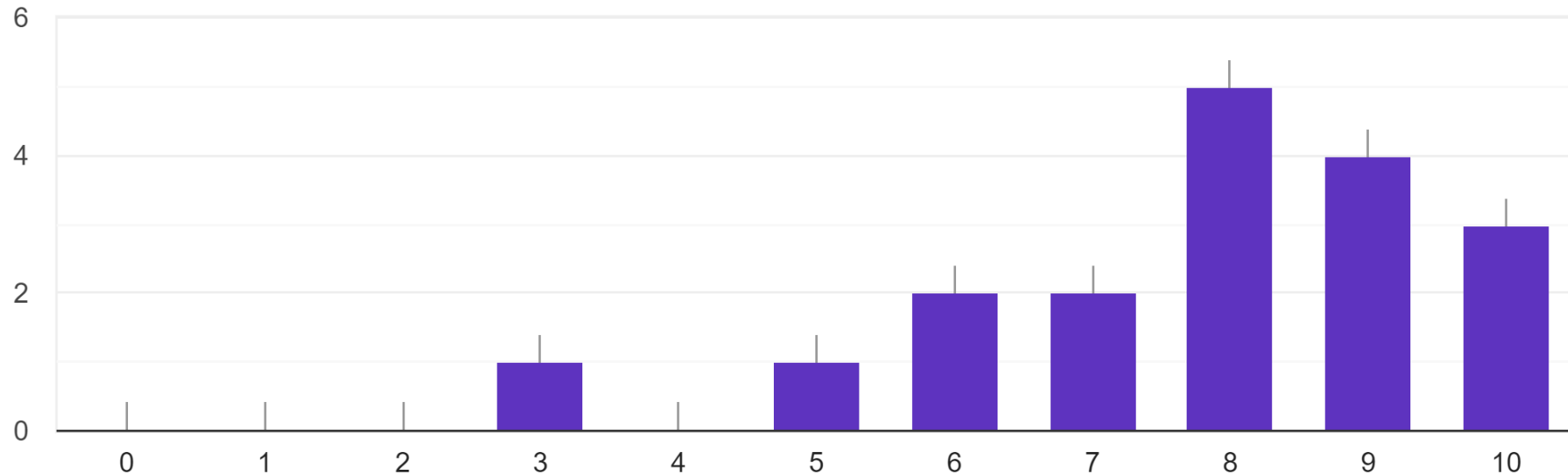
El curso a pesar de los cambios que se tuvieron que hacer a causa de la pandemia estuvo bien. En relacion a la inclusion de R, estuvo interesante pero siento que se deberia de hacer un enfoque de hacia que va a ir el curso. Si el enfoque es aprender a utilizar R tal vez sería mejor si el curso fuera como los de programacion normal que se da en la carrera. Porque al final lo que se realiza es un copia y pega de los algoritmos vistos en clase con el unico cambio de los datos a ingresar. Si el objetivo es realizar los análisis de los resultados, seria mejor en la evaluaciones en lugar de hacer todo un ejercicio desde cero, que venga ya planteado.

Por último, si se decide seguir llevando la herramienta R, que no es mala, sería una buena idea asociarse con los cursos de bases de datos, ya que ambos estan muy correlacionados y en el caso de nosotros, terminamos utilizando R en el curso de bases para correrlo con python y con jupyter para hacer un análisis final.

# El cambio: percepción del estudiante

En la escala se 0 a 10, cómo cambió su percepción en relación con la importancia tener conocimientos en un programa como R para hacer análisis de datos

18 respuestas

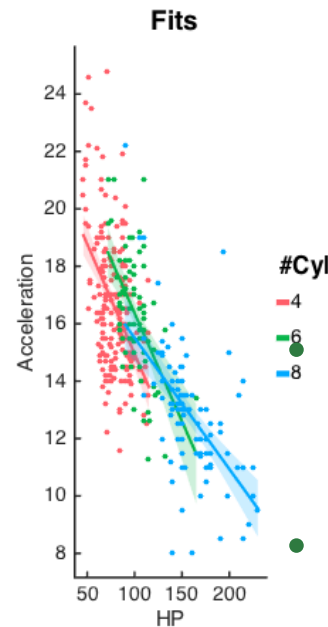
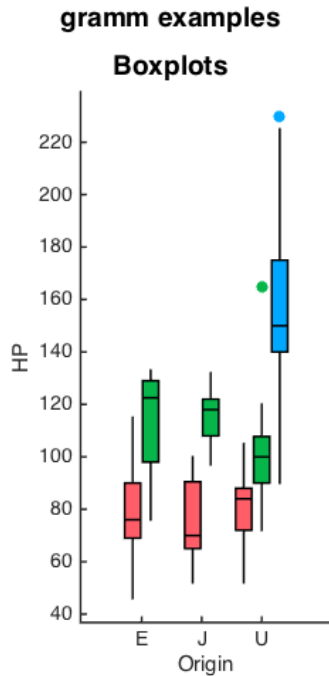
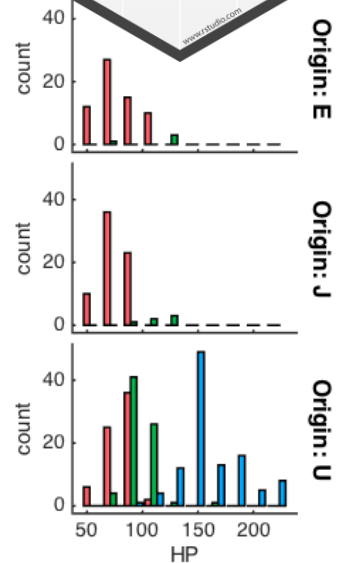
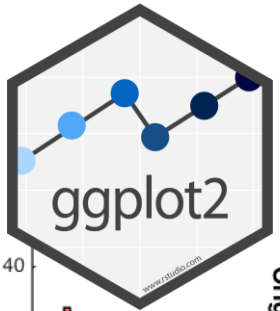


# ¿Por qué R?



- **R es el software básico de referencia en el mundo de la estadística, una de las herramienta más potentes y eficientes del mercado.**
- Es un software OPEN SOURCE, es decir que no usa encriptación (código abierto), por tanto, no tiene coste de adquisición y además está disponible para todos los sistemas operativos
- Es estable, confiable, flexible y a la vanguardia, debido a que está en permanente actualización con nuevas funciones y paquetes accesibles en tiempo real.
- **R está pensado para cálculos estadísticos, cálculos matemáticos y creación de gráficos de una manera fácil y rápida**, con mucha documentación disponible gracias a la gran comunidad activa que posee por ser un proyecto colaborativo.
- **R es capaz de conectarse a cualquiera de las bases de datos más utilizadas hoy en día**, leerla, procesarla y calcularla con una sola línea de código para todas las variables del dataSet.

# ¿Por qué R?



- R tiene un paquete para casi todas las aplicaciones para el tratamiento de **variables cuantitativas, cualitativas, datos estructurados y no estructurados** (vídeo, audio, imágenes, texto, etc), estadísticas en general, estadística bayesiana, métodos para Deep y Machine Learning, Minería de Datos, GIS, Web Mining , entre otros.
- R también maneja el **Álgebra** de matriz particularmente bien y es utilizado para el trabajo de modelos recursivos propios de los **Métodos Numéricos**.
- La visualización de datos es una fortaleza clave con el uso de bibliotecas como **ggplot2** bajo su filosofía de **Gramática de Gráficos** basado en capas.

# ¿Por qué R?



- Tiene la **versatilidad de vincularse a otras plataformas**, lo que permite desde la IDE de RStudio, manejar un repositorio de versiones de código en GitHub, se puede trabajar online y generar código práctico desde Jupyter, trabajar programación en paralelo usando Hadoop o Spark, para el abordaje de problemas complejos de Big Data.
- La IDE de RStudio cuenta con **herramientas para generar reportería e informes de alta calidad en diferentes formatos, con gráficos interactivos para los informes gerenciales o de investigación**, donde se puede tener acceso al código y realizar las pruebas de reproducibilidad, comúnmente exigidas en la actualidad por las revistas de publicaciones científicas.
- Muchas **empresas y universidades utilizan el software R para sus análisis estadísticos, por su robustez, su gran modularidad e infinitas posibilidades**, que se adaptan a todo tipo de necesidades para el manejo de datos complejos, además podemos ir ampliando nuestro entorno R con aspectos que realmente vayamos a trabajar y que nos permitirán tomar decisiones concluyentes y resultados que suelen ser realmente satisfactorios.





# ¿Por qué R en Ing. En Computación?

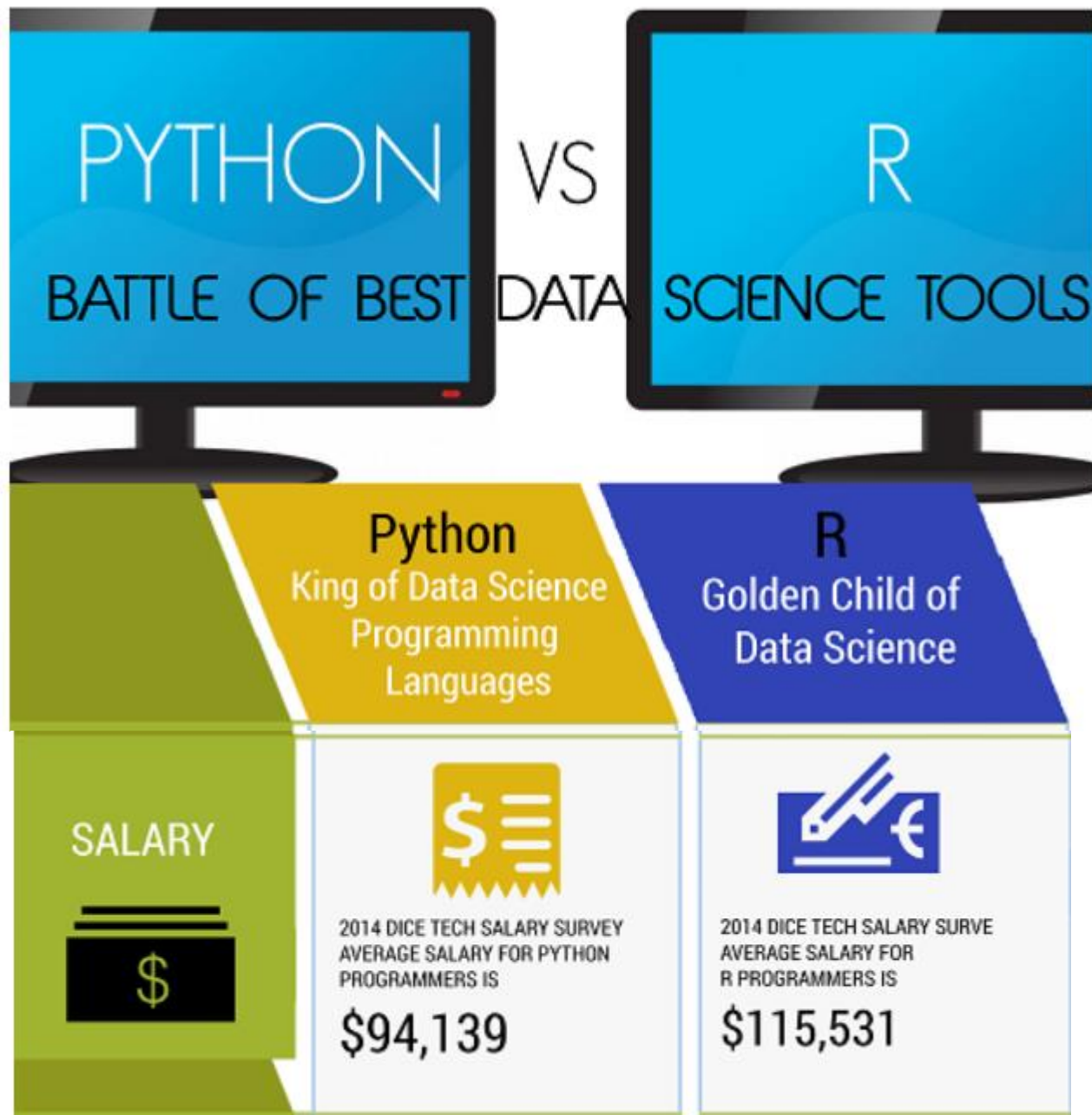


- La carrera cuenta desde hace dos años con una **Comunidad** para los estudiantes de computación en **Data Science**.
- La carrera de Computación en San Carlos, inició con la implementación de una **especialización en Ciencia de Datos a través de la FUNDATEC**, pero tiene como objetivo junto con la Comunidad de Data Science, promover este tipo de temáticas a través de eventos académicos y oportunidades de capacitación en la zona, de manera que se vaya generando capital humano con competencias en esta disciplina.
- La carrera ha tenido sus primeras experiencias en proyectos de investigación, colaboró en agosto de 2019 con el primer **Datatlón** donde mediante modalidad de competencia grupal se analizaron bases de datos públicas del país.

# ¿Por qué R en Ing. en Computación?



- En Julio de 2019 la carrera de Computación organizó el **Primer Simposio Data Science** dentro del reconocido congreso Iberoamericano de Computación para el Desarrollo (**COMPDES**). En este evento se contó con temáticas de aplicación de Data Science en la academia, la banca, empresas analistas de datos, cooperativas, entre otras.
- Desde el **Seminario de Actualidad Científica y Tecnológica** se han traído especialistas nacionales e internacionales, que han impartido conferencias en temáticas como Big Data, Minería de Datos, Aplicaciones de Big Data en la Física y Computación Cuántica.
- La **empresa privada** ha venido incrementando el uso de R en los servicios que ofrece, por lo que en los últimos años se ha evidenciado una **creciente solicitud de profesionales en el campo de la computación con dominio de R**.



Este estudio de 2014, mostraba la diferencia salarial anual promedio, entre un profesional con conocimiento de Python vs conocimiento en R.

Esa tendencia se mantiene en la actualidad, por asuntos de oferta y demanda, es más sencillo encontrar un profesional en computación con conocimientos en Python, pero en con competencias para programar en R, es más escaso.

# Profesiones asociados al profesional de computación que demanda conocimiento de R



- Data Analyst
- Data Scientist
- Marketing
- System Integration Engineer
- Data Engineer
- CTO Office
- Big Data Engineer
- BI Developer
- Data Architect
- Otros derivados



## Data Analyst

Grandes Ideas Empresariales SW S.A. • San José, CR

Publicado: hace 1 semana •  Sé de los primeros 25 solicitantes

[Guardar](#) [Solicitar](#)

Requerimos de analistas de datos con alta experiencia en Interpretar datos, analizar resultados utilizando técnicas estadísticas. Desarrollar e implementar análisis de datos, así como sistemas de colección de datos y otras estrategias que ayuden a optimizar la eficiencia estadística y la calidad. Esta persona transformará los datos en información, información en hechos y estos hechos en decisiones de negocio. Conocimiento fuerte y experiencia en paquetes de reporte Business Objects, Power BI, etc., bases de datos SQL  etc, programación XML, JavaScript o ETL frameworks

Conocimiento en estadística y experiencia utilizando paquetes de análisis estadístico para analizar sets de datos Excel, SPSS, SAS, etc

Altas capacidad de análisis

Alta habilidad de colección, organización, análisis y distribución de una alta y significativa cantidad de información con alta atención al detalle y precisión. Altamente identificado con queries, levantamiento de reportes y presentación de hallazgos.

Altamente comprometido

Planificación y control

Orientado a resultados

Orientado hacia la resolución de problemas

Capacidad de organización

Con iniciativa

Buen comunicador

**Nivel de experiencia**  
Sin experiencia

**Sector**  
Servicios y tecnologías de la información,  
Software, Internet

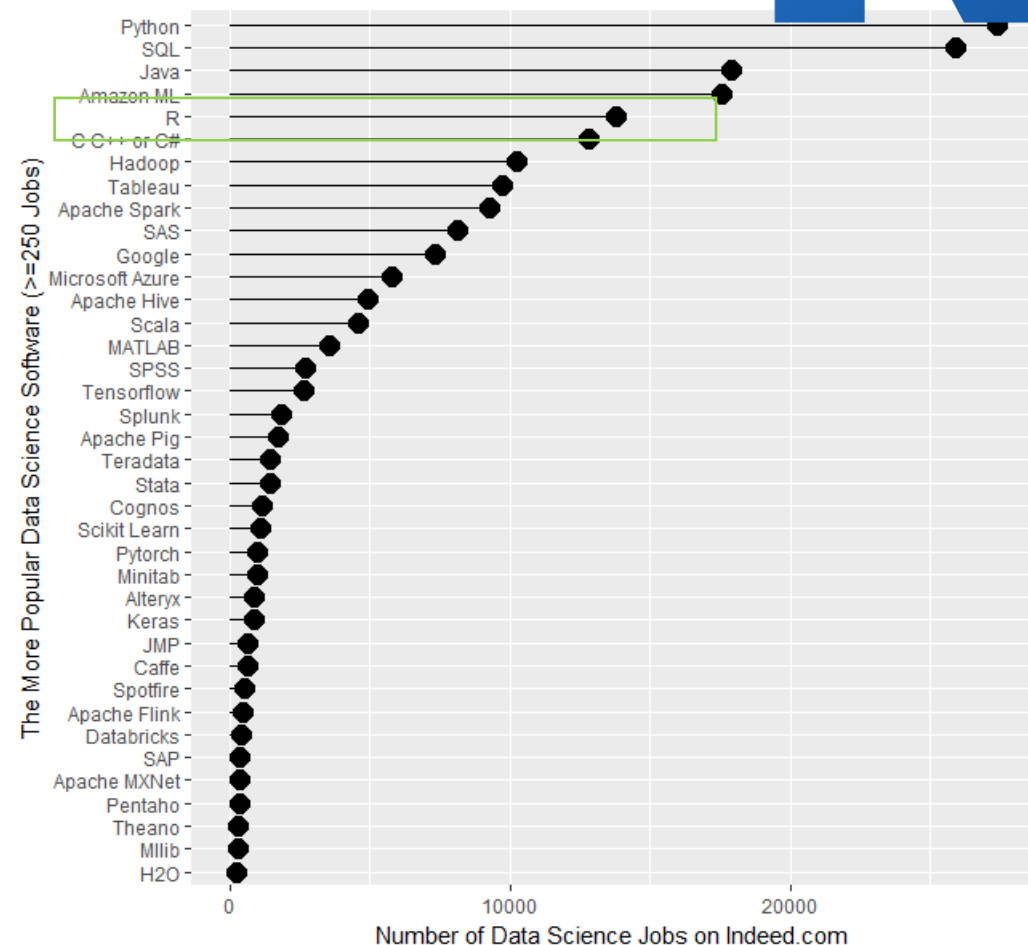
**Tipo de empleo**  
Jornada completa

**Funciones laborales**  
Tecnología de la información

# Demanda laboral

El siguiente gráfico muestra la demanda en trabajos relacionados con Ciencia de Datos en Indeed.com, según la función de las herramientas tecnológicas requeridas por el empleador.

Se puede notar que R se encuentra dentro de las 5 más demandadas.



# Uso de R en el mundo

El índice de TIOBE, es un informe mensual que elabora y publica la empresa TIOBE Software BV basado en el un cuantificación del uso de los distintos lenguajes de programación a nivel mundial, según los mayores motores de búsqueda del mundo (Google, MSN, Yahoo!, Baidu, Wikipedia, YouTube, entre otros)

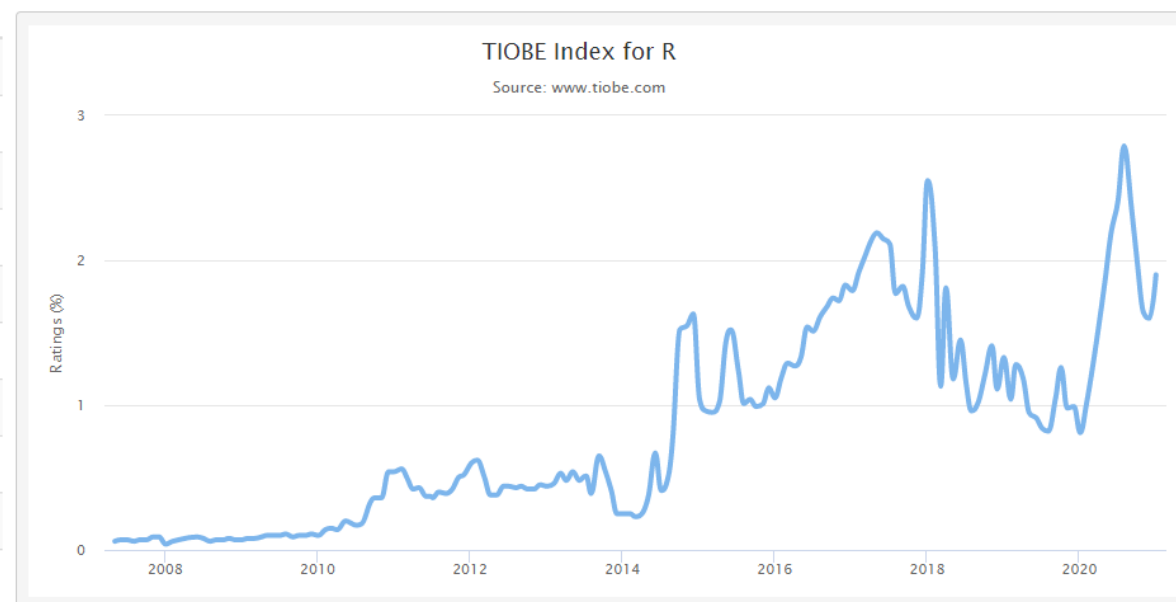


Some information about R:

▲ Highest Position (since 2007): #8 in Aug 2020

▼ Lowest Position (since 2007): #73 in Dec 2008

Jan 2021	Jan 2020	Change	Programming Language	Ratings	Change
1	2	▲	C	17.38%	+1.61%
2	1	▼	Java	11.96%	-4.93%
3	3		Python	11.72%	+2.01%
4	4		C++	7.56%	+1.99%
5	5		C#	3.95%	-1.40%
6	6		Visual Basic	3.84%	-1.44%
7	7		JavaScript	2.20%	-0.25%
8	8		PHP	1.99%	-0.41%
9	18	▲	R	1.90%	+1.10%
10	23	▲	Groovy	1.84%	+1.23%





# Uso de R en el mundo



**15.000 desarrolladores responden en Twitter cuál es su lenguaje de programación favorito, cuál odian y cuál recomiendan**

 **Victoriano Izquierdo**   
@victorianoizquierdo

Unas 15K personas de todo el mundo han creado un meme en Twitter de cuáles era sus lenguajes de programación favoritos, odiados o recomendados para aprender a programar... Los he recopilado y analizado con [@graphext](#) y aquí un hilo con los principales insights que encontré 🧵



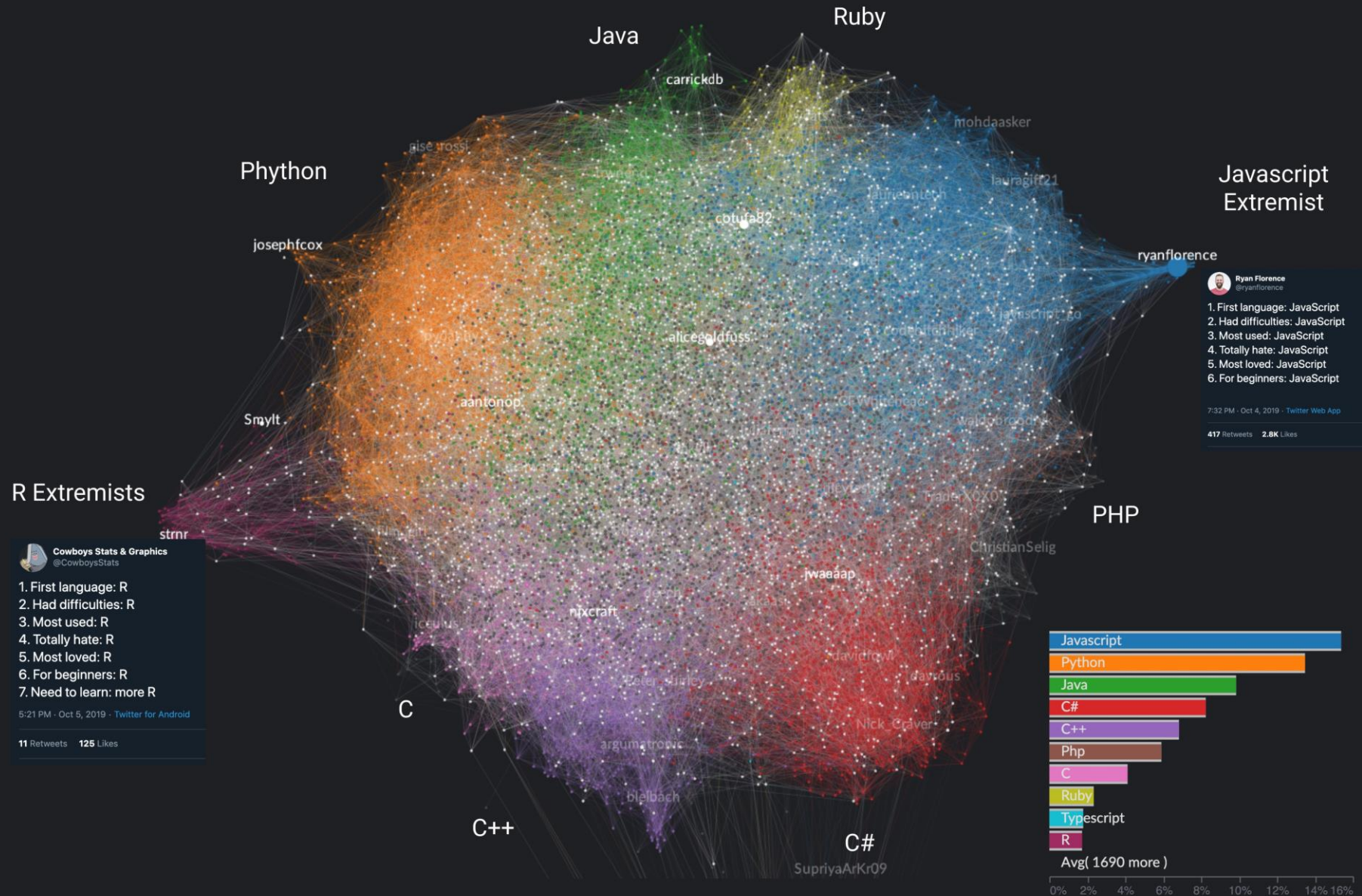
The four network graphs show the relationships between programming languages based on user preferences. The top-left graph, titled 'WHAT ARE THE MOST USED PROGRAMMING LANGUAGES', shows a dense cluster of languages including Java, Python, JavaScript, and Ruby. The top-right graph, 'WHAT ARE THE MOST LOVED PROGRAMMING LANGUAGES', shows a similar cluster but with a stronger emphasis on Python and JavaScript. The bottom-left graph, 'WHAT ARE THE MOST HATED PROGRAMMING LANGUAGES', shows a more dispersed network with a notable cluster for Java, indicating it is widely disliked. The bottom-right graph, 'WHAT IS THE BEST PROGRAMMING LANGUAGE FOR BEGINNERS', shows Python as the dominant language, with a clear 'Python wins by far!' annotation.

9:37 a. m. · 8 oct. 2019

👍 2 mil    💬 1,3 mil personas están twitteando sobre esto



# WHAT ARE THE MOST USED PROGRAMMING LANGUAGES



# Uso de R en el mundo

Top Analytics, Data Science, Machine Learning Software

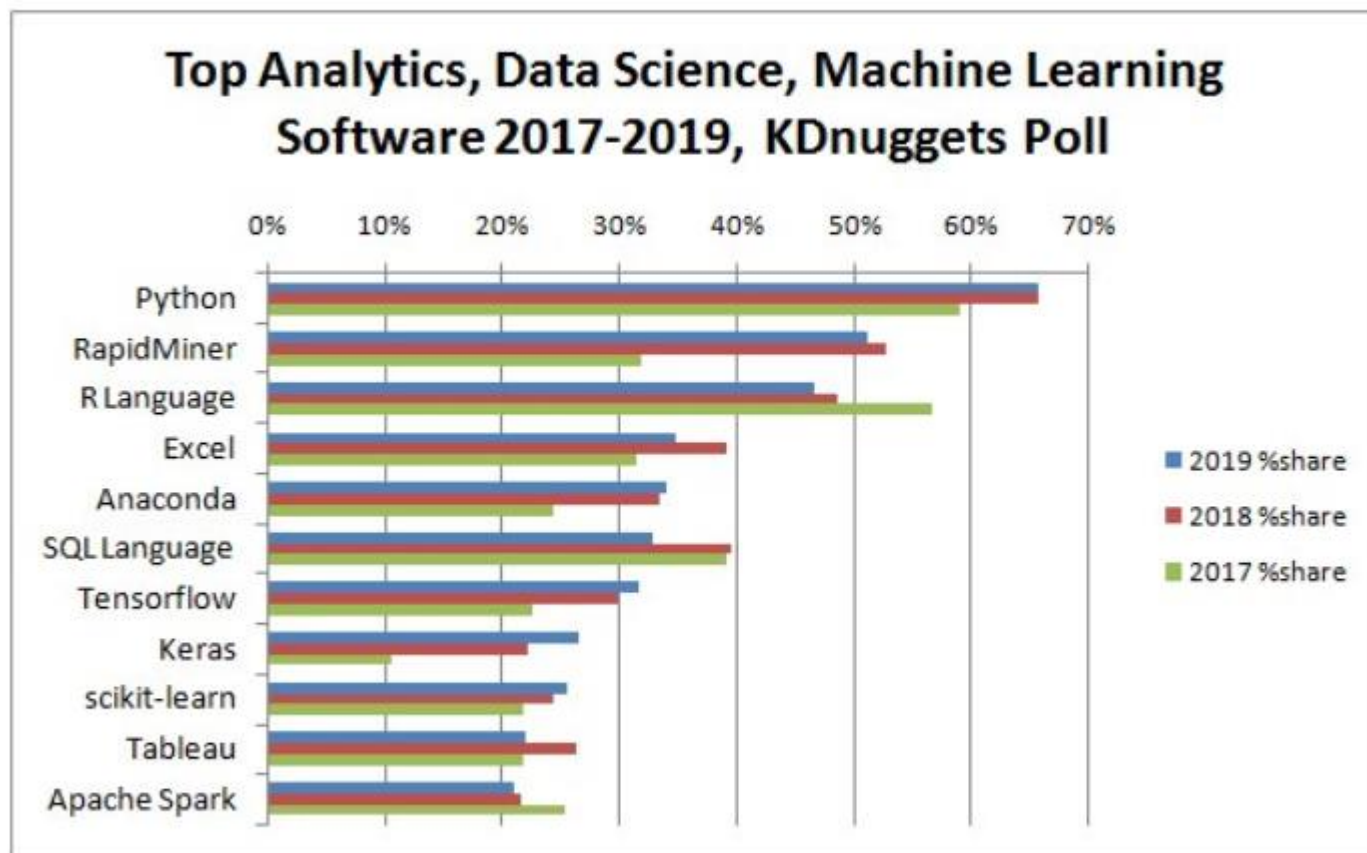


Fig 1: KDnuggets Analytics/Data Science 2019 Software Poll: top tools in 2019, and their share in the 2017, 2018 polls

# Fuentes consultadas



- <https://www.sintetia.com/8-grandes-motivos-para-comenzar-a-utilizar-r/>
- <https://www.enaes.es/blog/lenguaje-r-la-herramienta-de-los-cientificos-de-datos#gref>
- <https://www.maximaformacion.es/blog-dat/esta-claro-tenes-que-aprender-r-ya/>
- <https://www.kdnuggets.com/2015/10/data-science-programming-python-vs-r.html>
- <https://www.tiobe.com/tiobe-index/>
- <https://www.genbeta.com/desarrollo/15-000-desarrolladores-responden-twitter-cual-su-lenguaje-programacion-favorito-cual-odian-cual-recomiendan-1>