Data Visualization Dashboard

User Guide

April 2017 – Matthew Smith 12004210

This project was completed as part of the BSc (Hons) Computing Dissertation Module at The University of the Highlands and Islands

Setup

Requirements

- Linux Server (Ubuntu 16.04 used in development)
- Git
- NodeJS
- NPM
- Python 2.7
- Python Pip
- Apache Spark
- PM2 (Production Node Server)

Installation

To install git run

```
sudo apt-get install git
```

To install NodeJS and NPM run the following commands

```
sudo apt-get install nodejs
sudo apt-get install npm
sudo ln -s /usr/bin/nodejs /usr/bin/node
```

To install Pip run the following

```
sudo apt-get install python-pip
```

To install Apache Spark follow the following tutorial

https://www.tutorialspoint.com/apache_spark/apache_spark_installation.htm

then add the following to ~/.profile

```
export SPARK_HOME=/usr/local/spark
export PYTHONPATH=$SPARK_HOME/python:$SPARK_HOME/python/build:$PYTHONPATH
export PYTHONPATH=$SPARK_HOME/python/lib/py4j-0.10.4-src.zip:$PYTHONPATH
```

To install PM2 run the following command

```
sudo npm install pm2 -g
```

When the above has been completed, create a folder for the application then navigate into the folder. Then get the application from GitHub by running the following command.

```
git clone https://github.com/matthew110395/DataViewer.git
```

Then run the following command

```
Npm install
```

Then navigate to the Twitter folder and run the following command

```
pip install -r requirements.txt
```

Navigate to the parent directory and run the following command to start the server

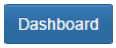
```
Pm2 start server.js
```

User Functions

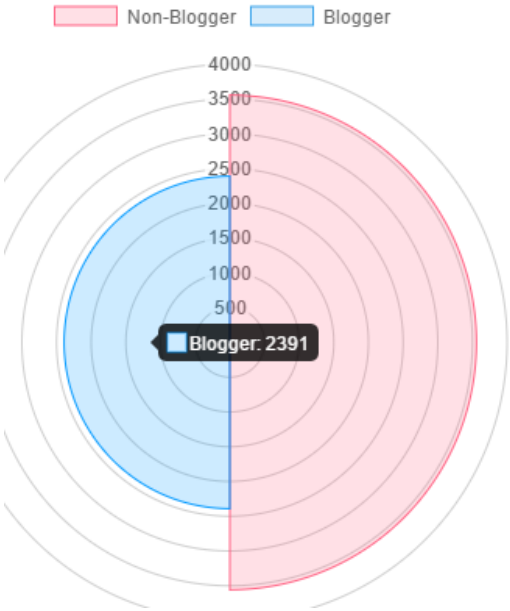
To Load the application navigate to <Server Name>:8000.

Viewing Charts

Charts will be displayed when the loading screen disappears, if the loading screen remains please contact the application administrator.

To view charts, from the configuration screen click 

To view the exact value of a point on a graph hover over it as shown below

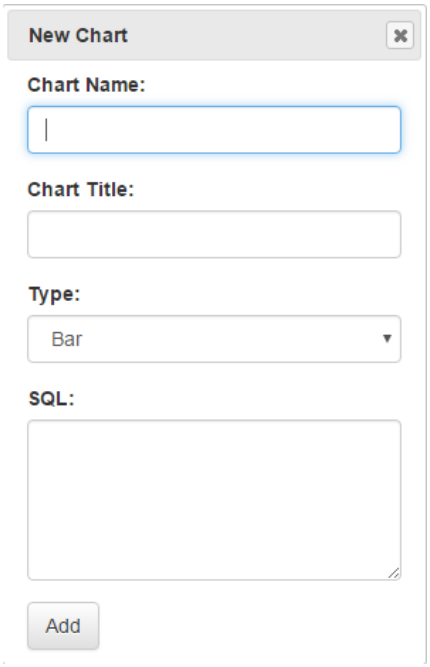


Adding Charts

Charts are added using the configuration screen, which can be accessed by selecting config on the main page or by starting the application when no charts exist.

1. Click 

and the following screen will appear

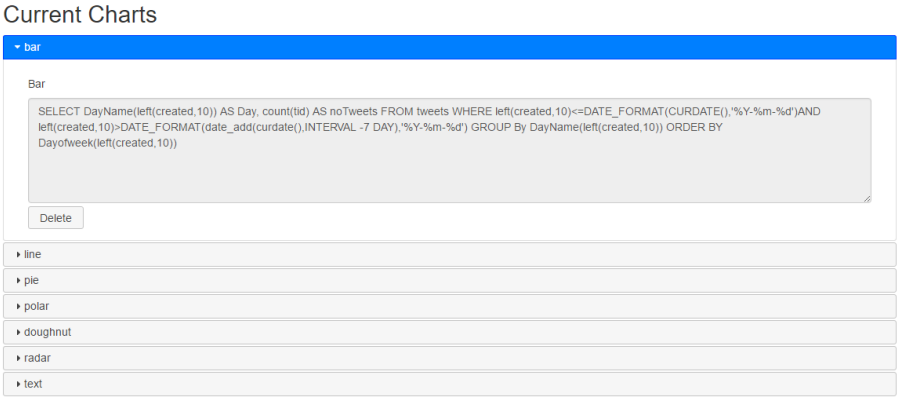


2. Enter a chart name, this is just to identify the chart and must be unique. No spaces are accepted
3. Enter a chart Title, this is displayed above the chart when it is Displayed
4. Select the type of chart you wish to create
5. Enter the SQL for the data which you want to view. This must be in the below format

```
SELECT <labelField> AS <DataLabels>, count(<CountField>) AS <Description>
FROM <Database>
```

6. Click Add

The chart will now exist in the current charts section of the configuration page as shown below
At present when charts are added, the server must be restarted to pull new charts.



Admin Functions

Database/Data Collection

This application requires a MySQL database to read data from. This is specified and edited in the current database section of the configuration screen.

If the collect data from Twitter checkbox is checked, data is collected from Twitter containing the word whisky. Machine learning is carried out on this data to determine if the tweet is by a blogger. To carry out machine learning, training data is required, this has been pre-populated for the current search Term, this can be found in the Twitter folder and is called mlData1.csv. Ensure tweets used for training contain no commas or new line characters as this will cause an error.

A log is produced as data is collected from Twitter, this can be found on the configuration screen and viewed by pressing the 

button

If the Collect data from Twitter check is a data base must use the same structure as the database in the example creation script shown below.

```
CREATE TABLE `tweets` (
  `tid` varchar(45) COLLATE utf8mb4_unicode_ci NOT NULL,
  `created` varchar(45) COLLATE utf8mb4_unicode_ci DEFAULT NULL,
  `tweet` varchar(450) COLLATE utf8mb4_unicode_ci DEFAULT NULL,
  `time_zone` varchar(45) COLLATE utf8mb4_unicode_ci DEFAULT NULL,
  `user` varchar(45) COLLATE utf8mb4_unicode_ci DEFAULT NULL,
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `tweetscol` varchar(45) COLLATE utf8mb4_unicode_ci DEFAULT NULL,
  `blog` int(11) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=1 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_unicode_ci;
```

To connect to the Twitter API, the app must be registered with Twitter.

1. Browse to https://apps.twitter.com/
2. Click Create an Application
3. Enter Information about your application
4. Click Create Application
5. Click the Keys and Access Tokens Tab
6. Click Create my access token at the bottom of the page
7. Note the consumer key, consumer secret, access token and access token Secret
8. Enter the above details into the configuration screen of the data viewer application