CRANFIELD UNIVERSITY

ALEX WOOLFORD

VISUALIZATION FOR COMMODITY NEGOTIATIONS IN THE ELECTRONICS MANUFACTURING SERVICES INDUSTRY

SCHOOL OF MANAGEMENT

Logistics and Supply Chain Management

MSc THESIS

Academic year: 2011-12

Supervisor: Melvyn Peters

June 2012

CRANFIELD UNIVERSITY

SCHOOL OF MANAGEMENT

Logistics and Supply Chain Management

MSc THESIS

Academic Year 2011-2012

ALEX WOOLFORD

Visualization for commodity negotiations in the electronics manufacturing services industry

Supervisor: Melvyn Peters

June 2012

This thesis is submitted in partial fulfillment requirements for the degree of Master of Science

ABSTRACT

In the electronics manufacturing services industry (EMS), materials procurement negotiations have a significant impact on the profitability of the company. This requires negotiators to rapidly interpret large amounts of procurement data. The purpose of the project was to utilize open source data visualization technologies to support supplier negotiations at a large EMS company. A prototype dashboard is included.

ACKNOWLEDGEMENTS

I am extremely grateful to my beautiful wife, Victoria, for her unwavering support and positive energy; to my boys, Miles (6) and Bradley (3), who make everything worthwhile; and to my dad, Ted, who made this academic journey possible.

I would also like to thank my advisor, Melvyn Peters, for his patience, guidance and insight; Andy Scanzio and Brent Frame, my previous managers at Flextronics, who encouraged me to learn and experiment; and Hadley Wickham, creator of many excellent R packages that have saved me countless hours (most notably ggplot, plyr, reshape, stringr, and lubridate).

CONTENTS

1 INTRODUCTION 1

1.1 About the EMS Industry 3

1.2 Flextronics 5

1.3 Commodity Negotiations 7

1.4 Structure of the thesis 11

2 BACKGROUND 11

2.1 Visualization 11

2.2 Tools 15

2.3 Business Intelligence 16

2.4 Visualization for Negotiations 17

3 METHODOLOGY 20

3.1 Previous Methods 20

3.2 Options Available 20

3.3 Justification of Choice 21

4 ANALYSIS 24

4.1 Stages in method 24

1) Definition 24

2) Literature review 24

3) Build prototype 24

4) Deploy prototype 34

5) Solicit feedback 34

6) Refine tool 35

7) Assessment 37

5 DISCUSSION 39

5.1 Data Quality 39

5.2 Alternative Part Identification 39

5.3 Lead Times 43

5.4 Transportation Cost 43

5.5 Non-Disclosure Agreements 44

5.6 Open Source Software 44

5.7 Scalability 45

5.8 A Passive Tool 45

6 CONCLUSIONS AND CRITIQUE 47

6.1 Findings with particular respect to objectives 47

6.2 Limitations of analysis 49

6.3 Future directions / recommendations for further work/research 49

6.4 Critique of the work (approach, analysis, etc…) 50

7 RECOMMENDATIONS 51

7.1 Short term 51

7.2 Medium term 51

7.3 Long term 51

7.4 Further Research 51

8 REFERENCES 53

APPENDIX A – FEEDBACK 57

1) Commodity Manager, Global Procurement, Mexico 57

2) Sr Director, Commodity Management, Global Procurement, Mexico 57

3) Sr Director, Commodity Management, Integrated Network Solutions, Mexico and Commodity Manager, Global Procurement, Mexico 57

4) Sr Global Commodity Manager, Global Procurement, Austria 57

5) Sr Director, Commodity Management, Global Procurement, USA 57

6) Vice President of Sourcing, Global Procurement, USA 57

7) Director, Information Technology, St. Petersburg, Russia 57

8) Sr Director, Commodity Management, Global Procurement, Mexico 57

9) Sr Director of Sourcing, Global Procurement, USA 58

10) Commodity Manager, Global Procurement, Mexico 58

11) Sr Purchasing Manager, Global Procurement, Israel 58

**FIGURES**

Figure 1 Stan Shih's smiling curve (Source: Author) 3

Figure 2 Profit margins across the EMS supply chain (Source: Author; Data: Google, 2011) 5

Figure 3 Spend: top 20 customers; commodity groups; Flex controlled % (Source: Author) 6

Figure 4 Scatterplot: Flex controlled % versus PPV % (Source: Author) 7

Figure 5 Target versus actual pricing reduction from suppliers (Source: Author) 9

Figure 6 Flextronics' sourcing control for spend with the top 30 suppliers (Source: Author) 11

Figure 7 Treemap example: materials procurement cost savings (Source: Author, 2012) 15

Figure 8 Approach to methodology (Source: Author, 2012) 23

Figure 9 Plotting function template, written in R (Source: Author, 2012) 25

Figure 10 Dashboard function (Source: Author, 2012) 25

Figure 11 Create dashboards for a list of parts (Source: Author, 2012) 26

Figure 12 Prototype ggplot2 faceted dotplot (Source: Author, 2012) 29

Figure 13 Query performance: a time based process map (Source: Author, 2012) 32

Figure 14 Multi-step queries with platform specific optimization (Source: Author, 2012) 33

Figure 15 Step plot to show contract pricing history (Source: Author, 2012) 36

Figure 16 Multiple ERP platforms at Flextronics (Source: Author, 2012) 39

Figure 17 AML cross using a self-join (Source: Author, 2012) 40

Figure 18 Differences in attributes of two items that have identical descriptions (Source: Author, 2012) 40

Figure 19 Parametric search for alternative parts (Source: Author, 2012) 42

Figure 20 Cumulative demand percent versus the number of parts (Source: Author, 2012) 45

Figure 21 Prototype dashboard (Source: Author, 2012) 48

ABBREVIATIONS

AML Approved Manufacturer List

API Application Programming Interface

ASIC Application-Specific Integrated Circuit

BATNA Best Alternative to Negotiated Agreement

EMD Electro-Mechanical Device

EMS Electronics Manufacturing Services

ERP Enterprise Resource Planning

ETL Extract Transform Load

GCM Global Commodity Manager

GUI Graphical User Interface

HES High End Semiconductor

LLD Linear Logic Device

NTS Negotiation Toolset

ODBC Open Database Connectivity

OEM Original Equipment Manufacturer

PCB Printer Circuit Board

PDF Portable Document Format

PO Purchase Order

PPV Purchase Price Variance

QBR Quarterly Business Review

SEC Securities and Exchange Commission

SMI Supplier Managed Inventory

SQL Structured Query Language

XML eXtensible Markup Language

# INTRODUCTION

In the fiscal year of 2011, Flextronics spent $26 billion on materials to build electronic products that returned $29 billion in revenue, indicating that the material cost of sales was almost 90 percent. Electronics manufacturing services (EMS) is a mature industry with low profit margins. Over the last decade our manufacturing facilities have driven out a lot of cost through Lean initiatives. Assembling electronics is something that Flextronics does extremely well but, unfortunately, this has become a commodity. When faced with excess manufacturing capacity, EMS companies sacrifice profit margins to cover their fixed costs. For example, Flextronics’ biggest competitor, Foxconn, manufactures the iPhone that retails for $560 for just $14 ($7 to assemble, $7 profit) (The Economist 2011).

When profit margins are so thin and materials make up such a large proportion of the costs, procurement becomes a critically important function. Around 50 Flextronics commodity managers meet face-to-face with suppliers to negotiate pricing, and they specialize in different areas.

All this procurement activity generates a lot of data. Although large companies are often awash with data, they often fail to transform it into actionable knowledge (Sackett and Williams 2003). There were over 13 million receipts for more than 700,000 different parts from more than 14,000 suppliers in fiscal year 2011. Although most of the commodity managers are familiar with Microsoft Excel, they typically don't have the skill set to transform the millions of records in the Oracle data warehouse into information that can support negotiation arguments. To bridge this gap, Flextronics built the ‘Negotiation Toolset’ (NTS). This distills data from 28 different tables into a single table, called the Negotiation Item Master, with one record per part. The records are then parsed into a collection of spreadsheets called ‘masterfiles’ for each of the commodity managers.

The current system attempts to represent everything there is to know about a part in a single record. Many of the source tables have one-to-many relationships; one part may have many pricing records, many receipts, and many open orders. When the tables are ‘flattened,’ we sacrifice some of the information richness for the sake of simplicity.

The masterfiles currently contain more than 200 columns and this number grows each quarter. Many users report difficulty in navigating the file because there are too many columns, and that the file size can cause computer performance issues when trying to manipulate data.

Information asymmetry is a source of competitive advantage. Prior to the face-to-face negotiations, the commodity managers sort, slice and sift through their spreadsheets to formulate their arguments. Certain parts are identified, flagged for discussion, and sent to the supplier before the face-to-face negotiation. The Texas Instruments negotiation, for example, takes three days and usually around 300 parts are discussed each day. There are a series of short dialogues averaging less than two minutes each, where the individual parts are discussed. Within each of these mini-dialogues, material prices are determined. It is therefore critical that the commodity manager has all the relevant facts to support his arguments at his fingertips.

Commodity managers are not statisticians. The effective use of statistical analysis depends on choosing the right model and being able to interpret the results. Commodity managers may find statistical terminology cryptic and inaccessible, and they may not be able to easily communicate the meaning to the person they are negotiating with. Analysis is currently centered around arithmetic and aggregations from the masterfile data.

Visualizations, by contrast, are intuitive and rich. Most people, without knowing what a p-value was, could look at a scatter plot and identify the outliers. Visualizations can sometimes uncover subtle details that may be missed by analysis since the information density is much higher than text or a table of data. Data graphics also help transcend language and cultural barriers. Visualization technologies have grown rapidly over the past few years

The purpose of this project is to evaluate the effectiveness of the addition of data visualization technologies to supplier negotiations at Flextronics. Because commercial business intelligence software license fees are typically expensive and IT resources are scarce, a decision was made to limit this project to open source solutions. There were three objectives:

1. Evaluate open source technologies visualization and select an appropriate platform to make visual dashboards to support supplier negotiations at Flextronics.
2. Build a prototype negotiation dashboard.
3. Evaluate the dashboard to see if it is beneficial to commodity negotiations at Flextronics.

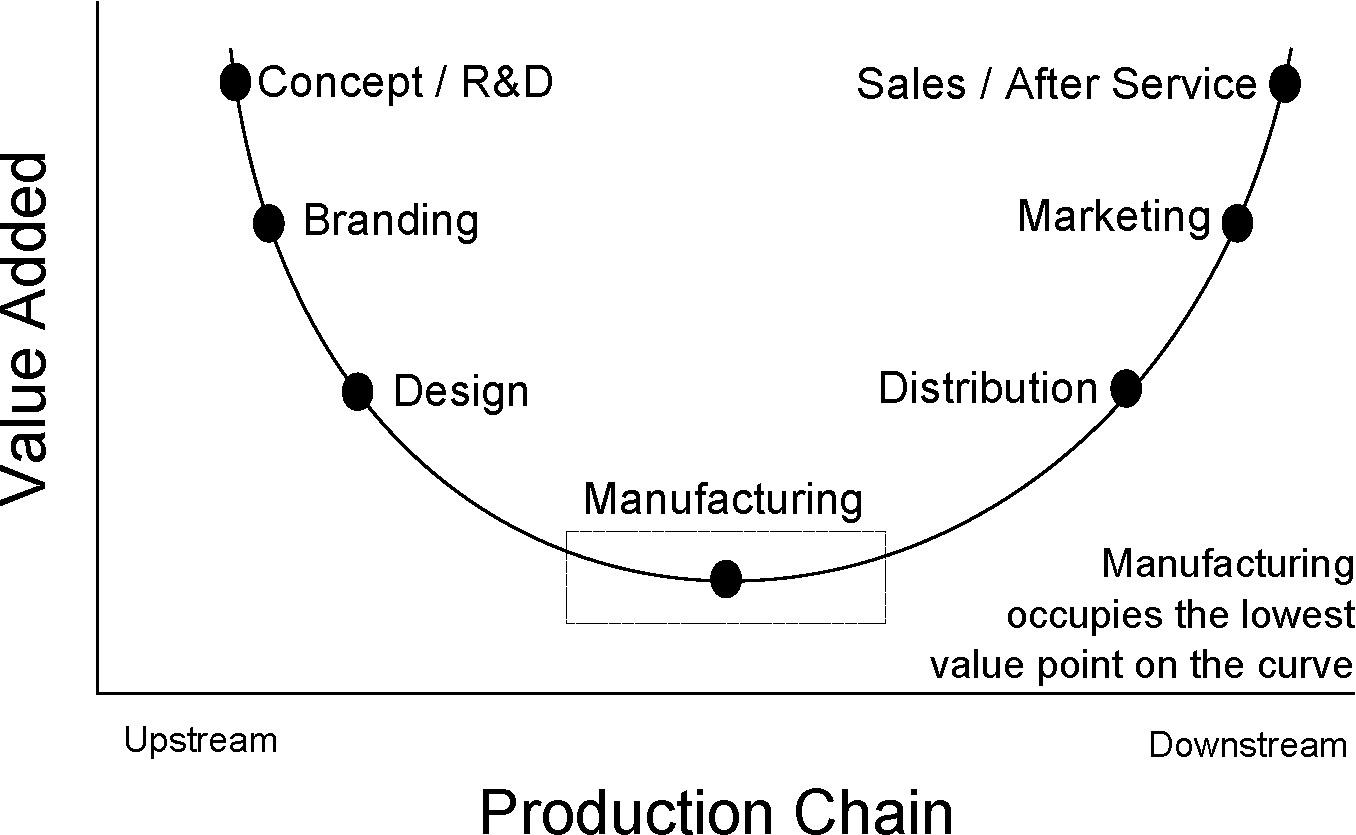
## About the EMS Industry

The Electronics Manufacturing Services (EMS) industry is comprised of contract manufacturing companies that build products for Original Equipment Manufacturers (OEMs). For example, Cisco (an OEM) specializes in the design and marketing of computer networking equipment. The role of the EMS industry is to purchase the materials from the manufacturers/distributors of the components and build the product for the OEM.

The EMS industry creates value for its customer in two ways. Firstly, EMS companies benefit from the commonality of parts across many products. This consolidation of materials demand enables the contract manufacturer to negotiate pricing with suppliers more effectively than the OEMs. In addition to pricing, the suppliers are more likely to offer a Supplier Managed Inventory (SMI) program with short lead times. Ultimately this allows the EMS companies to offer greater flexibility to their customers with regard to changes in demand. Secondly, EMS companies mitigate the risk of investment in manufacturing capacity through maintenance of a diverse portfolio of OEM customers across multiple markets. In this way, the workload at the manufacturing facilities remains somewhat stable because an increase in demand within one industry may compensate for a decrease in another. The stability maintains quality standards at the factory because it eliminates the necessity to periodically hire and fire production line workers (Barnes, et al. 2000).

Stan Shih, the founder of Acer, created a smiling curve (see figure 1) which highlights the idea that of all the activities in the production chain, manufacturing adds the least value (Shih, 2005, as cited in Yan & Islam, 2011).

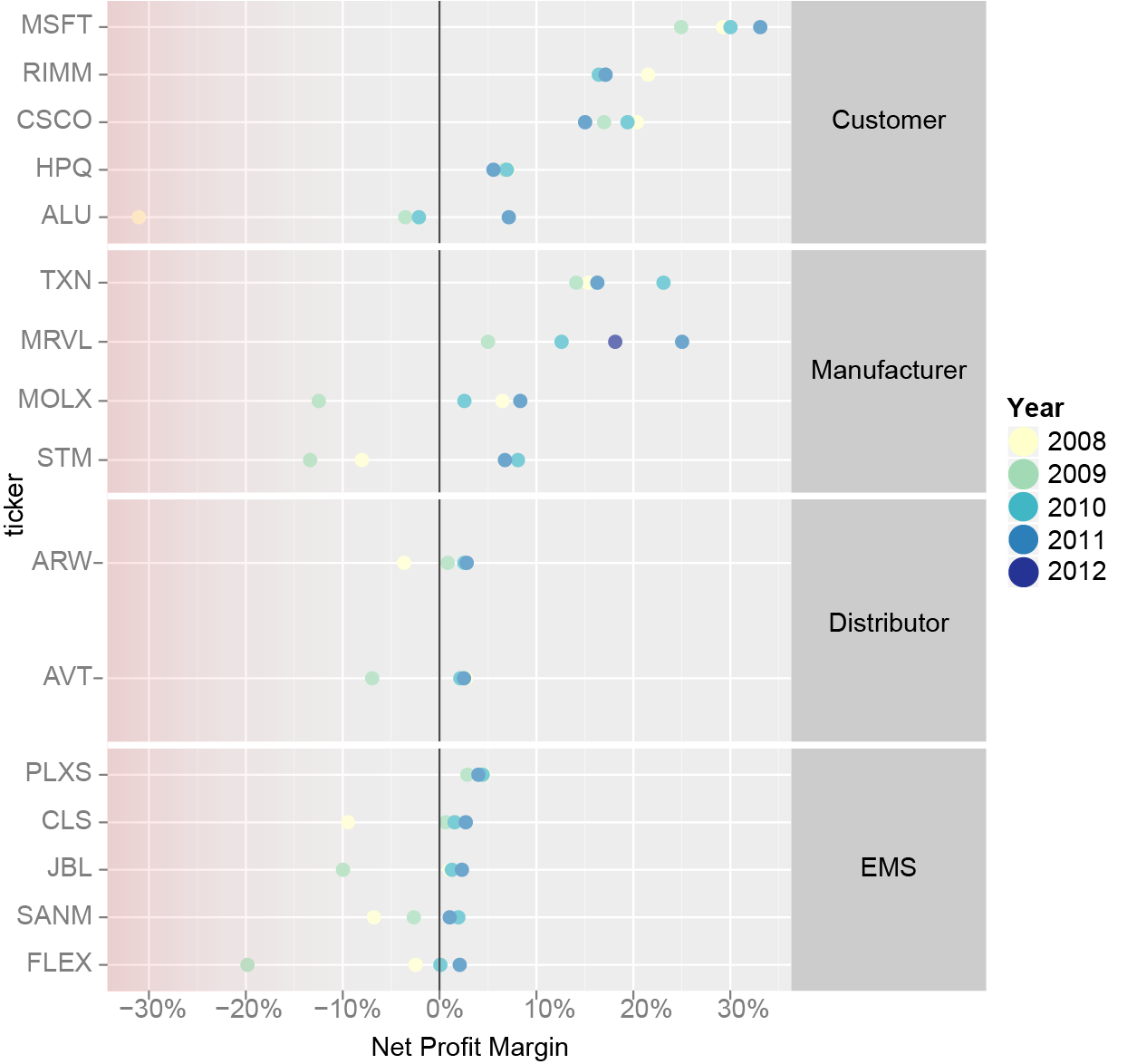
Figure Stan Shih's smiling curve (Source: Author)



Consequently, profit margins are low in the EMS industry (Barnes, et al. 2000).

The chart below (see figure 2) shows a comparison of profitability both across and within the major players in the production chain (OEM, component manufacturer, component distributor and EMS) in descending order of profitability. The black vertical line highlights the zero percent net profit margin as a reference point. Points to the left of the zero line made a net loss that year, whereas points to the right made a net profit. The publicly traded companies represented by the ticker symbols were selected based on their relevance to the industry, and Flextronics in particular. It is important to note that public financial data for Foxconn, Flextronics’ largest competitor, was unavailable in a reliable format because this Taiwanese company is not accountable to Securities and Exchange Commission (SEC) reporting standards.

Figure Profit margins across the EMS supply chain (Source: Author; Data: Google, 2011)



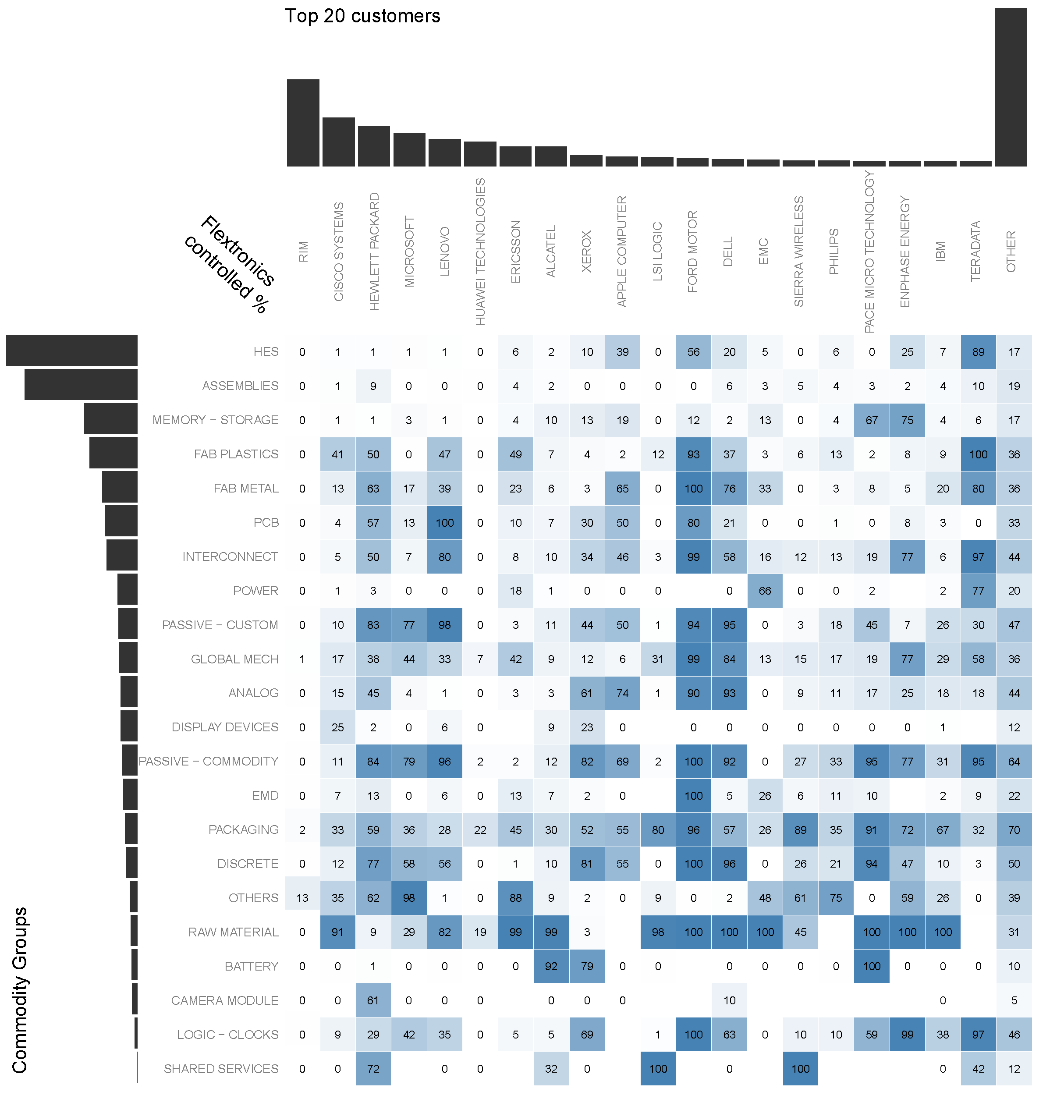
The findings from the chart are consistent with Stan Shih’s smiling curve, assuming that adding value drives profit. Clearly, the EMS industry is subject to very thin profit margins. The importance of negotiation becomes clear when the low profit margins of the industry are considered.

## Flextronics

Each OEM decides how much autonomy it is prepared to give the EMS with respect to sourcing decisions. Suppliers are often unaware of the level of autonomy that the EMS company has in its ability to source materials.

The following graphic shows materials spend for Flextronics’ top 20 customers and commodity groups. The barcharts represent the relative spend for each customer/commodity and are ordered in descending spend value for each customer and commodity. The heatmap section of the graphic represents the percentage of sourcing that is controlled by Flextronics. The darker tones indicate greater autonomy for Flextronics.

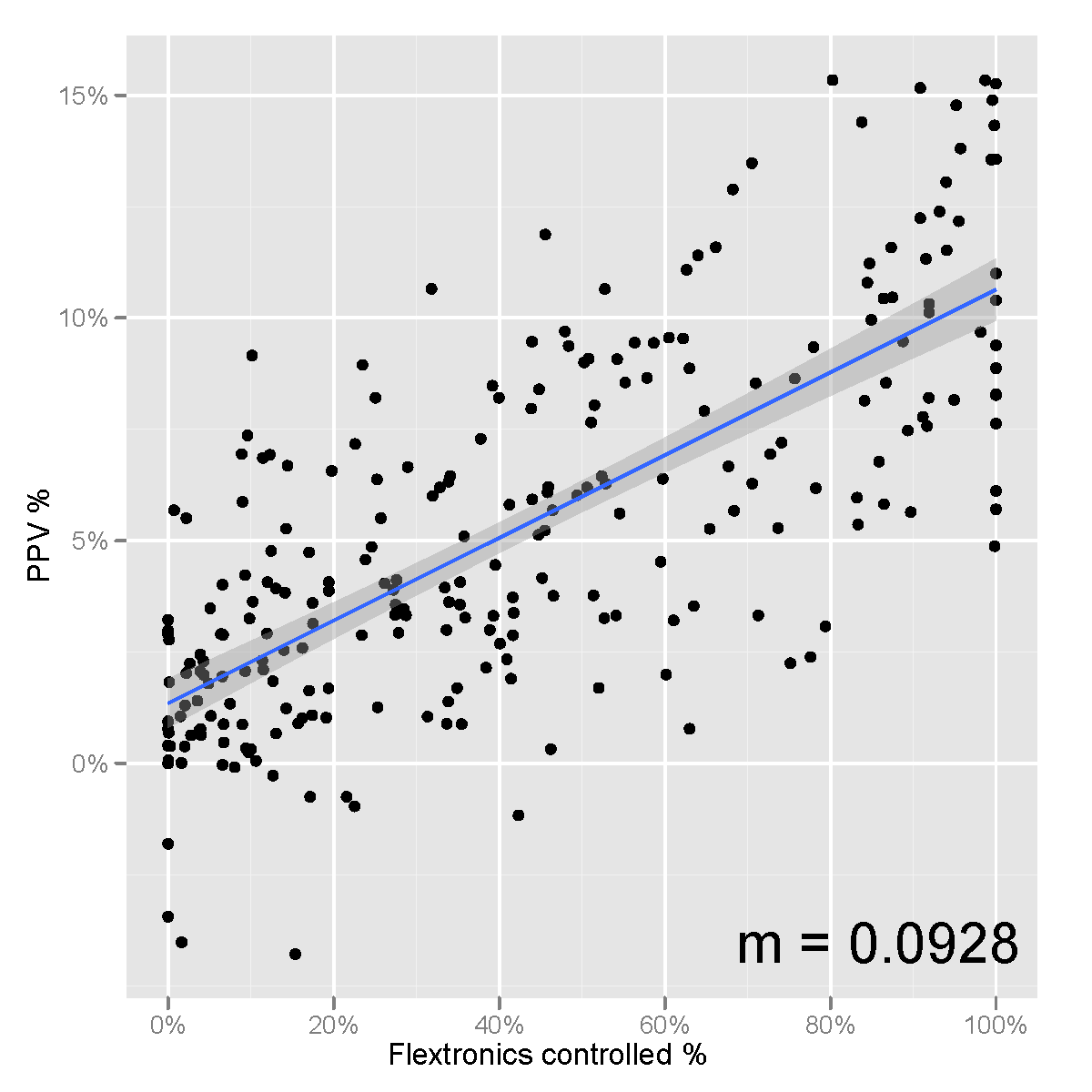
Figure Spend: top 20 customers; commodity groups; Flex controlled % (Source: Author)



The terms, conditions and relationship between the OEM and EMS vary widely. Some customers, such as Research in Motion (the company famous for the Blackberry device), give the EMS very little autonomy, while others, such as the Ford Motor Company, are hands-off with respect to sourcing decisions. The OEMs decide which components they will control. They negotiate pricing directly with suppliers for these parts and provide this pricing to the EMS to that the components can be purchased on behalf of the OEM.

Purchase Price Variance (PPV) is the difference between the materials purchase price and the cost to the customer. There is a positive correlation between Flextronics controlled spend and PPV, which is a major source of profit for the company. Each point on the scatterplot below represents an OEM customer.

Figure Scatterplot: Flex controlled % versus PPV % (Source: Author)



The gradient of the fitted line suggests that a 10.6%[[1]](#footnote-1) increase in source control yields 1% of additional PPV.

Commodity pricing is a service that Flextronics provides for its customers, and it is a source of profit for Flextronics. Global Commodity Managers (GCMs) have the power to award business to suppliers on behalf of those OEMs who cede control over sourcing. It is important that these GCMs are able to negotiate effectively with suppliers in order to generate PPV for Flextronics.

## Commodity Negotiations

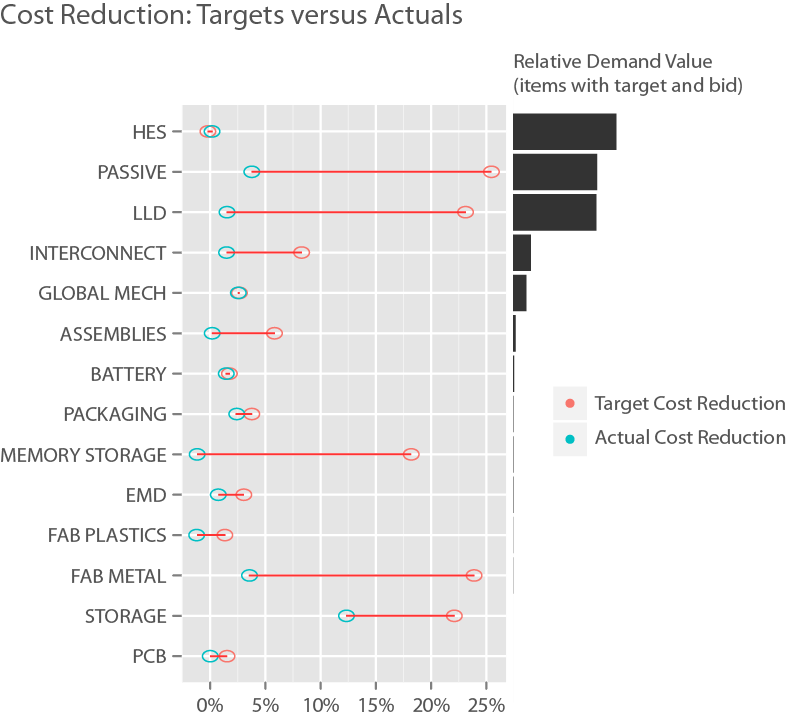
In commodity negotiations, representatives from the EMS and the suppliers communicate in order to review current agreements and make changes that result in a better outcome for both parties. One of the main tenets of negotiation is that both parties choose to negotiate because both believe they will get a better outcome by negotiating than they would by not doing so (Lewicki, Barry and Saunders 2010).

Negotiators usually engage the other party with either a distributive or integrative approach. A distributive approach is adversarial since the goals of both parties conflict. The focus is on dividing a fixed resource (e.g. simple price determination) where each party attempts to discover and influence the others resistance point. Information is only shared when it provides a strategic advantage. An integrative approach attempts to trade off differences in interest, assessment of the future, risk tolerance and time preference in order to create value for both parties. It is expected that both parties will share information and move away from their initial positions. Concessions are often packaged to avoid getting hung-up on single issues. The distributive approach, which is ideally suited to situations where the parties do not expect to work together in the future, is often overused and this leads to sub-optimal outcomes (Lewicki, Barry and Saunders 2010).

Since the supplier relationships are ongoing, it is important to maintain trust and minimize uncertainty. When an integrative approach is taken toward the negotiation, it indicates that both parties value the relationship and want to preserve goodwill. The quality of the supplier relationships vary from one supplier to the next: some are symbiotic while others are confrontational. In order to preserve healthy supplier relationship and obtain the best possible outcome, positions should be justified and realistic alternatives should be identified prior to the negotiation so that the negotiator knows how firm to be. Arguments generally perceived as fair include historical pricing data, competitor pricing (within the constraints of the non-disclosure agreements), and arguments on behalf of the ‘ (Lewicki, Barry and Saunders 2010). The best alternative to negotiated agreement (BATNA) is the minimum standard that a proposed agreement must meet (Fisher and Ury 1999).

It is important that negotiators set specific, verifiable, challenging yet realistic targets (Lewicki, Barry and Saunders 2010). The chart below shows the variance between target and actual price reductions for a recent negotiation period[[2]](#footnote-2).

Figure Target versus actual pricing reduction from suppliers (Source: Author)



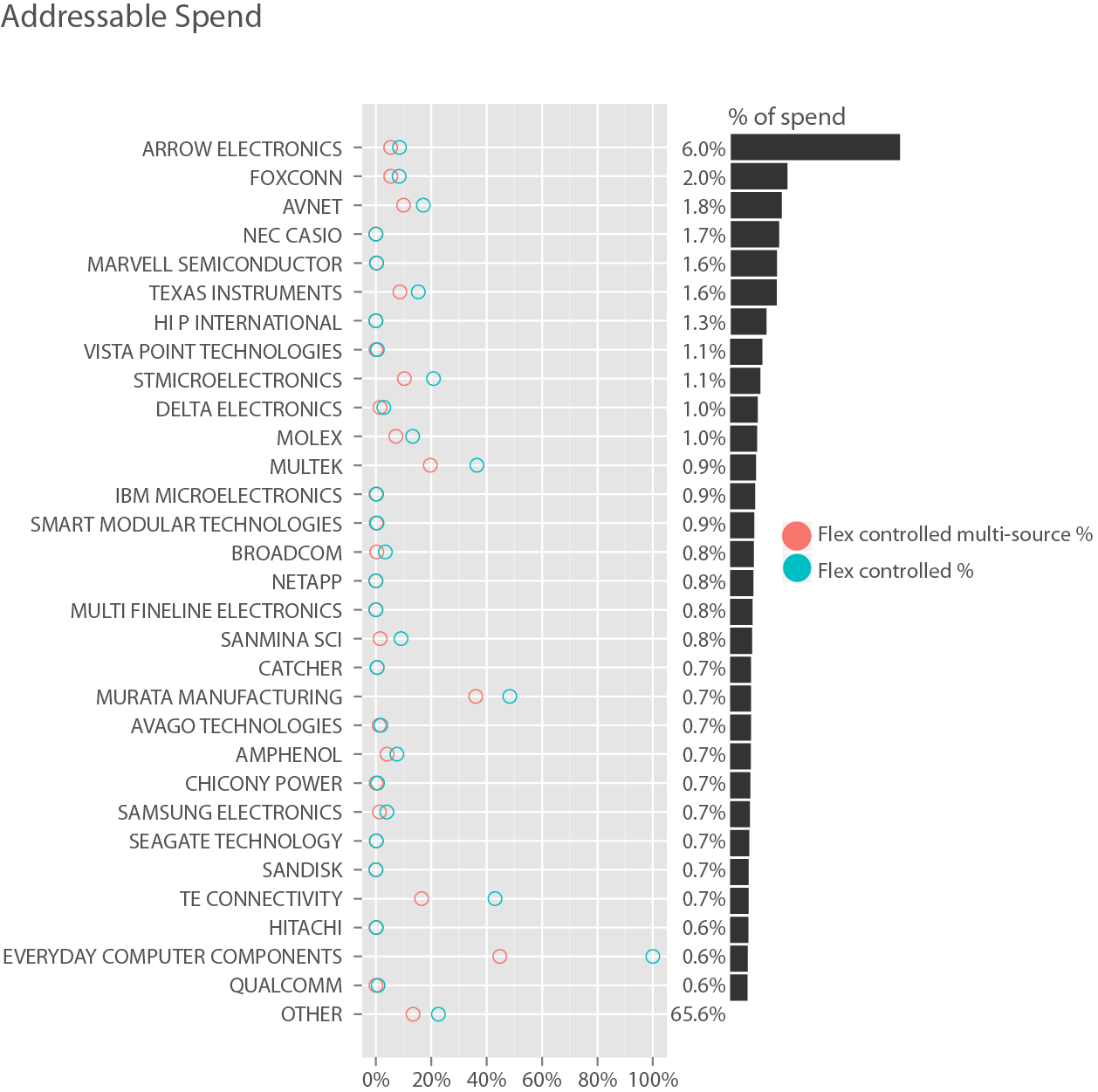
At Flextronics there is a significant disparity between target and actual negotiated prices. This may indicate that the target prices set by GCMs are not realistic, and suggests a need for more strategic target price protocol. An example of this was observed when one GCM described his current strategy for pricing targets. This negotiator uses the Excel function “randbetween” to create a random price reduction target between 12 and 18 percent. Although these targets are challenging, they are not verifiable or realistic. Consequently, the targets are often considered irrelevant by the supplier, the credibility of the commodity manager is diminished, and the target prices are disregarded altogether.

Negotiators need to carefully choose how much information to share, and how much they should trust the information they receive from the other party. They may selectively present facts that make them appear more powerful (Lewicki, Barry and Saunders 2010). Flextronics, for example, has a team dedicated to getting components designed into the OEM’s products, which increases the potential value of Flextronics to its suppliers.

When dealing with a supplier who has been appointed by the OEM as the single source for a part, the EMS has much less bargaining power. When the OEM has relinquished control and approved alternate sources, the EMS has greater influence with respect to terms and pricing. Within the top 30 suppliers for Flextronics, as with most of its suppliers, there is one portion of its spend that is controlled by the OEM and another that is controlled by Flextronics. However, the suppliers may be unaware of the level of control that the EMS has.

The chart below shows the level of sourcing control Flextronics has over its top 30 suppliers. The blue circles (Flextronics controlled %) refer to the percentage of spend with that supplier where the OEM has ceded source control to Flextronics. The red circles (Flextronics controlled % with multi-source) is the subset of the Flextronics controlled portion where an alternative supplier has already been approved by the OEM. Circles to the right of the chart (particularly red circles) indicate greater control over sourcing choices by Flextronics.

Figure Flextronics' sourcing control for spend with the top 30 suppliers (Source: Author)



## Structure of the thesis

In chapter 2, a literature review is provided about related topics such as business intelligence, and visualization tools for negotiations. In chapter 3, previous and alternative approaches are explored and the methodology for the project is described. In chapters 4 and 5, a visual negotiation dashboard tool is presented, analyzed and evaluated. Some limitations of the work are identified in chapter 6, and recommendations for improvement and further study are made in chapter 7. References are listed in chapter 8, which is followed by the appendix.

# BACKGROUND

## Visualization

Business people often have more difficulty with information communication than discovery (Wilkinson 2005). There is an idiom that says that a picture is worth a thousand words. Humans are, by nature, visual creatures and can easily differentiate between a variety of visual properties (Kosslyn 2006).

Although data visualizations can overcome constraints in our ability to process information, they are only effective if they can be interpreted (Kosslyn 2006). Before data can be effectively presented with graphics, the data must first be thoroughly explored in order to identify the salient facts (Yau 2011). Because data analysis is relatively inexpensive compared to data collection, we should not regularly fail to uncover the significant facts because we did not explore the data thoroughly (Cleveland, The Elements of Graphing Data 1994).

Statistical analysis often takes rich information and reduces it to a single number. When we do this, we lose some of the story (Cleveland, The Elements of Graphing Data 1994). A mean value of a distribution, for example, may be skewed to the right by outliers or may not represent a typical value if the distribution is bimodal (Leach 2006). Visualization provides us the flexibility to retain the original data, explore other quantitative relationships, and display statistical objects within the graphic (Wilkinson 2005). This enables us to spot relationships that we might not otherwise see.

The speed at which a graphic communicates a complete answer to a question is a measure of its efficiency (Bertin 2011). Although the information density can be extremely high (Cleveland, The Elements of Graphing Data 1994), it is important that the graphics are not unnecessarily complex (Booth, Colomb and Williams 2008). Combinations of text and graphics, whenever practical, are better than graphics alone since the points can be explicitly stated (Kosslyn 2006).

Readers of charts learn to interpret by recognizing patterns known as ‘graph schemata’. Conventions reduce the time taken for a user to read a chart. This has been likened to learning to drive: once a driver has mastered the controls nearly all their attention is focused on what is going on outside of the vehicle. For example, legend keys are traditionally located in the top-right corner of the chart. If they are placed elsewhere a chart may be less efficient since the user must create a new mental schema before looking at the data. For timeseries line charts, good news is traditionally shown as an increase, i.e. up and to the right, where bad news is shown as a decline. It is sometimes possible to express the data so that it conforms to this convention (Booth, Colomb and Williams 2008). Complex graphics can be used effectively providing that readers learn how to interpret them, though care must be taken not to make the graphic incomprehensible by overloading the short-term memory of the viewer (Kosslyn 2006).

Business users are more likely to trust and act on the data if they understand how the data was created, particularly when a graphic is extrapolated into the future (Kosslyn 2006).

Sorting categorical variables is a simple and effective way to visually provide insight into order. The position of a variable when ordered by size may provide more insight than arbitrary alphabetical ordering. For example, the ordered size categories [XS, S, M, L, XL, XXL] do not make sense when sorted lexically [L, M, S, XL, XS, XXL]. Most multivariate graphics are difficult to interpret when categorical variables are not ordered in a meaningful way (Wilkinson 2005).

The term ‘chartjunk’ refers to all non-data marks or redundant marks on a chart (Tufte 2000). Although Tufte asserts that we should try to maximize the data-ink ration, other authors refute this saying that, although Tufte’s graphics are beautiful, there is no scientific evidence to support his campaign against chartjunk (Kosslyn, 2006; Wilkinson, 2005).

In addition to the location in a two dimensional plane, data attributes can be expressed using a variety of ways: size, value, color, texture, orientation, and shape (Bertin 2011). Some attributes communicate better than others. Cleveland (1994) created a hierarchy of attributes to aid in the selection of data-to-object mappings, and the following list is ranked by ability to perceive differences in value from best to worst:

1. Position along a common scale
2. Position along nonaligned scales
3. Length
4. Angle/Slope
5. Area
6. Volume
7. Color

Although position along aligned scales ranked highest, bias and errors are introduced as points move away from the reference axis. Gridlines can help alleviate this issue (Kosslyn 2006).

Overplotting, where more than one point shares the same space, can be handled in a variety of ways. Cleveland (1994) experimented with a variety of shapes and determined that hollow circles had least problems with overplotting. Other approaches include binning, jittering and alpha blending (Wickham, ggplot2: Elegant Graphics for Data Analysis 2009).

Outliers can ruin graphics because they squash the bulk of the data into a small area of the chart. They should be investigated and may be removed if it is likely that they are erroneous values (Cleveland, Visualizing Data 1994). The frame of the chart should cover the range of the remaining values within the data (Wilkinson 2005). Unless zero is important, the scale should range from slightly below the lowest value to slightly above than the highest value (Kosslyn 2006).

If the data is extremely skewed with many overlapping values, the chart can be simplified by performing a logarithmic transformation (Cleveland, 1994; Wilkinson, 2005). There are three frequently used logarithm bases: 2, *e*, and 10. Base 2 is popular with computer scientists and mathematicians. Although natural logarithms, i.e. logarithms to base *e*, are the least intuitive because *e* (2.71828183) is a strange number, they enable rapid calculation of percentage differences between values. Base 10 is, most likely, the most intuitive logarithm for non-mathematicians because each increment represents an order of magnitude in the decimal number system (e.g. log10(10) = 1, log10(100) = 2, log10(1000) = 3, etc…).

The human eye is most sensitive to changes in vertical, horizontal and diagonal lines. It is easier to detect a difference between 44° and 45° to the horizontal than it is between 29° and 30°. To maximize the ability to differentiate changes in slope in line charts, axes should be adjusted to bank the absolute values to 45 degrees (Cleveland, Visualizing Data 1994).

The independent variable is usually placed along the x-axis and the dependant variable on the y-axis. The axes are sometimes transposed when the independent variable contains categorical data so that we are not forced to read the text labels vertically (Wilkinson 2005).

Color should be used to improve understanding of the facts within the graphic, and not to add glitz (Cleveland, Visualizing Data 1994). Color can be used to differentiate between categorical variables, and this is preferable to using different shapes in black and white (Cleveland, 1994; Kosslyn, 2006). Humans can generally only keep track of nine colors at once (Kosslyn 2006). Selection of color can be surprisingly difficult since each one is a combination of multiple attributes, e.g. hue, saturation and brightness. Selections should be scaled on one of the attributes (Wilkinson 2005). Wickham (2009) suggests using palettes that can be found online at ColorBrewer.org (Harrower and Brewer 2003).

Triangular coordinate plots can be used to compare combinations of three attributes e.g. price, lead time, and on time delivery (Wilkinson 2005). Step plots are used when values in a time series are held constant for periods of time, e.g. quarterly pricing, where the gradual undulations of a line chart would be misleading (Cleveland, 1994; Kosslyn, 2006). Line charts can be plotted using polar coordinates when making comparisons between seasons (Wilkinson 2005). Dot plots are used to display numeric values with descriptive labels. Graphical perception is more accurate with dot plots than other types of graphic for labelled data such as pie charts and bar charts. Dot plots can be enhanced by ordering the labels in a meaningful way. (Cleveland, 1994; Jacoby, 2006). Treemaps can be used to display a fairly large amount of hierarchical information in a relatively small space (Wilkinson 2005).

Figure Treemap example: materials procurement cost savings (Source: Author, 2012)



Although pie charts are popular, there is always a better way to present the data (Wickham, 2009; Booth, et al., 2008; Tufte, 2000; Cleveland, 1994). Three dimensional pie charts are even worse since the angular perspective makes it more difficult to compare wedge sizes (Wilkinson 2005). Three dimensional graphics should not be used to convey precise information (Kosslyn 2006).

## Tools

The vast majority of popular tools used to create statistical graphics do not contain sensible default options. Howard Wainer, the prolific statistician, puts this more bluntly:

“Thirty years ago I was enthusiastic and optimistic about the future of graphical use. I thought that software would be built with sensible default options, so that when the software was set on maximum stupidity (ours not its), a reasonable graph would result. The software should force you to wring its metaphorical neck if you wanted to produce some horrible, pseudo-3D, multicolored pie chart. Alas, I couldn’t have been more wrong.” (Bertin 2011)

Although the menus and dialogue boxes of a graphical user interface (GUI) are less bewildering to novices than the command line, many advanced users the command line because it offers greater control. If, in an exclusively GUI based system, all the options are visible to the user in forms and dialogue boxes, the user is constrained to a predefined framework. Since each feature requires its own dialogue box or options on a form, the functionality is often lacking (Wilkinson 2005).

## Business Intelligence

Since most supply chain management activities focus on exploiting information to improve processes, analytical capabilities can have a profound impact on operational performance (Valadares de Oliveira, McCormack and Trkman 2012; Trkman, et al. 2010).

Business intelligence returns value by making the right information available to operational decision makers in a timely manner. Data is acquired from various sources, integrated, and then analyzed to gain knowledge prior to taking actions. The ability to process information is therefore an important precursor to supply chain efficiency (Sahay and Ranjan 2008).

Despite capturing vast amounts of data and implementing sophisticated tools, there is no guarantee that companies can derive value from them. Generally speaking, companies who implement the most recent ‘bleeding edge’ technologies have below average profitability (Valadares de Oliveira, McCormack and Trkman 2012; Trkman, et al. 2010).

The cost and complexity of business intelligence technology deployment and acceptance is often underestimated for two main reasons. Firstly, solutions are often built from a variety of best in class components from different vendors that can be difficult to integrate (Sahay and Ranjan 2008). Secondly, changes to existing processes are often necessary to capture value from the investment and this needs to be managed (Williams and Williams 2003). Thirdly, technically savvy people often perceive technologies as being simpler to use than experience in the field would suggest (Lee, Kang and Kim 2007).

Information possession is not the same as information exploitation (Williams and Williams 2003). In order to have an impact, the acquired knowledge must translate into action, so there must therefore be a clear connection between the information and how it will be used to improve performance (Sahay and Ranjan 2008).

Operational data is collected and processed in various forms. The complexity of integrating all the disparate data sources, which increases with diversity and volume, often encumbers the organization’s ability to make sense of the data and transform it into information that can be used to support decision-making. Consequently, companies with more mature data integration and analytical capabilities are more profitable than their less mature competition (Trkman, et al. 2010).

## Visualization for Negotiations

Information visualization is the transformation of data into a visual form by mapping attributes of the data to geometric objects (e.g. points, lines, shapes, etc.). This creates an internal mental model in the individual that that provides insight and understanding that may not otherwise be immediately apparent from looking at the raw data in tabular form (Parush, Hod and Shtub 2007). The mental models are frequently used to aid in decision-making (Sackett and Williams 2003).

Although visualization has been around for many hundreds of years, there is still not a simple set of rules that can be followed to select the best method for visualizing a given dataset (Sackett and Williams 2003; Monzani, Bendahan and Pigneur 2004).

The human brain has separate memory storage for verbal, spatial, and visual objects (Lohse 1997). Visualization enables us to expand our working memory for a given cognitive task by transferring some of the mental workload used for information acquisition to the visual perception system. Cognitive bandwidth increases when mental processes are substituted with perceptual processes that take place almost automatically. This has the effect of compressing the information being processed by the reader, which, during a negotiation, reduces the amount of effort dedicated to problem solving (Eppler and Platts 2009; Weber, Kersten and Hine 2006).

Visualization enables us to identify hidden patterns and relationships by comparing different attributes (Monzani, Bendahan and Pigneur 2004). The additional memory can be used to track and compare more details. Patterns and structures are easier to recognize (e.g. maxima, minima, smooth curves and discontinuities). Directional order is more apparent in graphs when compared with tables. The reader can also focus more attention on problem solving instead of remembering facts (Weber, Kersten and Hine 2006; Larkin and Simon 1987).

Data visualization will improve decision-making only when cognitive resources are constrained (Lohse 1997). Because Enterprise Resource Planning (ERP) systems store data in many separate tables, users are often forced to collect data from several places in order to gather enough information to make a decision. Since expert ERP users can become quite adept at rapidly navigating the various screens and tables, data visualization is most appropriate for complex tasks and to support users with less domain knowledge (Parush, Hod and Shtub 2007).

Visualization reduces information overload: less time is spent searching and evaluating information. A study on the relative strengths and weakness of graphs versus tables showed that graphics enabled users to solve problems faster than tabular data. However, the increase in speed came at the expense of accuracy since the spatial judgment is more error prone than simply reading values from tables (Vessey and Galletta 1991).

Text is suitable for comparing two or three values, tables can be used to compare up to twenty values, but beyond that it is necessary to represent data graphically (Monzani, Bendahan and Pigneur 2004). Procurement negotiation data is both voluminous and complex. It is therefore an ideal candidate for visualization since a large textual explanation is necessary to communicate and present what is immediately obvious when presented graphically (In and Roy 2001; Weber, Kersten and Hine 2006; Eppler and Platts 2009).

Surprisingly, ERP systems rarely integrate data from multiple sources into coherent and comprehensive graphical displays in order to support decision-making (Parush, Hod and Shtub 2007). This is, most likely, due to the lack of a standardized approach to visualization. There are a myriad of different ways to represent any given dataset and the ‘best’ one very much depends the user’s expectations and what they want to convey (Monzani, Bendahan and Pigneur 2004).

Negotiators sometimes have a tendency to see only what they want to see. Prior knowledge has a strong influence on their perception. If each party in a negotiation has a different understanding of the current situation, there is a potential for conflict that may be irreconcilable within the negotiation (Weber, Kersten and Hine 2006). Because visualizations are easier to understand than verbal and written information, there is less opportunity for conflicting interpretations (Weber, Kersten and Hine 2006). Visualization was shown to be an effective method for understanding complex negotiation issues (In and Roy 2001).

Shared mental models however, which can be built through the presentation of information prior to and during the negotiation, are useful because they create a common and unambiguous perception of reality. These models also have a positive impact on the social climate, which increases the likelihood of agreement and a satisfactory outcome. In addition, there are social and emotional benefits that may lead to a better outcome in a negotiation, can reduce the dominance of the more imposing negotiators, and reveal areas of contention (Eppler and Platts 2009).

There is clearly a competitive advantage to the negotiating party with the greatest influence over the availability of information and its presentation in a shared mental model. Visualization plays a major role in building a mental model, which in turn enhances group cohesion, promotes consensus, and increases negotiator satisfaction (Swaab, et al. 2002).

Negotiation support systems are designed to make decision-making easier and to support communication. A decision support system should increase the capacity for information processing (Swaab, et al. 2002). Because visualization increases the amount of information that can be tracked and compared, there is a benefit to including visualization in a negotiation support system.

While graphs offer many advantages over tables, users may be reluctant to deviate from the reports they are used to receiving (Vessey and Galletta 1991). Microsoft Excel has become the de facto standard in many companies for the analysis and distribution of data.

Color is preferable to symbols to identify items in a chart because the reader can focus on a specific color and mentally disregard the others. If a good color scheme is chosen, users can differentiate very quickly between shades. Colors can be decoded three times faster than monochrome symbols (Lohse 1997).

In addition to color, gridlines improve speed and accuracy when reading a chart. They also reduce the amount of effort required to interpret the chart (Lohse 1997).

In order to fully realize the social value, visualizations should be designed to be interactive (Eppler and Platts 2009). Google has some excellent visualization objects that can be interactive. Flextronics’ commodity management team, however, felt that the negotiations were too fast-paced to use interactive visual analytics during the negotiation.

# METHODOLOGY

## Previous Methods

In 2008, Flextronics purchased the SAP Business Objects business intelligence suite as a replacement for its aging IBM Cognos toolset. At the time, there was a perception that the old system was outdated, lacking in functionality and due for replacement. Flextronics’ IT department gathered requirements from business users prior to selecting Business Objects and there was general consensus that the new system could do everything that the previous system could do, as well as offer some new capabilities.

In reality, there were some architectural constraints that would have resulted in a loss of functionality, had Flextronics switched from their legacy system to Business Objects. The scope of the project to migrate to Business Objects was cut and the current perception is that the toolset was to blame. Since the failed implementation, the company has been evaluating and exploring other business intelligence tools. Flextronics is currently engaged in a proof of concept with Microstrategy. The cost is expected to exceed $1 million.

## Options Available

Although Excel is an extremely popular tool, it is not suitable for statistical analysis when results matter. For example, there are major flaws with Solver, exponential smoothing, some distributions, the random number generator, the t-test, and order of precedence[[3]](#footnote-3) (McCullough and Heiser 2008).

Excel also has inferior graphical capabilities. The default options make it difficult for non-expert users to produce quality graphs because they tend to use solid shapes to mark points in scatterplots and bubble charts instead of circles or semi-transparent points, allow the user to easily create distorted 3D charts at the click of a mouse, and clutter the graphic with chartjunk (e.g. coloring, shadow, crowded axis labels). The poor default options rely on the user, who may not have a solid understanding of the principles of statistical graphics, to make good choices (Su 2008).

The existing negotiation toolset at Flextronics publishes pricing data to commodity managers in large Excel spreadsheets. The spreadsheets are extremely large and, so far, no attempt has been made to guide the commodity managers to the facts that will support their negotiation arguments. Given the amount of data and the limitations of working memory in the human brain, the negotiation dataset was a good candidate for visualization.

## Justification of Choice

Commercial business intelligence tools are expensive and time consuming to deploy. Decisions that impact thousands of employees and have large capital expenditures attached, especially in an economic downturn, may take a long time. Open source software was a good fit for this project because it is free to download and use. The purpose of the project was to evaluate the effectiveness of data visualization technologies to supplier negotiations at Flextronics. It might have been more difficult to get approval to proceed from management if the project had a large associated cost and no guarantee that it would return value to the company.

The open source statistical analysis language, R, has a number of extremely versatile packages for visual analysis such as ggplot2 (Wickham, ggplot2: Elegant Graphics for Data Analysis 2009), lattice (Sarkar 2008), and googleVis (Gessmann and de Castillo 2011). A growing number of large companies such as Shell, Google, Facebook, Pfizer, Merck and Bank of America use R for analytics. Revolution Analytics commercially support a subset of the functionality. Oracle’s new Exalytics BI Machine performs statistical analysis using R (Information Week 2011).

Because Business Objects and R target different groups of end users, it is not entirely fair to compare the two. The comparison, however, does show that sometimes businesses spend a lot of money and fail to obtain the desired results. It also shows that it is sometimes possible to spend no money and get exactly what you want.

At the time of writing, there were 3,675 contributed packages to extend the basic functionality of the R language to perform various analytical tasks (R Development Core Team 2012). It is often the case that, when some obscure analytical functionality is needed, a developer has already contributed code to the archive of packages. There is an active user community that can be accessed via the R-Help email list[[4]](#footnote-4), local Meetup groups and LinkedIn groups[[5]](#footnote-5). In addition to the excellent built-in documentation, there is a vast amount of documentation available online and over 100 books on R have been published (R Foundation for Statistical Computing 2012).

Although R presents a steeper learning curve than analytical tools with a graphical user interface, the effort is rewarded with some very advanced statistical and graphical capabilities, thanks to the plethora of contributed packages. There are several online galleries that display various infographics created in R and their associated source code[[6]](#footnote-6).

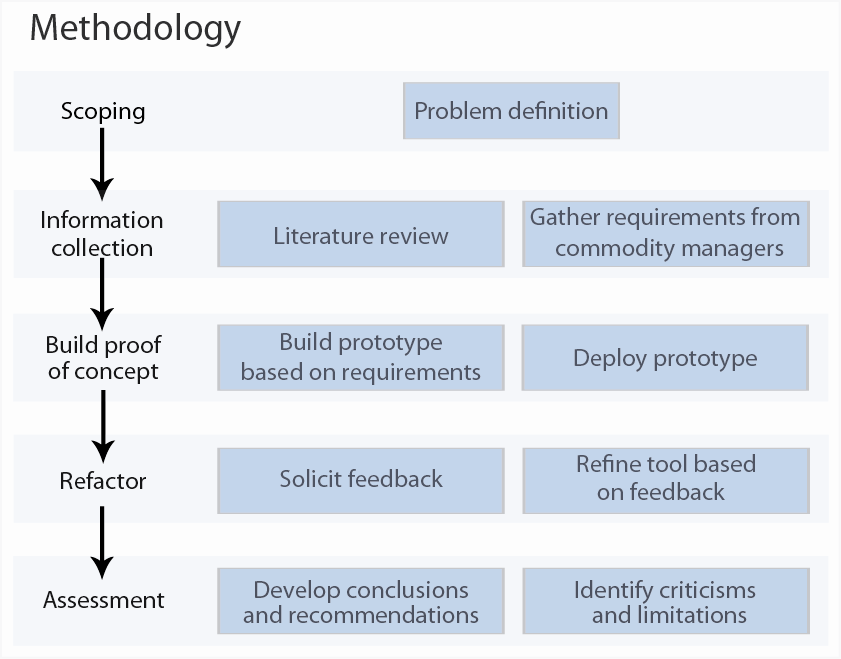
Another benefit of using R over Business Objects is the ability to intelligently combine data from multiple datasets without downloading large tables from multiple sources and joining them at the client. In the existing architecture at Flextronics it is not possible to join large tables. In R, an SQL query can be dynamically created based on the results from a previous query. There is currently a million-cell limit imposed in our Business Objects implementation and, although might sound like a lot of data, in reality it is less than 16 columns of a spreadsheet with 65,000 rows. Most of the tables in the data warehouse that contain transactional data have millions of records.

In addition to the built-in graphics functionality in R, there are a number of excellent packages that simplify and extend the basic functionality. The two most popular packages are ggplot2 (Wickham, An implementation of the Grammar of Graphics 2012) and lattice (Deepayan 2011). ggplot2 was chosen because the default settings do not contain unnecessarily vibrant colors and the logic behind the syntax is better documented.

The Adobe Portable Document Format (PDF) format was selected as the publication format because all the commodity managers have the Adobe PDF reader installed, and the graphics are vector based which means that images can be scaled to any size without any negative impact to the quality.

The approach to the methodology is shown below (see figure 8).

Figure Approach to methodology (Source: Author, 2012)



# ANALYSIS

## Stages in method

### Definition

The objective was to produce a tool using R to determine whether visualization is useful for negotiations at Flextronics. The idea to investigate visualization stemmed from dissatisfaction from the existing negotiation toolset, shortfalls of Flextronics’ Business Objects implementation, and awareness of some of the technical capabilities of the R language.

### Literature review

A review of the existing literature on data visualization, visualization for negotiations, and business intelligence was carried out. Most of the academic papers and journal articles were found by using various relevant search terms on the EBSCOhost, ScienceDirect, and Google Scholar databases. A number of relevant books were also read.

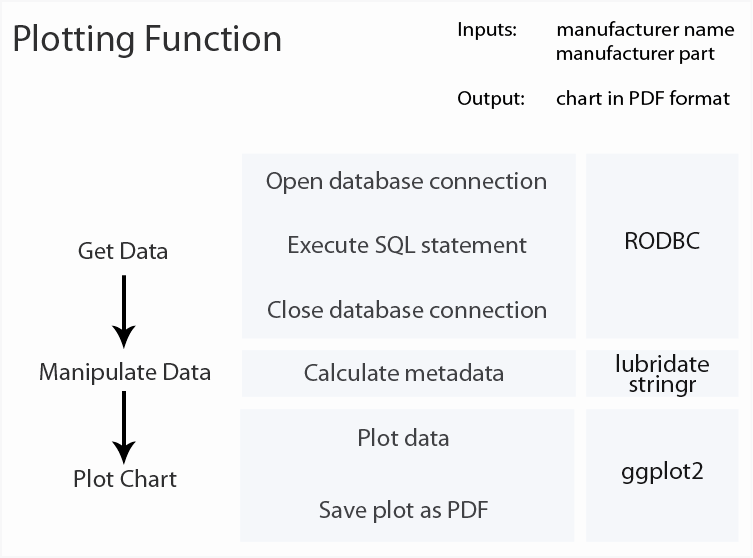
### Build prototype

The prototype dashboard was built and, in the process of building the prototype, a number of problems and solutions to those problems were uncovered. In this section we will start with an overview of the prototype dashboard design and then provide some of the solutions to common issues.

#### Overview of prototype

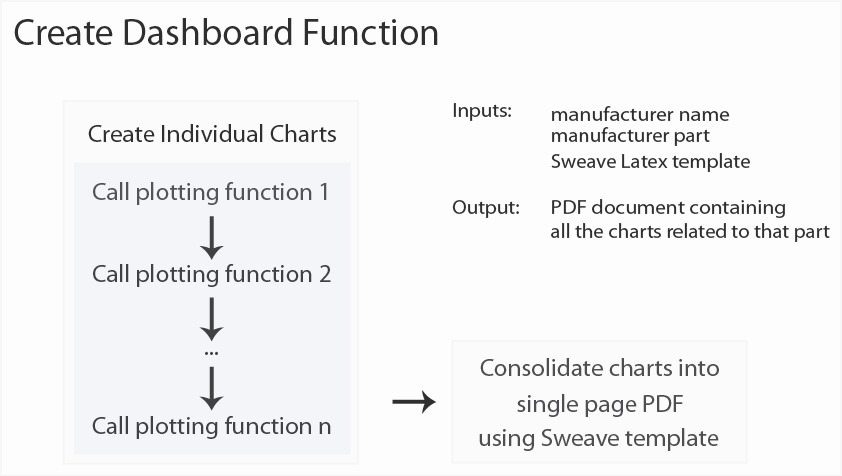
The creation of a graph from a datasource is a basic building block of the dashboard tool. A template function design was created with components to connect to a database, perform calculations, plot a chart, and save the chart to a PDF. The diagram below (see figure 11) shows the data-to-chart transformation process and identifies the packages used.

Figure Plotting function template, written in R (Source: Author, 2012)



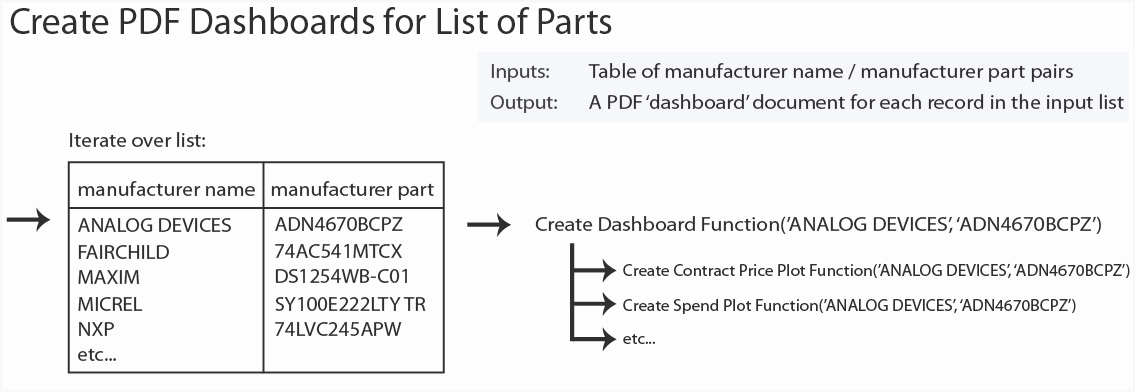
The dashboard for any given part contains multiple charts. The function that creates the dashboard iteratively calls the various plotting functions. Once the individual charts have been created they are published into a single PDF document using an R package called Sweave[[7]](#footnote-7), which uses a Latex[[8]](#footnote-8) template to combine all the individual components into a single PDF document for a given manufacturer part (see figure 11).

Figure Dashboard function (Source: Author, 2012)



During the course of a negotiation, Flextronics’ commodity managers may discuss hundreds of individual parts. The dashboards need to be prepared ahead of time for two reasons. Firstly, the data warehouses do not currently respond quickly enough to prepare the dashboards at the time of consumption since it currently takes more than two minutes to produce a dashboard for a single item. Secondly, the negotiators might not have access to Flextronics’ network. The commodity manager may specify the parts in a list, or they may provide a set of rules, e.g. top 100 parts by spend, to identify the parts to include.

Once the target list of parts has been identified, the individual dashboards are creating by iteratively calling the dashboard creation function and passing the manufacturer name and part as arguments, which in turn get passed as arguments to the individual plotting functions (see figure 11).

Figure Create dashboards for a list of parts (Source: Author, 2012)

#### Solutions to common issues

When a SQL query is executed the results are returned to a table, known as a data frame, in the R application. Once the data is in the R application, it can be transformed and manipulated into a form suitable for plotting. These data manipulation activities may entail joining dataframes from different sources, the calculation of metadata (e.g. ranking within a group, calculations based on existing columns), the creation of factors where non-lexical sorts are required, and subsetting to remove irrelevant records. The following section will provide specific details on how to perform these tasks using R.

R code can be stored and executed from text files. Although it is possible to write code in any text editor, there are a number of free text editors available with syntax highlighting for the R language[[9]](#footnote-9).

##### How to get the data

ODBC (Open Database Connectivity) is a common interface that is used to connect to a variety of different data sources. All the popular operating systems (i.e. Windows, Mac, Linux) and databases (e.g. Oracle, SQL Server, MySQL, DB2, Informix, Sybase, Teradata, etc…) support ODBC connectivity. The data source name (DSN) for each connection is defined along with any special connection parameters. Once ODBC connections have been configured in the operating system, they are available to R using the RODBC package (Ripley 2012).

The code below opens an ODBC connection, executes a query, and returns the results to a data frame:

#Connects to a database, executes a query, and returns the results.

library(RODBC)

database\_connection <- odbcConnect("ODBC\_CONNECTION\_NAME", uid = "USER", pwd = "PASS")

query <- paste(

"select ",

" \* ",

"from table\_name ",

sep = "\n"

)

query\_result <- sqlQuery(database\_connection, query)

close(mdss\_connection)

Code comments are added by putting hash symbol (#) at the start of the line. This tells the interpreter to ignore the rest of the line. The library(RODBC) line loads all the functions and classes from the RODBC library into memory.

The less-than sign followed by a hyphen (<-) is an assignment operator. In the example below, the odbcConnect object is being assigned to the name database\_connection, the SQL statement is assigned to the name query. The sqlQuery function takes the connection object and query as arguments and executes the query against the connection. The results are assigned to the query\_result data frame.

Although whitespace such as carriage returns and tabulations are meaningless in R, it is good coding practice to use whitespace to keep code readable. The time spent formatting code in a logical and legible manner is usually recouped when debugging. Because the parentheses around the SQL statement are aligned, it is obvious where the block of SQL starts and ends.

##### Write functions to reduce repetition

As a general principle, when looking at pages of code that have been written, if there are similar lines of code repeated throughout, it may be possible to create a general case function and eliminate the repetition. Because each line of code has the potential to contain an error, fewer lines of code result in fewer errors. Also, code is easier to maintain when behavior is defined in just one place rather than being scattered throughout.

Functions also simplify the design of the application: they can be thought of as a black box that contains zero or more inputs and outputs. This enables the developer to abstract chunks of functionality and not be distracted by the lower level details.

There are several user-defined functions that produce PDF charts for a given manufacturer part. The generic structure for the plotting function is as follows:

createSomePlot <- function(manufacturer\_name, manufacturer\_partno)

{

[code removed]

}

The two arguments, the manufacturer name and the part number, are passed to the plotting function. The call to the function is made:

createSomePlot(‘ANALOG DEVICES’, ‘ADM202JRM’)

Inside the plotting function, a SQL query is dynamically written by pasting the manufacturer name and manufacturer part arguments into the criteria literals of the ‘where’ clause in the query:

query <- paste(

"select ",

" \* ",

"from some\_table ",

paste("where manufacturer\_name = '", manufacturer\_name, "'", sep = ""),

paste("and manufacturer\_part = '", manufacturer\_part, "'", sep = ""),

sep = "\n"

)

##### Ranking over groups

The charts sizes in the dashboard are hard coded in order to keep the dashboard format consistent. If every dashboard had a slightly different layout the commodity manager would need to spend additional time searching for information and would be less likely to use the tool. The number of records returned by a SQL query may vary and, if more records are returned than can be typeset in the fixed space, the display may end up being overcrowded and illegible.

At the time of writing the YAGEO resistor RC0603FR-0710KL was sold to 161 different OEM’s for 136 different prices. It is neither feasible nor desirable to show this level of detail in the standard costs dotplot since only the lowest prices could be used as a negotiation argument.

In order to return, for example, the four lowest prices for each category a split-apply-combine strategy is used to add the ranking over each group:

data <- ddply(data, .(PRICE\_TYPE), mutate, RANK = rank(USD\_PRICE, ties.method = "random"))

The ddply function first splits the data dataframe into groups by price type. Then the mutate function, which adds the rank to each group, is applied. Finally, all the individual groups are reassembled into a dataframe and this is assigned to the name of the original dataframe, data.

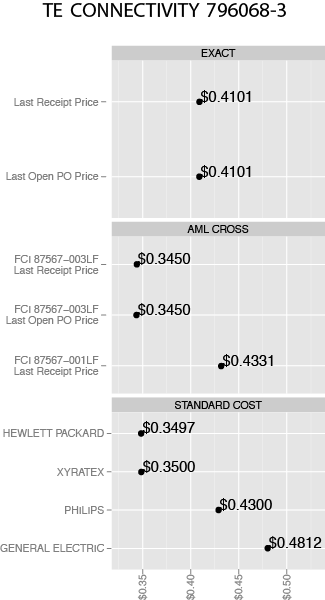
Once the rankings have been applied, a filter can be applied to the dataframe using the subset function to return the four lowest prices for each group:

data <- subset(arrange(data, MATCH\_TYPE, RANK), RANK <= 4)

##### Plots with ggplot2

Once the data has been extracted and transformed into a format suitable for plotting, a chart is plotted using the qplot function. In its most basic form, qplot requires 3 arguments: x, y, and data. Most plots are more complex and often have many additional parameters applied to build a plot. Figure 12 shows a multi-faceted dotplot containing recent purchase prices (top panel), recent prices of possible alternate parts (middle), and the prices that Flextronics charges its OEM customers for this part (bottom). A commodity manager could rapidly determine that this part is being purchased for $0.4101, that there may be an alternative equivalent part that is currently being purchased for $0.3450, and that Flextronics currently sells this part to Hewlett Packard for $0.3497 and Xyratex for $0.3500.

Figure Prototype ggplot2 faceted dotplot (Source: Author, 2012)



The qplot command to produce figure 12 was:

qplot(USD\_PRICE, DESCRIPTION, data = data, geom = "point", label = dollar.4(USD\_PRICE)) +

facet\_wrap(~MATCH\_TYPE, ncol = 1, scales = "free\_y") +

scale\_x\_continuous(formatter = "dollar”) +

opts(axis.text.x = theme\_text(angle = 90, hjust = 0, colour = "grey50")) +

labs(x = "", y = "") +

geom\_text(hjust = 0, vjust = 0)

ggplot2 has an enormous array of functions. To learn more about the practical use of this package consult the book written by its creator, Hadley Wickham, entitled ‘ggplot2: Elegant Graphics for Data Analysis’.

Once a graph has been plotted it can be saved as a PDF document. The dimensions of the graphic are specified as arguments to the save command:

ggsave("aml\_cross\_dotplot.PDF", height=8, width=5)

##### Handling Errors

Because the source data was often incomplete or contained bad data, it was not always possible to produce every chart for all the parts processed by the tool. By default, Latex ignores missing graphics and, because the charts are typeset to fit neatly together, a missing chart causes the dashboard to look jumbled. Also, a consistent layout will reduce the search time for the reader.

Error handling was added to maintain consistent dashboard formatting. The try function attempts to run some code, and allows for a graceful recovery if an error message is generated. In the code snippet below an attempt is made to plot and save an 8 inch by 5 inch chart and, in the event of an error, a blank 8 inch by 5 inch PDF file is put in its place:

try(

{

[code removed]

ggsave("aml\_cross\_dotplot.PDF", height=8, width=5)

},

)

createPdfH8W5PlaceholderFile("aml\_cross\_dotplot.PDF")

##### Reordering Factors

Factors in R are used to store categorical data and, by default, factors are sorted alphabetically. It rarely makes sense to perform lexical sorts because the order is pretty arbitrary. There is no reason that we should always consider parts from 3COM first and those from Zarlink Semiconductor last. In order to minimize the search time and show the order, the categories in the dotplots are sorted by price:

data$DESCRIPTION <- with(data, reorder(DESCRIPTION, USD\_PRICE))

##### Monitoring Progress and Debugging

Unexpected results sometimes occur during development, and this slows down the dashboard creation process. The browser function in R can be used to pause code execution at some point and provides an opportunity for the user to interrogate the values of the variables. This is useful when trying to diagnose some unexpected behavior. To pause execution and return an R command prompt simply place insert the function browser() in the desired place.

Status updates can be written to the console window to show the progress of the job. For example, a commodity manager might send a list of parts to process and ask for an estimated completion time. Functions can provide run-time status updates using some variation of the following code snippet:

cat(paste(now(), ": Now running createSpendTrendPlot...", sep = "", "\n"))

flush.console

The dashboards generate a lot of SQL queries against the various databases since each plot is based on the results from one or more SQL queries, each dashboard has multiple plots, and each part has its own dashboard. The volume of additional queries could, potentially, put a significant load on the data warehouse. The computational cost of running a query can be influenced by a number of factors such as the presence of indexes in tables, the size of the tables, and the SQL syntax. Queries can often be rewritten in many different ways so that they are logically identical but syntactically different. It is possible to automate the rewriting and performance testing, i.e. ‘tuning’, of SQL statements using SQL Optimizer[[10]](#footnote-10) by Quest Software.

As a first step towards targeted query optimization, some query performance logging was added. A dataframe was created to hold a start time, an end time, and a description:

startTime <- now()

endTime <- now()

description <- "Starting"

performanceLog <- data.frame(startTime, endTime, description)

Then timestamps were recorded at the start and end of blocks of code, before appending the record to the performance log dataframe:

startTime <- now()

[code removed]

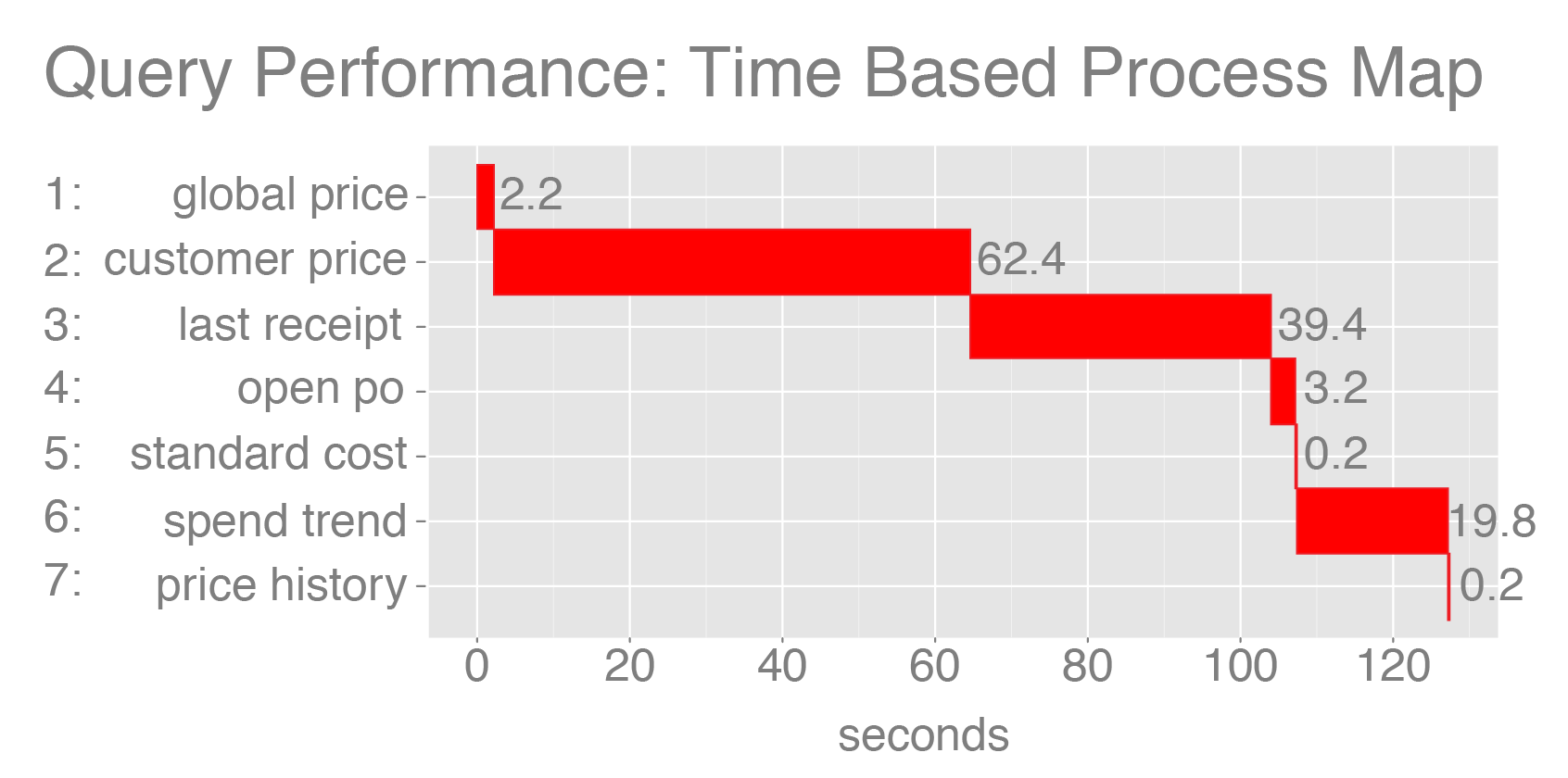
endTime <- now()

description <- "query: global price"

performanceLog <<-[[11]](#footnote-11) rbind(performanceLog, data.frame(startTime, endTime, description))

The performance log was analyzed and published in the form of a time based process map (see figure 14). This shows the sequence of events, the end-to-end time, and identifies the slower parts of the process.

Figure Query performance: a time based process map (Source: Author, 2012)



From the process map we can see that, for the prototype dashboard, it takes just over two minutes to run all the queries for a single part.

##### Joining Dataframes

It is sometimes necessary to join dataframes from different sources. At Flextronics, the negotiation dataset is stored in a SQL Server database and the data warehouse is Oracle. In order to get a complete picture of historic and future pricing, results from both sources need to be joined by a common key. The key fields can be explicitly specified or, in the snippet below, by all the common columns:

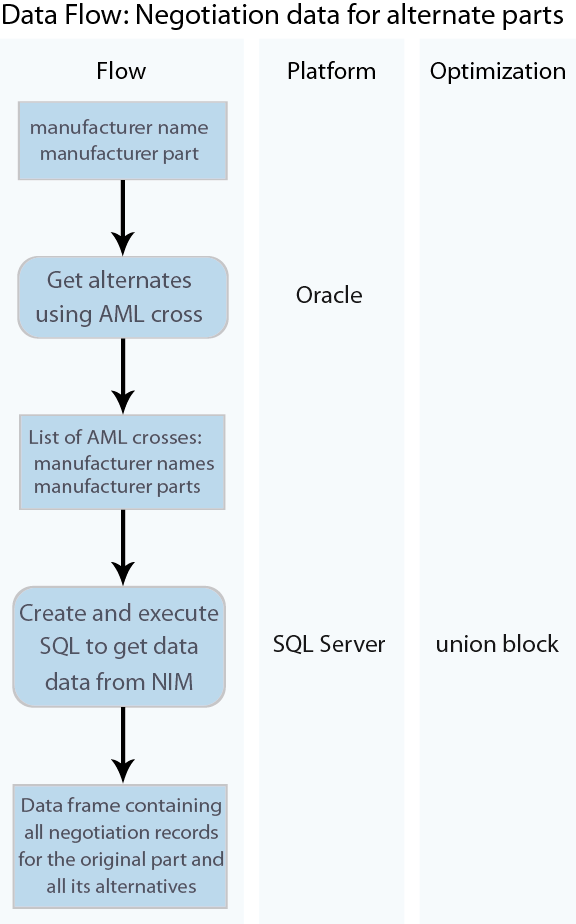
data <- join(data\_mdss, data\_nim, by = intersect(names(data\_mdss), names(data\_mdss)), type = "full"[[12]](#footnote-12), match = "all")

##### Multi-step Queries

In Flextronics’ Business Objects implementation, which is typical of most business intelligence tools, joins between different data sources are done at the client[[13]](#footnote-13). Sometimes there are restrictions on the amount of data that can be downloaded, and even if the tool allows the user to download large tables, it can take an extraordinarily long time.

The R language, with its enormous collection of packages, offers a lot more flexibility than purely SQL (Structured Query Language) or graphical user interface based query tools. A query can execute against one database and then use the results to write an optimized SQL statement against another database. This eliminates the need to move lots of data around in order to process a subset of parts. The query optimization might take advantage of platform specific capabilities. Flextronics’ SQL Server, for example, is extremely slow when filtering using the ‘in’ keyword, but an equivalent statement can be written with a series of unions and an inner join (see figure 15). Although Business Objects can use the results of one query as a filter for another, the ‘in’ keyword would be used and, in Flextronics SQL Server environment, this would result in unacceptably slow execution times and therefore could not be used to produce individual dashboards for thousands of parts.

Figure Multi-step queries with platform specific optimization (Source: Author, 2012)



##### Applying a Function to Elements in a List

The dashboard function creates a number of charts for just one part. When commodity managers meet with suppliers they may discuss hundreds of parts. The sapply function can be used to apply a function to all the elements in a list.

sapply(manufacterNamePartList, createDashboard)

### Deploy prototype

The initial prototype was built using five simple plots that were based on frequently used negotiation arguments:

1. Dotplot: most recent active pricing, i.e. current negotiated price, current price negotiated by the OEM, open purchase order (PO) price, last receipt price.
2. Dotplot: most recent active pricing for competitor parts.
3. Dotplot: standard costs, i.e. the price that OEM customer pays Flextronics.
4. Line chart: contract price trend.
5. Line chart: quarterly spend trend.

The prototype was distributed to commodity managers in order to solicit suggestions for improvement. Internal social media tools were also used to generate feedback within the company from a broader audience than immediate colleagues. A sample dashboard along with a description of its purpose was posted on Flextronics’ internal versions of Facebook[[14]](#footnote-14) and Youtube[[15]](#footnote-15).

Requests for dashboards came in two forms: a list of parts supplied in an Excel spreadsheet, or a selection of parts based on a rule, e.g. 'the top 100 parts for the interconnect commodity'. The form of the request, downstream of the initial part selection, had no impact to the tool since both the spreadsheet and the rule-based query enter the downstream process as a dataframe containing a list of parts.

### Solicit feedback

The prototype dashboard was shared with immediate colleagues via several calls using Flextronics’ Cisco MeetingPlace conferencing system. Although it was expected that most of the feedback would come from immediate colleagues who negotiate with supplies, feedback from other sources was not precluded. The use of social media resulted in feedback from colleagues that would have otherwise not been involved in this project.

It would have been interesting to solicit feedback, possibly using LinkedIn groups, from outside the company. A decision was made not to do this because Flextronics, like most companies, are sensitive to their employees posting information on social media[[16]](#footnote-16).

All the feedback was consolidated into a document (see Appendix A) and evaluated. Where possible, the ideas, suggestions and comments were translated into requirements, though not every suggestion could be translated into a legible graphic.

The nature of the project was exploratory. It was not clear, at the outset, what the dashboard would contain or look like. This made it difficult to capture feedback using a standardized form or questionnaire. Because the responses were less structured, the feedback was of a more qualitative rather than quantitative nature.

### Refine tool

If time had permitted, there would have been several iterations refactoring the dashboard tool.

#### Consolidating PDF’s into a single file

The original dashboard sample comprised of 50 individual PDF documents in a zip file. The intention was that the commodity manager could sort the files alphabetically and open the PDF for a specific part when they need to. One of the commodity managers requested that all dashboard files be consolidated into a single file for the sake of convenience. For the sake of expediency, this was done using Adobe Acrobat Pro. Although it would have been possible to consolidate the files programmatically and eliminate the dependence on third-party software, it was more important at the time to respond to the commodity managers' requests. It seemed more likely that commodity managers would be forthcoming with their feedback if their input was implemented quickly even if, behind the scenes, the solution was not as neat and self-contained.

The consolidated document took slightly longer to open initially, but once open, commodity managers were able to navigate the data faster than they could with individual files by using the 'find' function to search for parts. The five charts in the prototype dashboard took about 30k per part. The consolidated sample with 50 dashboards was about 1.5MB. A file containing dashboards for a few hundred parts could easily be transferred via email. The relatively small file size was another benefit of using vector rather than raster graphics.

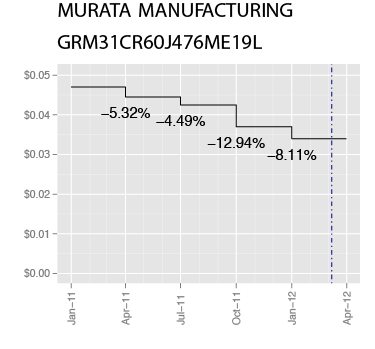
#### Future prices: bids and contracts

The initial prototype contained only historical data. This was a deliberate choice because the system that holds the future pricing is fragile and has a lot of visibility. If something went wrong, the negotiation dashboard project could easily have been scuttled before any potential benefit could be shown. Predictably, the commodity managers asked for the future pricing to be added to the dashboard and this was done. There was no detrimental impact to the performance of the SQL Server database.

#### Annotating Charts

One of the commodity directors, Ran Franco, suggested that the charts would be easier to read if the key features were annotated on the chart. The cost reduction percentages were added to the historic pricing step plot (see figure 15). The vertical blue dotted line marks the current date as a reference point.

Figure Step plot to show contract pricing history (Source: Author, 2012)



The pricing column was duplicated, offset by one position, and then combined with the original dataframe. The percentage change column was then calculated. Zero values were replaced with nulls to avoid littering the chart with superfluous zeros:

data <- cbind(data, c(NA, head(data$PRICE, nrow(data) - 1)))

names(data)[6] <- "PREVIOUS\_PRICE"

data <- mutate(data, PERCENT\_CHANGE = (PRICE - PREVIOUS\_PRICE) / PREVIOUS\_PRICE)

data$PERCENT\_CHANGE[data$PERCENT\_CHANGE == 0] <- NA

The percentage change labels were then added to the chart by mapping the PERCENT\_CHANGE column to a text object in the plot:

qplot([code removed]) + geom\_text(aes(label = [code removed]))

The x axis scale was based on a fixed rolling window so that, as commodity managers moved from one chart to the next, they could intuitively decipher the timeline without the need to interpret the x axis each time.

### Assessment

The original intention of the project was to evaluate the effectiveness of visualization on commodity negotiations at Flextronics. Having built and refined the visualization tool, unstructured interviews with commodity managers were carried out in order to assess the potential impact.

#### Data collection and validation

Because the project was based on a dynamic business environment it was not feasible to quantify the impact of the tool in monetary terms. There were relatively few participants and it was not possible to separate the impact of the tool from external factors.

#### Feedback source

Feedback came from a variety of sources:

* Sr Global Commodity Manager, Global Procurement, Austria
* Sr Global Commodity Manager, Global Procurement, Austria
* Commodity Manager, Global Procurement, Mexico
* Sr Director, Commodity Management, Global Procurement, USA
* Sr Director, Commodity Management, Global Procurement, Mexico
* Sr Director, Commodity Management, Integrated Network Solutions, Mexico
* Vice President of Sourcing, Global Procurement, USA
* Sr Director of Sourcing, Global Procurement, USA
* Sr Purchasing Manager, Global Procurement, Israel
* Director, Information Technology, St. Petersburg, Russia

#### Bias

The positive feedback that was received was certainly subject to bias because of social pressure. The report developer sought feedback from colleagues where the relationships have both a history and a future.

#### Initial interpretation of results

From a qualitative perspective, the initial reaction was extremely positive, and one Sr. Global Commodity Manager went so far as to say, “Approaching negotiation via visualization is definitively a good idea.” The quality of the graphics and typesetting were well received. Although the positive comments were encouraging to receive, the more critical comments were easier to translate into improvements.

The dashboard itself could have been developed further. Most of the participants could have specified additional content they would have liked to see added.

#### Sensitivity of results to assumptions

The objective was to determine whether visualization was a suitable tool for supplier negotiations at Flextronics. If the implementation of the prototype was executed poorly, it seems likely that the participants might come to the conclusion that visualization added no value and was therefore a bad idea.

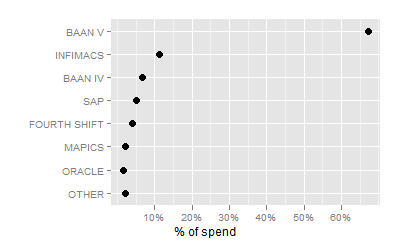
# DISCUSSION

## Data Quality

The initial prototype was built using a small sample of parts that had complete data and some interesting features. These parts were specifically selected because they were more likely to generate interest from the commodity management community. These parts were not typical of the population and, although the tool worked well using the test data, in reality there were all sorts of data problems that that resulted in confusing charts and jumbled dashboard layouts.

A major source Flextronics revenue growth has been through acquisition. Although BAAN V is Flextronics’ preferred ERP, more than 30% of spend is spread across 32 non-BAAN V sites.

Figure Multiple ERP platforms at Flextronics (Source: Author, 2012)

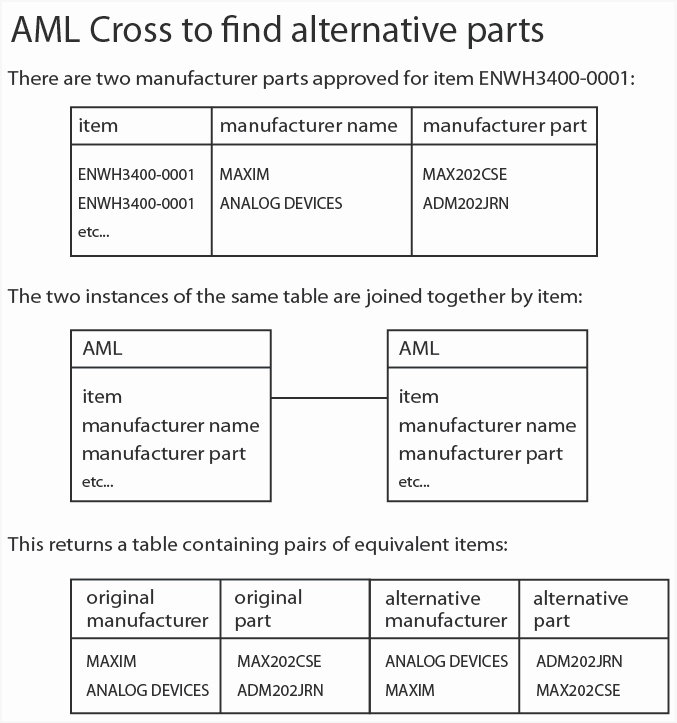


From a data governance perspective, it is difficult to control what data each site enters into its local ERP. The presence of certain columns and validation rules are inconsistent across the ERP variants. Data quality problems such as missing data and typographical errors, e.g. the year was entered as 3012 in the pricing table instead of 2012, often appear in the global data warehouse.

## Alternative Part Identification

One of the commodity managers had concerns about the method used to identify alternate parts. The Approved Manufacturer List (AML) is a table that contains the list of possible manufacturers and manufacturer parts for each item. A frequently used justification a commodity manager might make with a supplier is that a competitor sells a form, fit, and functionally equivalent item for less money. The AML table was joined to itself in order to identify parts that have been used interchangeably.

Figure AML cross using a self-join (Source: Author, 2012)



Although this works well most of the time, there are occasions where a higher specification part, e.g. a part with tighter tolerances, has been approved as a suitable replacement for a lower specification part. The AML self join associates the parts and suggests that they are interchangeable. Consequently, the AML cross alternatives sometimes have different tolerances and are therefore not always form, fit, and functionally equivalent.

For example, two capacitors from different manufacturers have an identical description: ‘CAPACITOR, TANTALUM, SOLID, POLARIZED, 10 V, 220uF, SURFACE MOUNT, 2917’. The AML cross suggested that the two parts can be used interchangeably, and revealed a price difference of about five percent. However, the parts were not truly functionally equivalent. While most of the attributes were identical, there were a few subtle differences:

Figure Differences in attributes of two items that have identical descriptions (Source: Author, 2012)

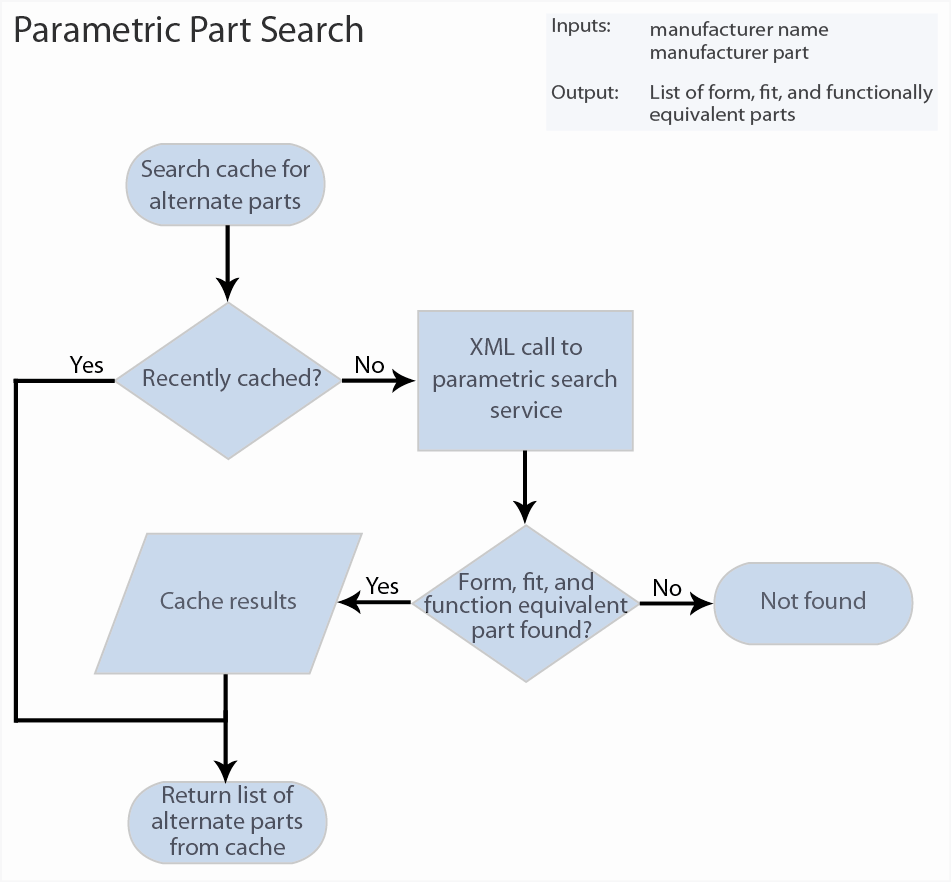
|  |  |  |
| --- | --- | --- |
|  | AVX/KYOCERA  TPSD227M010R0050 | KEMET  T520D227M010ATE040 |
| Positive Tolerance (%) | 10 | 20 |
| Operating Temperature-Max | 85 | 105 |
| Leakage Current | 22uA | 220uA |
| Ripple Current | 1.635A@100KHz | 2400mA@45cel@100KHz |
| Dissipation Factor | 0.08 | 0.1 |

The supplier of the more expensive part might argue that, although the parts are very similar, they are not identical and therefore this is not a fair comparison.

There are a number of electronics component data services that provide the capability to perform parametric searches to find alternative sources. The biggest player, IHS[[17]](#footnote-17), has an eXtensible Markup Language (XML) service that enable subscribers to enhance the functionality of their systems by making calls Application Programming Interface (API) calls to their extensive database. Flextronics would send an XML packet to the service provider that contains the username, password, manufacturer, and part number. The data service provider would respond by sending an XML packet containing the original part and a list of alternatives.

The XML service is relatively slow when compared with direct retrieval from an indexed table. The service subscriptions are also expensive: starting prices exceed $100K and increase with the number of API calls. In order to minimize the number of duplicate API calls and improve responsiveness, I propose that results are cached locally (see figure 19).

Figure Parametric search for alternative parts (Source: Author, 2012)



Manufacturing service agreements with customers often include a clause that commits to deliver ongoing cost reduction. Therefore, in order to reduce costs, a more formal and systematic process to identify alternate sources would be beneficial to the company. Most customers have quarterly business reviews (QBR) where performance issues are discussed. Because the industry is highly competitive, due to an oversupply of manufacturing capacity, customer agreements often commit to ambitious cost reduction targets. Flextronics’ supply chain managers are often left with the unenviable task of explaining why, due to economic realities, cost reduction targets have not been achieved.

There are often cases where cheaper alternative parts exists, but they have not been approved on the AML, and so Flextronics cannot take advantage of the lower pricing in order to reduce cost. In order to mitigate the claim that Flextronics failed to return the cost savings specified in the service agreement, it is helpful to have a pipeline of cost reduction proposals. In addition to providing more accurate pricing intelligence, the XML service could have a positive impact on customer relationships by systematically creating a pipeline of potential cost reduction opportunities.

## Lead Times

Although we know that lead time is an important sourcing characteristic, it is very much secondary to price because it is difficult to provide a common and consistent measure of lead time due to the patchwork of ERP systems that Flextronics uses to manage its business. Some Flextronics’ site ERP systems store lead time as an attribute of the item, while others store it at the item/supplier level. Some of the sites have internal processing times built in to the lead times while others do not. In addition, there are many parts on reduced lead time programs and these are not identified in a consistent way across all the sites. The order cycle times often seem completely unrelated to the nominal lead times. The order cycle times, i.e. the time elapsed between placing an order and receiving the parts, are frequently negative at the sites running SAP[[18]](#footnote-18).

## Transportation Cost

Because the value density of most of the materials used to build electronic products is relatively high, there has not been as much attention focused on transportation compared to many other industries. However, for certain items such as fabricated metal enclosures, the transportation cost could be a substantial cost driver and the delivery terms could make a big difference.

The cost and effort required to capture accurate lead time and transportation data are relatively high because of the piecemeal information infrastructure. One approach to manage the data issues has been to standardize the terms, where possible, across the supply base. This simplifies the sourcing decisions because the commodity managers are not forced to calculate trade-offs between price and delivery terms, for example. The benefit of this standardization is that it reduces complexity. A possible downside to the standardization of terms is that there are fewer opportunities to take advantage of the differences between the negotiating parties.

Data integrity is extremely important to Flextronics’ commodity managers and this has become a hot topic due to a series of interrelated issues. An in-house master data management (MDM) solution was replaced with the SAP MDM module in 2011. The new solution was built around some technologies that were new to Flextronics: Informatica ETL and SAP BI/BW data warehouse. The migration went badly and a decision was made to run both SAP and the legacy system in parallel until the differences could be reconciled. The negotiation datasource switched back and forth between the two different versions of the truth. At the same time, there was an ambitious push to implement numerous logic changes and add many new columns to the masterfiles. The constant concurrent changes caused a great deal of confusion and frustration and now there is consensus that a small, accurate dataset is better than a vast quantity of garbage.

The implication of the data quality issues to the visualization tool is that a few well-presented facts, primarily related to price, are more likely to result in adoption and further development of the tool.

## Non-Disclosure Agreements

One of the commodity directors noted that the prototype dashboard could not be shared with suppliers due to non-disclosure agreements. Suppliers typically mandate that Flextronics does not share their pricing with anyone who does not need their pricing in order for Flextronics to do business with them. It is acceptable to share a target price providing that the target does not contain any information that could be used to reveal another suppliers pricing. For example, if two suppliers make an equivalent part, it is acceptable to provide a target price, but unacceptable to state the purchase price of an alternative.

If the dashboard is to be used as a tool to communicate with suppliers, care must be taken not to violate any non-disclosure agreements. There is an outstanding requirement to for the commodity managers to have the ability to switch back and forth between two different versions of the dashboard: a complete version for commodity manager, and an obfuscated version for the supplier. This was not implemented because there was insufficient time.

## Open Source Software

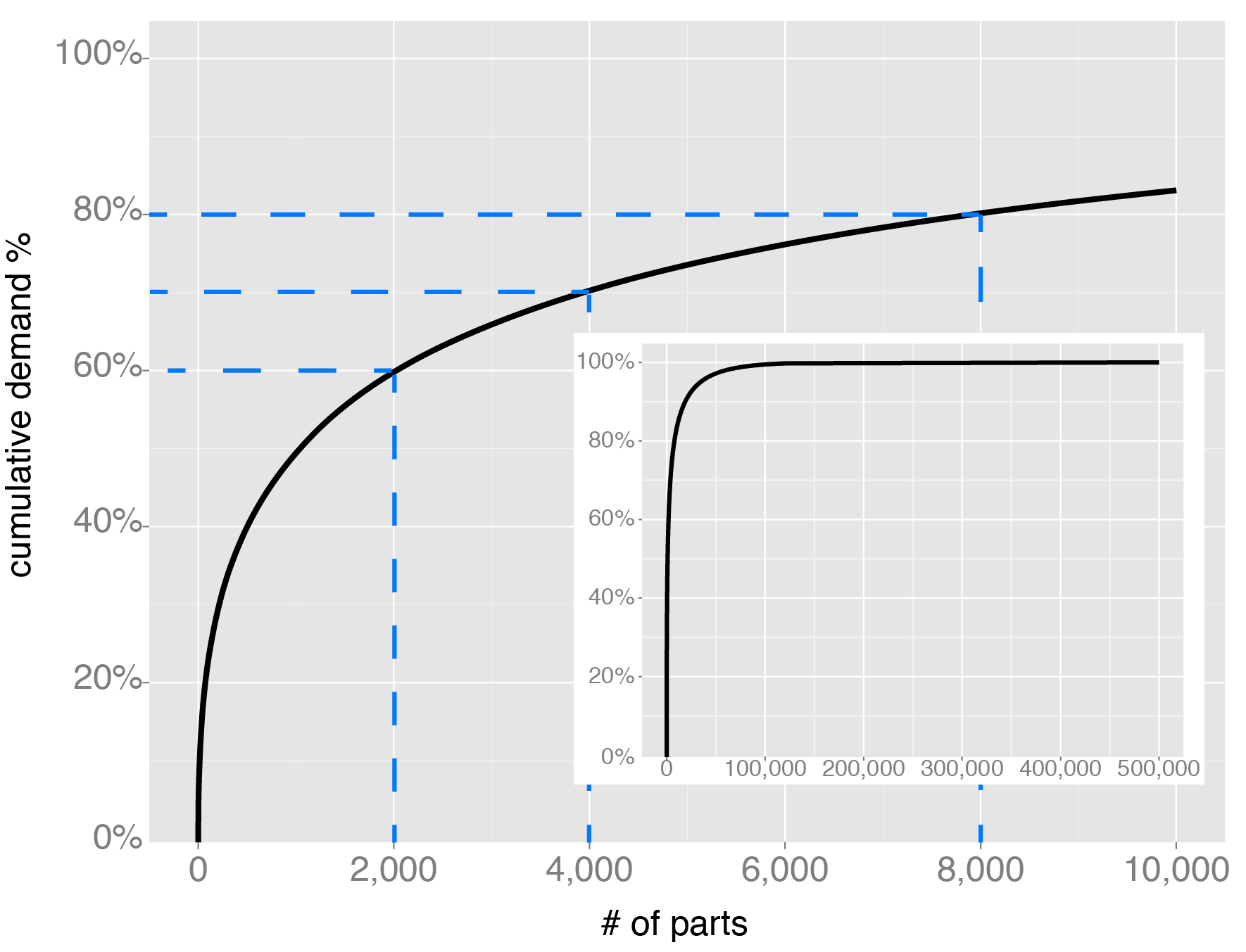
There were some reservations about the quality of open-source software and access to support in the event that something breaks. In reality, the R language has been around since 1993 and its user community is both large and active. Questions posted to the R-Help email list usually get answered very quickly, often within a few minutes of posting. Although it is possible to buy support for R from Revolution Analytics[[19]](#footnote-19), this is probably unnecessary since all the functionality used in the dashboard is based on packages that are both popular and mature.

Open source software enables the user to view and, if necessary, edit the source code. The cost savings treemap example (figure 7) was created using a new R package called ‘treemap’. Although the built-in color palettes did not have a red-white-green transition through zero, the single line that defined the color palette was obvious and trivial to edit.

## Scalability

In the current data warehouse architecture this will not scale to the entire negotiation dataset since each part takes two minutes to process, there are currently 540k parts, and the dashboards would need to be refreshed several times throughout the quarter. If the business requirement was to create dashboards for the entire negotiation dataset, then it could be done relatively cheaply and easily using a Hadoop[[20]](#footnote-20) cluster. This step was unnecessary to build the proof of concept. In reality, a commodity manager will rarely discuss more than 300 parts with a supplier in one day. The cumulative demand value curve has an extremely long tail: about 80 percent of the demand comes just from just 1.5 percent of the parts (see figure 15).

Figure Cumulative demand percent versus the number of parts (Source: Author, 2012)



## A Passive Tool

The time spent clicking around searching for data and recording the outcome should be minimized. The dashboard, in its current state, is a read-only passive tool. This means that the negotiator needs to constantly switch back and forth between the dashboard and the masterfile spreadsheet in order to record the outcomes of the discussions.

A future improvement would be for the PDF to be interactive and capture a journal of events for each part so that the negotiator can easily see the entire discussion history rather than just the most recent comment.

# CONCLUSIONS AND CRITIQUE

## Findings with particular respect to objectives

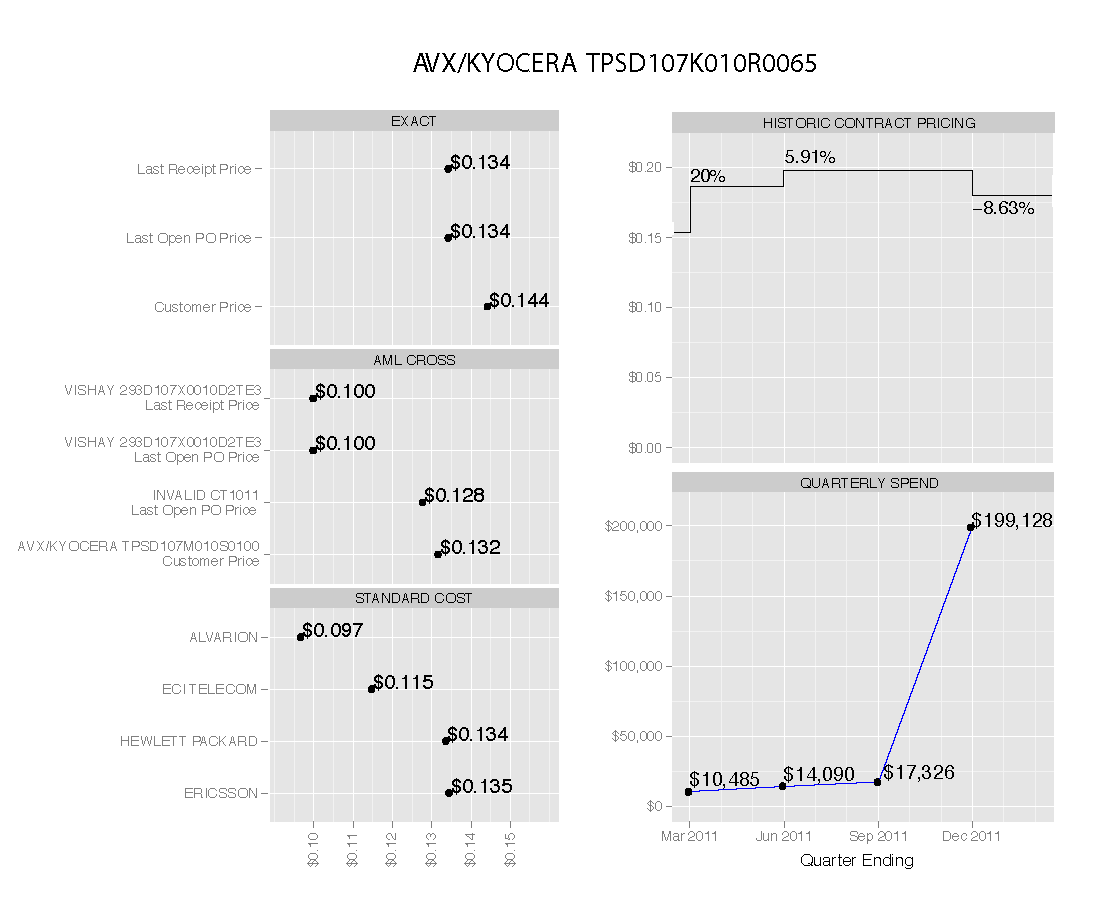
##### Objective 1: Evaluate open source technologies visualization and select an appropriate platform to make visual dashboards to support supplier negotiations at Flextronics.

An appropriate open source technology platform was identified. The pre-generated Adobe PDF dashboards worked well because they were extremely fast to navigate and were not dependent on any network connectivity. From the users perspective, no special software, expertise, or configuration was required to open and navigate the dashboards. The ggplot2 R package provided more control over the output than would have been possible using a graphical user interface.

##### Objective 2: Build a prototype negotiation dashboard.

A prototype dashboard was built using R and this was tested by running batches of a few hundred parts at a time.

Figure Prototype dashboard (Source: Author, 2012)



The prototype dashboard took approximately two minutes to process each part and this time was almost entirely spent waiting for queries to execute against the data warehouse. The time spent plotting charts and combining the graphical objects into the dashboard was negligible.

The potential utility of the dashboard improved throughout the project as input from commodity managers could be incorporated into the dashboard.

##### Objective 3: Evaluate the dashboard to see if it is beneficial to commodity negotiations at Flextronics.

The results of this evaluation were encouraging but inconclusive due to the qualitative and unstructured nature of the feedback, and the exploratory nature of the project. The dashboard, at the end of the project, was a prototype rather than a finished product. The visual negotiation arguments could certainly have been developed further given more time.

## Limitations of analysis

Because the analysis was conducted in an uncontrolled environment, it was not possible to exclude all the external factors, split the participants into two groups, and perform statistical tests determine whether visualization would be beneficial for Flextronics’ negotiation process.

The flexibility and speed of the exploratory and iterative approach to this project came at the cost of scientific rigor. It was impractical, given the time constraints, to capture reactions to dashboard components as they were developed using formal questionnaires.

## Future directions / recommendations for further work/research

The data that is currently captured from Flextronics’ negotiations is structured, e.g. price, minimum order quantity, lead time, etc..., and so this data can be neatly stored in a table. There is currently no attempt to capture and exploit unstructured data such as audio.

Many innovative products that we use today are based on a convergence of existing technologies. The popular iPod device, for example, is the combination of the MP3 audio codec and a storage device. Neither were new technologies when Apple developed the iPod. A speech-to-text service, such as Twilio[[21]](#footnote-21), could be used to transform the audio data into searchable text. There may also be an opportunity to augment the textual translation with voice stress analysis data which could, potentially, make it possible to search for contentious issues and see the dialogue history. Careful consideration needs to be given to the impact to the trust between the negotiators, and whether the benefit of the intelligence outweighs the cost to the relationship.

Interactive graphics are necessary in order to fully realize the social benefit of visualization in negotiations (Eppler and Platts 2009). Google has a number of excellent interactive graphical widgets[[22]](#footnote-22) that are accessible to R using the googleVis package. Tablet computers are rapidly gaining market share and there is certainly potential for an interactive negotiation dashboard designed to be used on an iPad or Android device. One of the negotiators compared the existing dynamic to the Hasbro board game called ‘Battleship’: both players sit on opposite sides of the table, look at their screens, one reads out a number, and the other one replies ‘hit’ or ‘miss’. A tablet is less cumbersome to turn around and the screen does not act as a physical barrier between the two parties.

It is possible to create interactive PDF’s using Latex. Although the tool is currently passive, it is certainly possible to create active content. The next generation of dashboard should incorporate the ability to capture negotiator input.

## Critique of the work (approach, analysis, etc…)

To evaluate the effectiveness of visualization on supplier negotiations, it is necessary to do the groundwork to support a definitive statement such as, “data visualization improved cost savings by x%.” Unfortunately, the quality of the prototype was subjective; although the qualitative feedback was generally positive and the design was based on documented graphical design principles, it was difficult to quantify. It was also impractical to isolate the effectiveness of the prototype from external factors such as market forces. The intention was for certain negotiators to use the dashboard for all of their parts and not just a randomly selected sample. Consequently, it was not possible to translate feedback into a quantifiable financial improvement.

# RECOMMENDATIONS

## Short term

Visualization is an appropriate tool for commodity negotiations at Flextronics. Additional graphics should be designed, coded, and added to the existing dashboard.

The IHS parametric search XML service would be beneficial for both negotiators and customer-facing account managers. The service, however, is relatively expensive and so it would make sense for the supplier facing team to share the tool and split the cost with the customer facing team. There should be a joint evaluation of the service.

## Medium term

The negotiation data is currently stored on a SQL Server that is not backed up and cannot support users querying the database directly. Flextronics has a robust SQL Server farm environment that is currently underutilized. The negotiation toolset should be moved to the new, robust environment and users should query the database directly instead of searching for the data they need in a vast collection of bloated Excel spreadsheets.

## Long term

Poor data integrity is rife at Flextronics, and this impedes management’s ability to make intelligent decisions. The IT department’s stated purpose is to provide “just enough” technology, though many internal customers would argue that the current level of service is not quite enough. Historically, Flextronics’ growth has largely been through acquisition. This, combined with “just enough” IT has resulted in a messy information infrastructure which impedes Flextronics’ ability to benefit from the economies of scale. There are currently 114 different ERP instances using 13 different platforms. If it is too complex and costly to replace the patchwork of ERP systems with one global system, then data integration needs to become a core competency and this activity needs to be resourced appropriately.

Flextronics should develop analytical capabilities internally since it is not always possible to satisfy business needs by simply purchasing top-tier Business Intelligence tools. Although the learning curve is steep compared with GUI based analytical tools, R is an extremely powerful tool and developing this skillset is worthwhile.

## Further Research

There are two major problems with analyzing the effectiveness of the dashboard for supplier negotiations. First, a visual dashboard contains many variables that may influence its effectiveness. These variables include which data is presented, dashboard layout, wording, chart types, markings on the charts such as the types of geometric objects, and the attributes of those objects (e.g. colors, fonts, symbols), etc. Second, it would be difficult to statistically isolate the influence of the dashboard from other factors such as market forces.

The approach would have been more effective if the dashboard was built on a platform suitable for rapid experimentation. For example, when an eCommerce company such as Amazon.com wants to know if layout A is better than layout B they create two versions of the page, randomly select the version when a page request is made, and track the results. This enables them to continually test hypotheses and refine their site.

There is an opportunity for a machine learning algorithm to generate hypotheses by experimenting with the dashboard variables. A ‘smart’ dashboard could learn, optimize itself, possibly reveal design principles that were not previously known, and move business intelligence graphics towards the science end of the art-to-science continuum

# REFERENCES

Barnes, E, et al. *Electronics manufacturing service industry.* Atlanta: Georgia Institute of Technology, 2000.

Bertin. *Semiology of Graphics: Diagrams, Networks, Maps.* Redlands, CA: ESRI, 2011.

Booth, Wayne C, Gregory G Colomb, and Joseph M Williams. *The Craft of Research.* 3rd. Chicago: University of Chicago Press, 2008.

Cleveland, William S. *The Elements of Graphing Data.* Summit, New Jersey: Hobart Press, 1994.

—. *Visualizing Data.* Summit, New Jersey: Hobart Press, 1994.

Deepayan, Sarkar. *Lattice Graphics.* 10 19, 2011. http://r-forge.r-project.org/projects/lattice/ (accessed 03 21, 2012).

Eppler, Martin J, and Ken W Platts. "Visual Strategizing - The Systematic Use of Visualization in the Strategic-Planning Process." *Long Range Planning* 42 (2009): 42-74.

Fisher, Roger, and William Ury. *Getting to Yes.* 2nd. London: Random House, 1999.

Gessmann, Markus, and Diego de Castillo. *googleVis: Interface between R and the Google Visualisation API.* 2011. http://cran.r-project.org/package=googleVis (accessed 2012 йил 1-1).

Google. *Google Finance.* 2011. http://www.google.com/finance (accessed 2011 йил 21-12).

Harrower, Mark, and Cynthia A Brewer. “ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps.” *The Cartographic Journal* 40, no. 1 (2003): 27-37.

In, Hoh, and Siddhartha Roy. "Issues of Visualized Conflict Resolution." *Proceedings of the Fifth International Symposium on Requirements Engineering.* IEEE. 1-2.

In, Hoh, and Siddhartha Roy. "Visualization Issues for Software Requirements Negotiation." *IEEE* (IEEE), 2001: 10-15.

Information Week. “Oracle Appliances: Ellison Touts In-Memory, Flirts With NoSQL.” 2011 йил 17-10: 13.

Jacoby, William G. *The Dot Plot: A Graphical Display for Labeled Quantitative Values.* East Lansing, MI: Michigan State University, 2006.

Kosslyn, Stephen Michael. *Graph Design for the Eye and Mind.* New York: Oxford University Press, 2006.

Larkin, Jill H, and Herbert A Simon. "Why a Diagram is (Sometimes) Worth Ten Thousand Words." *Cognitive Science* 11 (1987): 65-99.

Leach, Patrick. *Why Can't You Just Give Me The Number?* Gainesville, Florida: Probabalistic Publishing, 2006.

Lee, Kun Chang, Inwon Kang, and Jin Sung Kim. "Exploring the user interface of negotiation support systems from the user acceptance perspective." *Computers in Human Behavior* 23 (2007): 220-239.

Lewicki, Roy J, Bruce Barry, and David Saunders. *Essentials of Negotiation.* 5th. New York: McGraw Hill, 2010.

Lohse, Gerald. "The role of working memory on graphical information processing." *Behaviour and Information Technology* 16, no. 6 (1997): 197-308.

McCullough, B D, and David A Heiser. “On the accuracy of statistical procedures in Microsoft Excel 2007.” *Computational Statistics and Data Analysis* 52 (2008): 4570–4578.

Monzani, J S, S Bendahan, and Y Pigneur. "Decision and Visualization for Negotiation." *Proceedings of the 37th Hawaii International Conference on System Sciences.* Hawaii, 2004. 1-10.

Parush, Avi, Adi Hod, and Avy Shtub. "Impact of visualization type and contextual factors on performance with enterprise resource planning systems." *Computers & Industrial Engineering* 52 (2007): 133-142.

R Development Core Team. *R - Contributed Packages.* 2012. http://cran.r-project.org/ (accessed 2012 йил 16-03).

R Foundation for Statistical Computing. *R Books.* 3 19, 2012. http://www.r-project.org/doc/bib/R-publications.html (accessed 3 19, 2012).

Ripley, Brian. *RODBC: ODBC Database Access.* 2012. http://cran.r-project.org/web/packages/RODBC/index.html (accessed 2012 йил 16-03).

Sackett, Peter, and David Williams. "Data Visualization in Manufacturing Decision Making." *Journal of Advanced Manufacturing Systems* 2, no. 2 (2003): 163-185.

Sahay, B S, and Jayanthi Ranjan. "Real time business intelligence in supply chain analytics." *Information Management & Computer Security* 16, no. 1 (2008): 28-48.

Sarkar, Deepayan. *Lattice: Multivariate Data Visualization with R.* New York: Springer, 2008.

Su, Yu-Sung. “It’s easy to produce chartjunk using Microsoft Excel 2007 but hard to make good graphs.” *Computational Statistics and Data Analysis* 52 (2008): 4594–4601.

Swaab, Roderick I, Tom Postmes, Peter Neijens, Marius H Kiers, and Adrie C Dumay. "Multiparty Negotiation Support: The Role of Visualization’s Influence on the Development of Shared Mental Models." *Journal of Management Information Systems* 19, no. 1 (2002): 129-150.

The Economist. *Slicing an Apple.* 2011. http://www.economist.com/node/21525685 (accessed 2011 йил 16-12).

Trkman, Peter, Kevin McCormack, Marcos Paulo Valadares de Oliveira, and Marcelo Bronzo Ladiera. "The impact of business analytics on supply chain performance." *Decision Support Systems* 49 (4 2010): 318-327.

Tufte, Edward R. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press, 2000.

Valadares de Oliveira, Marcos Paulo, Kevin McCormack, and Peter Trkman. "Business analytics in supply chains – The contingent effect of business process maturity." *Expert Systems with Applications* 39 (2012): 5488-5498.

Vessey, Iris, and Dennis Galletta. "Cognitive Fit: An Empirical Study of Information Acquisition." *Information Systems Research* 2, no. 1 (1991): 63-84.

Weber, Michael, Gregory Kersten, and Michael Hine. "Visualization in e‐Negotiations: An Inspire ENS Graph is Worth 334 Words, on Average." *Electronic Markets* 16, no. 3 (2006): 186-200.

Wickham, Hadley A. *An implementation of the Grammar of Graphics.* 03 01, 2012. http://had.co.nz/ggplot2/ (accessed 03 21, 2012).

—. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer, 2009.

Wilkinson, Leland. *The Grammar of Graphics.* 2nd. New York: Springer-Verlag, 2005.

Williams, Steve, and Nancy Williams. "The Business Value of Business Intelligence." *Business Intelligence Journal* Fall (2003): 1-11.

Yan, Q, and S Islam. “Moving up the value chain: How to make the smiling curve smile?” Gothenburg, Sweden: University of Gotherburg, 2011.

Yau, Nathan. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics.* Kindle Edition. Wiley, 2011.

# APPENDIX A – FEEDBACK

### Commodity Manager, Global Procurement, Mexico

It is cumbersome to navigate a folder containing a PDF for each part. It would be much easier to search if all the individual dashboards were consolidated into a single PDF file that could be continually searched during negotiations.

### Sr Director, Commodity Management, Global Procurement, Mexico

There are non-disclosure agreements in place with suppliers and customers that prevent us from showing any specific pricing detail to the suppliers. If we are going to share this with suppliers we must be careful to stay within the rules. Supplier and customer specific details must be obfuscated.

### Sr Director, Commodity Management, Integrated Network Solutions, Mexico and Commodity Manager, Global Procurement, Mexico

Provide more samples for the passives, interconnect, and mechanical commodities.

### Sr Global Commodity Manager, Global Procurement, Austria

“Approaching negotiation through visualization is definitely an excellent idea. Lets have a follow-up call to bring the dashboard to the next level.”

### Sr Director, Commodity Management, Global Procurement, USA

“I like this a lot.”

The dashboard needs to be simple and provide instantaneous results. Face-to-face negotiations cannot be dependent on an internet connection because they often take place in lower cost regions where a stable, fast internet connection cannot be assumed.

### Vice President of Sourcing, Global Procurement, USA

“It looks cool.”

He suggested that personal relationships might be a more effective way to solicit feedback than social media and shared the idea with his extended team.

### Director, Information Technology, St. Petersburg, Russia

Dashboards allow us represent a lot of data and see the big picture, however, the creation of ‘infographics’ is more of an art than a science. Tools such as R can help process data, but creating something that is easy to understand and actionable is still a major challenge.

### Sr Director, Commodity Management, Global Procurement, Mexico

If we use open source software, who do we turn to if it stops working? If this software is so good, then why is it free?

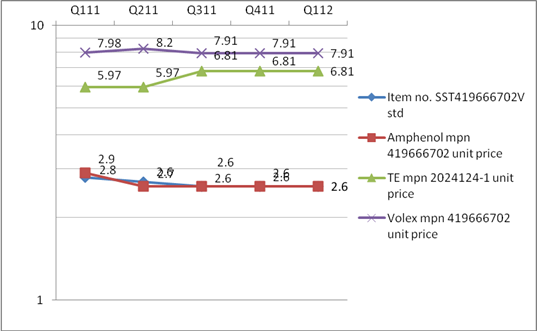
### Sr Director of Sourcing, Global Procurement, USA

Market trends should be added to the dashboard. Have a look at the data in iSuppli.

### Commodity Manager, Global Procurement, Mexico

The dashboard consolidates data that already exists in the data warehouse. It does not tell us anything that we could not already lookup elsewhere.

The dashboard should present pricing trends like this (data is not real):



The dashboard should show the price trend for each supplier over a period of time. The current dashboard does not give any indication of target costs. Raw material commodity price trends, e.g. gold and copper, should be included to provide a reference point for price changes[[23]](#footnote-23).

### Sr Purchasing Manager, Global Procurement, Israel

The actual prices in the contract history plot are not annotated on the chart. They are difficult to read from the y-axis.

The AML cross needs a health warning because:

1. AML table sometimes has missing/incomplete data.
2. Tolerances might not match: AML crosses are not necessarily form, fit and functionally equivalent. Crosses require expertise of the particular type of commodity to ensure that they are suitable equivalents.

The standard (selling) cost is only important when current price is higher than the standard cost. If there is no story to tell, do not show the graphic.

There are sometimes different standard costs in different geographical regions. The standard costs chart should include the region.

The dollar format should vary according to the number of decimal places in the price. Resistors and other extremely cheap items (typically passives) should show a price with a 5 decimal place precision, most other commodities should be displayed at 4 decimal places, and expensive ASICs may be displayed with a precision of 2 decimal places.

Although most of the negotiations are in the USD currency, some are negotiated in EUR. Where this is the case we need to show both currencies.

There is no need to look back further than 5 quarters of pricing history. All the time-series plots should span the same duration so that they can be easily compared.

The charts in the dashboard should be clearly labeled.

Future contract and bid pricing needs to be added for the original part and its parametric equivalents.

The tool should allow the commodity manager to perform calculations to immediately see what the impact of pricing changes.

The tool would be more valuable if it were active. Perhaps a commodity manager could submit pricing via a PDF form, and the prices get automatically submitted to the price book. If the tool could do something like this, then more people would want to use it.

1. 1 / 0.0928 = 10.6% [↑](#footnote-ref-1)
2. Quarter ending June 2011. [↑](#footnote-ref-2)
3. In Excel -(-1)2 = 1. [↑](#footnote-ref-3)
4. There are currently about 100 emails per day that ask and answer ‘how do I…’ type questions. [↑](#footnote-ref-4)
5. <https://stat.ethz.ch/mailman/listinfo/r-help>; <http://www.meetup.com/>; <http://www.linkedin.com/> [↑](#footnote-ref-5)
6. E.g. <http://addictedtor.free.fr/graphiques/> [↑](#footnote-ref-6)
7. R is an open source variant of a language called S, which was developed by Bell Labs. The S in Sweave is a legacy from R’s evolution from S. [↑](#footnote-ref-7)
8. There are some good examples of what is possible using the Latex markup language at: <http://www.tug.org/texshowcase/> [↑](#footnote-ref-8)
9. The author currently uses Notepad++ on Windows and TextWrangler on Mac. [↑](#footnote-ref-9)
10. <http://www.quest.com/sql-optimizer-for-oracle/> [↑](#footnote-ref-10)
11. The double assignment operator (<<-) is used to expand the scope of the assignment to include global variables. A single assignment (<-) cannot modify objects outside its current function. [↑](#footnote-ref-11)
12. In SQL, this would be described as a full outer natural join. ‘Full outer’ means include all records from both tables, regardless of whether there are any matches. ‘Natural’ means the join keys will be based fields with common names. Other types of join, e.g. left, right, and inner, are possible by changing the arguments to the join function. [↑](#footnote-ref-12)
13. These are sometimes known as heterogeneous queries. [↑](#footnote-ref-13)
14. Whisper [↑](#footnote-ref-14)
15. Flextube [↑](#footnote-ref-15)
16. This is not a new problem. Student’s t-test was first published in 1908. Its creator, William S Gosset, used the pen-name ‘Student’ because his employer, the Guinness brewing company, considered their analytical techniques to be a trade secret. [↑](#footnote-ref-16)
17. <http://www.ihs.com/industry/electronics/index.aspx> [↑](#footnote-ref-17)
18. This is probably caused by an ETL mapping error in Flextronics’ data warehouse. The order date is a mandatory field in the data warehouse receipts table. If the order date field is missing the current date is put in its place. [↑](#footnote-ref-18)
19. <http://www.revolutionanalytics.com/why-revolution-r/which-r-is-right-for-me.php> [↑](#footnote-ref-19)
20. <http://hadoop.apache.org/> [↑](#footnote-ref-20)
21. <http://www.twilio.com/docs/api/twiml/record#attributes-transcribe> [↑](#footnote-ref-21)
22. The author recently sat through a Microstrategy demonstration and was surprised to some of the Google chart widgets cited as a differentiator in their business intelligence suite. [↑](#footnote-ref-22)
23. Sr Purchasing Manager, Global Procurement, Israel thought this was a bad idea because we do not currently have reliable material content data. [↑](#footnote-ref-23)