# Azure Data Factory Hands-on Lab Guide

## Lab 1: Create ADF Instance

### Steps:

1. Sign in to the [Azure Portal](https://portal.azure.com).
2. Search for **Data Factories** in the search bar.
3. Click **+ Create**.
4. Provide the following details:
   * **Subscription**: Select your subscription.
   * **Resource Group**: Create new or choose existing.
   * **Region**: Choose a supported region.
   * **Name**: Enter a globally unique name for your Data Factory.
5. Select **Version V2** (recommended).
6. Click **Review + Create** → **Create**.
7. Wait for deployment to complete and click **Go to resource**.

## Lab 2: Linked Services Setup

### Steps:

1. Open your Data Factory resource and click **Author & Monitor**.
2. In the ADF Studio, go to the **Manage** hub.
3. Select **Linked Services** → **+ New**.
4. Choose **Azure Data Lake Storage Gen2**:
   * Provide account name, authentication method, and test connection.
   * Click **Create**.
5. Repeat step 3 for **Azure SQL Database**:
   * Enter server name, database name, authentication details.
   * Test connection and click **Create**.

## Lab 3: Dataset Creation

### Steps:

1. In ADF Studio, go to the **Author** hub.
2. Under **Datasets**, click **+ New Dataset**.
3. For **CSV Dataset**:
   * Select **Azure Data Lake Storage Gen2** as linked service.
   * Choose a sample CSV file location.
   * Provide schema if available or import from file.
   * Save as Input\_CSV\_Dataset.
4. For **JSON Dataset**:
   * Select **Azure Blob Storage** or ADLS as linked service.
   * Browse to JSON file.
   * Import schema if available.
   * Save as Input\_JSON\_Dataset.

## Lab 4: Pipeline Build

### Steps:

1. In ADF Studio, go to the **Author** hub.
2. Click **+ New Pipeline**.
3. Drag and drop a **Copy Data** activity onto the canvas.
4. Configure **Source**:
   * Select Input\_CSV\_Dataset.
5. Configure **Sink**:
   * Select **Azure SQL Database linked service**.
   * Choose or create a target dataset (Output\_SQL\_Dataset).
6. Map columns:
   * Click **Mapping** tab.
   * Auto-map or manually map source to target columns.
7. Validate pipeline and click **Publish All**.
8. Trigger pipeline manually by selecting **Add Trigger → Trigger Now**.

## Lab 5: Pipeline Monitoring

### Steps:

1. Navigate to the **Monitor** hub in ADF Studio.
2. Review pipeline run history:
   * Check status: **Succeeded**, **Failed**, or **In Progress**.
   * View activity run details.
3. For a failed run:
   * Click the failed pipeline instance.
   * Expand activity details to review error message.
   * Identify root cause (e.g., wrong schema, missing credentials).
4. Retry failed run:
   * Fix issue in pipeline/dataset/linked service.
   * Re-run pipeline using **Trigger Now**.