|  |  |
| --- | --- |
| 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 **09.04.01/05 Современные интеллектуальные**

**программно-аппаратные комплексы**

**ОТЧЕТ**

**по лабораторной работе №8**

|  |  |
| --- | --- |
| Группа | ИУ6-23М |

|  |  |  |  |
| --- | --- | --- | --- |
| Студент |  |  | И.И. Иванов |
|  |  | (Подпись, дата) | (И.О. Фамилия) |
|  |  |  |  |
| Преподаватель |  |  | П.В. Степанов |
|  |  | (Подпись, дата) | (И.О. Фамилия) |

*Москва, 2023 г.*

**Условие**

1. Выбрать любой датасет на kaggle.com
2. Cделать 10 выборок данных по выбранной предметной области

**Програмная реализация**

1

|  |
| --- |
| 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("TestingSpark").master("local[\*]").getOrCreate();  Dataset<Row> data = spark.read().format("csv").option("header", "true")  .load("src/main/resources/2023\_rankings.csv");  data.createOrReplaceTempView("VUS");  System.out.println("`````````````````````0st sample````````````````````");  spark.sql("SELECT \* FROM VUS;").show(30, false);  System.out.println("`````````````````````1st sample````````````````````");  spark.sql("SELECT rank\_order, name, scores\_overall FROM VUS ORDER BY scores\_overall DESC").show(30, false);  System.out.println("`````````````````````2st sample````````````````````");  spark.sql("SELECT rank, name, scores\_teaching, scores\_research FROM VUS WHERE scores\_teaching > 80 AND scores\_research < 80").show(30, false);  System.out.println("`````````````````````3st sample````````````````````");  spark.sql("SELECT rank, name, scores\_citations, scores\_citations\_rank FROM VUS WHERE scores\_citations\_rank > 200").show(30, false);  System.out.println("`````````````````````4st sample````````````````````");  spark.sql("SELECT AVG(scores\_teaching), AVG(scores\_research), AVG(scores\_citations) FROM VUS;").show(30, false);  spark.stop();  }  } |