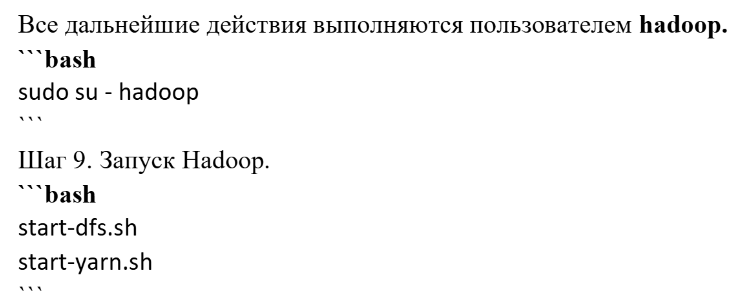
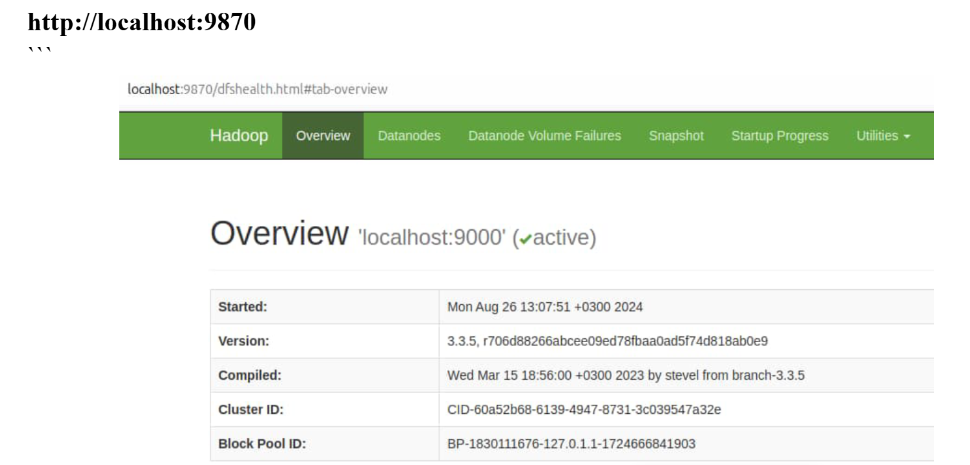
Заходим от пользователя devops и открываем терминал

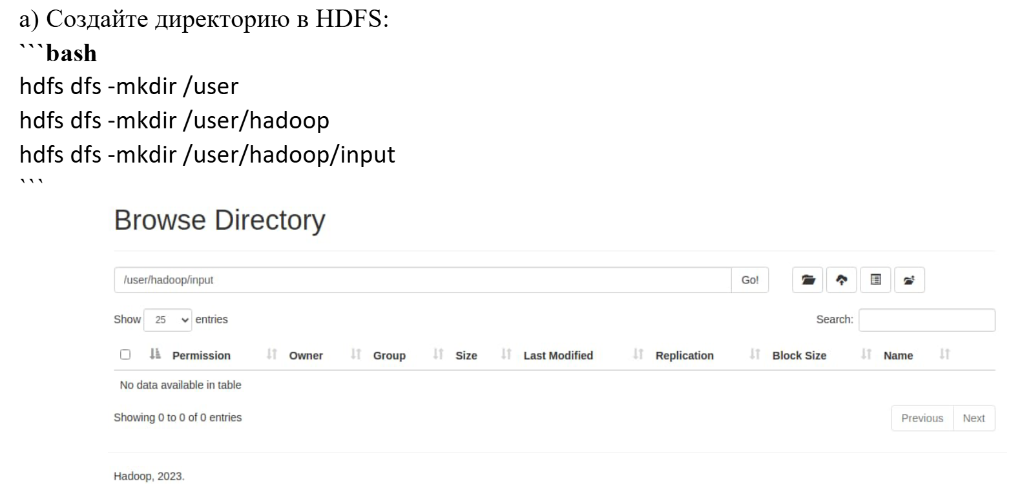
Необходимо выполнить sudo su hadoop (без -)



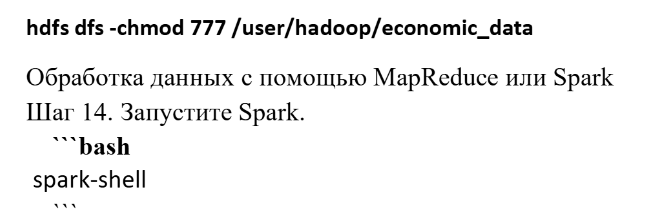
После запуска переходим в хост



Создаем новую директорию user 3, например



Далее необходимо предоставить доступ к директории

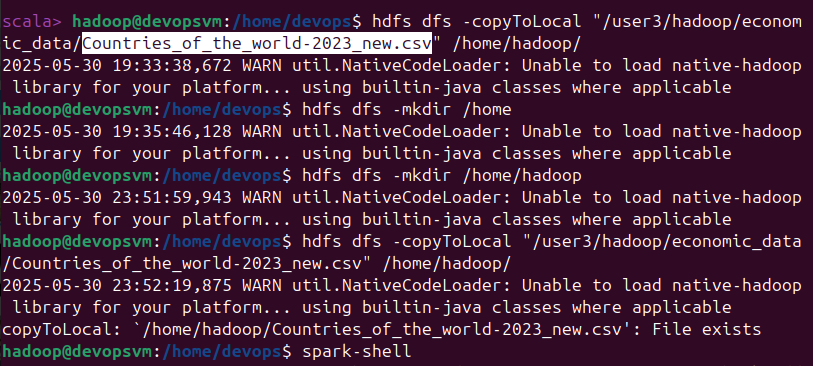


Можно загрузить файл вручную на сайте через upload

Создать еще одну локально /home/Hadoop

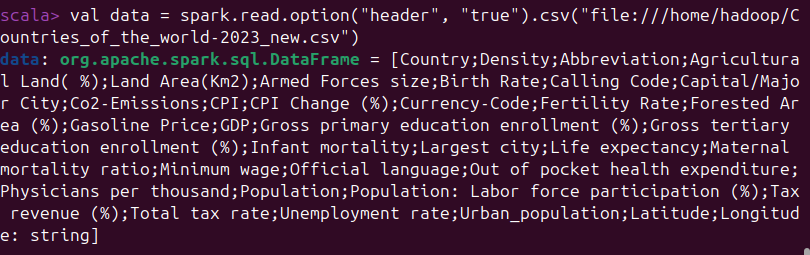
И выполнить команды для копирования туда файла

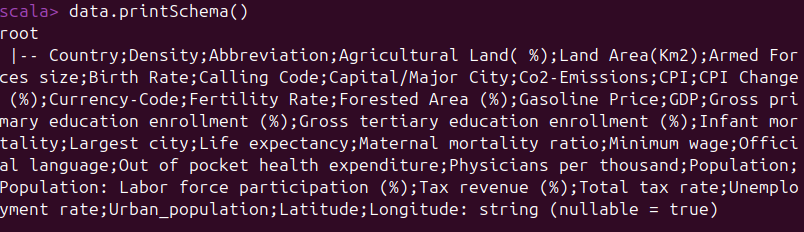
hdfs dfs -copyToLocal /3\_zadanie/hadoop/input/all\_stocks\_5yr.csv /home/hadoop/



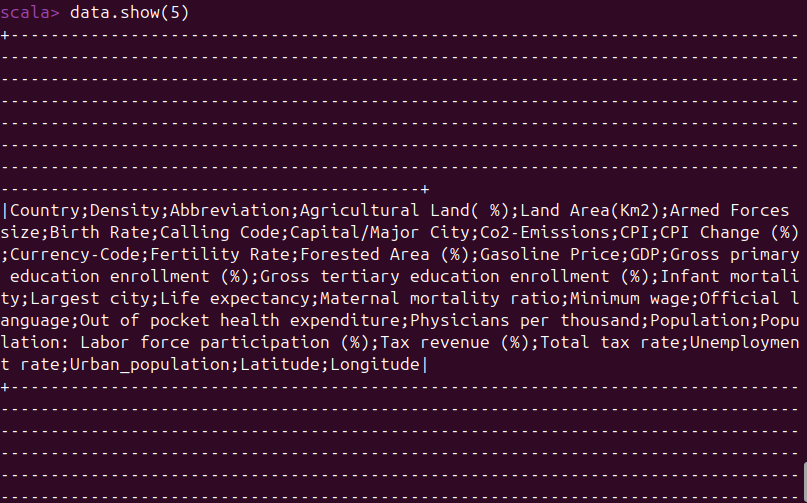
Теперь запускаем spark-shell для чтения файла и построения его схемы

val data = spark.read.option("header", "true").csv("file:///home/hadoop/all\_stocks\_5yr.csv")





Просмотр первых пяти строк



Country

Density (Плотность населения)

Abbreviation Сокращение стран

Agricultural Land( %) Сельскохозяйственные угодья(%)

Land Area(Km2) Площадь суши(км2)

Armed Forces size Численность вооруженных сил

Birth Rate Уровень рождаемости

Calling Code Телефонный код

Capital/Major City Столица/Крупный город

Co2-Emissions Выбросы Co2

CPI Индекс потребительских цен

CPI Change (%) Изменение индекса потребительских цен (%)

Currency-Code Код валюты

Fertility Rate Коэффициент рождаемости

Forested Area (%) Площадь лесов (%)

Gasoline Price Цены на бензин

GDP ВВП

Gross primary education enrollment (%) Общий охват начальным образованием (%)

Gross tertiary education enrollment (%) Общее число учащихся в высших учебных заведениях (%)

Infant mortality Младенческая смертность

Largest city Крупнейший город

Life expectancy Ожидаемая продолжительность жизни

Maternal mortality ratio Коэффициент материнской смертности

Minimum wage Минимальная заработная плата

Official language Официальный язык

Out of pocket health expenditure Расходы на здравоохранение из собственных средств

Physicians per thousand врачей на тысячу

Population Население

Population: Labor force participation (%) Население: доля рабочей силы (%)

Tax revenue (%) налоговый доход (%)

Total tax rate общая налоговая ставка

Unemployment rate уровень безработицы

Urban\_population городское население

Latitude широта

Longitude долгота