Service Shop Performance Insights from ERP Data

Shaun West[1(✉)], Daryl Powell[2,3], and Ille Fabian[1]

[1] School of Technology and Architecture, Lucerne University of Applied Sciences and Arts, Lucerne, Switzerland
{shaun.west,fabian.ille}@hslu.ch
[2] Department of Industrial Ecosystems, SINTEF Manufacturing AS, Horten, Norway
daryl.powell@sintef.no
[3] Department of Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway

Abstract. Enterprise Resource Planning (ERP) systems offer firms a wealth of readily available transactional data. However, deriving insights from such data often demands the examination of multiple issues simultaneously. In this paper we use simple data mining to analyze ERP data from 27 service shops over a period of 35 months. The data has been used to provide valuable business performance insights to the service shop managers. Though the granular ERP data needed to be supplemented by further data in some instances, we found it has the potential to provide real insights into a firm's performance. Such simple data mining approaches can be standardized and automated across service centers for insights that can be used to drive continuous improvement activities within and across sites. We also suggest that this initial, exploratory study opens exciting avenues for further research into business analytics and, business intelligence pipelines.

Keywords: Data mining · Enterprise Resource Planning · Business performance · Service center

1 Introduction

This exploratory study examines the improvement of service workshop operations management using existing Enterprise Resource Planning (ERP) data, and what insights could support decision-making processes. The authors were given a set of data from the ERP system of a service provider who repairs a range of industrial equipment. The company has provided these services for over 60 years, yet the service shop managers only had the ERP data in the form of a monthly cash flow statement. The system was set up to provide monthly financial statements based on the transactions logged in the ERP system, and workshop managers and the operations lead confirmed that they were using it to help with planning and control. This was outside the ERP system's initial requirements and outside the finance department's direct control, apparently being a grassroots initiative.

Published by Springer Nature Switzerland AG 2021
A. Dolgui et al. (Eds.): APMS 2021, IFIP AICT 631, pp. 162–171, 2021.
https://doi.org/10.1007/978-3-030-85902-2_18

Prior studies [1–4] confirmed that ERP data provides a log of business transactions as a basis for financial reporting, and firms such as SAP have invested in tools like HANA to extract insights from data. Studies have considered the data for decision support within businesses, in the area of customer service support [5], data mining for business analytics [6], as well as knowledge assimilation and more advanced decision support [7, 8]. Others have considered manufacturing firms or pure service businesses, but not an industrial repair and overhaul business that provides both workshop and field-based services. For these reasons, the research question for this paper is: "*what service center insights can be learnt from data mining of ERP systems?*".

2 Literature Review

The research question examines benchmarking and decision-making to support management planning, control, and performance in the context of operational excellence [9]. Using ERP data for operations management has been identified [10] as a strategical and a tactical approach in manufacturing firms and made-to-order businesses [11, 12].

Data mining can [13] reveal insights within a firm, and has been used in many different business areas, within the sales function, customer support and manufacturing [1, 3–5]. Increasingly, it is now being used for forward-looking business analytics [6], coupled with machine learning and other techniques. Studies [2, 14–16] describe using ERP systems to support lean or continuous improvements within firms in different manufacturing contexts, including in SMEs similar in size to typical service workshops. The maturity model [16] provides a five-level model with examples showing how and where ERP systems can support pull production in firms.

Benchmarking can improve organization performance by comparison with others, as described in the literature [17]. Often based on external benchmarks, it can be used internally to consider cross-business performance and used to support knowledge management and continuous improvement [18]. With the use of business analytics systems based on ERP data [19–21] further insights can be gleaned [7].

Decision-making is enhanced when data is presented in a form such as clear visual representation [17], that allows a team to assimilate it into knowledge and deliver actionable decisions and forecasts [23, 24]. ERP data can and should be used to support decision-making [8].

3 Methodology

Statistical analysis of an ERP data set of monthly cash flow and invoicing data from 27 service centers in one region. The data were provided as flat CSV files to be imported into MS Excel, SPSS, or Microsoft Power BI. Additional contextual management data to support the analysis was collected separately and additional tables were created. The steps applied to the raw ERP data were: data cleaning; collecting contextual information for the ERP data set; structuring the flat data to the structure of the business and its sub-business units; exploring the invoicing data and monthly cashflows.

4 Results

Around 3MB of raw data from 27 service shops were collected in a CSV file (later moved into Excel) over a 35-month period. Billing captured 40,000 lines of data whereas monthly cashflows provided 39 lines of data per month. Tables 1 and 2 provide an overview of the data fields.

Table 1. Billing data fields

Location code	Workshop location and sub-business unit
Customer code	The code for the customer
Zip code	The delivery Zip code
Net sales value	The value on the invoice
Net cost	The costs associated with the work

Table 2. Monthly cashflow data fields

Data field	Data field	Data field
Jobs sold	CM/sales	Amortization
Hours sold	Overhead costs	EBIT
External sales	Overhead recovery	Interest charges
Internal sales	Debt provisions	Profit before tax (PBT)
Total sales	Gross profit	BPT/Sales
Cost of goods sold	Gross profit/sales	WIP provisions
Labor costs	Admin costs	Bad debt provisions
Under absorption cost	Sales costs	Pre interest PBT
Employee social costs	Distribution costs	Internal rechanges
Material costs	Mgt overheads	Pre-exceptional PBT
Job expenses	Rent (nominal)	Exceptional costs
Provisions	EBITDA	PBT contribution
Contribution margin	Depreciation	No of Employees

4.1 Initial Analysis in Excel for Data Cleaning

The data was assessed in Excel for correction and validation and was generally consistent, though there were issues with the invoicing data, where the calculated margins could be very large (both negative and positive outliers). With no apparent reason for this, the outliers were removed for the analysis in this study. Discussions with the local individual workshop managers confirmed that there were quality issues with the data as often invoices were issued without them being linked directly to a repair project. This came about mostly due to additional repairs needed or to invoicing separately for out-of-scope items. At times invoicing was also used as a price adjustment mechanism, breaking a clear link between the cost of goods sold and the invoice value and creating a quality issue with the data. When analyzing this type of data is necessary to understand its contextual meaning.

The monthly cashflow data was used in the management control processes and every branch manager was sent a month end cash flow (Fig. 1). The sheet had evolved over time, allowing branch managers to talk knowledgably with the head of the operations, the business units and other branch managers, as well as with the finance departments. All branch managers were used to using the figures and could take meaningful actions based on the sheets.

	Jul-07 Actual	Aug-07 Actual	Sep-07 Actual	Oct-07 Actual	Nov-07 Actual	Dec-07 Actual	Jan-08 Actual	Feb-08 Actual	Mar-08 Actual	Apr-08 Actual	May-08 Actual	Jun-08 Actual	**2007 / 2008**
JOBS SOLD	46	39	41	49	39	31	54	51	34	57	41	34	**516**
HOURS SOLD	2,331	1,901	1,682	2,719	2,758	1,442	1,934	1,939	1,782	2,279	1,908	1,333	**24,008**
EXTERNAL SALES	73	83	80	109	101	118	78	80	69	146	63	59	**1,059**
INTERNAL SALES	20	6	5	2	4	4	4	5	2	9	6	5	**72**
TOTAL SALES	93	89	86	110	105	122	82	85	71	155	69	65	**1,132**
COST OF SALES :													
LABOUR COSTS	27	21	20	32	34	17	22	22	20	25	21	15	**277**
NON-PROD LABOUR	(3)	(1)	(1)	(4)	(3)	1	0	(0)	3	(0)	2	1	**(6)**
WORKS NI & SUP'N	3	3	2	3	3	3	3	2	3	3	3	3	**31**
MATERIAL COSTS	15	24	7	18	18	57	27	11	13	56	11	9	**266**
JOB EXPENSES	0	0	0	0	0	0	0	0	0	0	0	0	**0**
ADD BACK WIP PROV. MOVE.	0	0	0	0	0	0	0	0	0	0	0	0	**0**
MAN. GP	51	42	57	61	53	45	30	49	32	71	33	38	**563**

Fig. 1. Extract of the data provided to the managers from the monthly cashflows

4.2 Collection and Integration of Contextual Information

Initial analysis confirmed contextual information was missing: the floor space for each workshop and the hierarchical relationship with the business (e.g., its sub-business unit, based on business model, not location). Contextual information supports the wider interpretation of the data and helps to frame the data within the larger picture. The information was stored on a separate system within the finance department.

4.3 Structuring of the Flat Data

The business structure was used to create a synthetic consolidation of the invoice data and the cashflows. This created business level, and business unit perspectives of the data sets and individual branches. The data collected did not provide the consolidation as a business unit or business level on the same basis as the data from the individual branches. For this reason, a synthetic consolidation was needed, although this may have not fully considered the intra-business trading fully. This confirms that data cleansing to remove artifacts is an important activity where more focus should be given.

4.4 Exploration of the Invoicing Data

The invoicing data was explored using a set of scatter plots before moving to box plots, which offer more insights than the final scatter plots. Here, benchmarking shows high variance in both the size of individual jobs and the margins per job. Figure 2 shows the breakdowns from the business to the BU and finally to location 1 (LO01), in effect allowing consolation that was created as part of the structing of the flat data. Location 17 (LO17) has a Gross Profit (GP) margin that is slightly lower than the average for Business Unit (BU) B it has long tails, whereas LO19 with lower variance provided a more reliable project Gross Profit. Improvement in individual service shops' performance (e.g., no invoices with less than zero GP) could improve the firm's margin overall. The boxplots for the individual locations from on BU allow a comparison for each individual location and provides the comparisons with the other BUs as well as the business in aggregate.

Noise in the data from the scatter plots, required additional contextual information to understand the challenges with the data quality. Discussions confirmed that there were many invoices created where they were used for billings without matching the costs directly with them. Interestingly, anomaly data appears to be associated with the smaller invoice values and a filter applied at ±80% GP margin may provide a pragmatic approach to data cleansing.

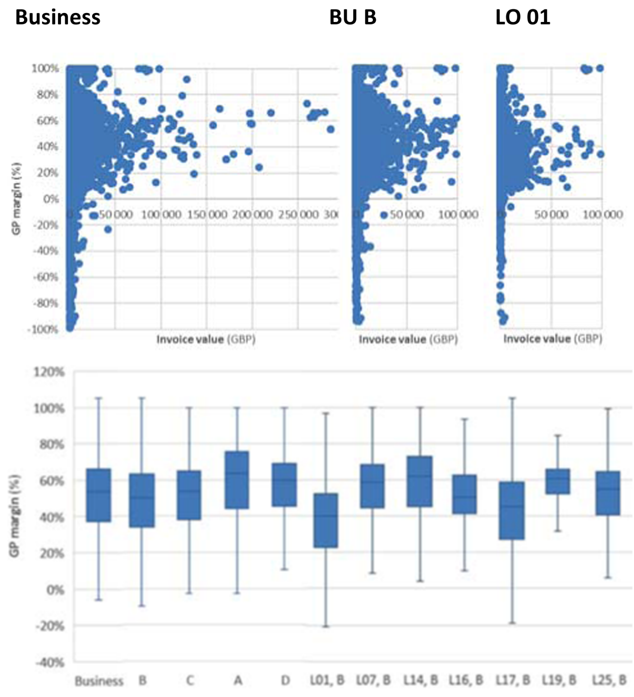

Fig. 2. Invoicing data provided statistical insights into the operations

Analysis of the customer base confirmed that the top ten customers on the consolidated business or at the business-unit level contributed less than 10% to the total sales volume (Fig. 3) but, when drilling down to individual service centers, the top ten customers often contributed 30% or more of the total sales, so the loss of two key accounts could create challenges for a service shop. This reliance on large orders is an insight into the underlying business model, as well as the risk of missing targets due to failing to win a large order. The firm focused on the overall business reliance on single customers and large orders rather than undertaking a detailed customer analysis at the business unit level providing new insights for the managers.

The use of customer post codes helped to visualize the geographic distribution of the customer base. However, we assumed that the invoice address was the ship-to address, but the data showed this was not always the case. Discussions within the firm confirmed that there could be problems with confusions between ship-to and bill-to addresses. This suggests a problem with both data quality and data structure.

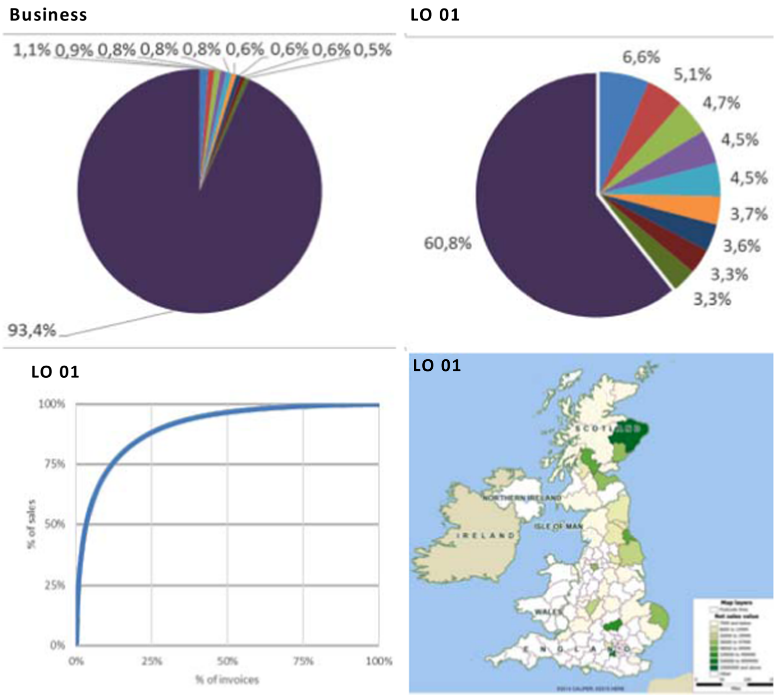

Fig. 3. Different perspectives of business sustainability (customers, project size, customer locations)

4.5 Exploration of the Monthly Cashflows

Trend data were plotted again at the three different levels of the firm (Fig. 4). In the past, only the spreadsheet had been used, but visuals presented the data in a more actionable form. The upper charts show data on a rolling 18-month basis, rather than a financial year basis, and lean approaches were used to convert the charts into control charts with a regression trend line to provide a form of forecast. The basic regression model was used. There was only a basic understanding of statistics, and the approach is a early approach with time-series data and commonly applied within firms with a lean/six-sigma improvement program. The upper and lower control lines were based on the ±2 sigma around the mean. The middle charts show the number of booked working hours in the month with LO 01 showing significant variance month-on-month. Compared with monthly under/over absorption, the weakness in management control was clear, as shown in the lower chart. Sales/hours show increases in value added and reflect the knowledge intensity, whereas costs/hours show the change in productivity. Other combinations of data (e.g., ROS%, Jobs Sold, Jobs sold/hours sold) were tested to understand the dynamics of the business. In several workshops there was under-absorption of labor one month and over-absorption the next, to the detriment of the ROS (in percentage and in absolute terms). These data could provide the basis for forecasting.

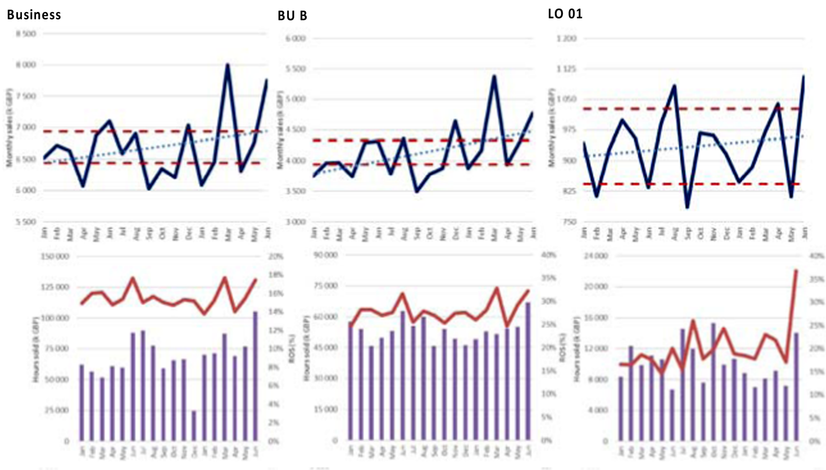

Fig. 4. Different perspectives providing different insights into the data

4.6 Benchmarking Metrics

To understand if the business is based on spares (materials), sales or labor, the materials cost/COGS was plotted on a monthly basis. Significant differences could be found between locations and business units, reflecting different business models. For example, one location was an outlier with high sales per square meter. Discussions with the business confirmed that the branch's sales contained a large volume of field service sales significantly greater than most locations.

5 Discussion

Our research question was "*what service center insights can be learnt from data mining of ERP systems?*" Drawing on ERP data from 27 service centers, we combined supplementary data and applied data mining to derive useful and useable insights. Using ERP data as a source was found to agree with the literature [10–12]. However, the use of visualizations to support decision-making, as seen in this study, was missing from the firm's approach. Data mining gave new insights into the firm's performance and again agrees with the literature [3–5]. What was missing was translating the financial data into actionable visual insights that could be used on an operational and tactical basis in the service shops. Also missing was an attempt to draw forecasts from the insights, or to use the insights as a benchmarking tool to support understanding of business performance [14, 16, 17].

5.1 Managerial Relevance

The approach moves away from traditional financial reporting and into the domain of business intelligence and analytics. Using ERP data integrated with contextual insights and additional information (ie., floor areas, employees etc.), lessons can be learnt and shared within the service network. This should improve performance, becoming the basis of continual improvement, and decision support (operational, tactical and strategic). Implementing such a tool needs careful consideration to ensure its usability and management use. This study confirmed that data collection (transactional data), derived parameters (hypercube dimensions) and insights should be structured around the individual operational locations and business units, as well as the overall business, so relevant information reaches different managerial levels to support business decisions. Doing so it became possible to gain new insights into the business.

5.2 Academic Implications

There are too few published examples of performance management of service business. Much of the available data is either at too high a level to allow detailed analysis or has critical data missing that prevents analysis. The data here confirms that granular ERP data, with some supplementary data, can provide real insights into a firm's performance. The lessons learnt from the analyses demand further investigation and could provide the basis for sharing lessons and experiences within a service business. Statistical-based models could be developed and integrated with business intelligence and analytics solutions to provide forecasting capabilities, based on the business structures that the data has described. The models could support decision-making on operational, tactical, and strategic levels.

6 Conclusions and Future Research

The analysis of the ERP data gave the firm new insights into the business's performance, and more insights could be gleaned when additional data was added to provide new KPIs.

The reports could be generated automatedly and shared monthly to support the branch managers, the business units, and the whole firm. The process could be represented as OLAP cubes to provide real-time business intelligence. Transforming insights into a visual form can offer actionable information that can support tactical and strategic decisions. For instance, a wide difference between GP% and the ROS% could identify a service center location where prices were maintained, albeit with a low total sales volume. Additional market data outside of the ERP would be required to confirm this.

This is an initial study and has not taken full account of the literature, the potential for operational improvements, or improved decision support. Advanced analytics has not been applied to the datasets. The possibility remains to create forecasts from the data and business simulations. The usability of the original uncleaned data set and the traditional management reports were not investigated in this study, nor were the use and useability of the new insights in terms of their support to the decision-making process. This should be investigated further, as should their integration into existing processes to ensure the impact within the business. There are therefore limitations to the study that demand further investigations in these areas.

References

1. Cohen, M.C.: Big data and service operations. Prod. Oper. Manage. (2018). https://doi.org/10.1111/poms.12832
2. Powell, D., Alfnes, E., Strandhagen, J.O., Dreyer, H.: The concurrent application of lean production and ERP: towards an ERP-based lean implementation process. Comput. Ind. (2013). https://doi.org/10.1016/j.compind.2012.12.002
3. Kusiak, A.: Data mining: Manufacturing and service applications. Int. J. Prod. Res. (2006). https://doi.org/10.1080/00207540600632216
4. Berry, M.J.A., Linoff, G.S.: Data mining techniques: for marketing, sales, and customer relationship management. Portal.Acm.Org (2004)
5. Hui, S.C., Jha, G.: Data mining for customer service support. Inf. Manage. (2000). https://doi.org/10.1016/S0378-7206(00)00051-3
6. Ledolter, J.: Data mining and business analytics with R. Data Min. Bus. Anal. R (2013). https://doi.org/10.1002/9781118596289
7. Elbashir, M.Z., Collier, P.A., Sutton, S.G., Davern, M.J., Leech, S.A.: Enhancing the business value of business intelligence: the role of shared knowledge and assimilation. J. Inf. Syst. (2013). https://doi.org/10.2308/isys-50563
8. Holsapple, C.W., Sena, M.P.: EP plans and decision-support benefits. Decis. Support Syst. (2005). https://doi.org/10.1016/j.dss.2003.07.001
9. Ivanov, D., Tang, C. S., Dolgui, A., Battini, D., Das, A.: Researchers' perspectives on Industry 4.0: multi-disciplinary analysis and opportunities for operations management. Int. J. Prod. Res. (2020). https://doi.org/10.1080/00207543.2020.1798035
10. Chiarini, A., Vagnoni, E.: Strategies for modern operations management. Benchmarking Int. J. (2017). https://doi.org/10.1108/bij-11-2015-0115
11. Aslan, B., Stevenson, M., Hendry, L.C.: Enterprise resource planning systems: an assessment of applicability to make-to-order companies. Comput. Ind. (2012). https://doi.org/10.1016/j.compind.2012.05.003
12. Kakouris, A.P., Polychronopoulos, G.: Enterprise resource planning (ERP) system: an effective tool for production management. Manage. Res. News (2005). https://doi.org/10.1108/01409170510784878

13. Larose, D.T.: Discovering Knowledge in Data: An Introduction to Data Mining (2005). https://doi.org/10.1002/0471687545
14. Powell, D., Alfnes, E., Strandhagen, J.O., Dreyer, H.: ERP support for lean production. In: Frick, J., Laugen, B.T. (eds.) APMS 2011. IAICT, vol. 384, pp. 115–122. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33980-6_14
15. Powell, D.: ERP systems in lean production: new insights from a review of lean and ERP literature. Int. J. Oper. Prod. Manage. (2013). https://doi.org/10.1108/IJOPM-07-2010-0195
16. Powell, D., Riezebos, J., Strandhagen, J.O.: Lean production and ERP systems in small- and medium-sized enterprises: ERP support for pull production. Int. J. Prod. Res. (2013). https://doi.org/10.1080/00207543.2011.645954
17. Agrahari, A., Srivastava, S.K.: A data visualization tool to benchmark government tendering process: insights from two public enterprises. Benchmarking (2019). https://doi.org/10.1108/BIJ-06-2017-0148
18. McGinnis, T.C., Huang, Z.: Rethinking ERP success: a new perspective from knowledge management and continuous improvement. Inf. Manage. (2007). https://doi.org/10.1016/j.im.2007.05.006
19. Shi, Z., Wang, G.: Integration of big-data ERP and business analytics (BA). J. High Technol. Manage. Res. (2018). https://doi.org/10.1016/j.hitech.2018.09.004
20. Gupta, S., Qian, X., Bhushan, B., Luo, Z.: Role of cloud ERP and big data on firm performance: a dynamic capability view theory perspective. Manag. Decis. (2019). https://doi.org/10.1108/MD-06-2018-0633
21. Chou, D.C., Bindu , H., Chou, A.Y.: BI and ERP integration. Inf. Manage. Comput. Secur. (2005). https://doi.org/10.1108/09685220510627241
22. Erasmus, P., Daneva, M.: An experience report on ERP effort estimation driven by quality requirements. In: CEUR Workshop Proceedings (2015)
23. Karsak, E.E., Özogul, C.O.: An integrated decision making approach for ERP system selection. Expert Syst. Appl. (2009). https://doi.org/10.1016/j.eswa.2007.09.016
24. Grover, P., Kar, A.K., Dwivedi, Y.K.: Understanding artificial intelligence adoption in operations management: insights from the review of academic literature and social media discussions. Ann. Oper. Res. (2020). https://doi.org/10.1007/s10479-020-03683-9