

## Facebook Data Analysis with SAS® Visual Analytics

Piyush Singh, Ghiyasuddin Mohammed Faraz Khan, Vikrant Bisht, Prasoon Sangwan  
TATA Consultancy Services Ltd

### ABSTRACT

Nowadays, social media data is playing a key role in understanding market trends and user sentiment. Organizations are interested more than ever to know user choices, their needs, feedback and social media like Facebook is very common platform to gather this type of data. This paper demonstrates how SAS® Visual Analytics can be used to analyze Facebook Data and create new reports based on this data which helps the management to understand the market and users' preference. It also demonstrates the Word Cloud analysis on Facebook data using SAS® Visual Analytics. This paper contains some techniques which can be very helpful for SAS Coders, where they can write their own SAS Code for analytics. Users can bring Facebook data to SAS® Enterprise Guide or Base SAS and find all sorts of interesting things.

### INTRODUCTION

There are some tools and techniques available to read social media data but for SAS users, it would be great if they can read and analyze the social data with SAS itself. SAS Visual Analytics provides platform for SAS users and provides the functionalities to process the social media. There are many different features and functionalities available but this paper explains two key functionalities available with SAS Visual Analytics (SAS® VA):

- ❑ Facebook link/caption with published date and likes count
- ❑ Word Cloud and sentiments analysis

### 1. FACEBOOK LINK/CAPTION WITH PUBLISHED DATE AND LIKES COUNT

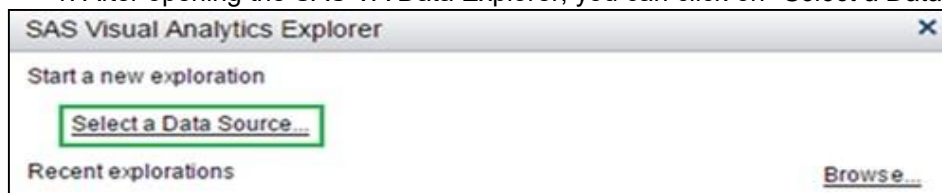
Before doing any analysis, first we need to read the Facebook data to SAS VA. SAS VA uses LASR internal server which contain the data and helps SAS VA to analyze the data in a fast and efficient manner. Once we load the Facebook data to LASR server, it is ready for user query in SAS VA or in other SAS client like Enterprise Guide.

#### 1.1 READ/ LOAD DATA TO SAS VISUAL ANALYTICS AND LASR SERVER

When we read Facebook page to SAS VA, it reads and loads the data to LASR server. When you create any graph or do any statistics, SAS VA reads the data to and from the LASR server. During SAS VA installation and configuration, it creates many different web link to do different kind of activities. To import Facebook data, you can use SAS VA Hub link:

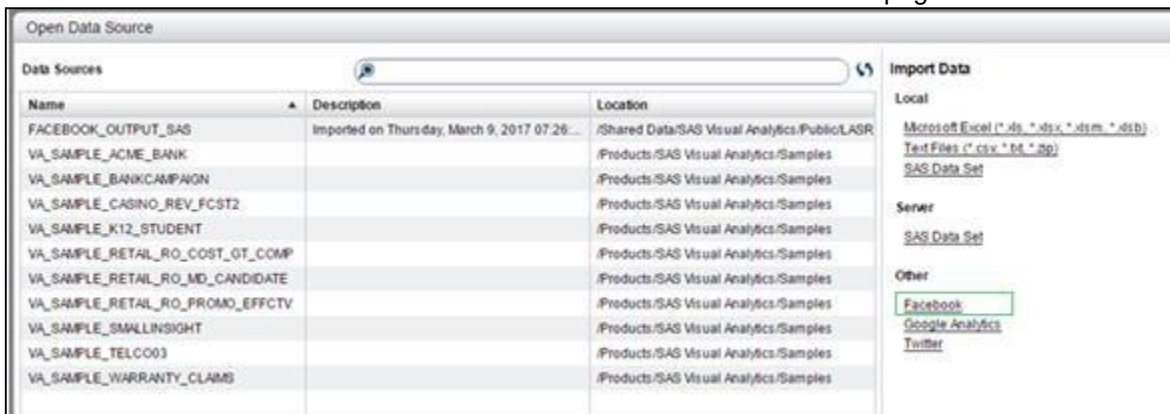
[http://<sasva\\_server:7980>/SASVisualAnalyticsHub](http://<sasva_server:7980>/SASVisualAnalyticsHub). You can perform following steps to read the Facebook page in SAS VA:

1. After opening the SAS VA Data Explorer, you can click on “*Select a Data Source*”.



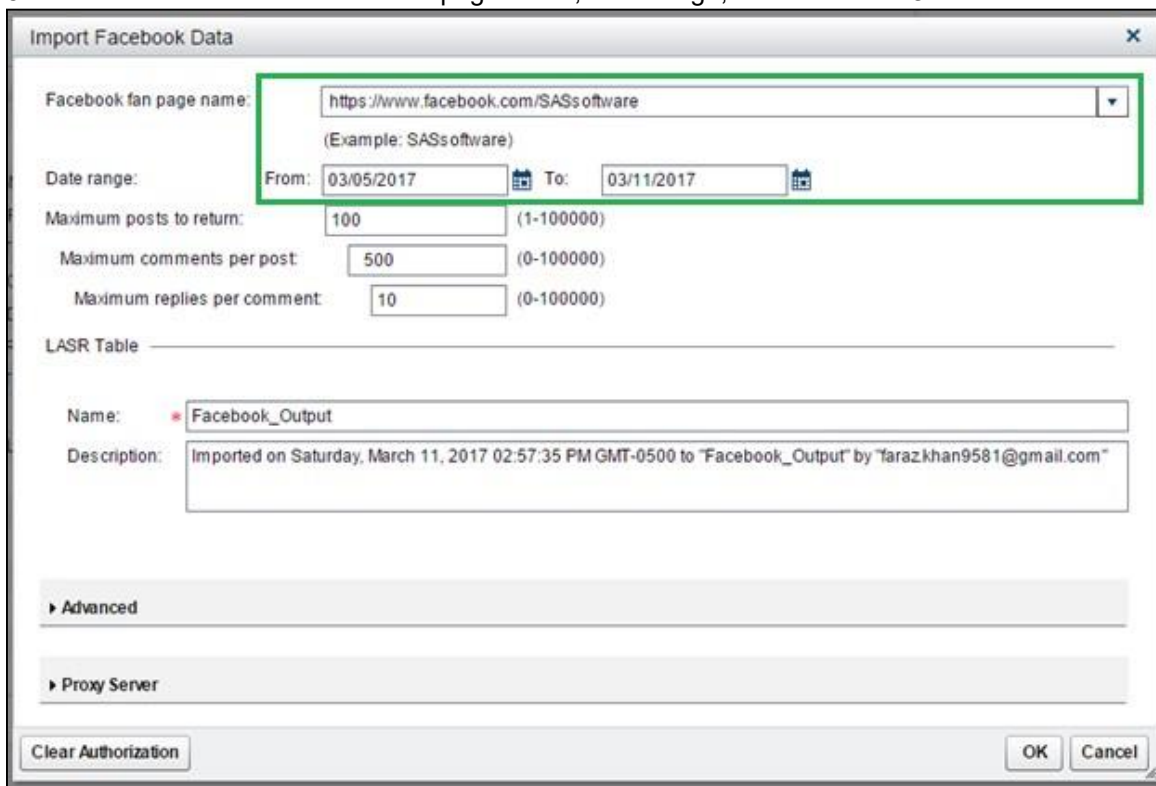
### Display 1. Select Data Source

- It opens a new window which shows the available tables and different source of data which SAS can use.
- You can select "Facebook" to launch a window to read the Facebook page.



### Display 2. SAS Visual Analytics Explorer

- A popup appears asking permission to import the Data from Facebook.
- You can enter the Facebook fan page name, Date range, etc. and click "OK".



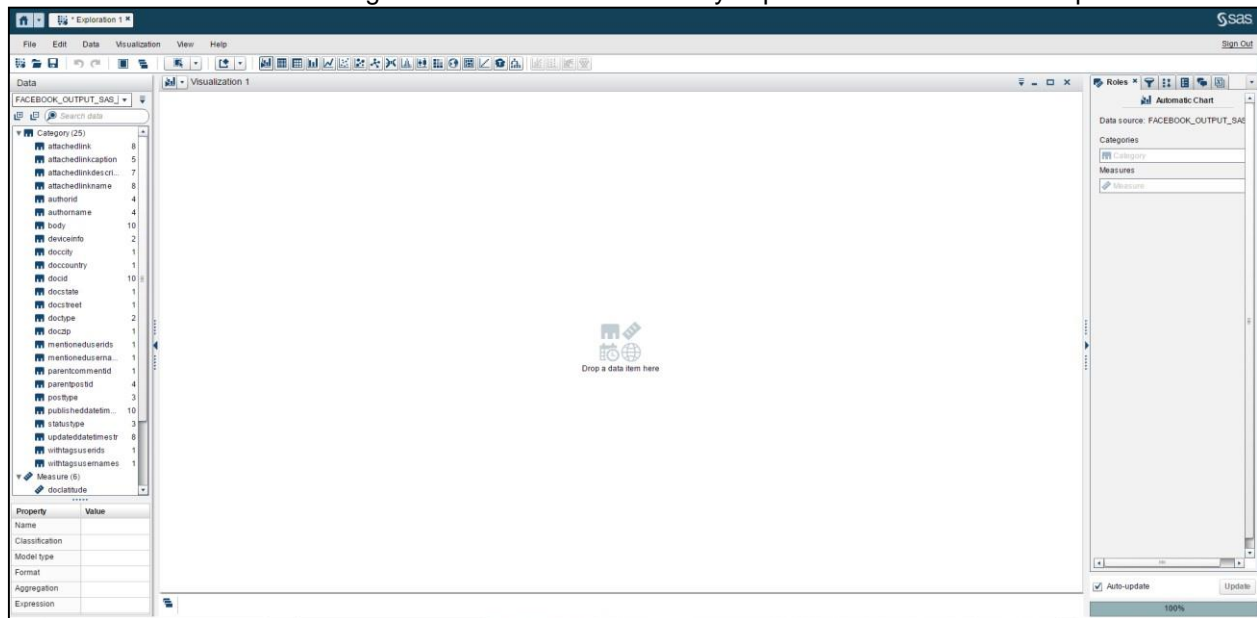
### Display 3. SAS VA UI for Facebook Import

**Fan Page** – A fan page is a public page created by organizations to represent themselves. We can either enter the full path of the fan page or just the name of the fan page from which the data is to be imported.

**Date Range** – This option is used to extract data from a specific time frame to SAS LASR Server

**LASR Table** – Name of the LASR table you want to create for the imported data. LASR table can help you if you want to process the same data from a previous import. You won't need to import data again and again. As discussed earlier, it helps to get quick response from visual analytics parameter changes.

6. The Facebook Fan Page Data has been successfully imported into the SAS VA Explorer.

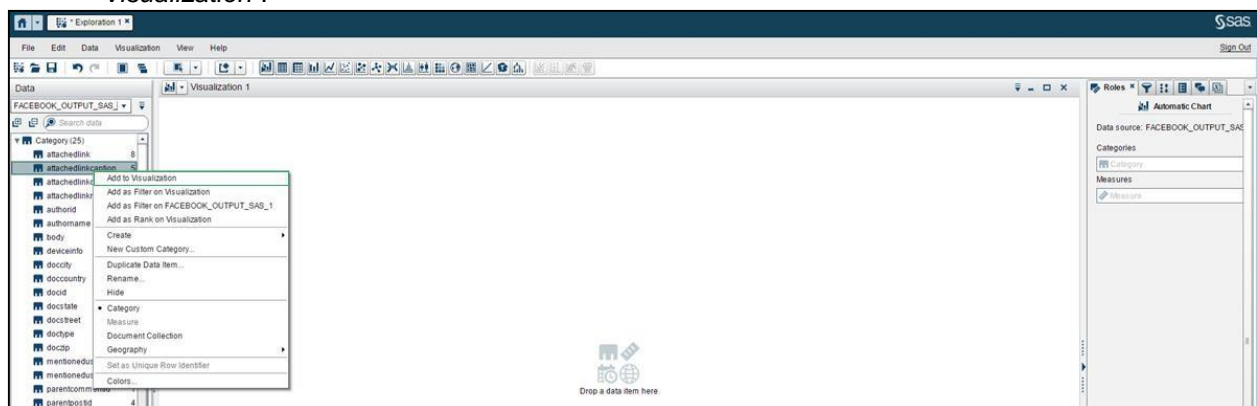


Display 4. Successful Read of Facebook Page

## 1.2 LIKES ANALYSIS FOR PUBLISHED LINKCAPTION

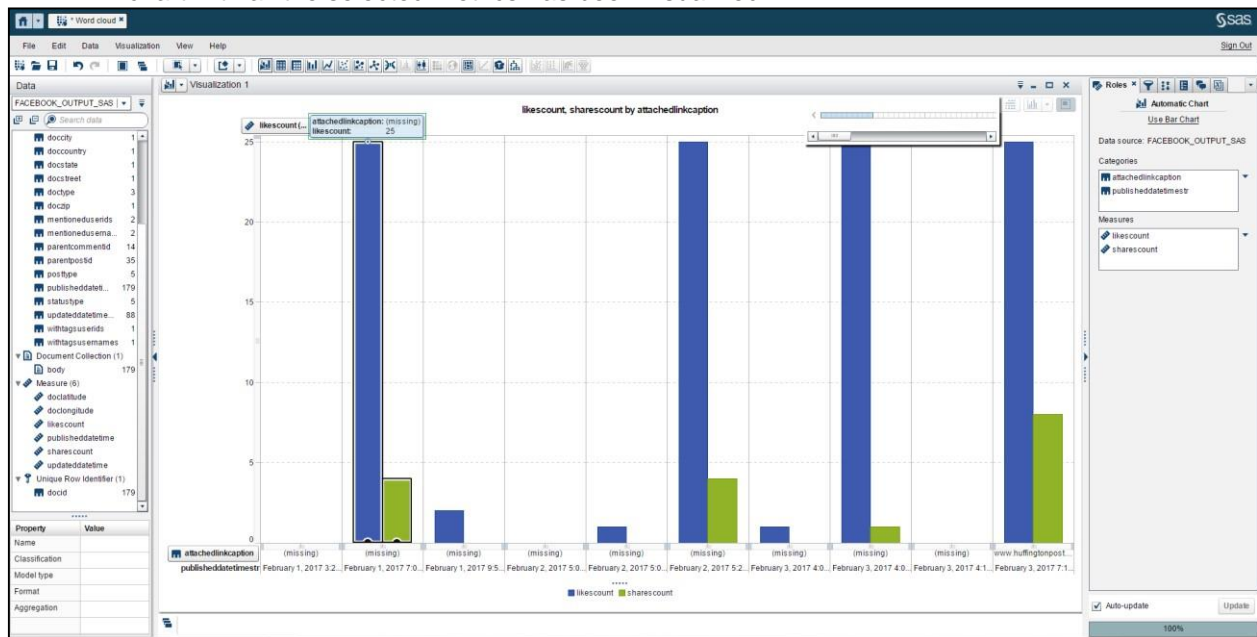
Reports can be created by visualizing the Facebook fan page data imported in SAS VA. These reports can be used by the management in order to understand the market trend, users' mood and interest. It is only able to show the incorrect value of 25 for the data Facebook data visualization which was imported into SAS VA. You can perform the following steps to create a visualization from the Facebook data imported in SAS VA:

1. Right Click on "*attachedlinkcaption*" (or any other Data Source from Category) and click "*Add to Visualization*". Similarly, right click "*likescount*" from the measures and click "*Add to the Visualization*".



## Display 5. Creating Visualization

2. A chart with all the selected Metrics has been visualized.



## Display 6. Successful Visualization of Data imported from Facebook.

In the visualized chart, the imported data has a problem with "*likescount*" measure. The visualized charts do not show the correct value for the "*likescount*" measure for the data which has more than 25 likes.

**Note** - If your SAS VA environment doesn't have the latest installed hotfix, data source with the likescount measure of 25 or less is displayed perfectly but the data source with more than 25 for instance say 45 likes is also showed as 25 instead of 45 likes.

This issue with the "*likescount*" measure is fixed via Hot fix which can be obtained by contacting SAS Technical Support.

## 2. WORD CLOUD AND SENTIMENT ANALYSIS

The Facebook fan page data uploaded into SAS VA can be used for Word Cloud Analysis. Word Clouds are used to represent a group of words from a character data item. The size of words in a Word Cloud represent the frequency or significance of the words and other such information. Word Clouds helps to analyze the sentiment of users from Facebook shares and updates. This also helps the business to read and understand the opinions, expectations and experience of the customers over multiple channels.

There are two types of Word Clouds available for SAS VA. They are:

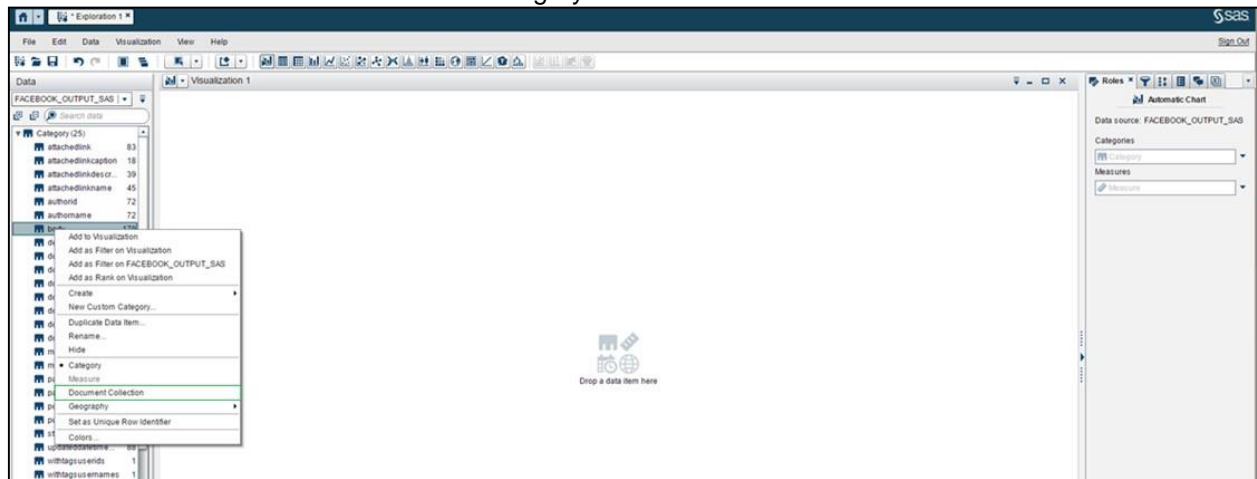
- **Word Clouds that uses Text Analysis** - For Text Analysis Word Cloud, first we need to create a new data role called Document Collection. In the Document Collection data role, each value is referred to a separate word document which could contain multiple words. The words which usually come together are called as Topics. This type of Word Cloud can contain multiple Topics and the size of the words in these clouds represents the significance(relevance) of the words.
- **Word Clouds that uses Category Values** - This Word Cloud can use the color and size of the words in order to depict the value of the measure items. This cloud generally recognizes each

value in a Category Data item as text. This Word Cloud usually represents text with the most occurrences or with the highest value for a measure item.

## 2.1 WORD CLOUDS THAT USES TEXT ANALYSIS

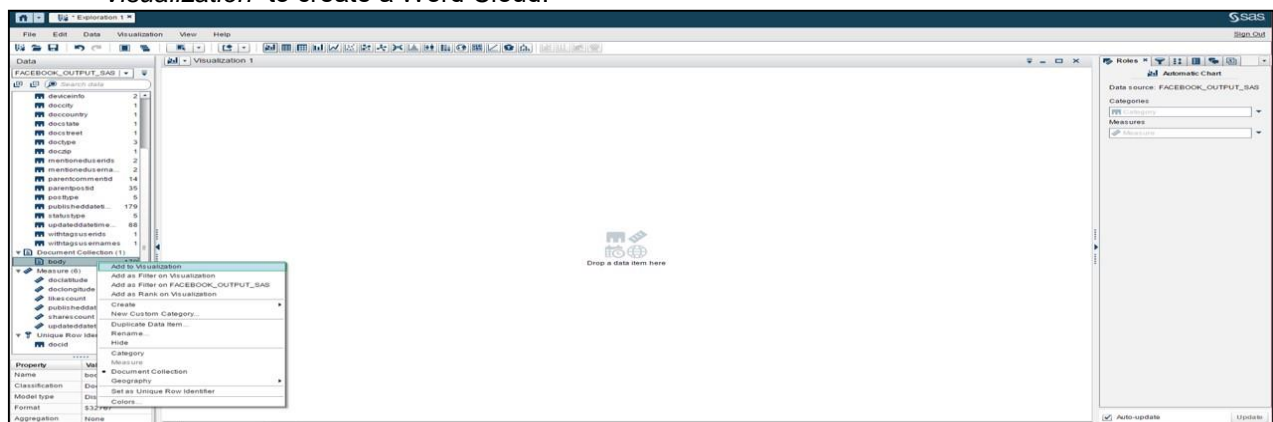
You can perform following steps to create a Word Cloud from the Facebook data imported in SAS VA:

1. After importing the data into SAS VA, you can right click on any category and select “*Document Collection*”. This Document Collection data role will be used for analysis in these clouds. The newly created document collection will have separate topics with their own Word Clouds. As shown in Display 7, once you chose to create a Document Collection from a data source, that Data Source is moved from the Category Data Role to the Document Collection Data role.



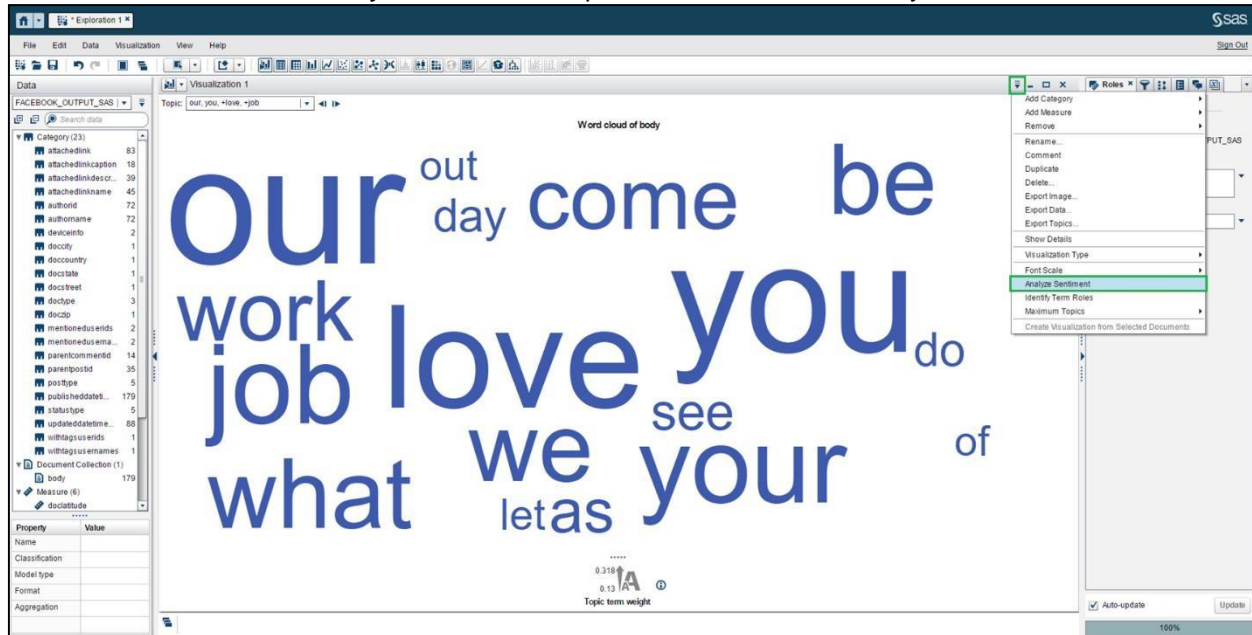
Display 7. Creating a Document Collection for Word Cloud.

2. Select a unique row identifier column comes up. Generally, a data source or measure is selected as a unique identifier for each row. An exclusive value should be assigned for every unique row identifier and repetition in these values can result in erratic results. You can click “OK” to continue after selecting a data source as unique identifier column.
3. Time taken to create the visualization is based on size of data source used. If the data source used is of bigger size, another popup comes up confirming the size of the characters in data source. You can click “OK” to proceed with creating the Document Collection.
4. The new Document Collection has been created. As shown in the screenshot below, you can right click on the data source whose document collection has been created and click “*Add to Visualization*” to create a Word Cloud.



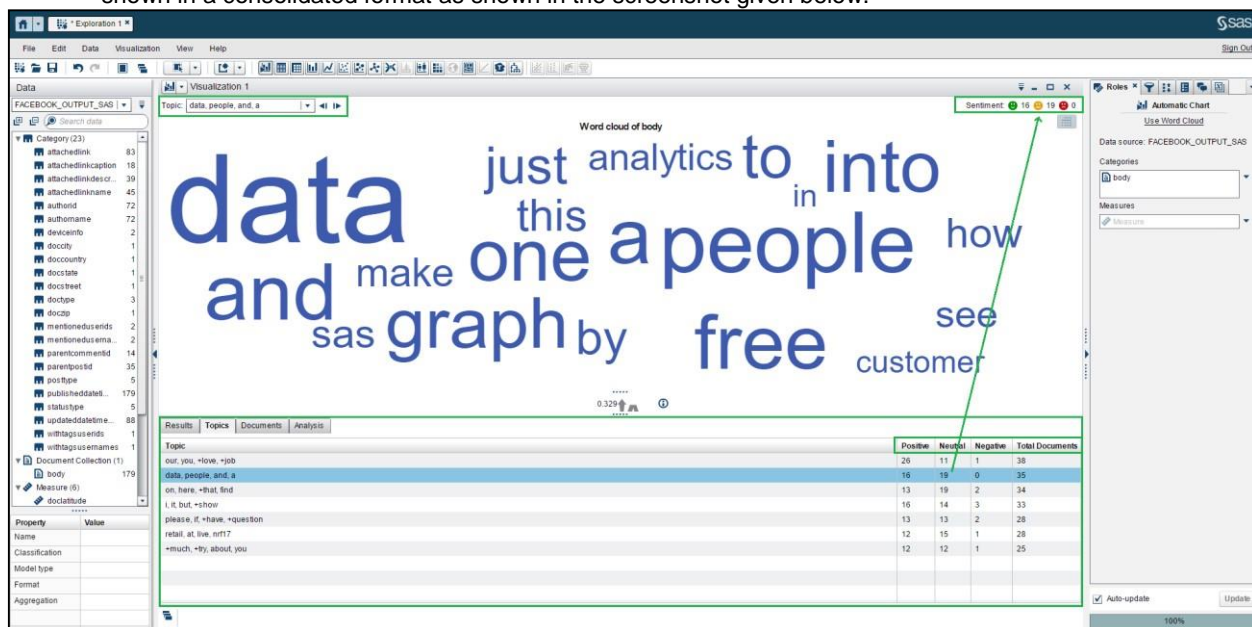
## Display 8. Document Collection

- The Word Cloud generated by SAS VA is a Text Analysis Word Cloud by default. After the Word Cloud has been created, you can right click on the arrow highlighted in the screenshot given below and select “Analyze Sentiment” to perform the Sentiment Analysis.



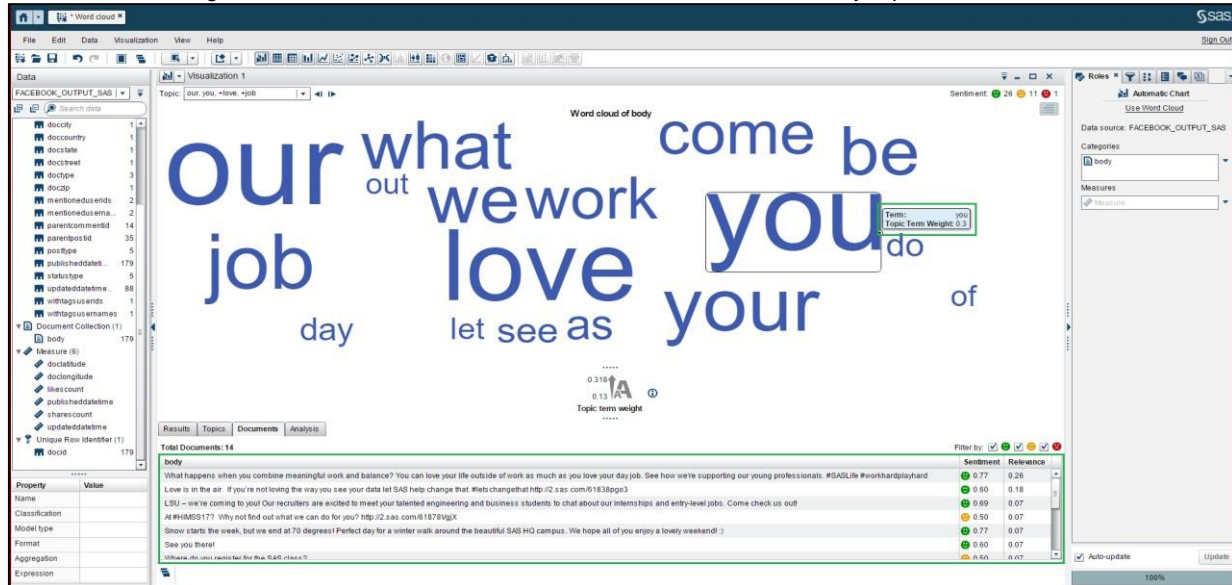
## Display 9. Word Cloud

- The Sentiment Analysis has been completed. As stated above, a Word Cloud based on Text analysis usually contains multiple Topics. Each of these Topics has its own Word Cloud and you can switch between different topics by using the “Topic” drop down menu or from the “Topics” Tab located in the highlighted parts in the below screenshots. The sentiments of all the Documents in all the Topics are analyzed and shown in a consolidated format as shown in the screenshot given below.



## Display 10. Topics Tab

- The “*Documents*” tab in the bottom of the screen contains the list of all the documents contained in any specific topic. This tab also shows the Sentiment Analysis of each document along with the Relevance of the document. You can also hover your cursor over a word in the Word Cloud in order to displayed the Topic Term Weight of the word. The contents of this tab are exclusive to every topic has its own set of documents.



## Display 11. Documents Tab

- The “*Results*” Tab consists of a table of all the Words in the Word Cloud and their Topic Term Weight as shown in the screenshot below.

Term	Topic Term Weight
data	0.329
people	0.246
a	0.232
and	0.232
free	0.228
one	0.221
into	0.22
graph	0.215
to	0.194
just	0.175

## Display 12. Results Tab

- The “*Analysis*” Tab contains a list of all the relevant terms associated with the document. The contents of this tab are generally constant for all the topics in the Word Cloud.

Term	Definition
Term Role	A function that is performed by a term in a particular context. A term can function as a part of speech, entity type, or other purpose that is user-defined.
Topic Term Weight	An indicator of the relative importance of a term in a topic as compared to other terms. A term with a value above a specified cutoff value contributes to the assignment of a document to the topic.
Relevance	A score that indicates how well a document satisfies a rule or model. The best match has a score of 1 and reflects a perfect (100%) match.
Sentiment	An attitude that is expressed about an item that is being analyzed, which can be a segment of text, a grouping of text segments, or a specific subject of interest.
Sentiment Analysis	The use of natural language processing, computational linguistics, and text analytics to determine the attitude of a speaker or writer with respect to a topic, document, or other item of analysis. Sentiment analysis results in a positive, negative, or neutral score on the target of analysis.

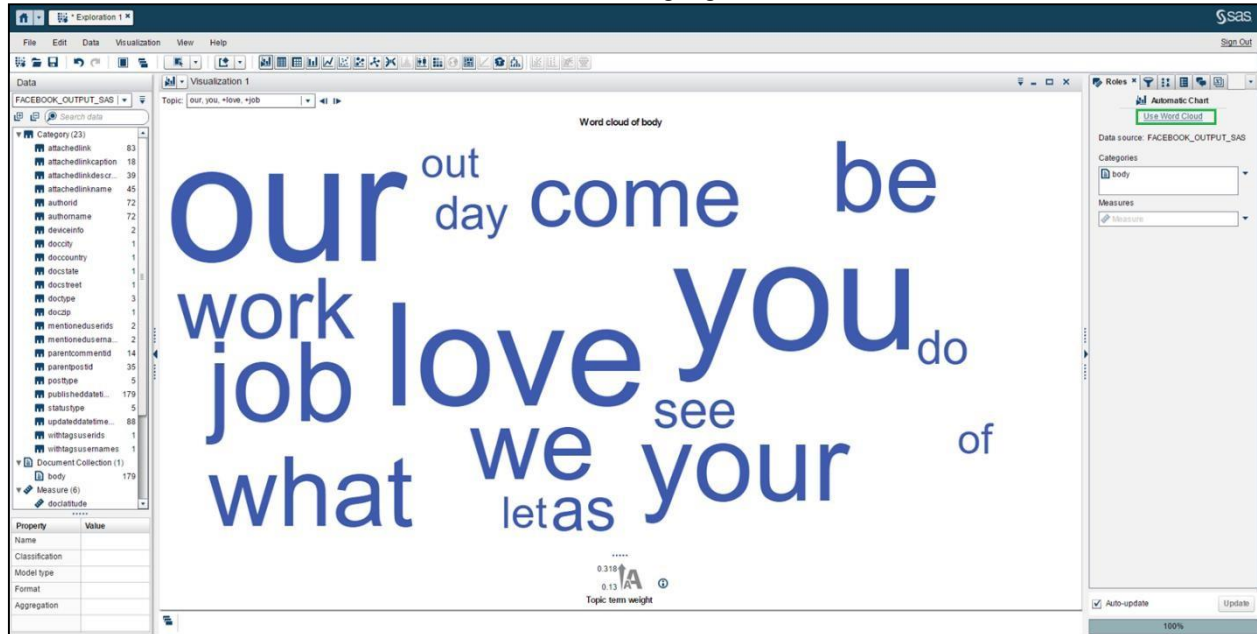
## Display 13. Analysis Tab

In order to create a Category Values Word Cloud, you can follow the steps given in the below section.

## 2.2 WORD CLOUDS THAT USES CATEGORY VALUES

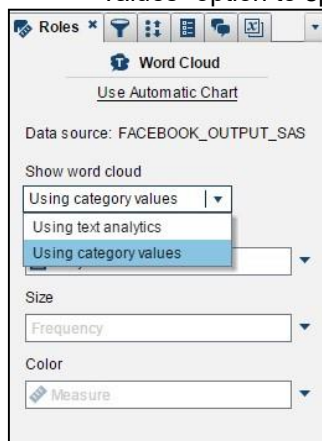
Category Values Word Clouds is created from a basic Word Clouds. The steps mentioned in previous Section can be followed to create a Word Cloud from the Facebook fan page. After creating the Word Cloud, you can follow the below steps in order to analyze the Category Values:

1. You can click on “Use Word Cloud” button as highlighted below.



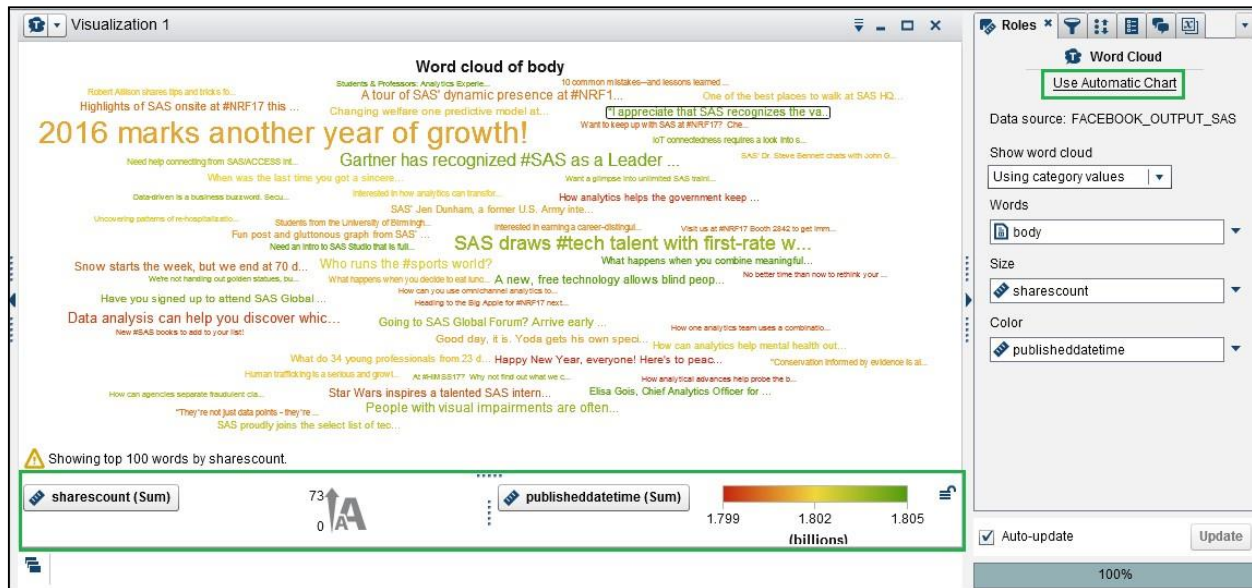
Display 14. Word Cloud

2. Clicking on “Use Word Cloud” button will create 4 drop down options. The “Show word cloud” option is used to select the type of Word Cloud to be created. You can select “Using category values” option to specify SAS VA to create a Word Cloud based on the Category Values.



Display 15. Use Word Cloud Option

3. You can select the appropriate Category Data Source value for the “Words” dropdown menu. “Color” and “Size” dropdown menus can be populated with the measures, whose values need to be used as parameters.



Display 16. Category Values Word Cloud

As shown above, the "publisheddatetime" is the measure for the color of the word and "sharescount" is the measure for the size. The data range for the size and the color of the words is given in the bottom of the screen (The first range is for the size of the words represented by the arrow and the second range is for the color of the words).

Below screenshot shows, when we hover over the words "2016 marks another year of growth" we see that "sharecount" value is 73 which is highest (as seen in Display 16 size data range) and results with bigger size of the word. Also the color of the word depends on the value for "publisheddatetime" i.e. 1800973293, which is in the middle of the data range for color, and hence the color is Orange.



Display 17. Sample Word Data in Word Cloud

Similarly as shown in the below screenshot, when you hover over the word "Need an Intro to SAS Studio", we see that the "sharecount" value for this word is 4 which is low and results with smaller size of the word. The color of this word is Green because the value for "publisheddatetime" is 1804266003 which close to the upper limit of the data range.



## Display 18. Sample Word Data in Word Cloud

This type of analysis can be done in conjunction with different Category Data Sources with varied measures to study and understand the market trends.

## CONCLUSION

This paper provides an introduction of importing facebook data in SAS Visual Analytics, how it can be analyzed, understand market trend and user preferences. It also provides an introduction to Word Cloud Analysis and Sentiment Analysis, which helps organizations to understand customer needs, feedback, and other such market like/dislike. These different kind of analysis with social media helps to plan and act as per their users need/ interest.

## REFERENCES

Singh, Piyush. Sangwan, Prasoon. Kumar, Veera. and Sanger, Todd. Nov-2016. "SAS® Enterprise Guide for SAS® Visual Analytics LASR Server" *Proceedings of South Central SAS User Group 2016*. San Antonio, TX, USA: SCSUG. Available at:

<http://www.scsug.org/wp-content/uploads/2016/11/SAS%C2%AE-Enterprise-Guide-for-SAS%C2%AEVisual-Analytics-LASR-Server.pdf>

Sangwan, Prasoon. Gupta, Tanuj. and Singh, Piyush. Apr-2016. "Key Requirements for SAS® Grid Users" *Proceedings SAS Global Forum*. Las Vegas, NV, USA: SAS Institute. Available at:

<http://support.sas.com/resources/papers/proceedings16/7140-2016.pdf>

Singh, Piyush. and Pohl, Gerhardt M. Mar-2014. "Enhancing SAS® Piping Through Dynamic Port Allocation" *Proceedings of SAS Global Forum 2014*. Washington, DC, USA: SAS Institute. Available at:

<http://support.sas.com/resources/papers/proceedings14/1826-2014.pdf>

## RECOMMENDED READING

- SAS® Visual Analytics 7.1 - User's Guide

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Piyush Singh

Ghiyasuddin Mohammed Faraz Khan

[piyushkumar.singh@tcs.com](mailto:piyushkumar.singh@tcs.com)

[ghiyasuddin.farazkhan@tcs.com](mailto:ghiyasuddin.farazkhan@tcs.com)

Vikrant Bisht

Prasoon Sangwan

[vikrant.bisht@tcs.com](mailto:vikrant.bisht@tcs.com)

[prasoon.sangwan@tcs.com](mailto:prasoon.sangwan@tcs.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. indicates USA registration. Other brand and product names are trademarks of their respective companies.