

Twitter Sentiment Analysis

(Polarity Identification)

Nick Oliver

Biljana Jovanova

Ashok Kumar Haralekere Motaiah

Mounica Urity

Murat Ozan Aydin

Jacob Reyes

Sercan Demir

Rajesh Chodavarapu

# Project Objective

The main purpose of this project is to analyze Twitter “tweets” on a specific topic and see the user overall opinion by calculating the polarity of the tweets. For this project we used Python to get data from Twitter’s API, scrape a website for news article content, analyze polarity with NLTK, and store data in an Azure SQL DB. We then created a web front end dashboard that would dynamically show the polarity of data we are gathering.

The two topics we selected to analyze are Cryptocurrency and Burger Store chains.

# Software Requirements

1. Gather the live twitter tweets for the specific topics.

2. Do the text analysis to find the polarity of the tweets.

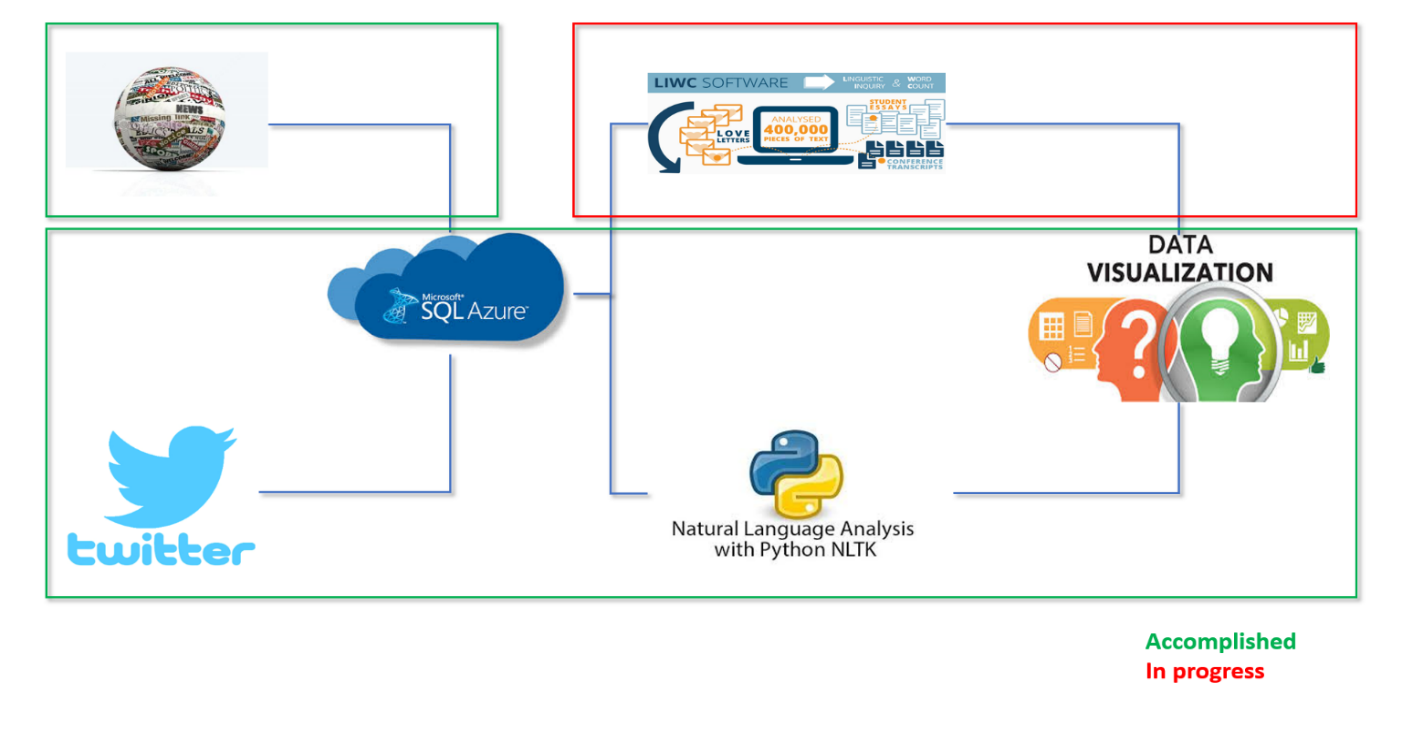
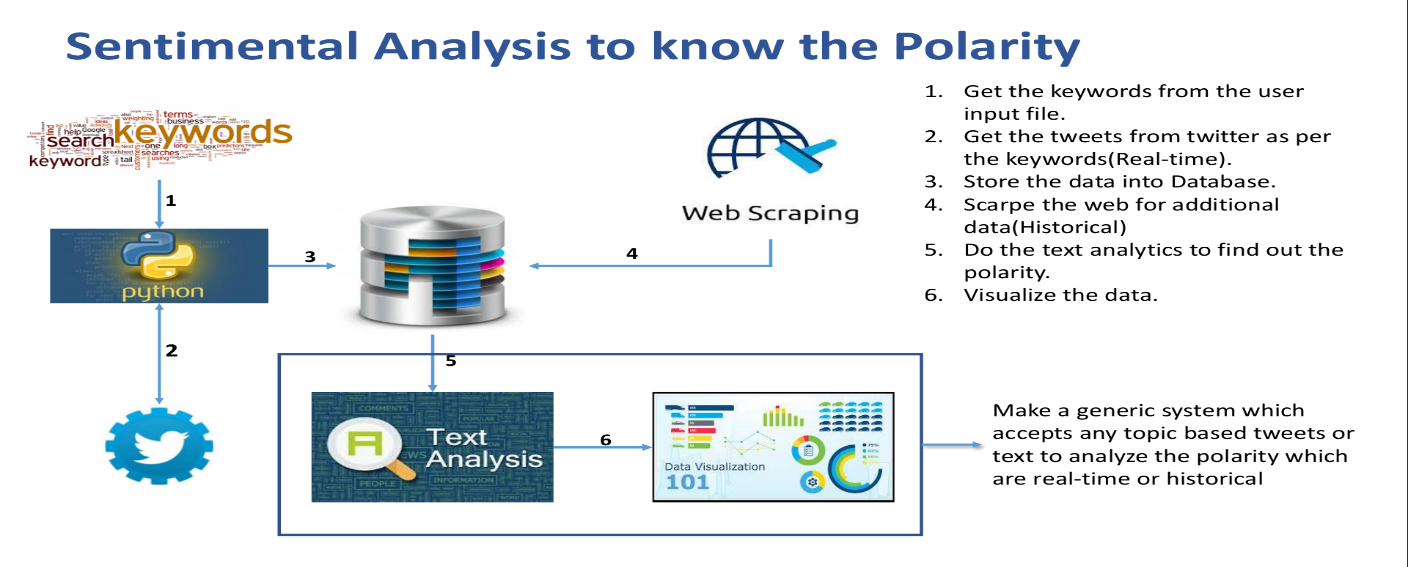
3. Store the data and results into the Database.

4. Scrape any website to get the data about the topic selected.

4. Do the text analysis and identify the meaningful results.

# High Level Design Diagram

***Figure 1***



***Figure 2***

# Technical Specifications

|  |  |
| --- | --- |
| Programming Language | Python |
| Database | MSSQL (Microsoft Azure) |
| Data Visualization | PHP, D3JS |
| Operating System | Linux, Windows |
| Tools Used | PyCharm, Spyder, Microsoft SQL Client, Dbeaver SQL Client |

**Table 1**

# Installation

**Python:**

<https://wiki.python.org/moin/BeginnersGuide/Download>

How to configure Microsoft Azure Database

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-get-started-portal>

|  |  |
| --- | --- |
| **Server Name** | scrapersmsda2017.database.windows.net |
| **Database Name** | msdatwitter |
| **User ID** | scrapersmsda |
| **Password** | msdascrapers@2017 |
| **Port Number** | 1433 |

**How to install PHP:**

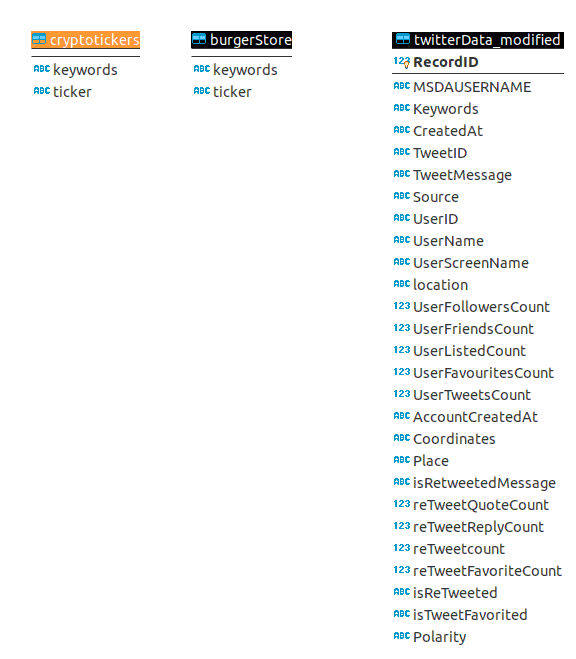
<http://php.net/manual/en/install.php>

you need to install PECL extensions

# Database Design

|  |  |
| --- | --- |
| 1. **Table Name** | 1. **Description** |
| 1. TwitterData\_Modified | 1. Used to store the twitter data collected |
| 1. Cryptotickers | 1. Contains the Topic ticker and keywords used to get the tweets |
| 1. BurgerStore | 1. Contains the burger company short form and the keywords used to get the tweets |

**Table 2**



***Figure 3***

# Code Explanation with Comments

AzureDBConnection\_Final.py - The above file has utilities to connect to the database and insert the data to twitterData\_modified table.



Streaming Twitter Tweets

TwitterStreaming\_Final.py - This file is used to get the tweets from Twitter Streaming based on the keywords specified.



Web Scraping

The following files are used to scrape the coindesk website to get the article content based on the search words.

CoinDeskScraping.py – This file is used to get the meta data of the articles searched by keywords in coindesk website.



getEachContent\_Final.py -This file is used to get the content of each article provided by the article link.



TwitterScraping\_Final.py

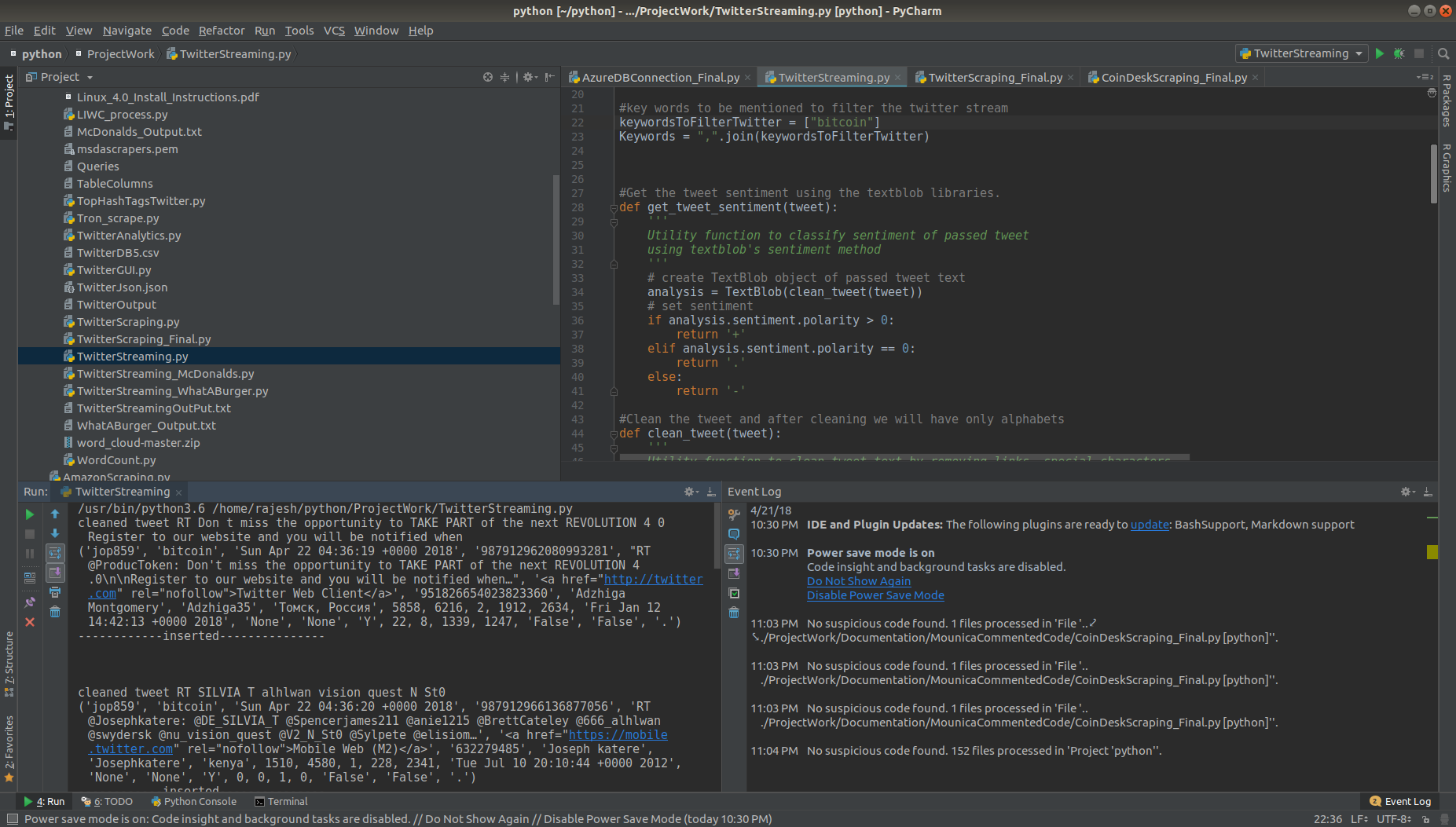
This is the supporting file to get tweet content if any of the tweets were truncated by the Twitter API.



# How to run:

TwitterStreaming\_Final.py → Insert into Database.

TwitterScraping\_Final.py → Update the Database if any tweet is truncated.



***Figure 4***

**Get the article details**

CoinDeskScraping.py → article\_df.csv → getEachContent\_Final.py → Text Files → Folder

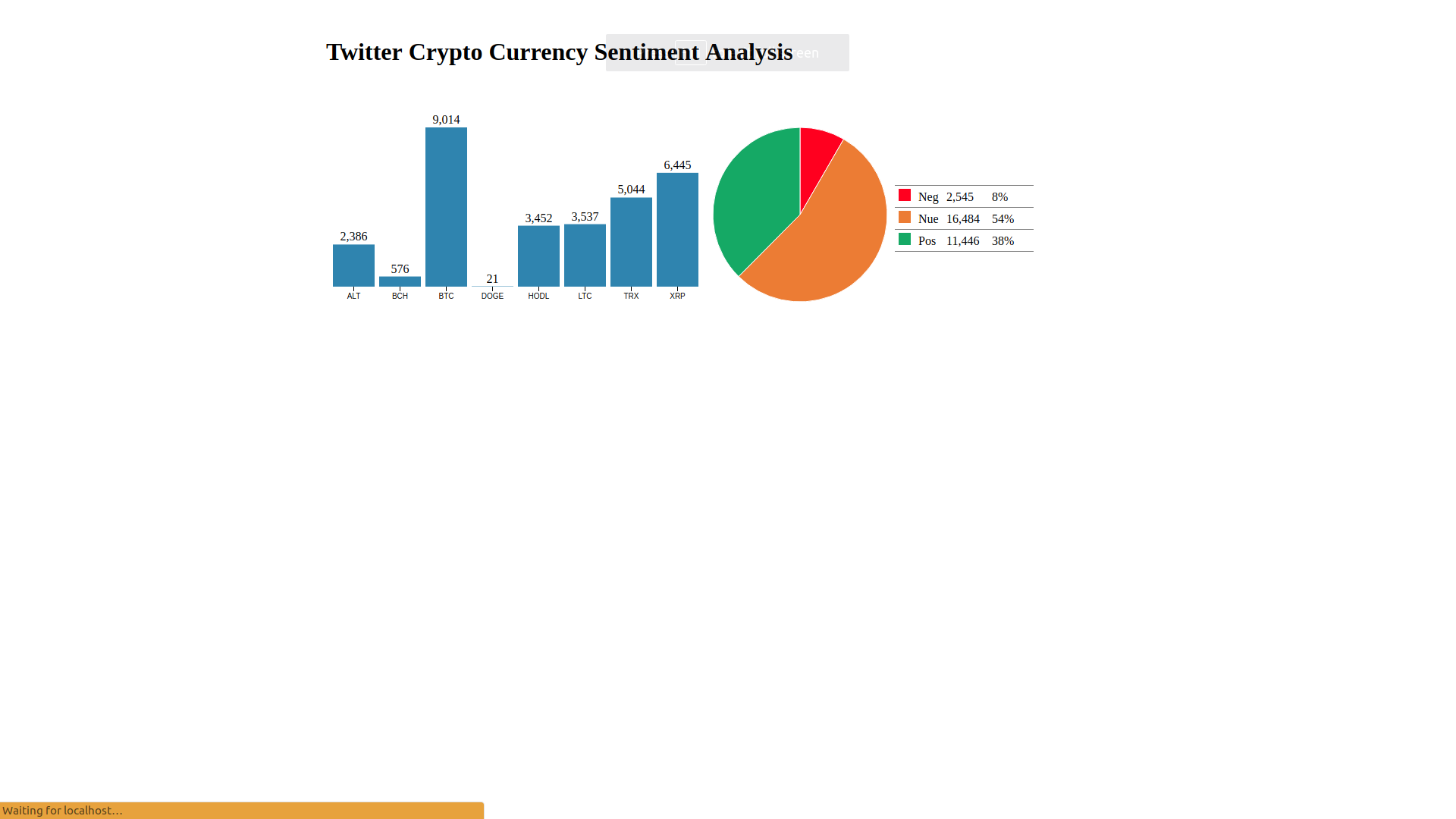
PHP Code for Data Visualization

polarity\_dashboard.php



# Results:

If you want to see the dynamic dashboard in action, please reach out to us. The dashboard is hosted locally.

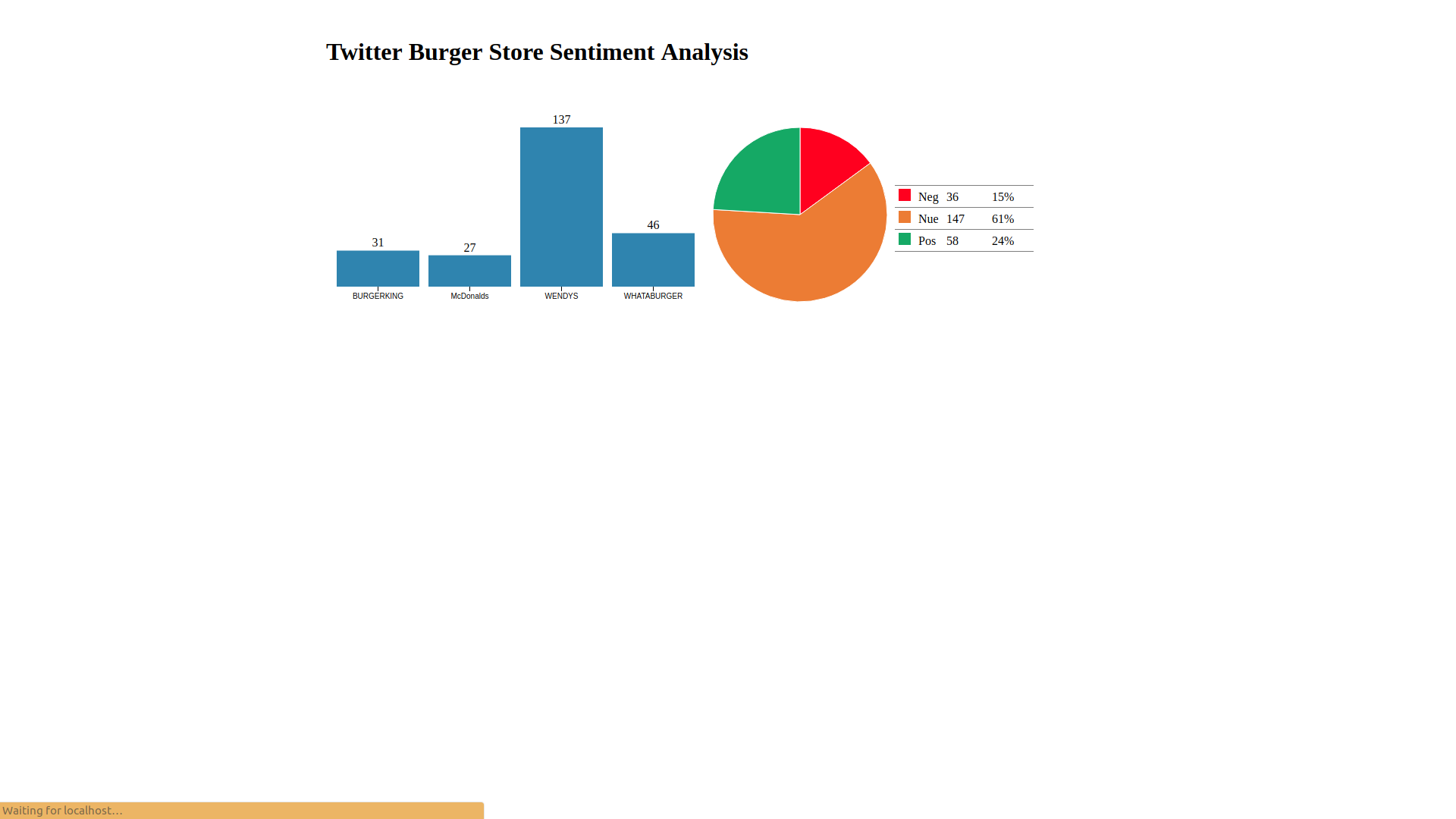


***Figure 5***

# Conclusion

The above system can be customized to see the Polarity of the tweets or contents with small tweaks.

For example, we tested the analysis on burger chain tweets. Below is the output that was customized for the burger store polarity tweets.



***Figure 6***