|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ **Информатика и системы управления**

КАФЕДРА **Компьютерные системы и сети (ИУ6)**

НАПРАВЛЕНИЕ ПОДГОТОВКИ **09.04.01 Информатика и вычислительная техника**

МАГИСТЕРСКАЯ ПРОГРАММА **09.04.01/12 Интеллектуальный анализ больших**

**данных в системах поддержки принятия решений.**

**Отчет**

**по лабораторной работе № 10**

**Вариант № 2**

**Название:** Spark

**Дисциплина:** языки программирования для работы с большими данными

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Студент | ИУ6-23М |  |  | А.С. Бояринов |
|  | (Группа) |  | (Подпись, дата) | (И.О. Фамилия) |
|  |  |  |  |  |
| Преподаватель |  |  |  | П.В. Степанов |
|  |  |  | (Подпись, дата) | (И.О. Фамилия) |

Москва, 2023

**Цель:** ознакомиться с работой Spark в языке программирования Java.

**Задание:** сделать 10 выборок данных по выбранной предметной области

Был выбран Dataset с лучшими кассовыми играми, с их рангом, суммой продаж, принадлежности к игровой серии и платформам.

Код класса Main:

package org.example;

import org.apache.spark.sql.Dataset;

import org.apache.spark.sql.Row;

import org.apache.spark.sql.SparkSession;

public class Main {

public static void main(String[] args) {

SparkSession spark = SparkSession.builder()

.appName("RandomSplitExample")

.master("local[\*]")

.getOrCreate();

Dataset<Row> dataset = spark.read().format("csv")

.option("header", "true")

.load("/home/user/Lab10/src/best-selling video games of all time.csv");

dataset.createOrReplaceTempView("games");

spark.sql("select \* from games").show();

spark.sql("select \* from games where Rank <= 5").show();

spark.sql("select \* from games where Rank%5=0").show();

spark.sql("select \* from games where Sales >= 100000000").show();

spark.sql("select \* from games where Title like '%Grand Theft Auto%'").show();

spark.sql("select \* from games where Series = 'Wii' and Sales >= 40000000").show();

spark.sql("select \* from games where Series like '%ario%'").show();

spark.sql("select \* from games where Publisher = 'Nintendo' and Series = 'None'").show();

spark.sql("select \* from games where Date like '%2017%' and Sales >= 30000000").show();

spark.sql("select \* from games where Platform = 'Multi-platform' and Series like '%itcher%'" +

"and Sales >= 20000000").show();

spark.stop();

}

}

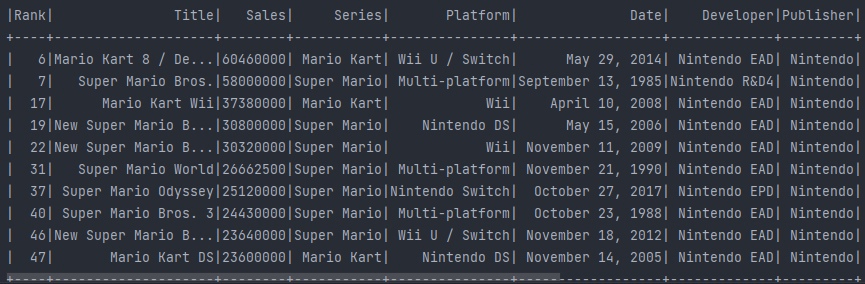


Рисунок 1 – Результат работы программы

**Вывод:** был изучен Spark в языке программирования Java.