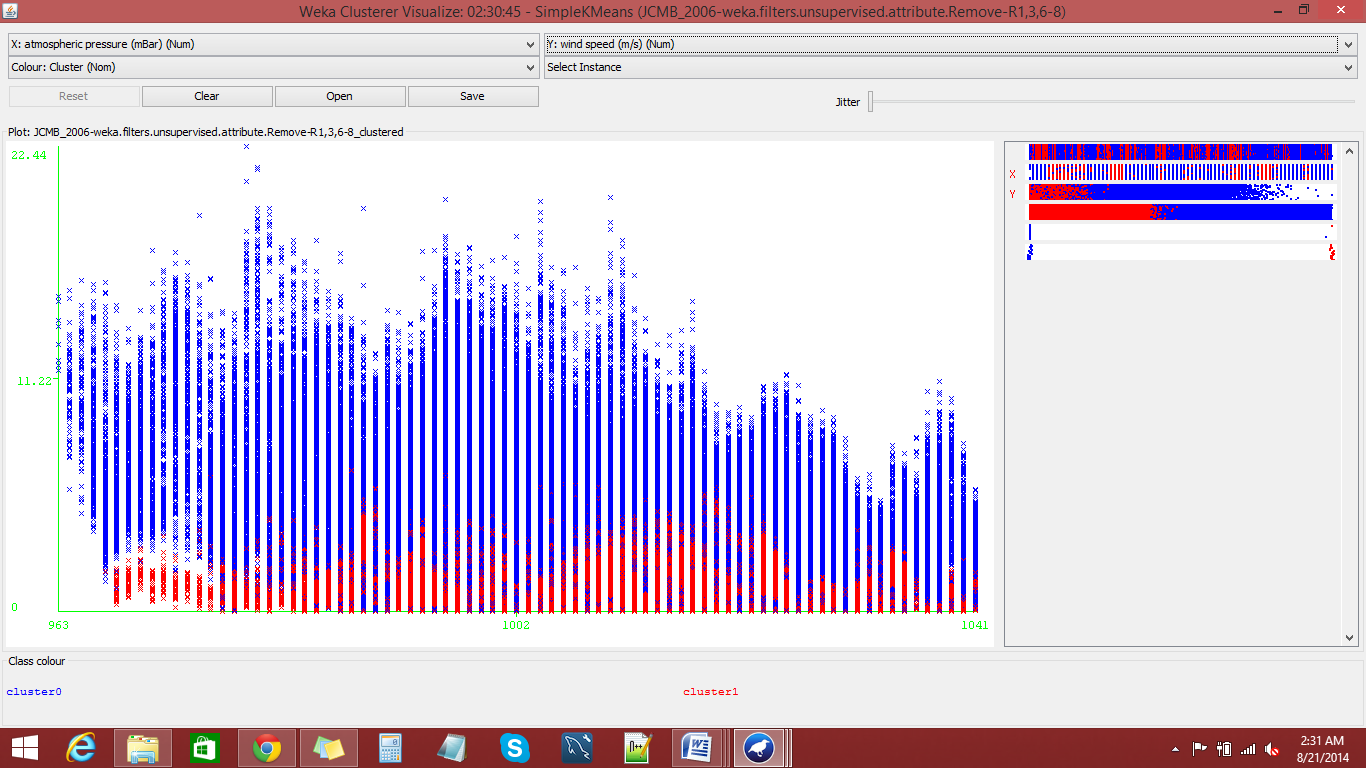
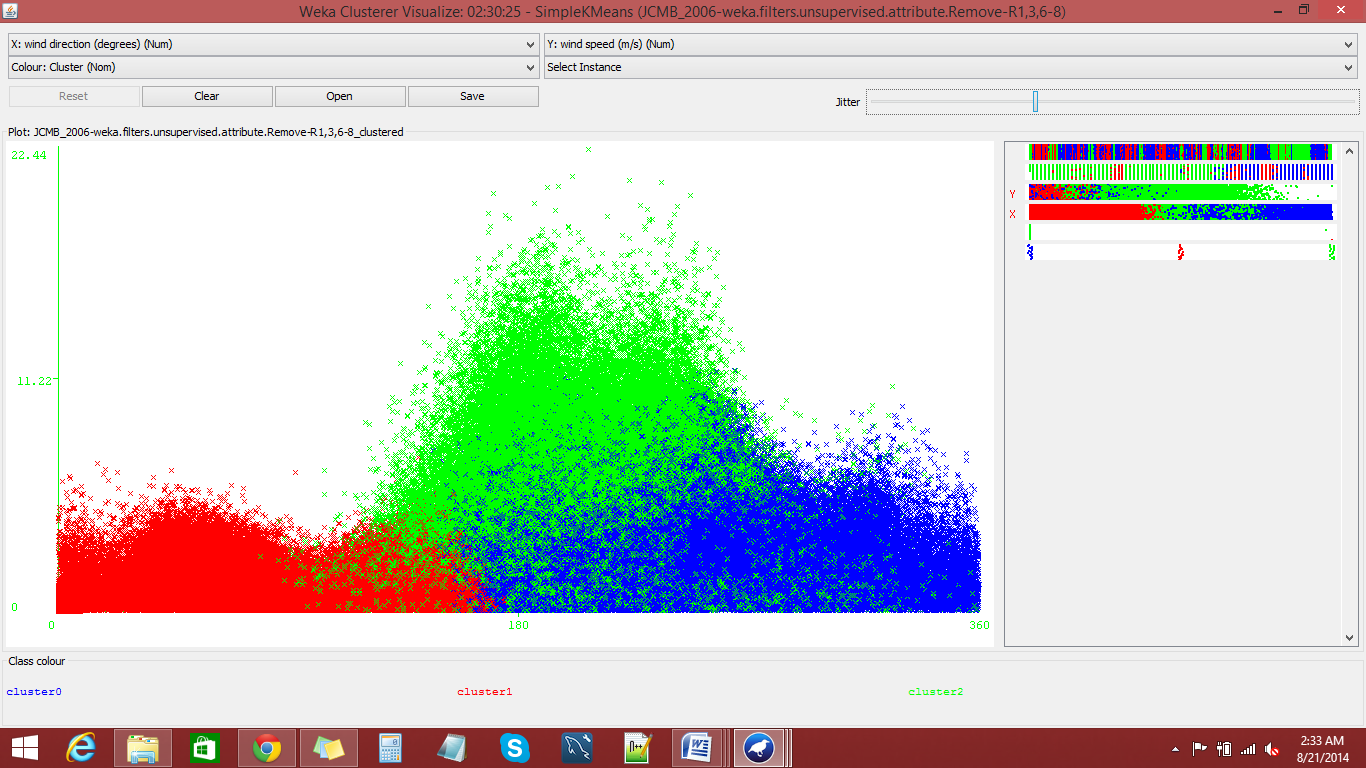
**CLUSTERING - Weka**

Applied clustering on weather data of size 19.4 MB and 351675 records of data. The Weather data was cleaned up before using it i.e. ad data or missing data was handled. K-means clustering was the method chosen and different cluster sizes were also used.

Below is the output from Weka Tool, where x-axis is Atmospheric Pressure and y-axis is Wind speed and cluster size is set to 2.



When cluster size is set to 3 and x-axis has wind direction and y-axis has wind speed is as follows:



**Visualizing the cluster by using D3.js**

The output of weka which is in .arff format should be converted into a .csv format and provided as an input to the Java Web Project created. Since we are trying to visualize the clusters in the cloud. We need to create an EC2 instance and install Eclipse tool in the EC2 environment. Import the html and the csv file into the Java Web project under the Web Content Folder. Now you will be able to see the output in the cloud (as shown below).

The output of Weka and the converted csv file is also attached for reference. The html file with the JavaScript code to visualize the cluster is also attached.

Below is the visualization of atmospheric Pressure on the X-axis and Wind Speed in the y-axis and the Clusters size is set to 2.

