

Get the Full Picture With SAS

Andy Bayliss, SAS, Marlow, United Kingdom

ABSTRACT

How do you turn a billion data points into one actionable insight, how do you improve the visibility and understanding of your data, and how do you provide trustworthy insights to a range of different audiences with data visualization? This paper explore how SAS can be used to deliver the full picture of your data.

INTRODUCTION

This paper, and the supporting presentation in this year Data Visualization Stream at Phuse Connect 23, discuss some of the data visualization capabilities available with SAS today and how these can be used to deliver insight into your data:

Topics covered include:

- The creation of visualizations with SAS code using SAS GRAPH, highlighting some of the key features required for clinical trial analysis and reporting.
- How SAS can provide fast insight, data exploration, and dashboarding for large and small datasets and multiple data types without the need to write any code.
- How these capabilities can be extended to incorporate open-source visualizations.
- How the visualizations available in SAS can be used to provide deeper understanding of data allowing you to concentrate on what your data picture is saying.

A PICTURE SAYS A THOUSAND WORDS!

"A Picture Says a Thousand Words" is a phrase we all know and one we all recognize to be true. While some believe this is a Chinese Proverb and attribute it to Confucius, its first use was, according to Wikipedai [\(1\)](#), at a banquet on journalism and publicity in 1911!

These days we are bombarded with images and trying to make an image memorable and meaningful is hard. Hopefully some of you will remember the image of a Tree Frog I've included on my previous Phuse presentations and papers from the last 10 years, even if you don't remember any of the things I presented. But is a picture enough?



A Picture Says A Thousand Words

Figure 1 A Picture Says A Thousand Words

LEONARDO DA VINCI – A GREAT ARTIST BUT DID HE DELIVER?



Figure 2 Leonardo Da Vinci

Leonard Da Vinci is a bit of a hero of mine for lots of reasons and many of you will be familiar with his work including, probably his most famous work, the Mona Lisa which you can now see in the Louver in Paris. This is certainly a memorable image but did you know that it took Leonardo over 16 years to paint and was never actually delivered to the person who commissioned it [\(2\)](#). So while it's a memorable image it's not a great example of delivering what is needed when required.

SO HOW DO YOU USE AN IMAGE TO SHOW ACTIONABLE INSIGHT FROM DATA FAST

We collect data from everywhere and everything and every day we get more and more data from more and more sources. So, what happens when you are collecting millions or billions of data points from processes, from sensors or from patients?

You need an approach that:

- Takes the data you collect.
- Can apply the right visualization to the data.
- Can apply the right analytics to the data.
- And can do it in seconds.
- Without having to write any code (unless you really want to).

SAS TO THE RESCUE

With SAS you have always been able to produce complex graphs from large amounts of data and many of you reading this paper will have far more experience in doing this than me to do this using SAS code but this can take time and effort. While the SAS code required was very simple creating the chart show opposite this still took around 26 minutes to produce when executed against 370 Million rows data.

```
PROC GCHART DATA=WORK.RealWorldData;
VBAR
  datamonth
  /
  CLIPREF
    FRAME
    TYPE=FREQ
    COUTLINE=BLACK
    RAXIS=AXIS1
    MAXIS=AXIS2
;
RUN; QUIT;
```

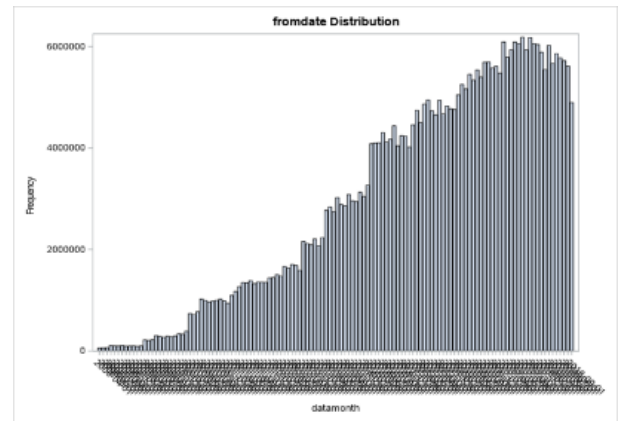


Figure 3 Age Distribution 1

For those of you who want to know more about coding visualizations in SAS then I've included some useful references below [\(3\)](#). And, as an aside, if you are interested in producing clinical graphs with SAS Code then I'd recommend you get this book or look at the multiple Phuse papers from Sanjay Matange, on one of my former colleagues at SAS, and available at Lexjansen.com [SAS Proceedings and more \(lexjansen.com\)](http://lexjansen.com)

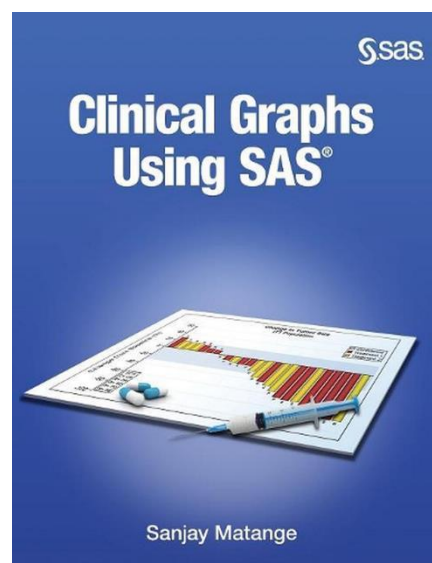


Figure 4 Clinical Graphs Using SAS

What if you don't want to write any code but still need to produce visualization from your data! That's where SAS Viya can help.

If you have never heard of SAS Viya, then where have you been! In a nutshell SAS Viya combines all the features and power of the SAS Programming Language we all know and love, with the latest in-memory, parallel processing capability in a cloud native architecture that can scale on-demand, and is simply:

“The quickest way from a billion points of data to a point of view”.

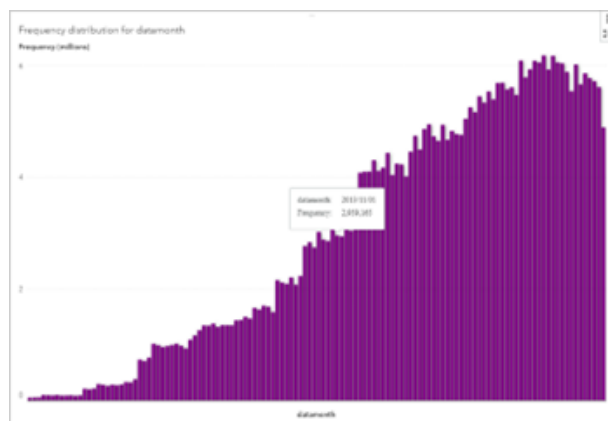


Figure 5 Age Distribution 2

Using SAS Viya, I was able to produce the same chart with the same 350 million rows of data in just **2 seconds** and I didn't need to write any code to do it.

You can find out more about SAS Viya and how it supports the full analytics lifecycle from to Questions to Decisions using the links provided below [\(4\)](#) but for the rest of this paper I'll focus on some examples of the sorts of visualization you can create in with SAS Viya.

SAS VIYA VISUALISATIONS THAT DELIVER INSIGHT

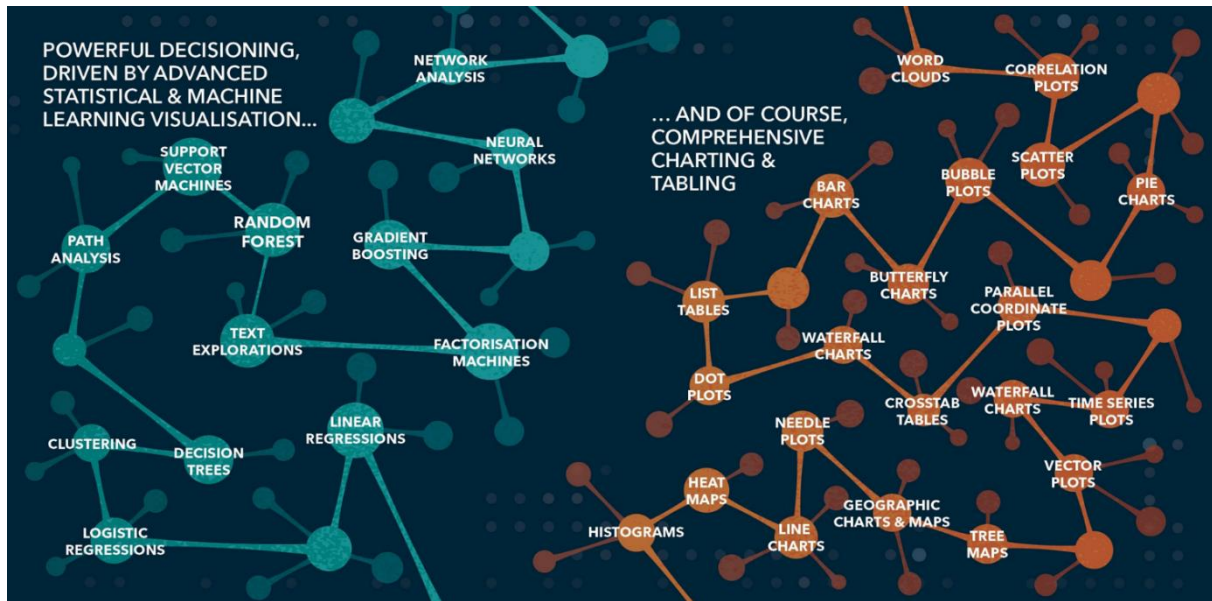


Figure 6 SAS Viya - Delivering Advanced Analytics and Visualization

SAS CODE:

As I mentioned earlier SAS Code is still fully supported in SAS Viya and you can produce all the visualizations you need using code if you want to. Here's a simple Bar Chart created using the standard tasks available within SAS Studio.

If you want to know more about the type of chart available in SAS Code then take a look here.

[SAS/GRAPH | SAS Support](#)

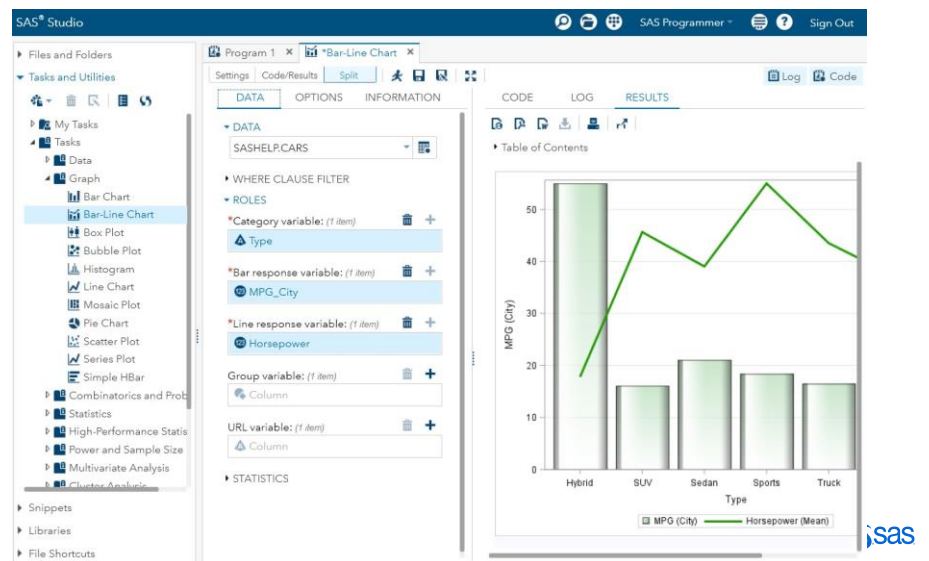


Figure 7 SAS Studio

ONE-CLICK VISUALISATION WITH SAS VISUAL ANALYTICS

The examples below are from SAS Visual Analytics. In the first image you can see an example of the visualization that SAS Viya automatically suggests based on the data available to help you provide instant insight. In this case the data is a 3.6m million rows of data relating to customer satisfactions and you can immediately see that while there are group of customers who are very satisfied with ratings in the 90-100% section and very few very dissatisfied customers in the 0-10% section, there are a larger number in the 30-50% range. All of this warrants further investigation and the interactive nature of the dashboard you can create in SAS Visual Analytics can help provide additional insight in just a few mouse clicks.

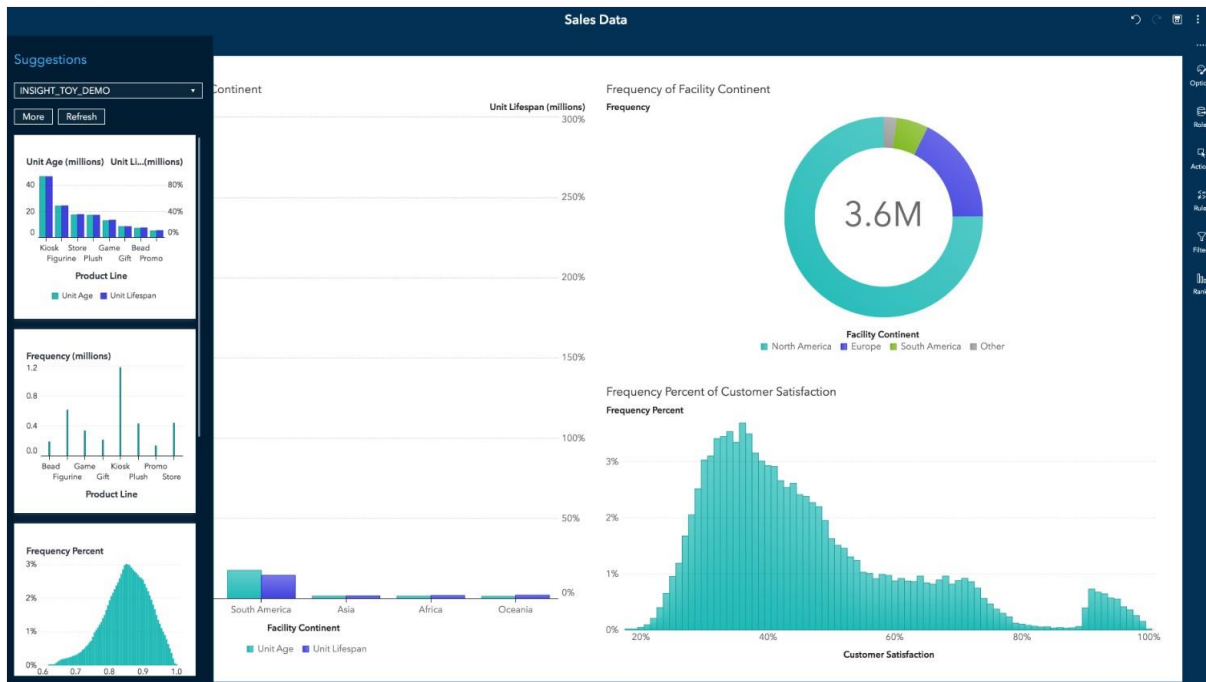


Figure 8 SAS Visual Analytics 1

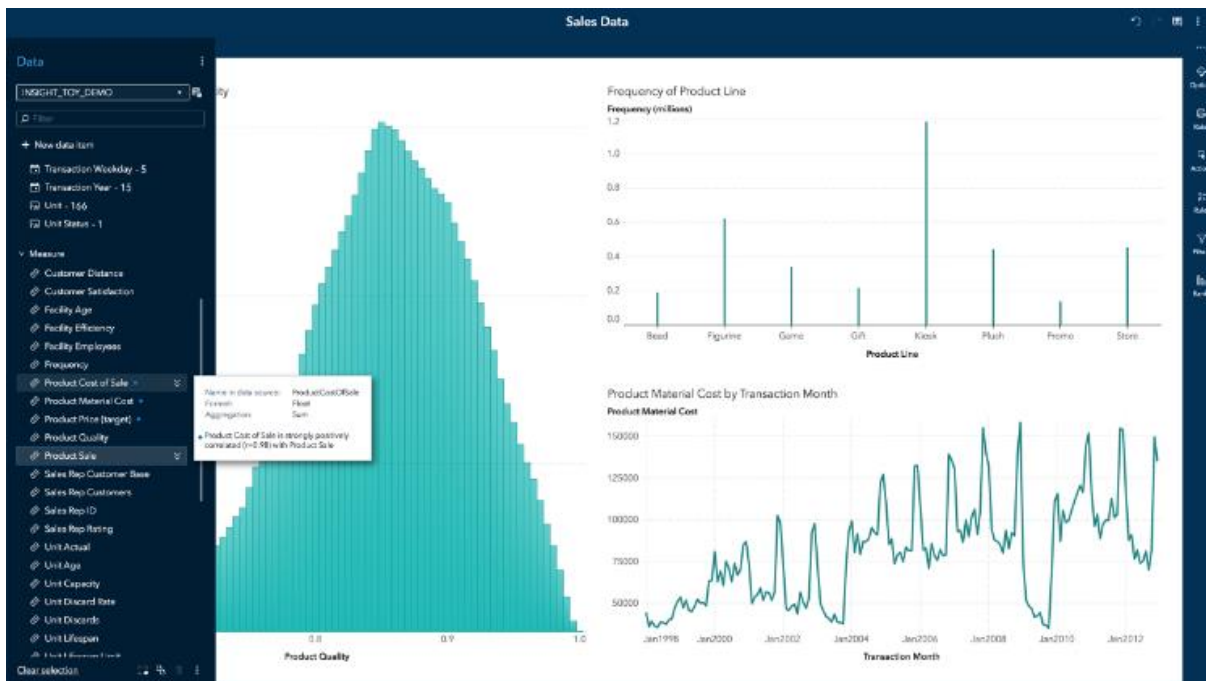


Figure 9 SAS Visual Analytics 2

FORECASTING THE FUTURE!

What about if you have data that represents a historical pattern, such as a sales history for your latest consumer product or block buster drug. SAS can help by providing not just visualizations based on the data you provide but insight into possible future behavior using accurate statistical forecasts and multiple advanced forecasting techniques that can taking into account a whole range of factors such as seasonal trends, product promotions, weather event and more.

The example below is based on the sale of millions of products from a product range by region over a 3-month period with a predicted sales forecast derived using a range of statistical forecasting and advanced machine learning capability. These are initially chosen by the system to provide the most accurate results and can be tuned by the use based on their specific knowledge. This is an example of “Human-In-The-Loop” analytics where the use of analytics and human knowledge are combined to add value and insight to the data. While this visualization provides a way to ‘see’ the impact of any changes the real power comes from the use of SAS’s advanced analytics to analysis the vast amount of data quickly so the users can make informed decisions based on the ‘actionable insight’ provided. All of this advanced insight is also available in just a few mouse clicks.

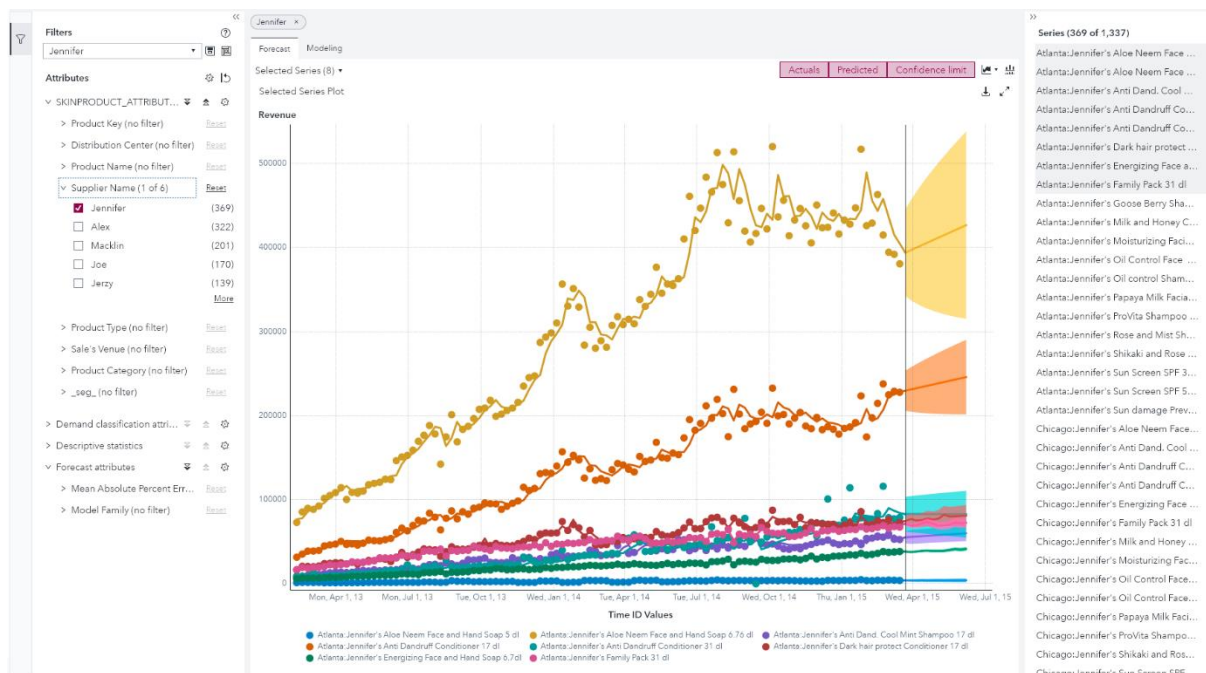


Figure 10 SAS Visual Forecasting

UNDERSTANDING RELATIONSHIPS

We all know that data items can be correlated but how can you assess the level of correlation when you have millions or billions of data points?

That's where the use of correlation matrices and cluster plots, two of the visualizations available in SAS Visual Statistics, can help. The first example below shows a correlation matrix with the darker rectangles indicating a stronger correlation between two columns in the data and the lighter ones indicate a weaker correlation. The second example shows the use of cluster plot which allows multiple relationships in potentially millions or billions of data points to be studied.

Both of these techniques and visualization can again be created in just a few mouse clicks as part of a data exploration exercise during the analysis of data. These techniques form the basis for our work on helping customer understand the relationship in the data to help eliminate bias when using data sources as input to advance AI and ML techniques and help companies make critical decisions in multiple area including finance (are you eligible for a loan or not) , healthcare (are you a good candidate for a particular treatment) and increasingly in Life Science companies (where should the initial launch of our latest treatment be to deliver maximum value to the patients who need it most). If you would like to know more on how SAS can help deliver responsible AI the please take a look here [Responsible Innovation | SAS](#)

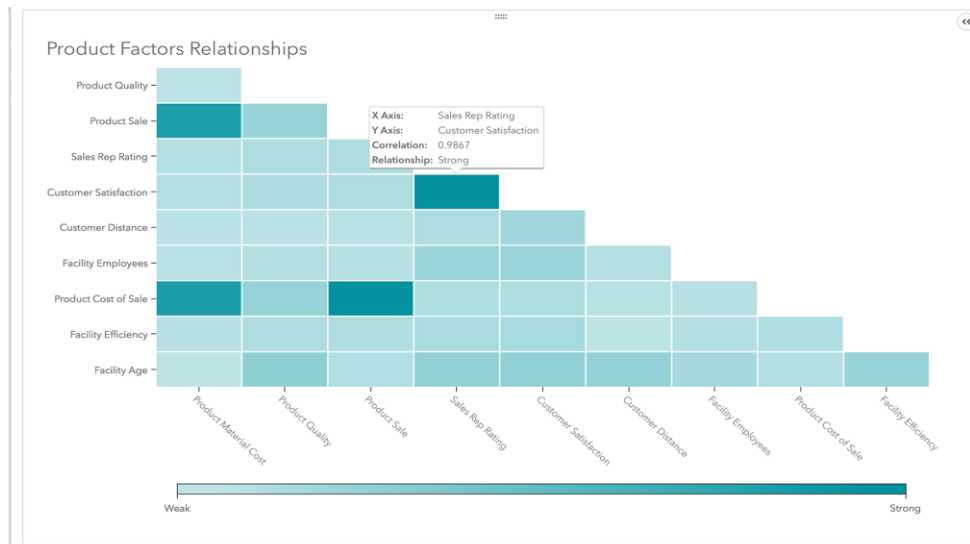


Figure 11 SAS Visual Statistics - Correlation

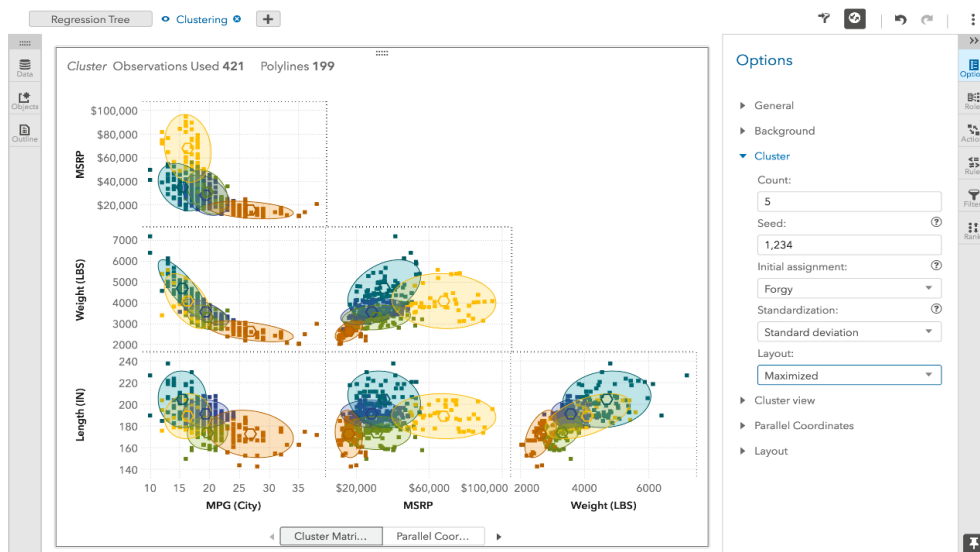


Figure 12 SAS Visual Statistics - Clustering

VISUALIZATION TO HELP UNDERSTAND MACHINE LEARNING AND ARTIFICIAL INTELLIGENT MODELS.

In the previous section we discussed the use of some of the SAS data visualization techniques to help explore and understand millions or billions of data points before they are used in as input for machine learning or artificial intelligence driven modelling. But SAS also provide a vast array of data visualizations that can help explain what a machine learning or artificial intelligence model is doing. While SAS always produces white box models based on SAS code, which can be reviewed by a programmer if required, we know that 'a picture paints a thousand words' and can help explain complex modelling in easy-to-understand visualizations.

The following examples are from our SAS Visual Data Mining and Machine Learning capability, available on the SAS Viya platform I mentioned earlier, and show a visualization of a Decision Tree which can be used to explain data relationship that drive a particular outcome or behavior. Decision Trees can be used as a technique to model data themselves but are also frequently used to help non-technical users understand the outcome of more complex machine learning and artificial intelligence driven techniques.

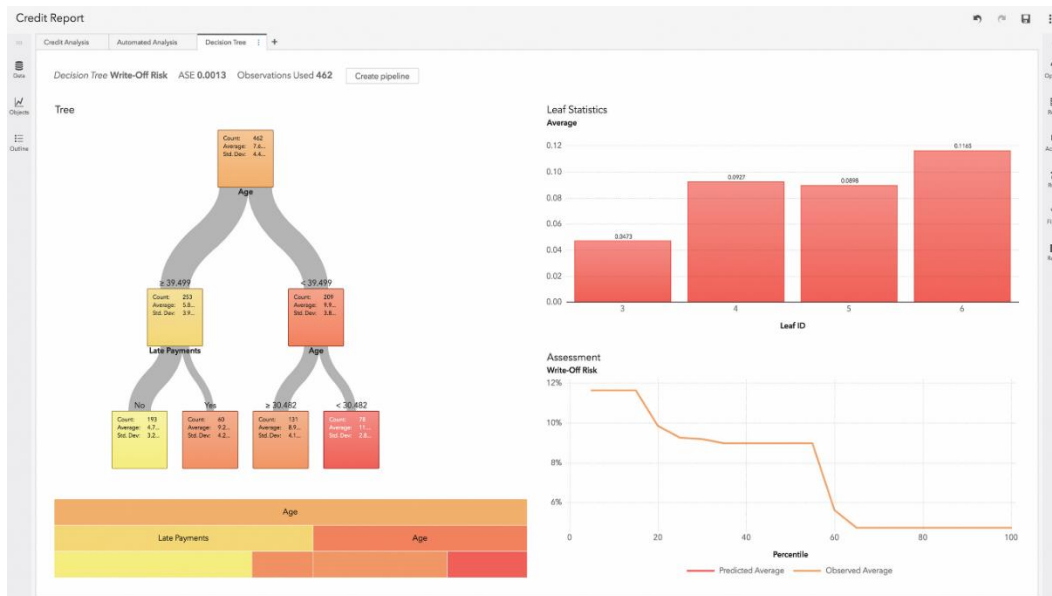


Figure 13 SAS Decision Tree

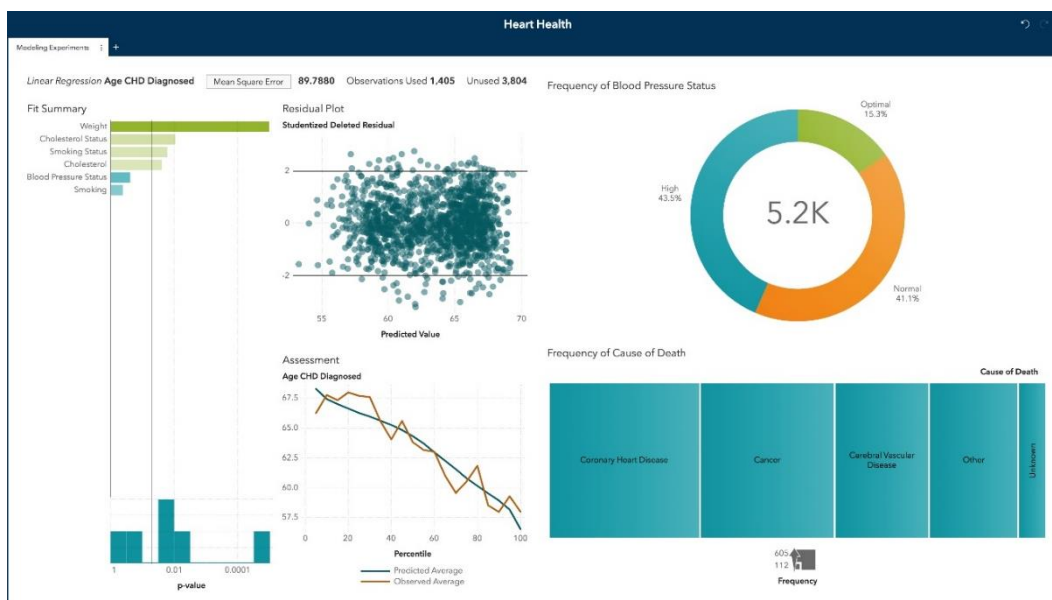


Figure 14 SAS Machine Learning

VISUALIZATION AND UNSTRUCTURED DATA

A vast amount of data is not available as structured data in nice rows and columns but comes as unstructured data in text documents and images. In an article in Baseline Magazine from 2023 it's estimated that "around 80 percent of the data available is in an unstructured form such as text, video files, audio files, images and social media posts" (5) and while some of these are already 'pictures' or in the case of video collections of pictures presents in a sequence, they can also be viewed as digital information and can be processed using SAS's advanced analytics to provide additional insight. I won't go into the use of images and video with SAS as that's a whole paper in itself, but I've personally worked on several projects which take images and video as a source and use SAS Analytics and Visualisation to help explain and detect issues and then provide additional actionable insight using some of the SAS Visualisation techniques I was already discussed above.

If we look just a textual information available in documents, web site and social media, as an example then by using of SAS Text Analytics we can extract valuable information from these which can then be shown visually. The example below shows how a large set of complaints data, captured via on on-line systems, can be automatically summarized and categorized to quickly provide actionable insight which would have taken may hours or days if an individual were to read all of these.

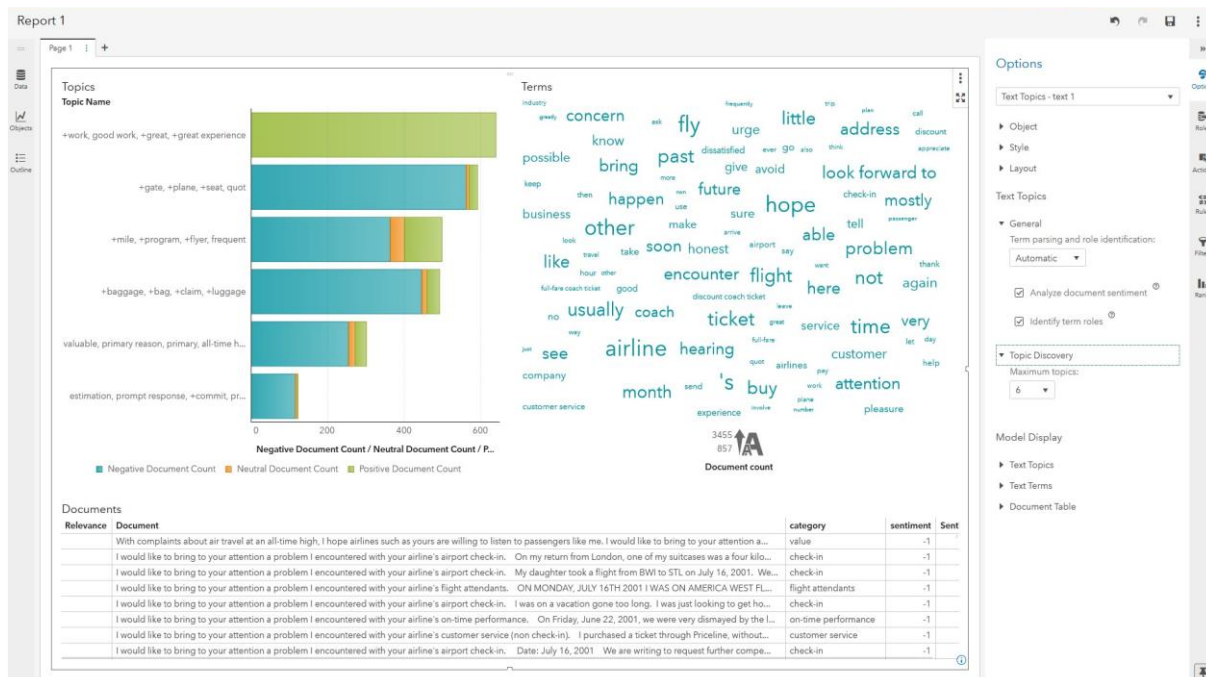


Figure 15 SAS Text Analysis

UNCOVERING HIDDEN NETWORK

As a final stop on this tour of the SAS visualization that can help you ‘Get the Full Picture Of Your Data’ I’ve included an example of a network diagram which can be used to show relationships between items and individuals. Typically uses include the identification of individuals associated with an insurance claim or how individuals in a criminal investigation are linked, or how the complex bill of materials for a product of drug are related. With insurance data the datapoints come from a wide variety of source such as claims data, policy data, social media, and 3rd party sources and the powerful network analytics available with in SAS can easily process all these data types, typically outperforming the network databases normally associated with this type of analysis, and is able to display the results as a fully interactive network visualization allowing user to explore the data at will, follow a training-of-thought investigation, and allowing them to gain a unique actional insights into the complex data in seconds.

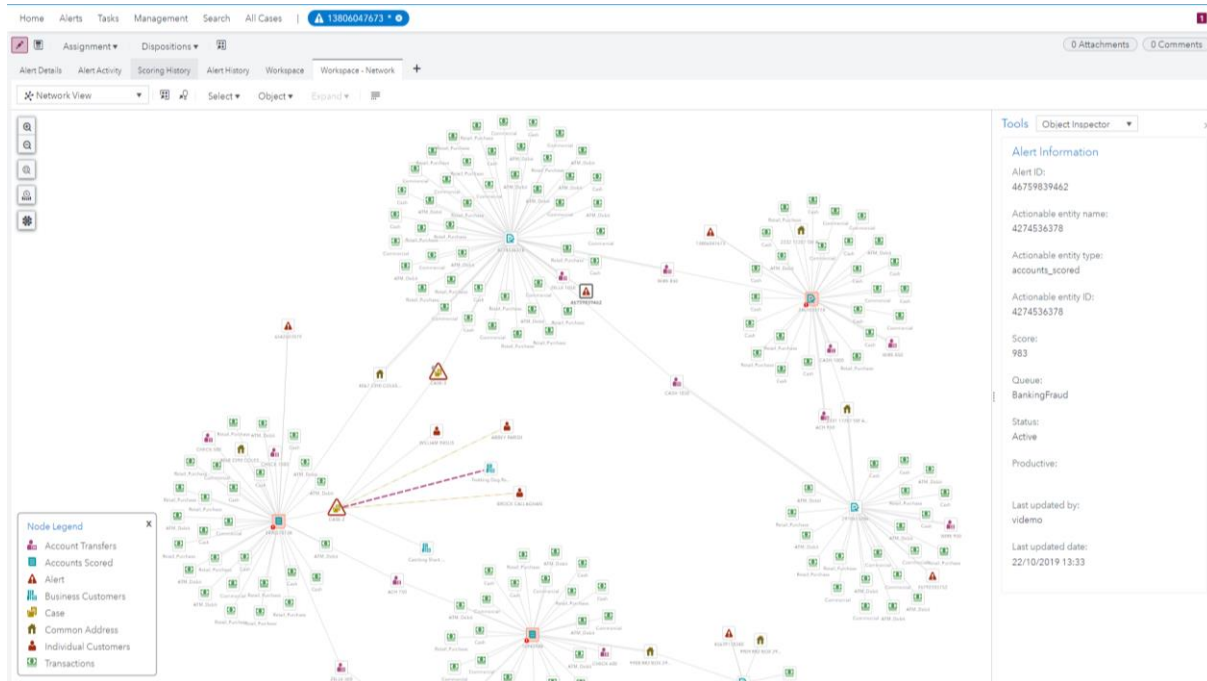


Figure 16 SAS Network Analytics

PLAYING NICELY IN THE VISUAL WORLD

So far I focused on the visualizations available within the SAS environment, but sometimes even SAS doesn't have just the right visualization for you need for you data and someone may have developed the perfect visualization for your data in an open source D3.js object [\(6\)](#). I'm not going to go into details on how this is set-up here as there are lots of tutorials and samples on the SAS Git Hub Page [\(7\)](#) and [\(8\)](#) so I'll just show you some example of diagrams you can populate with the data held in SAS [\(9\)](#).

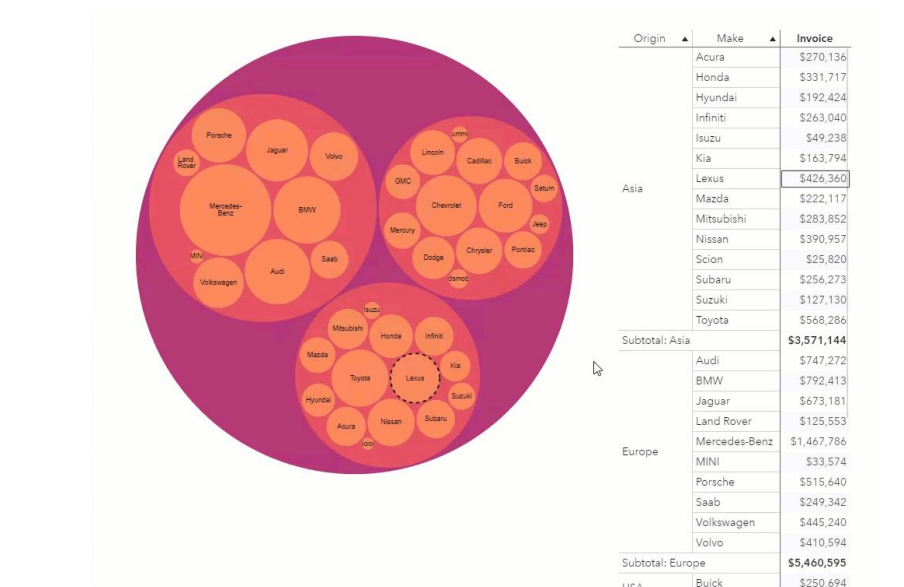


Figure 17 D3 Circle Packing Diagram

Process_Step_Name	Website_Session_ID (Distinct count)	Process Step
View Webinar Description	65	1
Select Webinar	26	2
Register	20	3
Start Webinar	9	4
Finish Webinar	7	5



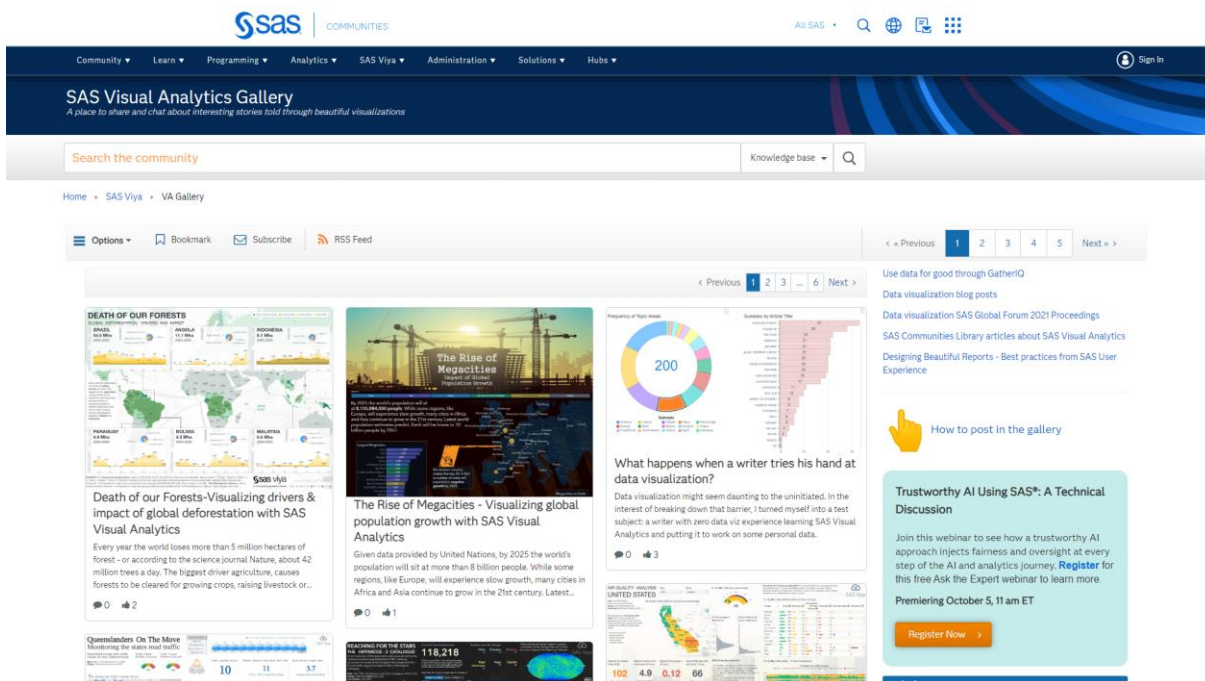
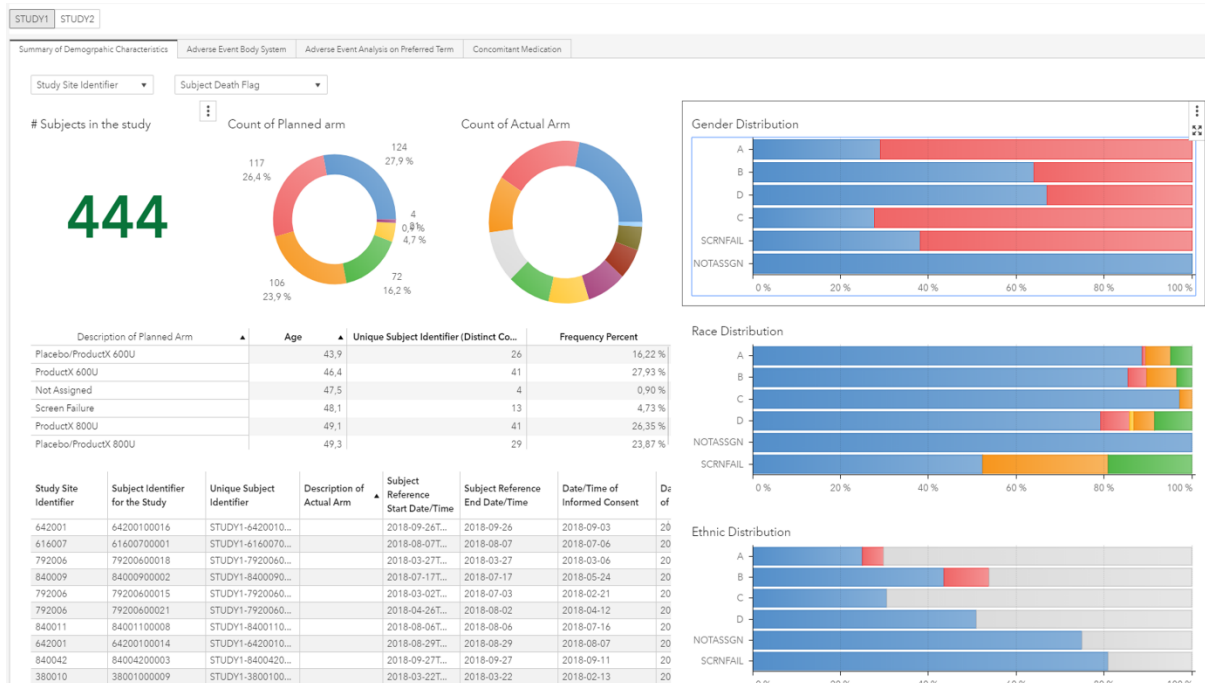
Figure 18 D3 Funnel Diagram

CONCLUSION

As conclusion I hope by now that we can see how SAS can be used to provide actional insight from billions of data points so you can **Get the Full Picture With SAS**.

If you would like to see more examples of how these powerful visualizations can be combined into fully interactive dashboards for executives and end users, using just a few mouse clicks in SAS Viya then take a look these

- [SAS Visual Analytics Gallery](#)
- [SAS YouTube Channel](#)



REFERENCES

1. [A picture is worth a thousand words - Wikipedia](#)
2. [Mona Lisa - Wikipedia](#)
3. [SAS Help Center: SAS/GRAPH 9.4: Reference, Fifth Edition](#)
4. [SAS Viya: The High-Performance AI & Analytics Platform | SAS UK](#)
5. [Structured Vs. Unstructured Data Including Examples of Both \(baselinemag.com\)](#)
6. [D3 by Observable | The JavaScript library for bespoke data visualization \(d3js.org\)](#)
7. [SAS Software · GitHub](#)
8. <https://github.com/sassoftware/sas-visualanalytics-thirdpartyvisualizations>
9. [Using samples within SAS Visual Analytics | sas-visualanalytics-thirdpartyvisualizations \(sassoftware.github.io\)](#)

CONTACT INFORMATION

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

Andy Bayliss
SAS
Wittington House
Marlow / SL7 2EB
Work Phone: +44 07818 046866
Email: andy.bayliss@sas.com
Web: www.sas.com

Brand and product names are trademarks of their respective companies.