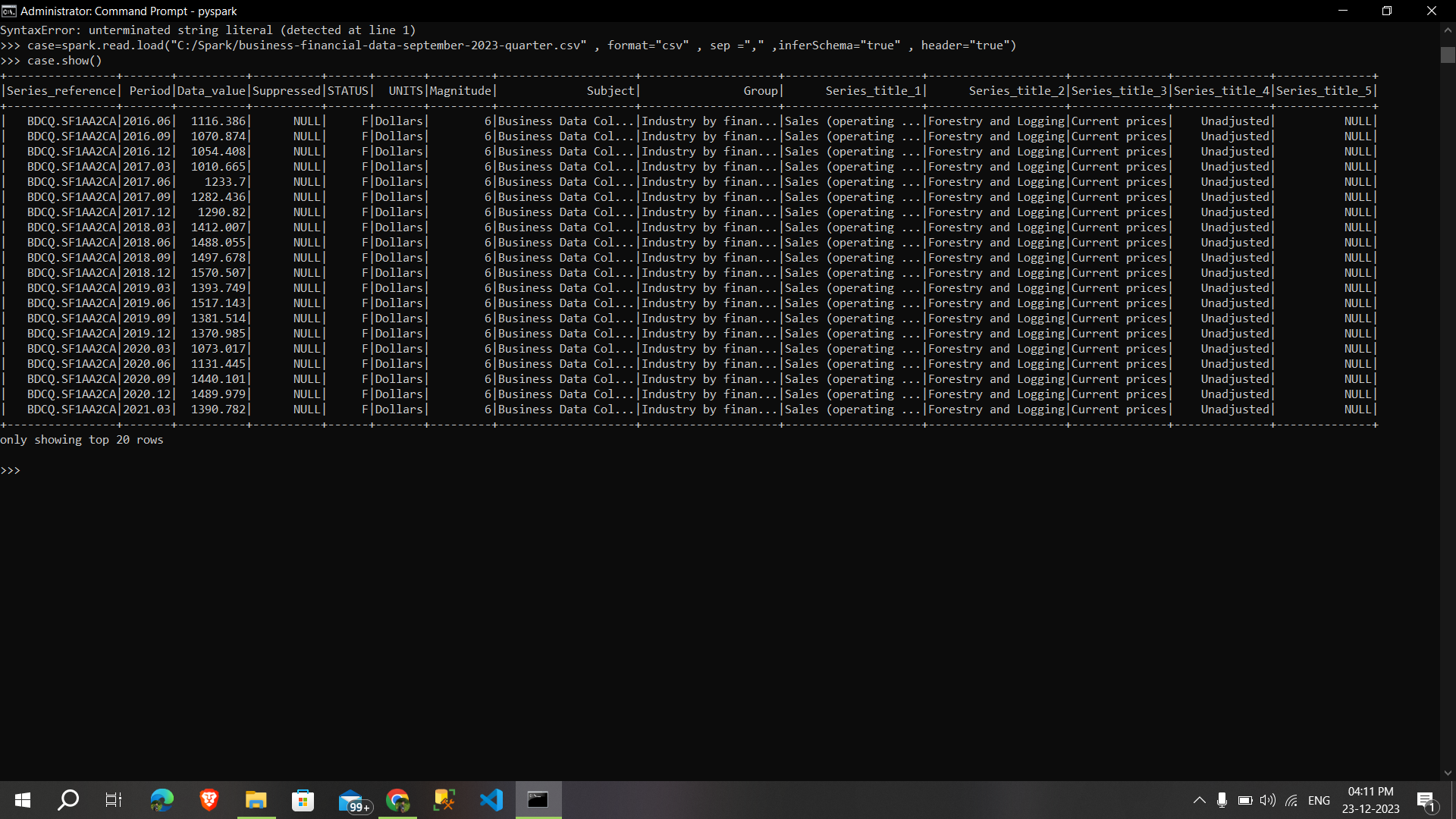
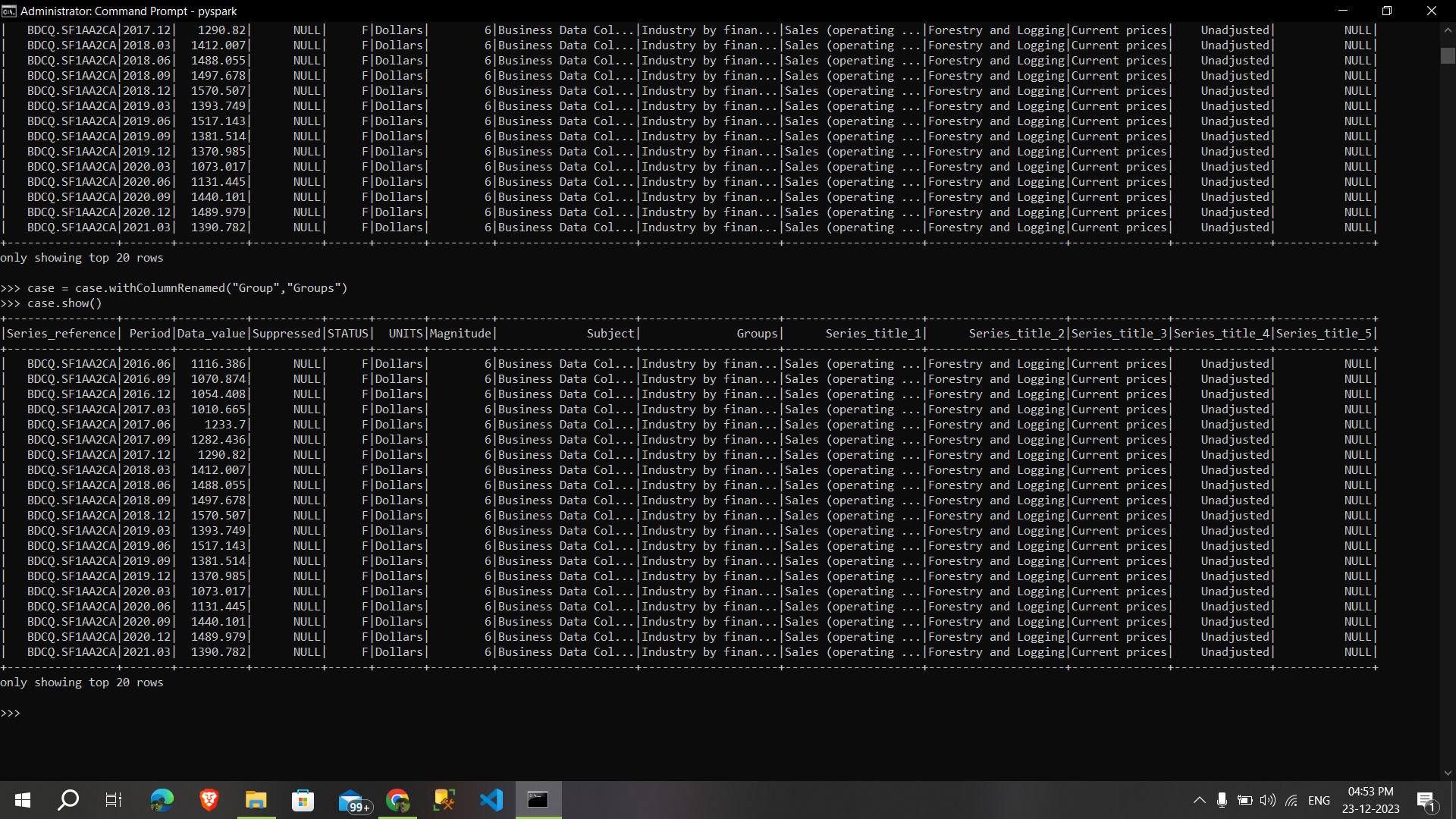
**Basic Functions of Spark**

With some random csv files using these pyspark commands,

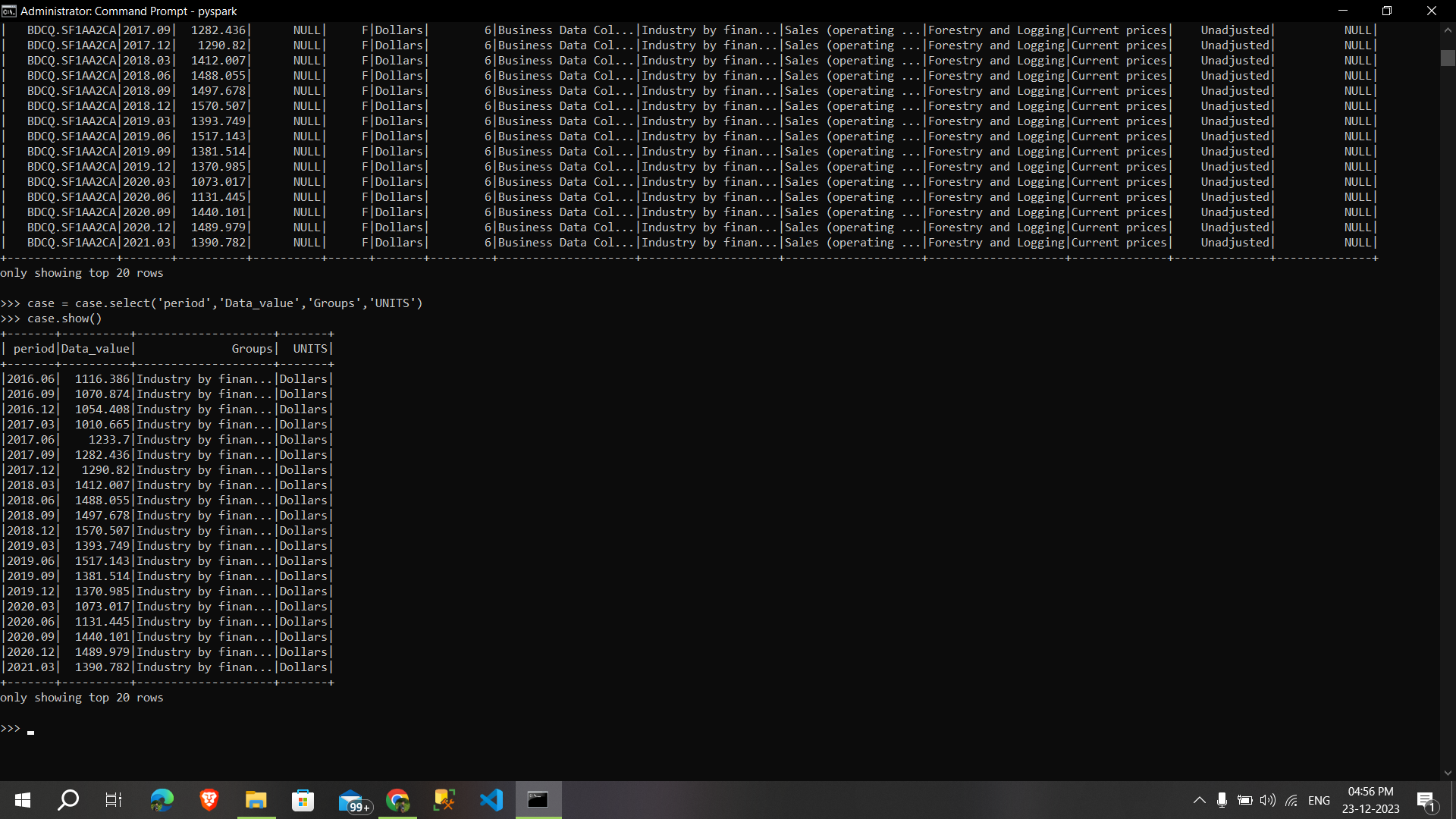
**Reading a file and show:**



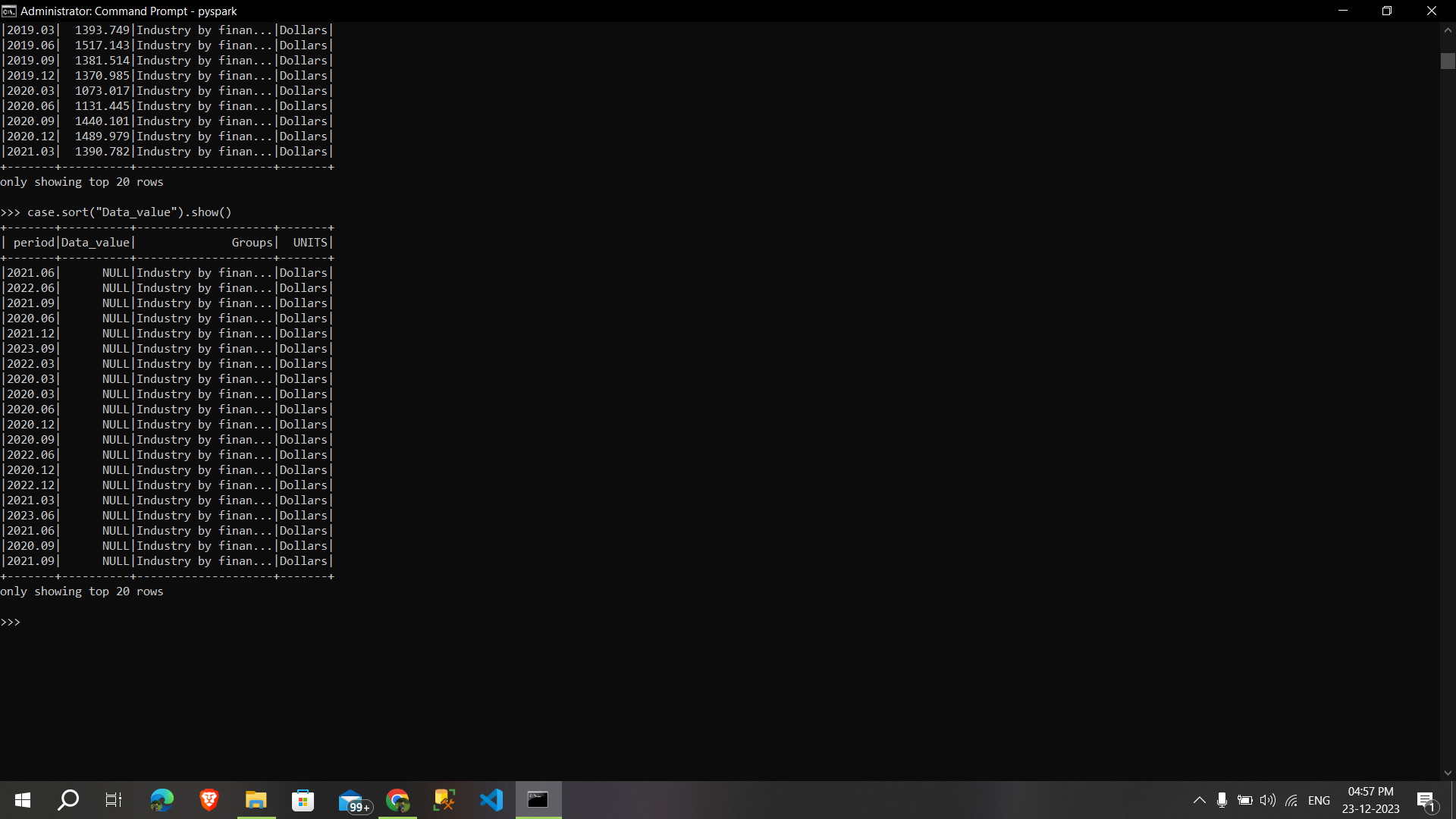
**Changing the column name:**

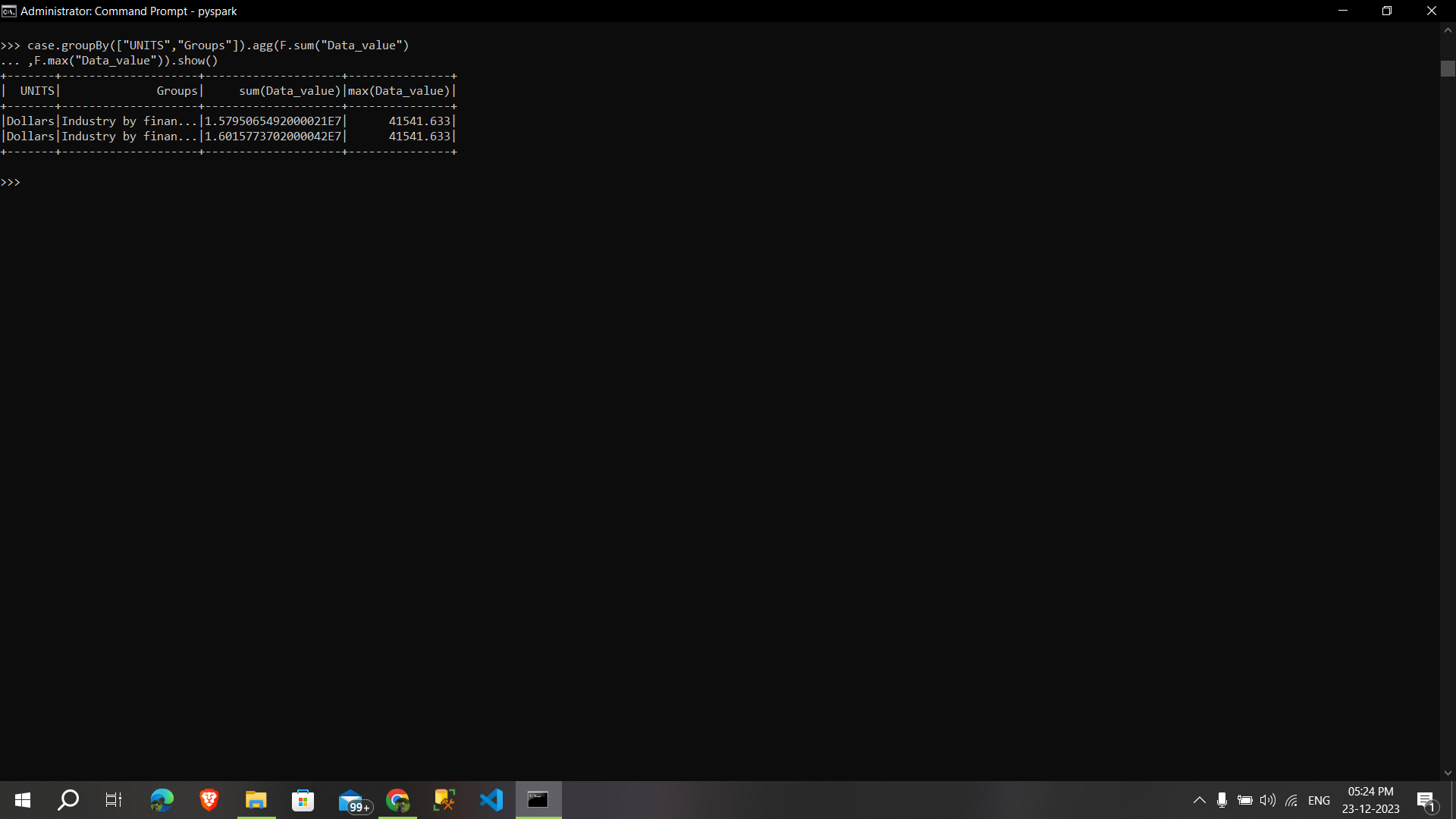


**Selecting particular column:**

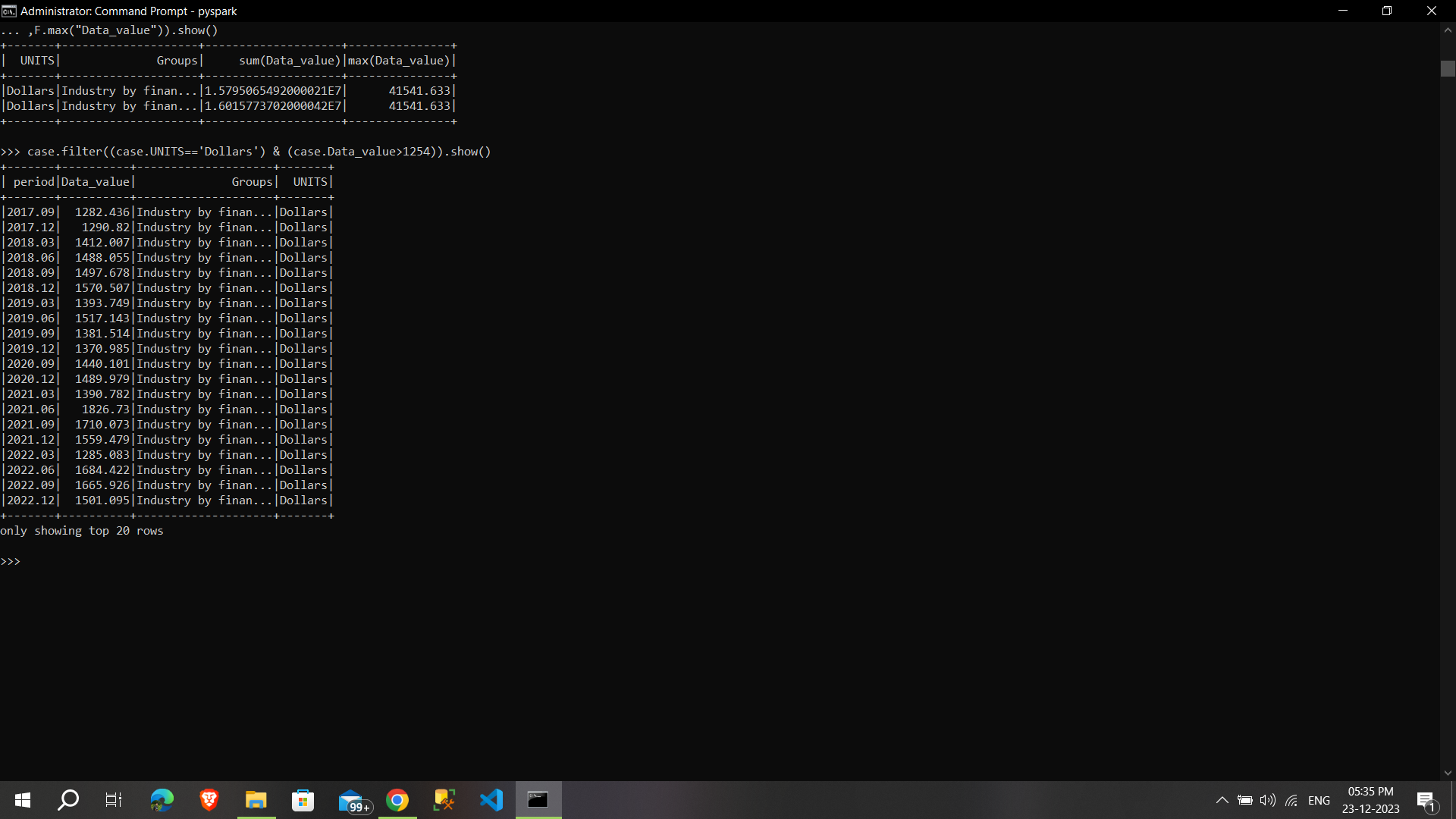


**Sorting the particular column:**



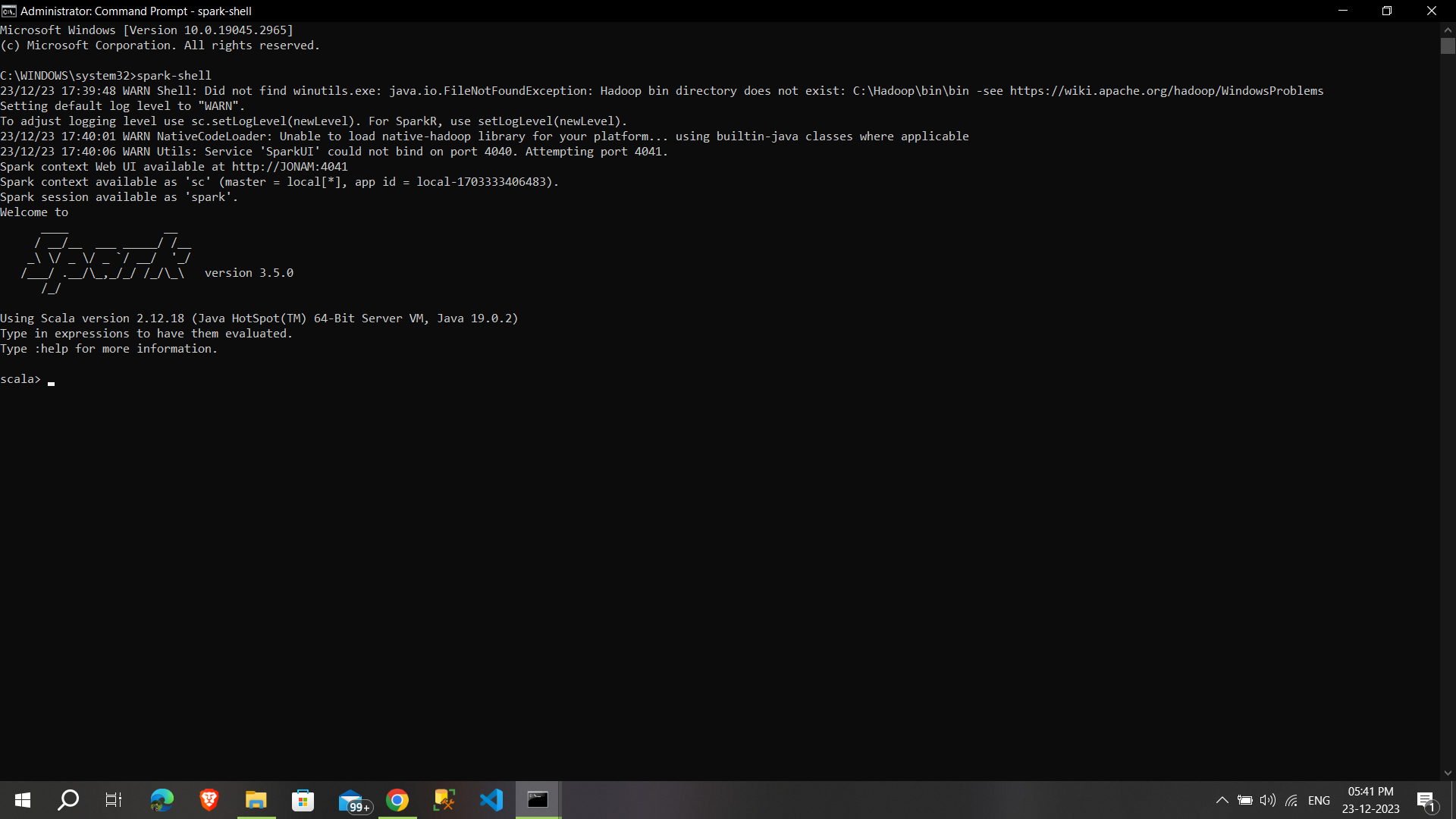
**Group by :**

**By using Filter:**

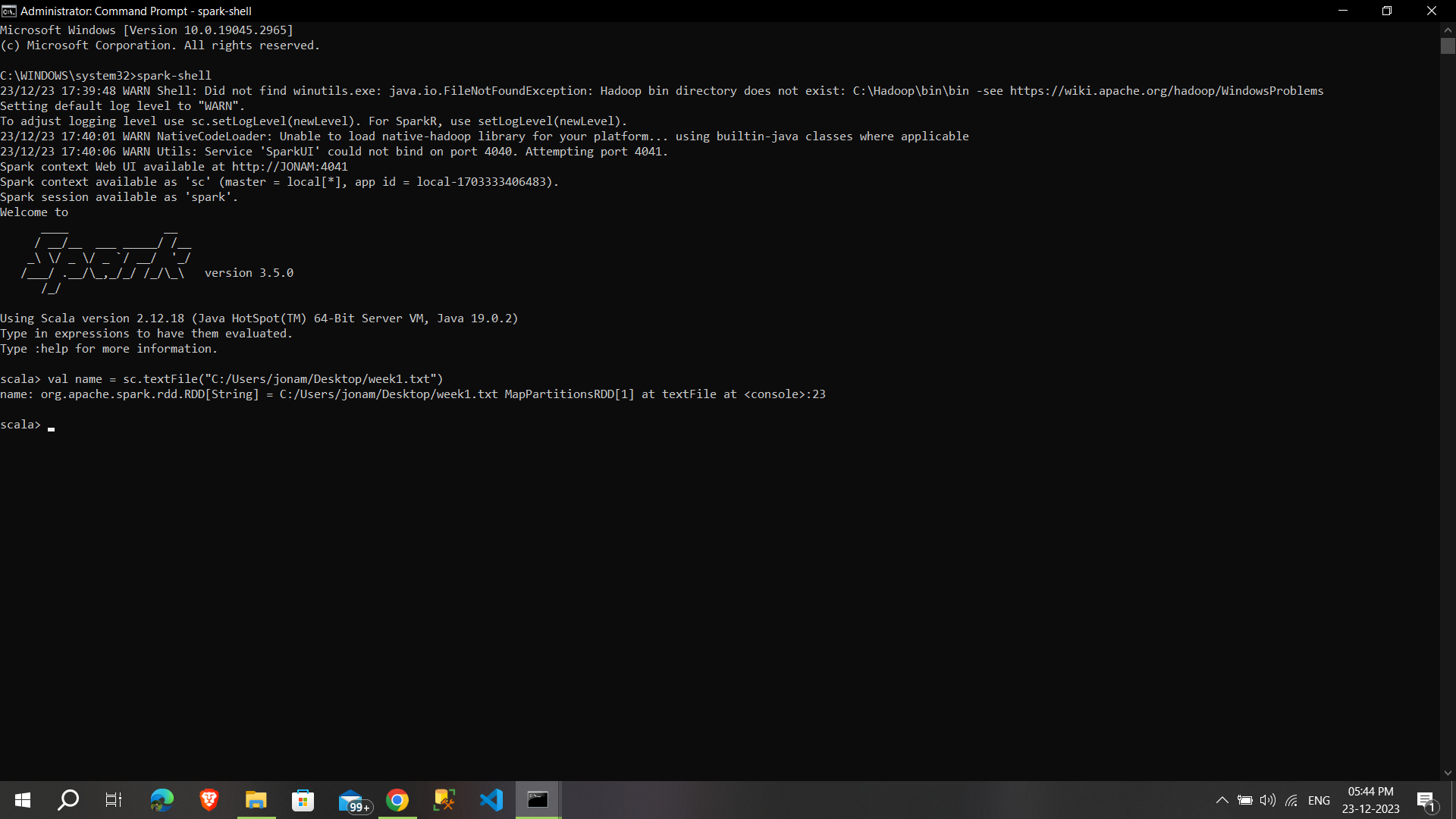


**Pyspark Commands**

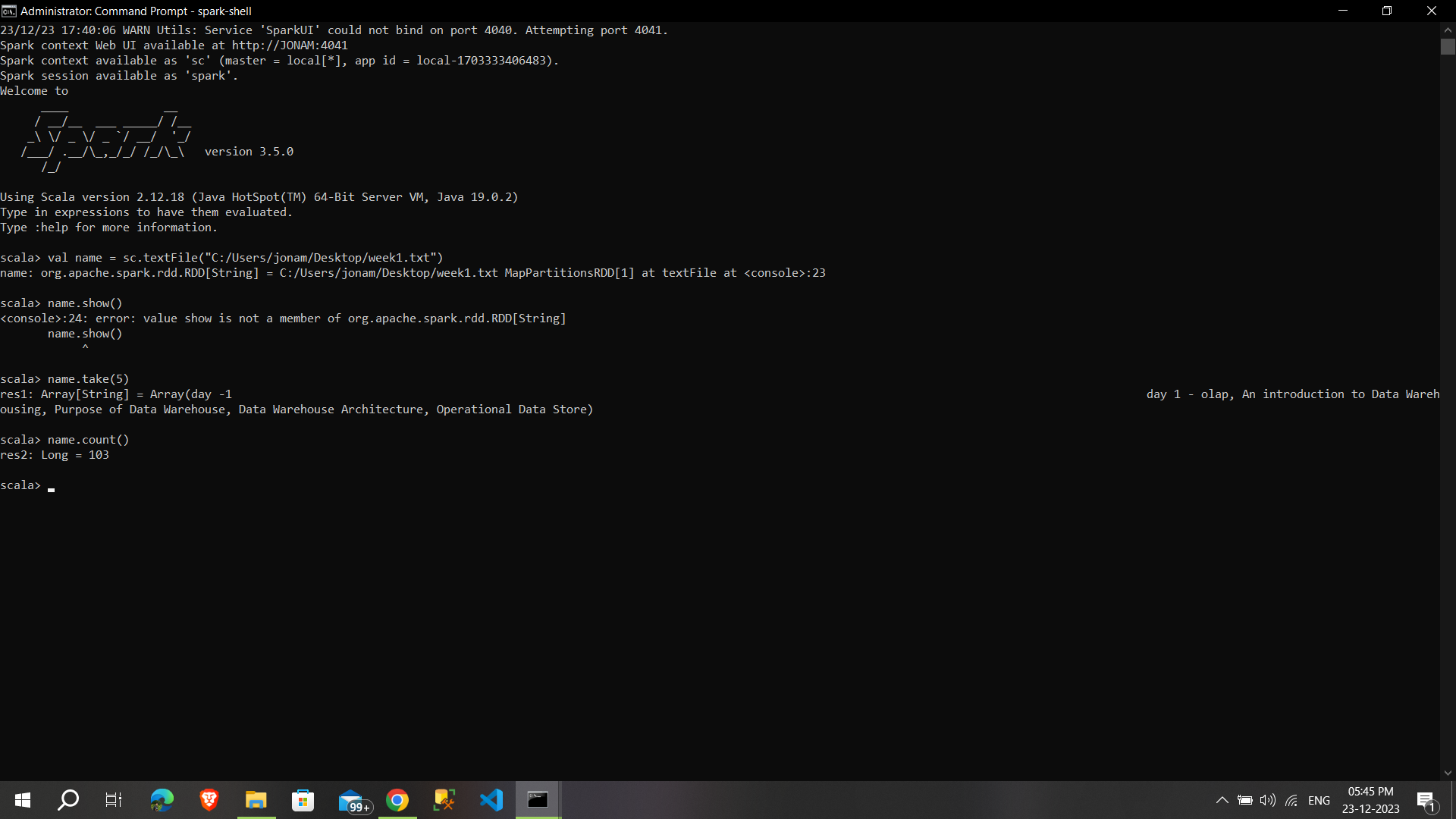
To start the **spark shell** open the command prompt type - **spark-shell**

****

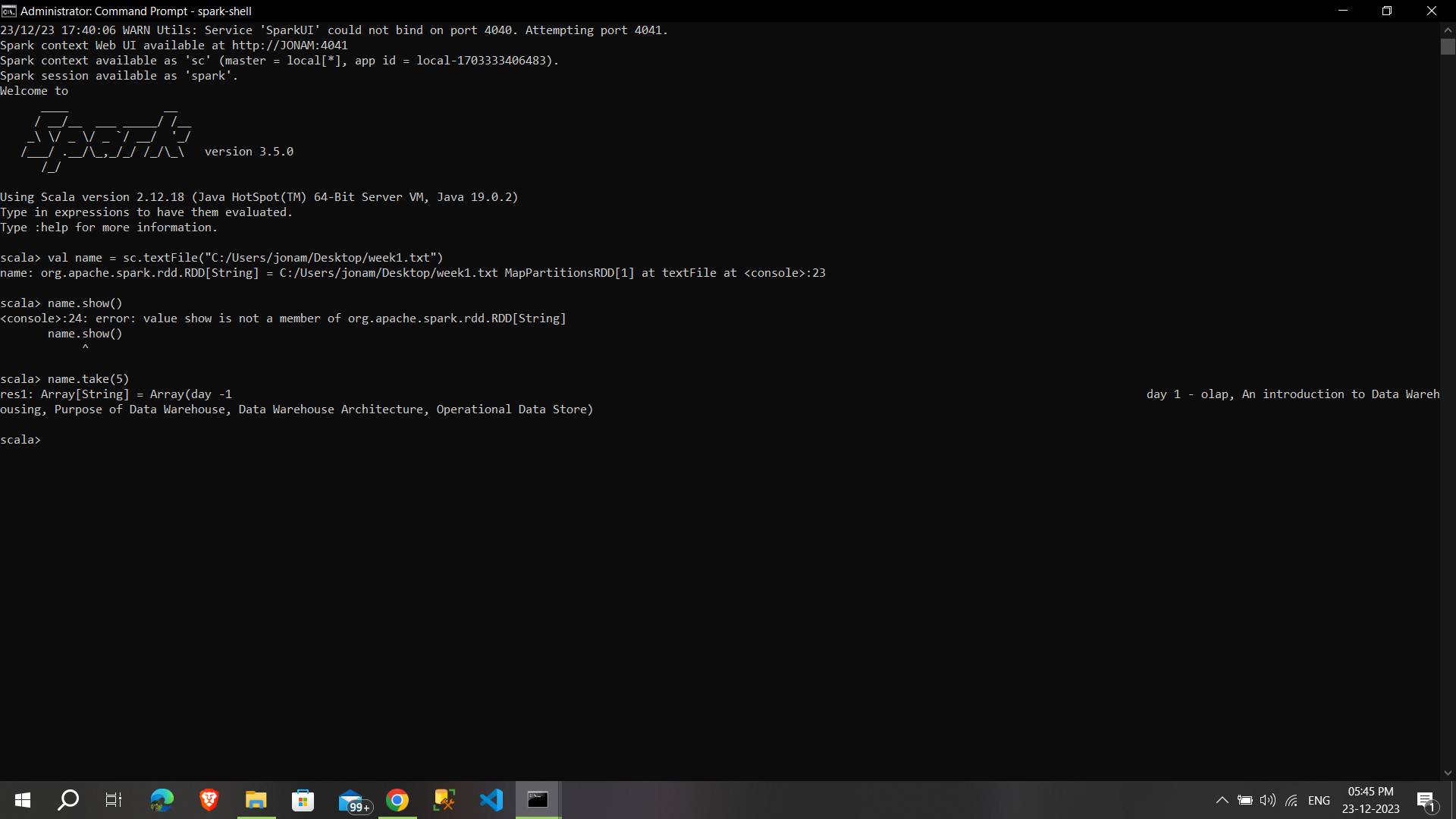
* To **Read a file** from the local system - **scala> val name = sc.textFile(“name.txt”)**

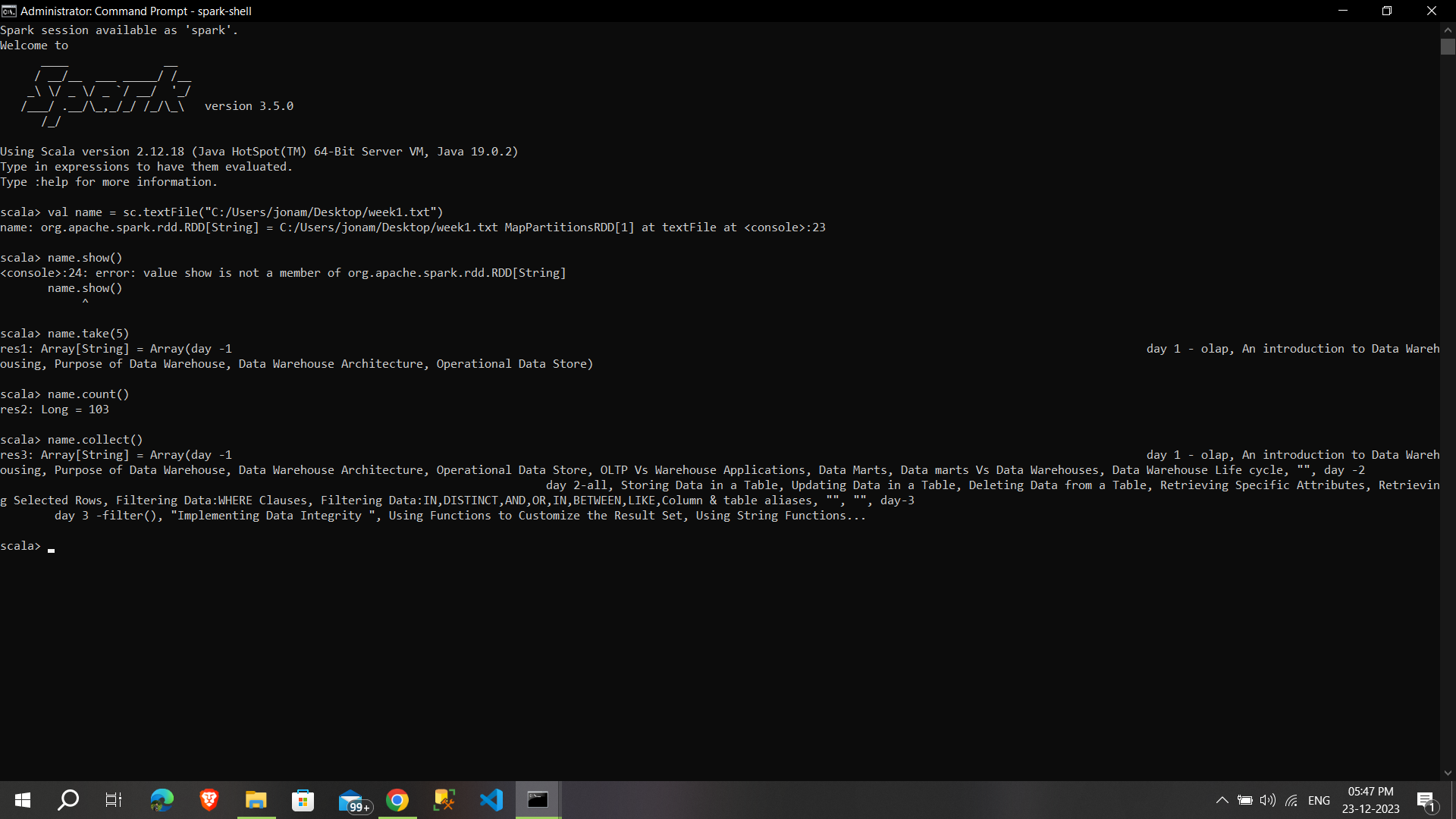
****

* To **count items** in RDD - **scala> newdata.count()**

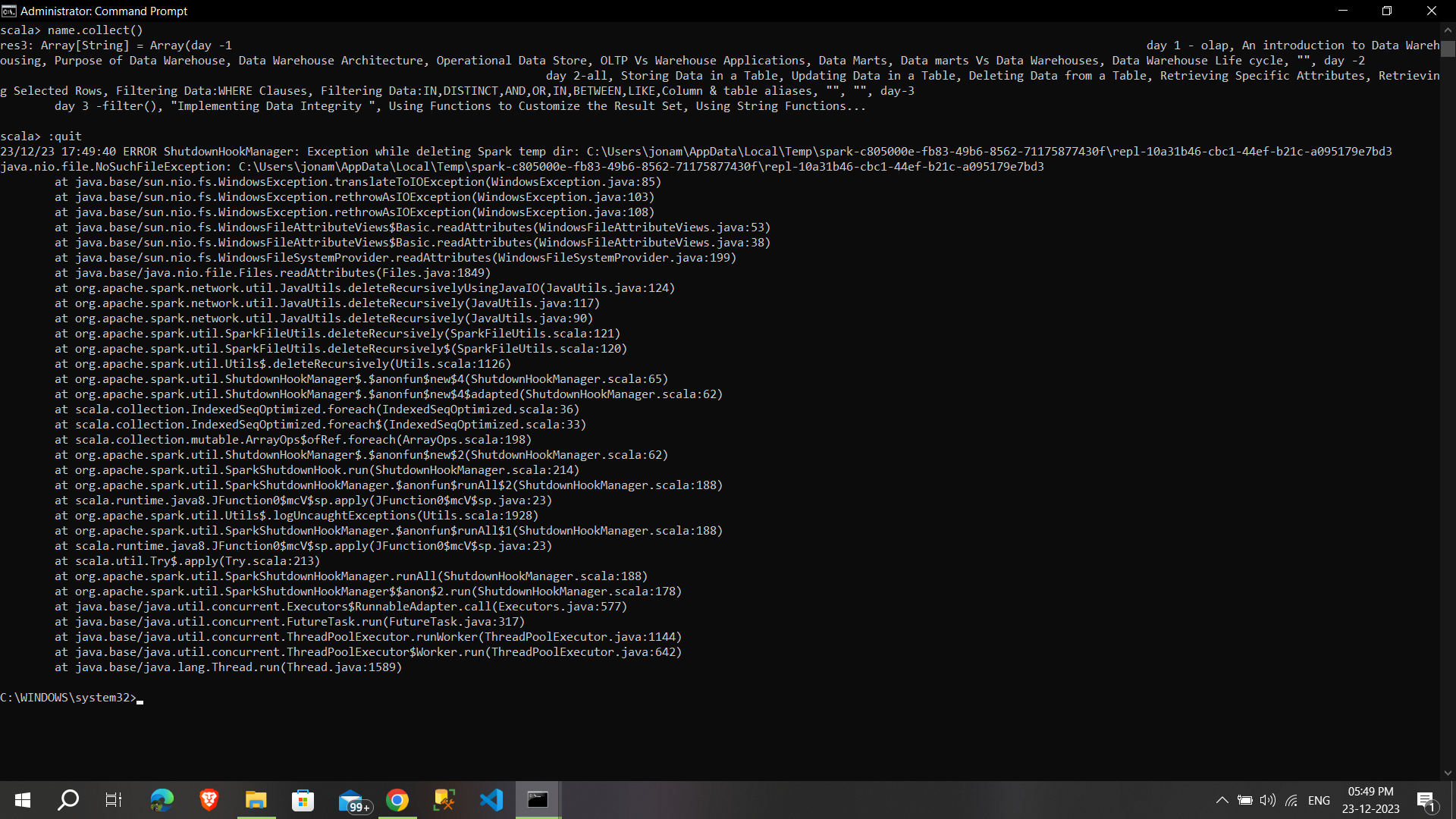
****

* To **read the first 3** items from RDD - **name.take(3)**

****

**Using collect() to show **

**Using :quit**

****