Functional Programming with Scala

Project Title: <u>Test Project</u>

Group Name: <u>Group 2</u>

ALQUIE, Emmy

IG5_Polytech

November 19, 2024

About the Project

- **Your Specific Contribution:**
  - Designed and implemented Scala code to connect Spark with Neo4j.
  - Created mock datasets for testing and validation of queries.

- **Methodology or Approach:**
  - Loaded data into Spark for transformation and processing.
  - Established a connection to Neo4j using the Neo4j Spark Connector.
  - Wrote nodes and relationships into the graph database.
  - Analyzed and filtered data using Spark SQL queries.

Challenges and Learning

- **Challenges Faced and Their Resolutions:**
  - Difficulty in configuring Neo4j and establishing a connection.
  - Resolved by troubleshooting the Neo4j URL and credentials.
  - Faced Spark job failures due to incorrect dataset formats; addressed by preprocessing the data.

- **Learning Gained:**
  - Gained understanding of graph databases and Cypher syntax.
  - Improved knowledge of Spark-Scala integrations for data processing.

- **Future Improvements:**
  - Automate data preprocessing to handle various dataset formats.
  - Optimize Cypher queries for better Neo4j performance.