Week 8 Assignment- Apache Solr

Guruprasad Velikadu Krishnamoorthy

College of Science and Technology, Bellevue University

DSC650-T301: Big Data

Professor. Nasheb Ismaily

January 29, 2024

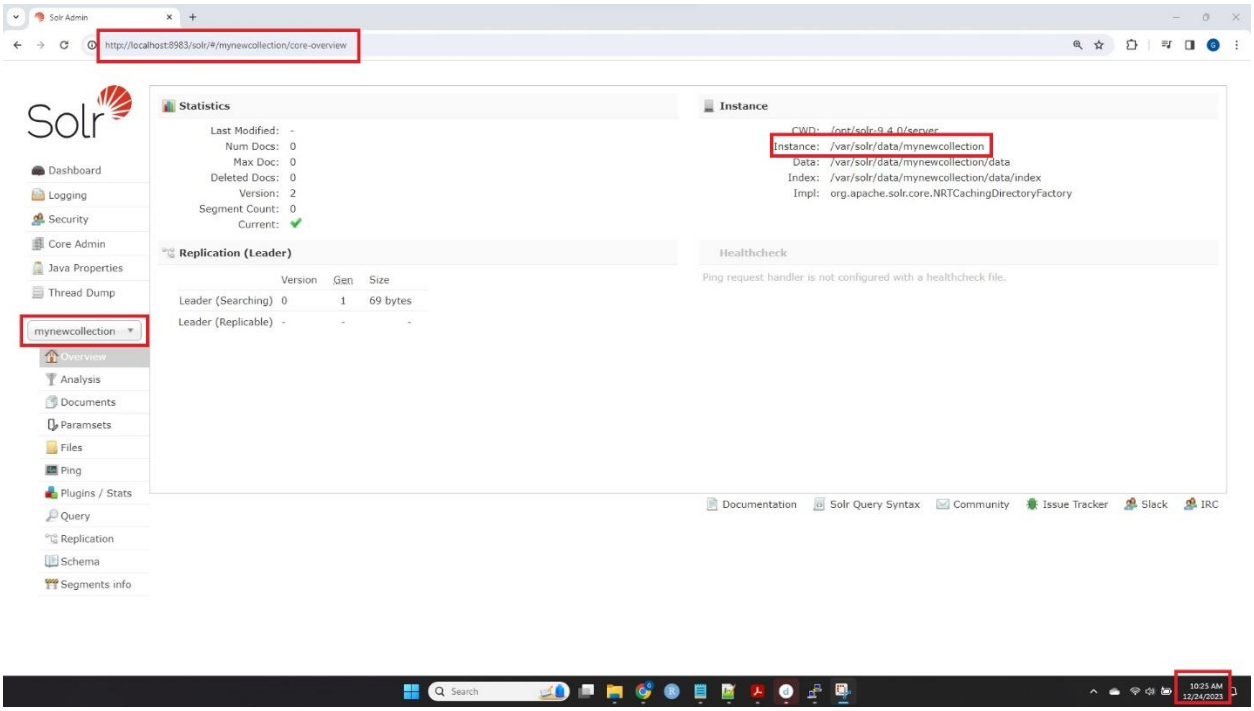
1. **Environment Initialization**

```
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$ docker-compose up -d
Creating network "solr_default" with the default driver
Creating solr_solr_1 ... done
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$ docker ps
CONTAINER ID   IMAGE     COMMAND                  CREATED         STATUS         PORTS                                       NAMES
49897ecb2126   solr      "docker-entrypoint.s…"   6 seconds ago   Up 5 seconds   0.0.0.0:8983->8983/tcp, :::8983->8983/tcp   solr_solr_1
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$
KRISUNICORN@bigdata-ds650:~/dsc650-infra/bellevue-bigdata/solr$ docker exec -it solr_solr_1 bash
solr@49897ecb2126:/opt/solr-9.4.0$
solr@49897ecb2126:/opt/solr-9.4.0$
```

2. **Creating a Solr Collection**

```
/opt/solr/bin/solr create -c mynewcollection
```

Verify that 'mynewcollection' has been successfully created.



3. **Generating and Adding Data to the Collection**

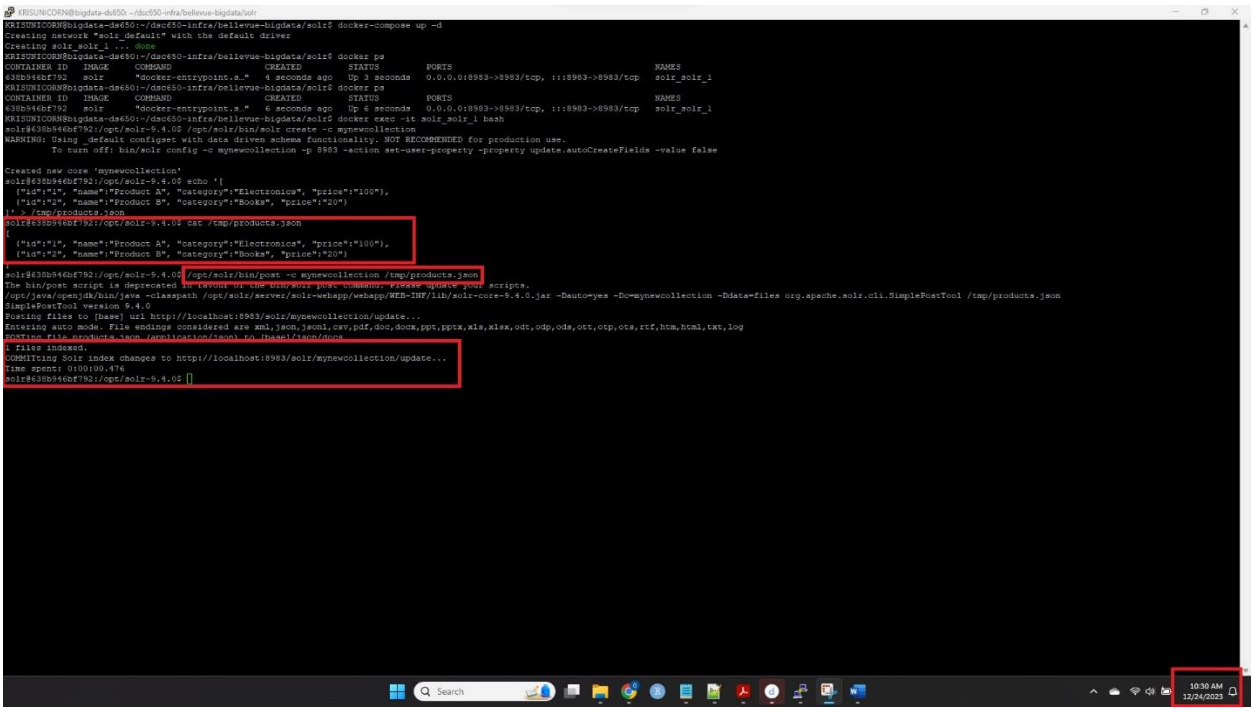
```
echo '[

 {"id":"1", "name":"Product A", "category":"Electronics",
"price":"100"},

  {"id":"2", "name":"Product B", "category":"Books",
"price":"20"}

]' > /tmp/products.json
```

```
/opt/solr/bin/post -c mynewcollection /tmp/products.json
```

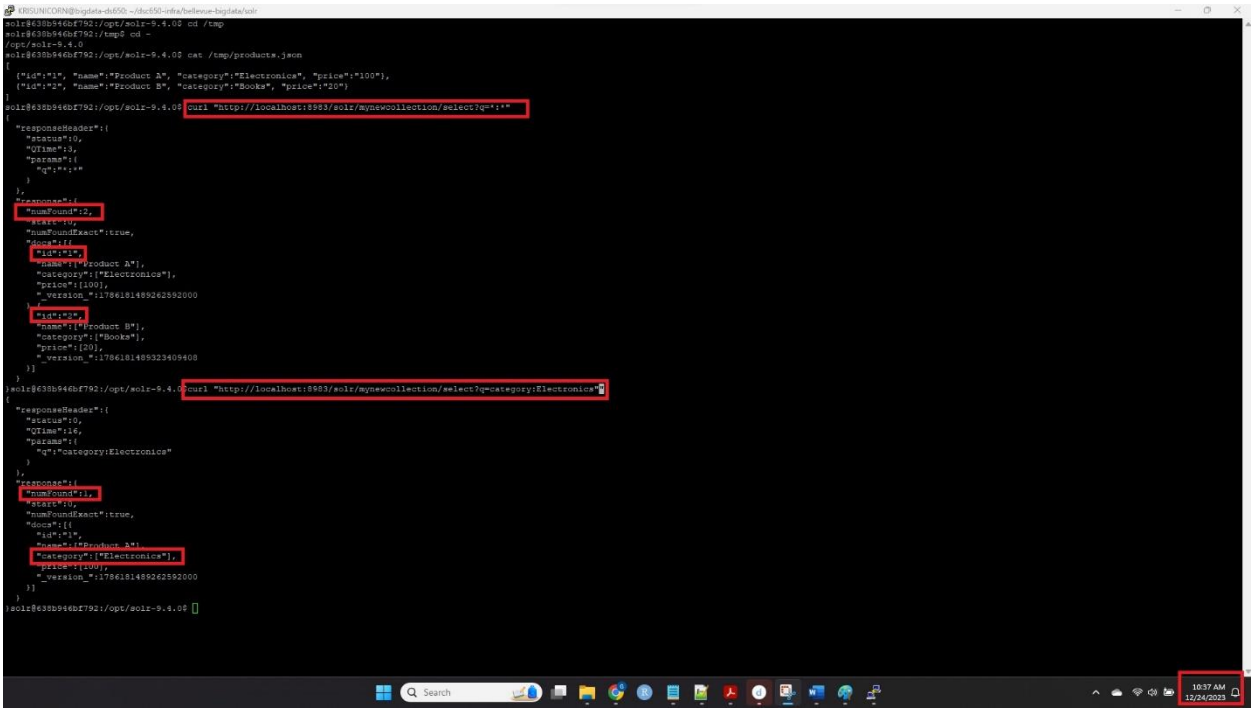


4. **Use Solr's query interface to retrieve all documents from 'mynewcollection'.**

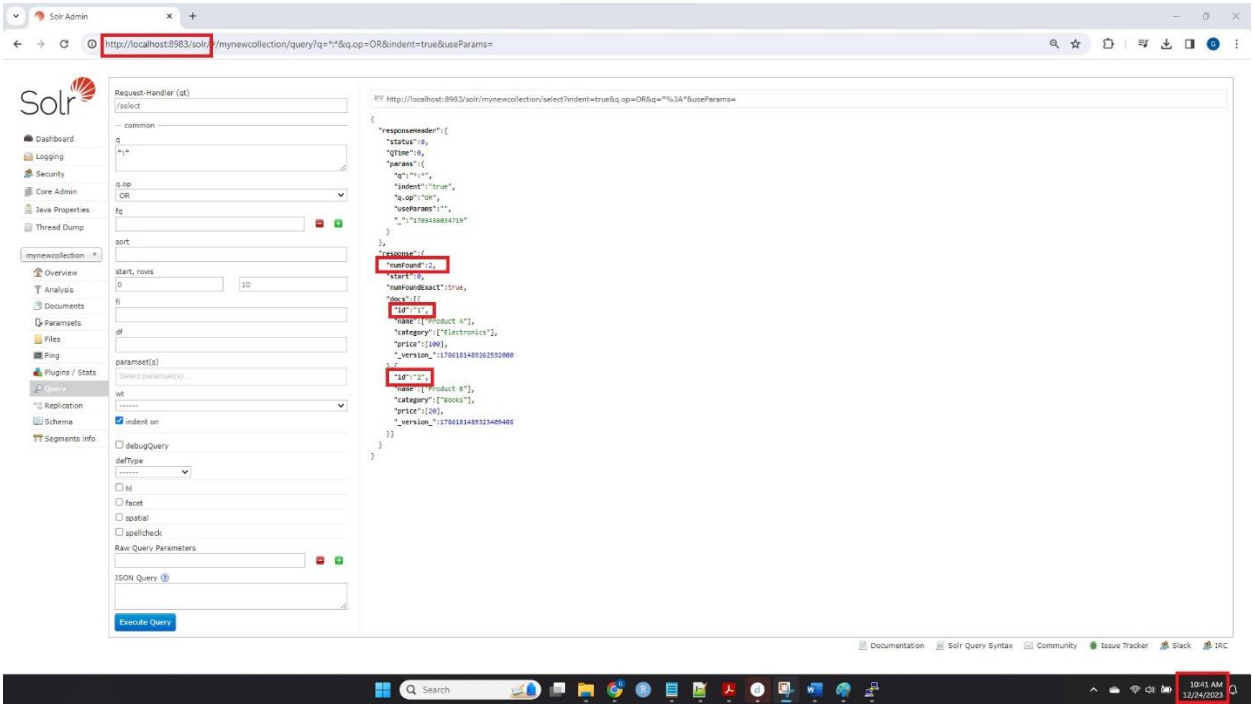
   curl http://localhost:8983/solr/mynewcollection/select?q=*:*

5. **Query for products in the 'Electronics' category.**

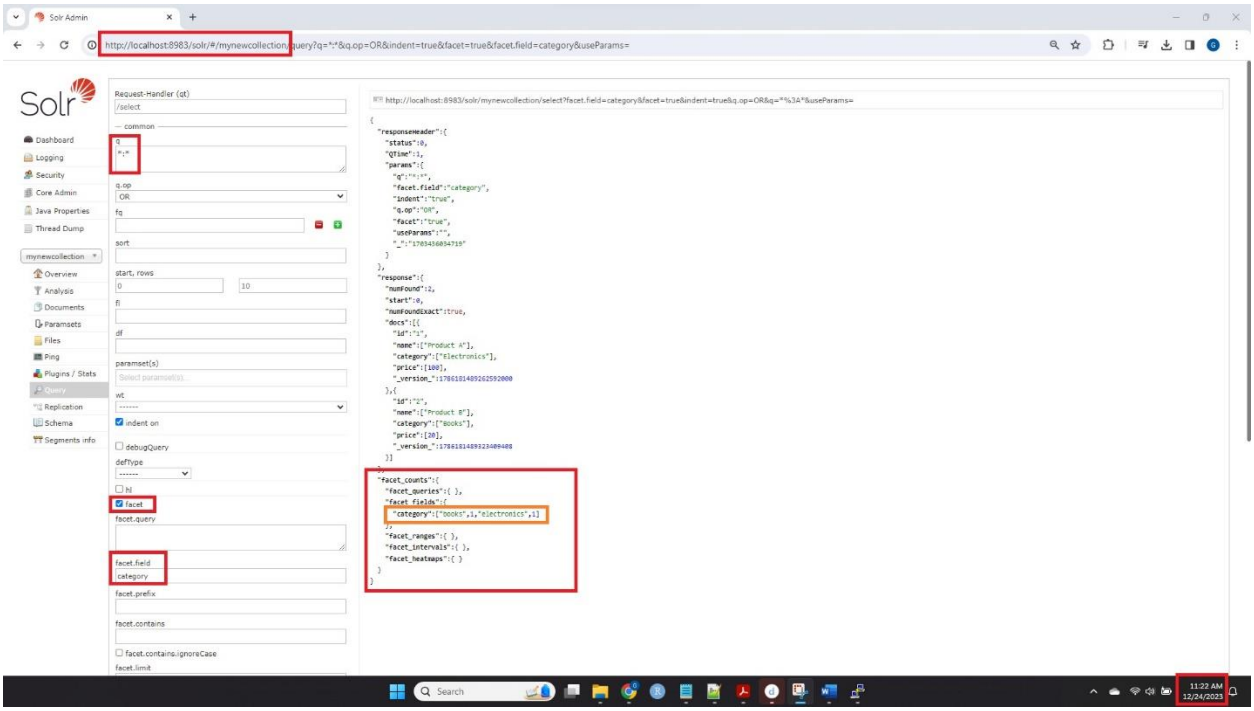
   curl http://localhost:8983/solr/mynewcollection/select?q=category:Electronics



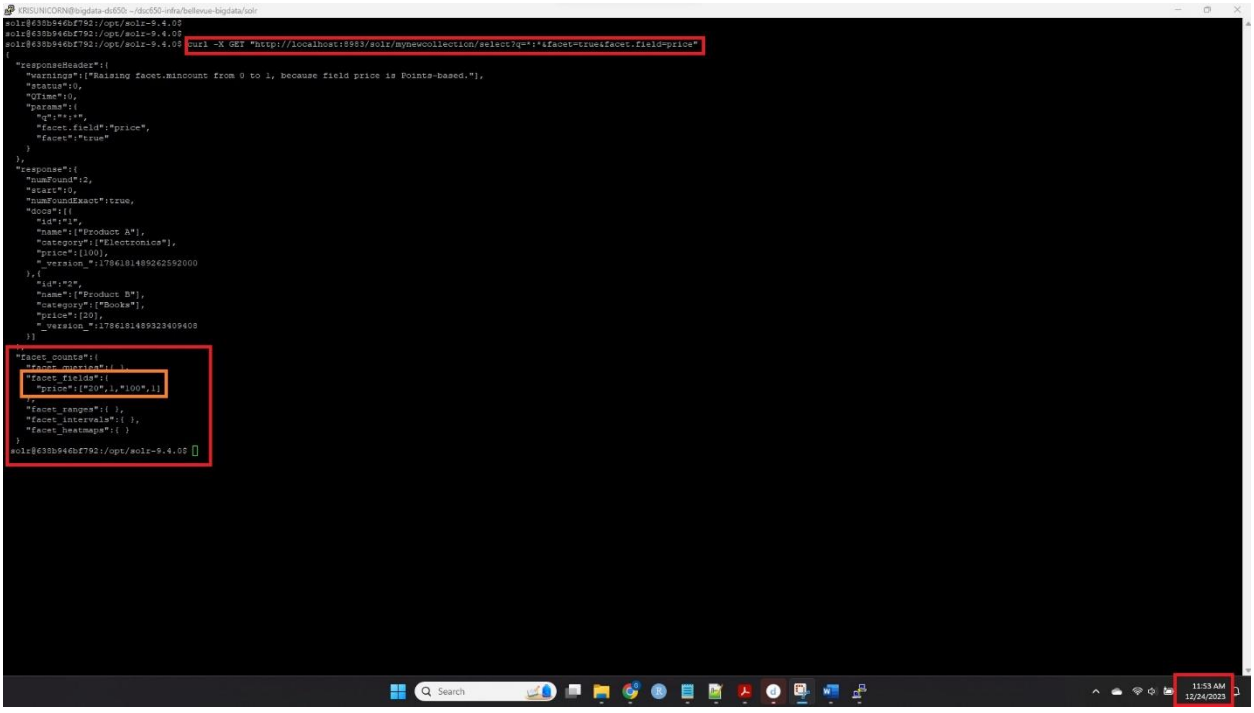
Query ‘mynewcollection’ in Solr Web Interface



6. **Try a faceted search to count the number of products in each category.**

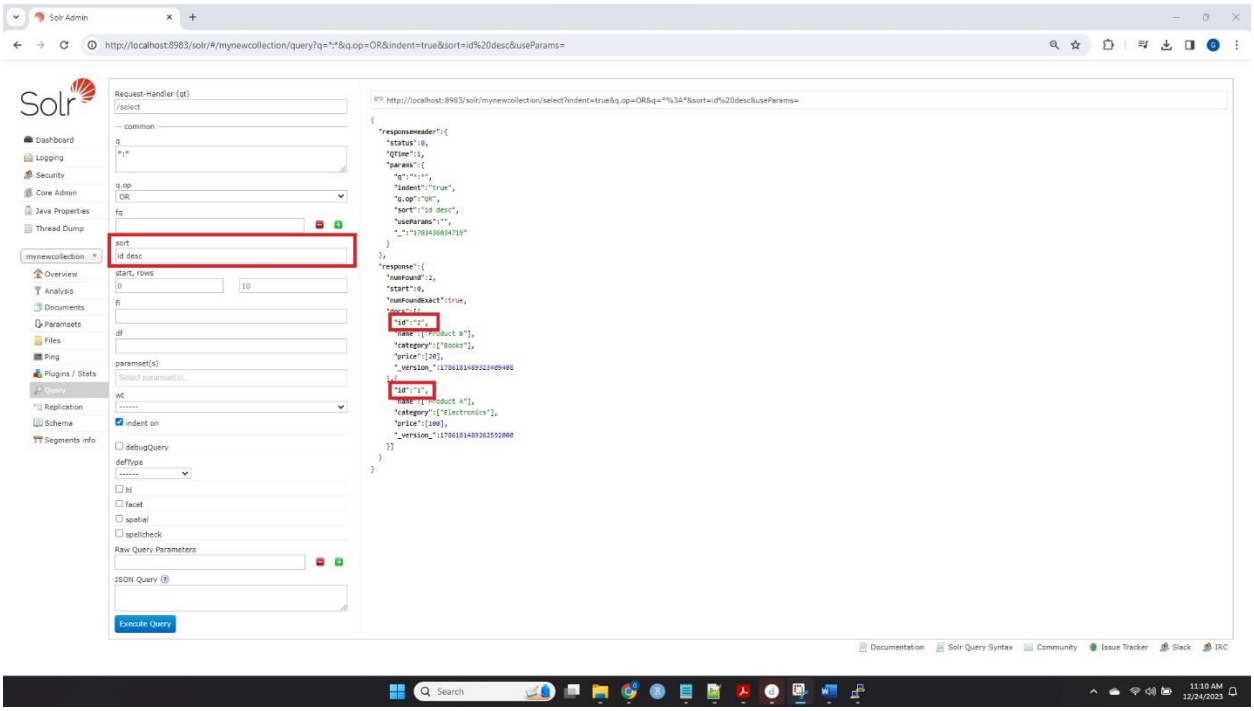


6.2. Faceting based on Price:

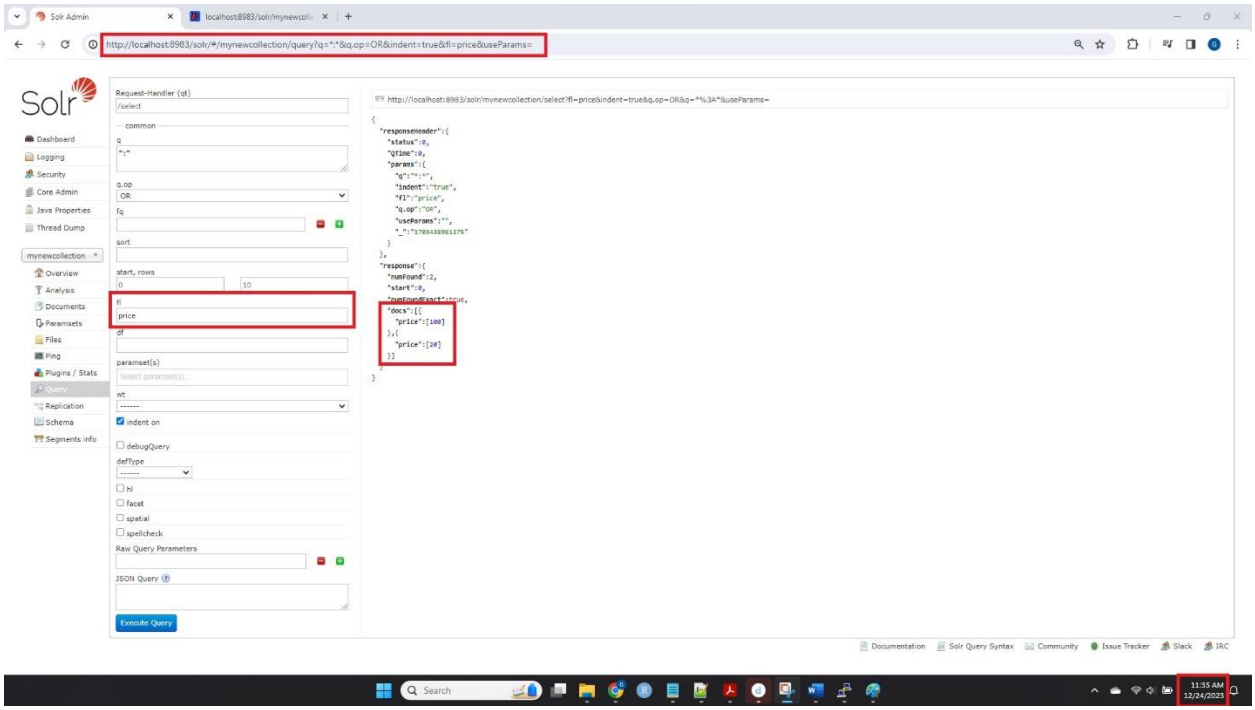


7. **Use the Web Interface to experiment further with various query parameters and filters.**

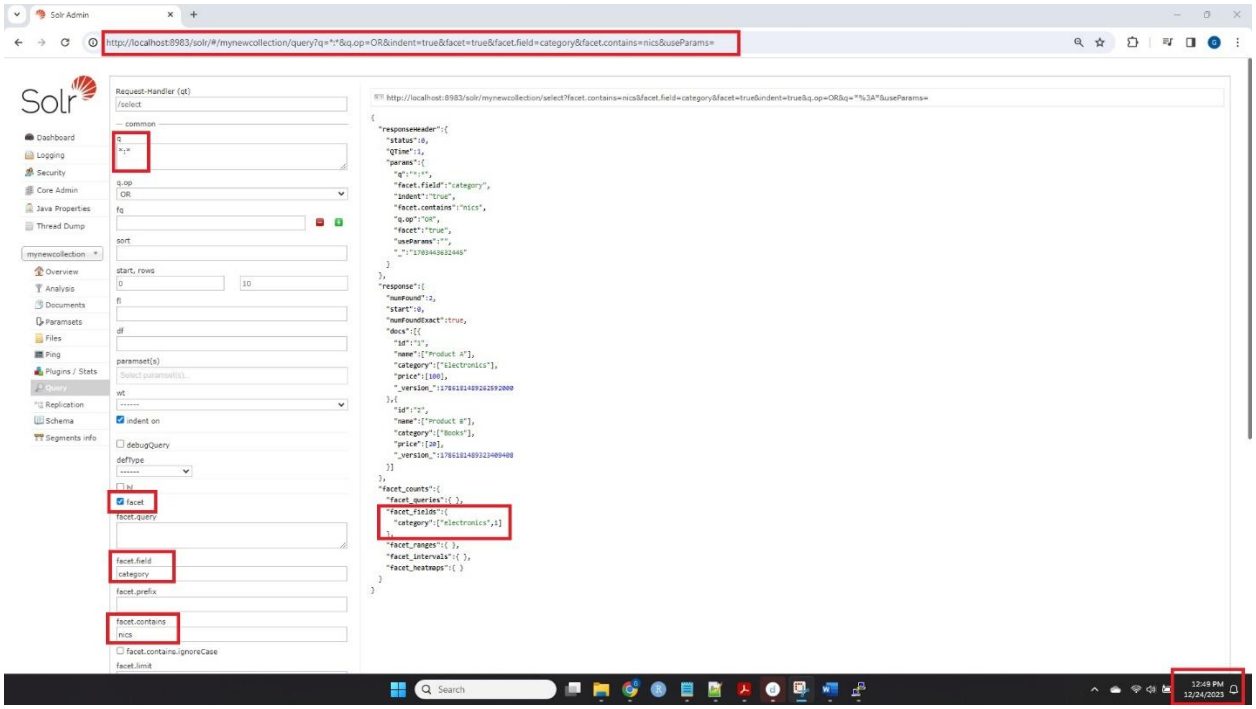
7.1. Sorting the data based on the “id” column in descending order:



7.2. Using “Price” in the field list filter:



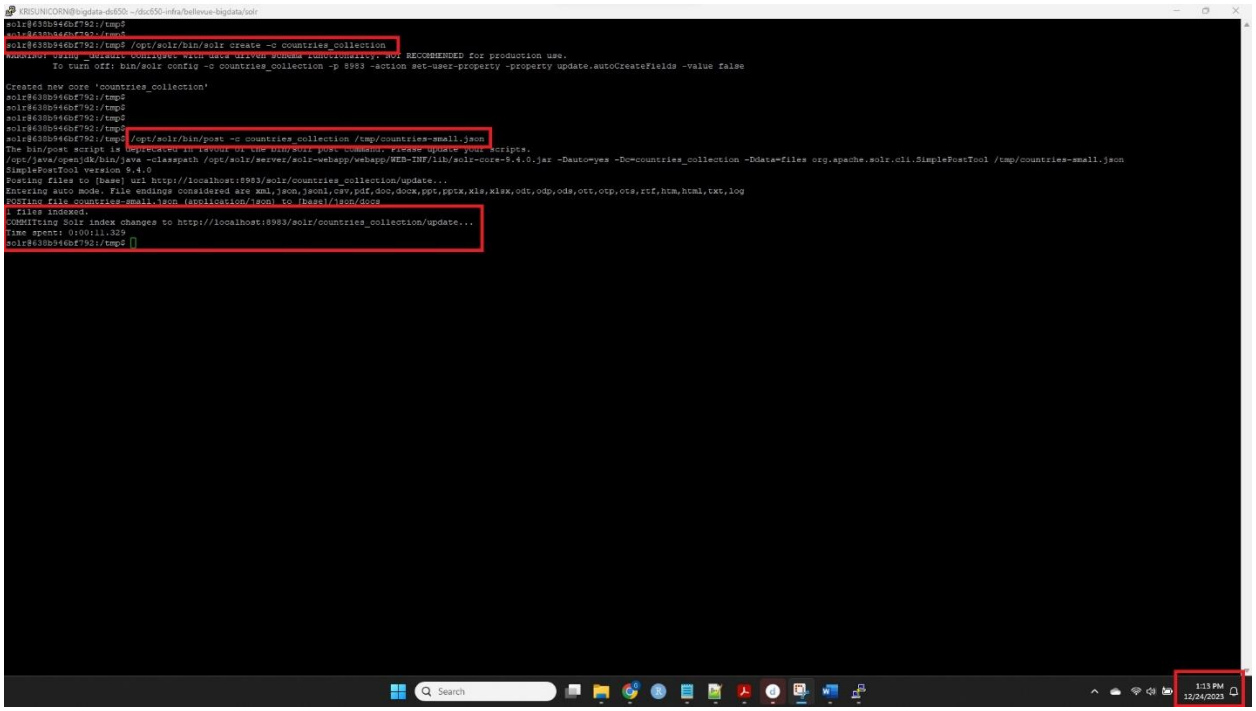
7.3. Using facet on “category” and facet.contains=”nics”



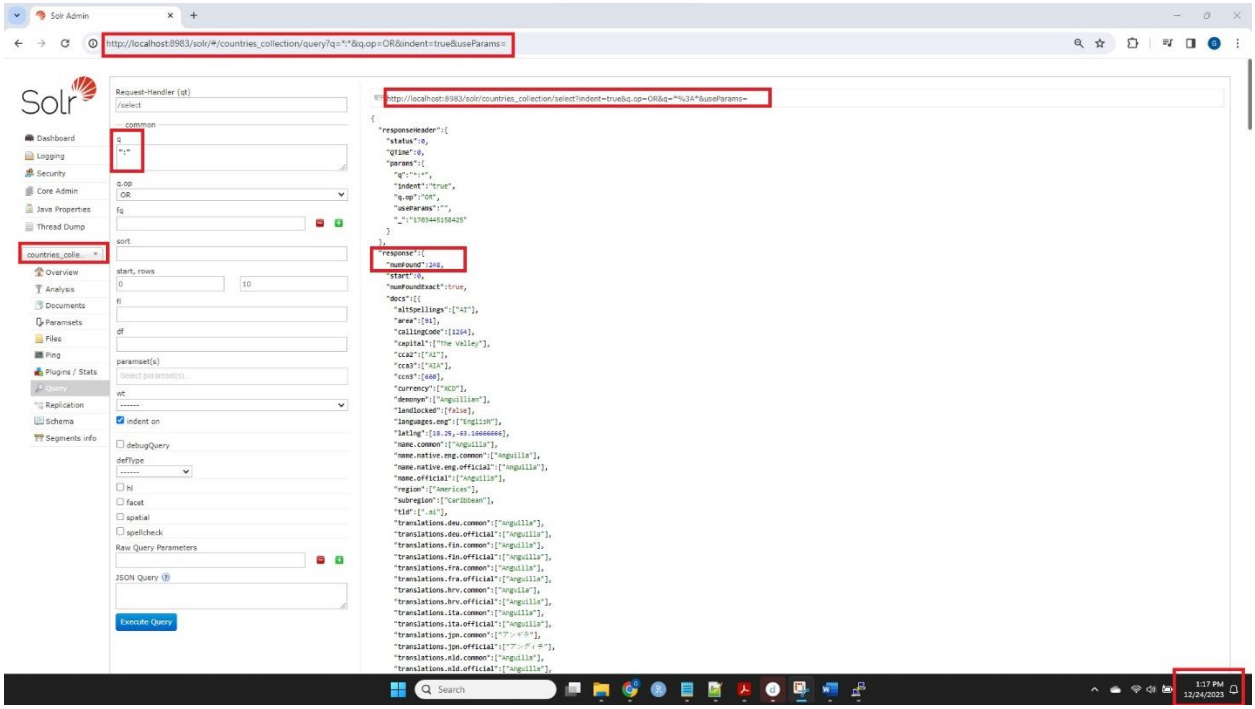
8. Custom JSON Data

8.1. Create a new collection in Solr

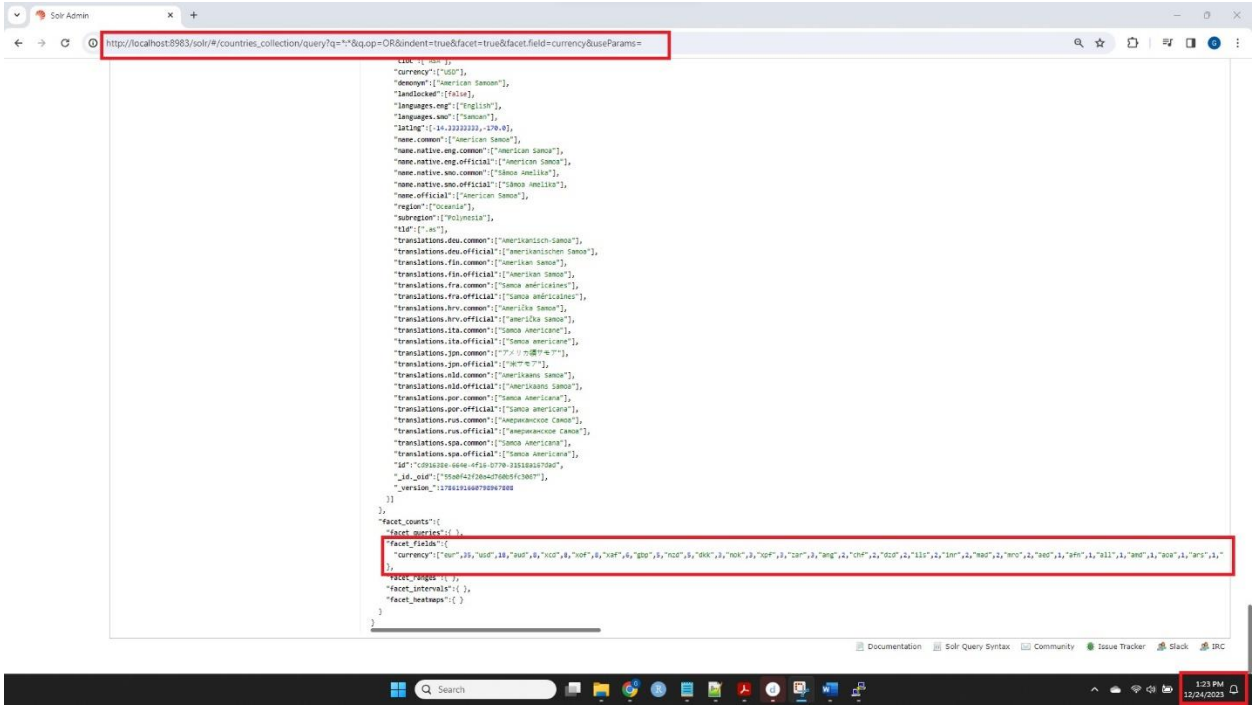
Creating a new collection named “**countries_collection**” to search the countries file.



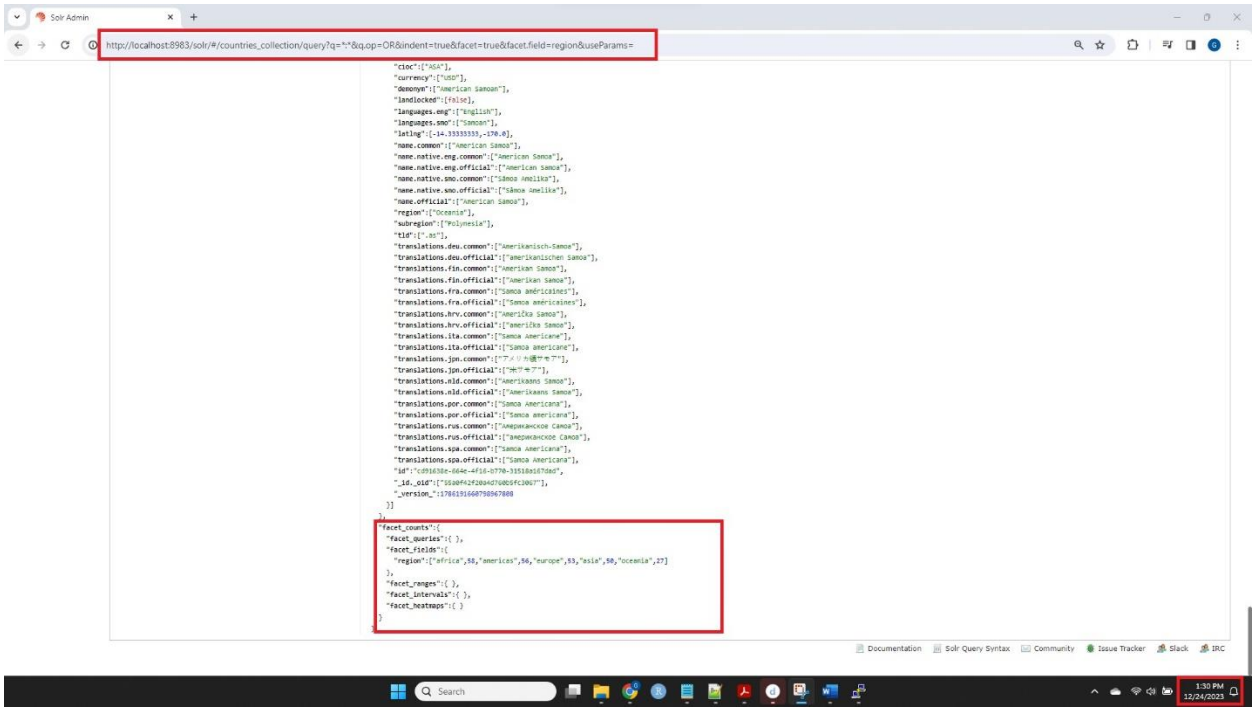
8.2. Querying the countries collection that returns 248 items



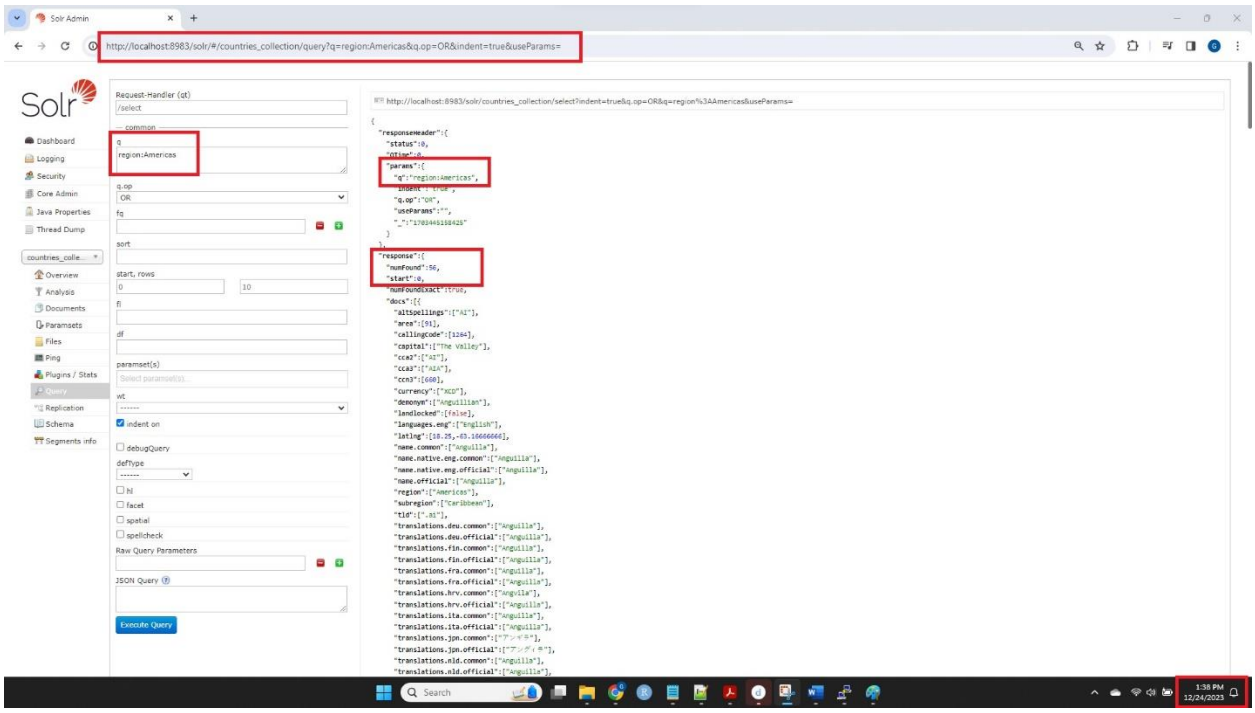
8.3. Faceting on the “currency” field in the collection



8.4. Faceting on the "regions" field in the collection:



8.5. Filtering for the region that contains “Americas”



8.6. Applying Multiple filters

In this example, the results are sorted by the area in ***descending order***. Only the top 5 rows from the results are extracted using the ***start,rows*** option. The results only returns selected fields mentioned in the "***field list***" box. The subregion is set as the "***default field***" and the default value of "America" is included in the ***"query"*** box.