

Introducción a Dataops

Información General

Fecha de inicio: 22/09/2026

Frecuencia: Semanal

Carga horaria por clase: 120 min

Presentación

El curso Introducción a DataOps propone una aproximación conceptual a las prácticas, principios y desafíos asociados a la gestión moderna de datos en organizaciones. Está orientado a personas que necesitan comprender cómo mejorar la confiabilidad, trazabilidad, calidad y disponibilidad de los datos, tanto desde una mirada técnica inicial como desde una perspectiva funcional o de gestión. A lo largo del curso se abordarán los fundamentos de DataOps, su relación con Data Engineering, DevOps, Analytics y MLOps, y su aplicación en el ciclo de vida completo de los datos: ingesta, procesamiento, almacenamiento, consumo, monitoreo y mejora continua. Se trabajará sobre problemáticas frecuentes como silos de información, falta de visibilidad, datos poco confiables, ausencia de ownership, pipelines frágiles, transformaciones no documentadas y mejora en la coordinación entre áreas técnicas y de negocio. El enfoque del curso es conceptual y aplicado, sin requerir programación ni trabajo técnico intensivo. Se busca que los participantes puedan incorporar vocabulario, criterios y marcos de análisis para participar en iniciativas de datos, comprender el funcionamiento general de pipelines, identificar riesgos operativos y reconocer buenas prácticas vinculadas con calidad, observabilidad, gobernanza, reproducibilidad y cultura DataOps. Al finalizar, los participantes contarán con una visión integral de DataOps y estarán en condiciones de identificar anti-patrones, proponer métricas básicas de madurez, reconocer responsabilidades clave y definir primeros pasos de mejora para entornos de datos reales.

Objetivos Generales

Brindar una introducción conceptual a DataOps, sus principios, prácticas y beneficios, para comprender cómo mejorar la calidad, confiabilidad, trazabilidad y operación de los datos en organizaciones.

Objetivos Específicos

- Comprender qué es DataOps, qué problemáticas aborda y cuál es su propósito dentro de una organización.
- Reconocer la relación entre DataOps, Data Engineering, DevOps, Analytics y MLOps.
- Identificar las principales etapas del ciclo de vida de los datos: ingesta, procesamiento, transformación, almacenamiento y consumo.
- Comprender el rol central de los pipelines de datos en una estrategia DataOps.
- Reconocer errores frecuentes en pipelines y su impacto en la operación y en el negocio.
- Incorporar criterios básicos de calidad de datos, incluyendo completitud, validez, consistencia, frescura y unicidad.

- Comprender la importancia del testing de datos, la observabilidad, el monitoreo, las alertas y los SLAs de datos.
- Identificar prácticas vinculadas con versionado, CI/CD, gestión de entornos, automatización y reproducibilidad.
- Comprender conceptos básicos de gobernanza de datos, incluyendo data lineage, catálogo de datos, roles, seguridad, control de accesos y compliance.
- Diferenciar DataOps de MLOps y reconocer sus puntos de contacto.
- Identificar anti-patrones comunes en entornos de datos y proponer acciones de mejora.
- Reconocer métricas básicas para evaluar la madurez y efectividad de prácticas DataOps.
- Analizar la dimensión organizacional y cultural de DataOps, incluyendo colaboración entre equipos técnicos y áreas de negocio.
- Aplicar los conceptos del curso a casos prácticos y ejercicios situacionales.

Destinatarios

- Perfiles técnicos y funcionales, pero no se requieren conocimientos avanzados de programación, infraestructura ni ciencia de datos.

Plan de Estudios

Módulo 1: Fundamentos de DataOps y Ciclo de Vida de Datos

- Definición y propósito de DataOps
- Problemáticas que aborda: silos, falta de visibilidad, datos poco confiables
- Relación con Data Engineering, DevOps y Analytics
- Principios de DataOps (automatización, colaboración, calidad, feedback continuo)
- Ingesta de datos (batch vs streaming)
- Procesamiento y transformación (ETL vs ELT)
- Almacenamiento (data lake vs data warehouse)
- Consumo de datos (BI, APIs, Machine Learning)

Módulo 2: Pipelines de Datos y Calidad de Datos

- Concepto de pipeline de datos y su centralidad en DataOps
- Orquestación vs ejecución
- Manejo de dependencias entre pipelines
- Errores comunes y su impacto en el negocio
- Dimensiones de calidad (completitud, validez, consistencia, frescura, unicidad)
- Testing de datos en pipelines
- Observabilidad: monitoreo, alertas y detección temprana
- SLAs de datos: definición e importancia

Módulo 3: Integración con Prácticas DevOps y Reproducibilidad

- Versionado aplicado a datos y transformaciones
- Conceptos de CI/CD en pipelines de datos
- Gestión de entornos (desarrollo, testing, producción)
- Automatización de procesos

- Por qué la reproducibilidad es crítica en datos
- Concepto de contenedores como mecanismo de portabilidad
- Entornos aislados y su importancia
- Infrastructure as Code: concepto y relevancia

Módulo 4: Gobernanza de Datos y MLOps

- Data lineage (trazabilidad): de dónde viene cada dato
- Catálogo de datos: qué es y qué habilita
- Roles y responsabilidades (data owners, stewards, engineers)
- Seguridad y control de accesos
- Compliance y regulaciones
- Diferencias y relación entre DataOps y MLOps
- Pipeline de Machine Learning: datos → modelo → despliegue → monitoreo
- MLFlow como herramienta homologada en la organización

Módulo 5: Anti-patrones y Métricas de Madurez DataOps

- Pipelines sin ownership claro
- Datos sin SLAs ni contratos
- Transformaciones no documentadas
- Ausencia de testing de datos
- Dependencias no mapeadas entre equipos
- Cómo medir si DataOps está funcionando: indicadores clave
- Frecuencia de incidentes de calidad de datos
- MTTR (Mean Time To Recovery) de pipelines
- Modelos de madurez DataOps

Módulo 6: Organización y Cultura DataOps

- Trabajo en equipos multidisciplinarios
- Interacción entre negocio y tecnología en contexto de datos
- Enfoque en time-to-value
- El rol de metodología en DataOps: cómo acompañar, qué facilitar
- Definition of Done para stories de datos
- Cómo integrar calidad de datos en el backlog y las ceremonias

Módulo 7: Casos Prácticos y Ejercicios Situacionales

- Problemas comunes en entornos de datos reales
- Buenas prácticas aplicadas
- Ejercicio práctico: dado un escenario, identificar anti-patrones, proponer métricas y definir primeros pasos de mejora
- Simulación de reuniones técnicas

Módulo 8: Integración, Repaso y Evaluación Final

- Repaso integral de todos los conceptos
- Consolidación del vocabulario técnico
- Discusión sobre aplicación en la organización
- Evaluación final de conocimientos

UTN.BA Campus Online — Universidad Tecnológica Nacional, Facultad Regional Buenos Aires

Mozart 2300, C1407 CABA, Argentina • +54 11 4867-7500 • campus@frba.utn.edu.ar

Documento generado el 24/07/2026 06:52 • UUID: b82cb389-822b-499e-8e7e-1693cd49b543