# Databricks fundamentals 101 Training Day

Agenda:

# **10:00 – 16:00 Databricks Fundamentals**

**Purpose**

The goal of this training session it to provide a high-level overview of the Databrick platform and an opportunity to get hands on the technology.

# **10:00 – 12:00 Presentation: Overview of Azure Databricks**

The presentation is designed to highlight some of the core values of using Databricks platform and features it provide for various personas: Data Engineer, Data Scientist (end-to-end) and Data Analyst.  Demo on ADF running a Databricks Job.

# **12:00 – 13:00 Lunch**

# **13:00 – 16:00** **Hands-on lab:  Databricks Fundamentals**

The lab session is tailored toward giving a hands-on experience on the platform.  Topics will be covered are as follows:

Core topics:

* Create/delete/collaborative workspace
* Folder structure
* Using notebook
* Various languages support (R, Python, SQL)
* Create/delete/run a job on a cluster
* Role-based access control
* Logs
* Various API
* Integration with Azure Devops

Data Engineer:

* Load data from Blob (make aware of the ADLS)
* Run simple ELT transformation using Spark (Spark SQL)/Databrick Delta

Data Scientist:

* Create a simple experiment (R/Python) (effort required to move to Databrick (Scikit learn)
* Train/test on a cluster

Data Analyst:

* Data visualization (various charts/graph)
* Visualize ML experiments

Discussion on follow-up use cases & scenarios around Databricks

# **Prerequisite**

* Azure Subscription