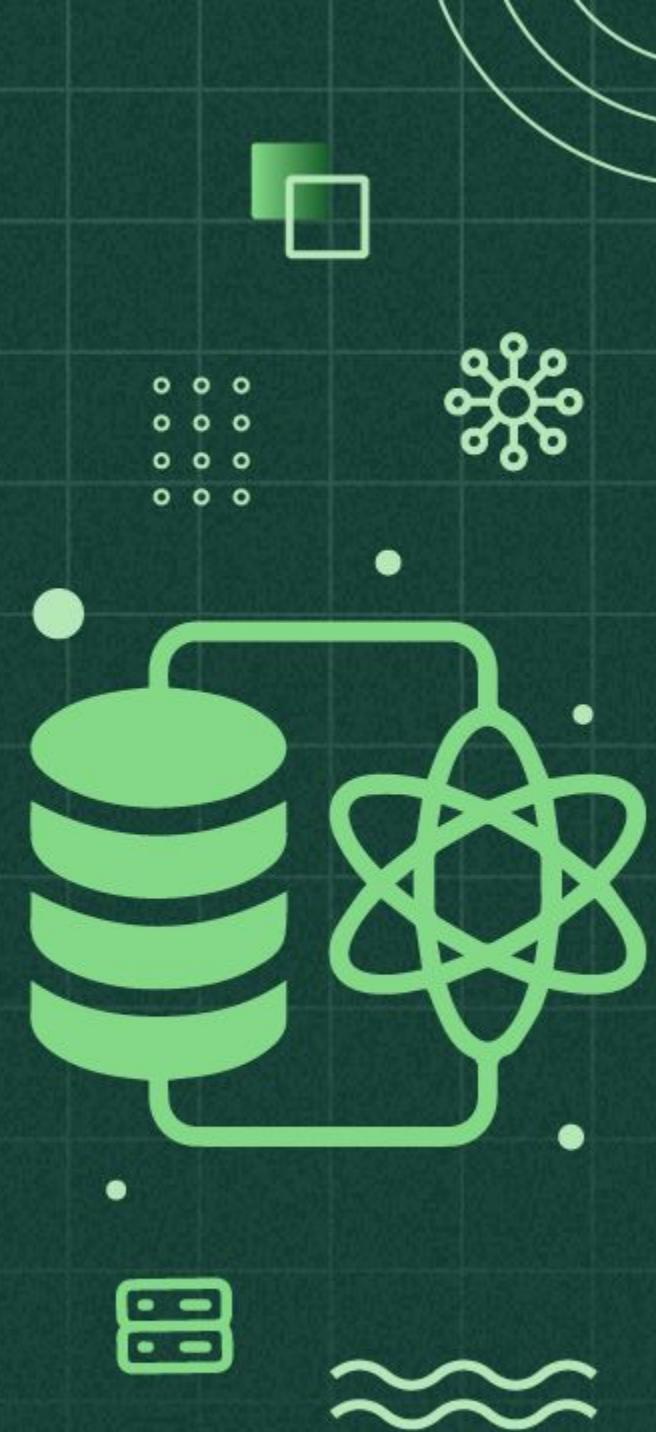
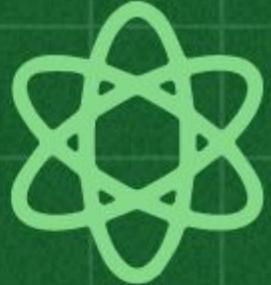


Cómo y por qué  
**Aprender Data  
Science e  
Inteligencia  
Artificial**

Miguel Torres y Romina Huamán





# Oportunidades laborales en Data Science e Inteligencia Artificial



LinkedIn search interface showing results for "data scientist".

Search bar: data scientist

Location: Todo el mundo

Filters: Empleos, Fecha de publicación, Nivel de experiencia

Results:

- Data scientist en En todo el ...** Alerta de empleo desactivada
- Data Scientist, Talent Acquisition** Atlassian Americas, Apurímac, Perú (En remoto) En busca de personal Promocionado
- Data Scientist** Toptal Perú (En remoto) 1 antiguo empleado trabaja aquí Hace 1 hora
- Data Scientist** TASA (Empresa del Grupo BRECA) Lima, Perú (Híbrido) Tienes una insignia de aptitud preferida Promocionado • 15 solicitudes
- DATA SCIENTIST SR.** Rimac Seguros y Reaseguros Lima, Perú (En remoto) 3 contactos trabajan aquí Promocionado • 9 solicitudes • Solicitud sencilla

LinkedIn search interface showing results for "Inteligencia artificial".

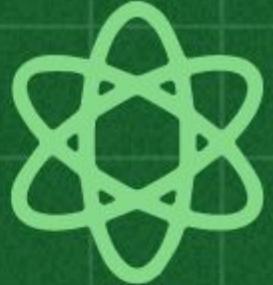
Search bar: Inteligencia artificial

Location: Todo el mundo

Filters: Empleos, Fecha de publicación, Nivel de experiencia

Results:

- Inteligencia artificial en En to...** Alerta de empleo desactivada
- Machine Learning Researcher** Recursive Tokio, Tokio, Japón (En remoto) Hace 17 minutos • Solicitud sencilla
- Data & Analytics Consultant** Sia Partners Ámsterdam, Holanda Septentrional, Países Bajos (Presencial) En busca de personal Hace 7 horas
- Data Analytics Consultant** Capgemini Engineering Bruselas, Región de Bruselas, Bélgica (Presencial) 1 antiguo empleado trabaja aquí Hace 2 días • 12 solicitudes
- Analyste de données, Omnia IA (CDC) 1** Deloitte Montréal-Ouest, QC (Presencial) 6 antiguos empleados trabajan aquí



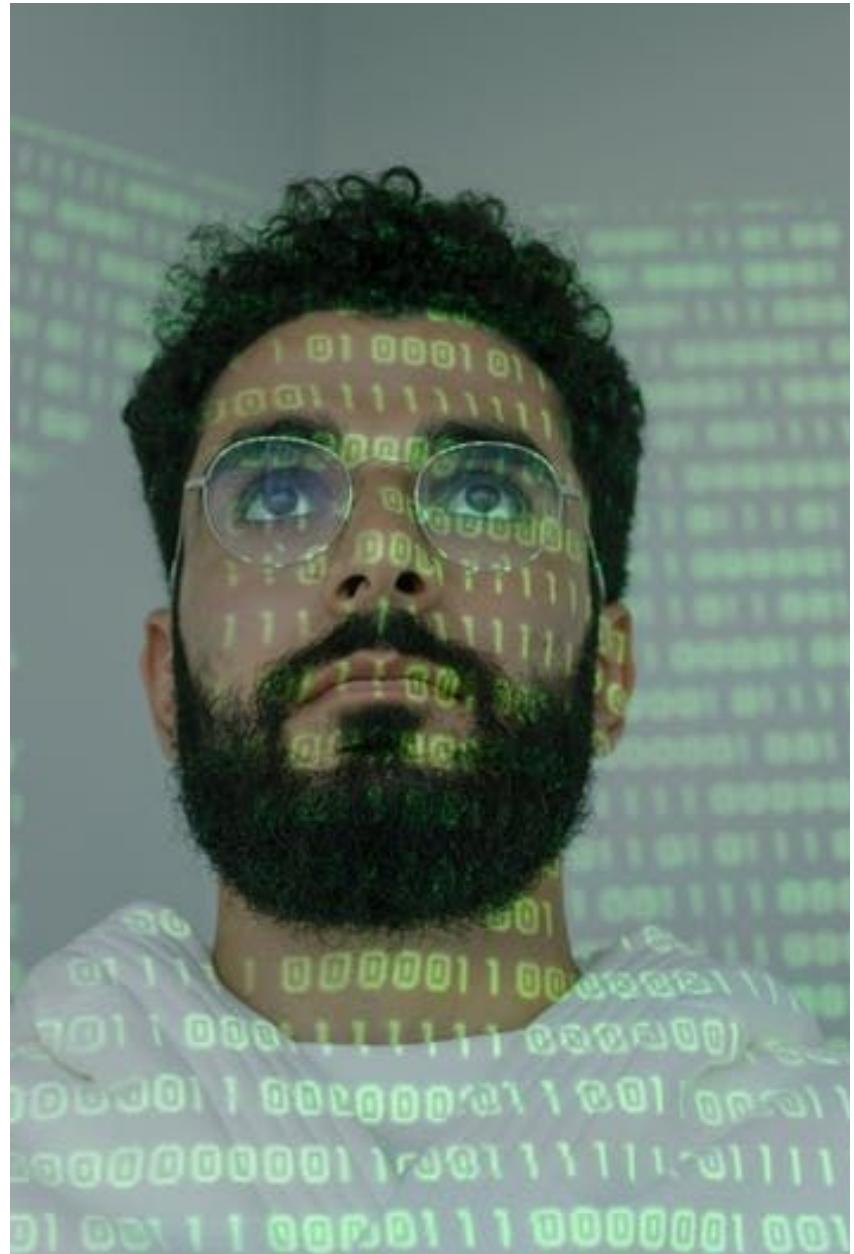
# ¿Qué es data science?





# ¿Qué es data science?

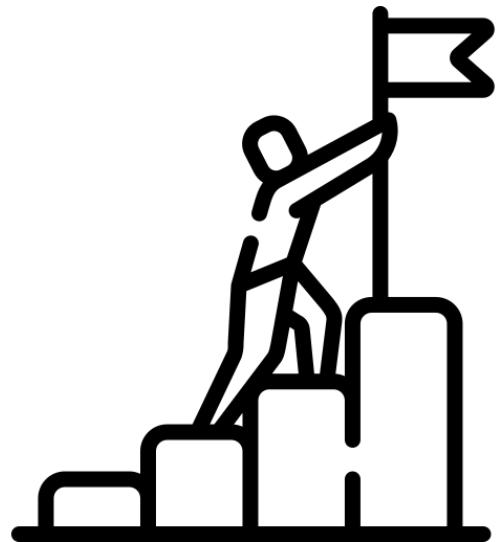
Data science o ciencia de datos **es el proceso** de descubrir información valiosa de los datos.

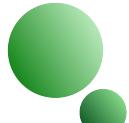




# ¿Cuál es su finalidad?

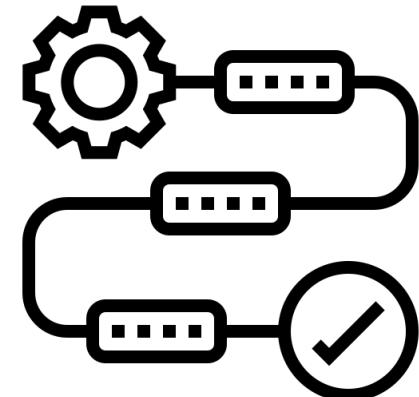
- **Tomar decisiones y crear estrategias de negocio.**
- **Crear productos de software más inteligentes y funcionales.**





# ¿De qué trata este proceso?

- Obtención de los datos.
- Transformar y limpiar los datos.
- Explorar, analizar y visualizar datos.
- Usar modelos de machine learning\*.
- Integrar datos e IA a productos de software.

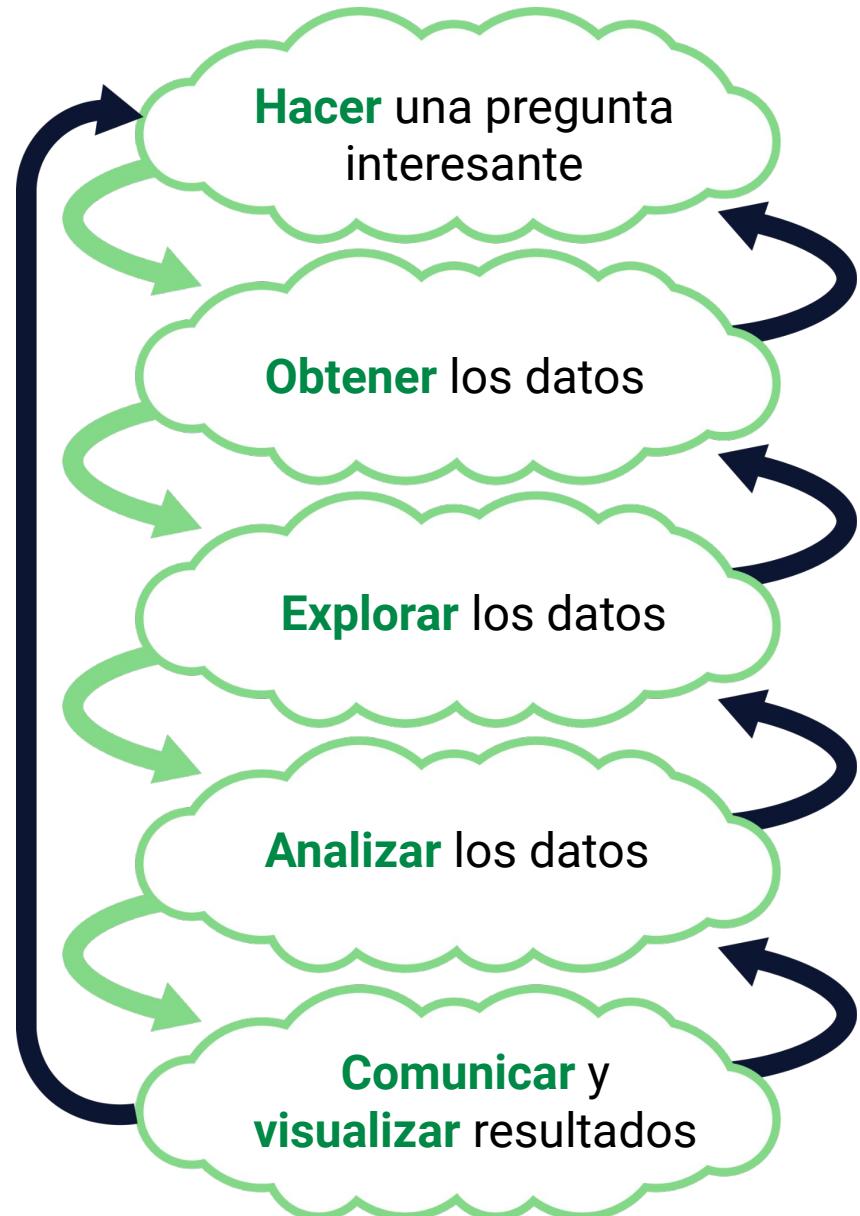


\*Inteligencia artificial. No siempre es necesario usarla.



# Proceso de la ciencia de datos

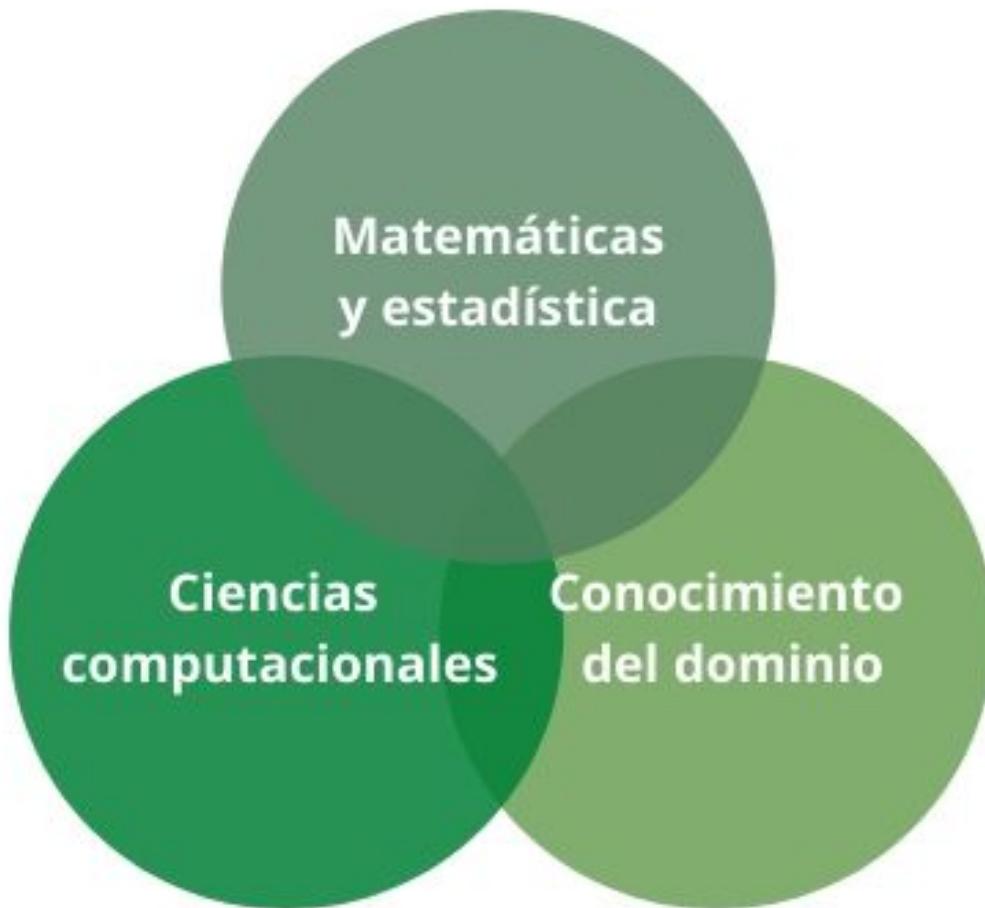
- El proceso entre proyecto a proyecto cambia poco.
- Es el proceso del método científico llevado al uso de datos.





# Proceso de la ciencia de datos





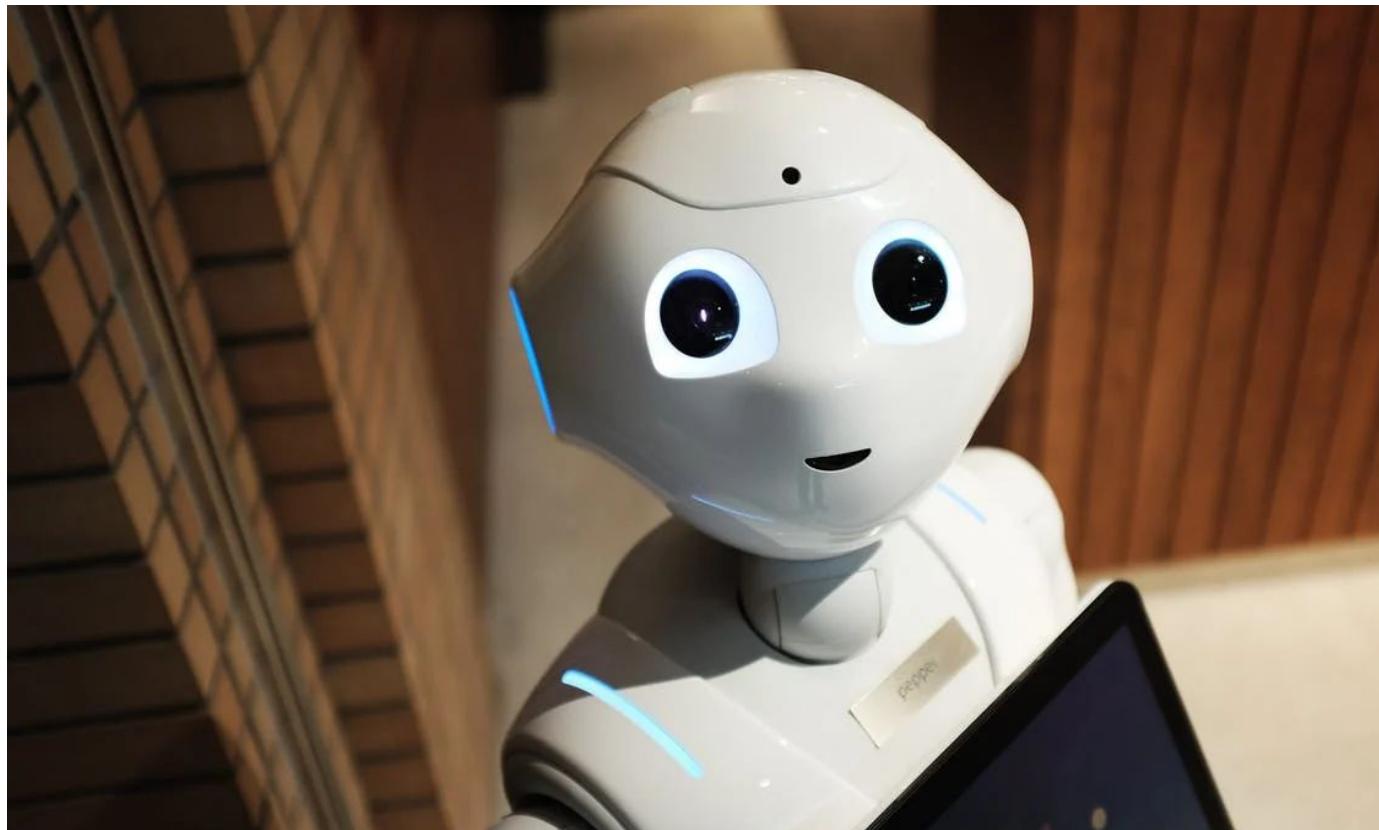


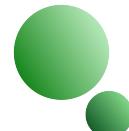
# ¿Qué es inteligencia artificial?



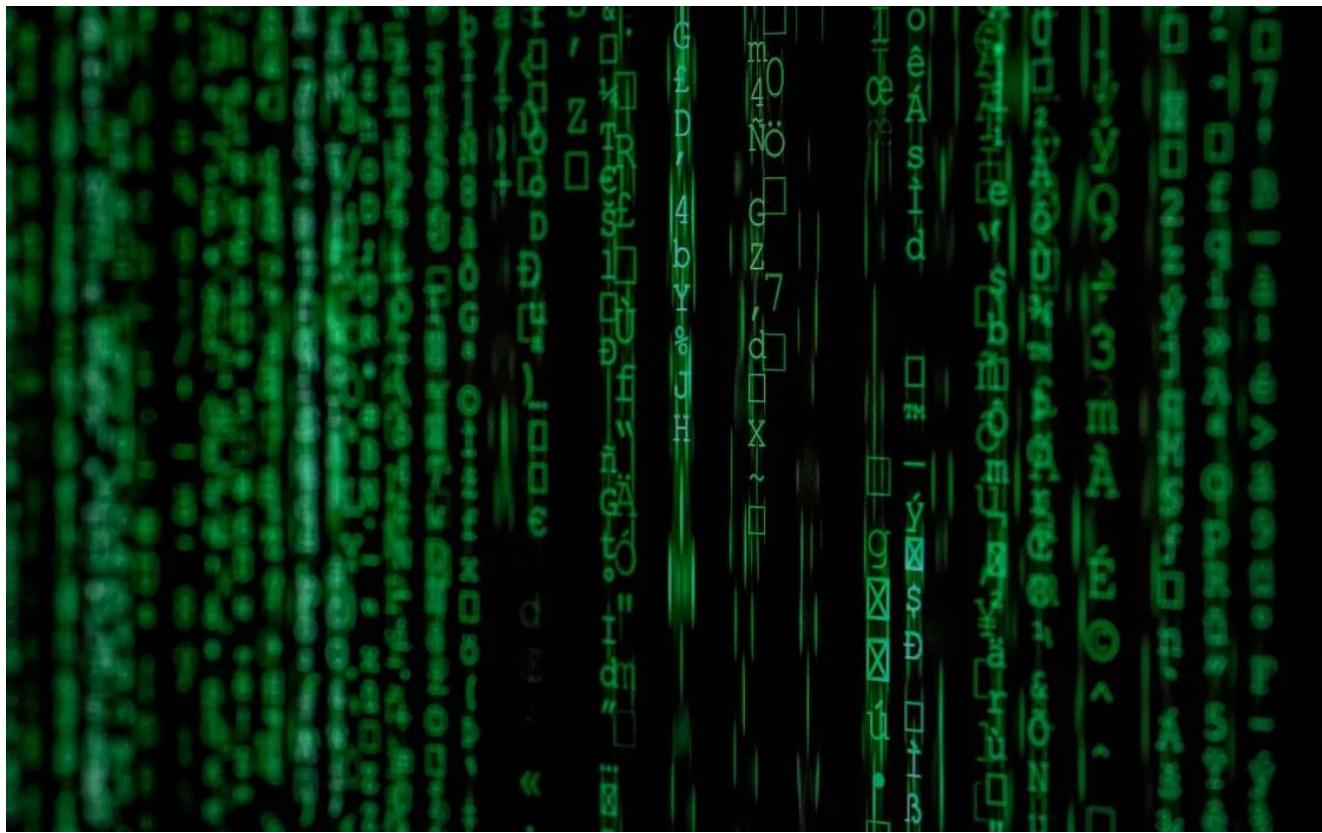


# Inteligencia artificial





# Inteligencia artificial





# Inteligencia artificial

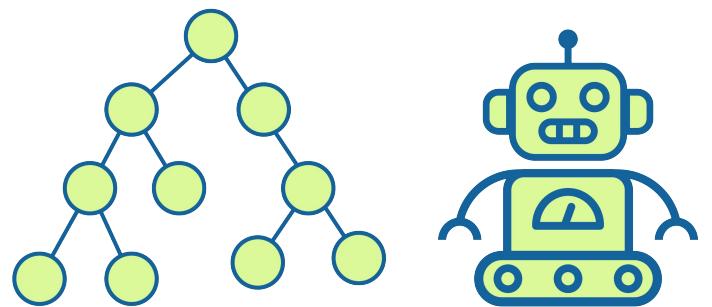
- Inteligencia artificial y data science no son lo mismo.



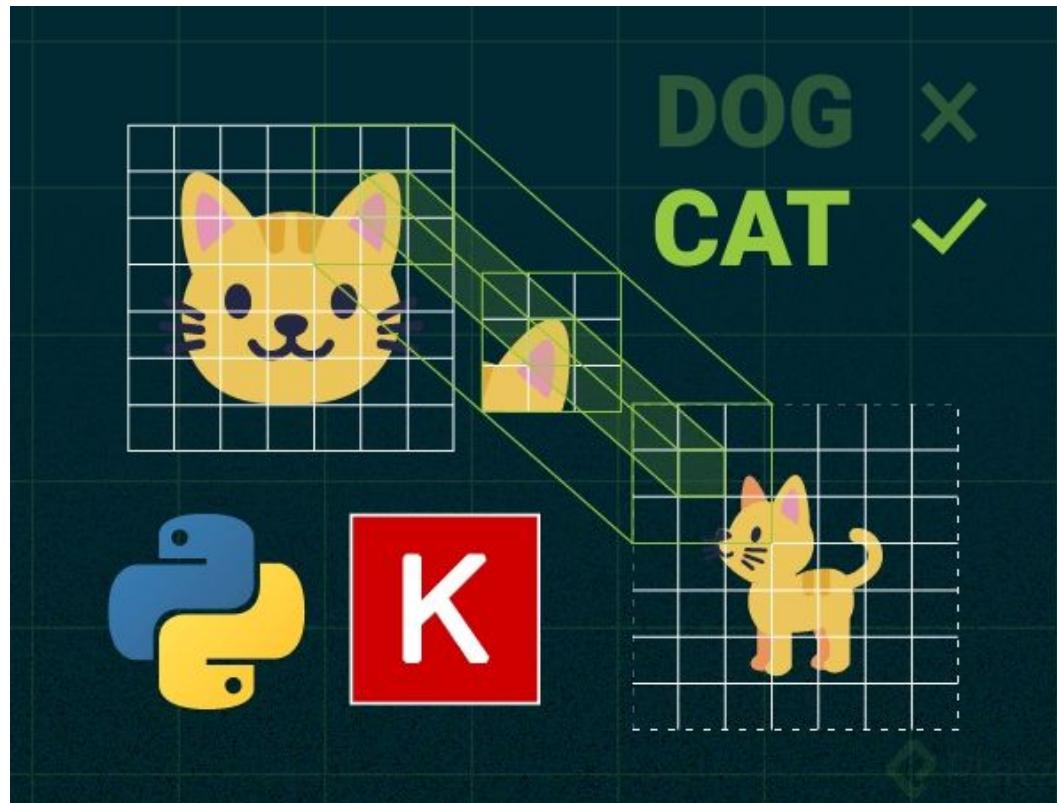


# ¿Pero qué es en realidad?

- Algoritmos para emular nuestra inteligencia natural.
- Reconocer patrones en grandes cantidades de datos.



# • ¿Pero qué es en realidad?



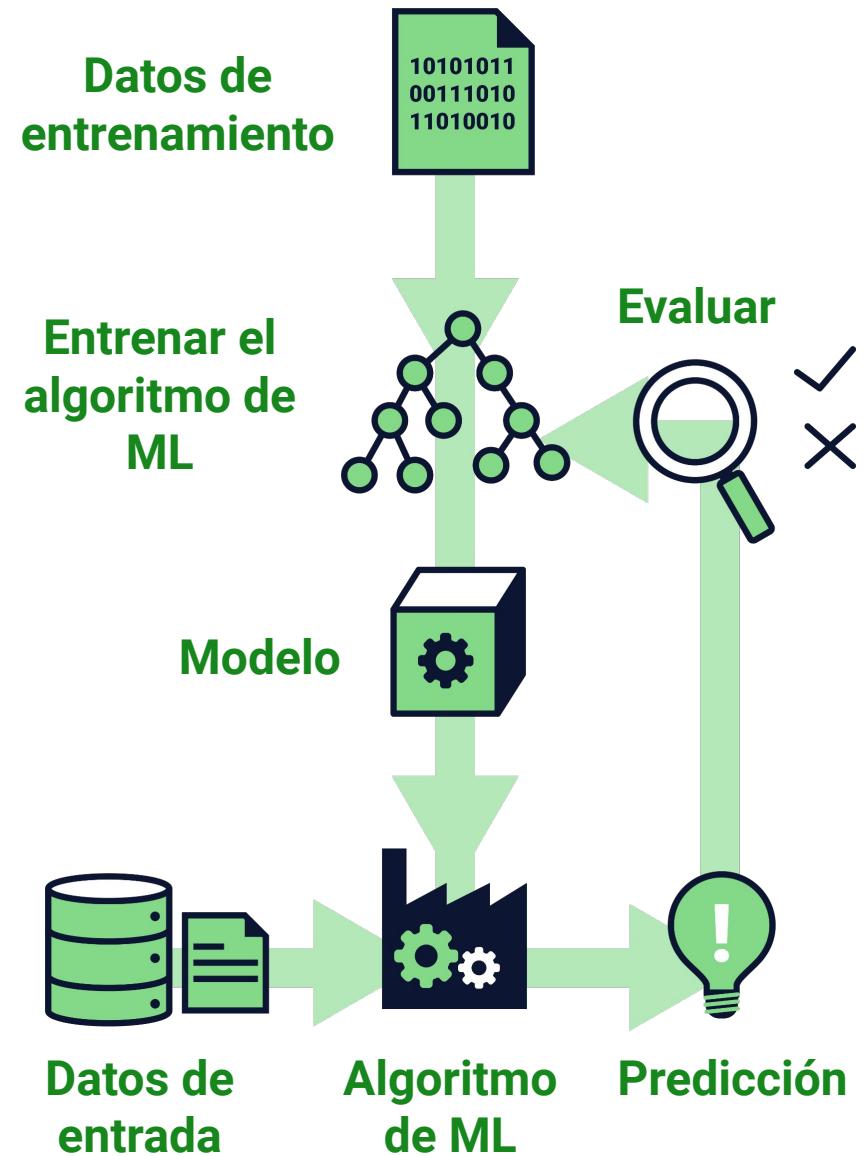


# IA en la actualidad



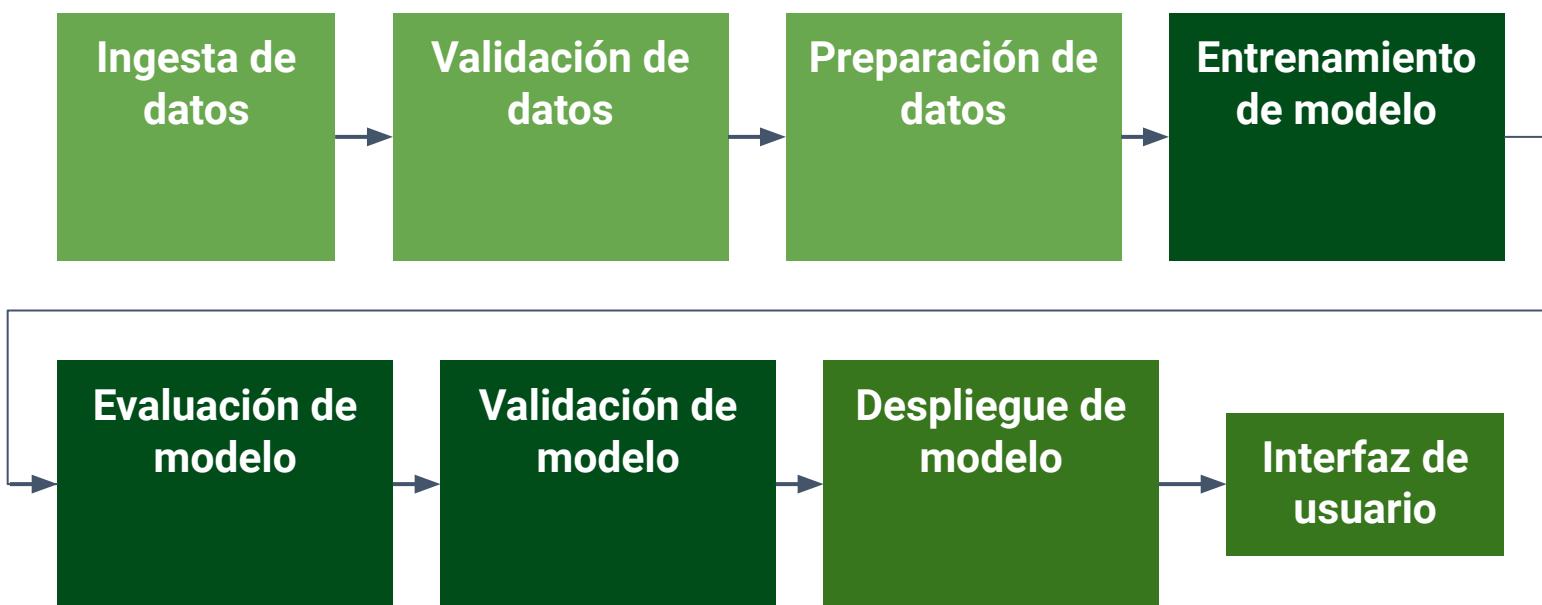


# Proceso de machine learning

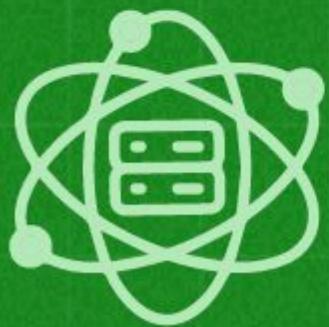




# Proceso de ciencia de datos: integrar ML a producto



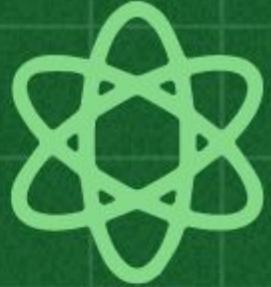
**¿Cuál es su diferencia  
con data science?**





# Data science + IA

Data science	Inteligencia artificial
Proceso para analizar datos y generar predicciones para toma de decisiones y crear productos con datos.	Algoritmos para predecir eventos futuros que emulan cognición.
En el proceso de data science utilizamos inteligencia artificial como una de sus herramientas.	



# ¿Qué es Big Data?





# Big Data

- Grandes volúmenes de datos muy variados y muy veloces.
- Resulta complicado procesarlos con métodos tradicionales.





# 5V de Big Data

5V de Big Data	Definición
<b>Volumen</b>	El almacenamiento de la masiva cantidad de datos que pueden ser recolectados de múltiples fuentes.
<b>Velocidad</b>	Los datos se generan en tiempo real gracias a las interacciones con las fuentes mencionadas, por lo que deben ser procesados con la misma velocidad.
<b>Variedad</b>	Todo tipo de datos, ya sea estructurados o no estructurados. Podrían ser tablas, texto, imágenes, videos, audio, bases de datos, etc.
<b>Veracidad</b>	Es la calidad y confiabilidad de los datos.
<b>Valor</b>	Los datos deben poder proporcionar un valor o beneficio a la empresa que los está usando.

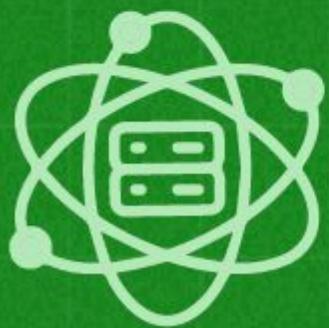


# Procesamiento de Big Data

- Se procesa al dividirla en partes pequeñas en varias máquinas.
- Tecnologías como Spark, Hadoop y servicios de cómputo en la nube.



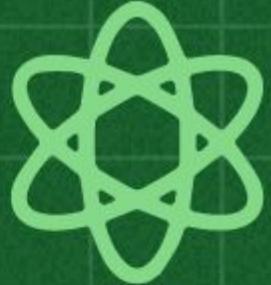
# ¿Cómo se une con data science e IA?





# Data science + IA + Big Data

Data science	Inteligencia artificial	Big Data
Proceso para analizar datos y generar predicciones para toma de decisiones y crear productos con datos.	Algoritmos para predecir eventos futuros que emulan cognición.	Grandes volúmenes de datos muy variados, muy veloces y difíciles de procesar.
<ul style="list-style-type: none"><li>• Big Data es materia prima que podemos usar en data science para hacer análisis más exhaustivos.</li><li>• Incluso podemos utilizar machine learning en ese mismo proceso para perfeccionar y evaluar los algoritmos de inteligencia artificial que creemos.</li></ul>		



**¿Qué NO es data  
science y por qué  
aprenderla?**





# ¿Qué NO es data science?

- Magia.
- Inteligencia artificial.
- Tener solamente métricas de algún dato sin hallazgos de valor.



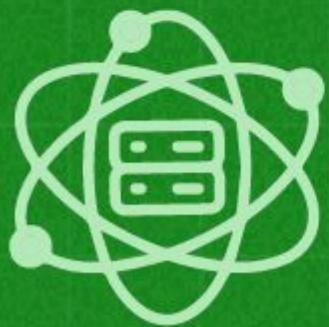


# ¿Qué NO es data science?

- Puras matemáticas.
- Trabajar solo con Big Data.



¿Por qué no aprender  
data science?



# ● ¿Por qué NO aprender data science?

- No quieres que el aprendizaje sea algo constante en tu vida.
- No disfrutas del “trabajo sucio”.



# ● ¿Por qué NO aprender data science?

- No te gusta comunicar y negociar lo que logras/encuentras.
- No te sientes motivado por ayudar a otras personas.



**CHALLENGE DENIED.**

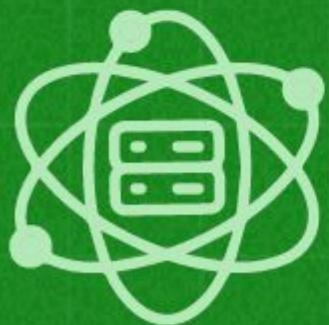
# • ¿Por qué NO aprender data science?

- No te gusta “hacer que las cosas pasen”.



**CHALLENGE DENIED.**

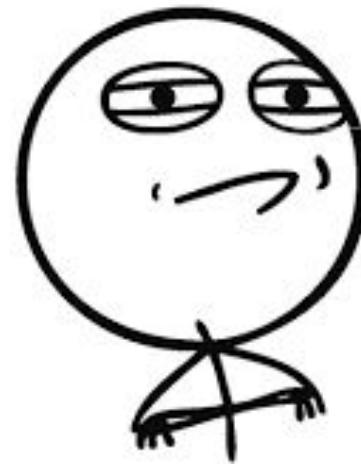
# ¿Por qué sí aprender data science?



# ● ¿Por qué aprender data science?

- Te emociona encontrar información de valor en datos/gráficas.
- Te interesa conocer cómo se aplican las matemáticas y algoritmos.

**CHALLENGE ACCEPTED**





# ¿Por qué aprender data science?

- Quieres trabajar en una organización data-driven y verla crecer.
- Quieres crear software/productos que cambien al mundo con datos e IA.

**CHALLENGE ACCEPTED**





# Áreas de aplicación de data science e IA

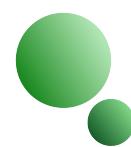




# Ramas de IA

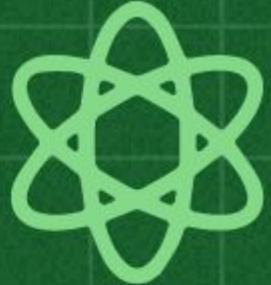
- Machine learning.
- Deep learning.
- RPA.
- Visión artificial.
- Procesamiento de lenguaje natural.
- Robótica





# Áreas de aplicación de data science

- Salud
- Procesos productivos
- Procesos comerciales
- Redes sociales



# Cómo funcionan los equipos de datos e IA



Roles en la industria

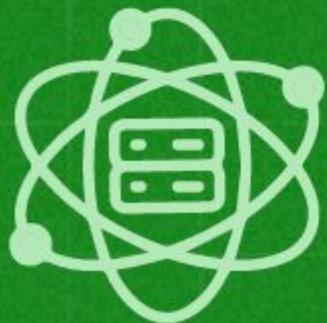


# Roles en la industria

- Data Scientist
- Data Analyst
- Data Engineer
- Machine Learning  
Engineer

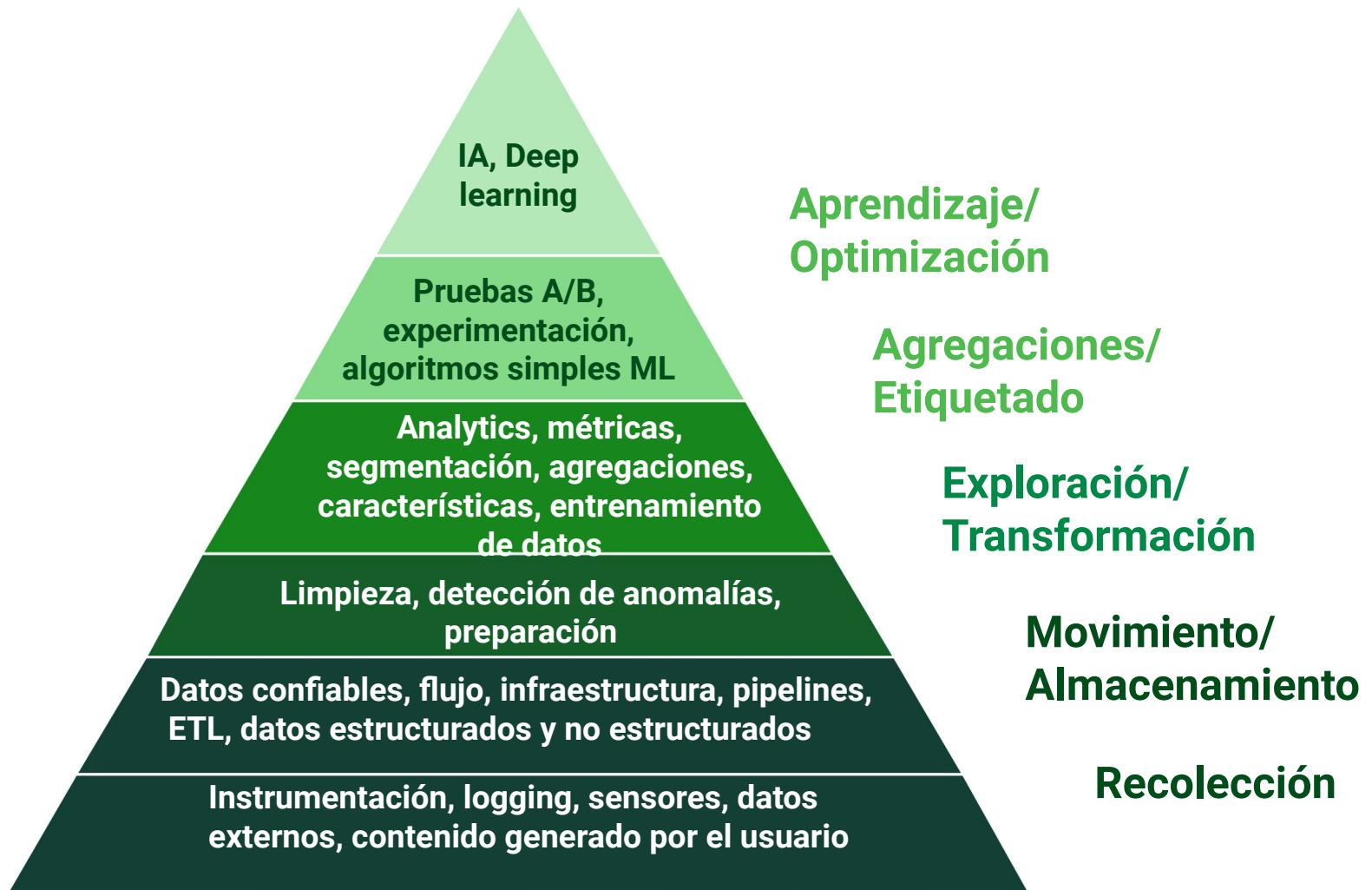


# Data science en las empresas y sus equipos

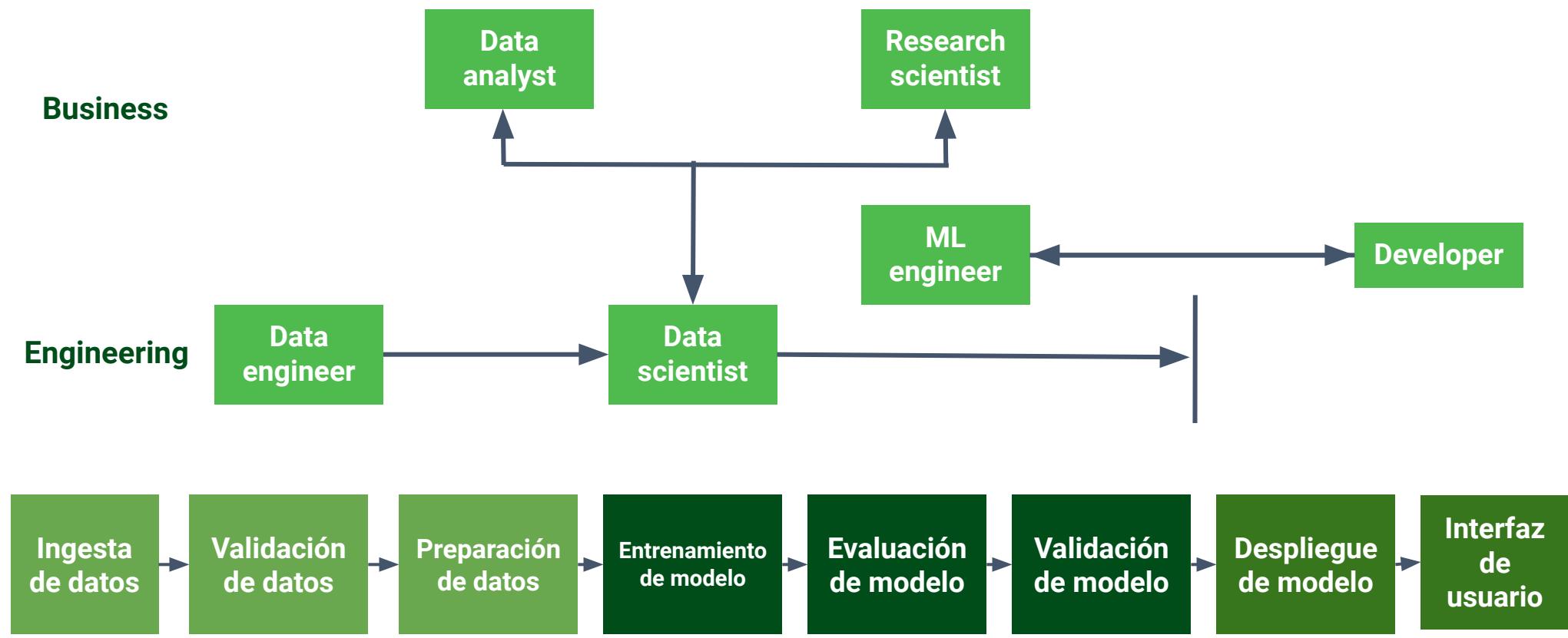




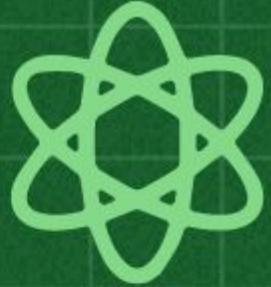
# La jerarquía de necesidades de data science



Referencia: 2. Data Science Hierarchy of needs (Monica Rogati – Hackernoon)



*Pasos y roles en el flujo de trabajo de data science  
(Design Patterns in Machine Learning).*

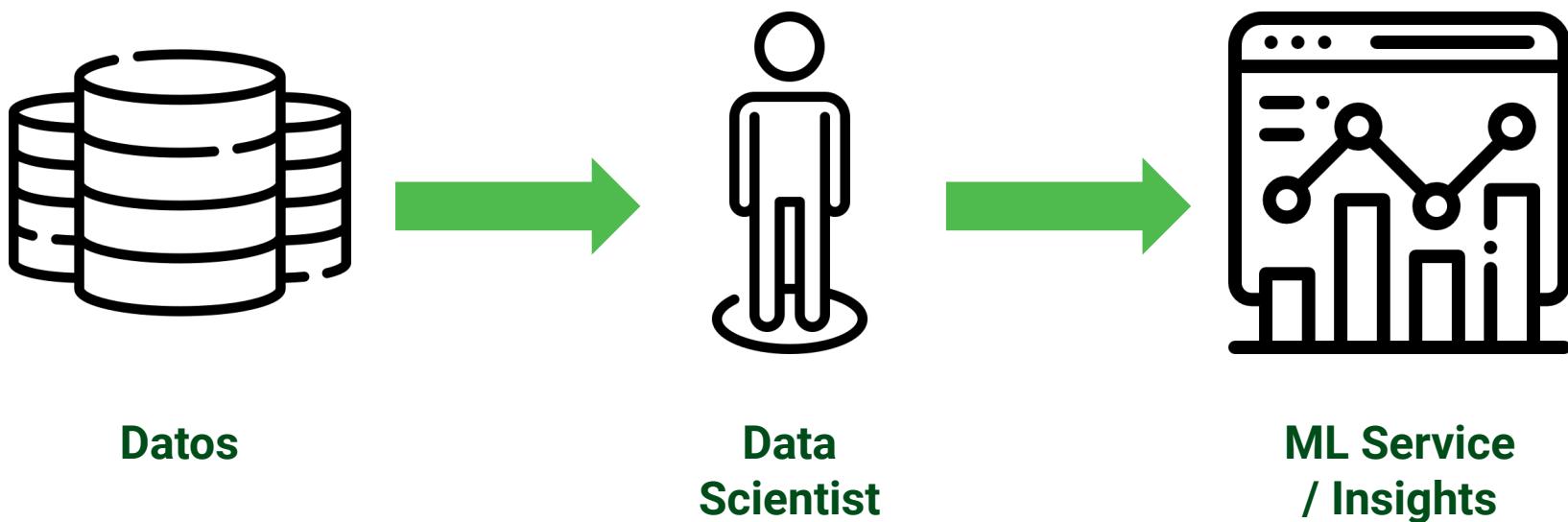


¿Qué hace una  
**Data Scientist?**





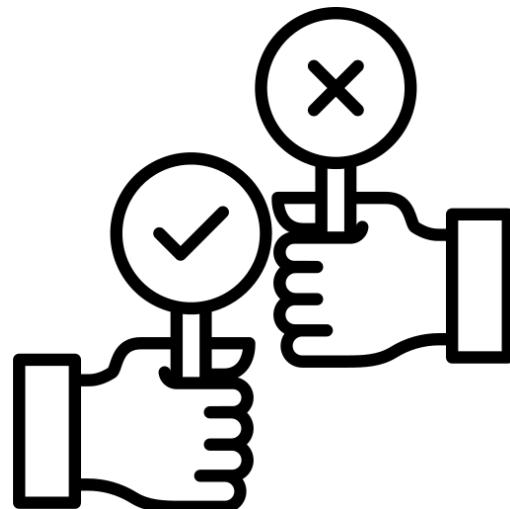
# ¿Qué hace una Data Scientist?



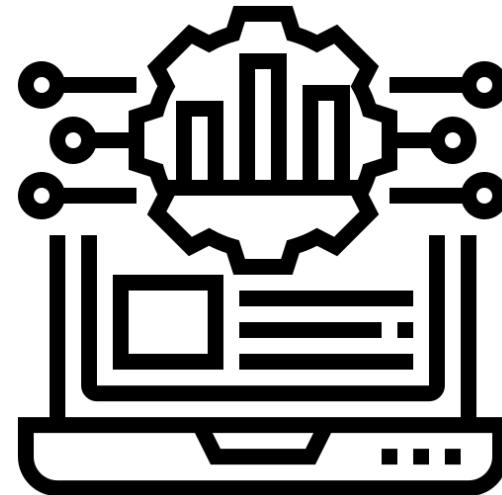


# ¿Qué hace un Data Scientist?

Toma de decisiones basadas en datos.



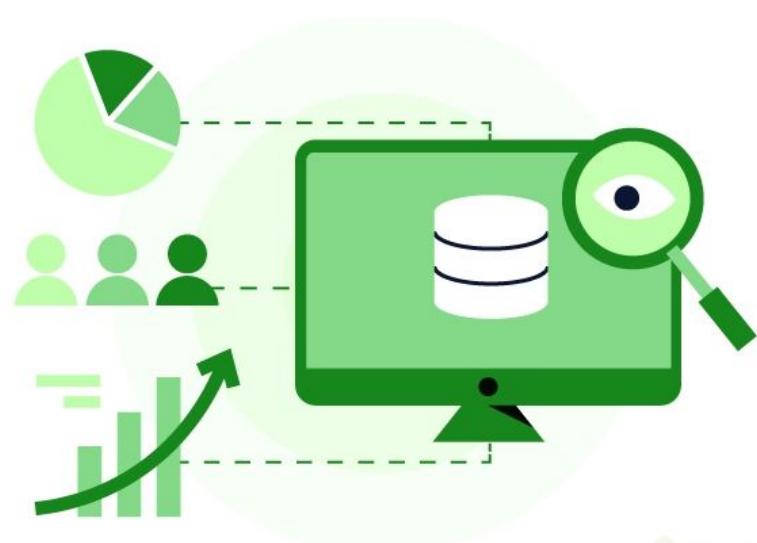
Incorporar datos a los productos de software.





# Día a día de Data Scientist

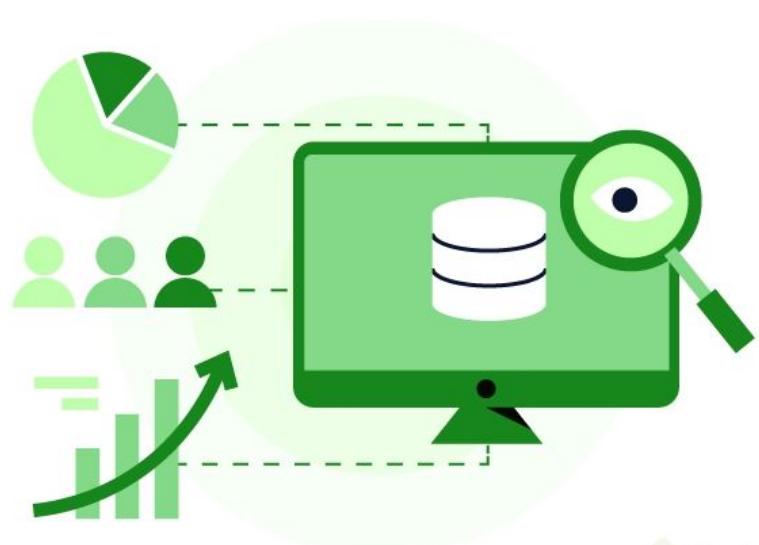
- **Obtener, limpiar y procesar datos.**
- **Diseñar y utilizar modelos de machine learning.**

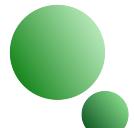




# Día a día de Data Scientist

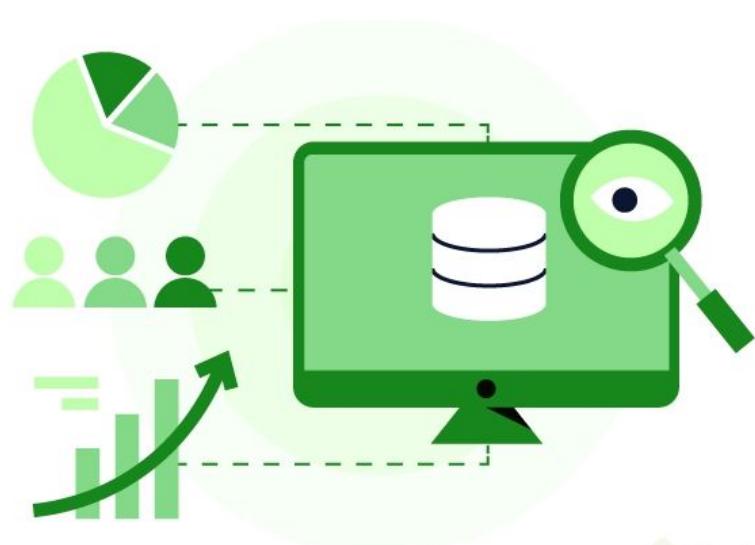
- **Monitorear la precisión de los datos.**
- **Automatizar procesos de recolección y transformación de datos.**





# Día a día de Data Scientist

- **Crear reportes de información en tableros.**
- **Incorporar datos a los productos.**





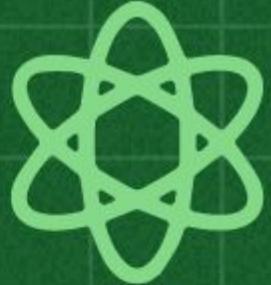
# Data Scientist vs. Data Analyst

## Data Scientist

- Incorporar datos a los productos - ingeniería.
- Programación avanzada con POO.
- Machine learning y estadística avanzada.

## Data Analyst

- Se enfoca en responder preguntas del negocio.
- Programación fundamental.
- Analizar el presente de la organización.



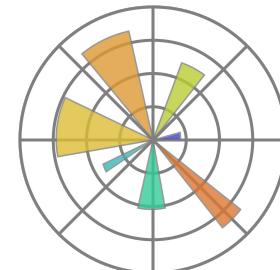
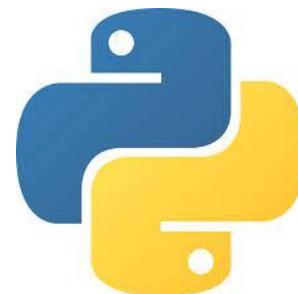
# Herramientas y tecnologías para Data Scientists





# Herramientas y tecnologías

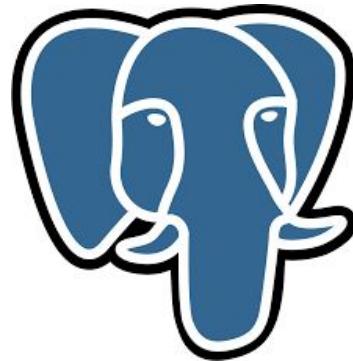
- Programación con Python o R (incluyendo POO).
- Jupyter Notebooks.
- Pandas, Numpy, Matplotlib.





# Herramientas y tecnologías

- Algoritmos y librerías de machine learning como scikit-learn y TensorFlow.
- Bases de datos SQL y NoSQL.



TensorFlow



# Matemáticas para data scientist

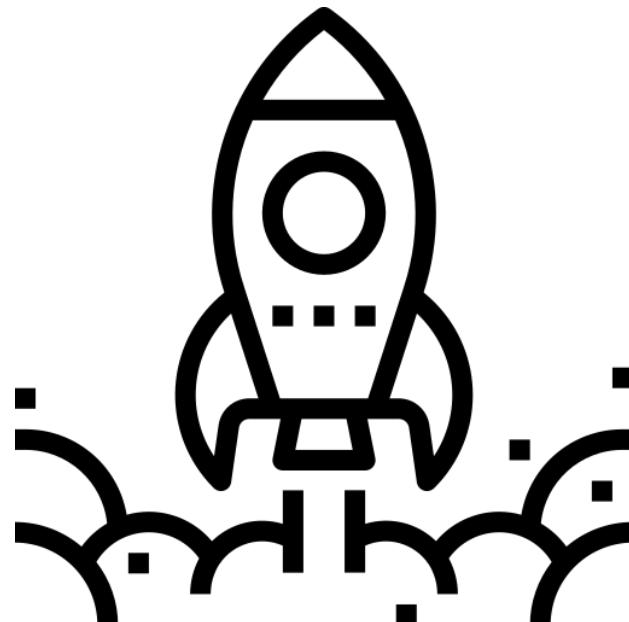
- Álgebra
- Estadística descriptiva e inferencial
- Probabilidad
- Álgebra lineal
- Cálculo





# ¿Cómo empezar?

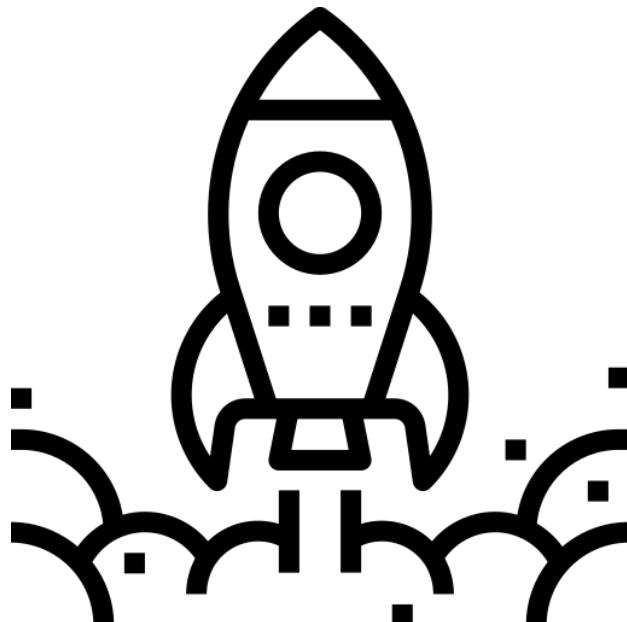
- Cómo utilizan los datos las organizaciones.
- Programación con Python y sus librerías para data science.





# ¿Cómo empezar?

- Usar Jupyter Notebooks.
- Estadística y probabilidad aplicada a data science.



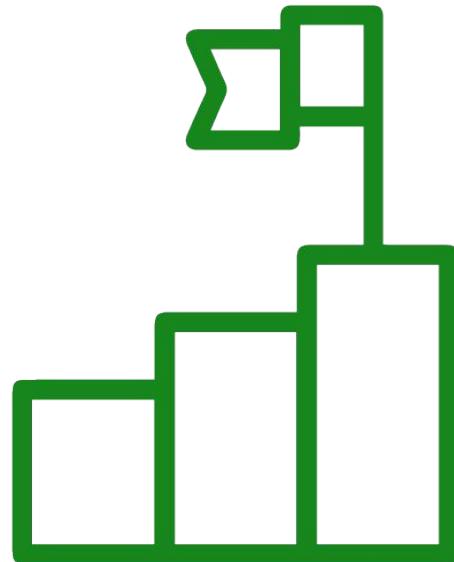


# ¿Cómo empezar?

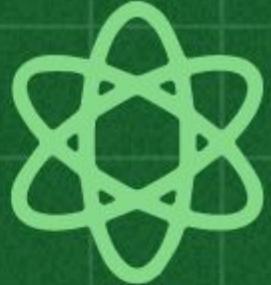




# Reto

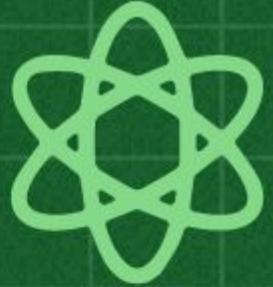


Busca en sitios de trabajo  
vacantes de Data Scientist.



# La experiencia de una Data Scientist



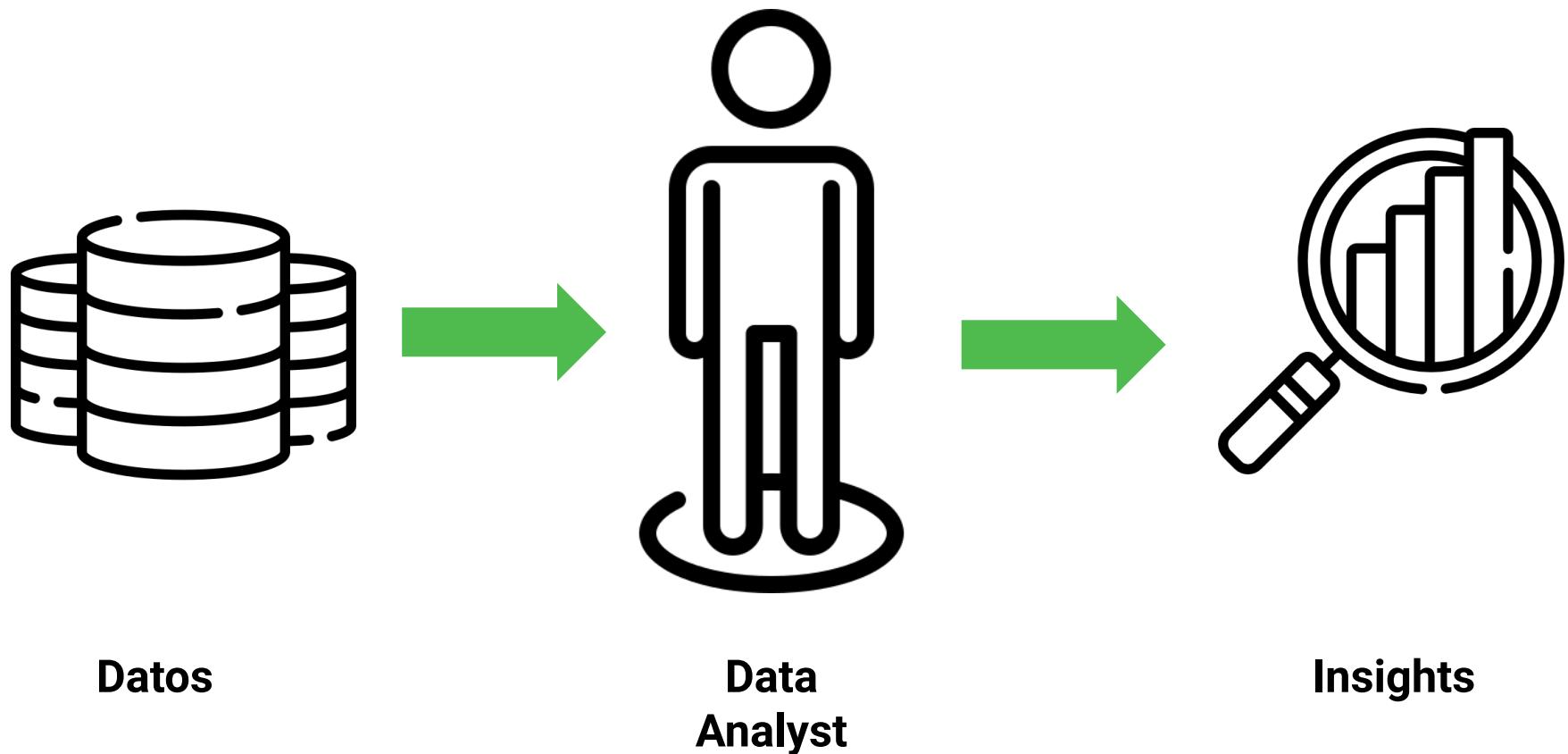


¿Qué hace una  
Data Analyst?





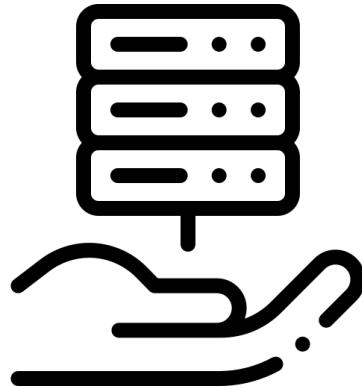
# ¿Qué hace una Data Analyst?





# ¿Qué hace un Data Analyst?

Extraer datos recolectados.



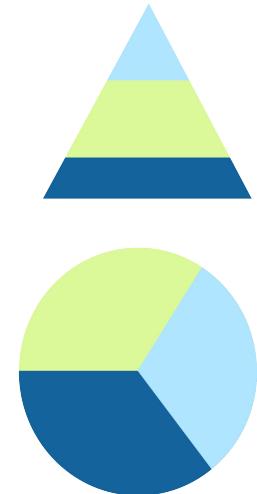
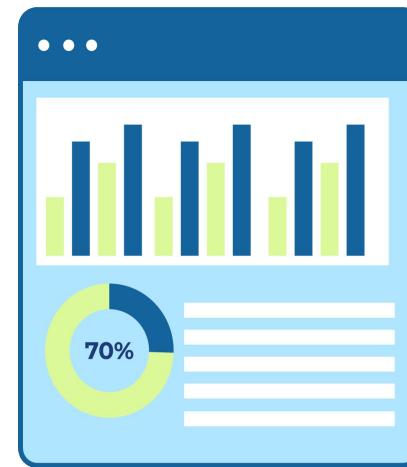
Analizarlos y reportar resultados.





# Día a día de Data Analyst

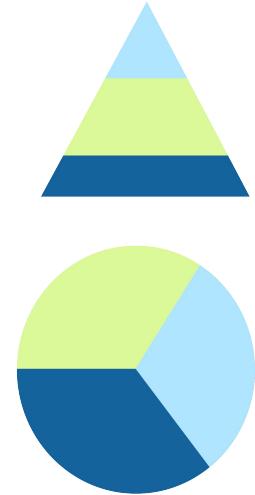
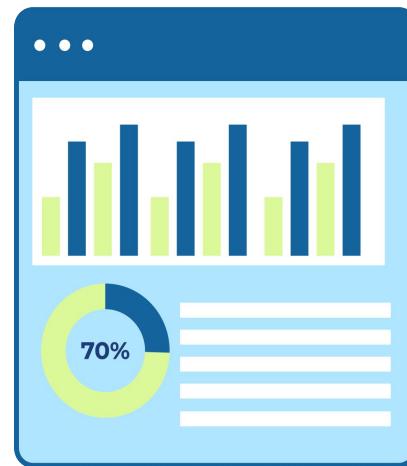
- **Identificar necesidades de información.**
- **Extraer datos de fuentes** con SQL o Python.





# Día a día de Data Analyst

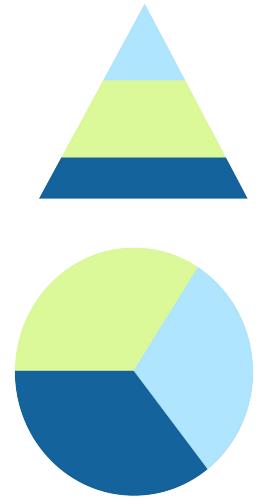
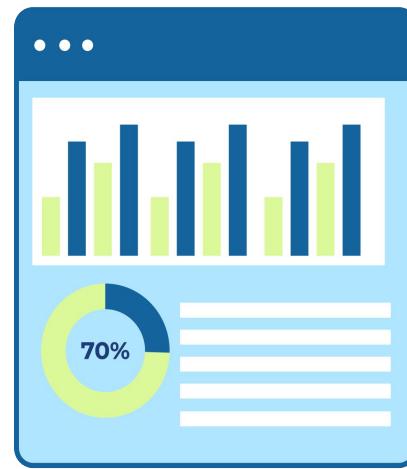
- **Limpiar y organizar los datos** para su análisis.
- **Analizar los datos** para identificar patrones y tendencias.





# Día a día de Data Analyst

- **Comunicar los hallazgos en tableros o dashboards.**



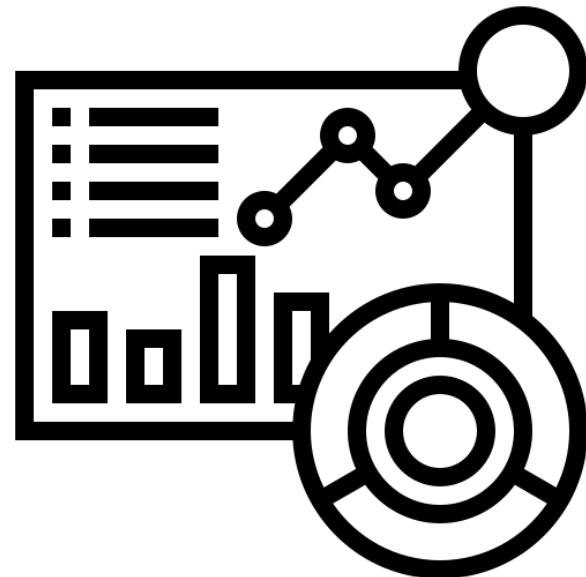
# Flujo de trabajo de Data Analyst

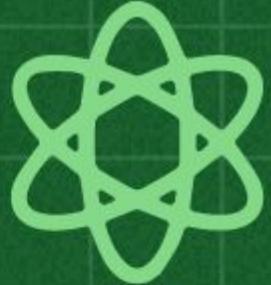




# Roles relacionados

- Business Analyst
- Data visualization specialist





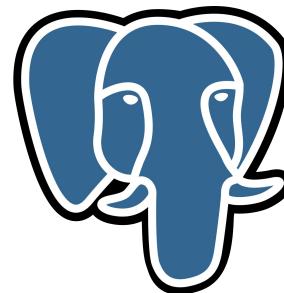
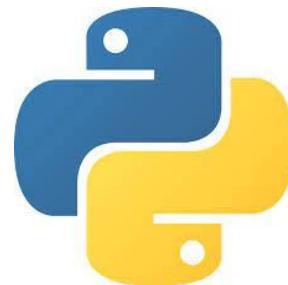
# Herramientas y tecnologías para Data Analysts





# Herramientas y tecnologías

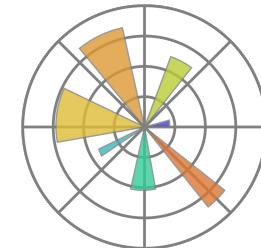
- Consulta de bases datos con SQL.
- Software de visualización de datos como Power BI y Tableau.
- Excel y Google Sheets.
- Programación con Python o R.





# Herramientas y tecnologías

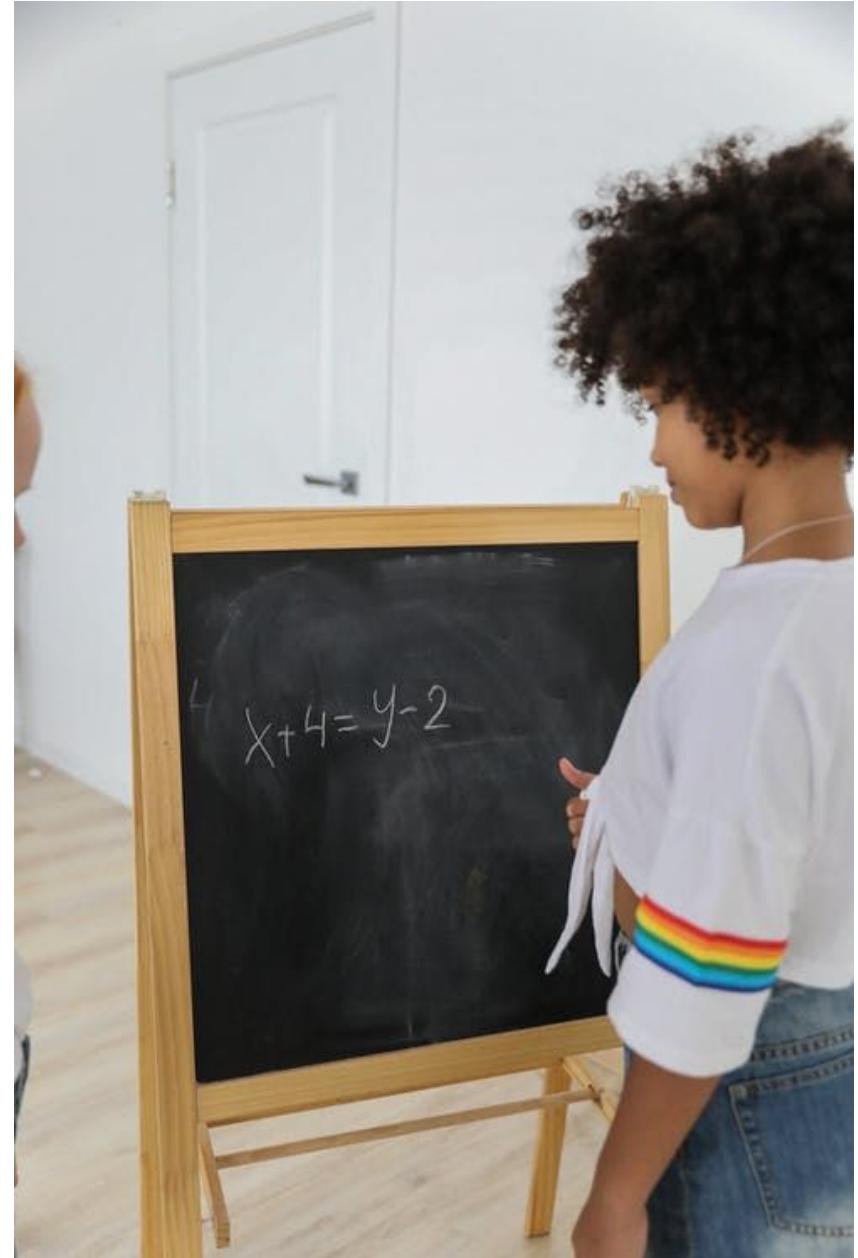
- Jupyter Notebooks.
- Pandas, Matplotlib, Numpy.





# Matemáticas para Data Analyst

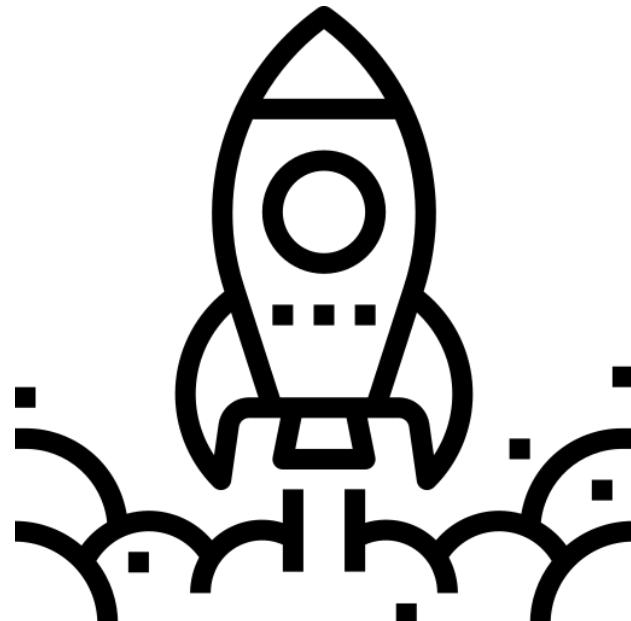
- Probabilidad
- Estadística  
descriptiva





# ¿Cómo empezar?

- Cómo utilizan los datos las organizaciones con Business Intelligence.
- Consultar bases de datos con SQL.
- Herramientas para análisis de datos como Excel, Power BI y Tableau.
- Estadística aplicada a análisis de datos.



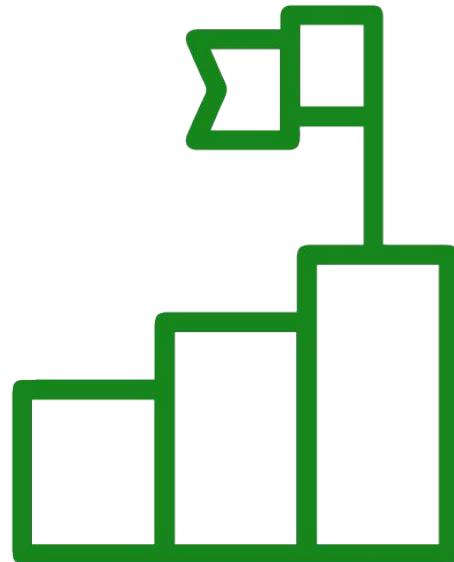


# ¿Cómo empezar?





# Reto



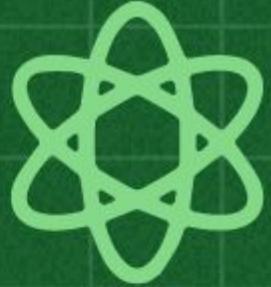
Busca en sitios de trabajo  
vacantes de Data Analyst.



# La experiencia de un Data Analyst

con Demian Arenas



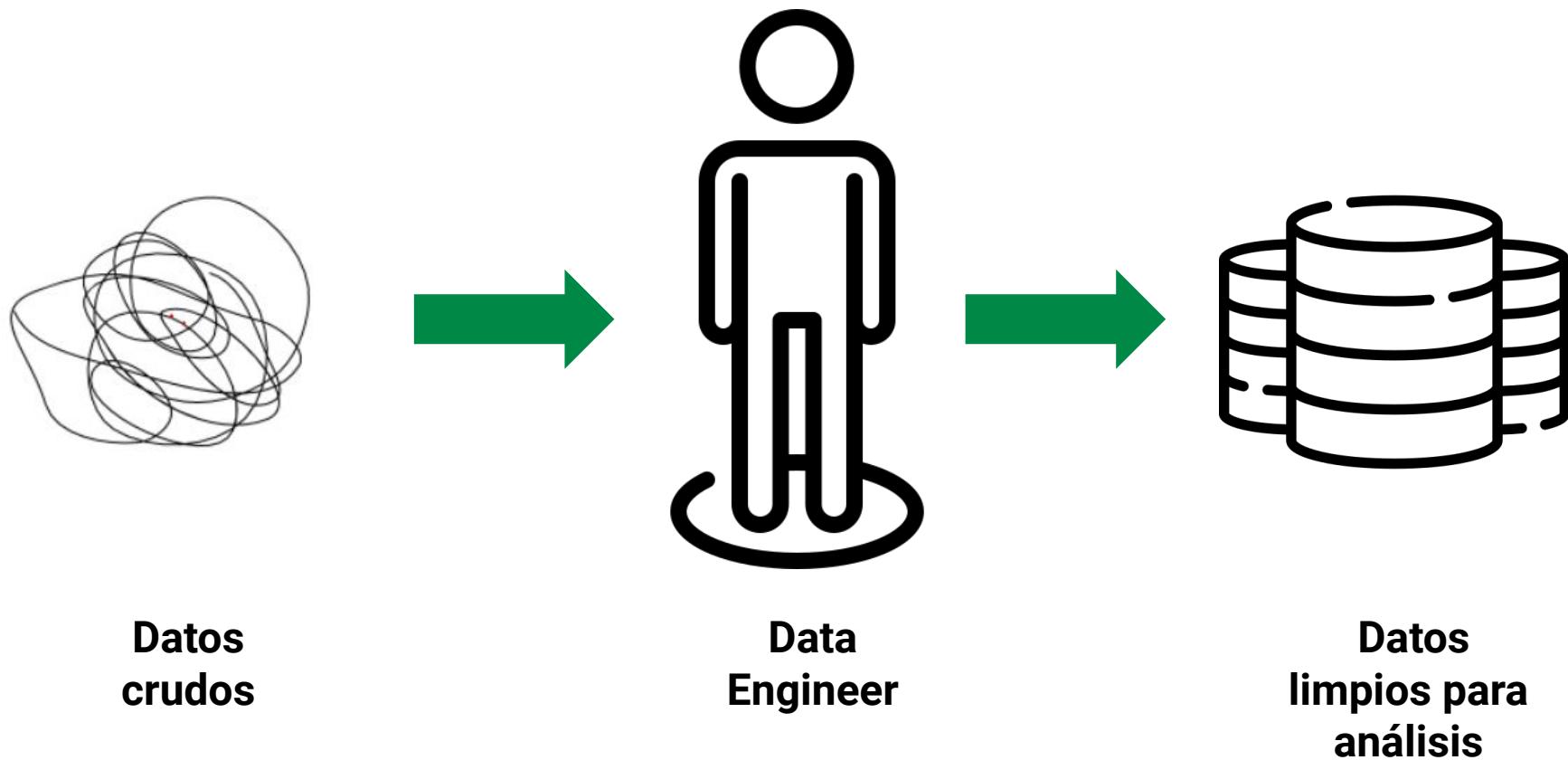


¿Qué hace una  
**Data Engineer?**





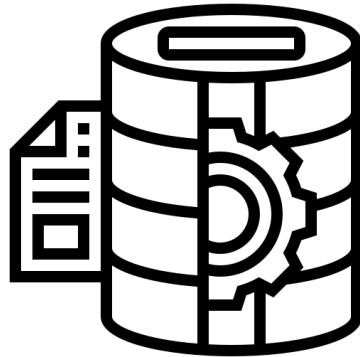
# ¿Qué hace una Data Engineer?



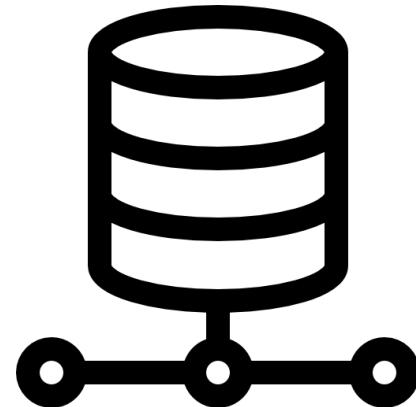


# ¿Qué hace un Data Engineer?

Trabaja para que el equipo tenga datos para análisis.



Crea pipelines ETL.

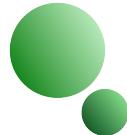




# Día a día de una Data Engineer

- **Data pipelines de ETL y bases de datos.**
- **Extraer datos de diferentes fuentes.**





# Día a día de una Data Engineer

- **Transformar los datos para análisis.**
- **Bases de datos especializadas para análisis.**





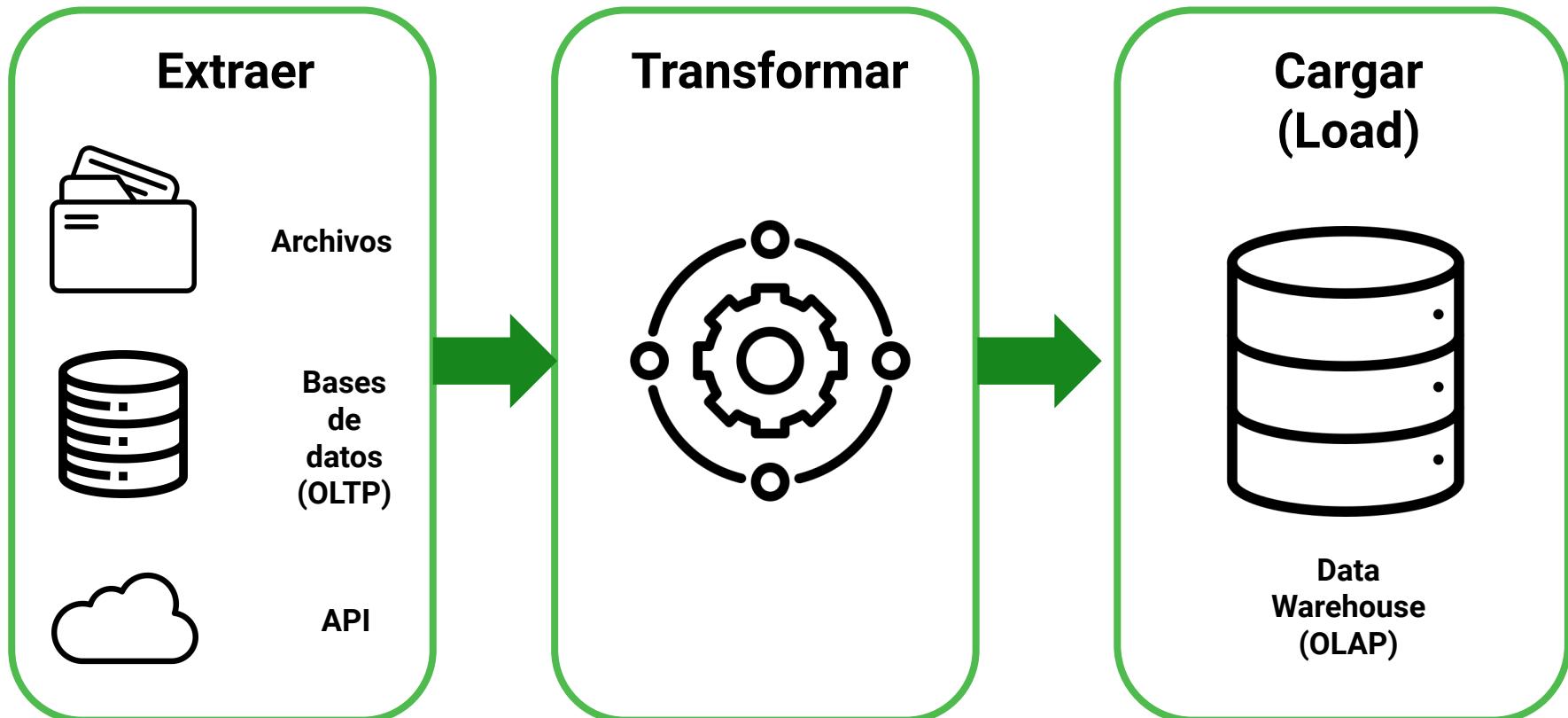
# Día a día de una Data Engineer

- **Crear automatizaciones para ETL.**





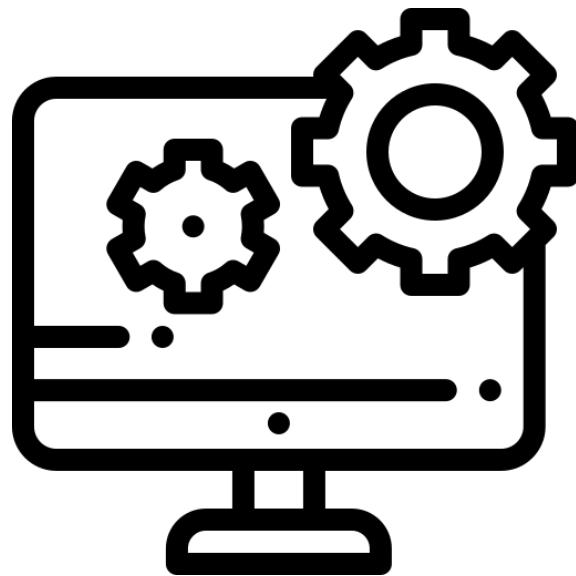
# Proceso ETL





# Roles relacionados

- Data Architect
- Big Data Architect





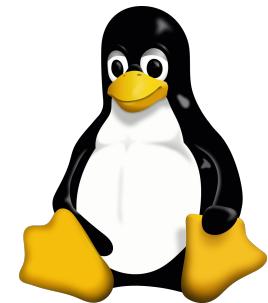
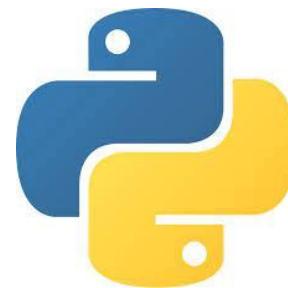
# Herramientas y tecnologías para Data Engineers





# Herramientas y tecnologías

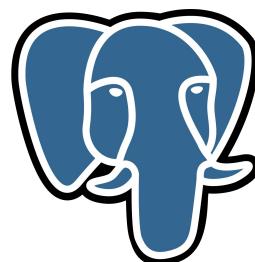
- Programación con Python y bases de ingeniería de software.
- Linux.
- Automatización y scripting.





# Herramientas y tecnologías

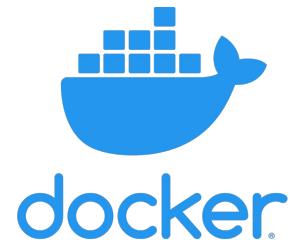
- Jupyter Notebooks y editores de código.
- Manejo avanzado de bases de datos SQL y NoSQL.
- Pandas, Dask y Apache Spark.





# Herramientas y tecnologías

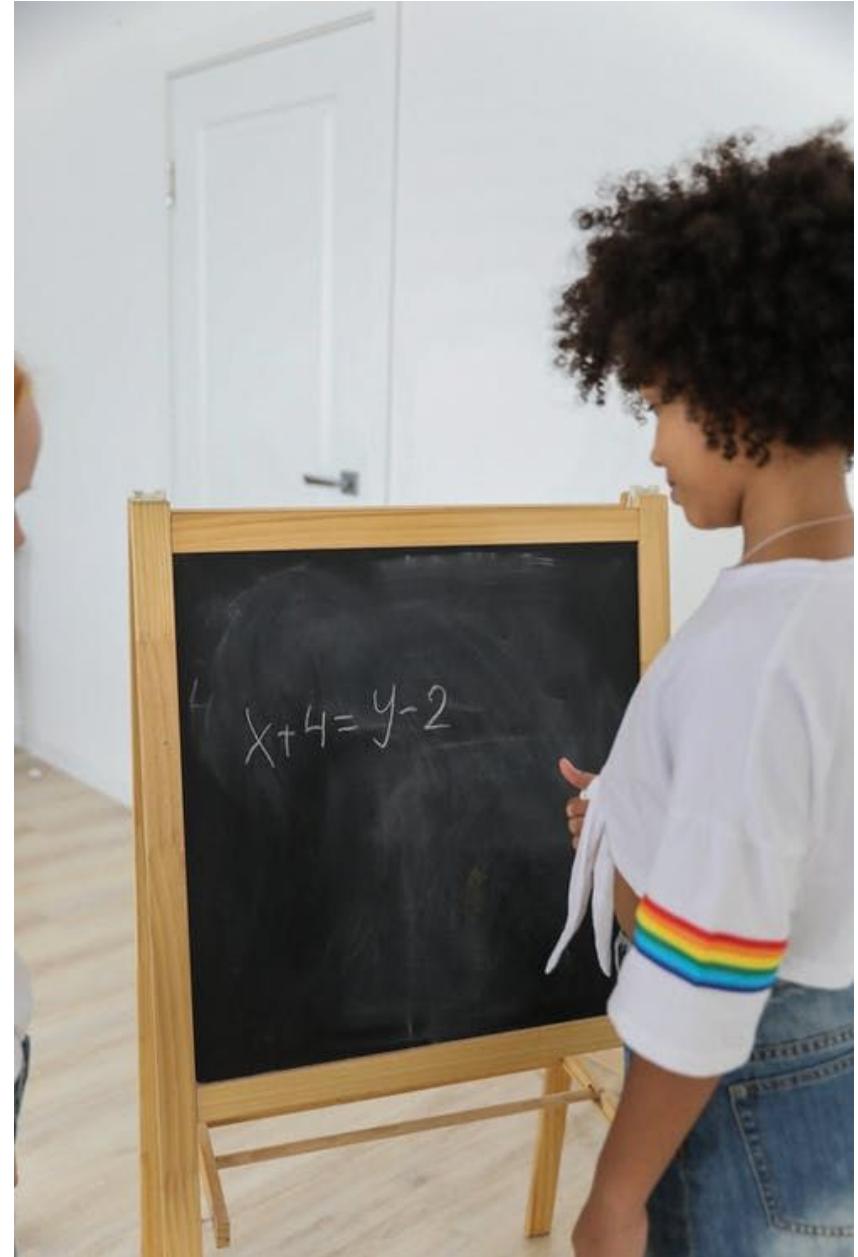
- Airflow.
- Tecnologías cloud.
- Contenedores Docker.
- Orquestadores Kubernetes.





# Matemáticas para Data Engineer

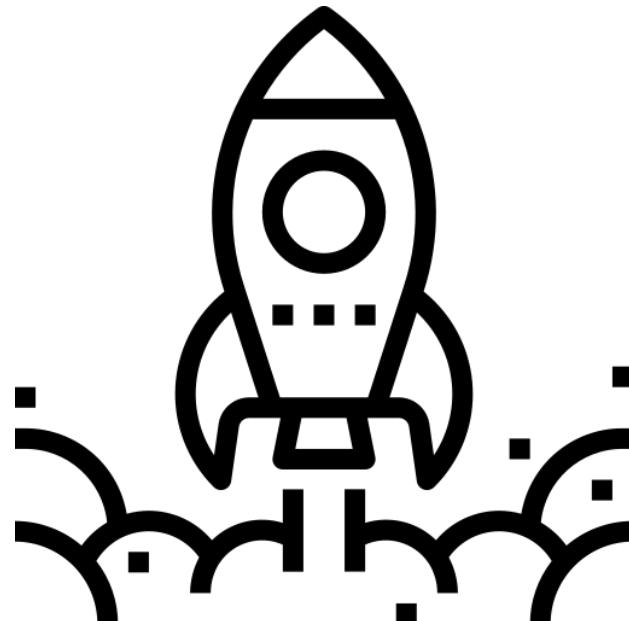
- Estadística descriptiva





# ¿Cómo empezar?

- Python y bases sólidas de ingeniería de software.
- Automatización y scripting.
- Pandas y Apache Spark.
- Bases de datos SQL y NoSQL.



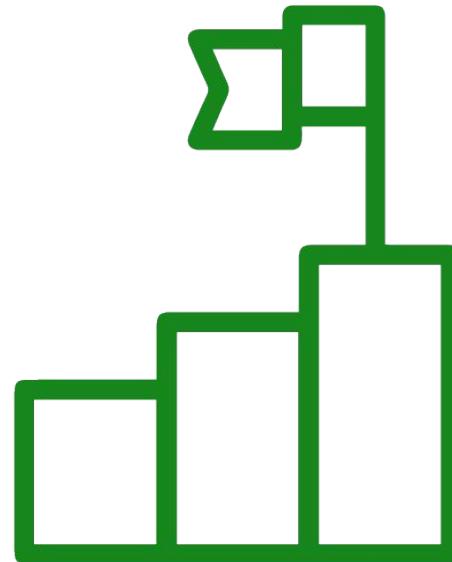


# ¿Cómo empezar?

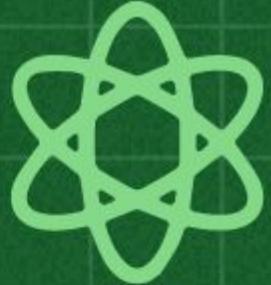




# Reto

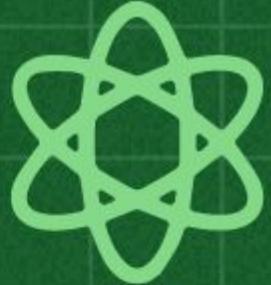


Busca en sitios de trabajo  
vacantes de Data Engineers.



# La experiencia de un Data Engineer



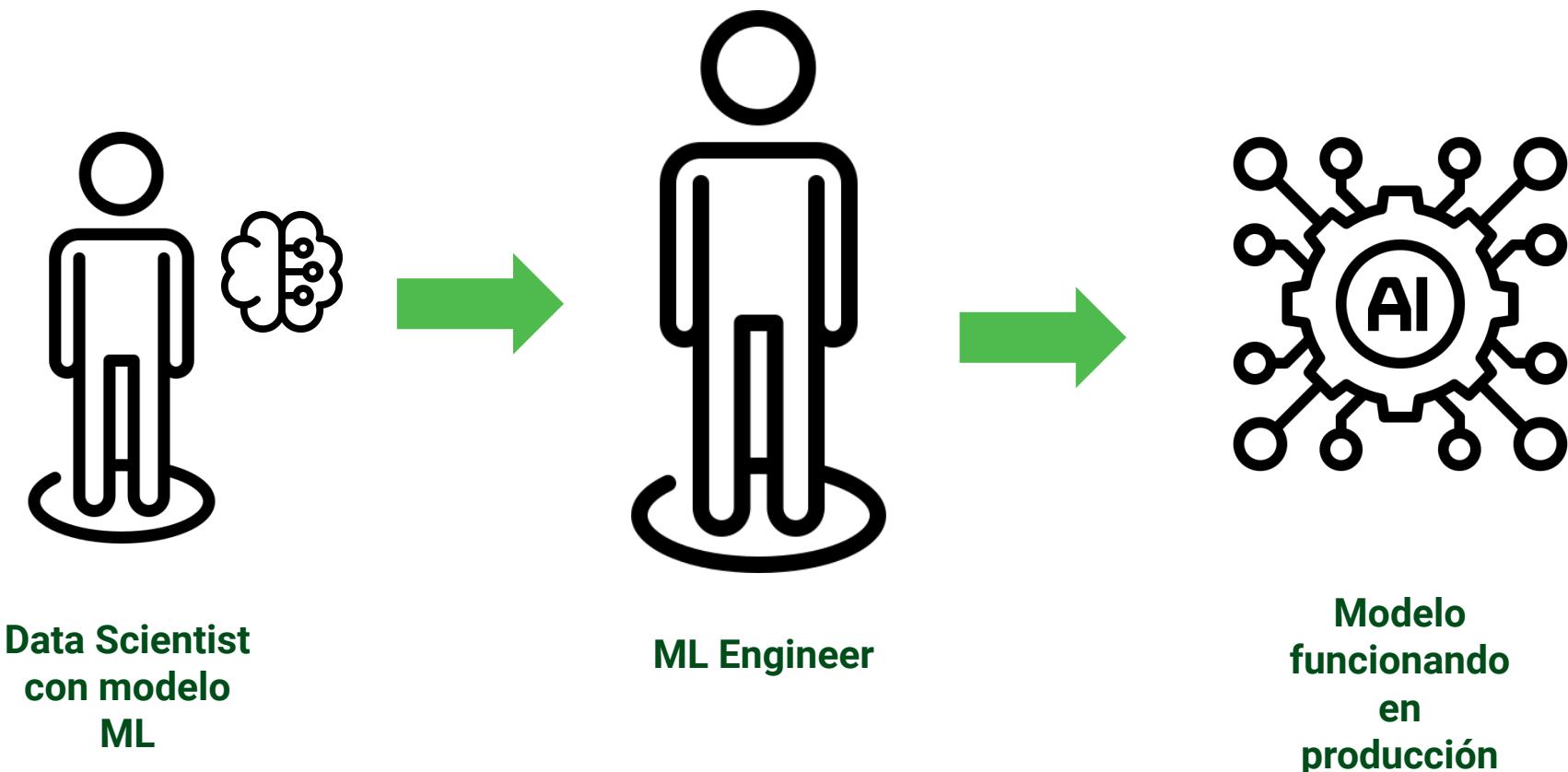


¿Qué hace una  
**Machine Learning  
Engineer?**



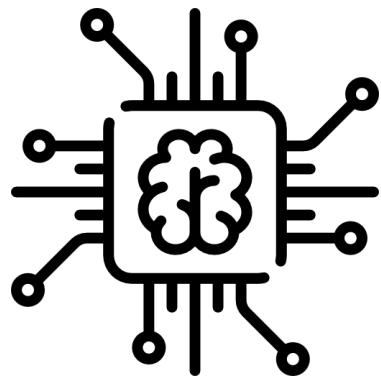


# ¿Qué hace una ML Engineer?

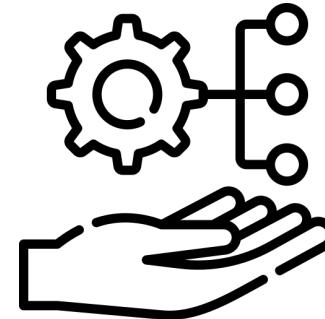


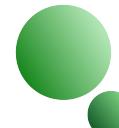
# • ¿Qué hace una ML Engineer?

Crear productos  
basados en IA.



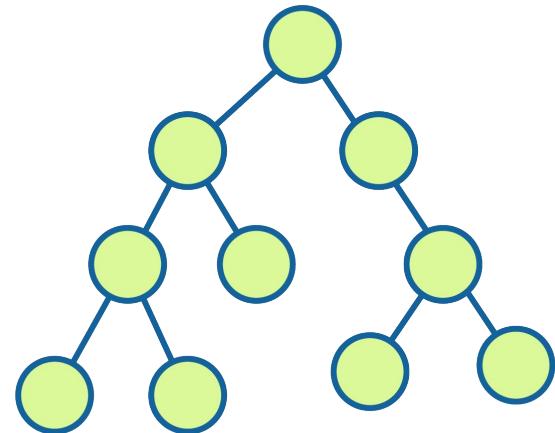
Escalar modelos de IA.





# Día a día de ML Engineer

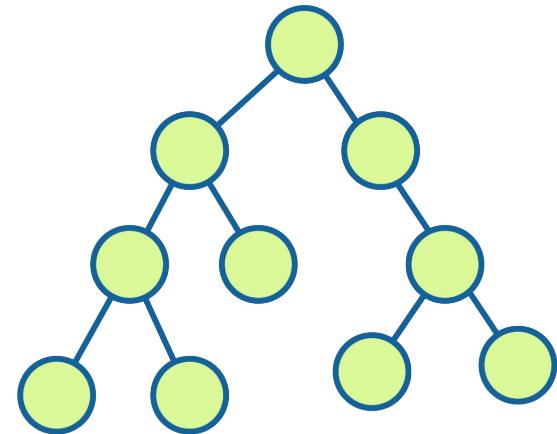
- Generar una **evaluación extensiva de métricas** de modelos de machine learning.
- Construir, escalar y robustecer sistemas de machine learning que **funcionen en producción**.





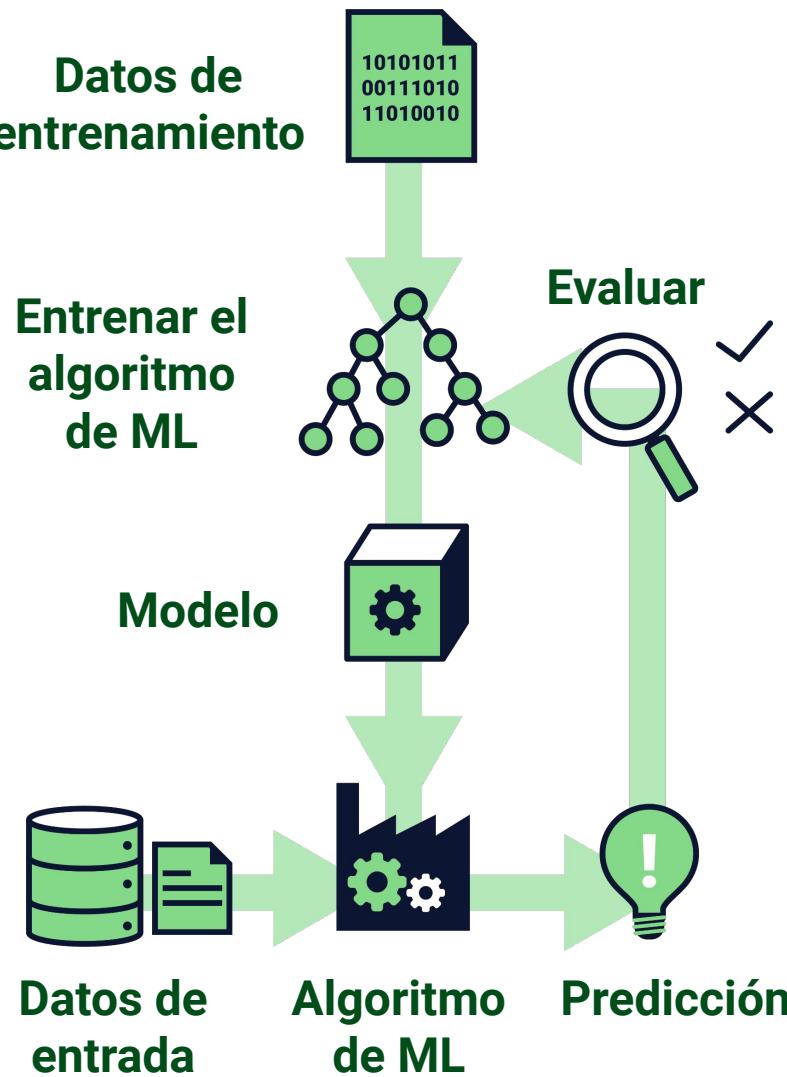
# Día a día de ML Engineer

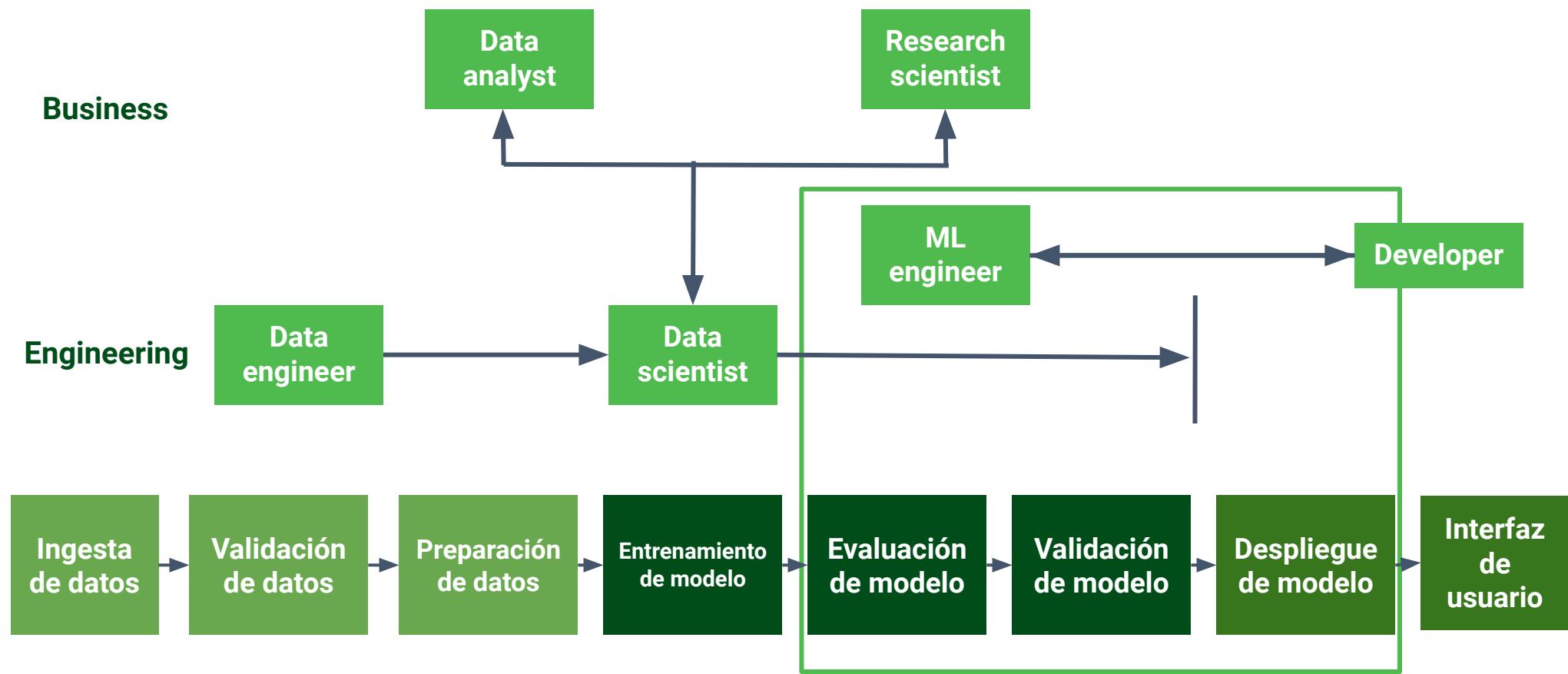
- **Colaborar** con Data Scientists y otras áreas de ingeniería de software.
- **Monitorear el desempeño y la funcionalidad** de los sistemas de machine learning.





# Proceso de machine learning





*Pasos y roles en el flujo de trabajo de data science  
(Design Patterns in Machine Learning).*



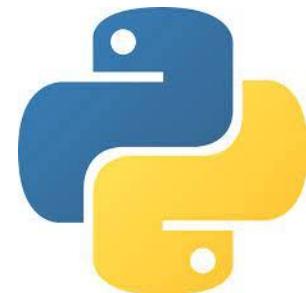
# Herramientas y tecnologías para Machine Learning Engineers





# Herramientas y tecnologías

- Programación avanzada con Python, Java y C++.
- Bases sólidas de ingeniería de software.
- Jupyter Notebooks.





# Herramientas y tecnologías

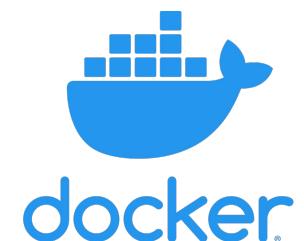
- Pandas, Numpy, Matplotlib, Seaborn.
- Uso extensivo de frameworks y librerías de machine learning.

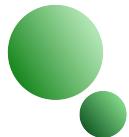




# Herramientas y tecnologías

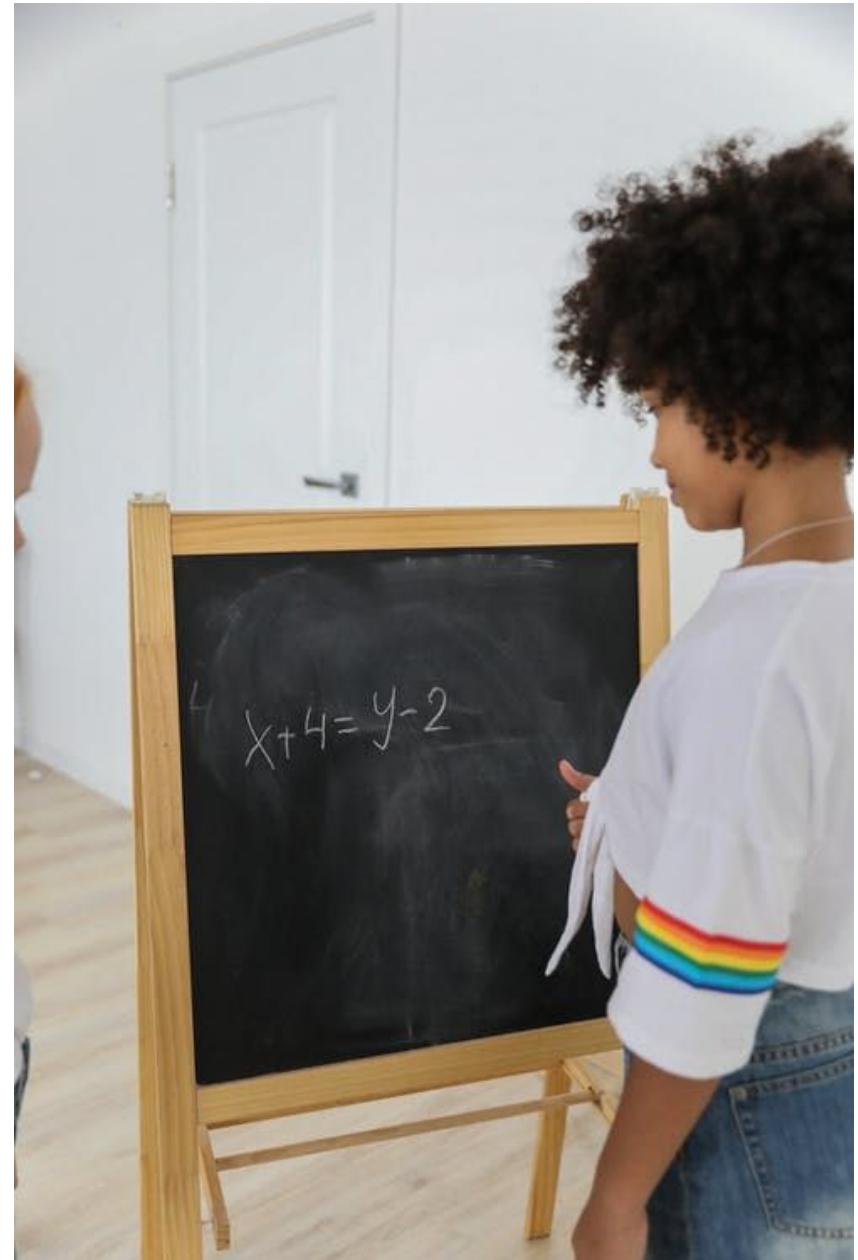
- Flask o FastAPI
- Tecnologías cloud
- Contenedores Docker
- Kubernetes

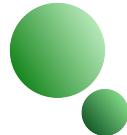




# Matemáticas para machine learning

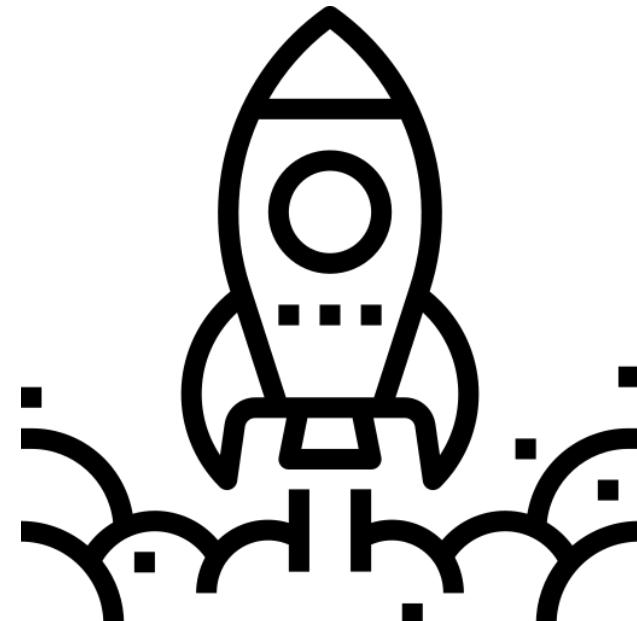
- Estadística descriptiva e inferencial
- Probabilidad
- Álgebra lineal
- Cálculo





# ¿Cómo empezar?

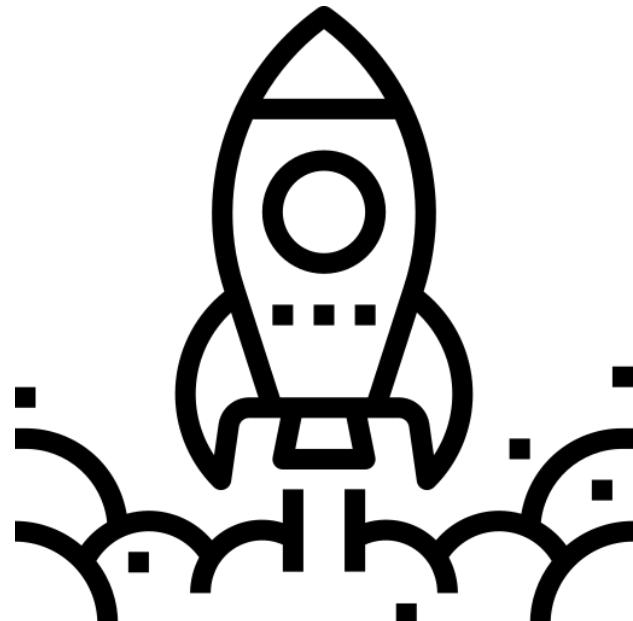
- Programación con Python y bases sólidas de ingeniería de software.
- Uso de librerías de Python para manipulación, análisis y visualización de datos.





# ¿Cómo empezar?

- Matemáticas aplicadas a data science e inteligencia artificial.
- Aplicación de modelos de machine learning con scikit-learn.



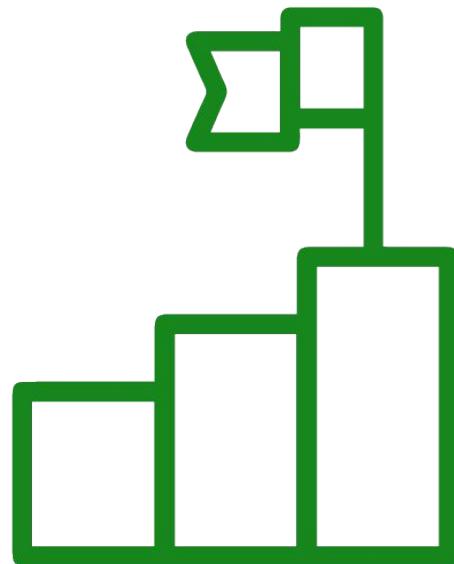


# ¿Cómo empezar?

**Platzi**  
**ESCUELA DE**  
  
**MACHINE LEARNING ENGINEER**



# Reto

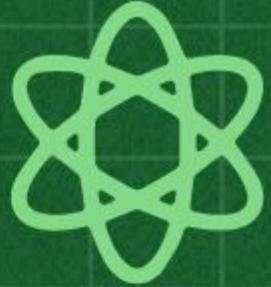


Busca en sitios de trabajo  
vacantes de ML Engineers.



# La experiencia de un **Machine Learning Engineer**





# Soft skills para profesionales de los datos e IA





En sus marcas,  
**listos, fuera**

