**A black text on a white background

Description automatically generated**   
 **Masterclass**

**Key Takeaways ✅**

**🔁 Databricks File System (DBFS)**

* DBFS is a distributed file system within Databricks.
* Allows use of file semantics (e.g., /FileStore/raw.file.csv).
* Integrates with Delta Lake for storage access via URLs.

**🔐 Accessing Azure Data Lake**

**Hardcoded Authentication (Not Recommended for Prod)**

* Use **Service Principal** for secure access.
* Setup via Azure Entra ID → App Registration.
  + Required: Application\_ID, Tenant\_ID, and secret.
* Assign role: **Storage Blob Data Contributor**.

**🛠️ Databricks Utilities**

* dbutils.fs.ls("<path>") – List files.
* dbutils.widgets.text("param", "default") – Add parameter.
* dbutils.widgets.get("param") – Get parameter value.

**🔐 Secrets Management using Azure Key Vault**

* Create Key Vault in Azure and store secrets.
* Connect to Databricks via **secret scopes**:
  + Create secret scope via https://<databricks-instance>#secrets/createScope
* Use: dbutils.secret.get(scope="scope\_name", key="secret\_key")

**📥 Reading and Transforming Data with PySpark**

* Read CSV with schema inference:
* spark.read.format("csv").option("header", True).option("inferSchema", True)
* Common transformations:
  + split(), withColumn(), lit(), cast().

**🔄 Delta Lake Concepts**

* **Delta Format** includes .crc, .json for transaction logs.
* %run "/path/to/notebook" to inherit variables/functions.

**✨ Key Delta Features**

* **Versioning**:  
   DESCRIBE HISTORY table\_name
* **Time Travel**:  
   RESTORE TABLE table\_name TO VERSION AS OF 1
* **Optimization**:  
   OPTIMIZE table\_name   
   OPTIMIZE table\_name ZORDER BY (column)
* **Cleanup**:  
   VACUUM table\_name

**📡 Autoloader – Streaming Data Ingestion**

* Read streaming data from Azure Data Lake using cloudFiles:
* spark.readStream.format("cloudFiles")...
* Write to Delta with checkpointing and schema merge:
* df.writeStream.format("delta")...

**❤️ I love data fam!**

Forever grateful for this growing community and always excited to keep learning and building!