

Repaso de R & Python

# Sesión 1

Fundamentos de Programación, Control de Flujo y Buenas Prácticas

Ricard Sierra Calls

# Índice

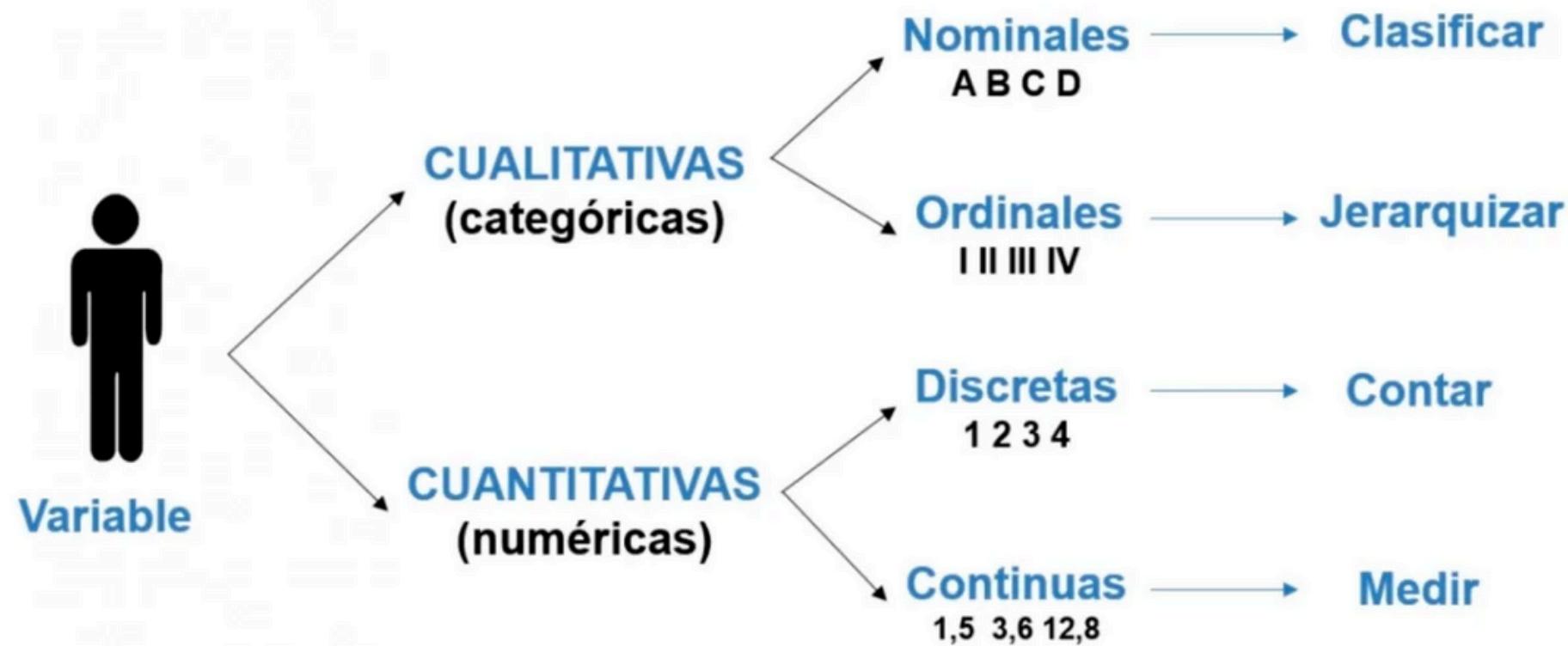
- 1 Introducción a R y Python
- 2 Tipos de datos
- 3 Control del Flujo
- 4 Buenas Prácticas
- 5 Archivos de trabajo
- 6 ¿Puedo usar ChatGPT?
- 7 ACTIVIDADES



En el 70-80% de los casos, **se pueden usar ambos**

- Manipulación de datos y visualización.
  - Investigación académica
  - Estadística en general
- 
- Lenguaje de propósito general, versátil
  - Proyectos de ML y DL
  - Más aplicable para aplicaciones externas

Para saber cuál usar, fijaros en la **sintaxis**, el manejo de **paquetes/librerías**, y la **orientación** de cada lenguaje

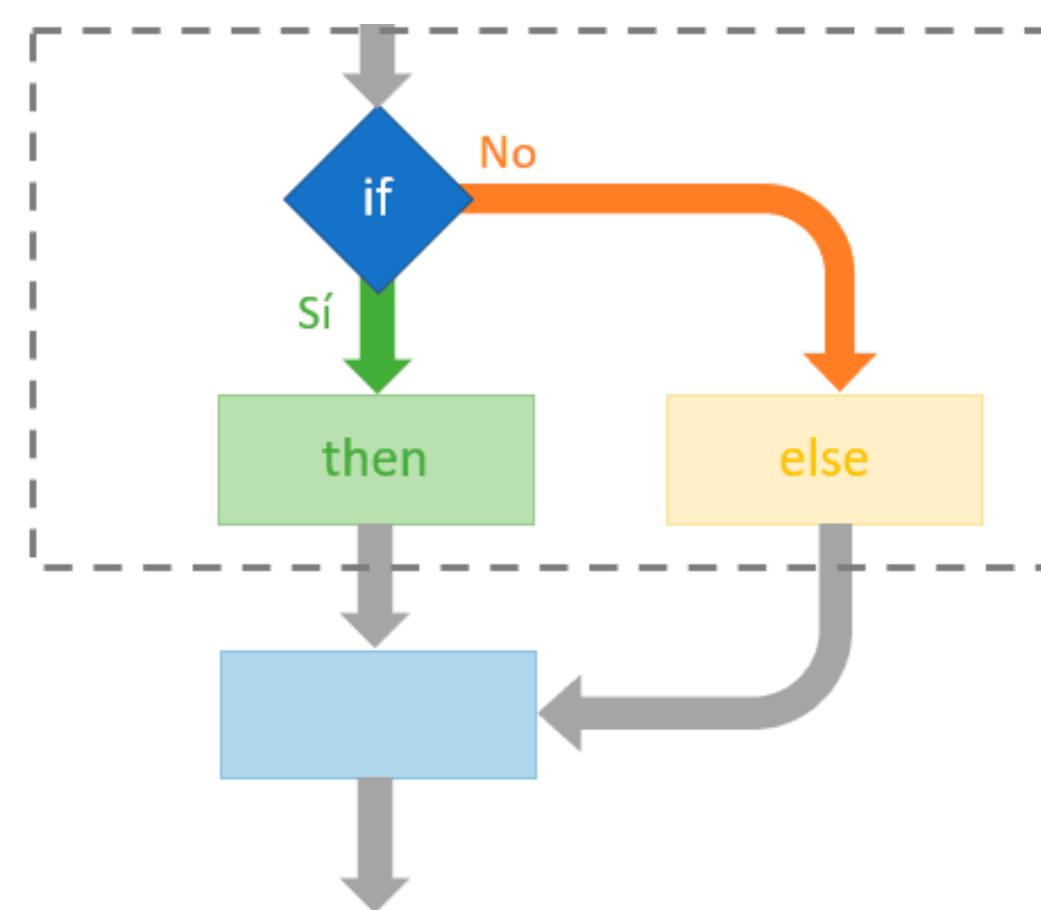


Numérico	Categóricos	Booleanos	Fechas
Altura Peso Clasificación	Género Nacionalidad Equipo	TRUE/FALSE Sano/Enfermo Aprobado/Suspendido	08/11/2024 11/08/2024 feb. 2024

# Control de Flujo

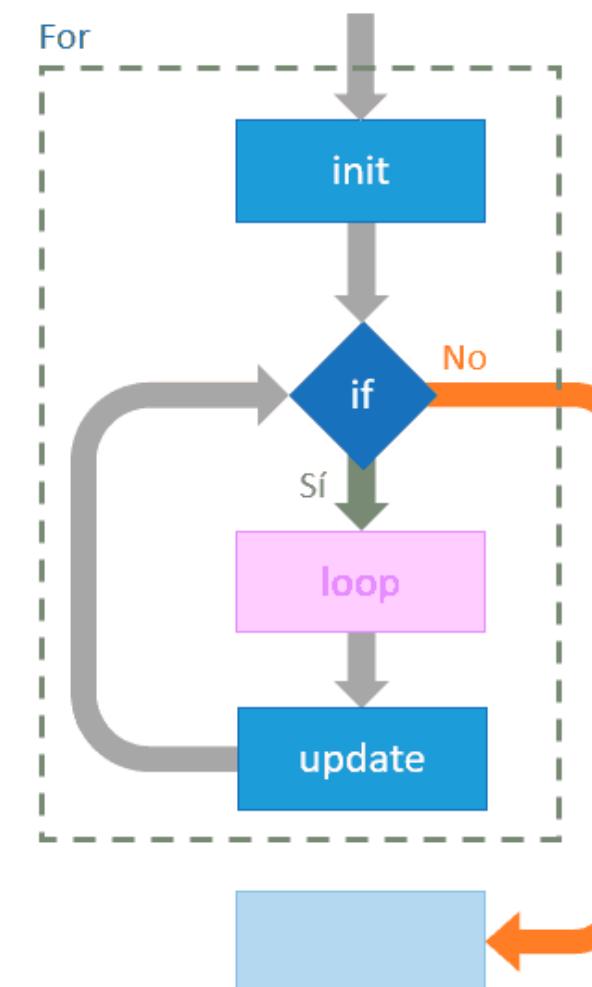
## Condicionales

(If, else, elif)



## Bucles

(for, while)



# Buenas Prácticas

- Estructura y organización de código
- Indentación (espacios a la izquierda)
- Nombres de variables y funciones descriptivos
- Comentarios y documentación de código



# Archivos de trabajo



## Cuadernos (.ipynb)

- Aprendizaje, exploración y pruebas rápidas.
- Código + texto + gráficos en una misma vista.
- Ejecución por celdas: flexible pero menos trazable.
- Prototipar y explicar ideas.



## Archivos de código (.py, .R)

- Proyectos "de verdad": módulos, paquetes, apps.
- Ejecutables, testeables y fáciles de versionar.
- Estructura clara: funciones, clases y tests.
- Usados para producción y colaboración.

### Flujo práctico:

Empieza en cuaderno → Consolida en módulos → Testeos de funcionamiento  
→ Ejecuta desde scripts

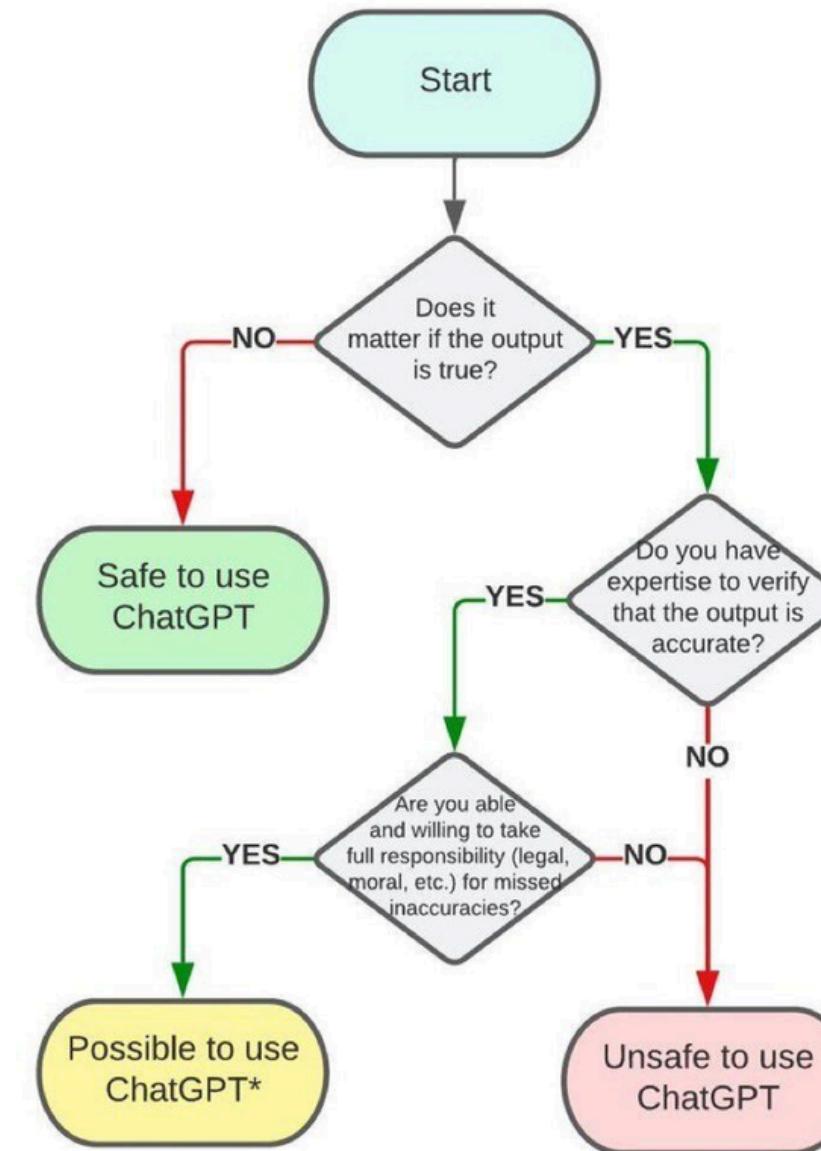
# ¿Puedo usar ChatGPT?

*Deberíamos tener miedo, no a las máquinas inteligentes, sino a las máquinas que toman decisiones para las que carecen de la inteligencia necesaria. Tengo mucho más miedo de la estupidez de las máquinas que de su inteligencia*

Sendhil Mullainathan

## Is it safe to use ChatGPT for your task?

Aleksandr Tiulkanov | January 19, 2023



\* but be sure to verify each output word and sentence for accuracy and common sense



# Material Adicional

- ([link](#)) **UCI ML Repository**: creado en 1987 por David Aha, +600 datasets donde filtrar por tipo de problema, tipo de datos, de características, etc
- ([link](#)) **OMS**: de las Naciones Unidas, datos de la salud de todo el mundo
- ([link](#)) **Data Is Plural**: newsletter semanal con todo tipo de datos actuales de diversas materias
- ([link](#)) **Data .gov**: base de datos publica del gobierno de EEUU donde podemos encontrar +300.000 conjuntos de datos
- ([link](#)) **NASA**: desde 1994, tanto de la tierra como de otros planeta
- ([link1](#)) ([link2](#)) ([link3](#)) **Google Dataset Search**
- ([link](#)) **DataHub Collections**
- ([link](#)) **KAGGLE**
- ([link](#)) **CERN Open Data**
- ([link](#)) GitHub - **Awesome Public Datasets**
- ([link](#)) DataSetsVerse: Metapquete que reúne una colección seleccionada de paquetes R que contienen conjuntos de datos específicos de cada dominio. Incluye datos de *series temporales, métricas educativas, registros de delitos, conjuntos de datos médicos y datos de investigación oncológica*.

# ACTIVIDADES

