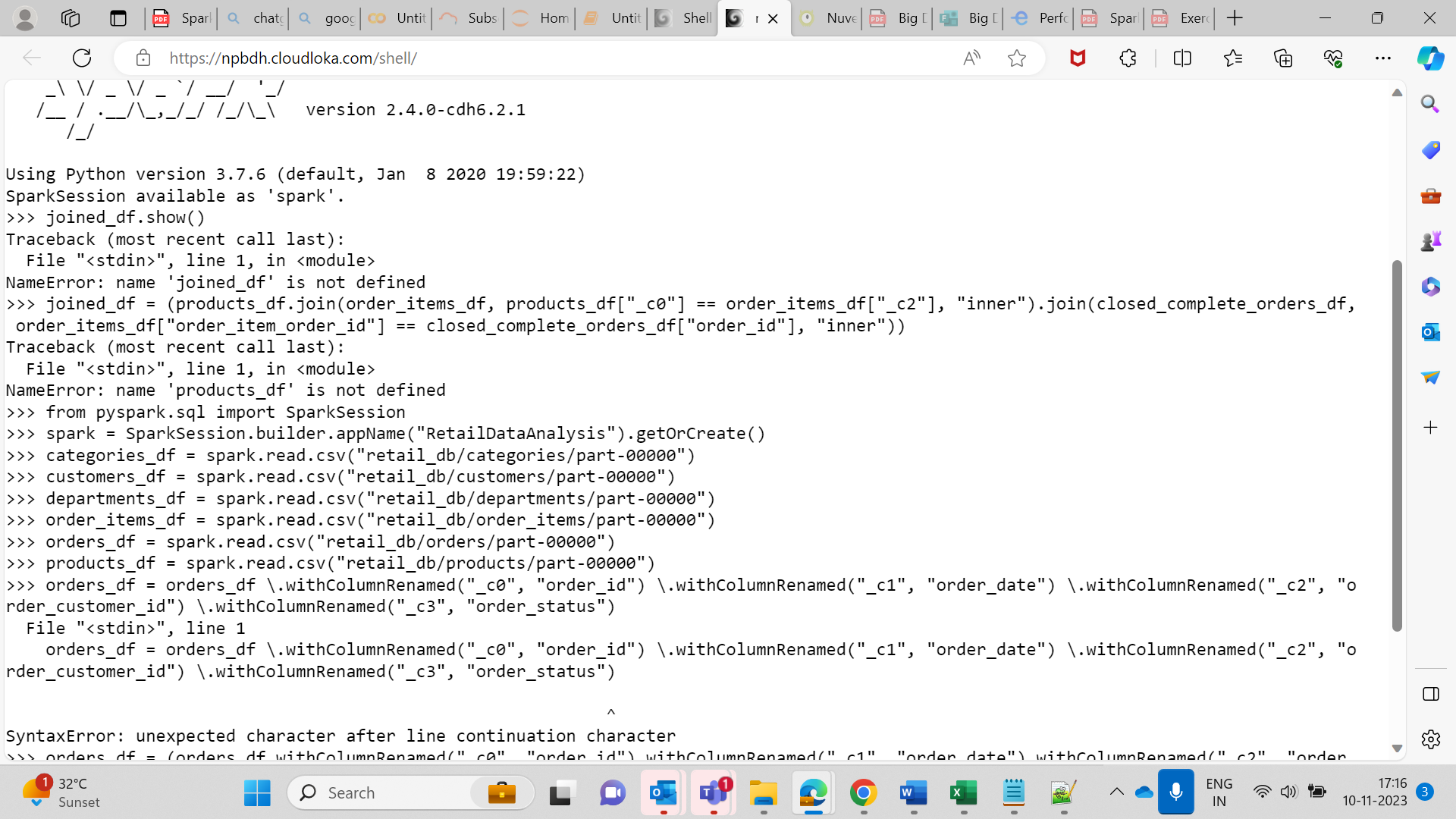
**Pyspark\_\_Assignment\_Task3\_Retail\_DB\_Analysis\_by\_Anup\_Patil**  
  
  
  
A screenshot of a computer

Description automatically generated  
  
  
**Q1 Get the count for each order status**  
  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer

Description automatically generated

order\_status\_count = orders\_df.groupBy("order\_status").count()

order\_status\_count.show()  
A screenshot of a computer

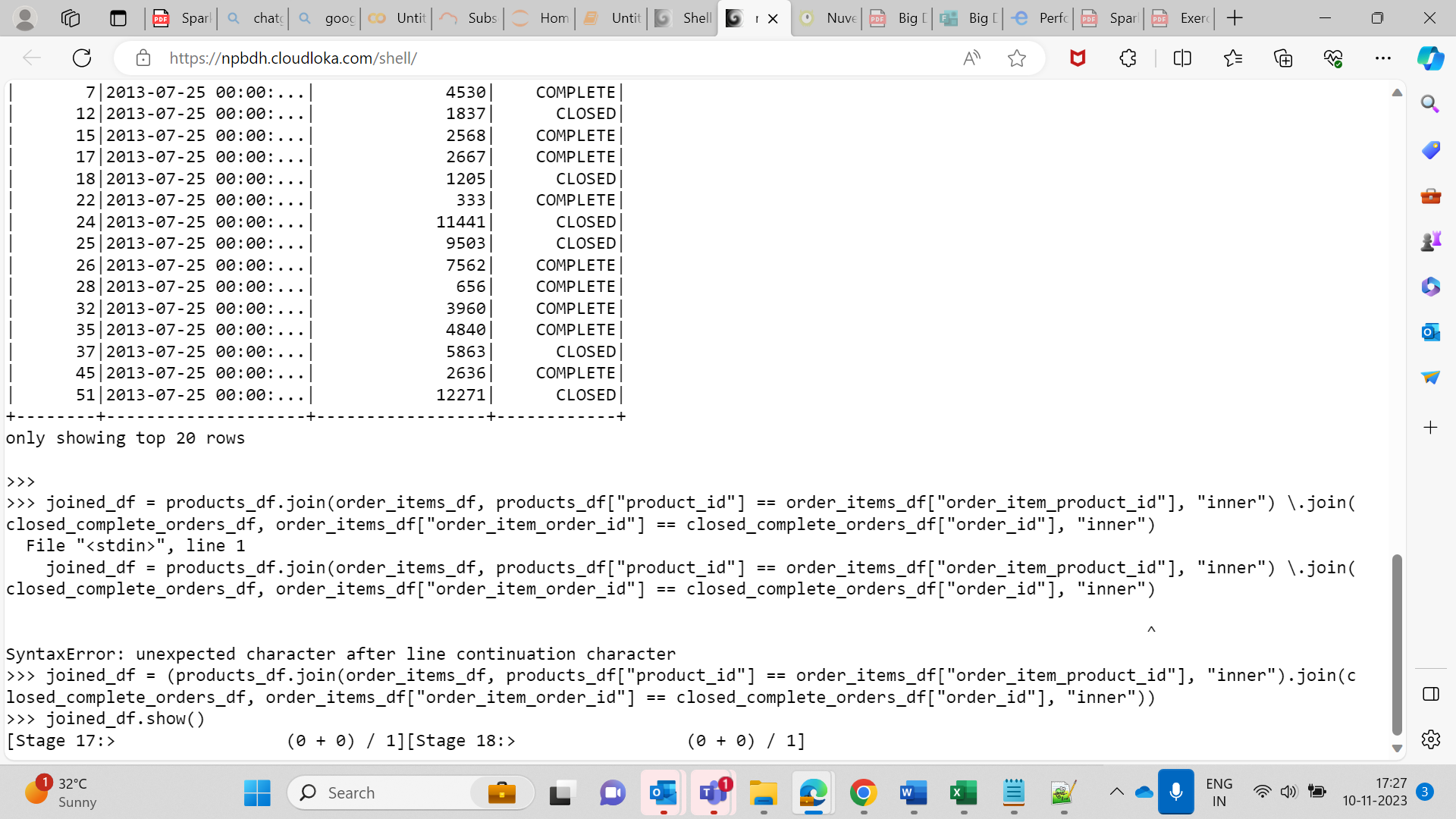
Description automatically generated  
  
  
  
**Q2. only COMPLETE or CLOSED orders**

closed\_complete\_orders\_df = orders\_df.filter((orders\_df["order\_status"] == "COMPLETE") | (orders\_df["order\_status"] == "CLOSED"))

A screenshot of a computer

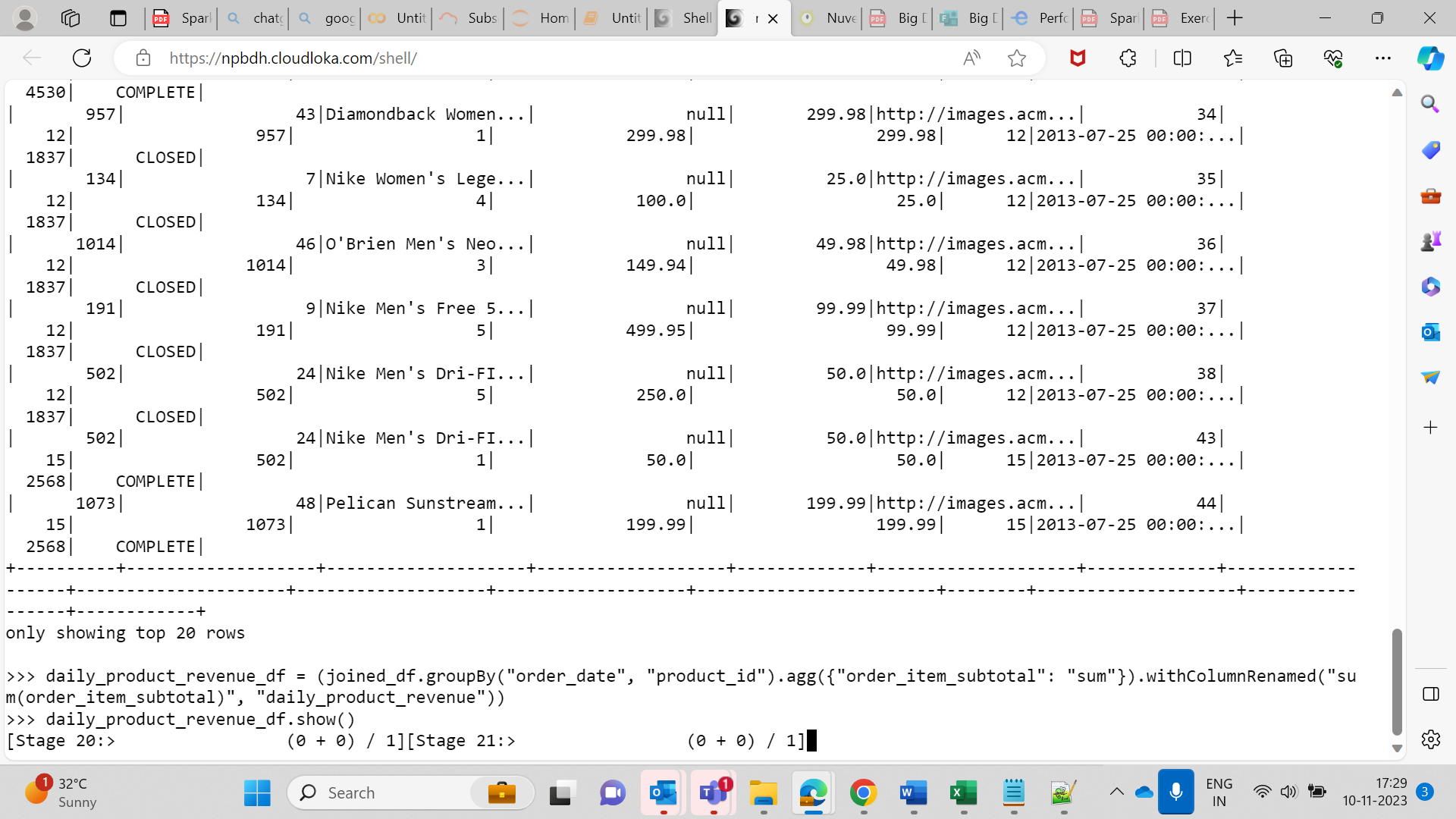
Description automatically generated  
  
A screenshot of a computer

Description automatically generated  
  
# Join the products, order\_items, and closed\_complete\_orders DataFrames

joined\_df = products\_df.join(order\_items\_df, products\_df["product\_id"] == order\_items\_df["order\_item\_product\_id"], "inner") \.join(closed\_complete\_orders\_df, order\_items\_df["order\_item\_order\_id"] == closed\_complete\_orders\_df["order\_id"], "inner")  
  
  
  
  
A screenshot of a computer

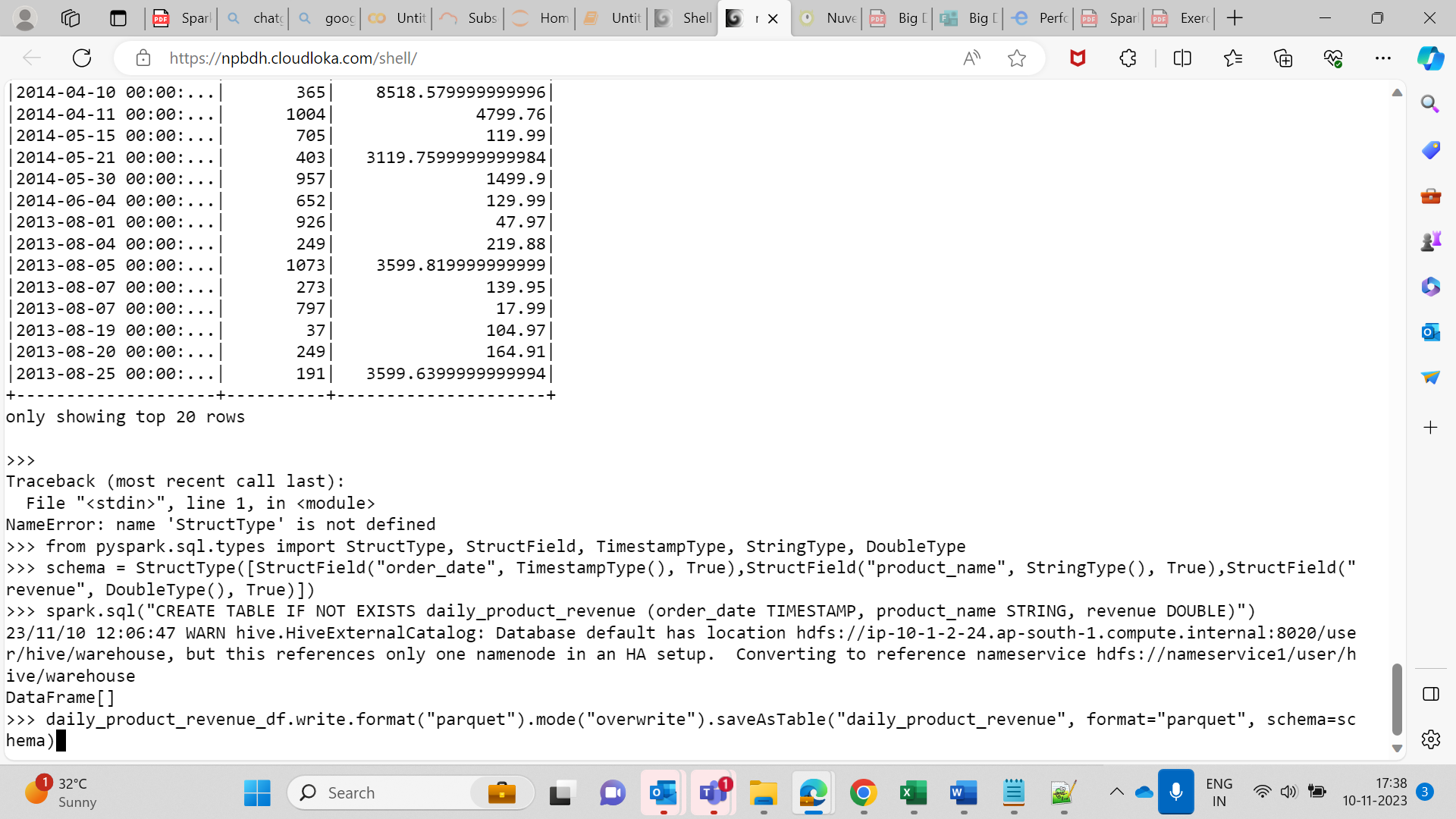
Description automatically generated  
  
  
**Q4.Calculate daily product revenue**

daily\_product\_revenue\_df = (joined\_df.groupBy("order\_date", "product\_id").agg({"order\_item\_subtotal": "sum"}).withColumnRenamed("sum(order\_item\_subtotal)", "daily\_product\_revenue"))  
  
  
A screenshot of a computer

Description automatically generated  
  
  
  
  
  
A screenshot of a computer

Description automatically generated  
  
 **Create o/p Table as Hive**  
  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer

Description automatically generated  
  
  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer

Description automatically generated  
  
  
