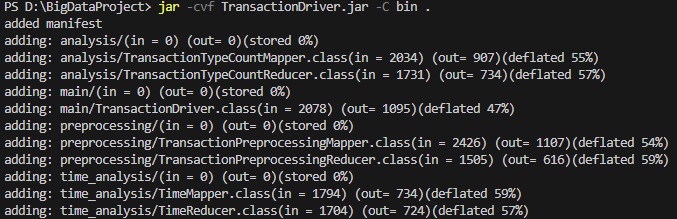
javac -d bin -classpath '.\lib\\*' src/analysis/\*.java src/time\_analysis/\*.java src/preprocessing/\*.java src/main/\*.java



jar -cvf TransactionDriver.jar -C bin .

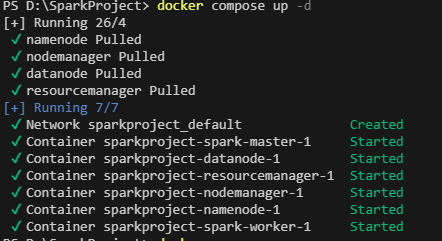


docker cp TransactionDriver.jar sparkproject-namenode-1:/tmp/TransactionDriver.jar

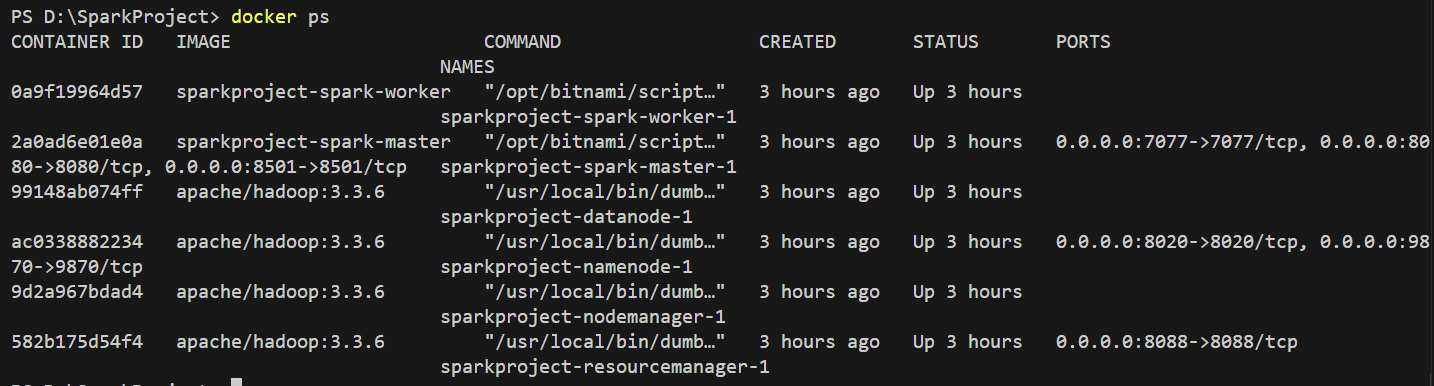


docker compose build

docker compose up -d



docker ps



docker cp .\transaction\_logs.csv sparkproject-namenode-1:tmp

docker exec -it sparkproject-namenode-1 bash

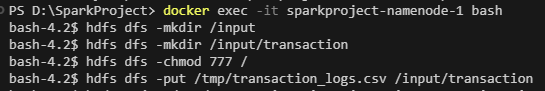
hdfs dfs -mkdir /input

hdfs dfs -mkdir /input/transaction

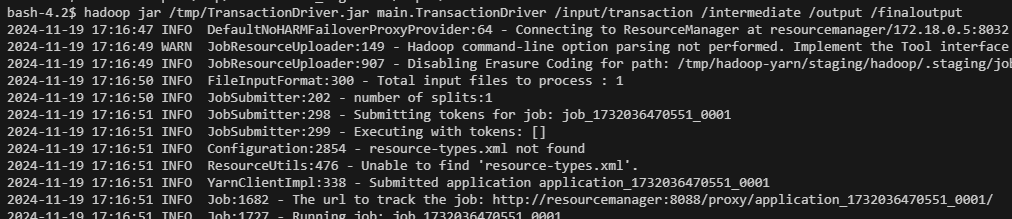
hdfs dfs -ls /

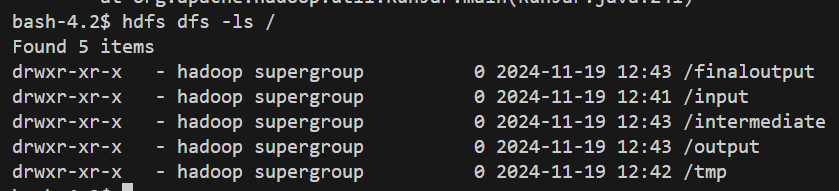
hdfs dfs -chmod 777 /

hdfs dfs -put /tmp/transaction\_logs.csv /input/transaction



hadoop jar /tmp/TransactionDriver.jar main.TransactionDriver /input/transaction /intermediate /output /finaloutput

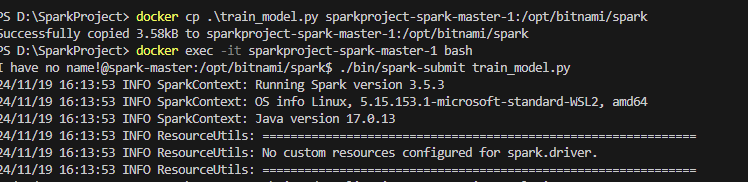




docker cp .\train\_model.py sparkproject-spark-master-1:/opt/bitnami/spark

docker exec -it sparkproject-spark-master-1 bash

./bin/spark-submit train\_model.py>> submit and train the train\_model





MAE: 271.0959136555383, MSE: 104259.43080545151, RMSE: 322.8922897894149

to verify go to hdfs and list the directory >> hdfs dfs -ls /

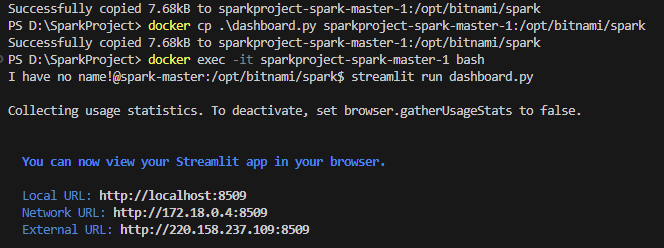
docker cp .\dashboard.py sparkproject-spark-master-1:/opt/bitnami/spark

docker exec -it sparkproject-spark-master-1 bash

pip install py4j

python -c "import py4j; print('Py4J is installed')"

streamlit run dashboard.py



Dashboard  
