

Introducción a Databricks y el ecosistema Spark

Agenda



IMMUNE
TECHNOLOGY INSTITUTE



Agenda



Databricks Onboarding

- Fundación y desarrollo de Databricks: historia, capitalización, número de clientes, valoración y ranking.
- Qué es Databricks y cómo se diferencia de otros entornos de Spark.
- Arquitectura general y componentes clave.
- Interfaz de usuario y navegación básica.

Agenda



Databricks - Data Engineering

- Repaso conceptos básicos de Spark
- Introducción a Delta Lake: conceptos y beneficios.
- Gestión de metadatos y versionamiento de datos con Delta Lake.
- Orquestación de procesos



Databricks – Data Analytics

- Herramientas de análisis de datos en Databricks.
- Creación de dashboards y visualizaciones interactivas.
- Integración con herramientas externas de BI y visualización.

Agenda



Laboratorio

- Ingesta de datos en Databricks Bronze, Silver, Gold
- Creación de dashboards



Conclusión

- Repaso de los conceptos clave cubiertos en el workshop.
- Mejores prácticas para aprovechar Databricks en proyectos de Big Data.
- Recursos adicionales para seguir aprendiendo.
- Q&A

Jesus Arnau Villar



Senior Cloud Data Engineer



Apache Spark 3 Certified



Databricks Solutions Architect Champion



<https://www.linkedin.com/in/jesus-arnau-84494441/>



<https://github.com/jarnawer>



databricks

IMMUNE
TECHNOLOGY INSTITUTE

