



# CURSO DATA ENGINEERING ON MICROSOFT AZURE

Plan de estudio



educación 





---

## Nuestro propósito

**Transformar positivamente la vida de las personas.**


Queremos que seas protagonista en la transformación que estamos viviendo. Por eso, nos comprometemos a capacitarte para que estés al día con las necesidades digitales actuales.

Te invitamos a trabajar en conjunto para que descubras tu mejor versión y la potencies. Anímate, toma las riendas de tu futuro.







Code your future!



## Contenido del curso



Aprende sobre la ingeniería de datos en base al trabajo con soluciones analíticas en tiempo real utilizando tecnologías de plataforma de datos de Azure.



## ¿Qué aprenderás?

- Procesamiento y Almacenamiento.
- Consultas interactivas SQL sin servidor.
- Transformación de datos en Azure Databricks.
- Carga de datos con Apache Spark.
- Ingesta en el almacén de datos.
- Data Factory y Synapse Pipelines.
- Integración de datos.
- HTAP con Azure Synapse Link.
- Seguridad de un extremo a otro.
- Procesamiento de transmisiones en tiempo real.
- Solución de procesamiento de transmisión.
- Marcas de agua para eliminar datos obsoletos.

## Plan de estudios

1

### Explore las opciones de computación y almacenamiento para cargas de trabajo de ingeniería de datos

- Introducción a Azure Synapse Analytics
- Describir Azure Databricks
- Introducción al almacenamiento de Azure Data Lake
- Describir la arquitectura de Delta Lake
- Trabajar con flujos de datos mediante Azure Stream Analytics
- Explore las opciones de computación y almacenamiento para cargas de trabajo de ingeniería de datos
- Combine la transmisión y el procesamiento por lotes con una única canalización
- Organice el lago de datos en niveles de transformación de archivos
- Indexación de almacenamiento de lago de datos para la aceleración de consultas y cargas de trabajo
- Describir Azure Synapse Analytics
- Describir Azure Databricks
- Describir el almacenamiento de Azure Data Lake
- Describir la arquitectura de Delta Lake
- Describir Azure Stream Analytics

2

### Ejecute consultas interactivas con grupos de SQL sin servidor de Azure Synapse Analytics

- Explore las capacidades de los grupos SQL sin servidor de Azure Synapse
- Consultar datos en el lago mediante grupos de SQL sin servidor de Azure Synapse
- Crear objetos de metadatos en grupos SQL sin servidor de Azure Synapse
- Proteja los datos y administre a los usuarios en los grupos SQL sin servidor de Azure Synapse
- Ejecute consultas interactivas utilizando grupos de SQL sin servidor
- Consultar datos de Parquet con grupos SQL sin servidor
- Crea tablas externas para archivos Parquet y CSV
- Cree vistas con grupos de SQL sin servidor
- Acceso seguro a los datos en un lago de datos cuando se utilizan grupos de SQL sin servidor
- Configure la seguridad del lago de datos mediante el control de acceso basado en roles (RBAC) y la lista de control de acceso
- Comprender las capacidades de los grupos de SQL sin servidor de Azure Synapse
- Consultar datos en el lago mediante grupos de SQL sin servidor de Azure Synapse
- Crear objetos de metadatos en grupos SQL sin servidor de Azure Synapse
- Proteja los datos y administre a los usuarios en los grupos SQL sin servidor de Azure Synapse

- Describir Azure Databricks
- Leer y escribir datos en Azure Databricks
- Trabajar con DataFrames en Azure Databricks
- Trabajar con métodos avanzados de DataFrames en Azure Databricks
- Exploración y transformación de datos en Azure Databricks
- Use DataFrames en Azure Databricks para explorar y filtrar datos
- Almacene en caché un DataFrame para consultas posteriores más rápidas
- Eliminar datos duplicados
- Manipular valores de fecha / hora
- Eliminar y cambiar el nombre de las columnas DataFrame
- Agregar datos almacenados en un DataFrame
- Describir Azure Databricks
- Leer y escribir datos en Azure Databricks
- Trabajar con DataFrames en Azure Databricks
- Trabajar con métodos avanzados de DataFrames en Azure Databricks

## 4

### **Explore, transforme y cargue datos en el almacén de datos usando Apache Spark**

- Comprender la ingeniería de big data con Apache Spark en Azure Synapse Analytics
- Ingesta datos con los cuadernos de Apache Spark en Azure Synapse Analytics
- Transforme datos con DataFrames en Apache Spark Pools en Azure Synapse Analytics

- Integrar grupos de SQL y Apache Spark en Azure Synapse Analytics
- Explore, transforme y cargue datos en el almacén de datos usando Apache Spark
- Realizar exploración de datos en Synapse Studio
- Ingesta datos con cuadernos Spark en Azure Synapse Analytics
- Transforme datos con DataFrames en grupos de Spark en Azure Synapse Analytics
- Integrar grupos de SQL y Spark en Azure Synapse Analytics
- Describir la ingeniería de big data con Apache Spark en Azure Synapse Analytics
- Ingesta datos con los cuadernos de Apache Spark en Azure Synapse Analytics
- Transforme datos con DataFrames en Apache Spark Pools en Azure Synapse Analytics
- Integrar grupos de SQL y Apache Spark en Azure Synapse Analytics

## 5

### Ingesta y carga datos en el almacén de datos

- Utilice las mejores prácticas de carga de datos en Azure Synapse Analytics
- Ingestión a escala de petabytes con Azure Data Factory
- Ingesta y carga datos en el almacén de datos
- Realice la ingestión a escala de petabytes con Azure Synapse Pipelines
- Importe datos con PolyBase y COPY usando T-SQL

- Utilice las mejores prácticas de carga de datos en Azure Synapse Analytics
- Después de completar este módulo, los estudiantes podrán:
- Utilice las mejores prácticas de carga de datos en Azure Synapse Analytics
- Gestión a escala de petabytes con Azure Data Factory

**6**

## **Transformar datos con Azure Data Factory o Azure Synapse Pipelines**

- Integración de datos con Azure Data Factory o Azure Synapse Pipelines
- Transformación sin código a escala con Azure Data Factory o Azure Synapse Pipelines
- Transformar datos con Azure Data Factory o Azure Synapse Pipelines
- Ejecute transformaciones sin código a escala con Azure Synapse Pipelines
- Cree una canalización de datos para importar archivos CSV con formato deficiente
- Crear flujos de datos de mapeo
- Después de completar este módulo, los estudiantes podrán:
- Realice la integración de datos con Azure Data Factory
- Realice una transformación sin código a escala con Azure Data Factory

**7**

## **Orquestar el movimiento y la transformación de datos en Azure Synapse Pipelines**

- Organice el movimiento y la transformación de datos en Azure Data Factory
- Orquestar el movimiento y la transformación de datos en Azure Synapse Pipelines
- Integre datos de portátiles con Azure Data Factory o Azure Synapse Pipelines
- Después de completar este módulo, los estudiantes podrán:
  - Organice el movimiento y la transformación de datos en Azure Synapse Pipelines



## Seguridad de un extremo a otro con Azure Synapse Analytics

- Proteger un almacén de datos en Azure Synapse Analytics
- Configurar y administrar secretos en Azure Key Vault
- Implementar controles de cumplimiento para datos confidenciales
- Seguridad de un extremo a otro con Azure Synapse Analytics
- Asegure la infraestructura de soporte de Azure Synapse Analytics
- Asegure el área de trabajo de Azure Synapse Analytics y los servicios administrados
- Proteger los datos del área de trabajo de Azure Synapse Analytics
- Después de completar este módulo, los estudiantes podrán:

- Proteger un almacén de datos en Azure Synapse Analytics
- Configurar y administrar secretos en Azure Key Vault
- Implementar controles de cumplimiento para datos confidenciales

9

## Soporte del procesamiento analítico transaccional híbrido (HTAP) con Azure Synapse Link

- Diseñar procesamiento transaccional y analítico híbrido con Azure Synapse Analytics
- Configurar Azure Synapse Link con Azure Cosmos DB
- Consultar Azure Cosmos DB con grupos de Apache Spark
- Consultar Azure Cosmos DB con grupos de SQL sin servidor
- Soporte de procesamiento analítico transaccional híbrido (HTAP) con Azure Synapse Link
- Configurar Azure Synapse Link con Azure Cosmos DB
- Consultar Azure Cosmos DB con Apache Spark para Synapse Analytics
- Consultar Azure Cosmos DB con un grupo de SQL sin servidor para Azure Synapse Analytics
- Diseñar procesamiento transaccional y analítico híbrido con Azure Synapse Analytics
- Configurar Azure Synapse Link con Azure Cosmos DB
- Consultar Azure Cosmos DB con Apache Spark para Azure Synapse Analytics
- Consultar Azure Cosmos DB con SQL sin servidor para Azure Synapse Analytics

10

## Procesamiento de transmisión en tiempo real con Stream Analytics

- Habilite la mensajería confiable para aplicaciones de Big Data con Azure Event Hubs
- Trabajar con flujos de datos mediante Azure Stream Analytics
- Ingesta flujos de datos con Azure Stream Analytics
- Procesamiento de transmisión en tiempo real con análisis de transmisión
- Utilice Stream Analytics para procesar datos en tiempo real de Event Hubs
- Utilice las funciones de ventana de Stream Analytics para crear agregados y resultados en Synapse Analytics
- Escale el trabajo de Azure Stream Analytics para aumentar el rendimiento mediante la partición
- Repartir la entrada de flujo para optimizar la paralelización
- Habilite la mensajería confiable para aplicaciones de Big Data con Azure Event Hubs
- Trabajar con flujos de datos mediante Azure Stream Analytics
- Ingesta flujos de datos con Azure Stream Analytics

11

## Cree una solución de procesamiento de transmisión con Event Hubs y Azure Databricks

- Procesar datos de transmisión con transmisión estructurada de Azure Databricks
- Cree una solución de procesamiento de transmisión con Event Hubs y Azure Databricks

- Explore las características y usos clave de la transmisión estructurada
- Transmita datos desde un archivo y escríbalos en un sistema de archivos distribuido
- Utilice ventanas deslizantes para agregar fragmentos de datos en lugar de todos los datos
- Aplicar marcas de agua para eliminar datos obsoletos
- Conectarse a transmisiones de lectura y escritura de Event Hubs
- Procesar datos de transmisión con transmisión estructurada de Azure Databricks

## Modalidad del Curso

### Duración

8 semanas / 32 h

### Frecuencia semanal

2 encuentros de 2 h

### Modalidad

Online en vivo

### Grupos reducidos

Promedio 15 personas

## Nivel: Principiante



- Principiante
- Intermedio
- Avanzado
- Experto

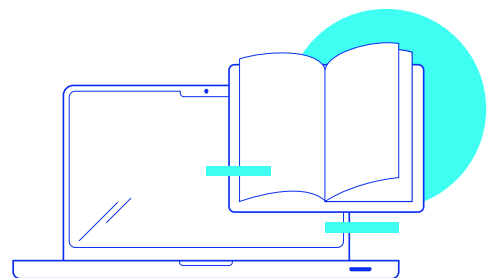
## Requisitos

Es recomendable tener una base sobre:

- **Microsoft Azure Fundamentals**
- **Microsoft Azure Data Fundamentals**

## Dedicación fuera de clase

Además de las horas de clase, recomendamos que inviertas 4 h semanales extras para realizar los desafíos complementarios, estudiar el material de lectura y completar los exámenes del Alumni.



## ¿Cómo será tu experiencia?



### Aprender haciendo

Ejercita y pon en práctica lo estudiado.



### Trabajo en equipo

Une esfuerzos y potencia los resultados.



### Clases grabadas

Consúltalas las veces que quieras.



### Profesores expertos

Aprende de gigantes de la industria.



### Asistente académico

Recibe soporte dentro y fuera de clase.



### Plataforma Alumni

Encuentra recursos, materiales y clases.

## ¿Por qué Educación IT?



### IT Créditos

Gana puntos al aprobar los exámenes de los cursos. Luego, podrás canjearlos por nuevos cursos sin costo alguno. Los IT Créditos que acumules no vencen ni se devalúan.



### Garantía de aprendizaje

Si necesitas reforzar conceptos, recuperar clases o no estás satisfecho, ¡vuelve a tomar el curso sin ningún costo! Puede ser de forma total o parcial.



### Comunidad en Discord


Mantente en contacto con la comunidad de EducaciónIT a través de nuestro servidor de Discord. Podrás hablar con tus compañeros, profesores, asistentes académicos y soporte.



## Preguntas frecuentes



**Si me pierdo una o más clases, ¿puedo recuperarlas?**




Todas las clases quedan grabadas de por vida en tu plataforma Alumni. ¡Siempre podrás volver a verlas cada vez que lo necesites!

**¿Cómo voy a aprender?**


Te enfrentarás a situaciones de trabajo reales, en donde tendrás que aplicar lo aprendido de forma individual y en equipo. Por medio de la prueba y el error, irás superando desafíos y obteniendo nuevas habilidades que luego podrás aplicar en el ámbito laboral.

**¿Cómo son las clases online en vivo?**

Las clases duran entre 2 y 3 horas de lunes a viernes (sábados 3 o 4 hs) y se desarrollan de forma online en vivo en aulas virtuales, donde vas a poder interactuar con el instructor y tus compañeros.



Manejamos cupos reducidos para que puedas tener un seguimiento más personalizado durante tu aprendizaje.





[www.educacionit.com](http://www.educacionit.com)



@educacionit

---