

TDWI RESEARCH

## TDWI CHECKLIST REPORT

# Moving Beyond Spreadsheets

Six Steps for Enabling Small and Midsize Businesses to Gain Data Insights

By David Stodder

Sponsored by:



[tdwi.org](http://tdwi.org)



AUGUST 2014

TDWI CHECKLIST REPORT

# MOVING BEYOND SPREADSHEETS

## Six Steps for Enabling Small and Midsize Businesses to Gain Data Insights

By David Stodder



555 S Renton Village Place, Ste. 700  
Renton, WA 98057-3295

**T** 425.277.9126  
**F** 425.687.2842  
**E** info@tdwi.org

tdwi.org

### TABLE OF CONTENTS

- 2 **FOREWORD**
- 2 **NUMBER ONE**  
Improve user productivity by overcoming the limitations of spreadsheets
- 3 **NUMBER TWO**  
Increase user confidence in data by establishing shared governance
- 3 **NUMBER THREE**  
Evaluate technologies that enable users to interact effectively with data
- 4 **NUMBER FOUR**  
Tap cloud computing options to support expansion of BI and analytics
- 4 **NUMBER FIVE**  
Gain uncommon insights from big data to create competitive advantages
- 5 **NUMBER SIX**  
Evaluate the potential benefits of implementing predictive analytics
- 6 **ABOUT OUR SPONSOR**
- 6 **ABOUT THE AUTHOR**
- 6 **ABOUT TDWI RESEARCH**
- 6 **ABOUT THE TDWI CHECKLIST REPORT SERIES**

© 2014 by TDWI (The Data Warehousing Institute™), a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. E-mail requests or feedback to info@tdwi.org. Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

### FOREWORD

Technology trends are moving in a favorable direction for small and midsize businesses (SMBs) seeking to improve their prowess in business intelligence (BI) and analytics. Tools are becoming easier to use for nontechnical users—that is, business users who may know their data but have neither the training nor the skills to write queries or provision their data. SMBs are generally defined as organizations with fewer than 1,000 employees and less than \$1 billion in revenue. Most SMBs do not have the large and experienced IT functions found in most large organizations to manage data, build data warehouses, and develop BI and analytics applications. Thus, the evolution of BI and analytics tools and platforms toward greater ease of use and deployment has the potential to level the playing field for SMBs as they compete with larger organizations.

Currently, the most prevalent tools in use within SMBs for accessing, analyzing, reporting, and presenting data are spreadsheet applications and reporting tools embedded in point solutions for finance, sales, or other operations and business management. However, with analytics becoming a competitive differentiator for SMBs, leadership in these organizations must determine whether users' current tools and platforms for data access, analysis, presentation, and management are adequate for meeting data-driven decision-making needs. Without the ability to interact with data effectively, SMB users are at a disadvantage, especially as business growth, change, and regulatory duties increase the size of (often scattered) data resources and the complexity of analytical demands.

Although some SMBs came into being with BI and analytics as part of their DNA, many firms are only now realizing that they need to reach higher levels of maturity in how they access, analyze, and share data. They are making plans to upgrade their capabilities. It can't be done all at once; the best approach is to tackle goals gradually, keeping the deployment of new technologies and methods aligned with business objectives. To ensure alignment, enterprises need good communication and collaboration among business and IT leadership.

This TDWI Checklist Report details six areas where SMBs should focus as they move beyond spreadsheets to improve capabilities for gaining data insights.



### NUMBER ONE

#### IMPROVE USER PRODUCTIVITY BY OVERCOMING THE LIMITATIONS OF SPREADSHEETS

Many SMB users begin their trek into data analysis and presentation with spreadsheet applications such as Microsoft Excel or Google Spreadsheets. It is not hard to understand why. Spreadsheets have long been a component of personal productivity applications and thus do not require additional software licenses. A recent TDWI Research survey<sup>1</sup> found that a majority of users employ spreadsheets rather than specialized BI or analytics tools as their primary tool for data access, analysis, and presentation. Acknowledging that spreadsheet use is unlikely to go away, most BI and analytics tool providers have improved integration with spreadsheets by making it easier to move data from BI and data warehousing systems to spreadsheets or from spreadsheets into BI and analytics tools.

However, as SMBs grow in number of users, lines of business, products and services offered, and sales channels, many run into significant difficulties if they depend on spreadsheets for data reporting, analysis, and presentation. As the business grows, spreadsheets become bigger and more numerous. Users pass them around, making their own copies and manually importing or cutting and pasting new data into them. Erroneous data gets baked into users' analysis. Eventually, disputes arise over who has the correct data and users lose trust in the analysis.

The biggest hit from this chaos is poor productivity. Spreadsheets become unwieldy documents that may feature hundreds of worksheet tabs and custom, manually created formulas for tracking budgets, financial performance, and other aspects of business operations. To perform more advanced analysis usually requires custom coding, which calls for skills that only a few power users have, making organizations vulnerable if those users leave the enterprise. Because users often do not share techniques or best practices, there can be considerable inconsistency and duplication of effort. Rather than collaborating effectively, users lose time just trying to understand what's in each spreadsheet and how to work with each spreadsheet's formulas.

Organizations should evaluate how they can improve the productivity of individual users as well as departments and divisions by moving beyond dependence on spreadsheets. SMBs should consider the business value of deploying BI and data management systems that can scale up to more users, more data, and more complex analytic needs.

<sup>1</sup> David Stodder [2014]. *TDWI Best Practices Report: Business-Driven Business Intelligence and Analytics*.

### NUMBER TWO

#### INCREASE USER CONFIDENCE IN DATA BY ESTABLISHING SHARED GOVERNANCE

The productivity problems that result from an overdependence on spreadsheets are related to a larger challenge facing SMBs as well as many larger organizations: lack of confidence in the data. Business growth and diversification often create numerous data silos that are difficult to integrate, and not just for technical reasons. Divisions, departments, and lines of business feel a sense of ownership of their data. Despite data quality problems that exist within their own silos, users in each group feel like they know and can therefore trust their data. They are reluctant to use other groups' data or consolidate their sources into a larger, shared resource.

To succeed in turning data into the valuable shared asset that supports informed strategic and operational decision making, SMBs need to build confidence in their data. Success with this objective requires addressing both people issues as well as technology and data management issues. Here are two initiatives that should be a priority:

- **Create a center of excellence to improve collaboration.** To improve shared leadership for BI, analytics, data integration, and other projects for generating value from data, organizations should create a dedicated center of excellence or competency center. Business leadership from across the organization should be represented along with IT, data management, and development. This group can articulate the business value of shared BI and data assets, including for regulatory data governance. This group can develop road maps for establishing companywide data access and directing consolidation of unnecessary data silos.
- **Define best practices for improving data quality.** Data quality and consistency are a constant challenge. Organizations must address the “garbage in, garbage out” problem caused by imperfect data entry at the source, sometimes due to application processes that force users to enter bad data to complete a task. Organizations can also employ data quality tools to help them remedy errors as they integrate and consolidate sources before bad data becomes part of users' BI, reporting, and analysis.

At this step, organizations should bring business and IT stakeholders together to provide visible, joint leadership for improving data quality, management, and governance. This will encourage greater user confidence in shared data assets.

### NUMBER THREE

#### EVALUATE TECHNOLOGIES THAT ENABLE USERS TO INTERACT EFFECTIVELY WITH DATA

SMBs will have an advantage over their competition if they can provide users with access to more data sources and give them greater confidence in the data by establishing leadership to resolve quality and consistency issues. However, users need more than just data. They need workspaces that enable them to interact with data effectively—for example, to compose queries, create reports, define metrics, explore data for analytic insights, and create and share visualizations. Spreadsheets confine most users to static reports and few visualization choices, and offer limited modes of data interaction. Small and midsize organizations should evaluate the potential benefits of BI and analytics tools for giving users a fuller range of interaction.

Even in large organizations, IT developers are stretched thin trying to respond to diverse user requirements for data interaction. With typically even less IT or BI developer resources at their disposal, many SMBs want technology solutions that enable users to drive data access and analysis themselves, without handholding from IT. The technology market has been responding with “self-service” BI, analytics, and data discovery tools that provide workspaces in which nontechnical users can go beyond canned reports to slice and dice data, perform exploratory discovery analysis, and create dashboards.

Some solutions offer fuller analytic platforms that employ in-memory computing to avoid having to fetch data from disk for every query. This enables users to work with much larger data sets on their own. Analytic platforms that feature in-memory computing can give users near-real-time access to data, which is better than spreadsheet-based systems can provide.

The self-service trend is fortuitous for users, in many cases enabling them to avoid long IT development cycles and produce at least basic, interactive data visualizations and dashboards. Users still need guidance, however; organizations should ensure proper training and provide professional guidance to get users off to a good start. Organizations should also evaluate whether their underlying data management and data platform infrastructure is prepared to handle the performance requirements of widespread data interaction, and whether having in-memory computing as a central architectural component could improve performance for data access and analysis.

### NUMBER FOUR

#### TAP CLOUD COMPUTING OPTIONS TO SUPPORT EXPANSION OF BI AND ANALYTICS

Cloud computing and software-as-a-service (SaaS) for BI, analytics, and data warehousing are maturing, giving SMBs broader options for meeting business needs. It is no longer axiomatic that SMBs are at a disadvantage against bigger competitors who have larger and more established IT functions. Cloud computing can enable SMBs to overcome their lack of on-premises IT expertise and infrastructure. They can “get big fast” by spinning up cloud BI and analytics services to meet dynamic needs.

Cloud computing is an important consideration for SMBs that view analytics as a core part of their business model. Internet-based firms that depend on providing services, for example, cannot wait until they have built their own on-premises data management infrastructure to analyze customer behavior data, Web logs, and other sources for insight into preferences and buying patterns. From day one, they need to innovate based on their knowledge of customers. With cloud options, firms can focus on analytics development rather than on configuring their data infrastructure.

Some SMBs are deploying cloud-based BI, analytics, and data management to serve users who need to interact with data from mobile devices. Organizations choosing this strategy will often provision users by loading data periodically into cloud-based systems from internal data warehouses or other business applications. IT personnel can manage data quality and address governance requirements before the data is loaded into the cloud.

Concerns about data security and privacy make many firms reticent to put data on public cloud-based systems; private clouds, which offer hosted services behind a firewall, are less of a concern. Organizations should review regulatory and governance policies before putting sensitive information outside firewalls.

Because few organizations will put all of their data on cloud-based systems, SMBs should develop comprehensive (sometimes called “hybrid”) data architectures that guide which data should be put on which platform. Along with a good understanding of business needs, factors to consider include how frequently data needs to be updated and the complexity and depth of analysis (including whether the analytics will require many passes through the data). Understanding these requirements will help organizations match user needs with platforms that can deliver the appropriate performance and availability.

### NUMBER FIVE

#### GAIN UNCOMMON INSIGHTS FROM BIG DATA TO CREATE COMPETITIVE ADVANTAGES

Spreadsheets and conventional BI applications deal mostly with data from financial management and transaction-oriented business applications. This data is usually highly structured and accords with data models built to handle expected data properties and relationships. Today, however, many organizations want to explore semi-structured and unstructured data to discover uncommon insights. Examples include clickstreams and page view records that capture online customer behavior, text from Web logs and social media comments captured across the Internet, and machine data generated by sensors. Structured transactional data can also be very large, but semi-structured and unstructured data sets can be truly enormous, leading to their popular designation as “big data.”

Users who work exclusively with spreadsheets and conventional BI tools cannot easily interact with big data because their applications are intended for structured data and employ rigid data models and schema that have been defined before the data enters the system. The desire to get beyond these limitations to explore other types of data has been a major driver behind the deployment of Hadoop, MapReduce, and related big data technologies. Organizations are implementing Hadoop files to store enormous “data lakes” that are not restricted to particular types of data or a priori schemas. Organizations can develop analytics to derive insights from this raw, freeform data, or they might choose to cleanse and transform it for loading into more structured data warehouses if users deem it important for ongoing BI reporting, analysis, and presentation.

Big data can play an important role for SMBs seeking new insights into customer behavior, social media sentiment, and market trends, or to correlate business performance with external developments impacting their business environment (such as news or weather). To improve operational efficiency, some SMBs may collect and analyze machine data for logistics or track the movement of inventory through supply chains.

Organizations should evaluate whether business objectives, particularly in marketing and customer engagement, could be furthered by pushing beyond structured data in spreadsheets and BI applications into the realm of big data. Organizations should consider cloud computing and online data services as alternatives to building on-premises expertise and technology platforms.



### NUMBER SIX

#### EVALUATE THE POTENTIAL BENEFITS OF IMPLEMENTING PREDICTIVE ANALYTICS

Predictive analytics methods and technologies enable organizations to examine data trends to explore what may happen next. Predictive analytics is aimed at getting organizations to insight that goes beyond reporting on historical data (what has happened), or even forecasting in spreadsheets or online analytical processing (OLAP) applications that base analysis on a defined set of “known knows.” Using models, predictive analytics helps organizations look into the unknown by using test-and-learn discovery processes. Organizations can gain insight into the likelihood of certain outcomes based on elements such as the performance of a model’s numerous variables and measurements of reactions to stimuli.

One of the biggest advantages of predictive analytics is reducing time to insight. Many organizations like to score predictive models against live data, which could include transactions, data streams, social media feeds, or other big data sources we’ve described. Whether using algorithms or additional software automation to score models against live or historical data, predictive analytics has the power to deliver insights that traditionally could take weeks or months of analysis to uncover. Marketing functions are major users of predictive analytics for sharpening marketing campaigns and improving the personalization of offers. Other major applications include fraud detection, resource allocation, risk management, and financial modeling.

SMBs can benefit from predictive analytics by using it to anticipate outcomes and therefore be proactive in how they prepare resources, such as customer service, if certain outcomes are expected. Gaining insights sooner can help organizations avoid unnecessary costs and reduce exposure to negative events such as fraud. SMBs have less room for error than larger organizations and must run operations as efficiently as possible; predictive analytics can deliver insights that will help SMBs be more efficient and effective.

SMBs should evaluate predictive analytics to determine where methods and technologies can offer the greatest business benefit. Although advanced methods such as predictive analytics have historically required specialists such as data scientists, the evolution of software tools and applications is making it easier for nontechnical business users to apply predictive analytics. Tools can reduce the time and expense of data preparation and model development. This technology evolution is making predictive analytics more realistic and affordable for SMBs.

### ABOUT OUR SPONSOR



[sas.com](http://sas.com)

Want to take your business to the next level? SAS has a road map to get you there. Our integrated capabilities in data management, analytics, visualization, and reporting are designed and priced for small and midsize businesses. Our solutions work together—or separately—to make you better, smarter, and faster.

One of our latest offerings is SAS® Visual Analytics, which combines data visualization with powerful, lightning-fast analytics. It enables anyone within an organization—regardless of technical ability—to easily explore and manage data, and create and share highly visual reports. Plus, users can access and explore information via mobile devices, anytime and anywhere—even when there's no Internet connectivity.

Your organization—or your data—may not be very big. But the value of the information hiding in your data is. For more information, visit [sas.com/smb](http://sas.com/smb).

### ABOUT THE AUTHOR

**David Stodder** is director of TDWI Research for business intelligence. He focuses on providing research-based insight and best practices for organizations implementing BI, analytics, performance management, data discovery, data visualization, and related technologies and methods. He is the author of TDWI Best Practices Reports on mobile BI and customer analytics in the age of social media, as well as TDWI Checklist Reports on data discovery and information management. He has chaired TDWI conferences on BI agility and big data analytics. Stodder has provided thought leadership on BI, information management, and IT management for over two decades. He has served as vice president and research director with Ventana Research, and he was the founding chief editor of *Intelligent Enterprise*, where he served as editorial director for nine years. You can reach him at [dstodder@tdwi.org](mailto:dstodder@tdwi.org).

### ABOUT TDWI RESEARCH

TDWI Research provides research and advice for business intelligence, data warehousing, and analytics professionals worldwide. TDWI Research focuses exclusively on BI, DW, and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of business intelligence, data warehousing, and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

### ABOUT THE TDWI CHECKLIST REPORT SERIES

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.