

---

# ANALÍTICA DE DATOS: BIGDATA – MACHINE LEARNING

- Dr. Ing. Gómez Avila José

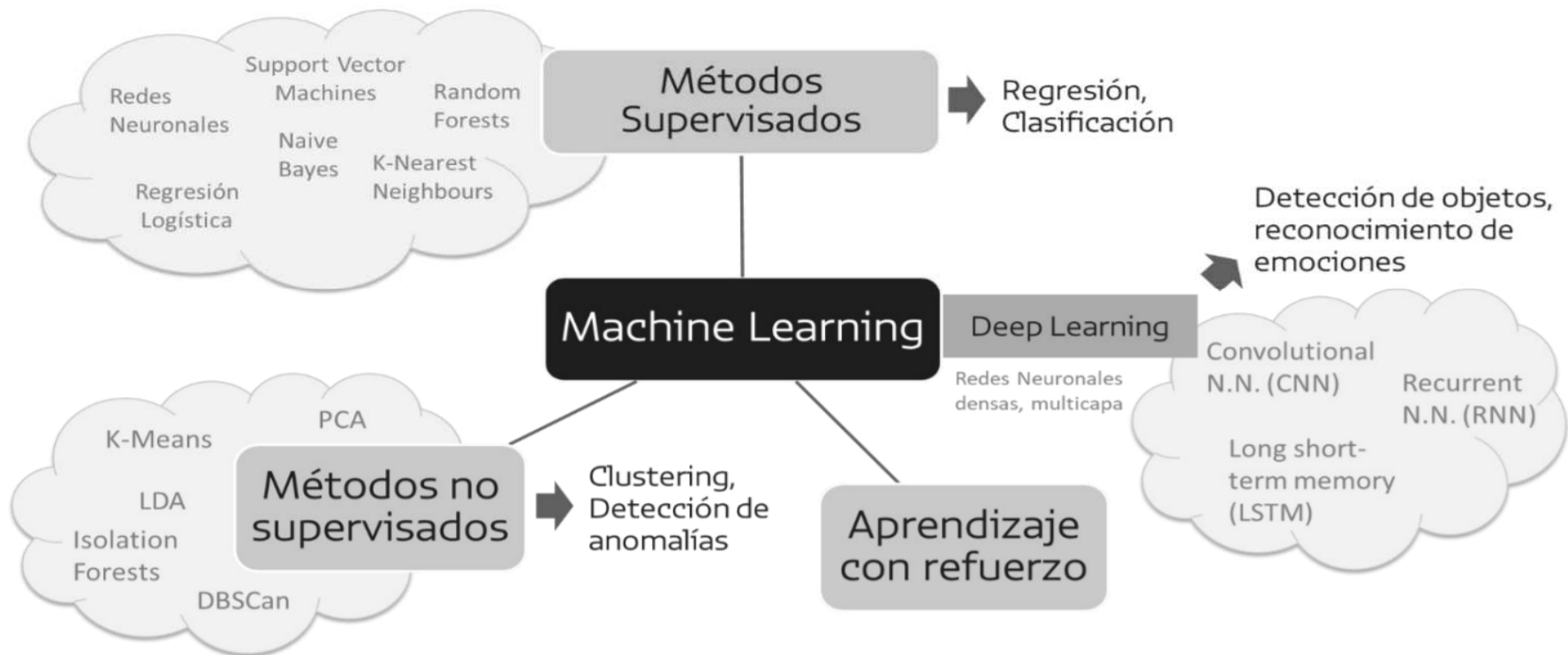
## INTERACCION MOTIVACIONAL

## Objetivo de la sesión



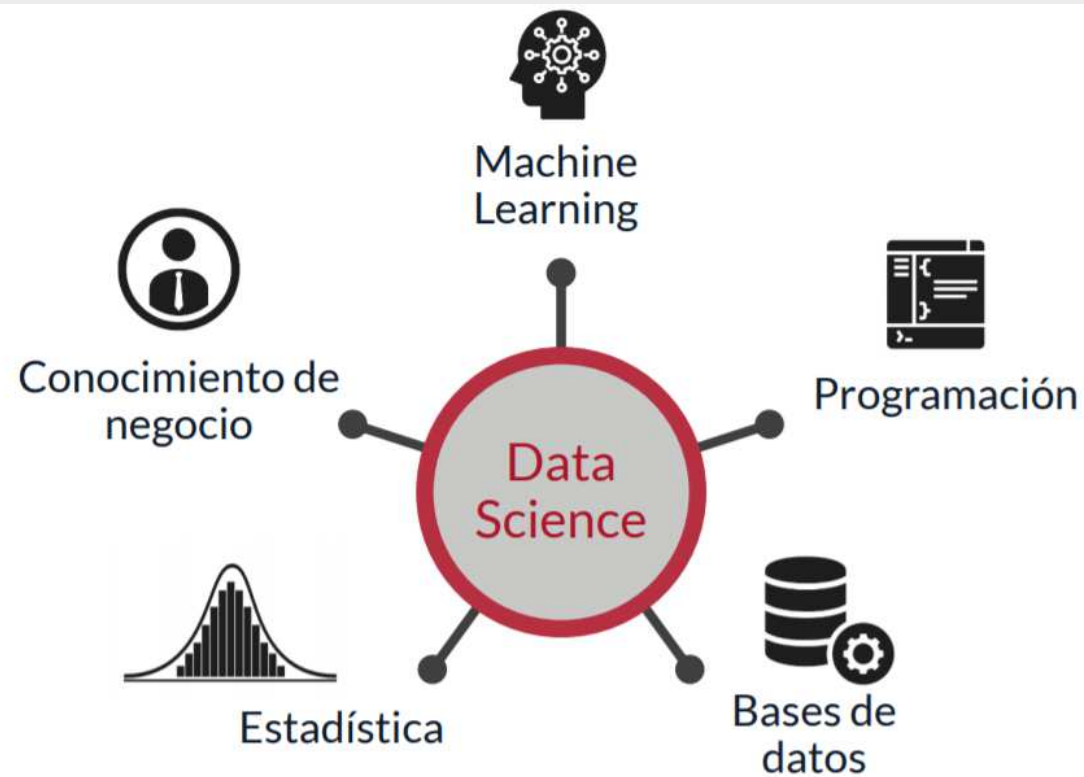
- Al terminar la sesión el estudiante comprende la analítica de datos y desarrolla un resumen ejecutivo de los principales términos utilizando los criterios impartidos por el docente.

## Hay distintos métodos de aprendizaje



- Uso del método científico para generación de insights a partir de los datos.

## Data Science es una disciplina que combina múltiples aspectos



## Analítica Avanzada



Patrones y tendencias con  
modelos

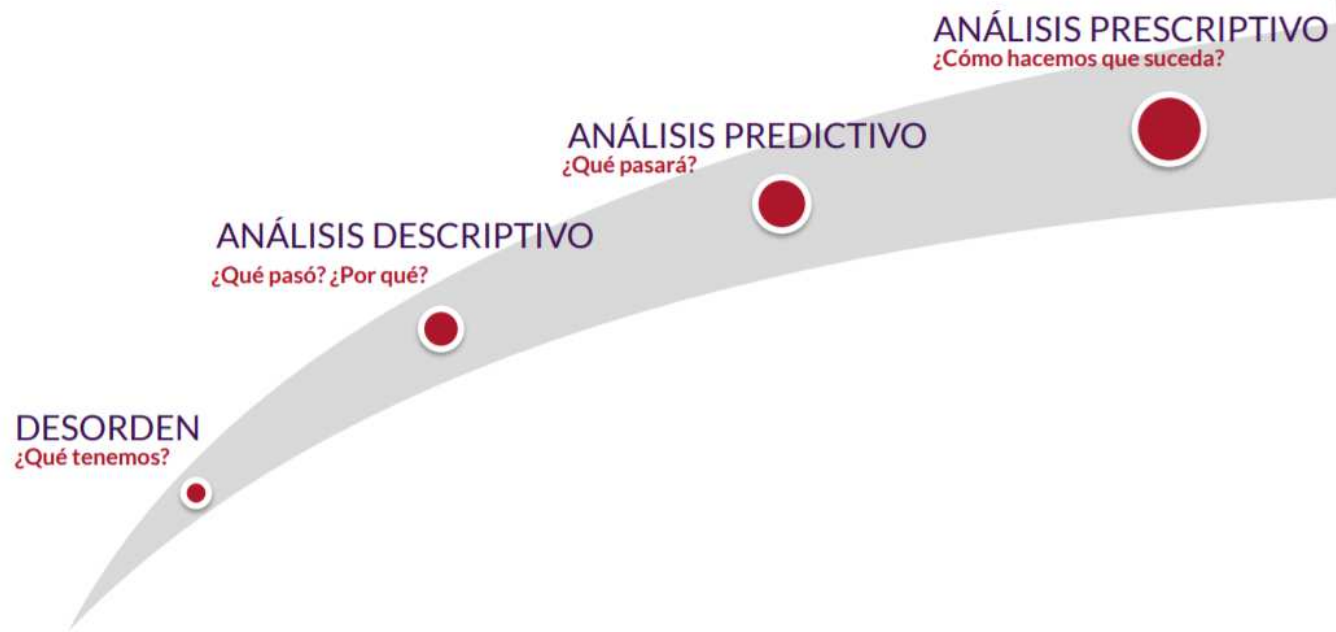


Algoritmos complejos y  
computación avanzada



Foco en predicción

## Las iniciativas con Analítica Avanzada depende del nivel de madurez





## Las iniciativas de la Analítica Avanzada / Data Science son iterativas



Estándar  
CRISP-  
DM

## Analítica Avanzada

VS

## BI Tradicional

Foco

Predicción, optimización y simulación

Reportes y visualizaciones

Uso de datos

Estructurados y no estructurados

Estructurados y algunos no estructurados

Preguntas

"Qué pasará"

"Qué pasó"

Implementación

Algoritmos complejos y/o  
arquitecturas tecnológicas especiales

Requerimientos cuantificables y  
normalmente predecibles

Usuarios

Data Scientists, Analistas de Negocio,  
Usuarios finales de negocio

Usuarios finales de negocio,  
Analistas

Uso

Proactivo

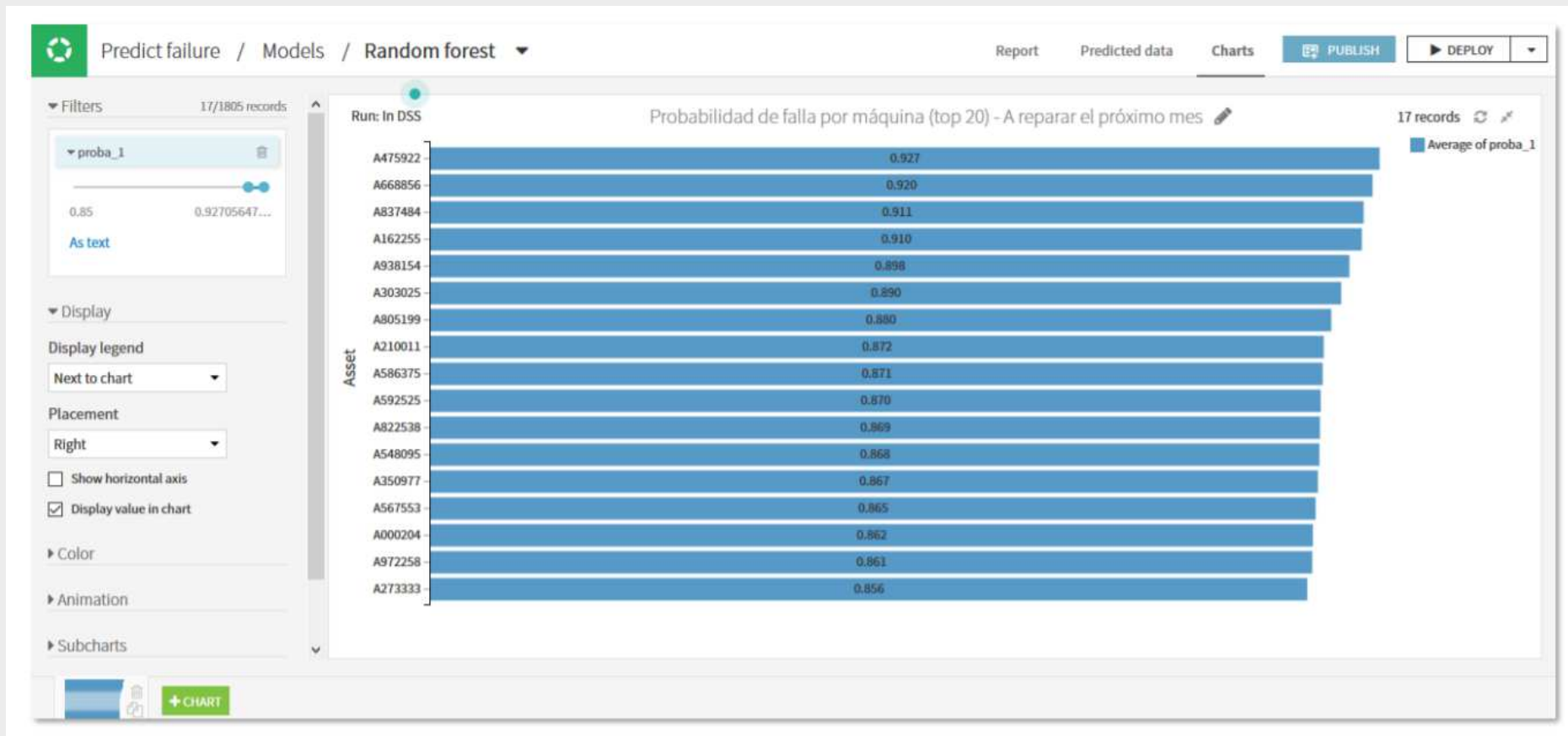
Reactivo

## Casos de uso en Analítica Avanzada

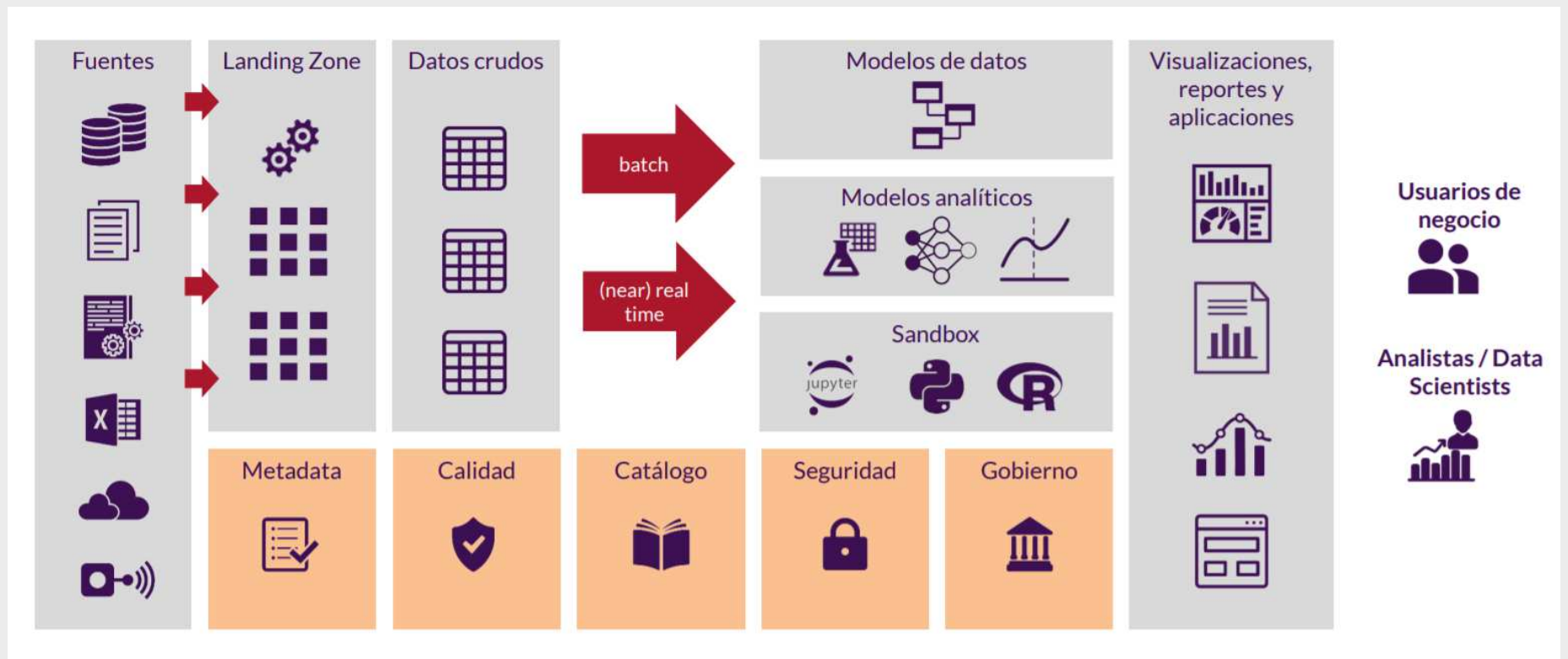


- Predicción de fuga
- Sistemas automáticos de recomendación
- Identificación de fraudes y anomalías
- Segmentación de clientes / audiencias
- Comprensión de comportamientos
- Detección de patrones complejos
- Optimización de precios, lugares, combinaciones
- Scores de riesgo

# Ejemplo: Mantenimiento Predictivo



# Arquitectura Típica Data Lake



## Data Lake

VS

## Data Warehouse

Foco

Integración de datos oportuna

Integración de datos correcta

Uso de datos

Estructurados y no estructurados

Estructurados y algunos no estructurados

Preguntas

"Qué pasa ahora y qué pasará"

"Qué pasó"

Implementación

Arquitecturas especiales, escalabilidad horizontal (muchas máquinas)

Escalabilidad vertical (más memoria, más disco, misma máquina)

Usuarios

Data Scientists, Analistas de Negocio, Usuarios finales de negocio

Usuarios finales de negocio

Uso

Proactivo

Reactivo

# Inteligencia Artificial



Contexto,  
adaptación

## Analítica Avanzada



Big  
Data



Data Science



Machine  
Learning

Optimización

## La verdad es que son temas de equipo



**Rol principal:**

Extraer conocimiento desde los datos y agregar valor con modelos

**Habilidades clave:**

Matemáticas, Estadística, Programación, Comunicación, Machine Learning

**Programas (ej.):**

SQL, Python, R



**Rol principal:**

Crear interfaces de consulta a datos desde múltiples fuentes y sistemas

**Habilidades clave:**

Programación avanzada, Sistemas distribuidos, Big Data, Flujos de información

**Programas (ej.):**

Hadoop, NoSQL, Python



**Rol principal:**

Crear consultas y visualizaciones para entender patrones

**Habilidades clave:**

Estadística, Comunicación, Conocimiento del Negocio, Visualización

**Programas (ej.):**

Excel, Tableau, SQL



## OTROS ROLES



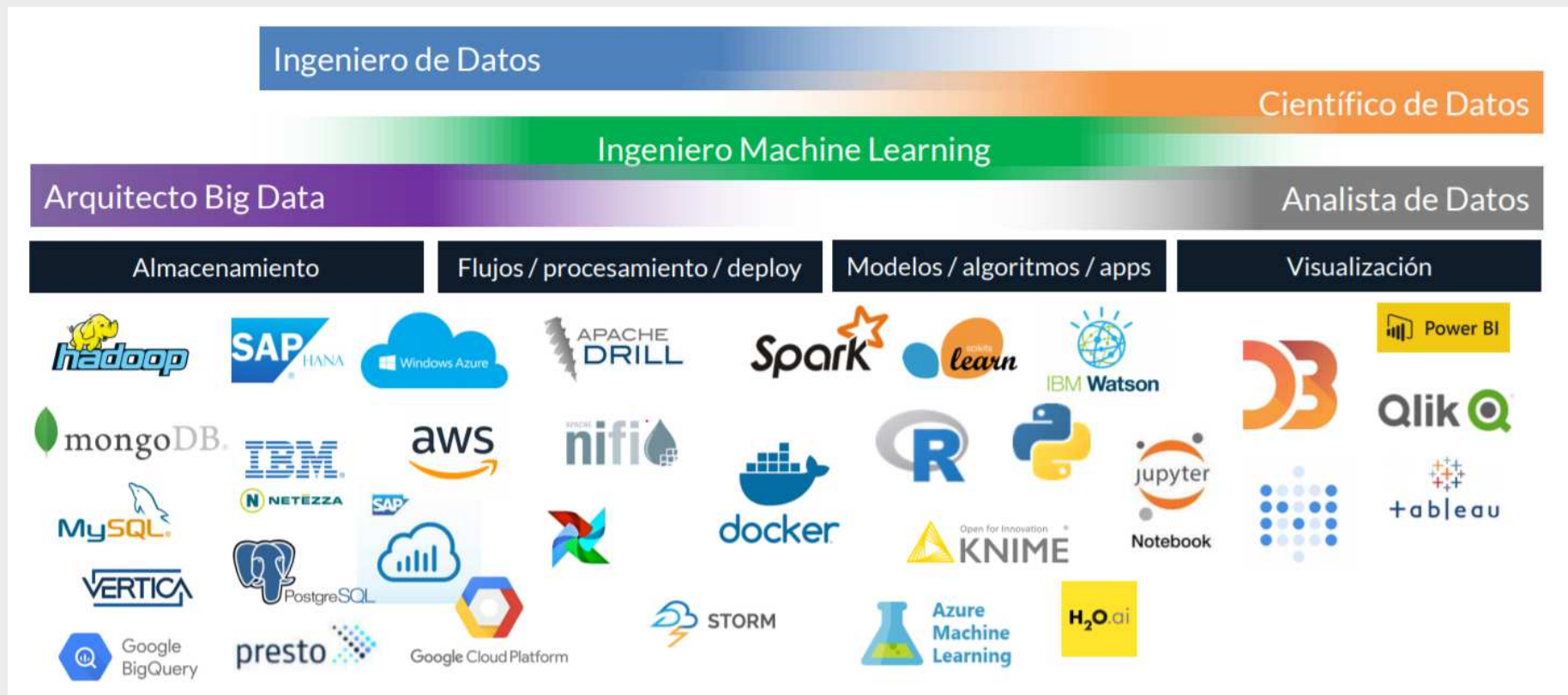
### **Machine Learning Engineer**

Rol principal: diseñar y desarrollar sistemas y modelos basados en Machine Learning

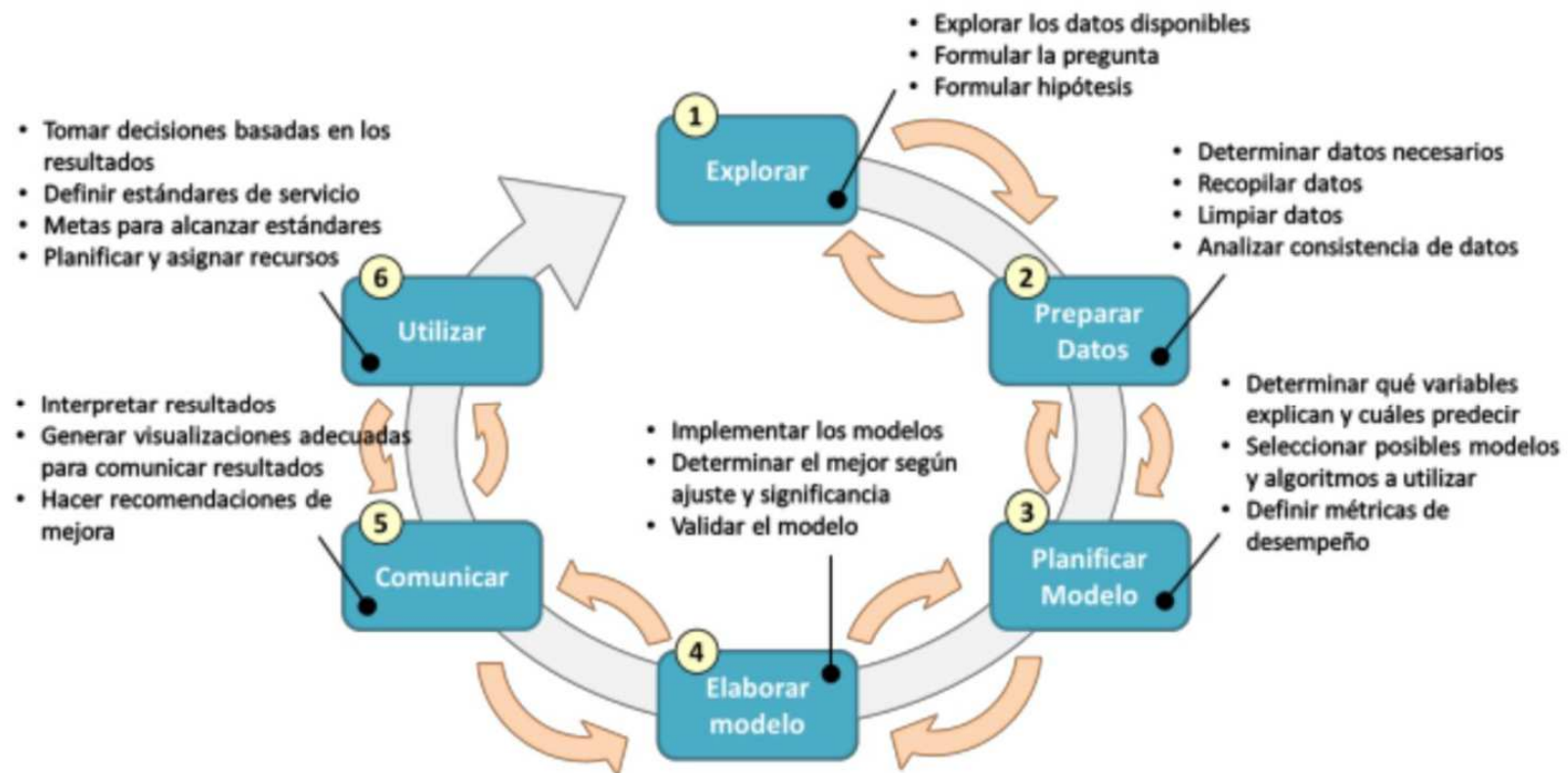
### **Big Data Architect**

Rol principal: diseñar componentes de almacenamiento, ingesta y consumo de datos de gran volumen, variabilidad y/o velocidad de generación

## ¿Dónde aplican los roles?



## Esquema conclusión



Ciclo de vida del análisis de datos

## PREGUNTAS



- GRACIAS POR SU ATENCIÓN