SPEND ANALYSIS: GETTING THE BASICS OF DATA RIGHT

Abstract

Data has received unprecedented significance in recent times and has emerged as a critical input powering transformational technologies and improving competitiveness. Although the power of data is undisputable, companies continue to grapple with the challenge of poor quality of data. This is particularly true with procurement organizations that are increasingly relying on data to spot patterns, predict spend, and improve processes and procurement performance. This paper highlights the fundamentals of data management and the best practices for getting the basics of data right.

Infosys®
Navigate your next

Fundamentals of data management

In order to have a high quality of data, the first step is to identify the data sources and to understand the processes around data generation. This will help identify data and/ or process gaps with the aim to automate regular reporting tasks and define a long term data strategy. This data strategy will further help in defining data cleansing strategy to ensure that only clean and accurate data is available for analysis.

The execution of data strategy primarily involves two phases:

1. One-time historical analysis including data cleansing, taxonomy development, classification, analysis, and visualization

2. Ongoing periodic refresh with regular analysis reports

In both of these phases, the activities are similar. Any spend analysis solution primarily has three foundation blocks:

1. **Data management** – This phase consists of the following activities which are required in order to have a clean and accurate data:

   a. Identifying data sources and extracting data from these

   b. Defining data cleansing rules per field (such as dates, currency, cost center, and business units)

   c. Cleansing and normalizing the information in relevant fields and making the data usable for analysis

Most of these processes can be automated with minimal human intervention. During normalization, special focus should be on assigning a common representative name to similar looking supplier names. External data sources can be used to assign parent supplier relationship aimed at identifying the ultimate parent supplier name making it easier for supplier consolidation.

2. **Data classification** - Spend data classification is one of the most important aspect for identifying the right opportunity. This depends on the following factors:

   a. Taxonomy: for classification,

   b. Availability of information: to drive accuracy of classification, and

   c. Tools / technologies: for classification such as artificial intelligence (AI) based self-learning classification engines and analytics on big data platforms

3. **Visualization** – This phase involves visualizing the data using a visualization engine. This engine should have a set of ready to use reports with a drill down capability. Similarly, it should also have the capability to customize dashboards with the help of a simple drag and drop functionality.

a. Visibility on various parameters including spend per category, supplier, business unit, cost center, general ledger (GL) code, geography, and plant

b. Suppliers per category and percent spend per supplier

c. Various views on payment terms, currency, price, and related variations

d. Drilling down the data with respect to one or multiple parameters

These dashboards can provide users with the following actionable savings opportunities:

a. Identifying the scope for corporate contracts

b. Reducing / rationalizing suppliers across categories

c. Identifying discounts opportunities with the spend aggregation

d. Aggregating demand across organization

e. Monitoring and rationalizing supplier payment terms

f. Capturing and analyzing contract compliance

g. Identifying savings opportunities from purchase-price-variance

Strategies to fix the basics of data

The basic step prior to any data analysis is to ensure that the data is error free. In order to have that, it is critical to know the common sources of error and the means to fix those. Some of the key strategies to manage this are:

1. **Drive accurate classification** - More often than not, purchase orders (PO) or invoice descriptions do not provide sufficient information to correctly identify what is being purchased. To ensure the quality, the following steps can be helpful:

   a. The description quality improves with the increase in cataloged purchases. Free text POs, being a manual process, still remain a big change management issue. Hence, improving catalogue buying is a key strategy in improving incoming data quality.

   b. Post facto POs mostly have inadequate descriptions as their sole purpose is just to make a payment. The "no PO, no pay" policy can help with this situation as the buyer needs to provide the necessary information while purchasing any product/service which makes it easier to classify the line items correctly. If implemented properly, it can make the analysis smoother and deliver accurate results.

   c. Analyze the existing processes to identify the root causes of poor quality descriptions. This analysis could help define and work out a strategy to improve the incoming data quality.

   d. Standardize the classification. It is a common practice for many buyers to create their own classification. For example, for many OEM components, an item is bought as a part of a machine such as a gasket. It can be classified as a part of the pump part as it is used for pump sealing or it can also be classified as a gasket itself. Hence, it is always advisable to classify an item on its own and not per its usage.

2. **Refine taxonomy** - It is imperative to review the taxonomy being used and refine it to correctly reflect the commodities purchased. Here is an example of how availability of a correct taxonomy node helps in identifying spend category correctly:

Sample Spend Transaction:

Vendor Name	PO description	SECT Code	SECT Code description.	Spend amount
Wheeler Machinery Co.	MS992 Repair Serial# 3131J513SP	E02003	Haul Trucks	$3,280

Possible Classification per UNSPSC version 17 taxonomy structure:

Segment	Family	Class	Commodity	UNSPSC
Building and Construction Machinery and Accessories	Heavy Construction Machinery and Equipment	Earth Moving Machinery	--	22101500
OR				
Industrial Production and Manufacturing Services	Manufacturing Support Services	Manufacturing Maintenance and Repair Services	Nothing specific for earth moving equipment	73152100

Suggested correct UNSPSC customization and hence classification:

Segment	Family	Class	Commodity	UNSPSC
Industrial Production and Manufacturing Services	Manufacturing Support Services	Manufacturing Maintenance and Repair Services	Earth Moving Equipment Maintenance or Repair Services	73152113

Such refinement of taxonomy ensures that an appropriate taxonomy node is available for getting the spend line item classified correctly.

3. Tools/technologies for classification:

Tools and technologies play a vital role in driving accurate data classification. the Data classification engine should involve the following:

a. Rule based engine to ensure that the high value and high volume line items get classified correctly

b. Machine learning / artificial intelligence based engine, trained on customer specific data patterns, to ensure that the remaining data gets classified to the optimum accuracy with speed

In addition to the classification engine, the classification process should also be designed to capture the voice of customer to enhance the accuracy of the classification. A joint classification review with the category managers helps capture specific feedback to refine the classification. This review should primarily be a onetime activity unless glaring accuracy issues are observed during ongoing refreshes.

Thus, a best in class tool and review process ensure the accuracy and completeness of the data fed to the visualization engine.

Conclusion

Power of data is undisputed in the era of analytics driven decision making. With data emerging as a competitive advantage, companies continue their relentless quest to gain access to high quality data analysis. Investments in various tools and procurement systems continue to rise as corporates try to enhance their spend analytics capabilities. However, in order to drive an effective spend control through analytics, it is of utmost importance for the organizations to exert emphasis on the data quality.

Author

Vineet Kulkarni – *Practice Lead, Sourcing & Procurement, Infosys BPM*
Vineet is a domain expert for Spend analysis and master data management (MDM) and helps clients resolve their data challenges and also identify areas for delivering greater spend value. With 30+ years of experience, including 14+ years in the ITES industry, he has been instrumental in leading many spend analysis and master data projects end-to-end across multiple industries.

References

https://medium.com/@SeattleDataGuy/good-data-quality-is-key-for-great-data-science-and-analytics-ccfa18d0fff8
https://www.orpheus-it.com/big-data/perfect-spend-data
https://www.supplychainbrain.com/articles/27627-taking-control-of-data-quality
https://www.sdcexec.com/home/article/10269433/bad-data-the-problem-with-procurement

For more information, contact infosysbpm@infosys.com

Infosys®
Navigate your next

Infosysbpm.com

Stay Connected   SlideShare