**NAME: HARISH CHANDRA JYOSHI**

**CLASS ID: 06**

**TEAM ID: 04**

**UMKC EMAIL ID:**[**hjddh@mail.umkc.edu**](mailto:hjddh@mail.umkc.edu)

**NAME: ATLURI VENKATA AKHILA KRISHNA**

**CLASS ID: 01**

**TEAM ID: 04**

**UMKC EMAIL ID:**[**vagq2@mail.umkc.edu**](mailto:vagq2@mail.umkc.edu)

**TASK 1:**

**OBJECTIVE** 1.**HIVE USECASE**

a)Create a Hive Table including Complex Data Types

b)Use built-in functions in your queries

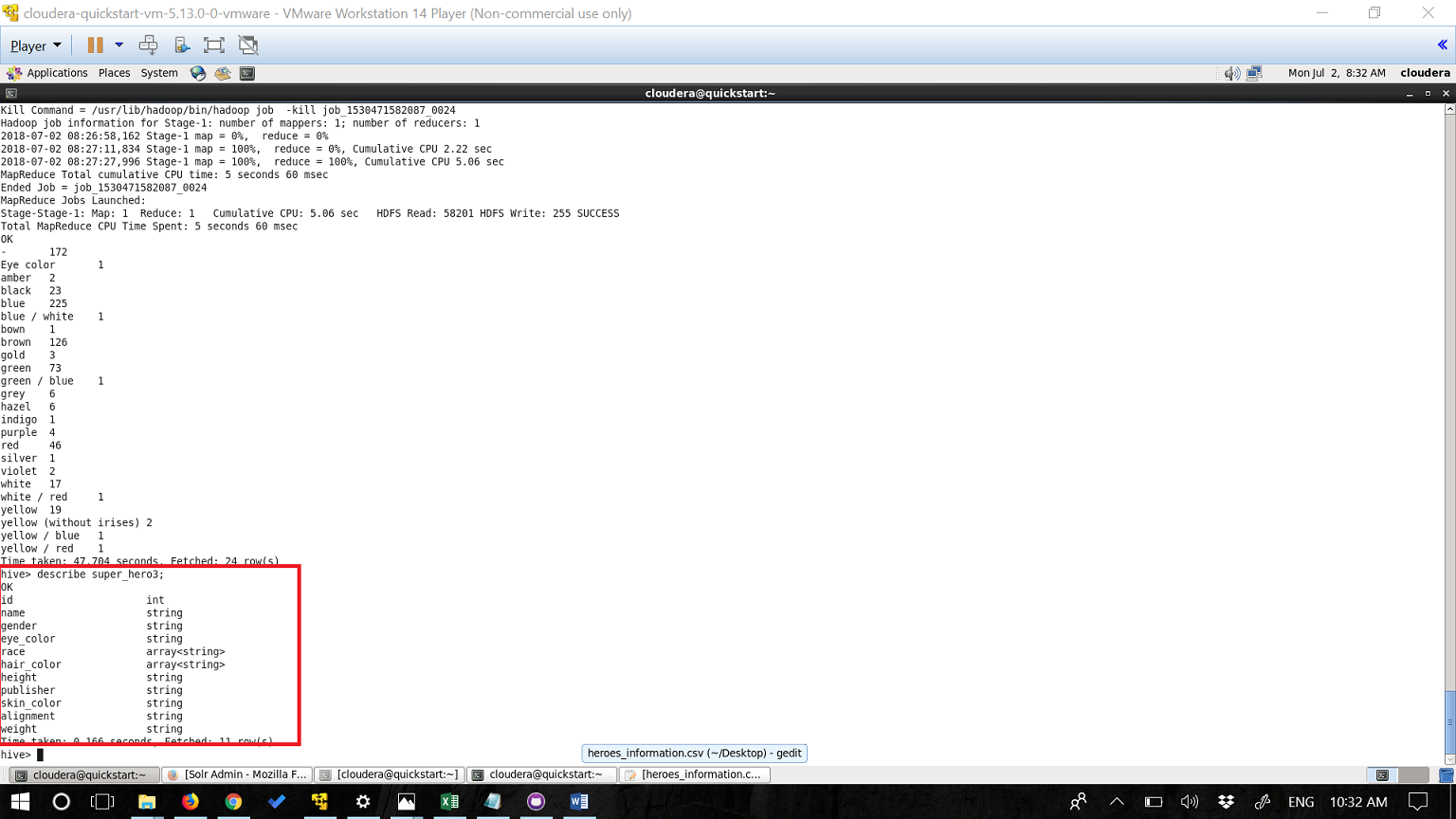
c)Perform 10 intuitive questions in Database

**DATASET used:** I have used superheroes as my dataset

**Queries:** I have created a hive table for superheroes dataset with array data type for haircolor and eye\_color column.

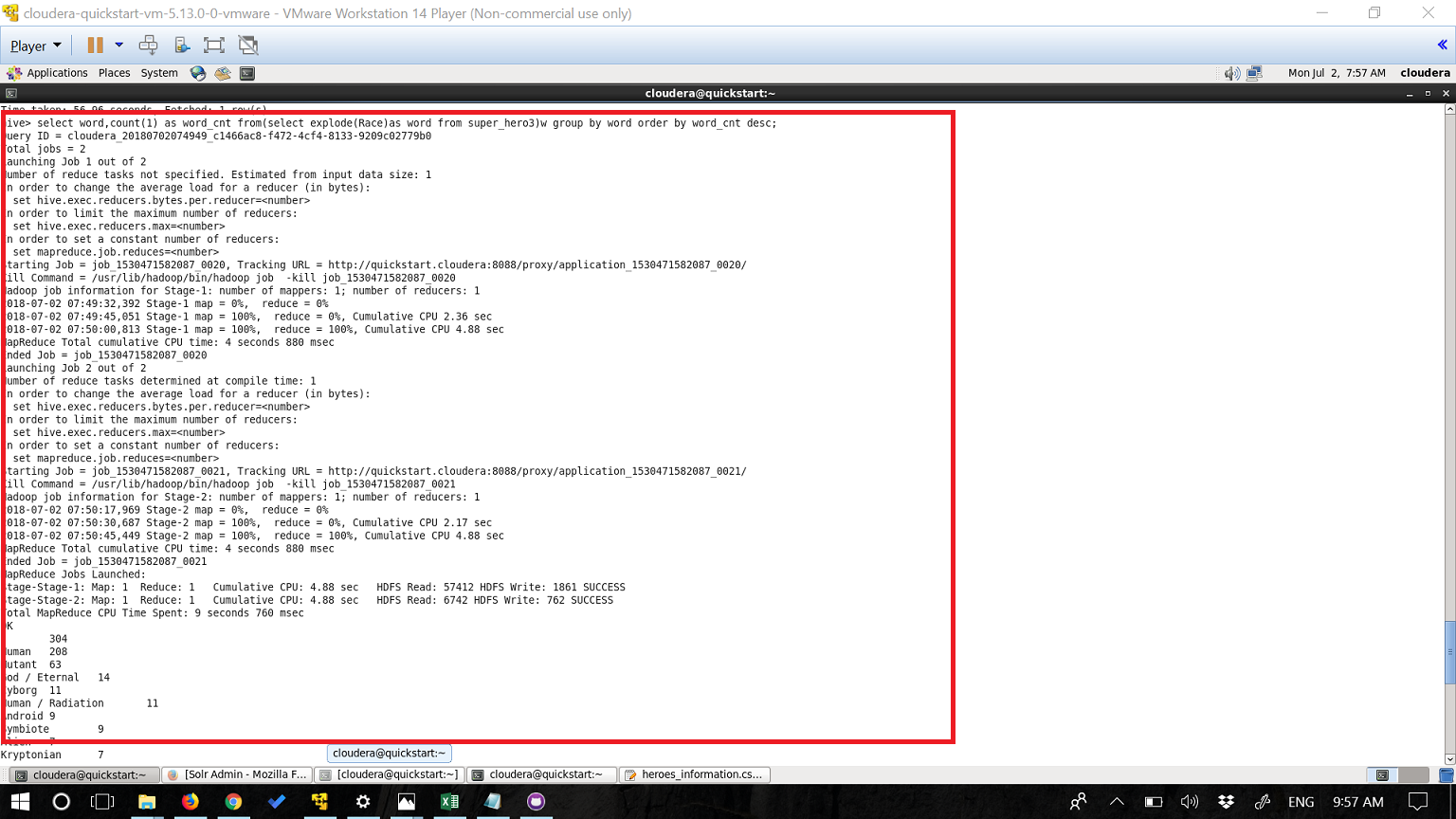
I performed the following queries.

**SUPER\_HERO TABLE:**



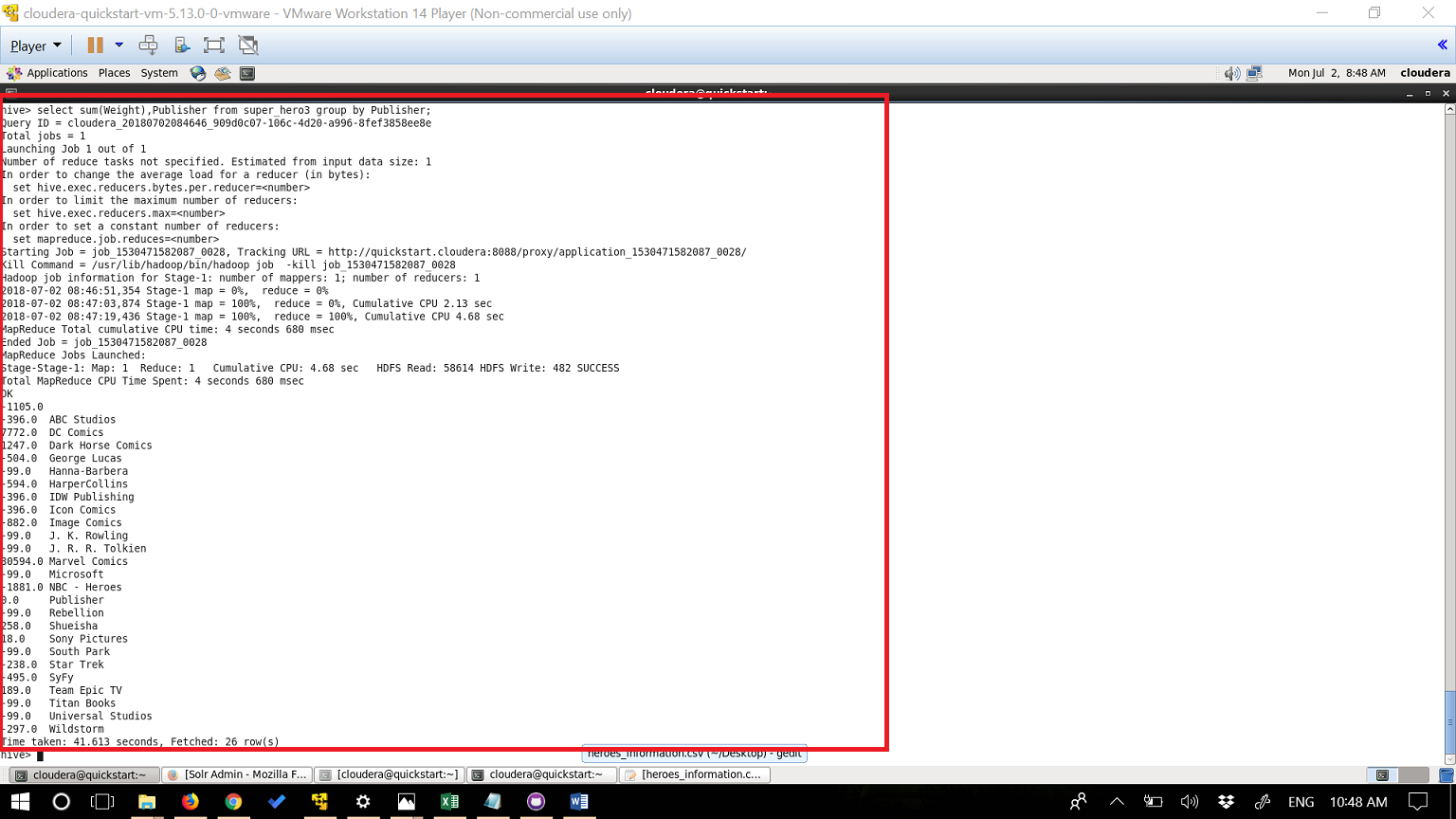
**Query1:**

**Wordcount**



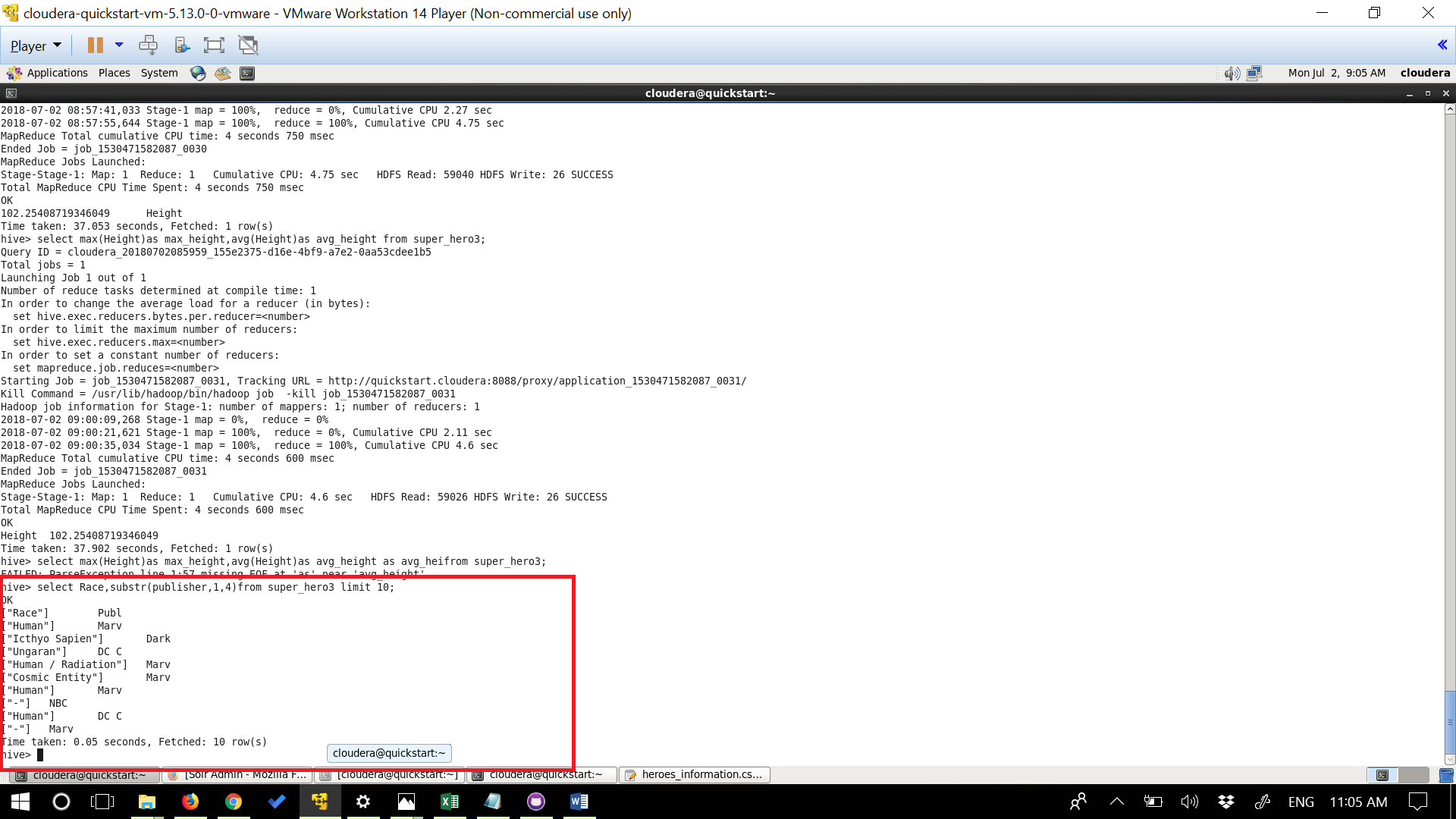
**Query2:**

**Sum of Weights grouped by Publisher**



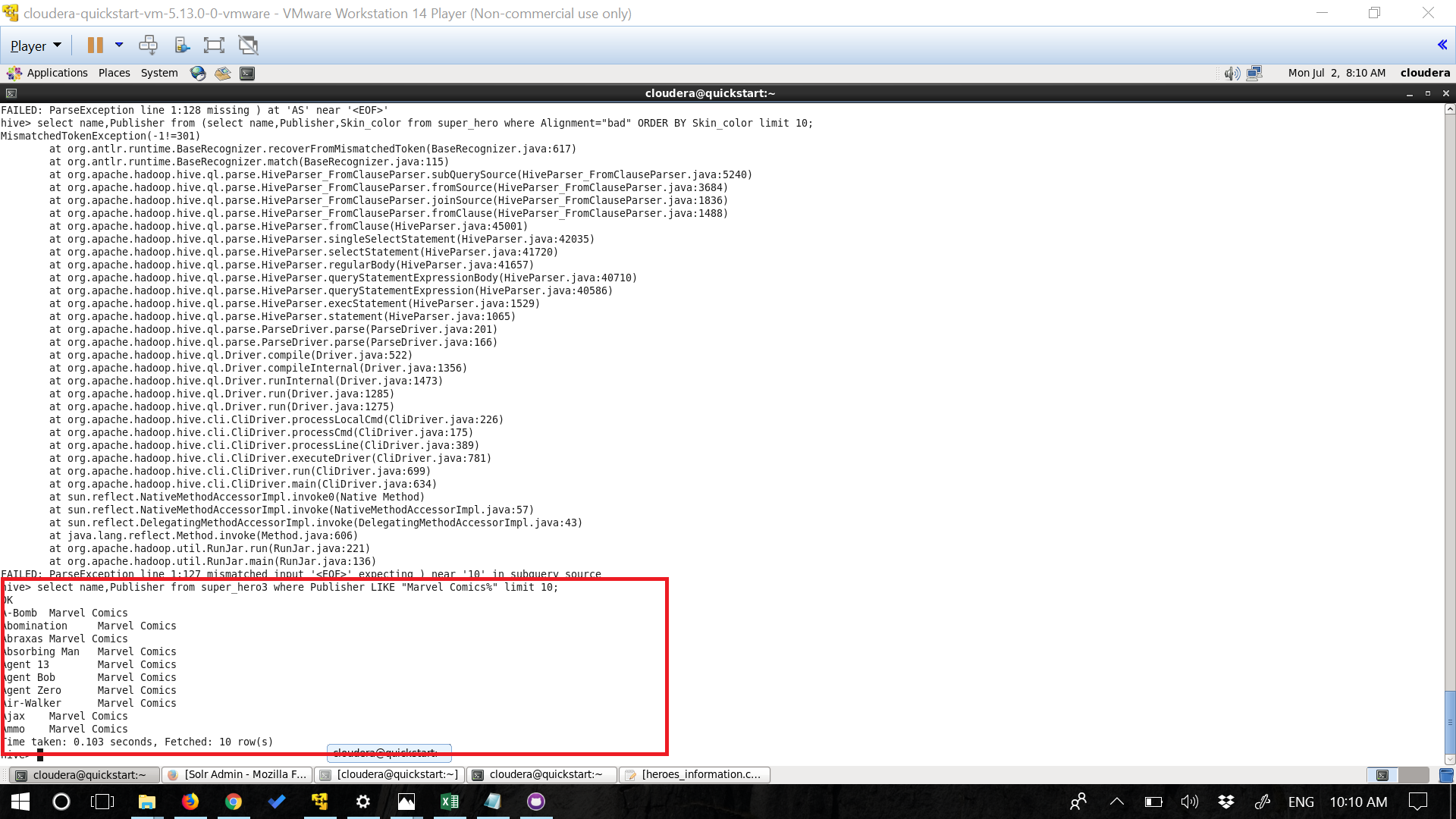
**Query3:**

**A substring of 1 t0 4 characters**



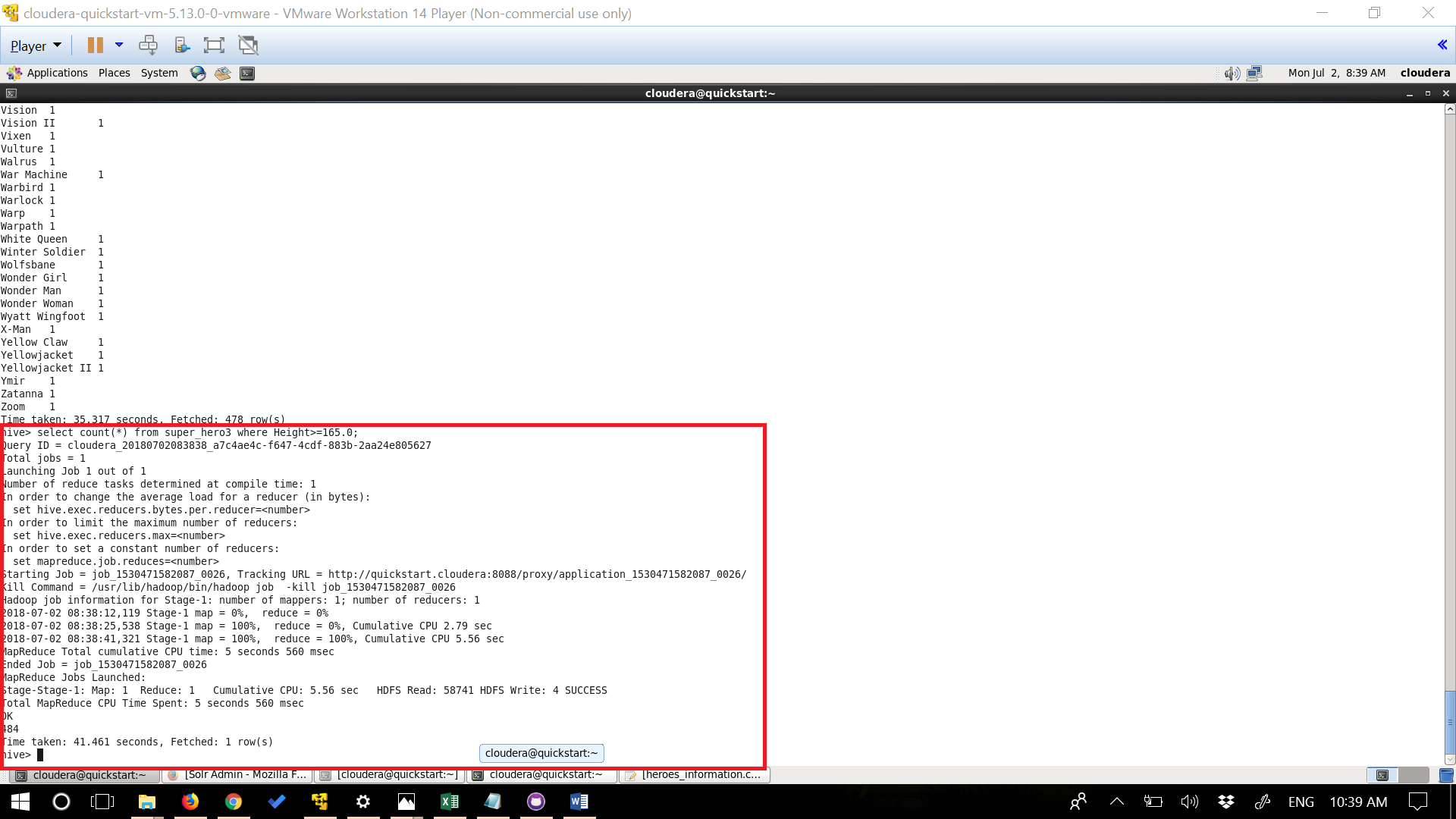
**Query4:**

**Pattern Matching**



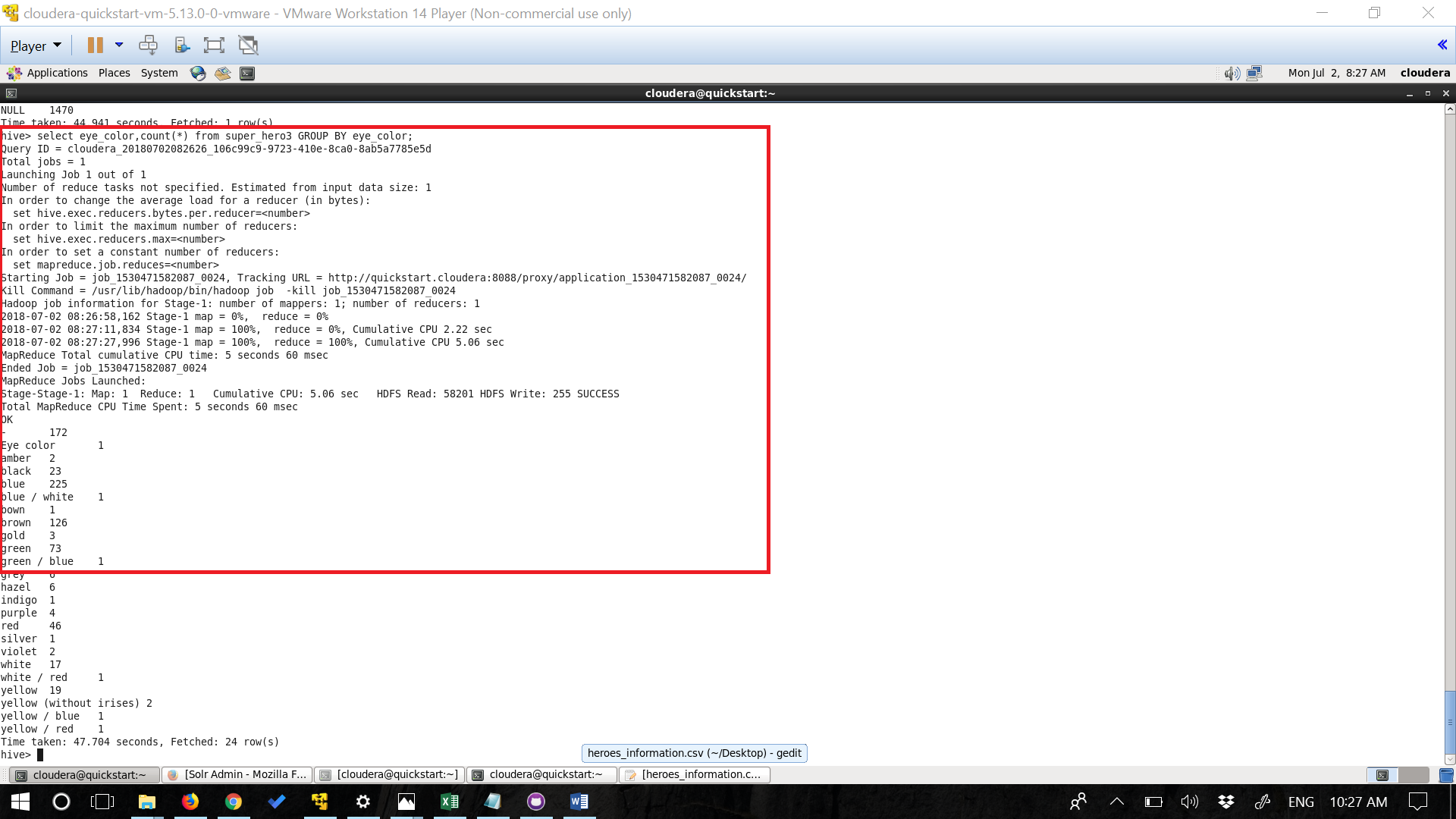
**Query5:**

**Counting whose height is greater than 165**



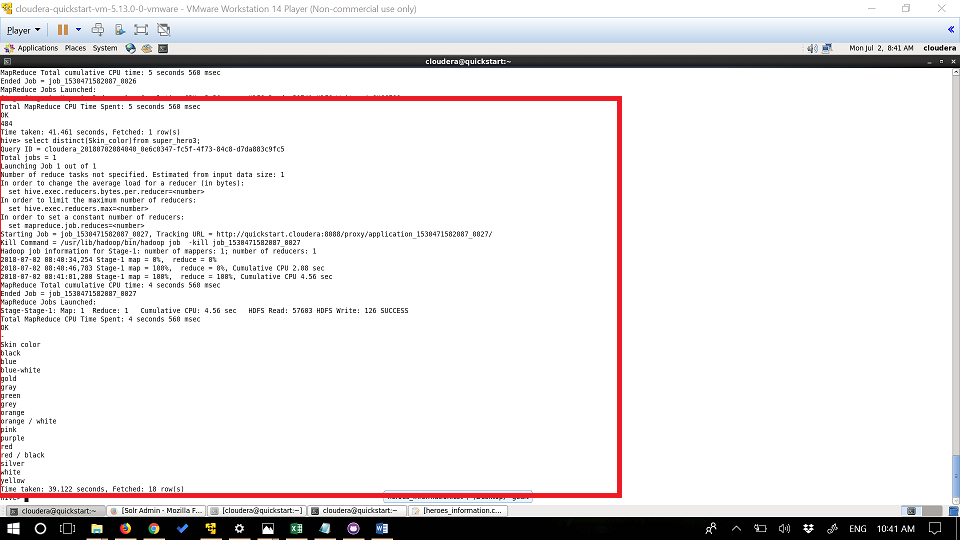
**Query6:**

**Counting different types of eye colors and grouping by eye color**



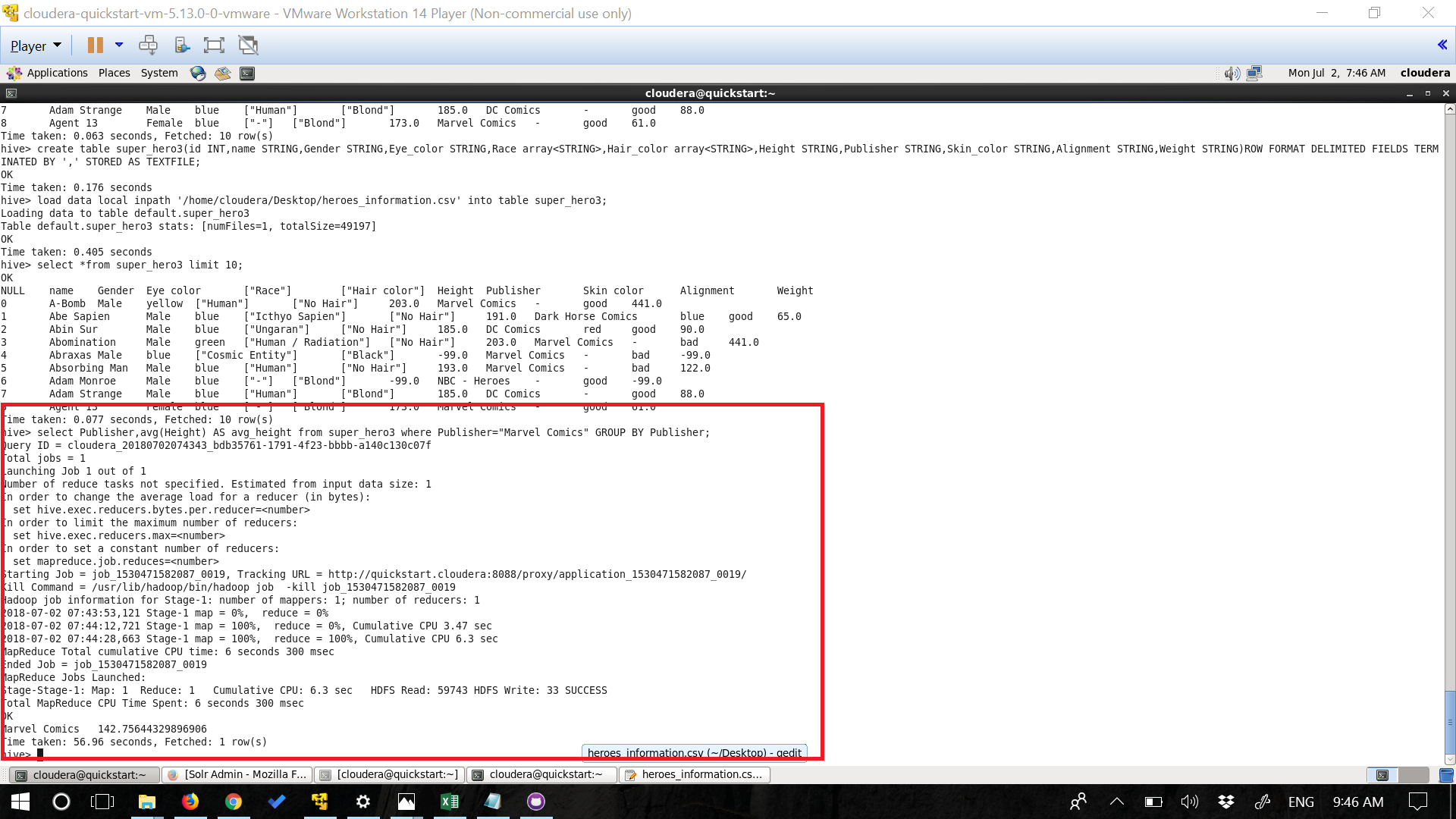
**Query7:**

**To get distinct Skin colors in the dataset**



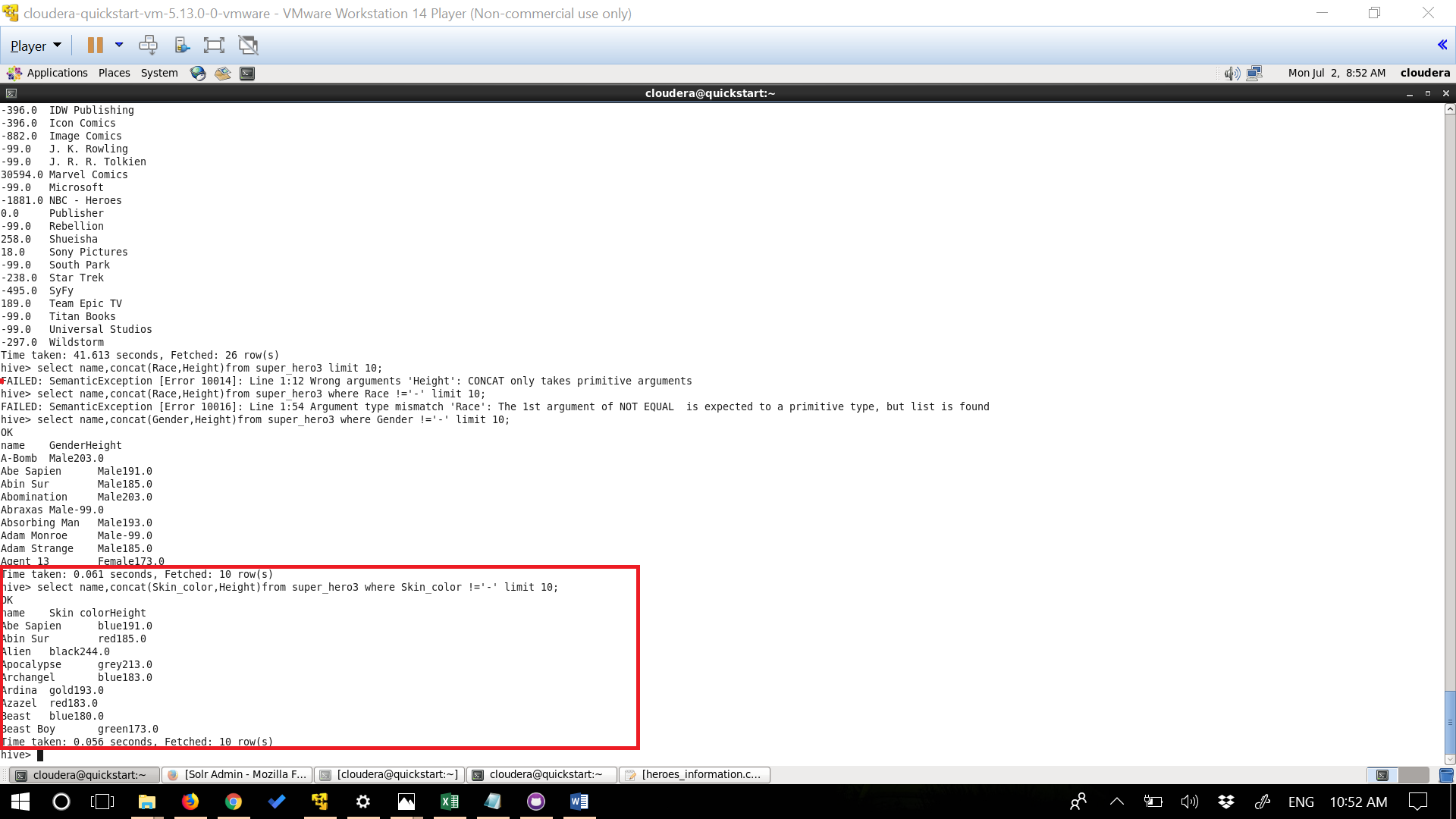
**Query8:**

**Retrieving average Height in the Dataset**



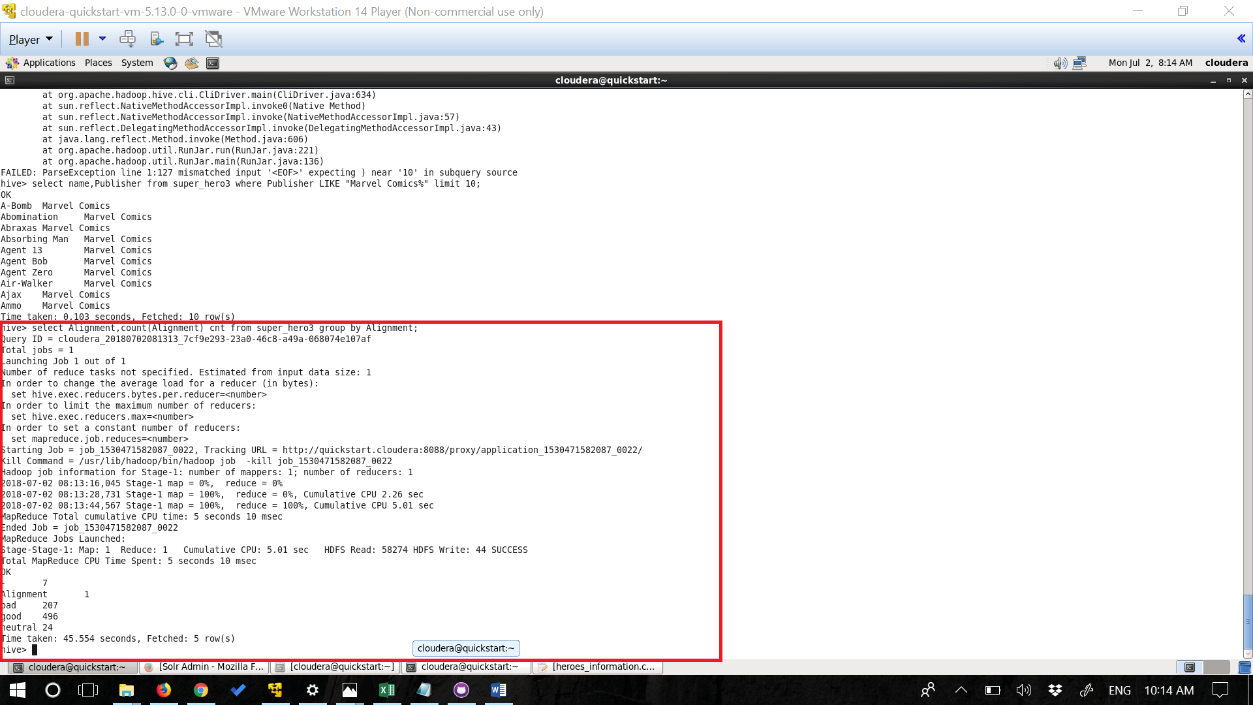
**Query9:**

**To get Name and Concatenating Skin color and height**



**Query 10:**

**Counting Different types of alignment in data set.**



**TASK 2:**

**SOLR USECASE**

a.Create a Solr Collection including our own Field Types

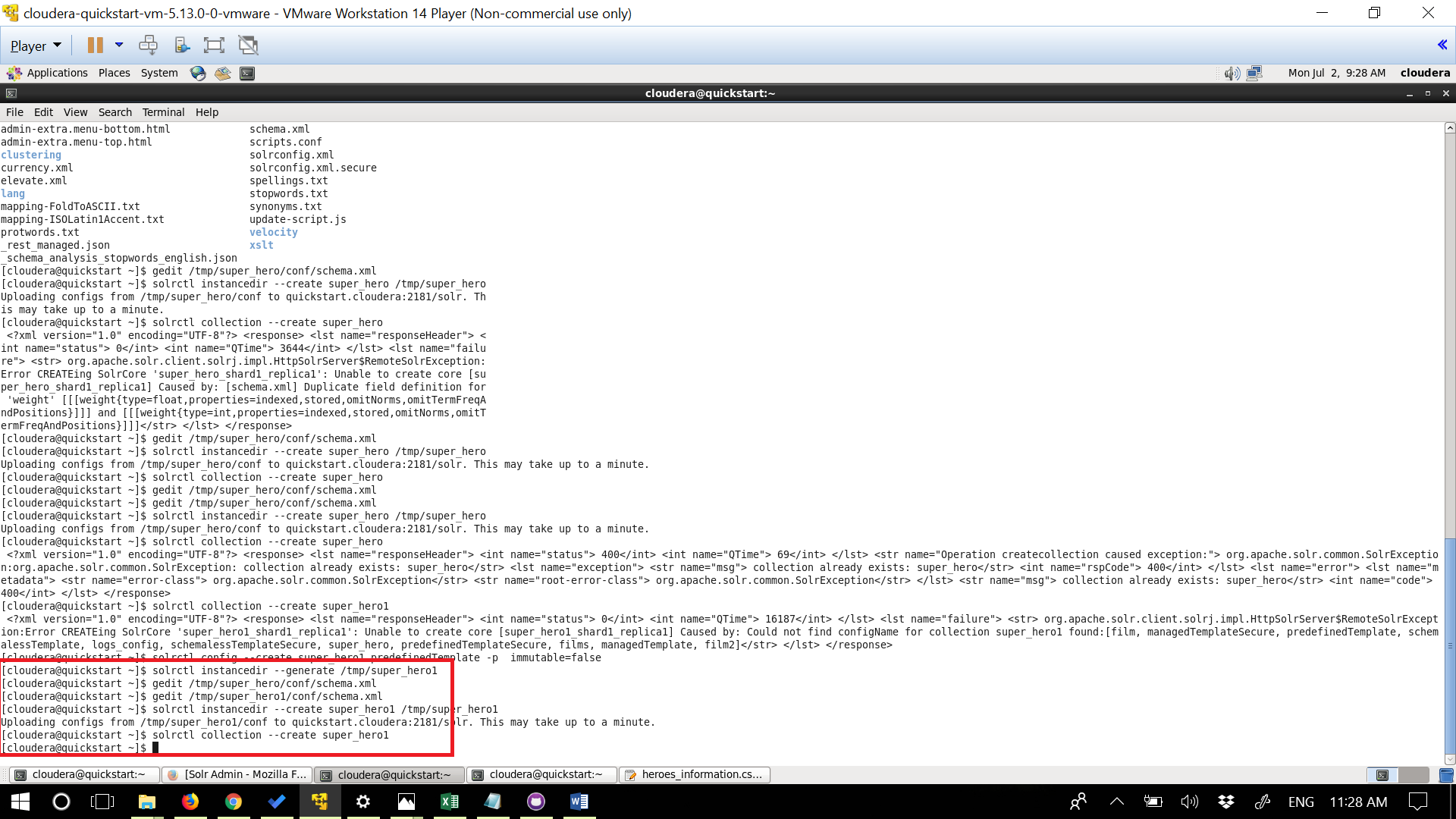
b.Perform 10 intuitive questions in Dataset .

Dataset used: I have used superheroes as my data set

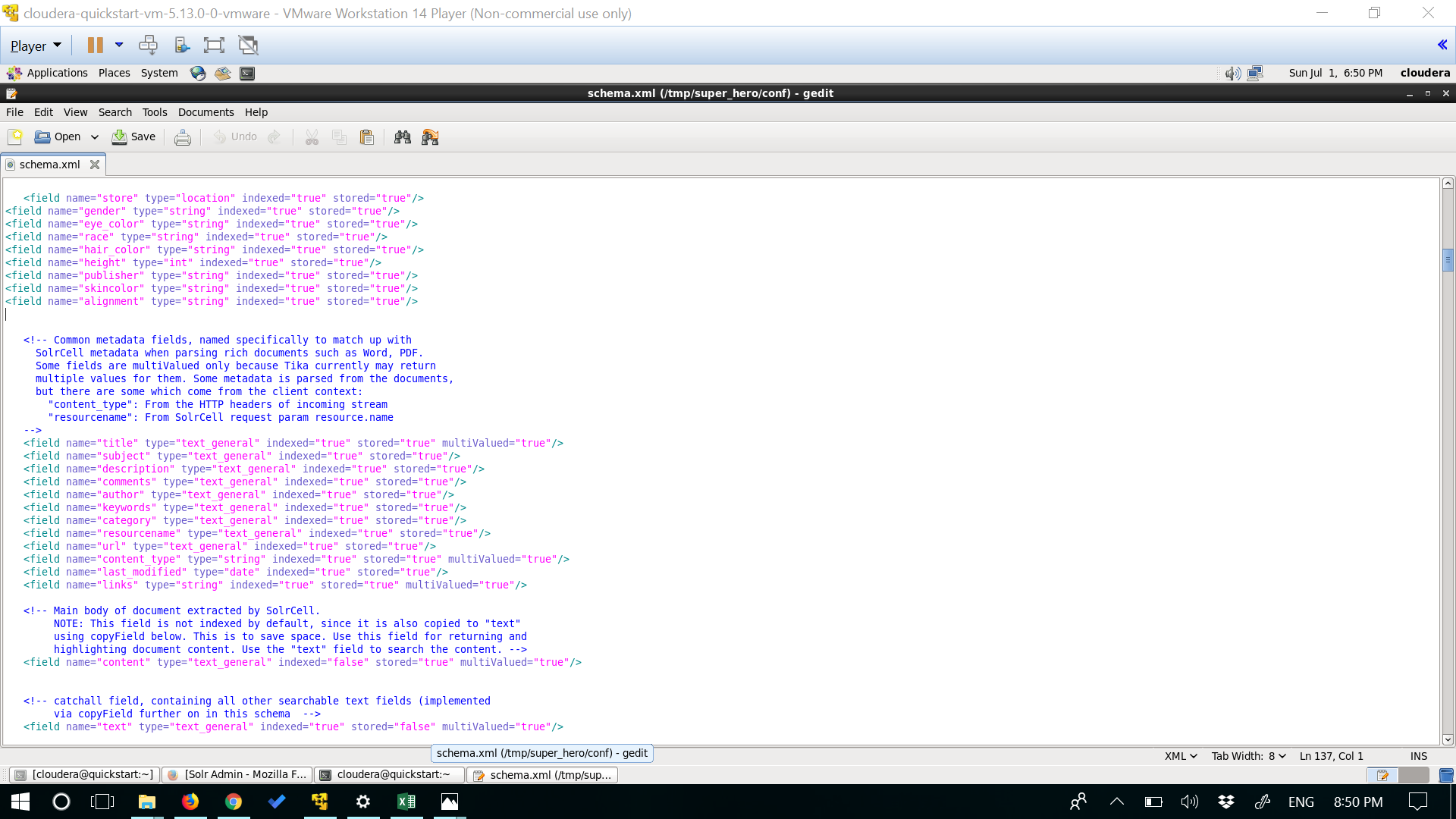
**Queries:**

I have performed a various different type of queries. created an instancedir and Solr collection and edited the schema and included various field types.

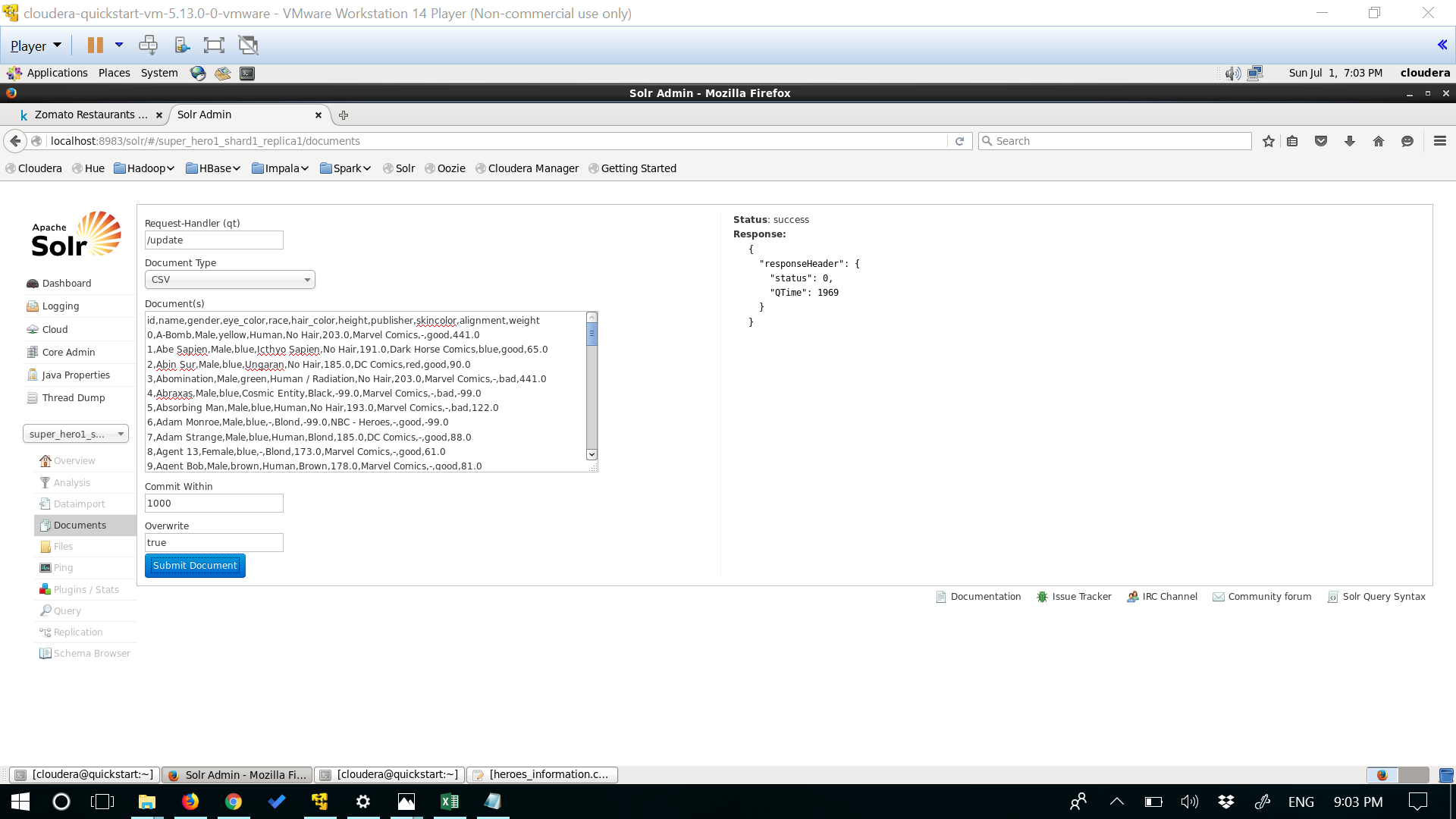
**Generated and cretaed a instace dir**



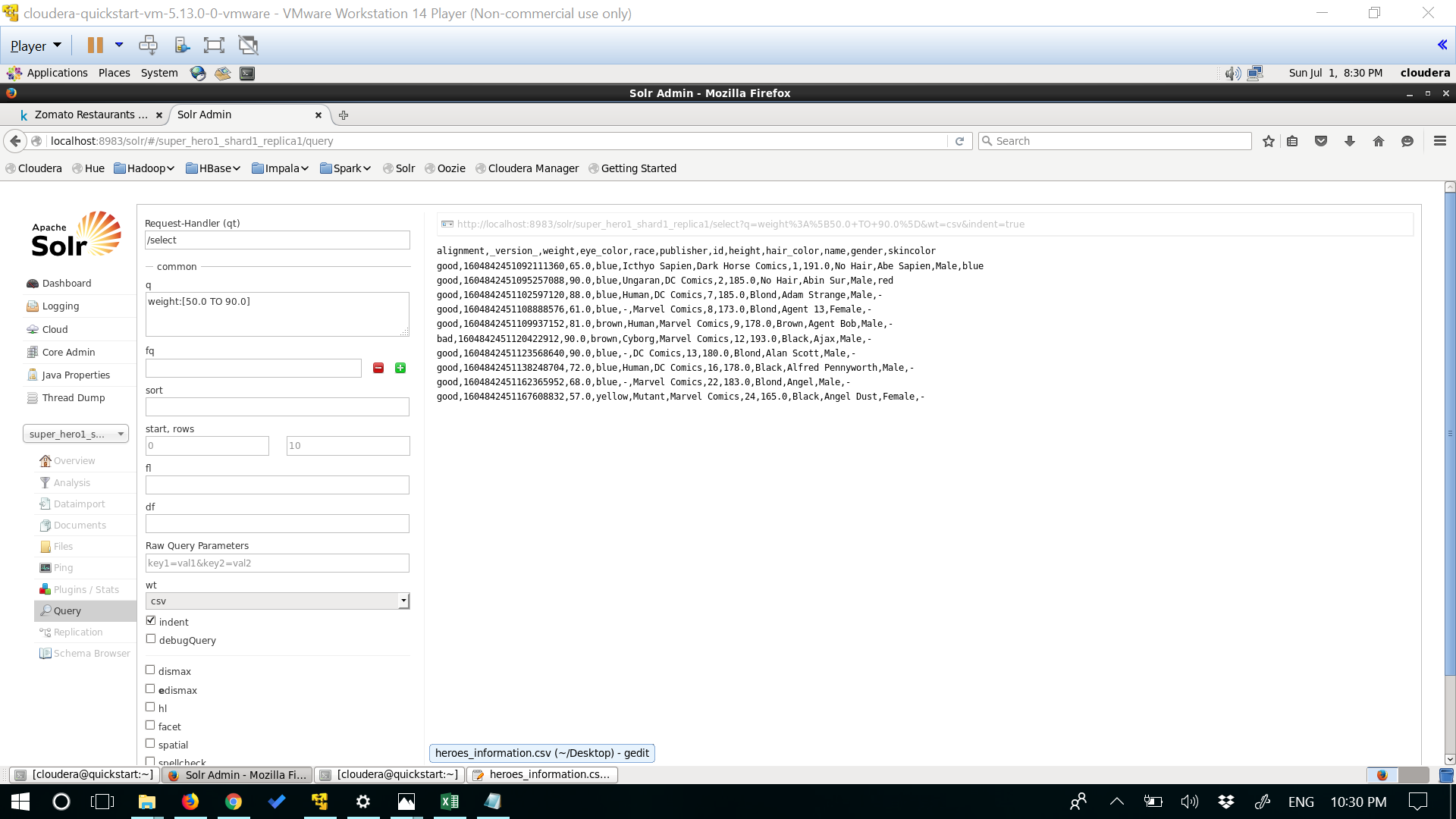
**Edited schema and include field types**



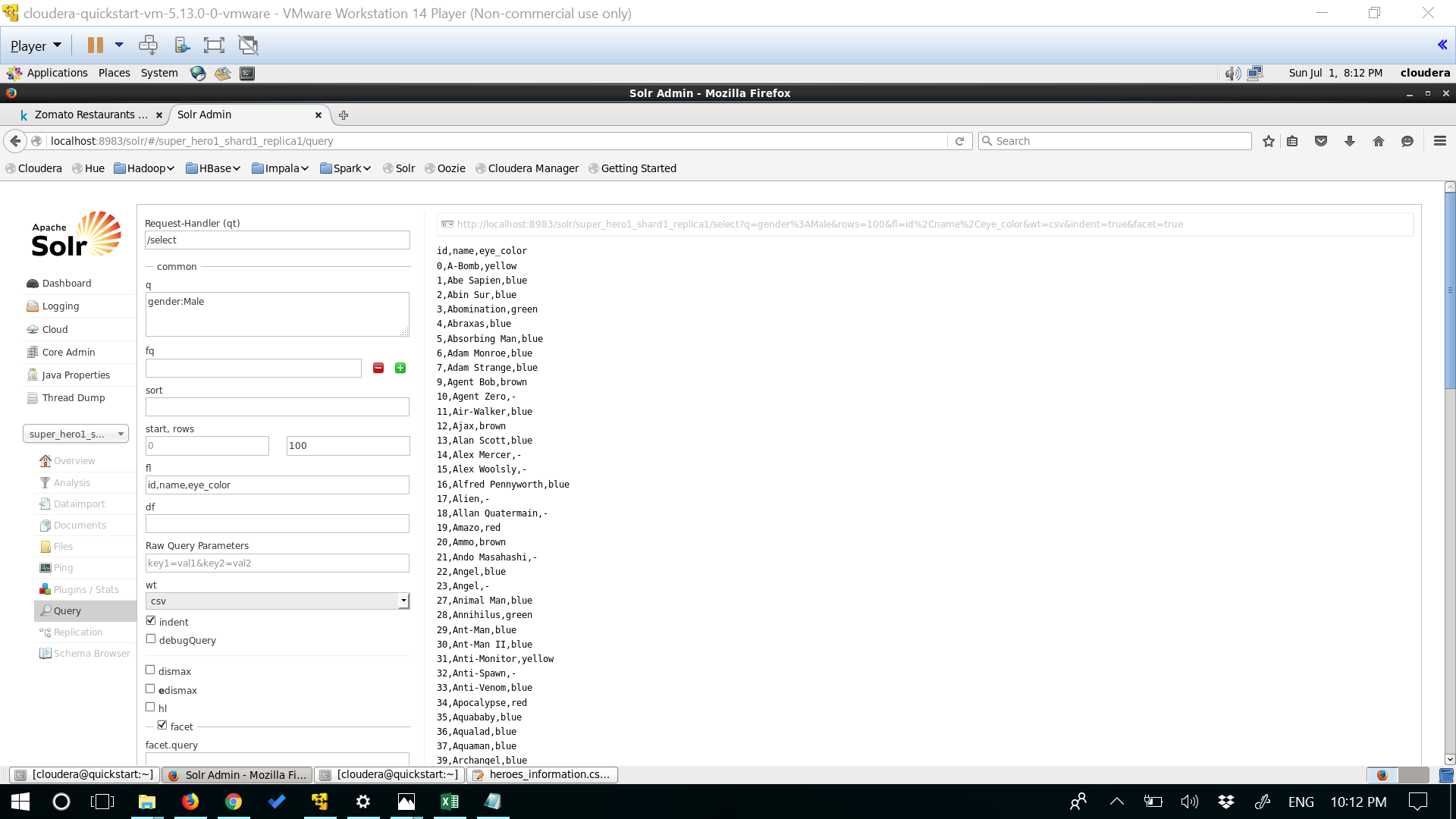
**Uploaded document into solr cloud**



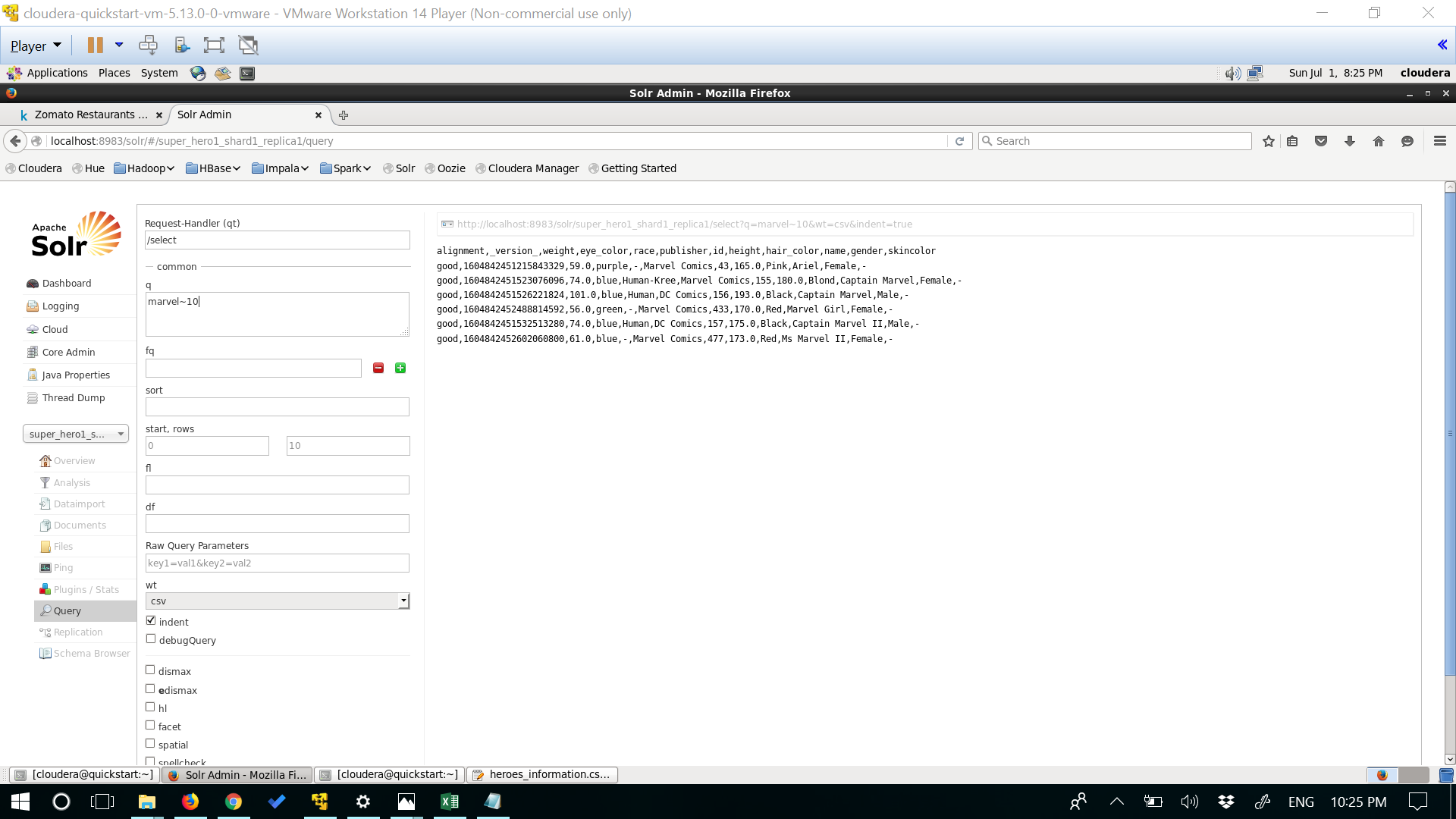
**Query1:**



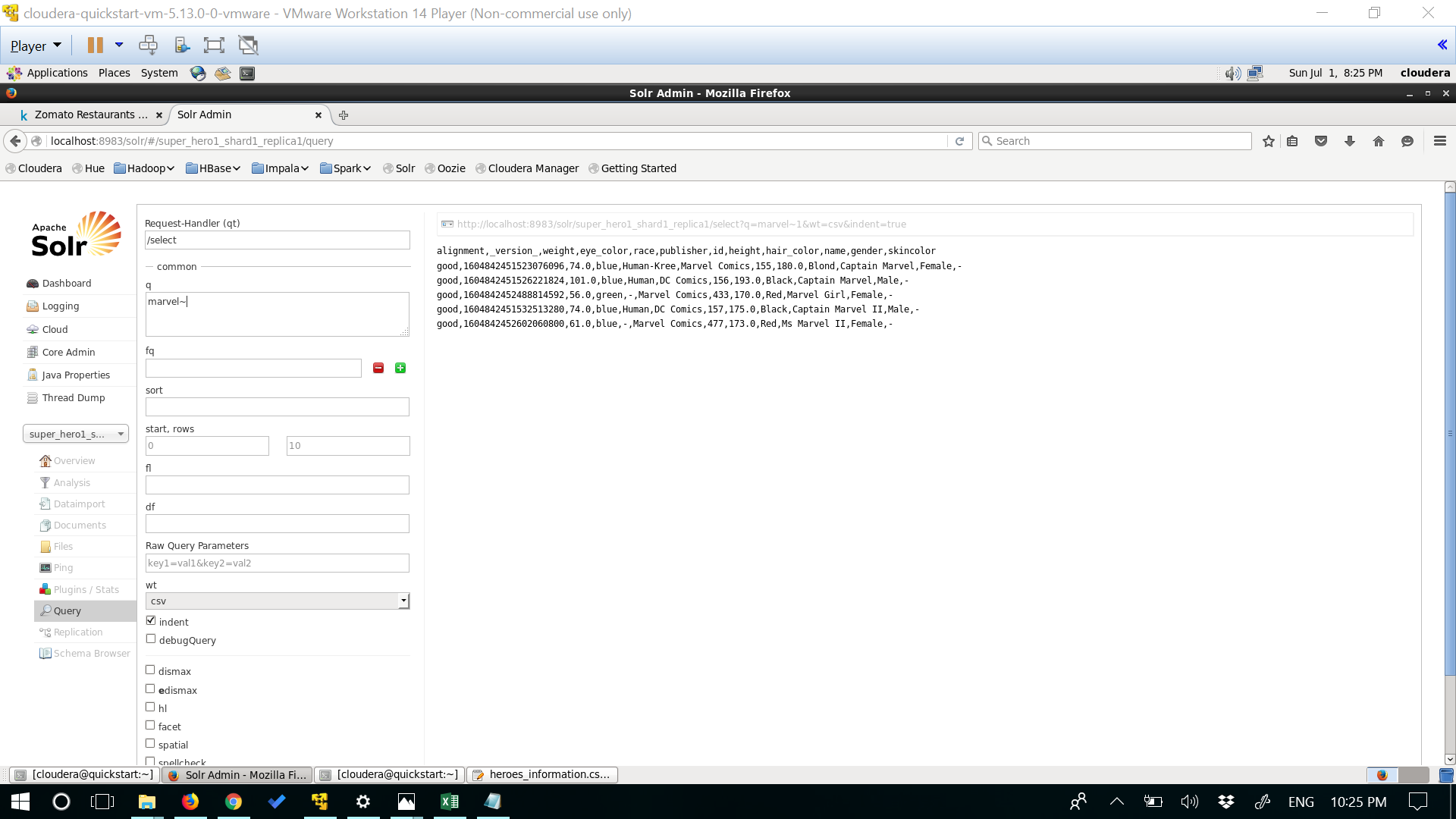
**Query2:**



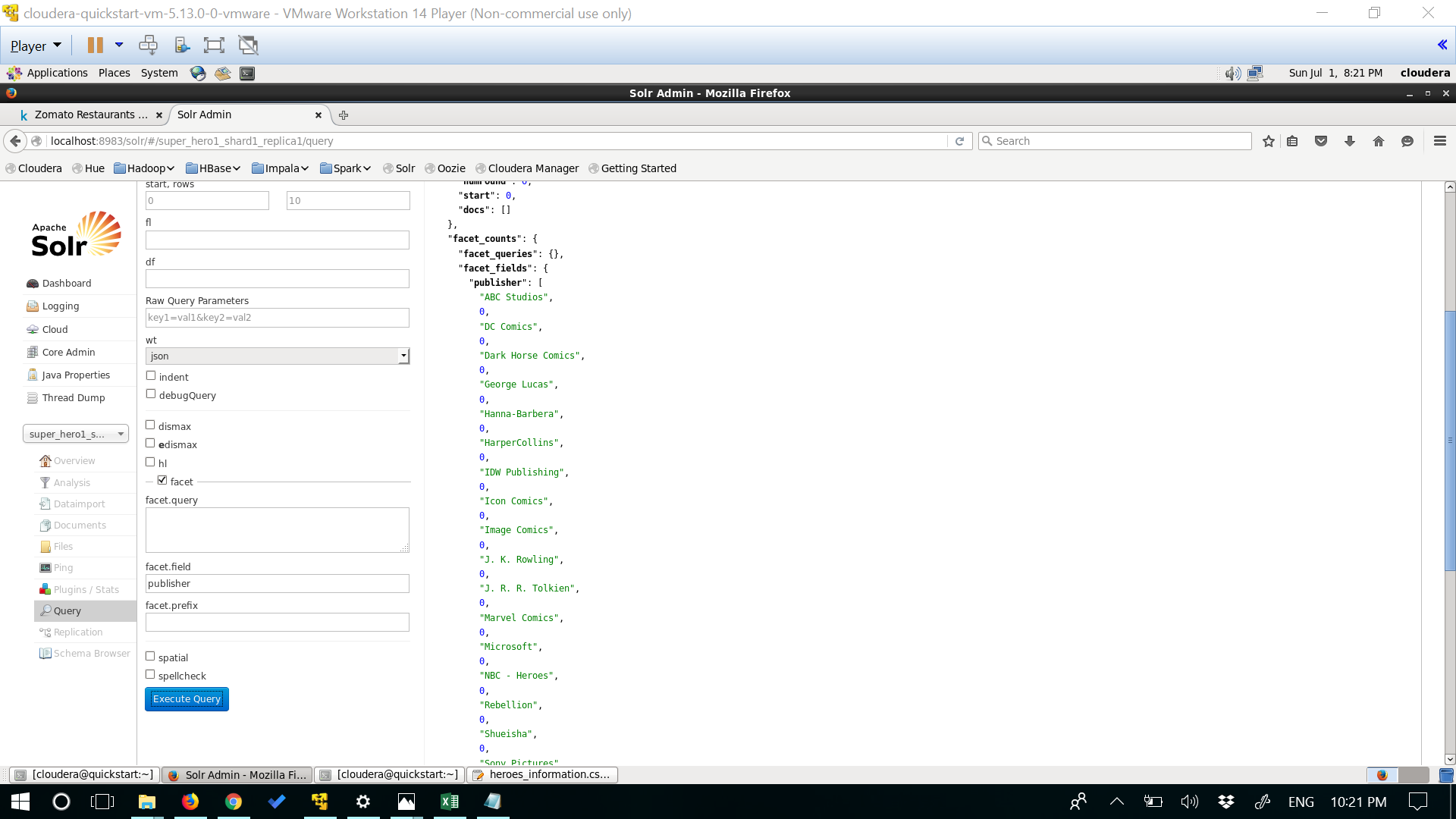
**Query3:**



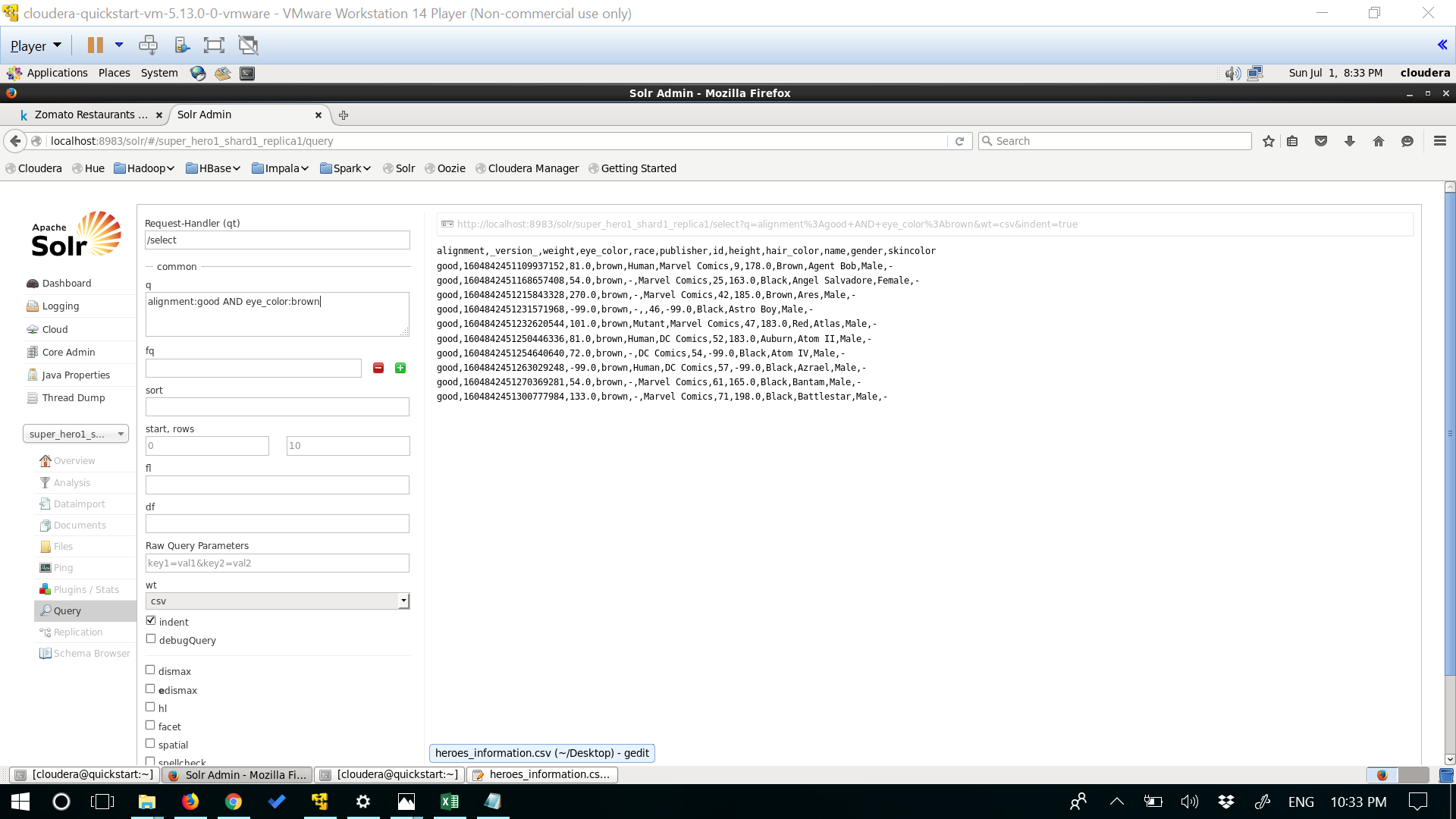
**Query4:**



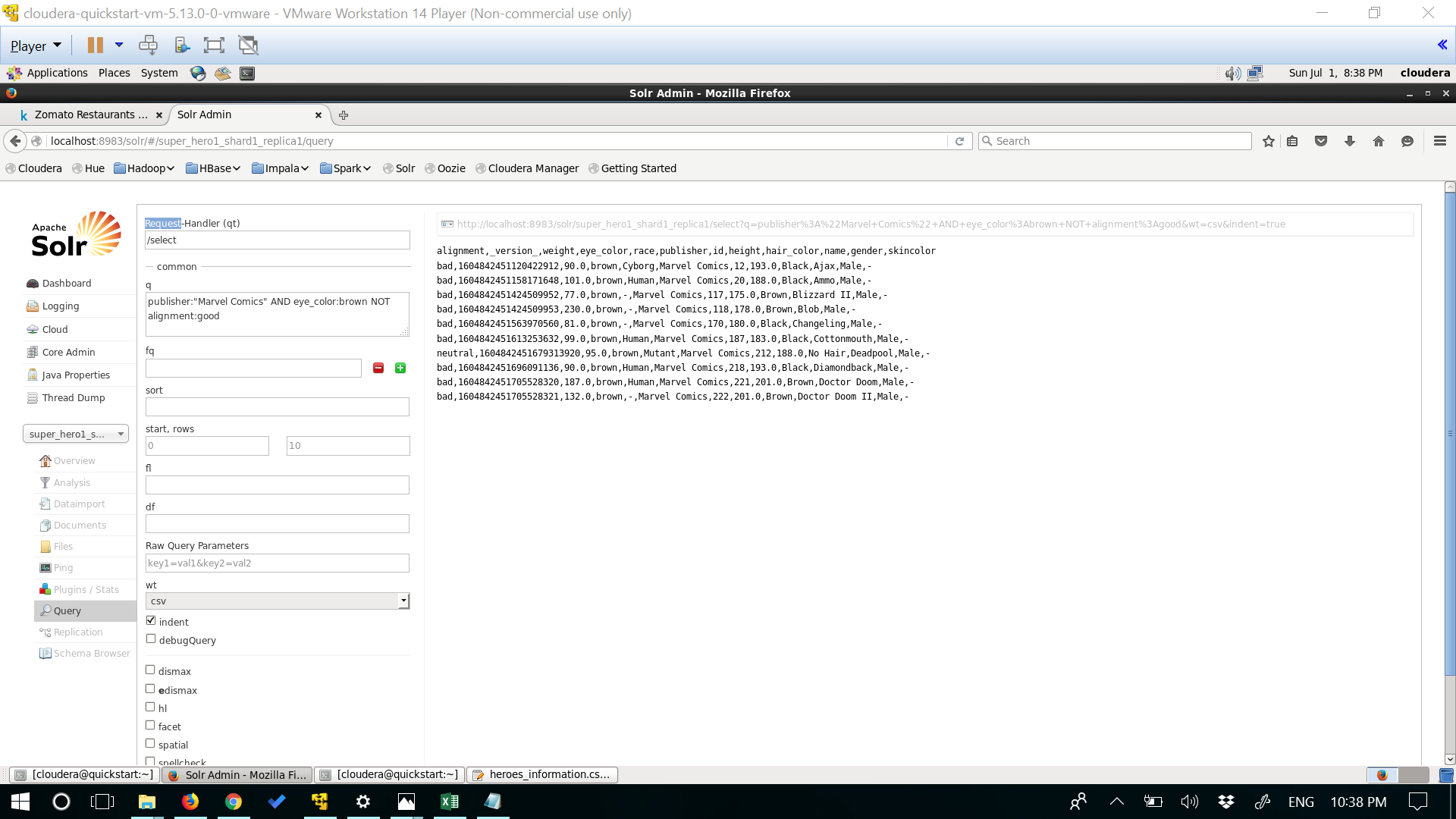
**Query5:**



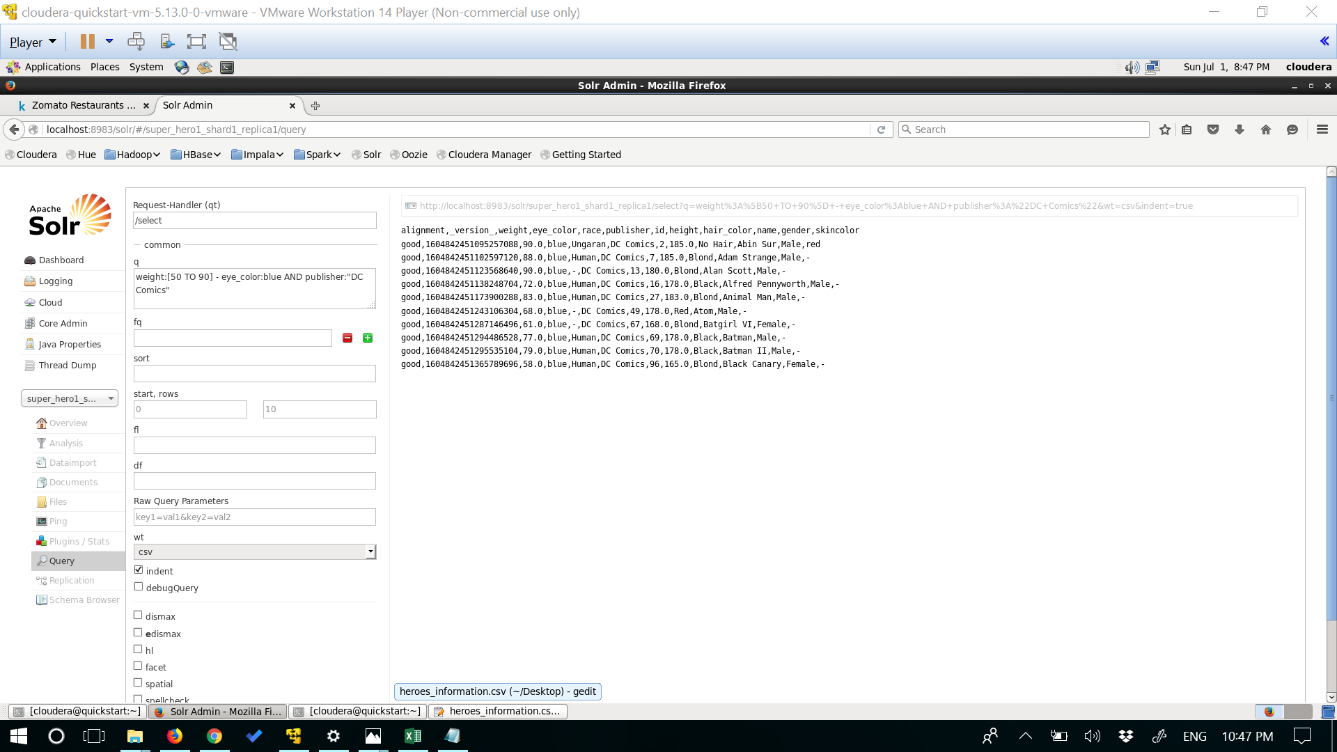
**Nested\_Query1:**



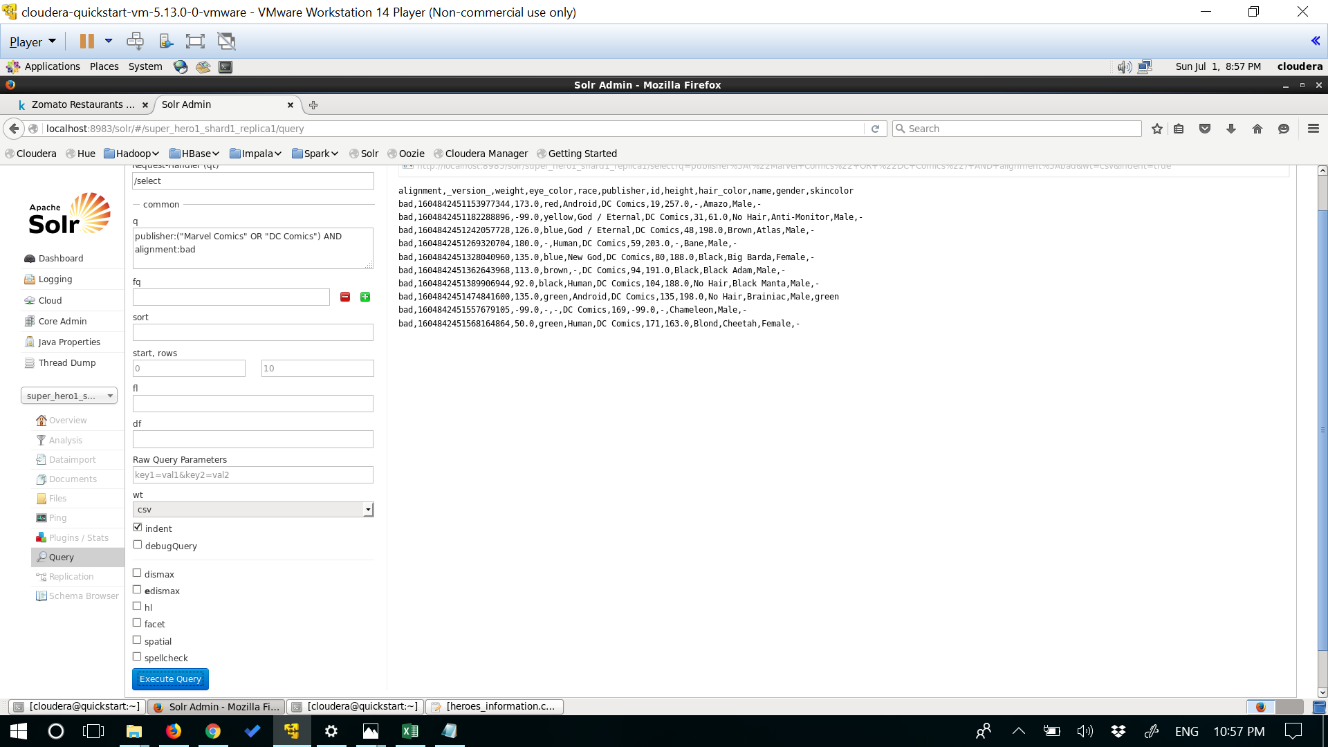
**Nested\_Query2:**



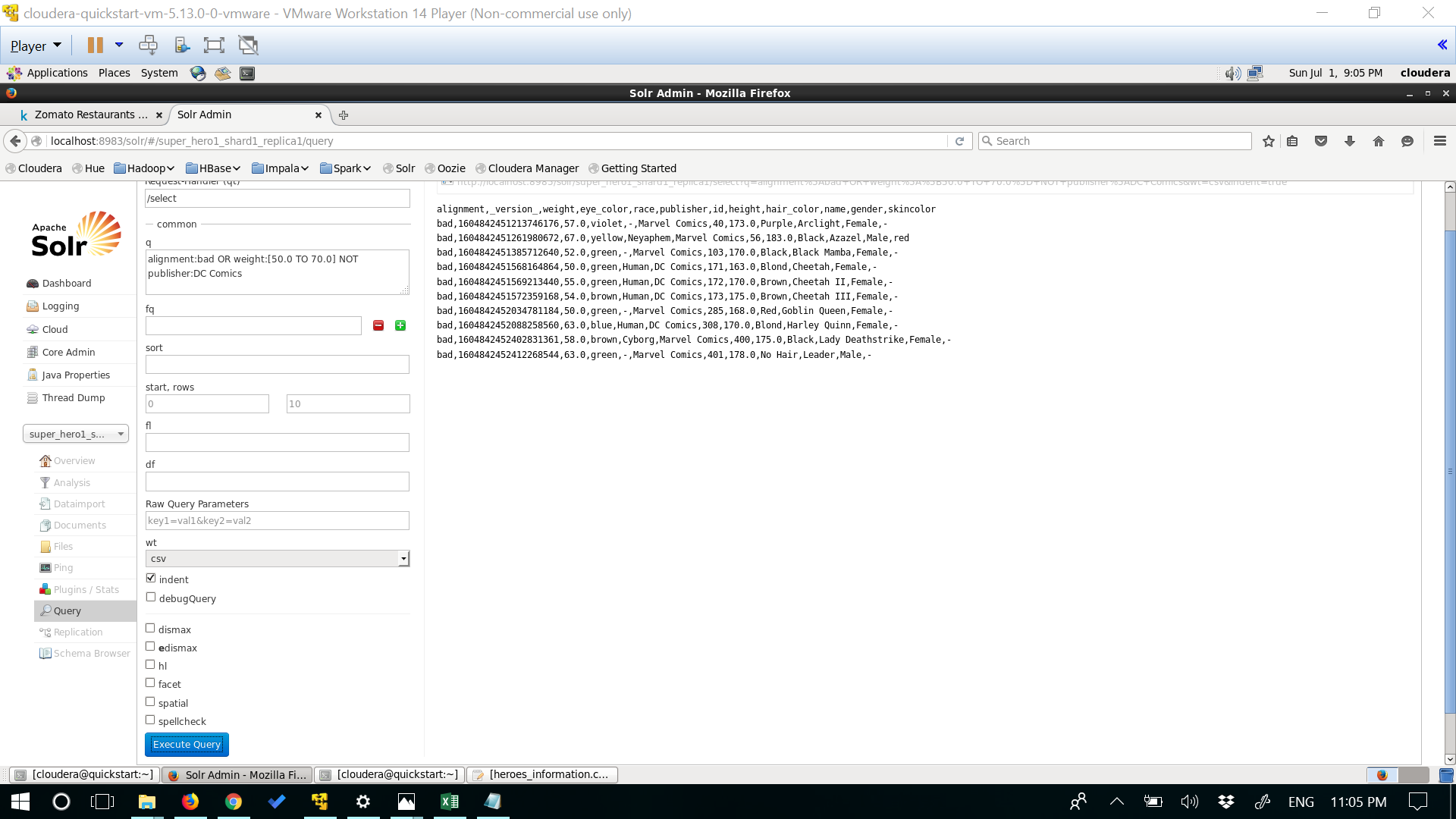
**Nested\_Query3:**



**Nested\_query4:**



**Nested\_query5:**



**# References:**

1.<https://www.tutorialspoint.com/apache_solr/apache_solr_querying_data.htm>

2.<https://community.modeanalytics.com/sql/tutorial/sql-subqueries/>

3.<https://lucene.apache.org/solr/guide/6_6/the-standard-query-parser.html>

4.<https://acadgild.com/blog/working-with-hive-complex-data-types>

5.<https://www.kaggle.com/claudiodavi/superhero-set/data>