**Group-3**

1. We have 4600 article id’s and each article is cited for 100 times.
2. Sort the articles by year and title and then perform the statistical analysis.
3. Under statistical analysis we need to collect authors, citations where it has been published ,intent(background, method and other types) and influential or non-influential.

**Part-2:**

Collect all the abstracts and separate them year by year and perform the analysis on each topic

**Tasks Done:**

1. Collected all the article id’s from the CSV file and appended it to the URL.
2. Read each URL from the CSV file and saved the data into a JSON file.

**Attendance:**

1. Akshitha Maddi
2. Akshatha Kamath
3. Vyshnav Kasa
4. Varun Chandra Dasari

You need to create several tables in the database

1. Frist table: article information

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Article\_id | Venue | Title | Year | Abstract | Doi | fieldsOfStudy | Topics (separate the topic by coma) | Download url (pdf\_url) | Total citation count | Total download | Influential citation count |
|  | SIGIR |  |  |  |  |  |  |  |  |  |  |
|  | SIGIR |  |  |  |  |  |  |  |  |  |  |
|  | SIGIR |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

1. Author information

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Article\_id | Author\_id | Author name | Author\_rank | Author\_url | Author\_institution | Author publication count (how many publications has the author published) |
|  |  |  | First |  |  |  |
|  |  |  | Second |  |  |  |
|  |  |  | Third author |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

1. Citation information

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Article\_id | Citing\_paper\_id | Citing\_paper\_doi | Citing\_paper\_title | Citing\_paper\_title | Citing\_paper\_venue | Citing\_paper\_year | Intent  (every intend one row) | Is influential |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

1. Reference information

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Article\_id | Reference\_paper\_id | Reference \_paper\_doi | Reference \_paper\_title | Reference \_paper\_title | Reference \_paper\_venue | Reference \_paper\_year | Intent  (every intend one row) | Is influential |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |