

**COLLECTING TWEETS USING TWITTER STREAMING       API’S**

**Sub: Principles of Big Data Management**

**Phase 2 of Project**

|  |  |
| --- | --- |
| Sirisha Rella | (16268829) |
| Vineeth Dudipalli | (16274444) |
| Rajeshwari Cholleti | (16272108) |

 Instructor

Dr. PRAVEEN RAO, Ph.D.

**Abstract:**

Phase 2 of this project deals with the following requirements:

1. Writing interesting analytical queries on twitter data that we collected.

2. Developing interesting visualizations.

**Title:**

Technology in different Domains.

**Technologies & Tools Used:**

1. Spark

2. Python

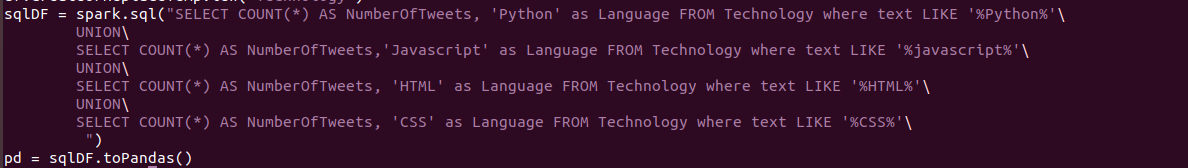
3. Tableau

**Queries and Analysis:**

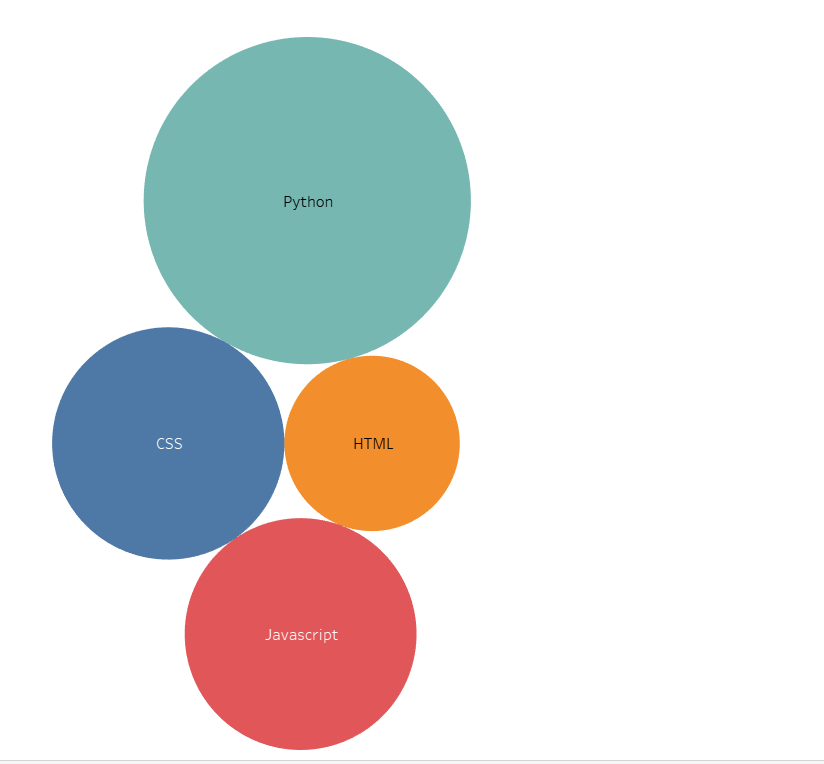
**Query-1:**

In this query we found the number of users tweeting about Python, JavaScript, HTML and CSS.

**Code:**



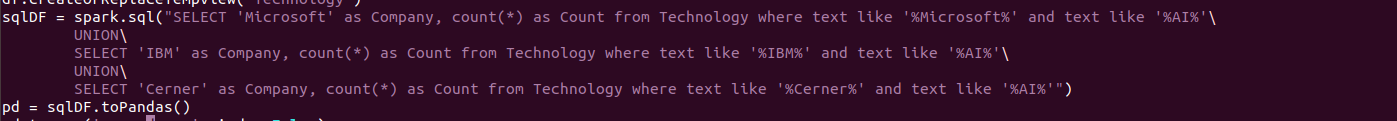
**Visualization:**



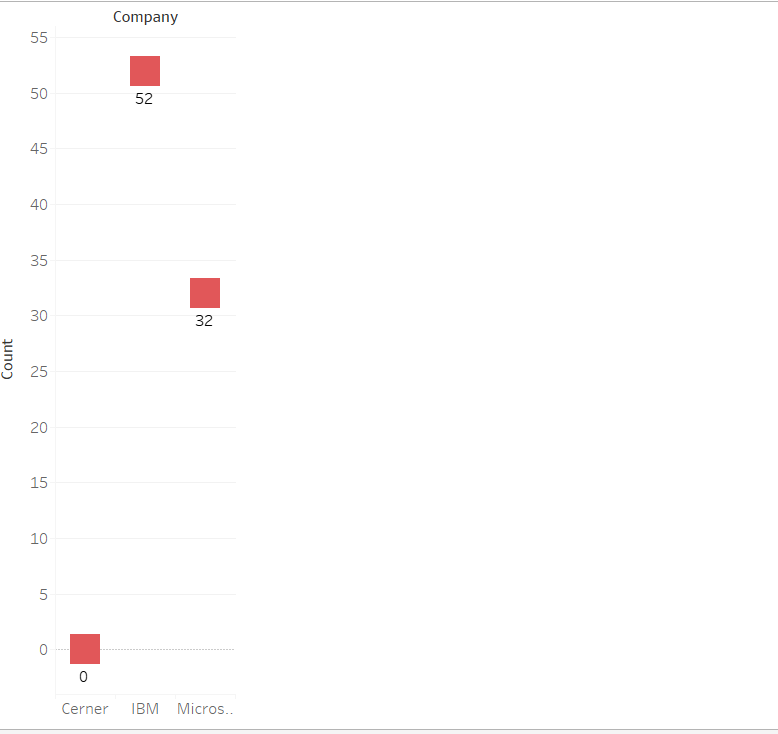
**Query-2:**

In this query we found how Microsoft, IBM and Cerner associates/employees are actively tweeting about AI.

**Code:**



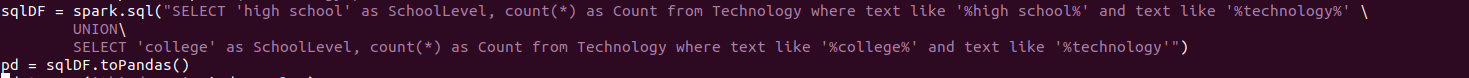
**Visualization:**



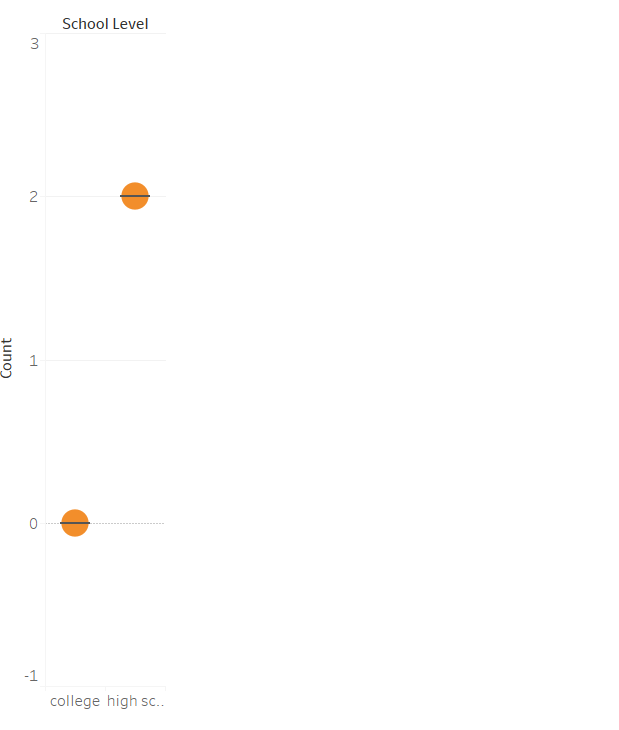
**Query-3:**

In this query we found count of active high school/college students’ achievements/participation in technology.

**Code:**



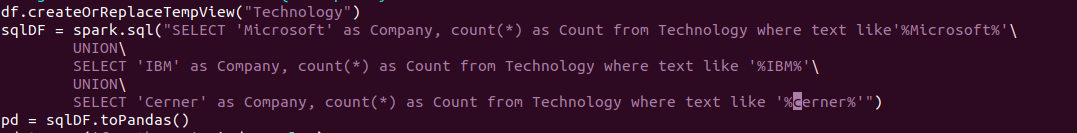
**Visualization:**



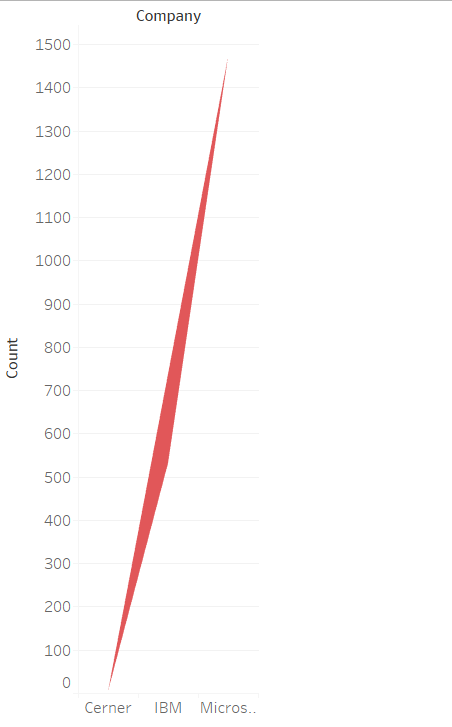
**Query-4:**

In this query we found the Count of people who are actively tweeting about Microsoft, Cerner and IBM

**Code:**



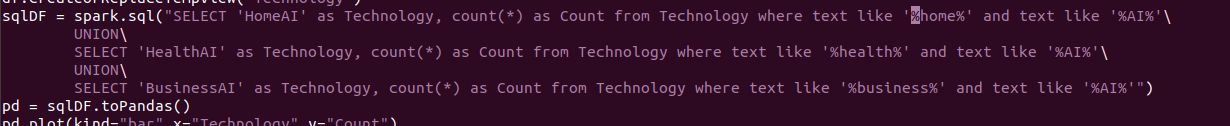
**Visualization:**



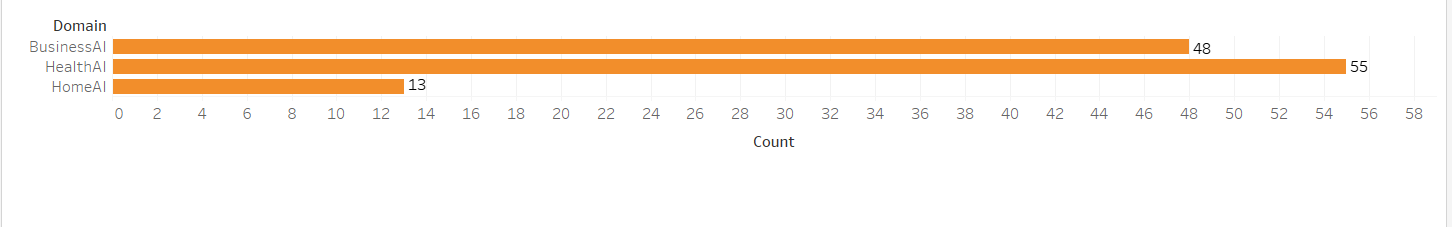
**Query-5:**

In this query we found tweets about how AI is impacting in health, home and business AI.

**Code:**



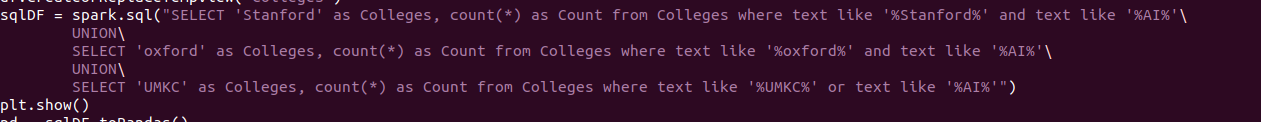
**Visualization:**



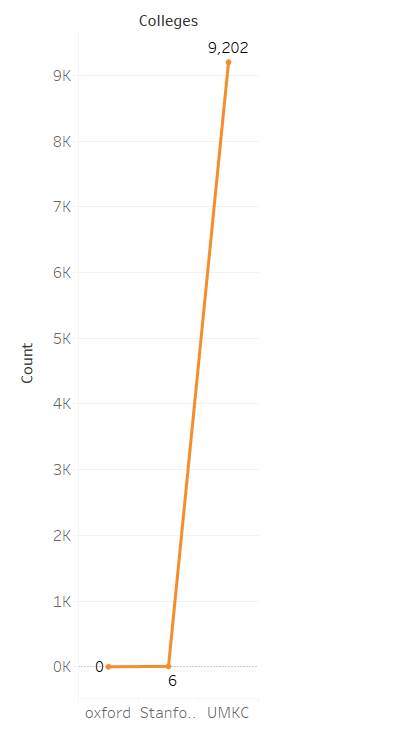
**Query-6:**

In this query we found which students who are actively participating/tweeting about python programming based on college name.

**Code:**



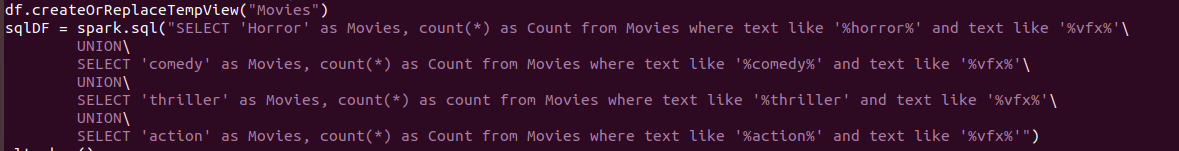
**Visualization:**



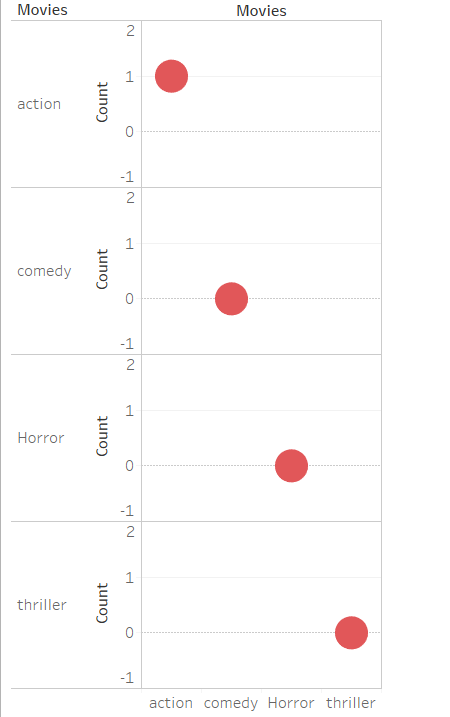
**Query-7:**

In this query we found tweets that talk about different kind of films that are using vfx effects.

**Code:**



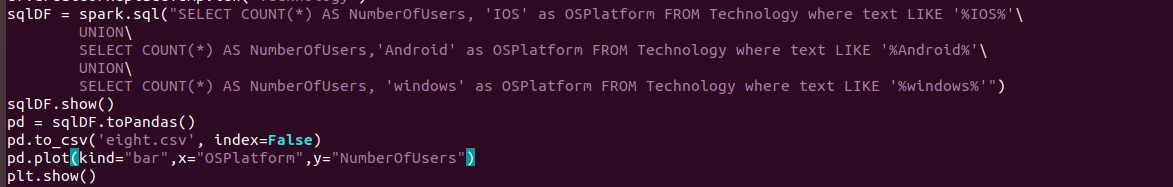
**Visualization:**



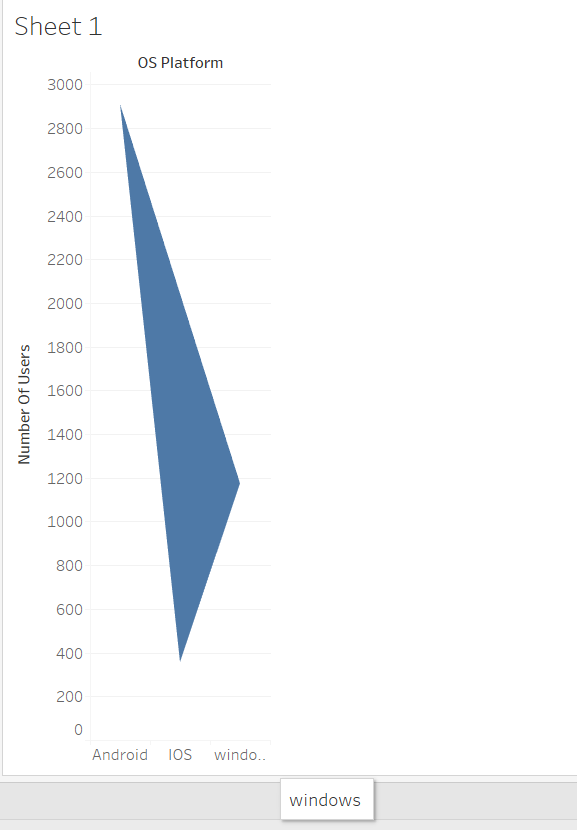
**Query-8:**

In this query we found number of people tweeting about IOS, Android and Windows operating systems.

**Code:**



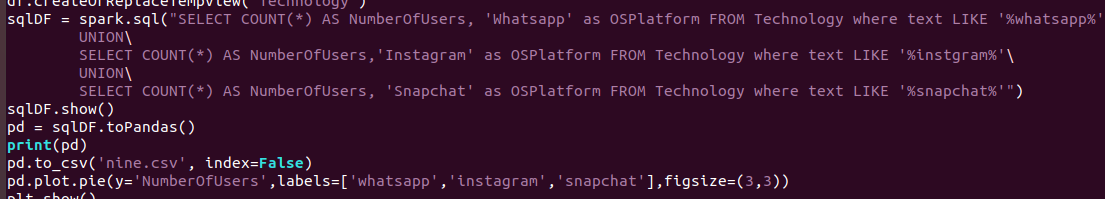
**Visualization:**



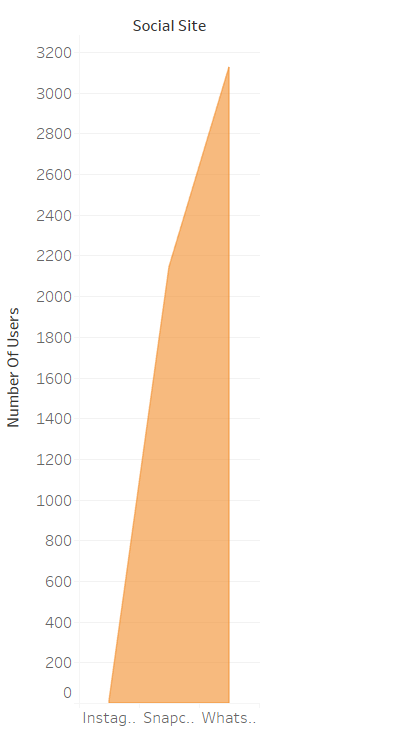
**Query-9:**

In this query we found number of people tweeting about WhatsApp, Instagram and snapchat

**Code:**



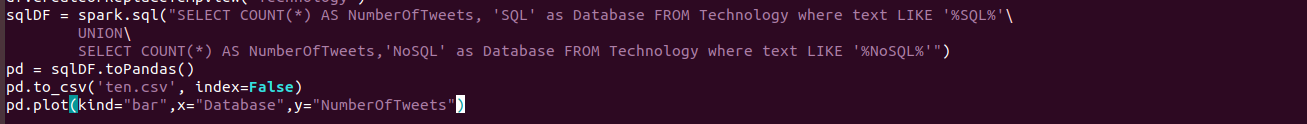
**Visualization:**



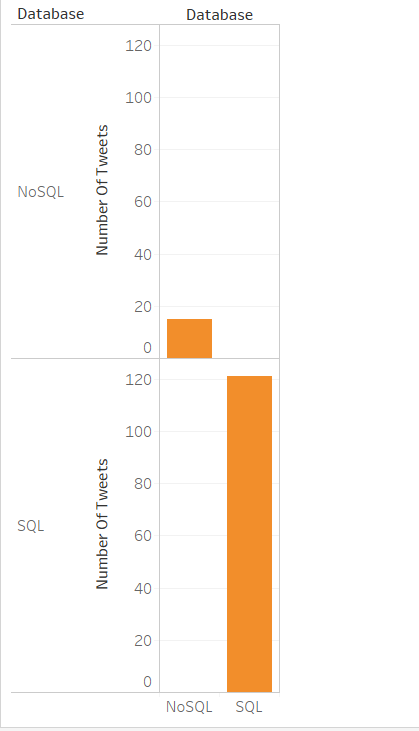
**Query-10:**

In this query we found number of people tweeting about SQL and NoSQL

**Code:**



**Visualization:**



**Code Link:**

<https://drive.google.com/drive/folders/1ByXhNlRRIBotiP601KsJhz63dEoO6Tmx?usp=sharing>