**eCommerce Data Analytics with watsonx.data**

**Introduction**

This repository showcases a powerful **data analytics** asset built around **Watsonx.data**, an advanced **lakehouse** platform for managing and analyzing large datasets. As a unified data platform, **Watsonx.data** combines the best of both data lakes and data warehouses, enabling businesses to efficiently store, process, and analyze structured and unstructured data. With native support for **open table formats** like **Delta Lake**, **Hudi**, and **Iceberg**, **Watsonx.data** provides a high-performance, scalable environment for data transformations using **Spark** as the transformation engine.

Focusing on a real-world **eCommerce use case**, this asset demonstrates how **Watsonx.data** can streamline data storage, transformation, and analytics. The final insights are visualized in an interactive **Cognos Analytics** dashboard, driving data-driven business decisions.

**Demo Video**

To get an in-depth look at the entire workflow, from data transformation to final analytics, watch the demo video:

[Watch the Demo Video](https://ibm.ent.box.com/folder/297661873530)

**Why Catalogs Are Required**

**Watsonx.data**, as a **lakehouse** platform, provides an integrated cataloging feature that allows for efficient organization, management, and governance of vast datasets. By enabling data consistency, discoverability, and security, **Watsonx.data** makes it easier for businesses to manage complex data and extract valuable insights across both structured and unstructured data.

**Key Highlights**

The demo covers the following essential features of the asset:

1. **Entity-Relationship Diagram (ERD):**  
   A visual representation of the eCommerce data structure and relationships.
2. **Architecture Overview:**  
   Explanation of how **Watsonx.data** integrates with **Spark** and supports open table formats (Delta Lake, Hudi, Iceberg) for data storage and transformation.
3. **Data Transformation with Python:**  
   Python scripts demonstrating how data is transformed using Spark, with the processed data stored and managed in **Watsonx.data**.
4. **Cognos Analytics Dashboard:**  
   A dynamic dashboard built using **Cognos Analytics**, visualizing key business insights from the eCommerce data.

**Repository Contents**

* **Source Code:**
  + [git link](https://github.ibm.com/ecosystem-engineering-si-lab/E-Commerce-with-watsonx.data/tree/main/src) : Scripts for data transformation.
  + [git link](https://github.ibm.com/ecosystem-engineering-si-lab/E-Commerce-with-watsonx.data/tree/main/src/db2srcfiles): Sample DB2 source files for use in the demo.
* **Documentation & Resources:**
  + A Word document detailing the architecture, ERD, and full workflow.
  + A PowerPoint presentation used in the demo video.

[architecture git link](https://github.ibm.com/ecosystem-engineering-si-lab/E-Commerce-with-watsonx.data/tree/main/architecture) and [Demo ppt and One pager git link](https://github.ibm.com/ecosystem-engineering-si-lab/E-Commerce-with-watsonx.data/tree/main/ppt)

**How to Use**

To execute the demo and explore the functionality, follow the detailed instructions provided in the [word document.](https://github.ibm.com/ecosystem-engineering-si-lab/E-Commerce-with-watsonx.data/tree/main/Word%20document)

**Key Features**

* **Watsonx.data Lakehouse:**  
  **Watsonx.data** is a unified **lakehouse** platform that supports both data lakes and data warehouses, providing seamless integration for storage, processing, and analytics. Native support for **open table formats** (Delta Lake, Hudi, Iceberg) enables efficient data management.
* **High-Performance Data Transformations with Spark:**  
  Use **Spark** as the transformation engine for fast, scalable data processing.
* **Interactive Analytics with Cognos:**  
  Create rich, interactive dashboards and business reports with **Cognos Analytics**.
* **Real-World eCommerce Use Case:**  
  The asset demonstrates the use of **Watsonx.data** in an **eCommerce** business context, showcasing how the platform helps solve real-world business challenges.

**Additional Notes**

* Ensure you have access to **Watsonx.data** and **Cognos Analytics** for full functionality.
* For any questions or issues, feel free to raise an issue in this repository.