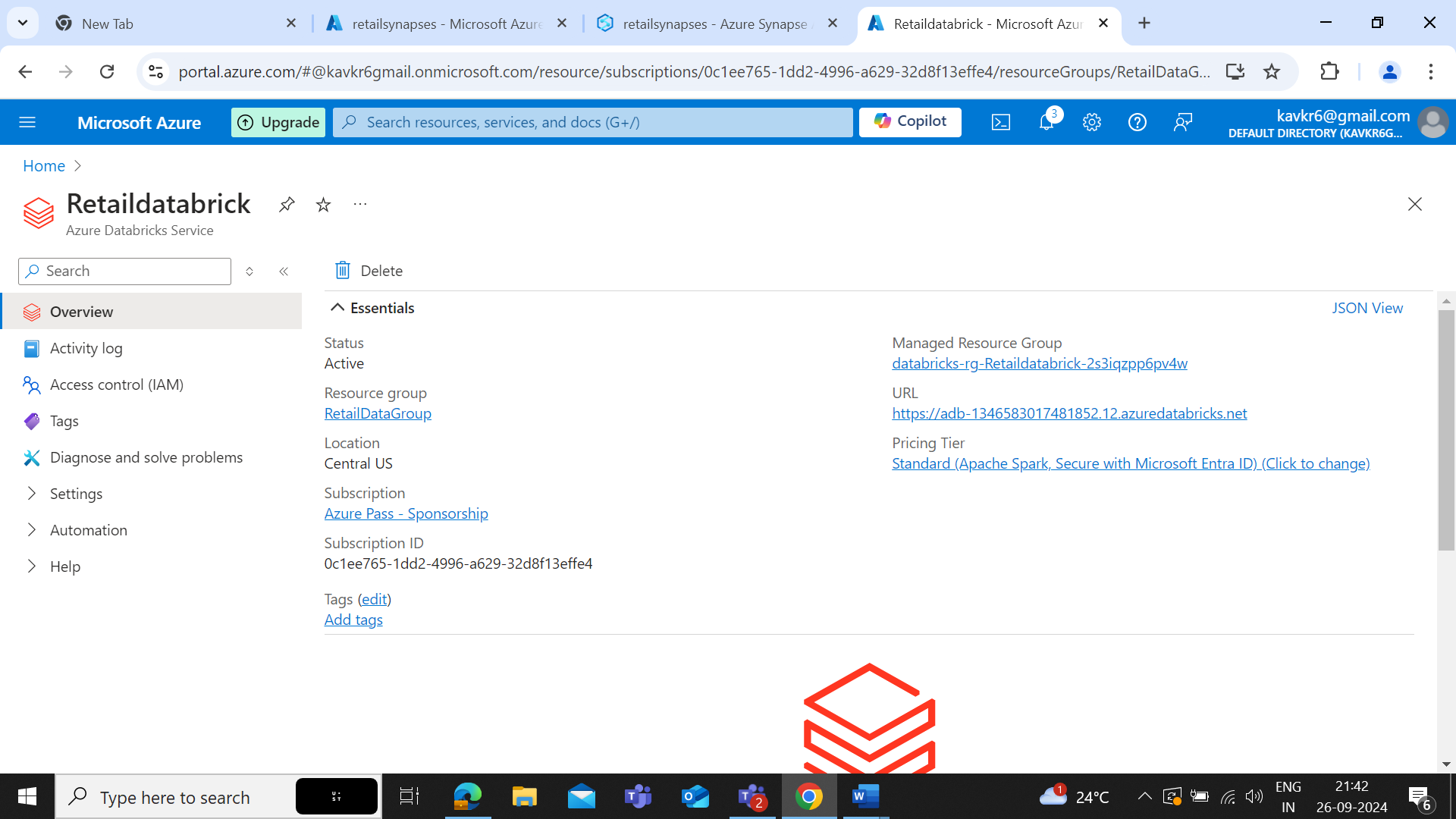
**CaseStudy-2**

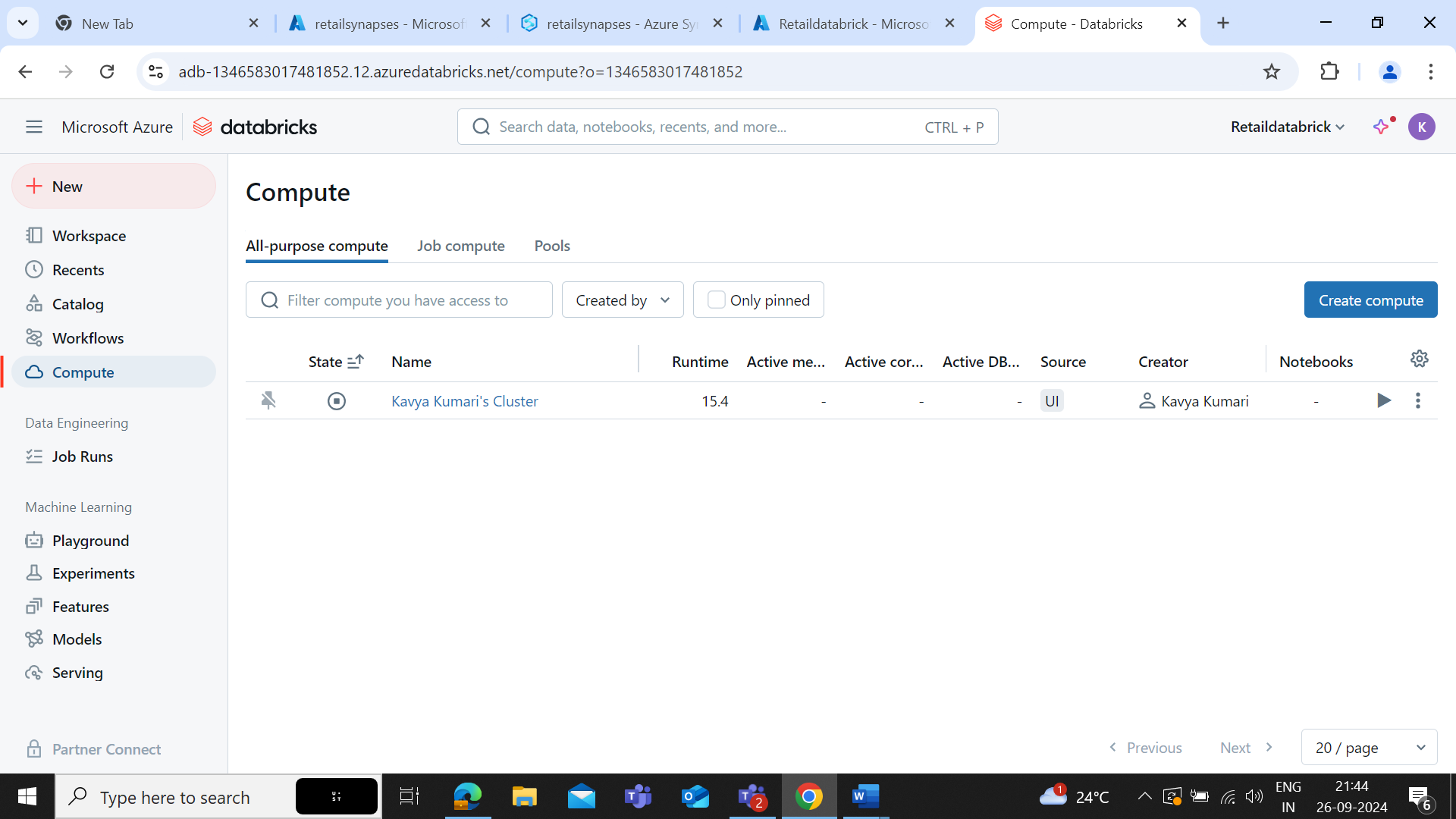
1. **Create Storage Accounts:**
   * **Storage Account Name:** xyzretailstorage
   * **Resource Group:** RetailDataGroup
   * **Region:** Choose an appropriate region (e.g., East US)
2. **Create Containers:**
   * **Source Container:** source-data
   * **Target Container:** transformed-data
3. **Upload Sample Data:**
   * Upload sales\_data.csv to the source-data container.

**2. Set Up Azure Databricks**

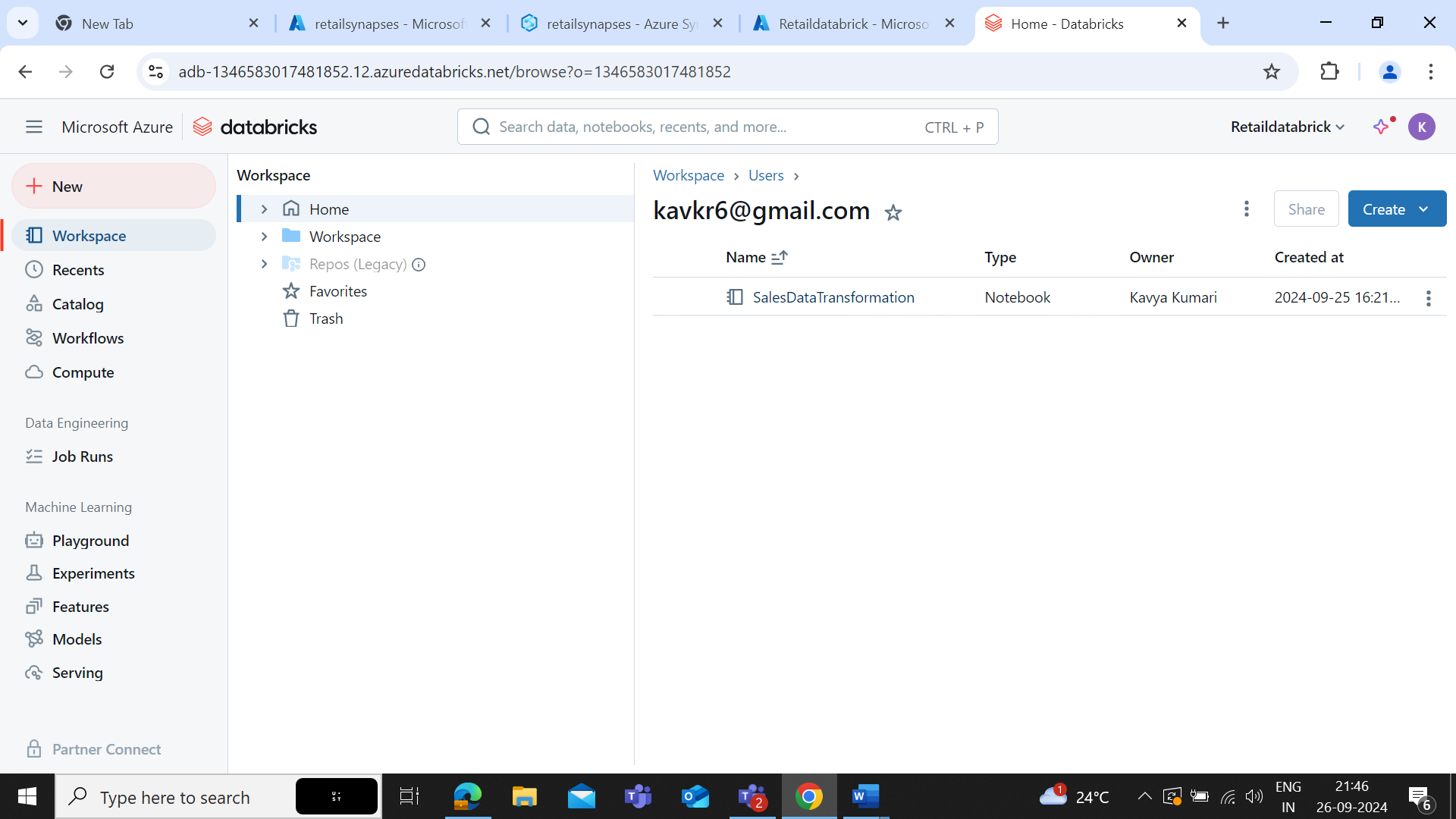
1. **Create an Azure Databricks Workspace:**
   * **Workspace Name:** RetailDatabricks
   * **Resource Group:** RetailDataGroup
   * **Region:** Same as storage account (e.g., East US)

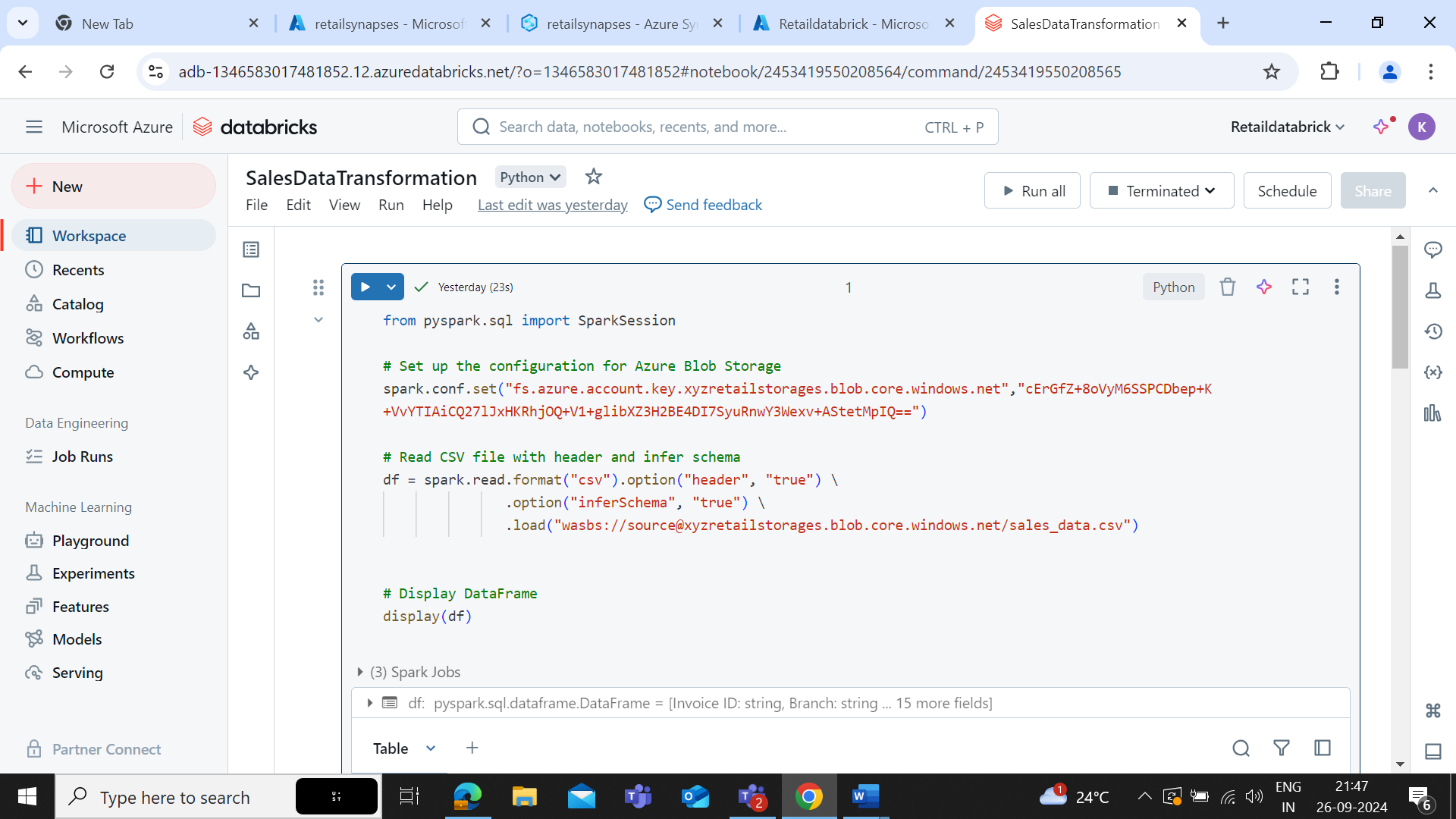


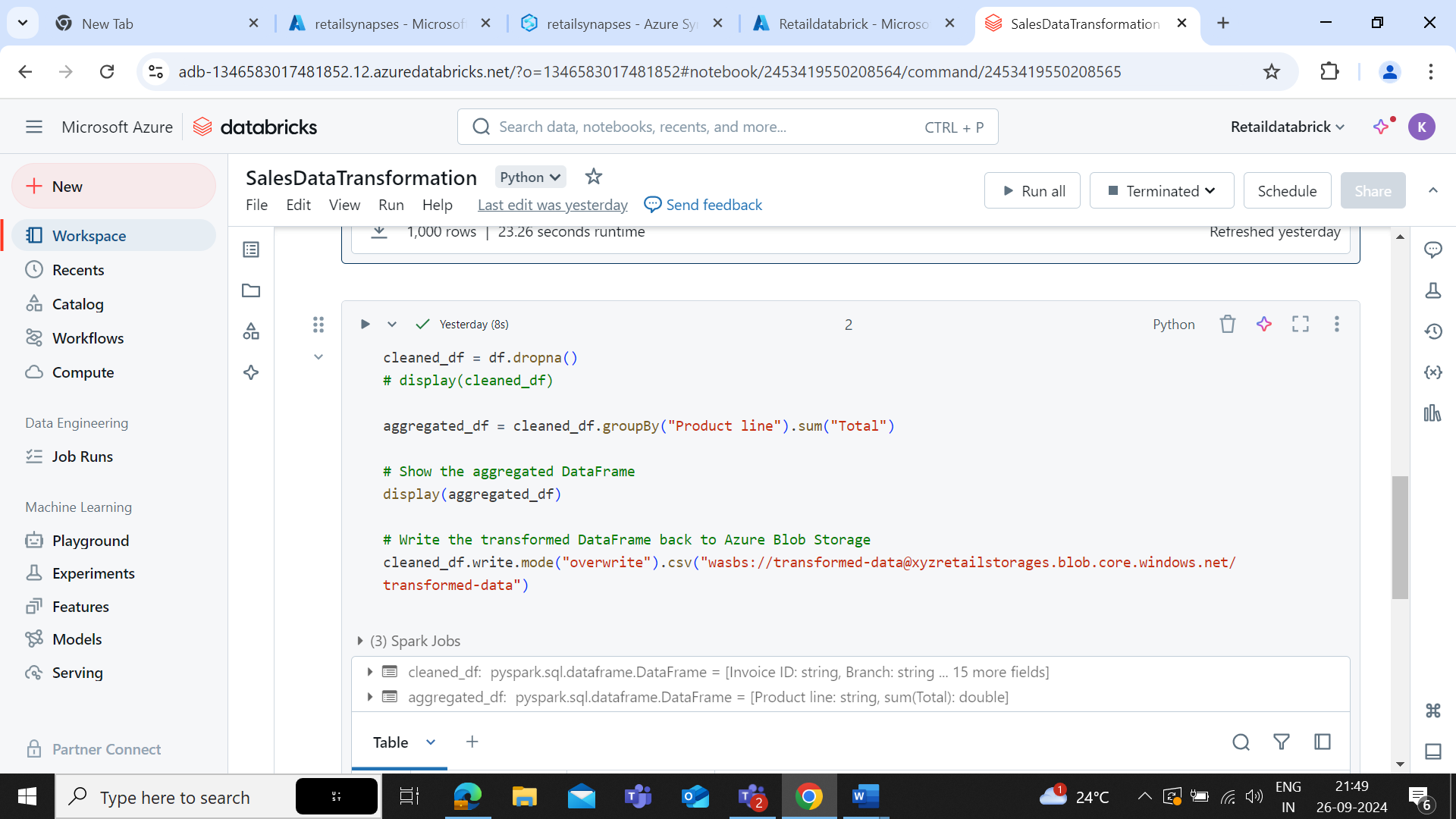
1. **Create a Cluster:**
   * Set up a cluster in Azure Databricks with appropriate configurations for processing the data.



1. **Create a Databricks Notebook:**
   * **Notebook Name:** SalesDataTransformation
   * Use PySpark or SQL to read the CSV file from Azure Blob Storage, perform transformations (e.g., cleaning, aggregating), and write the transformed data back to the transformed-data container.







A screenshot of a computer

Description automatically generated