

Департамент образования и науки города Москвы
Государственное автономное образовательное учреждение
высшего образования города Москвы
«Московский городской педагогический университет»
Институт цифрового образования
Департамент информатики управления и технологий

Агафонов Антон Александрович БД-241м

**Практическая работа 1. Введение в большие данные и их хранение.
Инструменты обработки больших данных (Hadoop)**

Направление подготовки/специальность
38.04.05 - Бизнес-информатика
Бизнес-аналитика и большие данные
(очная форма обучения)
Вариант 1

Москва

2025

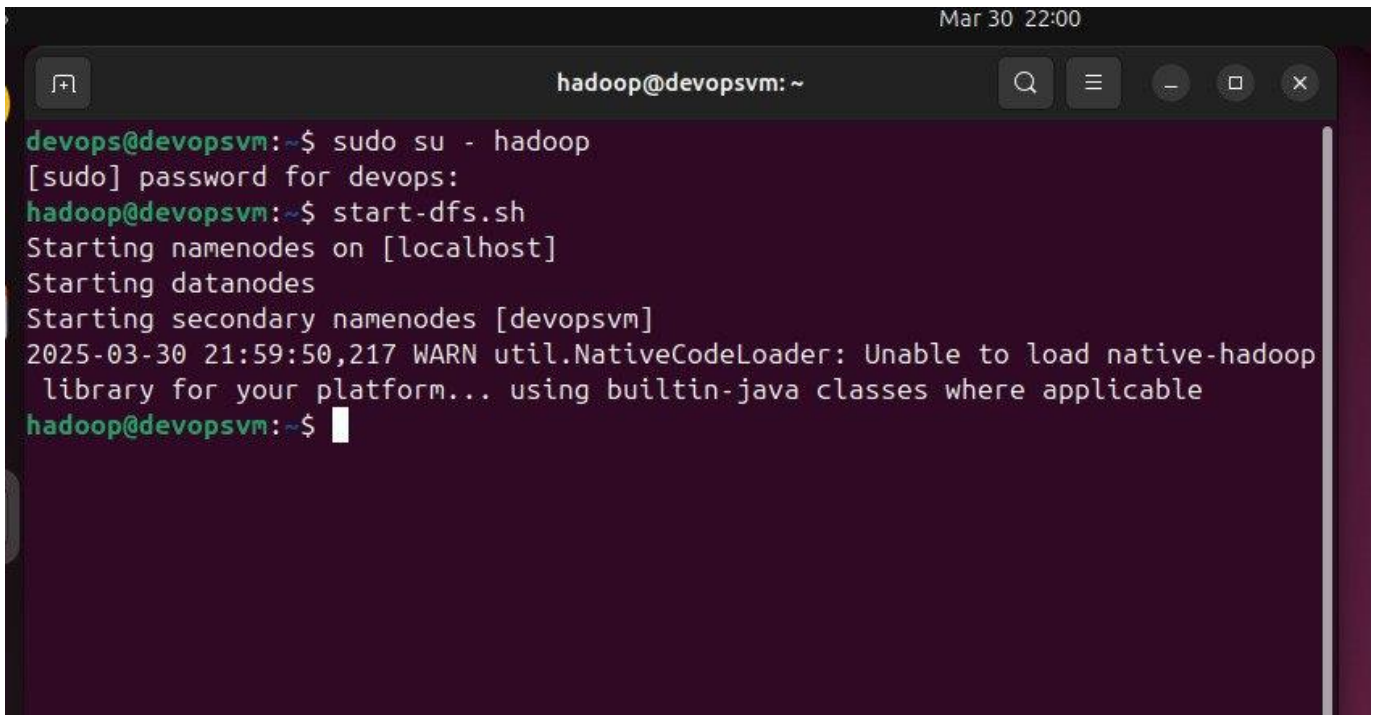
Цель

Изучить основные операции и функциональные возможности системы, что позволит понять принципы работы с данными и распределенными вычислениями.

Основная часть

Запускаем hadoop

Start-dfs.sh

A terminal window titled 'hadoop@devopsvm: ~' with a timestamp 'Mar 30 22:00'. The terminal shows the following commands and output: 'devops@devopsvm:~\$ sudo su - hadoop', '[sudo] password for devops:', 'hadoop@devopsvm:~\$ start-dfs.sh', 'Starting namenodes on [localhost]', 'Starting datanodes', 'Starting secondary namenodes [devopsvm]', '2025-03-30 21:59:50,217 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable', and 'hadoop@devopsvm:~\$' with a cursor.

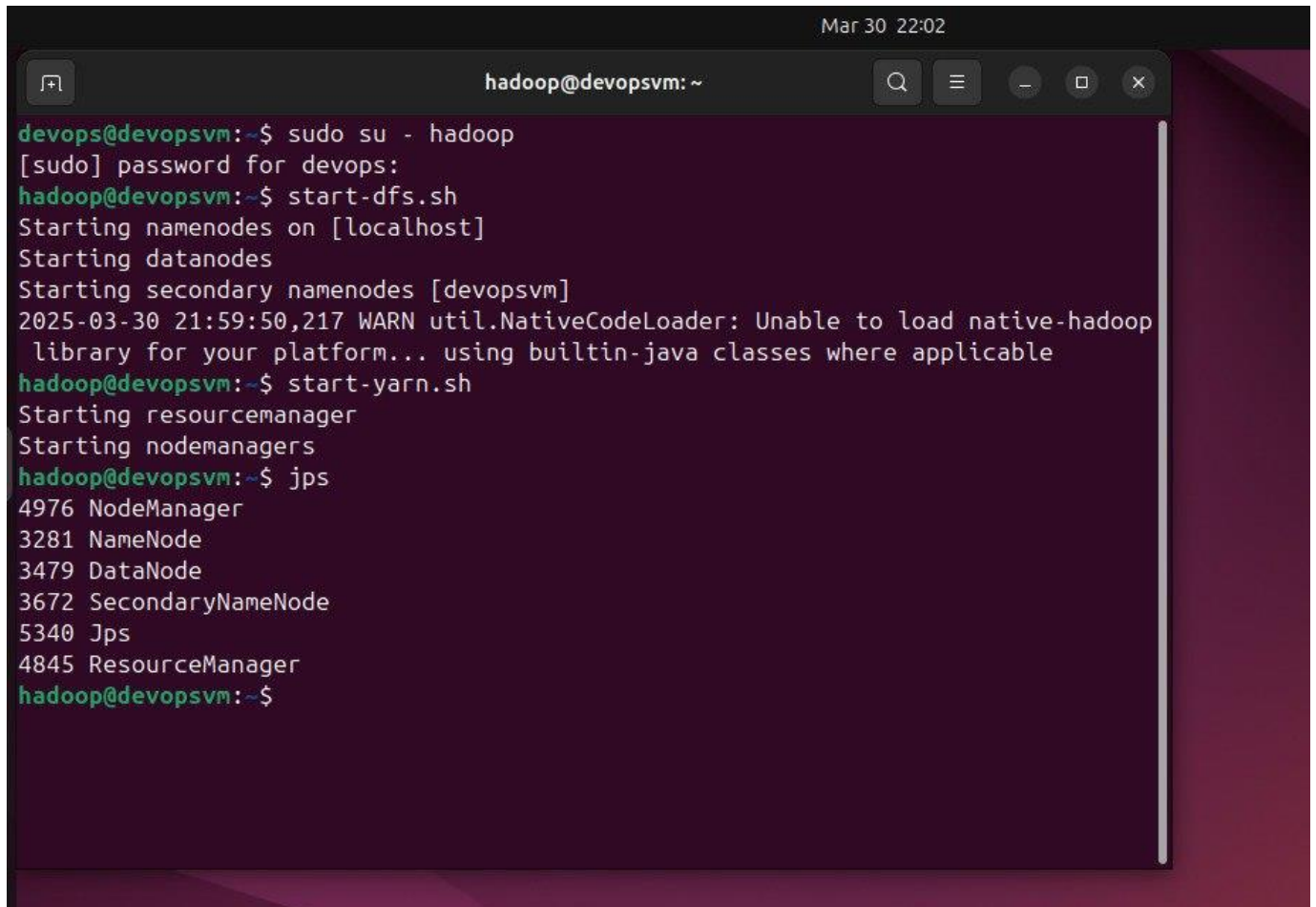
```
Mar 30 22:00
hadoop@devopsvm: ~
devops@devopsvm:~$ sudo su - hadoop
[sudo] password for devops:
hadoop@devopsvm:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [devopsvm]
2025-03-30 21:59:50,217 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$
```

Рис.1 Запускаем файловую систему

Start-yarn.sh

Проверяем запущенные службы командой

jps

A terminal window titled 'hadoop@devopsvm: ~' with a timestamp 'Mar 30 22:02'. The terminal shows the execution of several commands to start Hadoop services. The user switches to the 'hadoop' user, runs 'start-dfs.sh' to start the Distributed File System (DFS) services, and then 'start-yarn.sh' to start the YARN services. Finally, the user runs 'jps' to list the Java processes running on the machine. The output shows that the NameNode, DataNode, SecondaryNameNode, ResourceManager, and NodeManager are all running. A warning message is also displayed regarding the native Hadoop library.

```
hadoop@devopsvm: ~  
devops@devopsvm:~$ sudo su - hadoop  
[sudo] password for devops:  
hadoop@devopsvm:~$ start-dfs.sh  
Starting namenodes on [localhost]  
Starting datanodes  
Starting secondary namenodes [devopsvm]  
2025-03-30 21:59:50,217 WARN util.NativeCodeLoader: Unable to load native-hadoop  
library for your platform... using builtin-java classes where applicable  
hadoop@devopsvm:~$ start-yarn.sh  
Starting resourcemanager  
Starting nodemanagers  
hadoop@devopsvm:~$ jps  
4976 NodeManager  
3281 NameNode  
3479 DataNode  
3672 SecondaryNameNode  
5340 Jps  
4845 ResourceManager  
hadoop@devopsvm:~$
```

Рис.2 Проверка запущенных служб

Проверяем доступность запущенных систем

Переходим по ссылке для проверки запущен ли dfs по ссылке

Localhost:9870/dfshealth.html#tab-overview

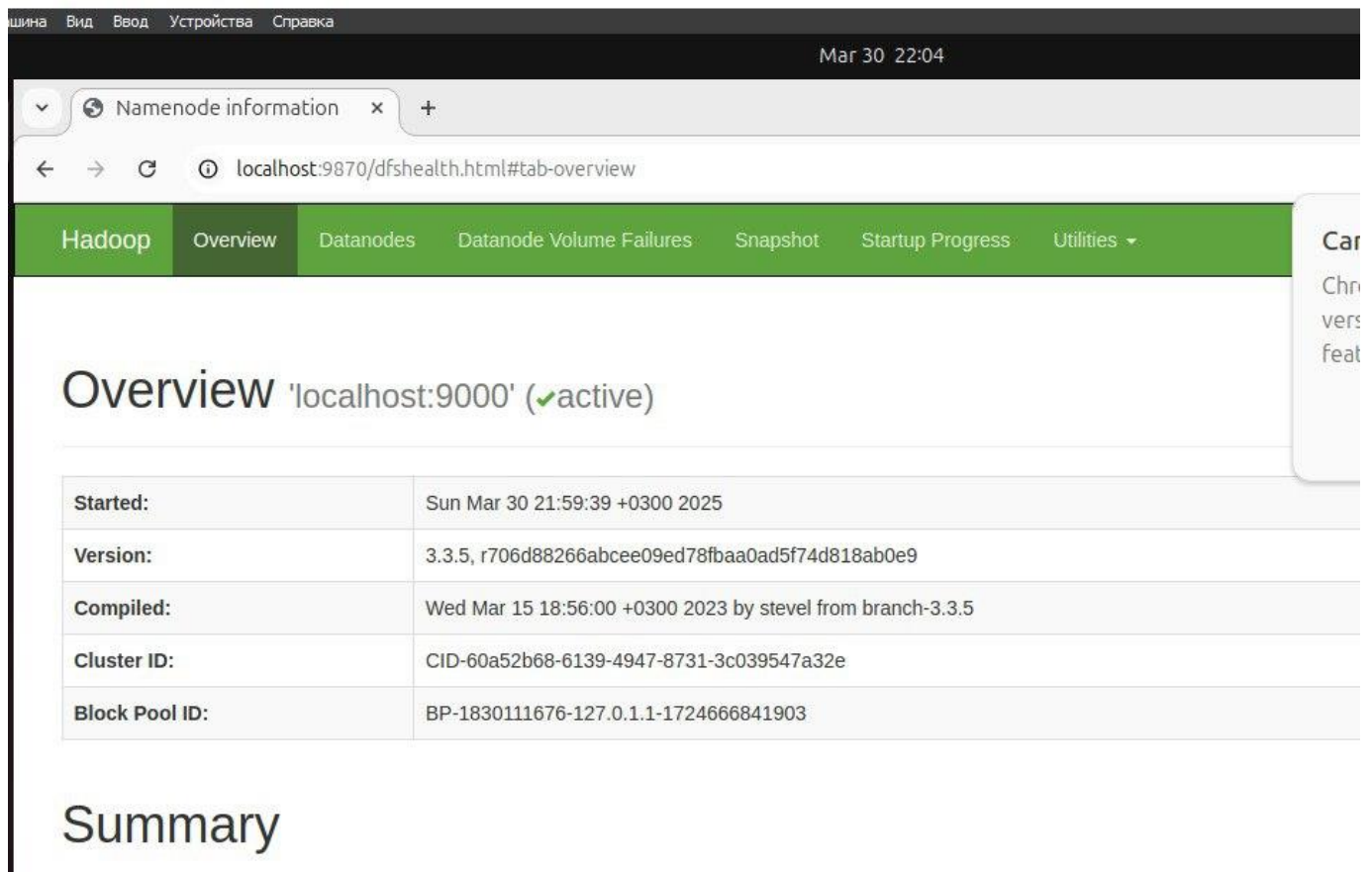


Рис.3 проверка доступности систем

Проверяем запущен ли yarn по ссылке


Localhost:8088/cluster

Mar 30 22:05

Browsing HDFS

All Applications

localhost:8088/cluster



Cluster

[About](#)
[Nodes](#)
[Node Labels](#)
[Applications](#)

NEW
NEW_SAVING
SUBMITTED
ACCEPTED
RUNNING
FINISHED
FAILED
KILLED

[Scheduler](#)

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed
0	0	0	0

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes
1	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type
Capacity Scheduler	[memory-mb (unit=Mi), vcores]

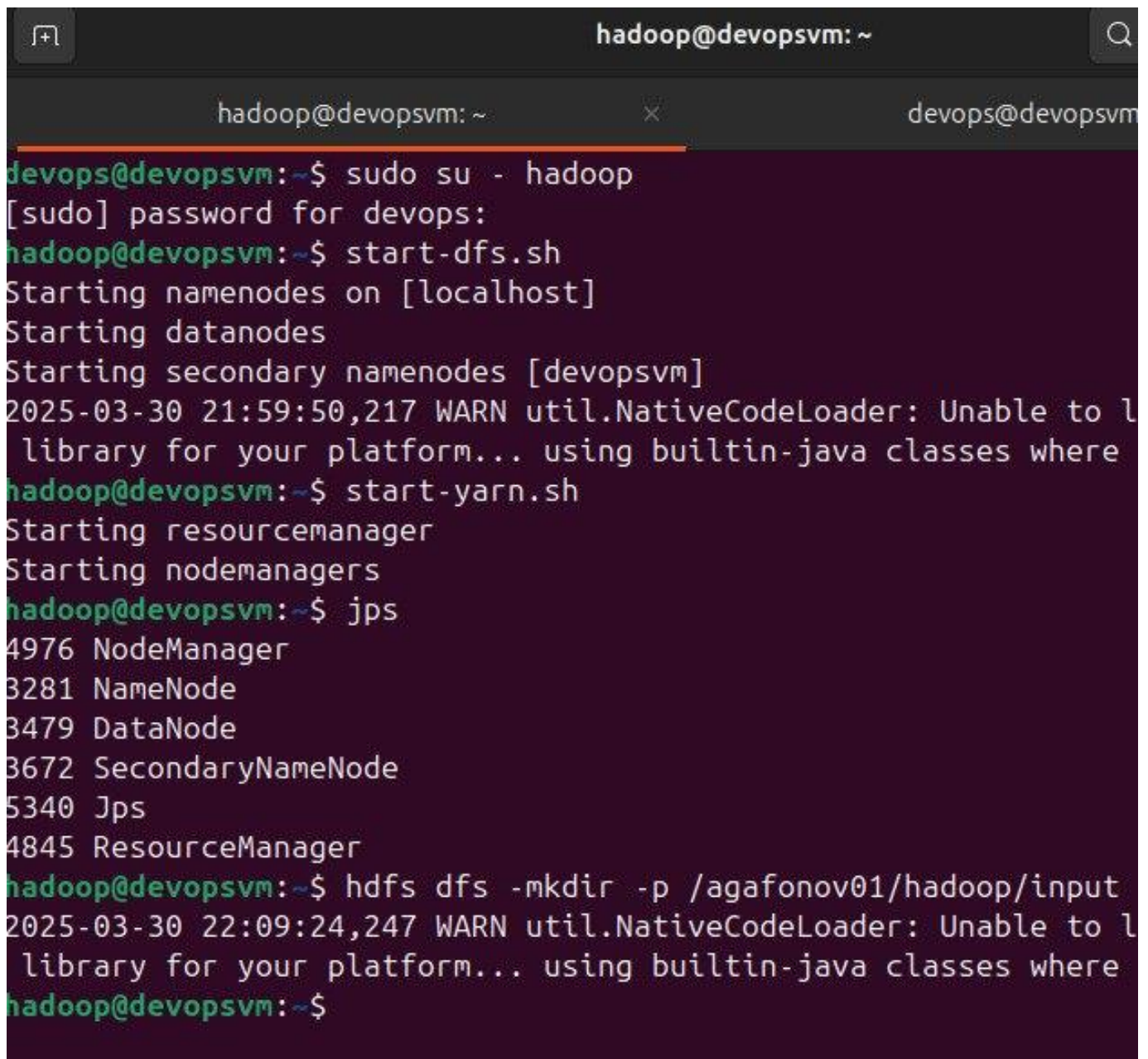
Show 20 entries

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime
Showing 0 to 0 of 0 entries							

Рис.4 Проверка работы yarn

Создаем пользователя и каталог командой

Hdfs dfs -mkdir -p /agafonov01/Hadoop/input



The image shows a terminal window with a dark background. At the top, there is a title bar with the text 'hadoop@devopsvm: ~' and a search icon on the right. Below the title bar, there are two tabs: 'hadoop@devopsvm: ~' (active) and 'devops@devopsvm'. The terminal content shows the following sequence of commands and outputs:

```
devops@devopsvm:~$ sudo su - hadoop
[sudo] password for devops:
hadoop@devopsvm:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [devopsvm]
2025-03-30 21:59:50,217 WARN util.NativeCodeLoader: Unable to load native-lib
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@devopsvm:~$ jps
4976 NodeManager
3281 NameNode
3479 DataNode
3672 SecondaryNameNode
5340 Jps
4845 ResourceManager
hadoop@devopsvm:~$ hdfs dfs -mkdir -p /agafonov01/hadoop/input
2025-03-30 22:09:24,247 WARN util.NativeCodeLoader: Unable to load native-lib
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$
```

Рис.5 Создание пользователя

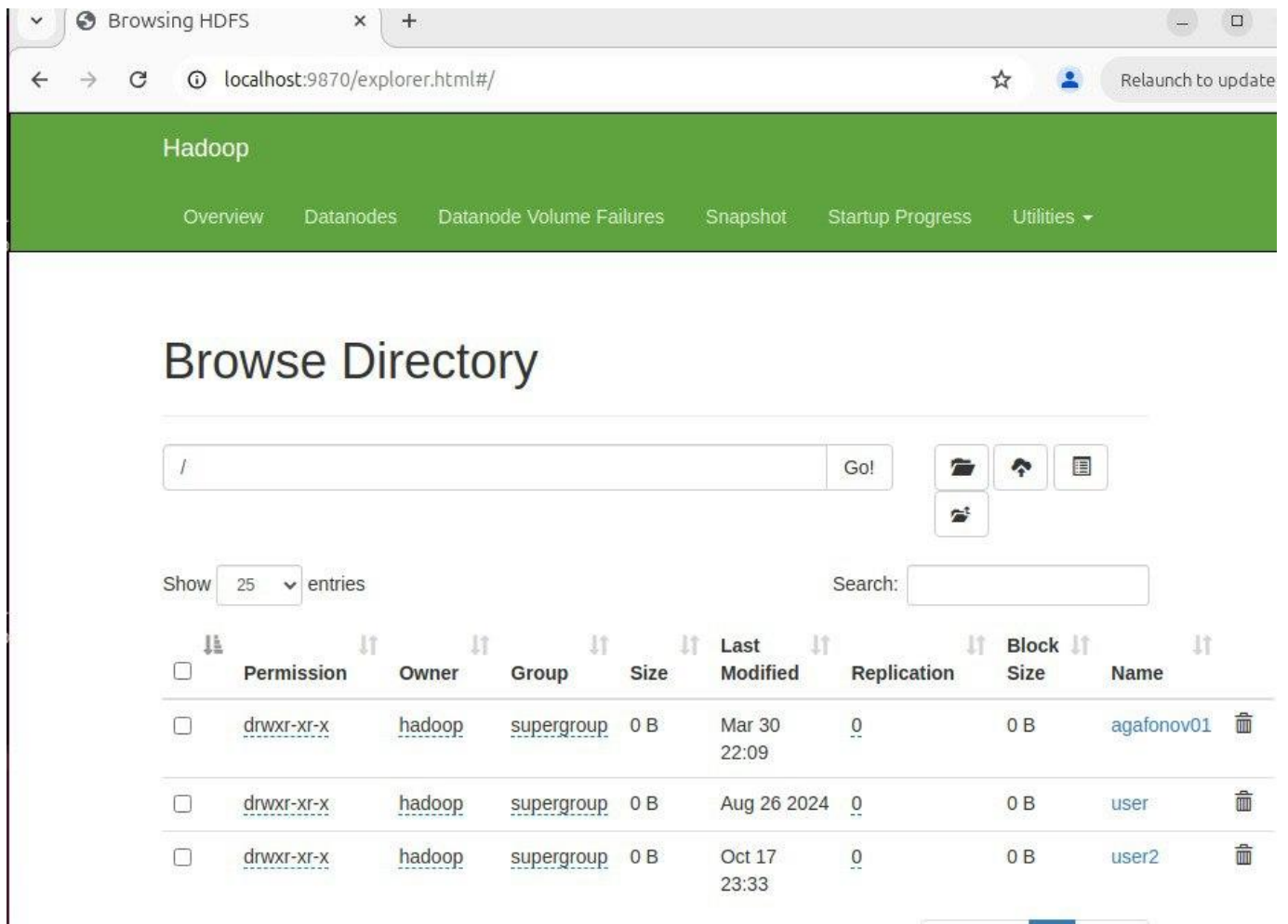


Рис.6 Проверка созданного пользователя

Скачиваем данные

wget

https://raw.githubusercontent.com/BosenkoTM/Distributed_systems/main/practice/2024/Iw_01/GDP.csv


```
Mar 30 22:18
devops@devopsvm: ~
hadoop@devopsvm: ~ x devops@devopsvm: ~ x
devops@devopsvm:~$ wget https://raw.githubusercontent.com/BosenkoTM/Distributed_
systems/main/practice/2024/lw_01/GDP.csv
--2025-03-30 22:18:16-- https://raw.githubusercontent.com/BosenkoTM/Distributed
_systems/main/practice/2024/lw_01/GDP.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.109.1
33, 185.199.108.133, 185.199.111.133, ...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.109.
133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 30268 (30K) [text/plain]
Saving to: 'GDP.csv'

GDP.csv          100%[=====>] 29.56K  --.-KB/s    in 0.02s

2025-03-30 22:18:16 (1.42 MB/s) - 'GDP.csv' saved [30268/30268]

devops@devopsvm:~$
```

Рис.7 Скачиваем данные

Создаем каталог для данных

Hdfs dfs -mkdir -p /agafonov01/Hadoop/input/economic_data


```
Mar 30 22:22
hadoop@devopsvm: ~
hadoop@devopsvm: ~
devops@devopsvm:~$ wget https://raw.githubusercontent.com/BosenkoTM/Distributed_
systems/main/practice/2024/lw_01/GDP.csv
--2025-03-30 22:18:16-- https://raw.githubusercontent.com/BosenkoTM/Distributed_
systems/main/practice/2024/lw_01/GDP.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.109.1
33, 185.199.108.133, 185.199.111.133, ...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.109.
133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 30268 (30K) [text/plain]
Saving to: 'GDP.csv'

GDP.csv          100%[=====>] 29.56K  ---KB/s    in 0.02s

2025-03-30 22:18:16 (1.42 MB/s) - 'GDP.csv' saved [30268/30268]

devops@devopsvm:~$ sudo su - hadoop
[sudo] password for devops:
hadoop@devopsvm:~$ hdfs dfs -mkdir -p /agafonov01/hadoop/input/economic_data
2025-03-30 22:22:03,053 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$
```

Рис.8 Создание каталога economic_data

Mar 30 22:23

Browsing HDFS

localhost:9870/explorer.html#/agafonov01/hadoop/input

Relaunch to update

Hadoop

Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/agafonov01/hadoop/input Go!

Show 25 entries Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
<input type="checkbox"/>	drwxr-xr-x	hadoop	supergroup	0 B	Mar 30 22:22	0	0 B	economic_data

Showing 1 to 1 of 1 entries

Previous 1 Next

Рис.9 Проверка созданного каталога economic_data

Переносим данные в каталог

Hdfs dfs -put GDP.csv /agafonov01/Hadoop/input/economic_data

```
Mar 30 22:31
hadoop@devopsvm: ~
hadoop@devopsvm: ~
systems/main/practice/2024/lw_01/GDP.csv
--2025-03-30 22:18:16-- https://raw.githubusercontent.com/BosenkoTM/Distributed
_systems/main/practice/2024/lw_01/GDP.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.109.1
33, 185.199.108.133, 185.199.111.133, ...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.109.
133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 30268 (30K) [text/plain]
Saving to: 'GDP.csv'

GDP.csv          100%[=====>] 29.56K  ---KB/s    in 0.02s

2025-03-30 22:18:16 (1.42 MB/s) - 'GDP.csv' saved [30268/30268]

devops@devopsvm:~$ sudo su - hadoop
[sudo] password for devops:
hadoop@devopsvm:~$ hdfs dfs -mkdir -p /agafonov01/hadoop/input/economic_data
2025-03-30 22:22:03,053 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$ hdfs dfs -put GDP.csv /agafonov01/hadoop/input/economic_data/
2025-03-30 22:31:22,343 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
```

Рис.10 Перенос данных в каталог

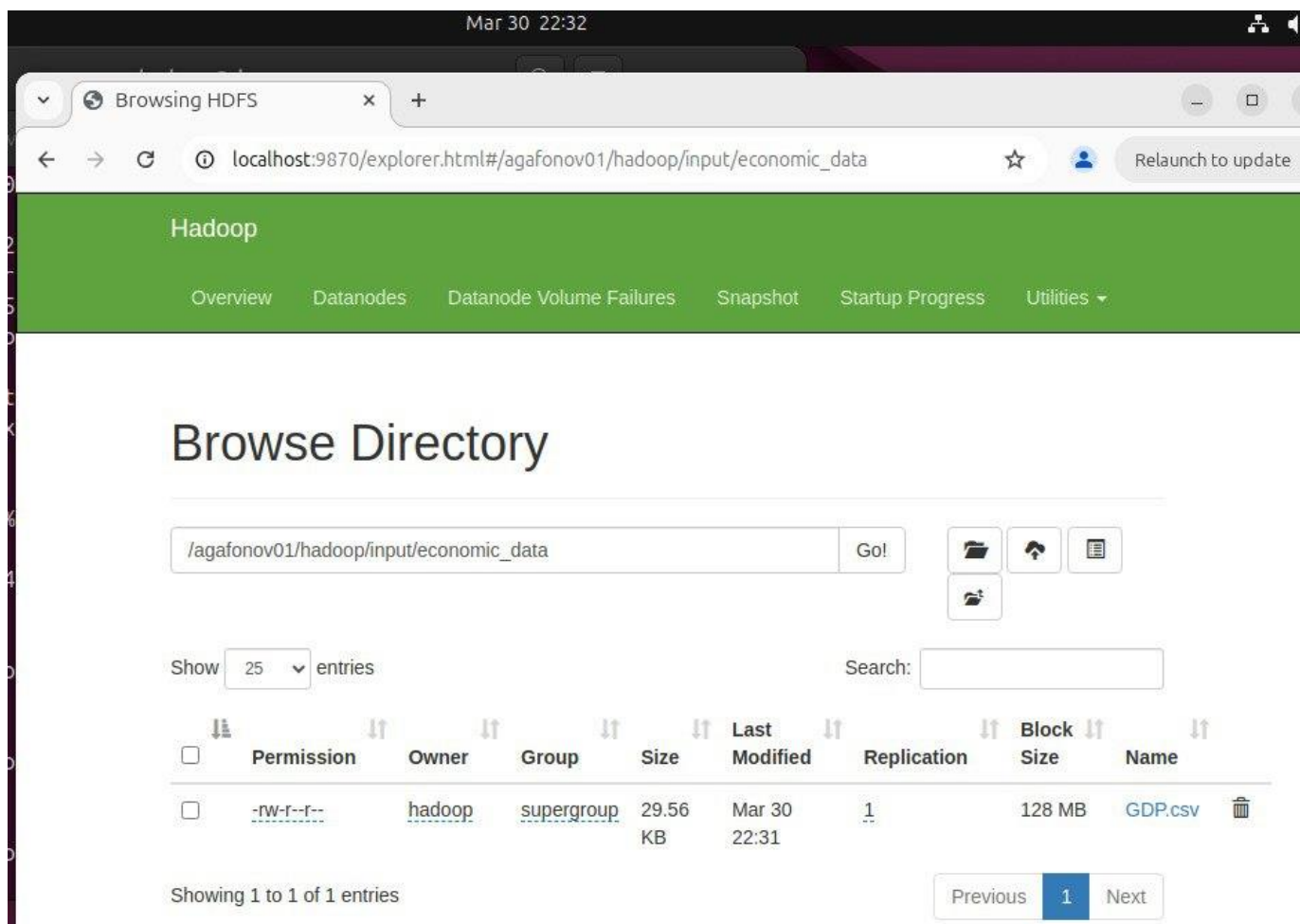
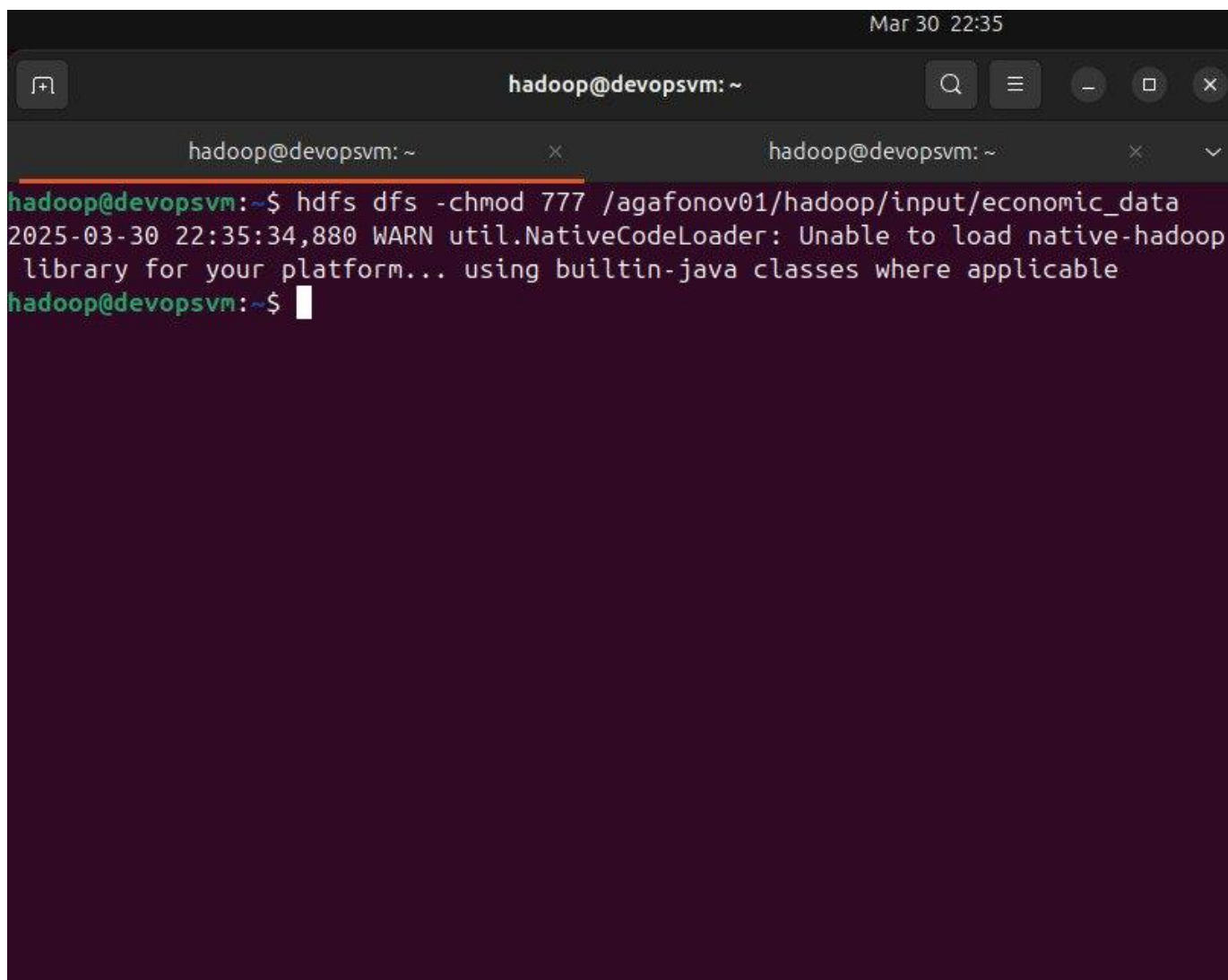


Рис.11 Проверка перенесенных данных

Задаем права доступа

hdfs dfs -chmod 777 /agafonov01/hadoop/input/economic_data



A terminal window titled "hadoop@devopsvm: ~" with a timestamp "Mar 30 22:35". The window shows a command being executed: `hadoop@devopsvm:~$ hdfs dfs -chmod 777 /agafonov01/hadoop/input/economic_data`. The output of the command is: `2025-03-30 22:35:34,880 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable`. The prompt `hadoop@devopsvm:~$` is visible again on the next line.

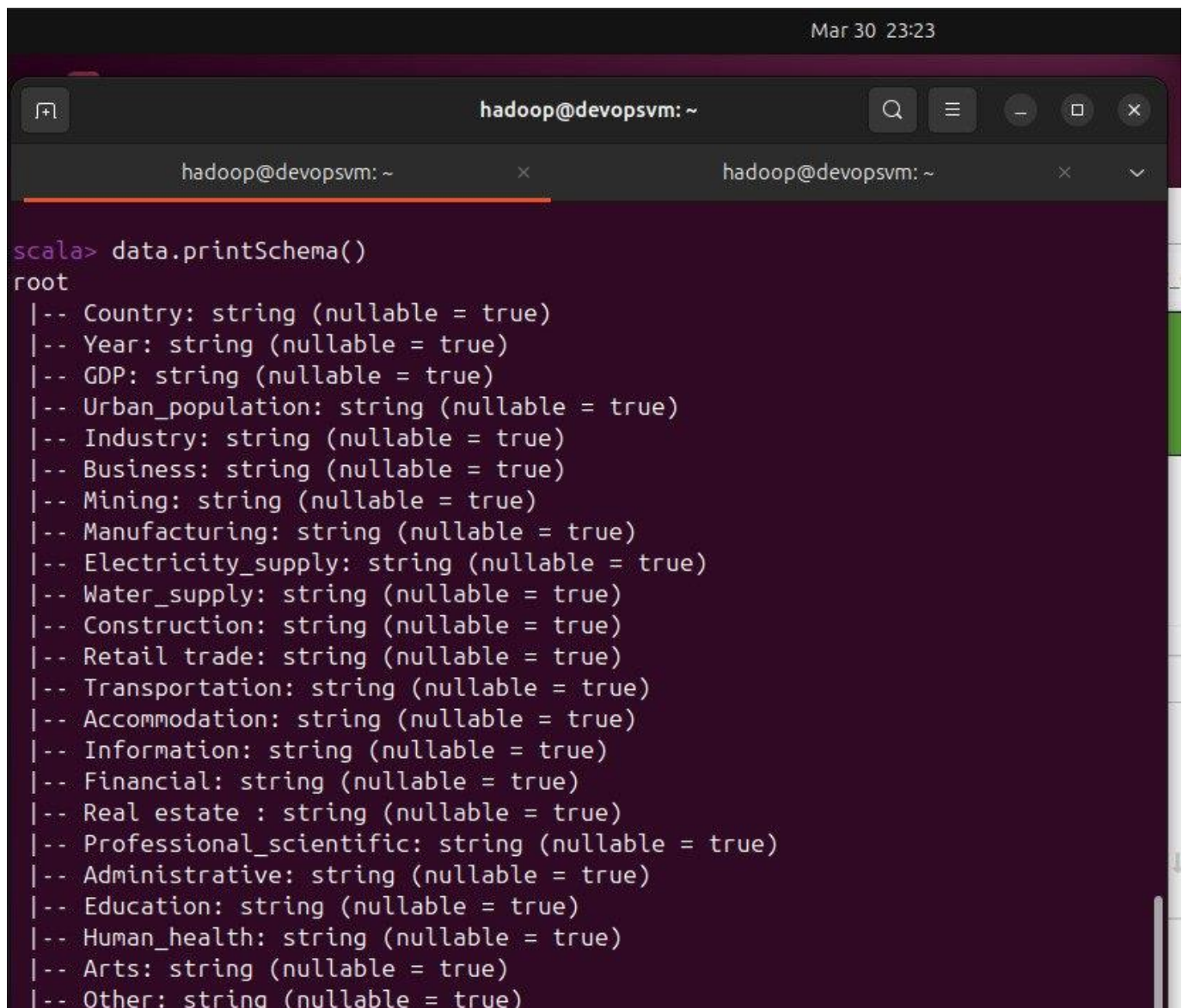
Рис.12 Устанавливаем права доступа

Обработываем данные при помощи Spark

Spark-shell

Проверка полученной схемы данных

data.printSchema()

A screenshot of a terminal window with a dark background. The window title is 'hadoop@devopsvm: ~'. The terminal shows the command 'scala> data.printSchema()' and its output. The output starts with 'root' and lists 20 fields, each with a comment indicating it is a nullable string. The fields are: Country, Year, GDP, Urban_population, Industry, Business, Mining, Manufacturing, Electricity_supply, Water_supply, Construction, Retail trade, Transportation, Accommodation, Information, Financial, Real estate, Professional_scientific, Administrative, Education, Human_health, Arts, and Other. The date and time 'Mar 30 23:23' are visible in the top right corner of the terminal window.

```
scala> data.printSchema()
root
|-- Country: string (nullable = true)
|-- Year: string (nullable = true)
|-- GDP: string (nullable = true)
|-- Urban_population: string (nullable = true)
|-- Industry: string (nullable = true)
|-- Business: string (nullable = true)
|-- Mining: string (nullable = true)
|-- Manufacturing: string (nullable = true)
|-- Electricity_supply: string (nullable = true)
|-- Water_supply: string (nullable = true)
|-- Construction: string (nullable = true)
|-- Retail trade: string (nullable = true)
|-- Transportation: string (nullable = true)
|-- Accommodation: string (nullable = true)
|-- Information: string (nullable = true)
|-- Financial: string (nullable = true)
|-- Real estate : string (nullable = true)
|-- Professional_scientific: string (nullable = true)
|-- Administrative: string (nullable = true)
|-- Education: string (nullable = true)
|-- Human_health: string (nullable = true)
|-- Arts: string (nullable = true)
|-- Other: string (nullable = true)
```

Рис.15 Выводим схему

Вычисление среднего значения GDP

val result = data.selectExpr("avg(GDP) as avg_GDP")

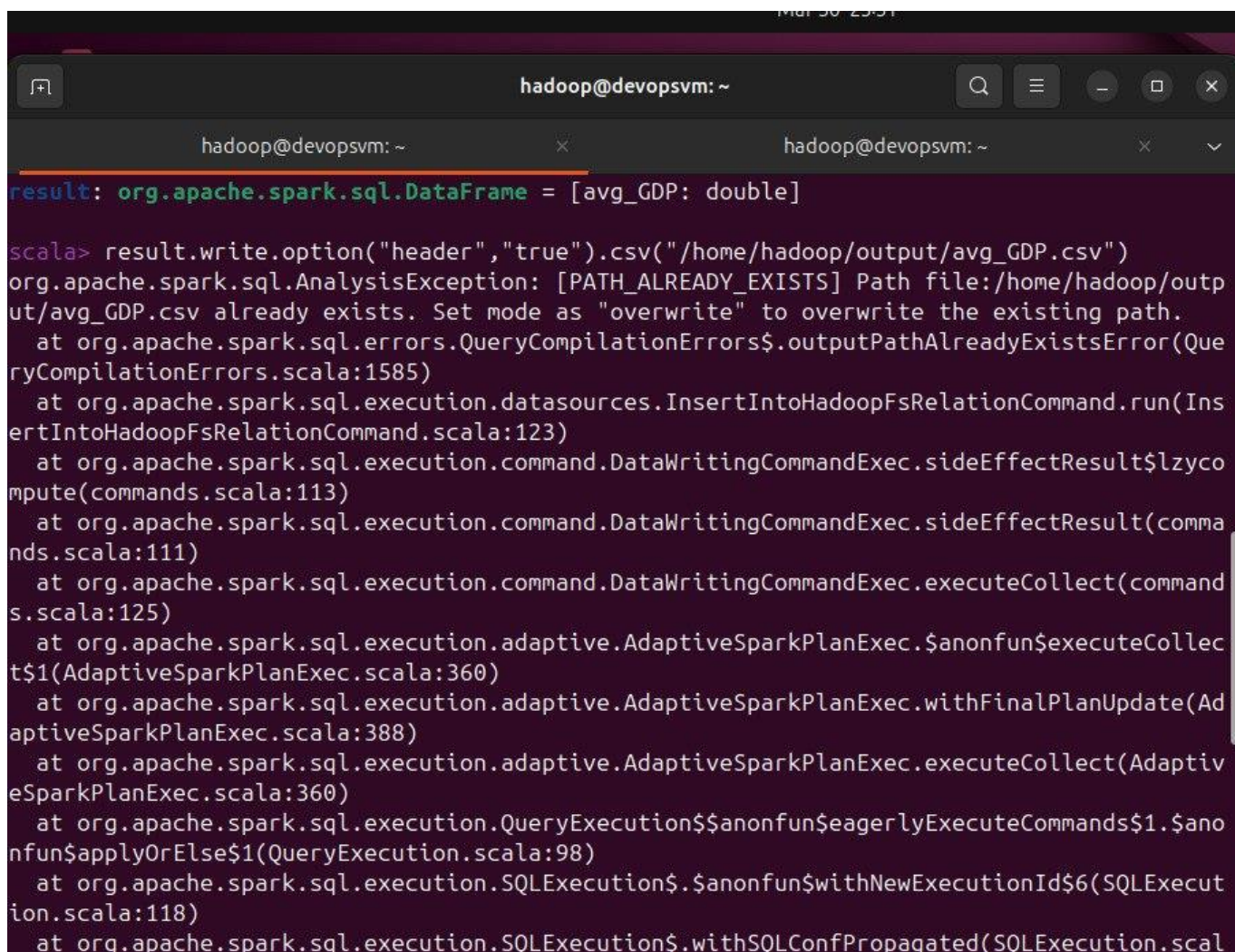
```
scala> val result=data.selectExpr("avg(GDP) as avg_GDP")
result: org.apache.spark.sql.DataFrame = [avg_GDP: double]

scala> 
```

Рис.16 Вычисление среднего значения GDP

Сохраняем результата в CSV файл

```
result.write.option("header", "true").csv("/home/hadoop/output/avg_GDP.csv")
```



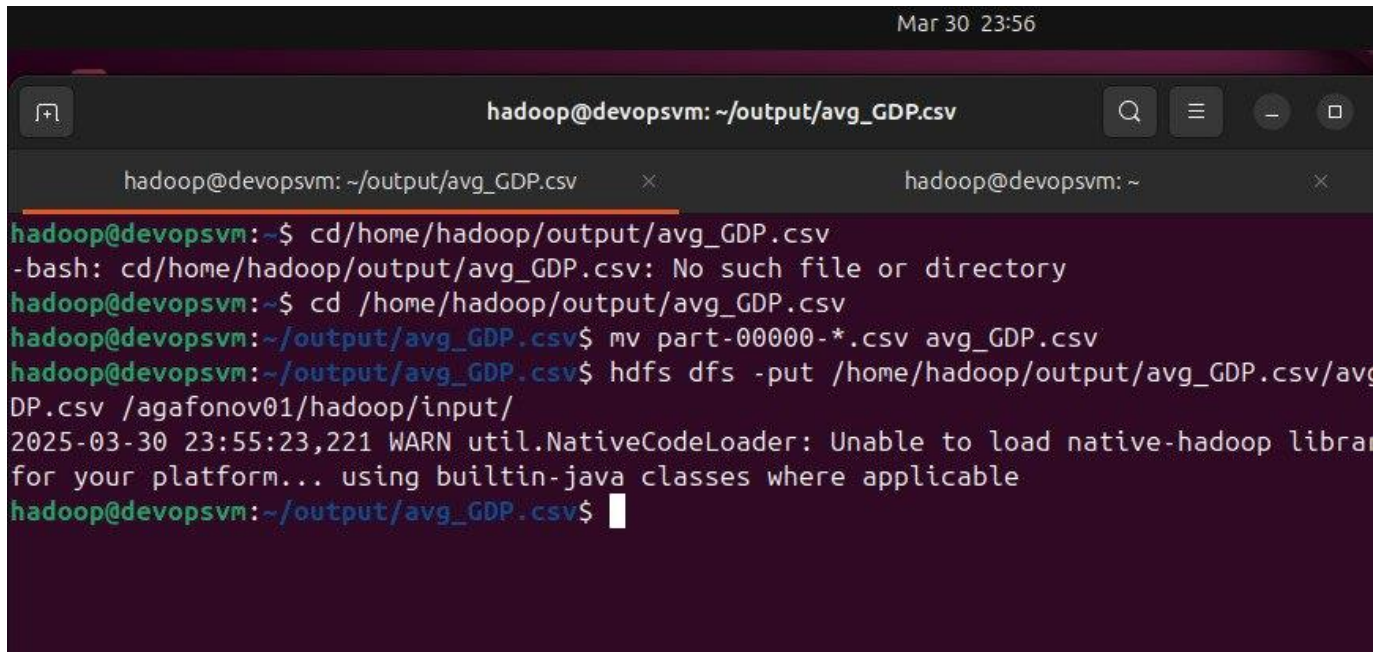
```
hadoop@devopsvm: ~
result: org.apache.spark.sql.DataFrame = [avg_GDP: double]

scala> result.write.option("header","true").csv("/home/hadoop/output/avg_GDP.csv")
org.apache.spark.sql.AnalysisException: [PATH_ALREADY_EXISTS] Path file:/home/hadoop/output/avg_GDP.csv already exists. Set mode as "overwrite" to overwrite the existing path.
    at org.apache.spark.sql.errors.QueryCompilationErrors$.outputPathAlreadyExistsError(QueryCompilationErrors.scala:1585)
    at org.apache.spark.sql.execution.datasources.InsertIntoHadoopFsRelationCommand.run(InsertIntoHadoopFsRelationCommand.scala:123)
    at org.apache.spark.sql.execution.command.DataWritingCommandExec.sideEffectResult$lzycompute(commands.scala:113)
    at org.apache.spark.sql.execution.command.DataWritingCommandExec.sideEffectResult(commands.scala:111)
    at org.apache.spark.sql.execution.command.DataWritingCommandExec.executeCollect(command.scala:125)
    at org.apache.spark.sql.execution.adaptive.AdaptiveSparkPlanExec.$anonfun$executeCollect$1(AdaptiveSparkPlanExec.scala:360)
    at org.apache.spark.sql.execution.adaptive.AdaptiveSparkPlanExec.withFinalPlanUpdate(AdaptiveSparkPlanExec.scala:388)
    at org.apache.spark.sql.execution.adaptive.AdaptiveSparkPlanExec.executeCollect(AdaptiveSparkPlanExec.scala:360)
    at org.apache.spark.sql.execution.QueryExecution$$anonfun$eagerlyExecuteCommands$1.$anonfun$applyOrElse$1(QueryExecution.scala:98)
    at org.apache.spark.sql.execution.SQLExecution$.anonfun$withNewExecutionId$6(SQLExecution.scala:118)
    at org.apache.spark.sql.execution.SQLExecution$.withSQLConfPropagated(SQLExecution.scala:118)
```

Рис.17 сохранение результатов

Переходим в директорию с результатами

```
cd /home/hadoop/output/avg_GDP.csv
```

A terminal window titled 'hadoop@devopsvm: ~/output/avg_GDP.csv' with a timestamp 'Mar 30 23:56'. The terminal shows a series of commands and their outputs. The user attempts to change the directory to '/home/hadoop/output/avg_GDP.csv' but receives an error: '-bash: cd/home/hadoop/output/avg_GDP.csv: No such file or directory'. They then successfully change to the directory using a space-separated path. Next, they rename a file 'part-000000-*.csv' to 'avg_GDP.csv'. Finally, they execute 'hdfs dfs -put /home/hadoop/output/avg_GDP.csv /agafonov01/hadoop/input/' to upload the file. A warning message from 'util.NativeCodeLoader' is also visible. The prompt returns to the user's home directory.

```
hadoop@devopsvm: ~/output/avg_GDP.csv
hadoop@devopsvm: ~/output/avg_GDP.csv
hadoop@devopsvm:~$ cd/home/hadoop/output/avg_GDP.csv
-bash: cd/home/hadoop/output/avg_GDP.csv: No such file or directory
hadoop@devopsvm:~$ cd /home/hadoop/output/avg_GDP.csv
hadoop@devopsvm:~/output/avg_GDP.csv$ mv part-000000-*.csv avg_GDP.csv
hadoop@devopsvm:~/output/avg_GDP.csv$ hdfs dfs -put /home/hadoop/output/avg_GDP.csv /agafonov01/hadoop/input/
2025-03-30 23:55:23,221 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~/output/avg_GDP.csv$
```

Рис.18 Проверяем полученный файл

Загружаем полученный файл в HDFS

```
hdfs dfs -put /home/hadoop/output/avg_GDP.csv /agafonov01/hadoop/input/
```


Mar 30 23:56

Home

Browsing HDFS

localhost:9870/explorer.html#/agafonov01/hadoop/input

Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/agafonov01/hadoop/input Go!

Show 25 entries Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	27 B	Mar 30 23:55	1	128 MB	avg_GDP.csv
<input type="checkbox"/>	drwxrwxrwx	hadoop	supergroup	0 B	Mar 30 22:31	0	0 B	economic_data

Showing 1 to 2 of 2 entries

Previous 1 Next

Рис.19 Проверка загруженного файла в hdfs

Проверка загрузки

hdfs dfs -ls /user01/hadoop/input/

```
hadoop@devopsvm:~/output/avg_GDP.csv$ hdfs dfs -ls /agafonov01/hadoop/input
2025-03-30 23:57:42,223 WARN util.NativeCodeLoader: Unable to load native-hadoop library
for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 hadoop supergroup 27 2025-03-30 23:55 /agafonov01/hadoop/input/avg
_GDP.csv
drwxrwxrwx - hadoop supergroup 0 2025-03-30 22:31 /agafonov01/hadoop/input/econ
omic_data
hadoop@devopsvm:~/output/avg_GDP.csv$
```

Рис.20 Проверка загрузки

Задание для самостоятельной работы

Подключиться к HDFS и убедиться, что файл доступен по пути
hdfs://localhost:9000/agafonov01/hadoop/economic_data/GDP.csv

Запускаем jupyterlab и загружаем файлы из hdfs

File_path= "hdfs://localhost:9000/agafonov01/Hadoop/input/economic_data/GDP.csv"

df = spark.read.csv(file_path, header=True, inferSchema =True)

df.show(5)

```
[38]: # Чтение данных из HDFS
file_path = "hdfs://localhost:9000/agafonov01/hadoop/input/economic_data/GDP.csv"
df = spark.read.csv(file_path, header=True, inferSchema=True)

# Просмотр первых строк данных
df.show(5)
```

[Country Year]	GDP	Urban_population	Industry	Business	Mining	Manufacturing	Electricity_supply	Water_supply	Constr
uction Retail trade Transportation Accommodation Information Financial Real estate	Professional_scientific Admini	strative Education Human_health Arts Other							
Austria 2010 35390	57.4	24.0	25.2	18.3		24.4		23.6	12.2
9.9 27.5	7.3	9.9	21.2	30.3		27.0			34.0
22.5 27.8	12.0 34.0 32.0								
Austria 2015 36140	57.72	21.8	23.4	13.7		22.7		17.6	9.3
8.2 23.3	11.6	6.4	22.4	30.3		28.0			31.3
20.0 24.2	12.9 26.2 28.3								
Austria 2016 36390	57.91	20.8	22.3	14.4		21.9		13.2	8.2
8.3 23.3	14.5	5.9	20.9	27.1		28.7			30.4
17.8 24.3	14.5 20.8 27.8								
Austria 2017 36980	58.09	20.7	22.3	10.9		21.7		13.0	8.4
8.3 23.2	12.4	5.7	20.6	28.4		29.0			29.4
17.4 23.7	15.0 19.1 26.9								
Austria 2018 37690	58.3	20.4	22.0	7.9		21.4			
8.3 23.2	11.7	5.4	20.7	28.2					
17.1 23.6	15.3 18.3 26.4								

Would you like to get notified about official Jupyter news?

Индивидуальное задание

Вариант 1

https://raw.githubusercontent.com/BosenkoTM/Distributed_systems/main/practice/2024/lw_01/AAPL.csv

Для начала создаем каталог в hdfs

Hdfs dfs -mkdir -p /agafonov01/hadoop/input/vol



Рис.21 Создание каталога surgutneftgaz в hdfs

Переносим файл из пользователя devops в пользователя Hadoop

sudo mv /home/devops/Desktop/AAPL.csv /home/hadoop

Переносим данные в каталог

Hdfs dfs -put AAPL.csv /agafonov01/Hadoop/input/vol

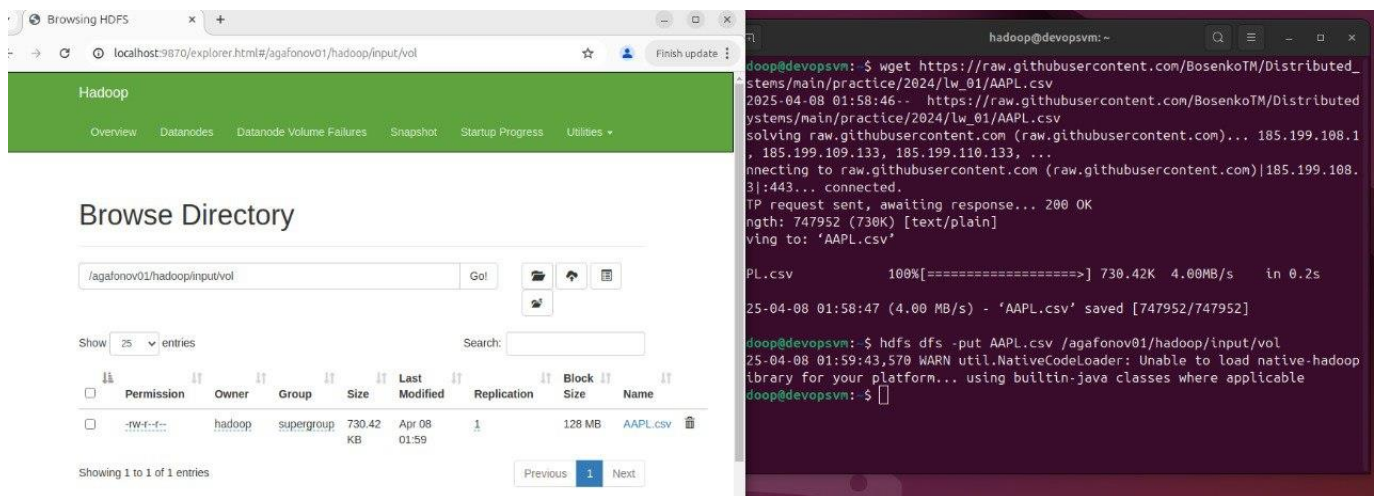


Рис.22 Переносим данные об акциях в hdfs

Устанавливаем права доступа

```
hdfs dfs -chmod 777 /agafonov01/hadoop/input/vol
```

```
hadoop@devopsvm:~$ hdfs dfs -chmod 777 /agafonov01/hadoop/input/vol
2025-04-08 02:02:08,618 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
hadoop@devopsvm:~$
```

Рис.23 Устанавливаем права доступа

Загружаем данные из hdfs

Import pandas as pd

Import matplotlib.pyplot as plt

From pyspark.sql import SparkSession

Spark = SparkSession.builder\

.appName("Economic Data Analysis")\

.config("spark.hadoop.fs.defaultFS","hdfs://localhost:9000")\

.getOrCreate()

Spark.conf.set("spark.sql.shuffle.partition", "50")

File_path = "hdfs://localhost:9000/agafonov01/Hadoop/input/vol/AAPL.csv"

Df= spark.read.csv(file_path, header = True, inferSchema =True)

Df.show(5)

```
# Создание SparkSession
spark = SparkSession.builder \
    .appName("Economic Data Analysis") \
    .config("spark.hadoop.fs.defaultFS", "hdfs://localhost:9000") \
    .config("spark.ui.port", "4050") \
    .getOrCreate()

# Установка количества разделов для shuffle операций
spark.conf.set("spark.sql.shuffle.partitions", "50")

# Чтение данных из HDFS
file_path = "hdfs://localhost:9000/agafonov01/hadoop/input/vol/AAPL.csv"
df = spark.read.csv(file_path, header=True, inferSchema=True)

# Просмотр первых строк данных
df.show(5)

j1: pandas_df = df.toPandas()
pandas_df.head()
```

	Country	Year	GDP	Urban_population	Industry	Business	Mining	Manufacturing	Electricity_supply	Water_supply	Accommodation	Information	Financial	Real estate	Professional_scientific	Administrative	Education	Human_health	Arts	Othe
0	Austria	2010	35390	57.40	24.0	25.2	18.3	24.4	23.6	12.2	9.9	21.2	30.3	27.0	34.0	22.5	27.8	12.0	34.0	32.0
1	Austria	2015	36140	57.72	21.8	23.4	13.7	22.7	17.6	9.3	6.4	22.4	30.3	28.0	31.3	20.0	24.2	12.9	26.2	28.0
2	Austria	2016	36390	57.91	20.8	22.3	14.4	21.9	13.2	8.2	5.9	20.9	27.1	28.7	30.4	17.8	24.3	14.5	20.8	27.0
3	Austria	2017	36980	58.09	20.7	22.3	10.9	21.7	13.0	8.4	5.7	20.6	28.4	29.0	29.4	17.4	23.7	15.0	19.1	26.0
4	Austria	2018	37690	58.30	20.4	22.0	7.9	21.4	14.4	8.1	5.4	20.7	28.2	29.2	28.3	17.1	23.6	15.3	18.3	26.0

5 rows x 23 columns

Заключение

В ходе проделанной лабораторной работы, были изучены основные операции и функциональные возможности системы, что позволило понять принципы работы с данными и распределенными вычислениями, также было выполнено задание по вариантам.