saasabi

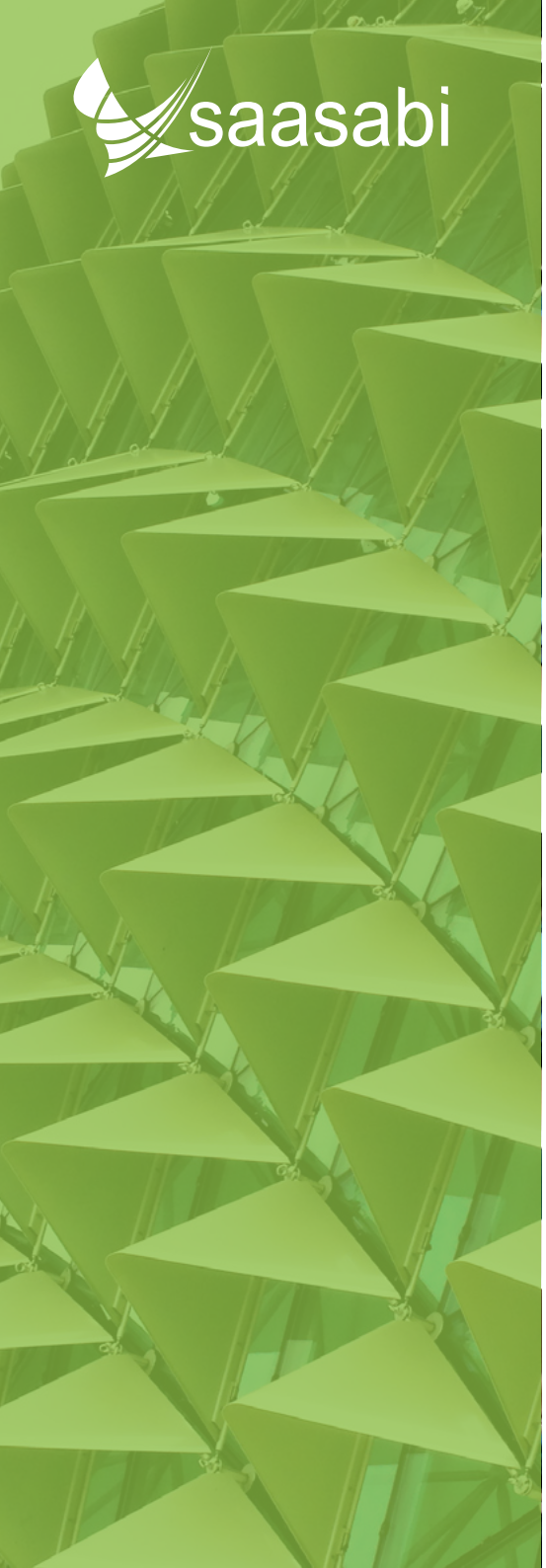
Saasabi White Paper

# Big Data Analytics at SMB Scale

## 4 Best Practices



"Big data" has been one of the most important catchphrases in the technology world over the past decade. What, exactly, constitutes big data? Is it relevant to small and medium-sized organisations? (Spoiler alert: Yes, it is.) How do SMBs take advantage of the new software tools that have emerged over the past few years to unlock the value of their data?



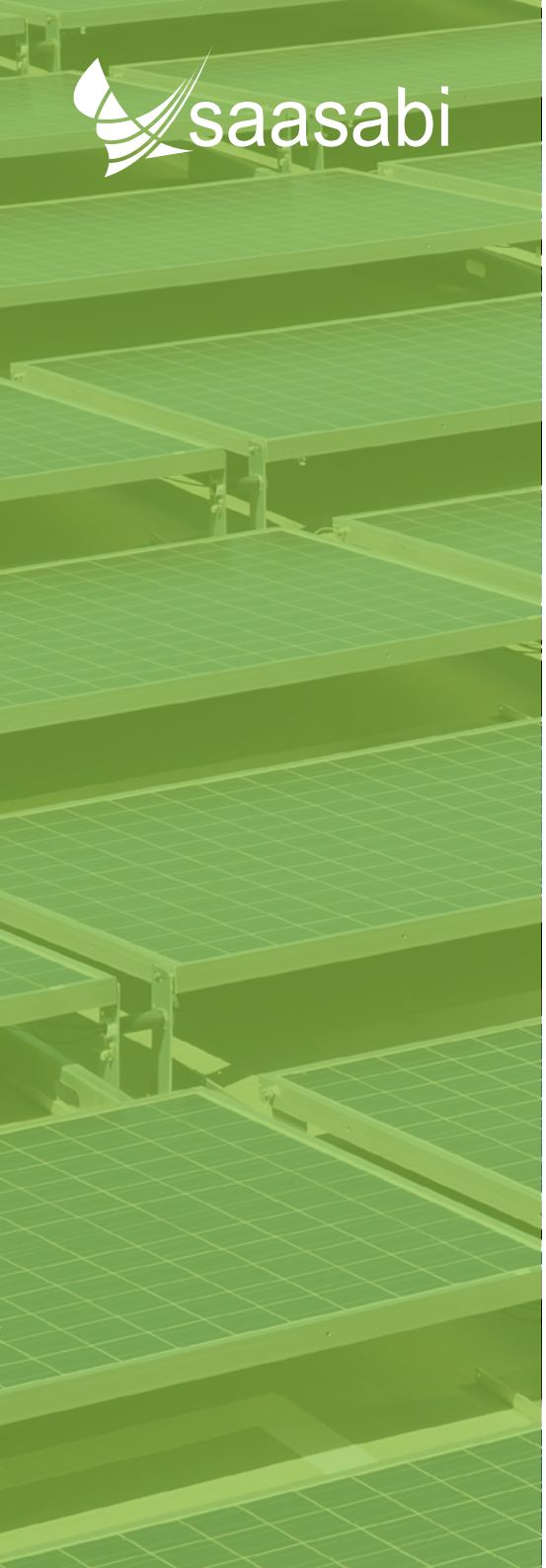


Let's start at the very beginning. What is big data? If you Google the term, you'll come up with a variety of definitions. Some people apply the term "big data" only to unstructured data. A good example of unstructured data is the raw content (in any format – text, photos or videos) that comes from customers' interaction with a business via social media platforms like Twitter and Facebook. Others define big data in terms of four V's: volume, velocity, variety and veracity.

**Volume** puts the "big" in big data. It is the amount of data in your organisation. The amount of digital content in the world is currently estimated to double every 18 months.[1] According to IBM, every single day we create 2.5 quintillion bytes of data[2] (for those who are not used to dealing with numbers of that size, that's 2.5 followed by 18 zeros!). SINTEF, an independent Norwegian research organisation, reports that the exponential growth of data means that 90% of the data that exists in the world today has been created in just the last two years.[3]

There is no specific threshold at which regular-sized data suddenly gets promoted to big data. For some organisations, the total volume of data they're managing might only be gigabytes, yet other qualities about the data qualify it as big data. For others, the amount of data could reach into the petabytes. That is, without a doubt, big data.

**Velocity** is the frequency of data updates. To wrap your head around the concept, think in terms of the number of credit card transactions processed each minute of each day by a single large online retailer such as Amazon, and then scale that up across all online and offline retailers across the globe. Or think of the number of Facebook status updates posted by your friends each day and multiply that across the social platform's nearly 1 billion active daily users.





**Variety** refers to the different types of data an organisation encounters. As mentioned above, this can include both structured and unstructured data. Structured data has explicitly defined properties. For example, you can define a bank balance as piece of data that will always be expressed numerically with a certain number of decimal places and that will be associated with a designated currency. Unstructured data, on the other hand, has no rules. It can include both text and multimedia content. Examples of unstructured data abound in our everyday lives. They include emails, Instagram photos, security video, audio files, PowerPoint presentations, etc. All of these digital files may have metadata associated with them (such as the time of creation, the location where the information was captured, the author, etc.), and they may have their own internal structure, but there are no rules that apply across all of these different types of content. Experts estimate that as much as 90% of the data in any organisation is unstructured.

**Veracity** is the final V. It addresses the quality of the data. Is the data accurate, complete, precise, representative, consistent and fresh? If not, your data analysis project won't get far. We will come back to the topic of data quality below.

No matter how you define big data, one thing is for certain: organisations – even the smallest ones – have access to more data than ever before. They are both generating data themselves, and they also can tap into huge publicly-available data sets that are available online. These include demographic information, traffic patterns, weather patterns, consumer purchase behavior and more.

It's easy to see how some might find this sea of data overwhelming. If that's the way you feel, you're not alone. Managing and harnessing data to drive better decision-making is as much a challenge for huge, multinational companies as it is for small businesses with only a few employees, though the scale of the challenge increases as the size of the organisation grows.



The good news? **SMBs are well positioned to take advantage of the big data revolution.** Not only are they smaller and therefore more nimble than their lumbering enterprise cousins, they can take advantage of a host of enterprise-grade, cloud-based, data-analysis tools that have been created in recent years. Just as technology infrastructure is available as a service through providers such as Microsoft, Amazon and Google, many of today's data analytics solutions are delivered as software as a service (or SaaS ) and give SMBs access to the same sophisticated analysis tools that were once only available to the largest organisations. And, of course, the price for storing today's increased data volumes has been drastically reduced over the past decade making it far more viable for any size organisation to hold onto its data.

The key to diving into big data is, paradoxically, to start small. Don't get paralysed by the infinite possibilities. Define an initial project scope that is both finite and manageable. Determine which business challenges you'll address first. Tackle them, learn from the experience and expand the scope of your data analysis projects as you go. As the tennis great Arthur Ashe once said, "Start where you are. Use what you have. Do what you can." Words to live by.

**Here are a few best practices you can follow to maximize the value of your organisation's data:**



## Best Practice #1: Identify the sources of data in your organisation

Today's SMBs have access to a variety of on-premise and cloud-based business systems that generate actionable data. These can include:

- Accounting packages
- Customer relationship management (CRM) tools
- Web site analytics
- HR and payroll systems
- Customer service platforms
- Manufacturing systems
- Social media platforms
- Job costing and management systems
- GPS tracking systems

Do not forget to include the less sophisticated systems (think Excel spreadsheets, email, historical transaction data, mailing lists, bank statements, etc.) used by your organisation.



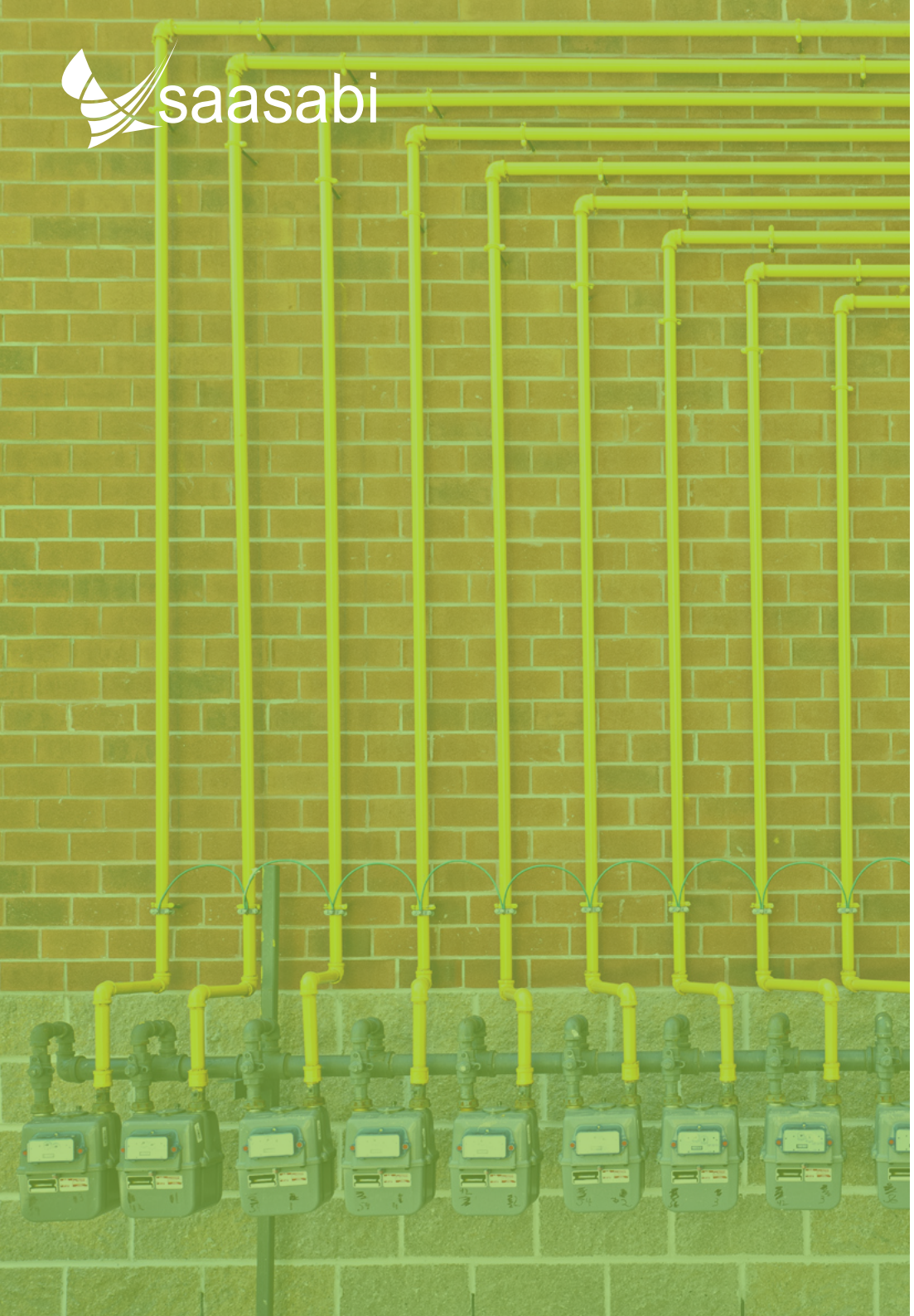
## Best Practice #2: Eliminate manual processes

In order for your organisation to have systems that are scalable and data that is current, avoid manual processes. Instead of tracking customer service interactions or software bugs via spreadsheets, invest in software such as ZenDesk, Groove, Kana or Atlassian that will help your organisation manage those processes. Likewise, move to an Internet-based telephone system so that all of your call activity can be easily tracked and measured. Talk to your customers and suppliers and see what data they can provide you electronically. This information can then be automatically added to your store of useful data.



## Best Practice #3: Institute data hygiene/ data governance policies

To become a data-driven organisation, you have to get your data in order. Everyone is familiar with the old adage “garbage in, garbage out.” It only stands to reason that your ability to gain useful insights from your data is dependent on the quality of that data. For instance, if you’re interested in finding out which regions are showing the most sales growth for a certain product line, you’ll need to make sure that your CRM system contains accurate information on your clients’ locations. Simply standardising client names in your CRM system (so that GE, General Electric and General Electric LLC refer to the same account) is a good start. There are many companies that can automate data validation and standardisation processes. For example, products like QFire Software are available as affordable monthly subscriptions and can be integrated with your CRM system to automate data quality improvements with your customer data.





## Best Practice #4: Determine which questions you need answered and where the data to answer those questions resides

Once you have most, if not all, of your processes automated and your data is cleaned up, you can begin to query it. The dashboards and reporting tools built into the software your organisation uses may already provide the answers to some of your initial business questions. However, one of the biggest challenges faced by organisations large and small is analysing data that resides within separate systems that are isolated from one another. For example, you can clearly see which sales rep has the highest volume of sales, but if your CRM system doesn't tie into your job costing system and your customer service systems, you may not know which sales rep brings in the most ***profitable*** clients. Or perhaps you'd like to know how weather patterns impact foot traffic and sales volume in your retail locations. In order to figure that out, you will need to integrate an external data source on local weather patterns with the sales data your organisation is generating internally. Again, there's good news here. Many affordable, cloud-based business intelligence software solutions enable you to pull in data from a wide variety of sources in order to perform analysis across the united data set. Another important point here is that the use of these systems today is vastly different from days of old where a specialist business analyst was needed to understand what was going on. Many modern solutions provide self-service tutorials to help you connect different data sources and determine how best to use that data to deliver value to business users quickly and easily.



## In Conclusion

Big data is now a fact of life for organisations of all sizes. With the right tools, it can be used to drive success for your SMB. The latest cloud-based solutions enable big data analytics at SMB scale – and budgets! There are thousands of free public data sets, and there are a number of very powerful and low-cost or free data visualisation tools – software such as Microsoft Excel, Microsoft Power BI and Tableau Public. Affordable SaaS solutions are available to support data quality initiatives, the unification of siloed data sets for integrated analysis, and optimized customer communications based on the insights you uncover.

The big data revolution is, in fact, a huge boon for SMBs because it has catalysed the development of these many data preparation, data analysis, data visualisation and data-driven marketing tools. So go ahead and identify your first big data project. Just remember to start slowly and build on your success. This may be the beginning of a much more profitable, data-driven era for your small business.

### Sources:

1. IBM, Cisco, Frost & Sullivan Analysis ; http://www.investinbsr.com/ipaforum/wp-content/uploads/Iain-Jawad-IPA-Forum-2014-Presentation.pdf

2. IBM; http://www.datamation.com/applications/big-data-analytics-overview.html

3. SINTEF; http://www.sintef.no/en/news/big-data--for-better-or-worse/

Saasabi delivers scalable, cloud-based business intelligence and analytics solutions that enable organisations to make better, data-driven decisions. Saasabi allows data from multiple, disparate sources to be combined and structured into a format that is easily accessed for visualisation and sharing.

**Go to www.saasabi.com to learn more**