Visualizing Business Process Based Costs

Ana Santos
ana.s.santos@tecnico.ulisboa.pt

Técnico Lisboa

September 2014

Abstract

Enterprises have a pressing need to know their costs. Having a cost model can get them so far, they need to visualize the data in a fast, accessible and efficient way. Using a research methodology we studied the efficiency of some forms of data communication in this context. Since it has already been proved that the human brain more efficiently processes visual information than text we proposed the application of a diverse range of visual representations specifically to visualize business process based cost data. The process of evaluation involved collecting feedback and testing with users as to ascertain the degree of usability and efficiency and to compare these results with those obtained using a tool familiar to them.

Keywords: Information visualization, Visualization, Knowledge, Business Process, Cost Efficient

1. Introduction

Presently, enterprises face a world that is globally connected in nearly every aspect. Information flows freely and international economic transactions have become very common. Globalisation not only means that enterprises have now a more open market for business, it also means that they have more competition. There are dangers associated with this level of openness, nowadays economic downfalls affect more enterprises than ever before. So to remain competitive and to assure their survival, enterprises have been focusing their attention increasingly more on costs[1].

Recently, Information Visualization has become crucial in the field of enterprise management[1]. Traditionally, cost management relies mostly on management experience with damaging effects that range from low-efficiency to incapability to perform real-time cost management. In addition to this, the use of non-interactive forms of visualization, e.g., spreadsheets or static graphics do little to nothing in order to give access to all information needed. Information Visualization can improve or even solve some of these problems.

2. Problem

Cost control has become, nowadays, one on the most prominent problems for every enterprise. Factors like fierce competition, globalisation and/or the difficult time that world economy lives nowadays, have contributed to this. Today, management not only needs to have access to an efficient cost model but they also have to understand the data generated from it so that they be able to comprehensively understand the company cost level and detect patterns, thus grasping an accurate decision-making direction and truly realizing an effective cost control. In summary, **enterprises generate a lot of cost information and they need a clear, fast and effective way to visualize their costs**.

Over the years, the technology and graphical performance of computers has been seeing its development grow exponentially. Despite this level of progress and the subsequently creation of refined visual representations, transmitting the information contained in these has not been quite as easy. Although some enterprises have embraced Information Visualization most of them use visualization tools that resort to reports with traditional visual representations, e.g., lines, bar or pie charts with a lot of textual elements in it and practically no interaction.

3. Motivation

This problem should and must be resolved because it has become essential for enterprises to have immediate access to information about their real costs. Even if a company adopts a cost model, it must have a way to view all the information in the fastest manner as possible so that correct and precise decisions can be made. Moreover, the correct visualization and the opportunity of interaction with the data can not only provide the opportunity to optimize resources but it can also make a company more competitive.

Additionally, the lack of new visual representations in a financial context when new ones are introduced to the world[2], more and more every day, gives us motivation to study the impact they could cause if used.

3.1. Research Question

When comparing with what is available concerning visualization techniques, one can not avoid the conclusion that these tools implement a rather small number of them and, though is considered to be effective for visualization of high dimensional data, parallel coordinates are rarely implemented[2]. Mindful of this fact, the research question used as a cornerstone for our research is: **can more sophisticated visual representations be useful and acessible for visualizing costs?**

We use the word accessible since we intend to make cost information visualization not only effective but also less demanding in terms of financial expertise.

4. Related Work

On its more general definition, Visualization can be described as *"the act or process of interpreting in visual terms or of putting into visible"*[3] or more specifically as the *"study of the transformation of data to visual representations for use in developing effective and efficient cognitive processes that make it possible to gain insight from that data"* [4].

Visual representations have been used as an effective channel of communication of ideas since the dawn of man. In fact, several researchers have declared Visualization as being vital to the resolution of intricate problems in diversified areas, from conventional science and engineering areas of expertise to key areas as financial markets, national security, and public health[7]. Visualizations take great advantage of multiple distinctive features of the human cognitive processing system. As stated by Colin Ware, *"the power of visualization comes from the fact that it is possible to have far more complex concept structure represented externally in a visual display that can be held in visual and verbal working memories"*[8].

Our innate capability of promptly detect and fathom visual patterns from visual displays makes the interpretation of data a great deal more perceptive. The cause for this is the automatic use of the human process of pattern finding by visualizations[9, 10]. The human race comes naturally wired in order to be able to make spatial inferences and decisions, the experience we gather through our lives only enhances our ability to do so. Visualizations map abstract elements and spatial relations onto elements and relations in a concrete visualization[7].

With the invention of computers, more forms of visual representations have appeared. The increasingly boundless opportunity to gain access to increasingly powerful and less expensive computers, combined with advances in computer graphics, has made it possible for a general access to systems that allow interaction and/or manipulation of visual representations in real time and exploration of the data in various forms of representation. There is, though, a consequence of its use. It can be difficult to define representations that effectively achieve their goal. Complex concepts should be communicated with clarity, precision and efficiency [6].

4.1. Information Visualization

Information Visualization is often described as being the mechanism of manipulating data, information and knowledge into visual displays taking advantage of humans' natural visual capabilities[11]. It can also be defined as "the computer assisted use of visual processing to gain understanding"[12]. The main goal is to add knowledge through the data or to gather information by analysing visually the data. In a business management context, Information Visualization assists in the process of decision-making.

Some theories claim that Information Visualization can intensify the procurement of knowledge originating at large and abstract data sets and that knowledge acquisition comes from the awareness and recall of abstract data and their mutual or reciprocal relation[13]. The objective of this process is to diminish the complexity of the analysis and understanding by making available proper techniques so that data can be displayed in a visual manner.

Ideally, through these methods, users will be able to delve into the data no matter the level of abstraction. In addition, this freedom will give them a greater sense of engagement with the data. With more interaction with the data can come a more deep understanding of it. Subsequently, this can give incentive for the users to discover more details and connections in it that otherwise would never come to light.

These methods are also important for the support of patterns recognition as they capitalise on the visual recognition skills of the users. However, human information processing capabilities, both visual and cognitive, are limited and systematically biased. Effective visualizations must take these facts into account, selecting and highlighting essential information, eliminating distracting clutter, and conveying ideas that are not inherently visual, such as transformations or causality, through visual channels.

Theories of computational efficiency defend that there are representational codes that help to infer more information than others. This means that the human being has limited and systematically biased information processing capabilities. In order to be effective, visualizations have to select and highlight crucial information, eliminating everything considered clutter, and visually transmit ideas that are not intrinsically visual, e.g., transformations or causality.

This was noticeable on Larkin and Simon work. They confirmed that diagrammatic representations aided more search process in physics than textual representations. In conclusion, different representations of the same data lead to different processing opportunities, which means different efficiency levels[14]. This is consequence of the fact that visual attributes like color, size, among others, are processed even before complex cognitive processes are triggered[15].

4.1.1 Taxonomy

There have been several attempts at constituting taxonomies to differentiate kinds of Information Visualization, one considered pioneer by analytically delineating visualization techniques was the proposal constructed by Shneiderman[11] where two dimensions are identified: data type and task[16]. Another well mentioned taxonomy for Visualization was brought by Keim[17]. It considers that a three step process in order to explore visual data: Overview first, zoom and filter, and the details-on-demand. This process is somewhat known as the *Information Seeking Mantra*[17]. In this work, the data to be visualized is cataloged in the following six categories:

One-dimensional It is data that customarily has one dense dimension.

Two-dimensional This kind of data has two distinct dimensions, e.g, geographical data has longitude and latitude, and normally is associated with X-Y plots.

Multi-dimensional Any data that subsists in more than attributes and that, in consequence, cannot be associated with simple visualizations as two-dimensional or three-dimensional plots.

Text and Hypertext Both data types cannot be easily depicted by numbers and so accordingly most visualization techniques that are considered standard cannot be enforced.

Hierarchies and Graphs This type of data reflects its relationship with others pieces of information, usually graphs are extensively used in order to represent such connections.

Algorithms and Software Data that reflects the flow of information in a program. This data can lead, for example, to the creation of graphs that represent the structure of the code lines, useful for debugging.

While visualization techniques may be separated into the following groups:

Standard 2D / 3D These displays include bar charts, x-y and x-y-z plots.

Geometrically Transformed This category includes techniques found in exploratory statistics, e.g. scatterplot matrices, and other techniques associated with the term "projection pursuit".

Icon-based The attributes values of multi-dimentional data are mapped into the features of an icon. The kind of icon used can vary from little faces to stars or even stick figure icons. The correlation between values and the features of the icon are what finally generates the visualization.

Dense pixel In dense pixel techniques each dimension value is corresponded to a colored pixel. Afterwards all pixels associated with a dimension are clustered together into adjacent areas. This process allows visualizations to hold incredible amount of data and several ways to display it. Different purposes lead to different arrangements.

Stacked Presents data hierarchically, includes techniques like Worlds-within-worlds and Treemap.

4.1.2 Color

One cannot talk about Information Visualization without mentioning the role that something as simple as color can take in it. Color helps us distinguish elements more easily, sometimes it is the only factor that makes elements different from one another[18]. The ability of human eye to detect colors is resumed basically to the possession of three color receptors, called cones. Cones can divided into three subclasses considering the wavelength: long (L), medium (M) and short (S)[19].

So different wavelengths mean different receptors, but not all are absorbed with the same degree of efficiency. For example, lower wavelengths, that correspond to the blue part of the color spectrum, have a low rate of absorbance. This explains why information encoded in true blue can be less successfully analysed. This is important because one of the many applications that color has in information visualization is associated with labelling, more technically called nominal information coding. It is generally considered the most effective way to classify visual objects [18]. But there are some guidelines to be considered:

- Colors should not elevate the complexity of the visualization. Fewer colors are preferable to countless;
- The objective is to use color as a tool to guide the eye to relevant information areas;
- The data should be represented by the color map;
- Visual representations should look appealing and relaxing, an user can easily lose interest if it does not grab his attention. On the other hand, in order to avoid stressing the eye, which can occur when analysing them for an extended period of time, garish colors and wide vibrant areas must be avoided at all cost.

4.1.3 Equivalency

Representations can be equivalent on two levels, informationally and computationally. A representation is said to be informationally equivalent to another when all its information is likewise verifiable from the other representation and vice-versa. In turn, the information in one representation can give origin to the other representation. Computational equivalency is verifiable when both representations are informationally equivalent and, in addition to this fact, any conclusion drawn effortlessly and expeditiously from one representation can be stressed from the information on the other representation, and vice-versa.

4.1.4 Interaction

So that the data analytical process can proceed with ease it is imperative that the user is able to interact with the visual display. This entails the use of focus, context and distortion techniques, e.g., operations like drilling down or up along the data to modify the focus of the view without losing the whole picture, and to connect and make available different views for the data appear from several perspectives[2]. The interaction techniques infer how achievable any level of synergy is between the user and the visual representations. These techniques are what makes possible for the user to explore the data more freely.

4.1.5 2D vs 3D

In three-dimensional representation (3D), Information Visualization is confronted with two problems: occlusion and spatial. Visual representations that use three dimensions can present difficulties caused by occlusion induced by elements that appear to be "concealed" by other elements in front on them. The difficulty to position the elements with respect to all axes involved in the representation[15] can add to difficulties already existent in the representation. Though some solutions have been presented, the occlusion problem persists and this kind of representation is only used in specific cases, e.g., when it is needed to produce an object in movement. In conclusion, between these two types of data representation, the default to be used should be 2D.

5. Proposal

The main objective of this proposal is to make available tested visual representations that have been proved to be useful and accessible in business process based costs data visualization. We want to not only to give the best visualization means of representation for this strain of data but also to evaluate the use of more sophisticated kinds of visual representations found in Information Visualization, specifically in the area of visualization of cost data.

5.1. Simple

The first category, Simple, includes the most common visual representations found on existent tools, e.g., bar-line-pie charts, scatterplots and heatmaps. For the purposes of this study, we analyse the Bar Charts and Pie Charts.

We maintain these representations on account of not only, as previously mentioned, them being standard presence on visualization tools but, furthermore, for their simplicity. This makes them ideal for simple visualizations, for example, visual representations of linear data.

5.1.1 Bars Charts

This kind of visual representation is one of the most familiar to the general public.With Bars Charts, questions involving the determination of the highest and lowest value within a category can be answered, e.g, "What is the resource with the highest cost?" or "What resource can be considered average?".

But there is a problem with Bar Charts, they are limited when hierarchy is called to question. For questions like "What are the elements that compose process B?" to be answered we would need a more complex Bar Chart with a lot more information (textual and visual), which can be confusing for the user.

5.1.2 Pie Charts

Pie Charts, like Bars Charts, are familiar to general public and have been around for quite some time in tools similar to Microsoft Office Excel for Mac 2011

(Excel). This type of charts can be used, very much alike Bars Charts, to answer questions about high and lows. Nevertheless using multiples instances of Pie Charts interconnected we intend to extend the domain of information a pie chart can present.

5.2. Sophisticated

The Sophisticated category includes, as the name suggests, more complex visual forms of representation. By using the term sophisticated we are addressing more rarely seen and more recently created, by comparison, forms of visual representations. Given the nature of cost data, there are some important aspects, for example, data flow and hierarchy of the entities involved, must be made available to the user. For a better understanding, even the composition of the cost must be available to the user. Cost can be a single value or result of an operation made over multiples values. To overpass that deficiency, we propose the addition of newer visual representations.

5.3. Sankey Diagram

Our use of Sankey Diagram is justified by its properties. This kind of visual representation is mostly used to envision how energy or materials progress through assorted networks and processes, indicating quantitatively information about how its contents relate with each other. In this proposal we intend to take full advantage of this property but, this time, by applying it to cost data.

The different entities involved in this type of data, i.e., processes, activities, resource pools and resources will be represented by visual elements, distinguished by different colors. These elements are connected through graphs that not only will reflect the flow direction but in addition will appease its conservation. This means that the sum of incoming values for each visual node is exactly the same value as the sum of its outgoing values[20].

This is ideal for any scenario that includes following a flow of values from start to finish, which includes cost data. By referring to hierarchy we are invoking all scenarios where ranking is important. For example, the total cost of an enterprise is composed by the costs of each company branch and so on.

5.4. Sunburst

This type of diagram in also used in order to represent these hierarchy, giving it special relevance. Radial visualizations have proven to be advantageous when illustrating information hierarchies[21].

Additionally, this kind of visualization also plays with proportion, a very important attribute that, like the thickness of any line in a Sankey diagram, provides us with a way to show the cost value and, at the same time, provide the user with an immediate comparison with any value in the same visualization.

5.5. Treemap

By mentioning Treemaps we refer to those with and without a zoom feature. Since Treemaps are also associated with hierarchies, we included this kind of visualization in our proposal. Due to the lack of its presence in tools, when considering financial data, we included it in the Sophisticated category.

5.6. Bubbles Charts

Similar to Sunburst diagrams, Bubble Charts can mainly make use of three aspects: proportion, color and hierarchy. The manipulation of the proportion aspect enable us to reflect the value of the entity, big values will infer big bubbles. The color, as always, will facilitate the task of entity identification by the user with different colors being used in order to distinguish different entities. The fact that bubbles can contain smaller one installs a sense of composition and hierarchy in this visual representation.

6. Demonstration

6.1. Sunburst

We start by establishing the use of Sunburst as a mean of visualizing data concerning resource pools where the visual representation acquires the depth of two layers. The inside layer corresponds to all existent resource pools, the area of each one is proportional to its cost, capacity or Capacity Cost Rate (CCR), according to the option selected. Subsequently the outer layer corresponds to the resources, these appear correlated with the resource pool with which they are associated. On **Figure 1** we demonstrate how resource information appears in the same visual representation.

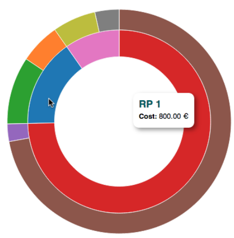

Figure 1: RP - Sunburst showing resource pool information

In addition, we also applied this visual representation to business processes and resources information. Concerning business processes, the center node represents the total average cost of business processes, the inner layer represents the business processes and the outer layer corresponds to the resource pools associated with the business processes.

Sunburst can be found in its more simple form when its representing resources data. Resources are a basic form of data, this means they cannot be divided into more basic entities. So for this fact, the visual representation presents only one layer of data, the resource layer. When representing business processes or resource pools, the Sunburst allows through the selection of any element represented an automatical drill-down, e.g, by selecting a resource pool we obtain a sunburst with only the resources associated to it.

6.2. Pie Charts

In similarity to Sunburst we adapted Pie Charts to the context of costs and hierarchy. In order to represent business processes and the resource pools associated to them we make use of two-steps progression pie charts. Considering a business process scenario, this means that in a first step once the user selects the visual representation he will be bestowed upon with a pie chart with all business processes available and their average cost. Subsequently, by selecting one business process, a second pie chart emerges with all resource pools associated to it and their average costs.

With a resource pool scenario an identical process transpires. The first pie chart will retain all resource pools available and their costs and the selection of a specific one will trigger the generation of the second pie chart with all resources associated with it and their costs. In addition to cost information, resource pools and resources can transpire capacity and CCR values through the selection of the corresponding option by the user.

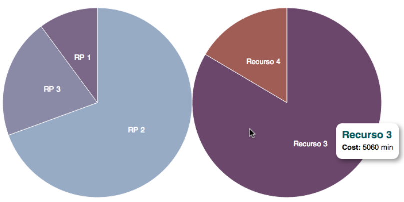

Figure 2: Pie chart displaying RP, RP3, information and its resources after its selection by the user

6.3. Bubbles Chart

Very much alike Sunburst, this kind of visual representation combined with a zoom feature allows to represent data with depth, e.g, with Business Processes we want to not only display all available but additionally the resource pools with which they are associated.

In **Figure 3** is represented an example of an implementation of Bubbles Chart with Business Processes average cost data. As can be seen, all business processes available are represented as bubbles with an area dependable on its average costs. This fact means that the bigger the cost value the more considerable the size of the bubble will be. As in Sunburst, the selection of a business process triggers a drill-down process and the visual representation zooms into the bubble and allows the visualization of the resource pools associated with it and their average cost values.

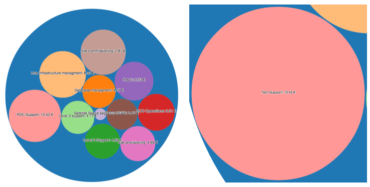

Figure 3: Bubbles displaying BP (left) and BP POC Support resource pools (right) after its selection by the user

Since Resources, as previously mentioned, can not be divided into more basic components, the zoom feature disappears and a more simple version of this visual representation is displayed, as can be seen in **Figure 4**.

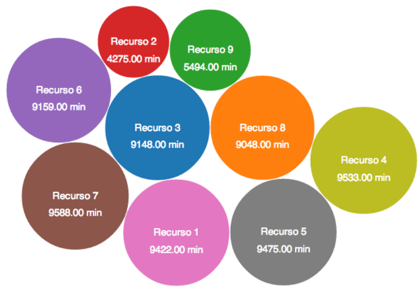

Figure 4: Bubbles displaying R capacity information

The proportion of the bubble of each resource is directly linked with the value of the attribute selected (cost, capacity or CCR).

6.4. Treemap

With Treemaps the element of proportion is also visible, either with with data from Business Processes, Resource Pools or Resources.

This visual representation allows another approach on how costs can be depicted. For example **Figure 5** illustrates an example of usage of Treemaps in a Resources context. All resources are represented as individual elements, rectangles, with different colors and their dimension is connected, as previously mentioned, to their value.

Futhermore these individual components compose a general element that symbolises the grand total of the values represented. This allows the observation of impact of an individual element relatively to the whole.

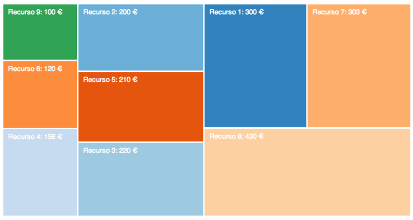

Figure 5: R - Treemap example

Additionally, Treemaps were adapted to contain the structure and values of Business Processes and Resource Pools. Since both retain more data than the immediate values associated to them, for example, Business Processes also have data concerning the Resource Pools associated to them, this visual representation carries, in consequence, more elements and is more complex. So to atone this there is a zoom feature in these scenarios.

6.5. Bars Charts

Bar Charts are one of the most recognisable forms of visual representations by the general public. In this thesis they were applied to data concerning Resources.

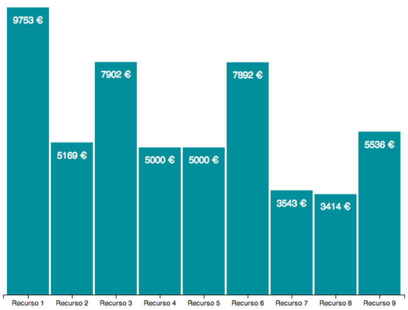

Figure 6: R - Bars Chart example

This kind of visual representation allows a slightly different kind of analysis and perhaps a more expeditious one. The eye no longer focuses its attention on the area of element buts its height.

6.6. Sankey Diagram

This visual representation is composed by three elements: business processes, cost flow connections and resource pools. Contrary to with previous visualizations, by using Sankey Diagrams to envision Business Processes we are able to not only see how much an resource pool costs in average in a business process but additionally to scrutinise the average grand total of each resource pool. An example can be seen in **Figure 7**.

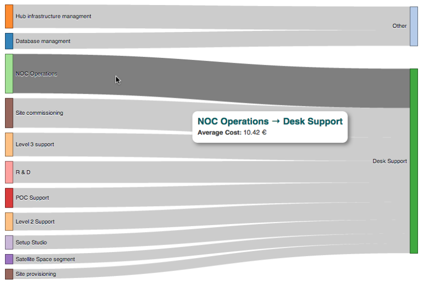

Figure 7: BP - Sankey Diagram example

Although the business process elements are arranged automatically this visual representation possess an interactive feature that allows a user to arrange them so that he can make a more specific analysis of his chosing.

7. Evaluation

The goal of this evaluation was to ascertain to which degree the solution proposed in Proposal (section 5) solved the problem stated in Problem (section 2). The evaluation method was divided into the following areas:

1. **Usability Tests and Survey -** Comparative analysis between time records of how long a user takes in order to answer a set of questions through the use of the visual representations and the time he or she takes through a tool already known to them, Excel. These tests were followed by a survey as to ascertain the usability of each visual representation.

2. **Feedback -** Feedback given by potential final users after analysis of the visual representations.

Our testing group consisted of 10 persons(6 female and 4 male) with ages between 18 and 30 years old. Only two users declared being vaguely daltonic.

7.1. Usability Tests

We present an analysis of the time records obtained through testing with users. We tested three categories of cost data: Business Processes, Resource Pools and Resources.

Business Processes (BP)

In order to test with data from BP we selected the following visual representations: Pie Charts, Bubble Charts, Sankey Diagram, Sunburst Diagram and Treemap. So to examine in contrast with tools that users already knew, we further tested the same data with Excel, a tool known by all of them. The following questions were asked so to register the time users took to give an answer using each of the visual representations provided and Excel:

1. Which business process has the highest cost value ?

2. Which resource pools belong to business process X ?

3. What is the average cost value of business process X ?

4. What is value of the resource pool with the hightest average cost value of business process X ?

Speaking in a general manner Excel had the best average time response of all considered, with 3,716 seconds. The highest average, and for the purposes of this thesis the worst result, was achieved by users using Treemap, 8,7 seconds. Nevertheless this changes slightly when going unto a question by question analysis. In Question 1, when required to identificate the business process with the highest average cost value, users were the most efficient in facilitating their answers by using Bubble Charts and the least with Treemap.

For the remainder of the questions, the best time records were obtained when users used Excel and the least with Treemap. Nevertheless it is worth mention that both the Sankey Diagram and the Bubbles Chart obtained good results, e.g, in Question 2 users using a Sankey Diagram took in average 4,273 seconds to answer, only 0,459 seconds more than with Excel, and in Question 3 the difference between the average time response using Bubbles Chart and Excel was even lower, 0,368 seconds.

Resource Pools (RP)

The visual representations tested with resource pools data were: Pie Charts, Sunburst, Bubbles Charts and Treemap, these two last ones with a zoom feature. Futhermore as a mean of comparison we also tested the data using Excel, since most users are familiar with it. As to test the data we requested answers to the following questions:

5. Which resource pool has the highest cost value ?

6. What resources belong to resource pool X ?

7. What is the value of the resource with the hightest cost value of resource pool X ?

As can be observed in **Table 1**, although Bubbles Chart comes as the visual representation with lower average response time there is little different between its value and the value of the second and third lowest averages, Excel and Pie Charts respectively.

Table 1: RP - Average Response Time

Average Response Time

Pie	3,707
Sunburst	5,156
Treemap	5,404
Bubbles	3,462
Excel	3,666

On a question by question analysis, Excel is the one that has the worst performance on the identification of the resource pool with the highest cost value, with users taking in average 5,087 seconds to answer. This duration is equal to twice as long as the best average obtained, Bubbles Charts.

This changes in Question 6. When confronted with a question concerning the composition of the Resource Pools, users had some difficulty answering through a Sunburst visual representation and took even longer to answer, in average 7,458 seconds. In this question Excel had the best time, 2,487 seconds.

In the last question, Question 7, the highest average response time was obtained by using Treemap. In contrast, Excel provided the means for the lowest.

Resources (R)

In Resources, users were confronted with data with level of depth of one, this meant that the entity, "resource", was not the result of the composition of other entities. The visual representations selected for testing were: Bars charts, Sunburst, Treemap and Bubbles Charts. In addition to these representations we also tested, as mentioned in previous subsections, with Excel. Within this theme we made the following questions:

8. Which resource has the lowest value ?

9. What are the resources that you consider to have average cost ?

10. What is the value of resource X ?

The best results overall are with Bars Charts with users taking, in average, 3,484 seconds to provide an answer. In the opposite side of the spectrum is Excel with an average response time of 7,107 seconds, the worst result of all options analysed.

It came to our attention that although Bubbles Charts does not have the best average response time achieved it came quite close to it. When comparing the results achieved with Bars Charts and Bubbles Charts, it is noticeable that users take longer to suggest which resources could be considered to have average cost value (Question 9), almost 2,5 seconds more in average. This fact suggests that Bubbles

Charts did slightly more inefficient when faced with questions of the same nature of Question 9. But the worst results, generally speaking, in this theme was Excel, as previously mentioned. With Excel an user takes slight more than three times more in Question 8 and twice as long to than it does using the best found alternative, Bar Charts.

While in Question 9 we registered additionally the answers given, an user could indicate an unlimited number of resources as being the ones considered with average cost. We detected that users had a tendency to facilitate more resources when using Treemap than with any other visualization. The best percentage recorded in this case was of 50%, which means only half of the users indicated that resource as being average. This improved slightly with Excel. It had fewer answers although its percentages divided themselfs between acceptable, 80% and not acceptable, 10%. The most consistent answers were given through Sunburst, not only the options given were fewer, 4 alternatives, but moreover their percentages ranged between 70% and 90%.

8. Conclusions

Having proposed, coordinated and conducted a set of tests to different kinds of visual representation when applied to business process based cost, we learned some interesting facts.

Treemap performed, in general, badly. It was the visual representation that lead users to take longer in order to answer questions about the data, it gathered somewhat mixed feeling from testers. Half of the testers remarked that they would probably not use it frequently when associated with a zoom feature. Some considered it unnecessarily complex. More than half of the users had great doubts regarding the difficulty of its usage and some admitted needing help in order to analyse it. This improved slightly with Pie Charts.

The time results obtained through Pie Charts were not distinguishable. Some users considered it a bit unnecessarily complex but otherwise it was remarked they would probably use it to some extend frequently, easy to use and with a quick learning process. Additionally, users remarked that they felt very assured in using it.

Sunburst was initially the visual representation that most piqued the interest of potential final users, it conveyed to be a good approach to data representation. This hypothesis was refuted through testing. Though its time records were not the most unfavorable, this visual representation was the one that received the most concerns from the users. When inquired about what resources they considered to have average cost value, users offered the less possibilities and most identified the same resources. Sunburst was considered by users as being marginally complex, of difficult usage, it required previous learning, assistance and did not garner much confidence from them.

Another visual representation that had picked interest in the initial phase of this thesis was Sankey Diagram. This diagram was exclusively applied to Business Processes data. It performed very well, however users demonstrated some doubt about the frequency of its use. They made note of a slight possibility of needing assistance in order to use it and some admitted being a bit difficult to use. In contrast, Bubbles Chart performed nicely while testing and users responded very well to it. It was remarked that the visual representation had a very quick learning process and that it was easy to use.

When dealing with Resources' data, Bars Chart the best time records by comparison. This visual representation was the most familiar to users and the one that gathered the most positive feedback. Users remarked it was easy to use, has a quick learning process and that they would use it frequently. However when doing a cross-examination between all visual representations and Excel, one cannot ignore that to some extend users can offer answers in an efficient manner through Excel but it can very quickly turn inefficient as the amount of data to present increases. This is proven by the fact that only when dealing in one of the three data categories, BP, Excel was considered the best.

In the beginning we had the objective to answer the following question, *"Can more sophisticated visual representations be useful and accessible for visualizing business process based costs?"*. The study conducted gaves us a glimpse that it can but not without some constraints. For example, as previously mentioned Sankey Diagram performed quite well but users did not felt quite as confident in using it as they felt when using a bars chart, partially because it was a more foreign visual concept to them. The tests conducted enabled us to draw some conclusions, however in retrospect, we learned that it would have been more preferable to have had a more extensive set of questions in order to conduct a more precise study though it would, most certainly, be a more dragging process for testers and results could be affected by this.

8.1. Limitations

As it normally occurs in a research study, there were found some limitations. The vast diversity of visual representations available infers indirectly a limitation in this study. Only a specific set of visual representations were tested with cost data related to business processes. Nevertheless we gathered enough data and feedback from users and potential final users that allowed us to draw conclusions regarding each visual representation. The unlimited length that the name that some business process can take proved to be a limitation to identification label that our visual representations have. It presents difficulties when presented with a very long string.

References

[1] Lizhi Kong. Cost prediction and visualization control system for enterprises. In *2010 International Conference on Management and Service Science (MASS)*, pages 1–4. IEEE, 2010.

[2] Leishi Zhang, Andreas Stoffel, Michael Behrisch, Sebastian Mittelstädt, Tobias Schreck, René Pompl, Stefan Weber, Holger Last, and Daniel A. Keim. Visual analytics for the big data era - A comparative review of state-of-the-art commercial systems. In *IEEE VAST*, pages 173–182. IEEE Computer Society, 2012.

[3] "Visualization". Merriam-Webster Online Dictionary. http://www.merriam-webster.com/dictionary/visualization, November 2013. Acessed on 29/11/2013.

[4] Theresa-Marie Rhyne and Min Chen. Cutting-edge research in Visualization. *Computer*, 46(5):22–24, 2013.

[5] Jarke J. van Wijk. The Value of Visualization. In *IEEE Visualization*, page 11. IEEE Computer Society, 2005.

[6] E.R. Tufte. *The visual display of quantitative information.* Graphics Press, 1983.

[7] Chris Johnson, Robert Moorhead, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry S Yoo. Nih/nsf visualization research challenges report. In *Los Alamitos, Ca: IEEE Computing Society.* Citeseer, 2006.

[8] Colin Ware. Visual queries: The foundation of visual thinking. In Sigmar-Olaf Tergan and Tanja Keller, editors, *Knowledge and Information Visualization*, volume 3426 of *Lecture Notes in Computer Science*, pages 27–35. Springer Berlin Heidelberg, 2005.

[9] Tanja Keller and Sigmar-Olaf Tergan. Knowledge and information visualization. chapter Visualizing Knowledge and Information: An Introduction, pages 1–23. Springer-Verlag, Berlin, Heidelberg, 2005.

[10] Tong Li, Shan Feng, and Ling Xia Li. Information visualization for intelligent decision support systems. *Knowledge-Based Systems*, 14(5):259–262, 2001.

[11] Luca Chittaro. Information visualization and its application to medicine. *Artificial Intelligence in Medicine*, 22(2):81–88, 2001.

[12] Stuart K Card and Jock Mackinlay. The structure of the information visualization design space. In *Symposium on Information Visualization (InfoVis '97)*, pages 92–99. IEEE, 1997.

[13] Tanja Keller, Peter Gerjets, Katharina Scheiter, and Bärbel Garsoffky. Information visualizations for knowledge acquisition: The impact of dimensionality and color coding. *Computers in Human Behavior*, 22(1):43–65, 2006.

[14] Jill H Larkin and Herbert A Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11(1):65–100, 1987.

[15] Riccardo Mazza. *Introduction to information visualization.* Springer, London, 2009.

[16] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, pages 336–343. IEEE, 1996.

[17] Daniel A. Keim. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[18] Colin Ware. *Information Visualization: Perception for Design.* Morgan Kaufmann Publishers, San Francisco, CA, USA, 2012.

[19] Reto Stauffer, Georg J Mayr, Markus Dabernig, and Achim Zeileis. Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations. Technical report, Faculty of Economics and Statistics, University of Innsbruck, 2013.

[20] Patrick Riehmann, Manfred Hanfler, and Bernd Froehlich. Interactive sankey diagrams. In *Symposium on Information Visualization (InfoVis 2005)*, pages 233–240. IEEE, 2005.

[21] John Stasko and Eugene Zhang. Focus+ context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *Symposium on Information Visualization (InfoVis 2000)*, pages 57–65. IEEE, 2000.