



MAX-PLANCK-INSTITUT  
FÜR DEMOGRAFISCHE  
FORSCHUNG

MAX PLANCK INSTITUTE  
FOR DEMOGRAPHIC  
RESEARCH

# Introducción a la Simulación a Partir de Modelos Basados en Agentes

Día 1

---

Daniel Ciganda

Marzo 2020

**Modelo:** Representación de un sistema o un proceso a través de sus **características principales** con el fin de manipularlo.

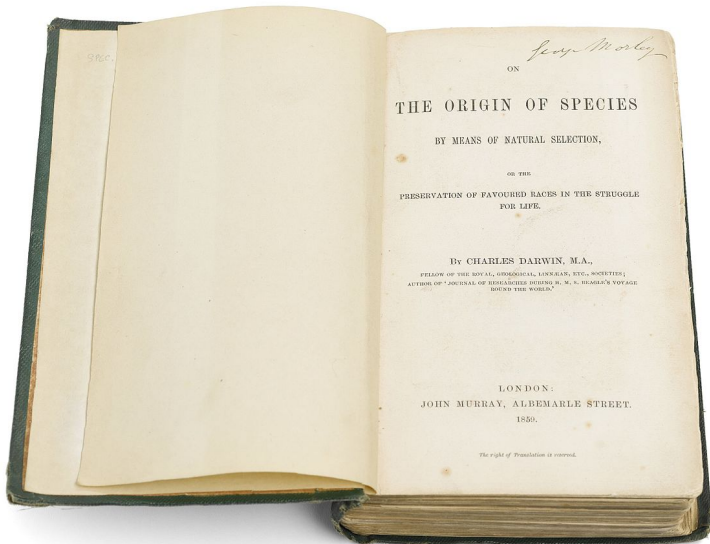
**Simulación:** Imitación del funcionamiento del sistema o proceso representado.

- Control de las **condiciones experimentales** en un ambiente de bajo riesgo.
- Construcción de escenarios contrafácticos.
- Observar el comportamiento del sistema **en un período de tiempo extenso** en un tiempo real acotado.

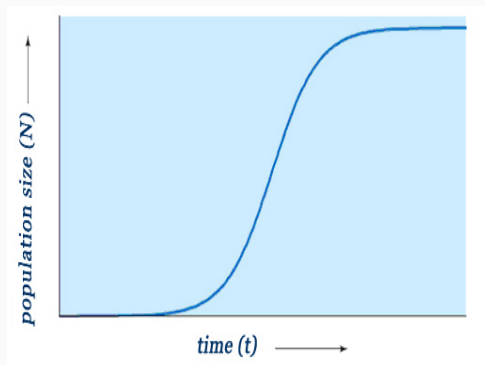
## **Tipos de Modelos**

---





$$\frac{dN}{dt} = rN \left( 1 - \frac{N}{K} \right)$$



# Modelos Computacionales

## Function

### Read Input Data

---

#### Initialization:

Define Initial Population Size; Define Attributes

---

#### Start Simulation:

#### **while** *time* < *end time* **do**

Choose Next Event  $\rightarrow$  *nEvent*

Update Clock, Ages and Waiting Times

**if** *nEvent* = *Complete Education* **then**

| Update Attributes and Waiting Times

**end**

...

**if** *nEvent* = *Death* **then**

| Update Attributes and Waiting Times

**end**

**if** *End of Current Year* **then**

| Compute and Store Aggregate Indicators

**end**

**end**

---

#### Result:

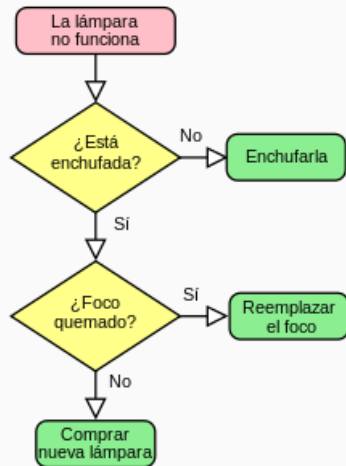
Collect Output in List for Analysis

**end**

**Definición:** Conjunto de instrucciones no-ambiguas, ordenadas y finitas que permiten, solucionar un problema, realizar un cómputo o llevar a cabo alguna otra actividad.

**Chazelle:** "algorithms as modelling devices".

Al igual que el hardware, los algoritmos son **tecnología**.





El conjunto de instrucciones contenido en el algoritmo se expresa en un determinado lenguaje de programación. La elección del lenguaje debería tener en cuenta:

- Flexibilidad, eficiencia.
- Velocidad
- Acceso a infraestructuras.
- Trabajo en equipo, comunicación dentro de la comunidad de referencia.
- Inversión

Ejemplos: Fortran, C++, Python, Julia, R, NetLogo, etc.

R es una implementación (dialecto) de S, creado por Ross Ihaka y Robert Gentleman en 1992 cuando trabajaban como estadísticos en la Universidad de Auckland.

De **CRAN**: "R is a language and environment for statistical computing and graphics."

- Un lenguaje de programación comprensivo, simple y efectivo que incluye condicionales, loops y funciones recursivas definidas por los usuarios
- Un espacio para almacenar y manejar datos eficientemente
- Un extenso, coherente e integrado conjunto de herramientas intermedias para el análisis de datos
- Herramientas gráficas para el análisis de datos

**Paquetes:** tidy, ggplots2, etc.

**RStudio** es el editor más popular (IDE - Integrated Development Environment).

# Popularidad de Lenguajes para “Ciencia de Datos”

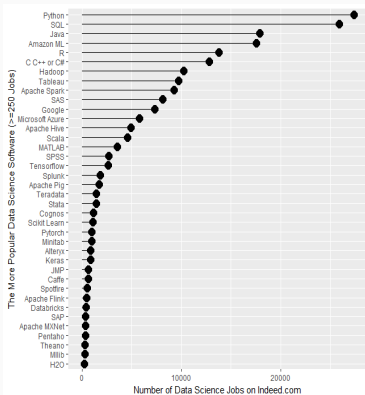


Figure 1: Trabajos ofrecidos 2018

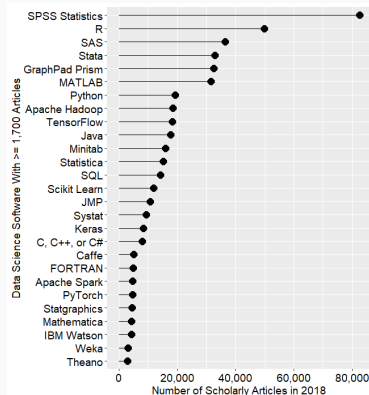
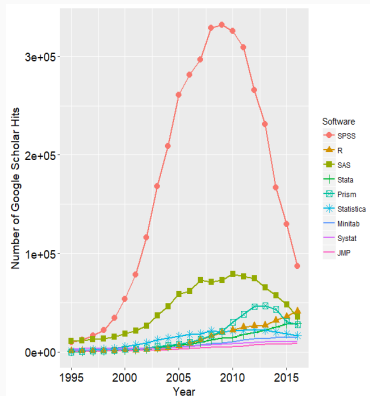


Figure 2: Artículos Científicos

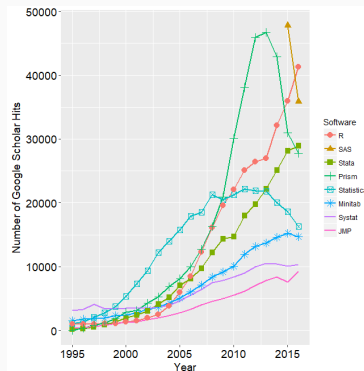
Fuente: Robert A. Muenchen, *The Popularity of Data Science Software*.

<http://r4stats.com/articles/popularity/>

# Popularidad de Lenguajes para “Ciencia de Datos”



**Figure 3: Artículos Científicos**



**Figure 4: Arts. Científicos (sin SPSS)**

Fuente: Robert A. Muenchen, *The Popularity of Data Science Software*.

<http://r4stats.com/articles/popularity/>

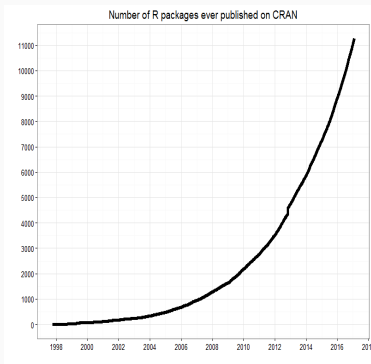


Figure 5: Crecimiento Paquetes

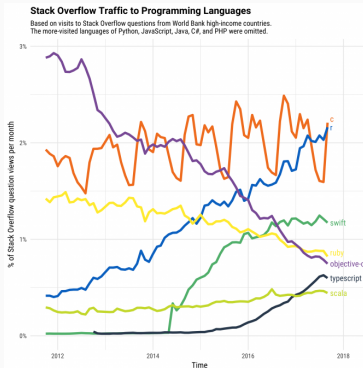


Figure 6: stackoverflow.com

Fuente: <https://blog.revolutionanalytics.com/>

## **Ventajas**

- Combina funcionalidad para el análisis de datos con un lenguaje de programación completo.
- Comunidad activa y en crecimiento
- Hegemónico en varias disciplinas científicas
- Gratuito y abierto

## **Desventajas**

- Dificultad para manejar grandes bases de datos
- Velocidad

- Explorar RStudio.
- Comenzar el aprendizaje de R en R con `swirl`

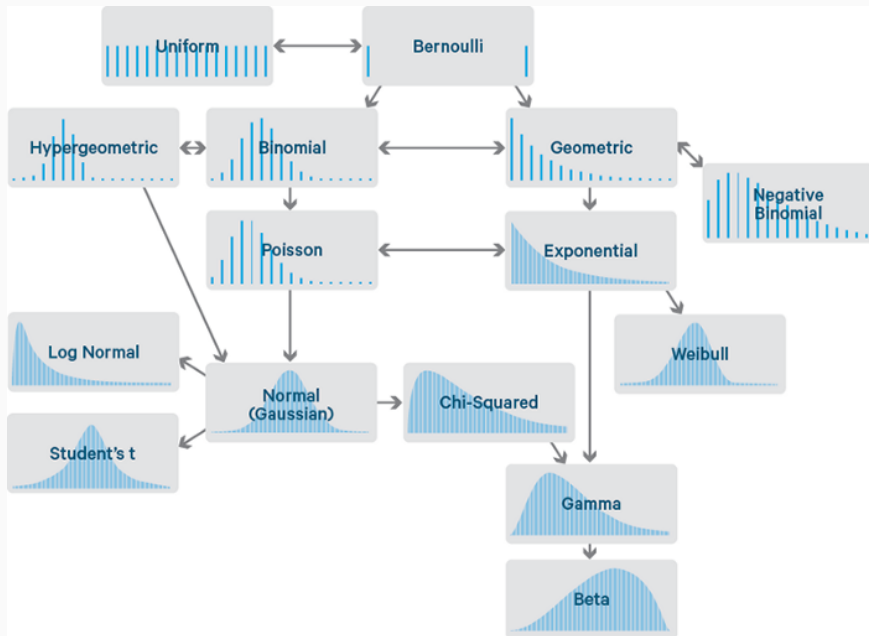
**Modelo Probabilístico:** El resultado no está completamente determinado por las condiciones iniciales, los parámetros y las especificación del modelo.

**Variable aleatoria:** Su valor es el resultado de un evento aleatorio. No conocemos el valor exacto que va a tomar en cada realización, pero podemos conocer la probabilidad de cada uno de los posibles valores. Pueden ser continuas o discretas.

**Distribución de probabilidad:** Probabilidad de ocurrencia de cada valor de una variable aleatoria.



# Distribuciones



- Ejercicios con distribuciones en R: generación de números aleatorios, muestreo.