

# Ciclo de Vida de un Proyecto de Ciencia de Datos

Luis Norberto Zúñiga Morales

29 de marzo de 2022

# Ciclo de Vida de un Proyecto de Ciencia de Datos

Consta de cinco fases principales:

- 1 Conocimiento del área de estudio
- 2 Adquisición y estudio de datos
- 3 Modelado
- 4 Implementación
- 5 Aceptación de la aplicación

# Ciclo de Vida de un Proyecto de Ciencia de Datos

## Data Science Lifecycle

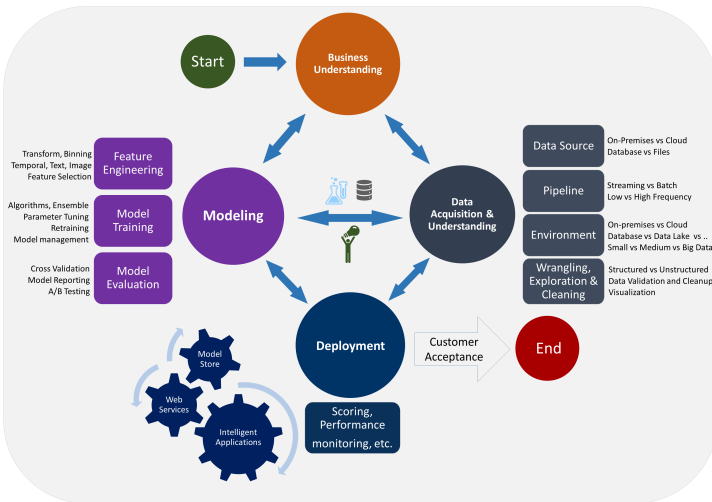


Figura: Ciclo de vida de un proyecto de ciencia de datos. Fuente: [Microsoft](#)

- **Definir objetivos**
- **Identificar el origen de los datos**
  - Datos pertinentes para la pregunta. ¿Tenemos indicadores para medir el objetivo y las características que están relacionados con él?
  - Datos que representen una medida precisa de nuestro objetivo de modelo y de las características de interés.

# Adquisición y Comprensión de los Datos

- **Introducción de los datos:** mover los datos desde las ubicaciones de origen a las ubicaciones de destino donde se ejecutan las operaciones de análisis.
- **Exploración de los datos:** Comprender qué dicen los datos por medio de técnicas de visualización y resumen.
- **Configuración de una canalización de datos:** ¿cómo se canalizan nuevos datos al sistema? ¿Lotes? ¿Tiempo real?

- **Ingeniería de características:** incluir, agregar y transformar variables sin procesar para crear las características que se utilizan en el análisis
- **Entrenamiento del modelo**
  - Elección del o los modelos
  - Entrenar el modelo con los datos proporcionados
  - Evaluar su desempeño y elegir el mejor
- **Evaluación del modelo**
  - ¿El modelo responde a la pregunta con la confianza suficiente considerando los datos de prueba?
  - ¿Debe probar algún planteamiento alternativo?
  - ¿Debe recopilar datos adicionales, rediseñar las características o experimentar con otros algoritmos?

# Implementación

- ¿Cómo se implementan (*deploy*) los modelos? ¿Cómo se hacen operativos?
  - Nube
  - Hojas de cálculo
  - Software propio o línea de comando
  - Aplicaciones de back-end como servicio web

¿Cumple las expectativas del cliente o del programa de apoyo?



# Bibliografía