



METROBI – Estudo de caso do uso da metodologia CRISP-DM na inteligência de negócios de prestador de serviços metrológicos.

Amorim, C.B¹, Margotti, E¹, Oliveira, A. L.¹, M. de, Demay M. B.¹.

¹ Fundação CERTI, Florianópolis, 88040-970, Brasil

bca@certi.org.br, elm@certi.org.br; aeo@certi.org.br; mbd@certi.org.br

Resumo:

É crescente a necessidade de adaptação das empresas diante do aumento de competitividade. O uso de Inteligência de Negócios fornece uma alternativa para a demanda crescente de adaptabilidade das empresas (Watson e Wixom, 2007). O presente artigo tem como objetivo, definir as informações que sirvam de apoio estratégico tanto para o laboratório, quanto para seus clientes. Como metodologia, trata-se estudo de caso aplicado em um laboratório de metrologia, com a abordagem quantitativa, com a construção de hipóteses para a análise dos dados. Foi possível avaliar que a mineração e análise de dados de um prestador de serviços tecnológicos podem ser usados para aumentar não somente a competitividade dos laboratórios, mas também das cadeias econômicas por ele atendidas.

Abstract:

There is a growing need for companies to adapt in the face of increased competitiveness. The use of Business Intelligence provides an alternative to the growing demand for adaptability from companies (Watson and Wixom, 2007). This article aims to define the information that serves as strategic support for both the laboratory and its customers. As a methodology, it is a case study applied in a metrology laboratory, with a quantitative approach, with the construction of hypotheses for data analysis. It was possible to assess that data mining and analysis of a technological service provider can be used to increase not only the competitiveness of laboratories, but also the economic chains served by it.

1. Introdução

Com a evolução das sociedades e a crescente competitividade que permeia o cenário empresarial global, as organizações têm buscado formas de se adaptarem. Neste contexto ocorre a busca por ferramentas tecnológicas que possam ajudá-las com a tomada de decisões assertivas nos mais diferentes níveis organizacionais. Assim, a Inteligência de Negócios (Business Intelligence – BI) surge como uma alternativa para a demanda crescente de adaptabilidade das empresas e respostas rápidas ao mercado (Watson e Wixom, 2007).

Julga-se que diante de um cenário de maior competição, as organizações necessitam repensar quatro pontos fundamentais: i) a imprevisibilidade de determinados eventos, bem como o impacto que eles produzem em seus negócios; ii) a ausência de dados para prever determinados eventos; iii) excesso de dados, o que dificulta a análise de cenários; e iv) a imprevisibilidade de determinados eventos (McGee, 2004). Nesse cenário a confiabilidade dos dados se torna uma propriedade indispensável. Desta forma

a grande quantidade de dados coletados e suas análises tendem a ser inúteis se forem armazenadas informações não confiáveis, que não ajudem a entender uma realidade complexa.

Inseridos no ambiente de imprevisibilidade e constante mudança estão os laboratórios metrológicos, empreendimentos prestadores de serviços de caráter técnicos especializados (Toyama, 2017). Como características, possuem: i) incertezas de execução; ii) padronização na execução; e iii) base em normas técnicas ou procedimentos validados. Destaca-se, então, que o método utilizado para a execução do serviço define como o produto do laboratório é produzido, assim como a discussão da questão mercadológica. Desta forma, compreender os aspectos do comportamento do consumidor de serviços, especificamente metrológicos, torna-se tarefa importante para o bom andamento dos prestadores de serviços da área (Toyama, 2017).

Neste contexto criou-se o “MetroBP”, framework apoiado na metodologia CRISP-DM, para concepção de um Business Intelligence para um prestador de serviço metrológico. Tal iniciativa visou realizar análise de dados de serviços metrológicos com o objetivo de definir as informações que sirvam de apoio estratégico tanto para o laboratório, quanto para seus clientes. Com estes dados cria-se a possibilidade de um impacto sistêmico na competitividade da cadeia, com informações valiosas a todos os envolvidos.

2. Business Intelligence

O termo Business Intelligence (BI) foi cunhado por Howard Dresner em 1989 e se refere a sistemas e métodos que dão suporte ao sistema de decisões a partir de fatos fornecidos (Watson e Wixom, 2007). Pode-se conceituar o BI como um conjunto de processos, tecnologias e ferramentas que consigam transformar dados em informações, estas em conhecimento, e este em planos que resultem em benefícios para o próprio negócio e para o ecossistema (Loshin, 2003).

Destacam-se dois pontos cruciais no desenvolvimento do BI. O primeiro se refere à adoção de Data Warehouse, possibilitando visões mais abrangentes e integrativas dos dados, assim a consulta aos dados de forma facilitada. O segundo ponto trata do desenvolvimento computacional e a disseminação dos dispositivos individuais, quando surgem as primeiras ferramentas de análises de dados e geração de relatórios, denominadas On-line Analytical Processing (OLAP).

3. Metodologia

O presente trabalho é um estudo de caso aplicado em um laboratório de metrologia, com a abordagem quantitativa para a análise dos dados. O processo de análise foi dividido em 5 etapas resumidas no quadro 1.

Quadro 1. Etapas metodológicas

Etapa 1	Análise das bases de dados disponíveis e identificação das variáveis em foco.
Etapa 2	Grupo de Foco com os coordenadores para avaliação dos achados e indicação de direcionamentos possíveis.
Etapa 3	Desenvolvimento de hipóteses de trabalho.
Etapa 4	Criação de matriz de decisão para a escolha das hipóteses (maior impacto e menor esforço).
Etapa 5	Implementação dos resultados dos testes de hipóteses.

A etapa 1 foi realizada a partir da análise de dados internos com o sistema que o laboratório de metrologia utilizava. Também foram avaliados dados externos da Receita Federal, mas em função da capacidade de processamento destes, não foi possível incorporá-los ao estudo inicial, sendo um ponto de melhoria previsto na operação.

A etapa 2 partiu da realização de um grupo de foco (*focus group*) com coordenadores do laboratório, visando a apresentação dos achados e a discussão de ações futuras e temas relevantes. A partir desta, foi possível seguir para a etapa 3, que tratou da criação de hipóteses de trabalho. Foram levantadas, na

ocasião, 15 hipóteses para que, na etapa 4, elas fossem alocadas em uma matriz de decisão para a escolha das mais prioritárias.

A última etapa foi a análise efetiva e testes destas hipóteses. A análise aconteceu com o apoio do software Power BI, da Microsoft, a partir da metodologia CRISP-DM (*Cross-Industry Standard Process for Data Mining*) (Wirth e Hipp, 2000). Esta metodologia fornece uma abordagem estruturada para processos de mineração de dados, sendo amplamente utilizada devido à sua poderosa praticidade, flexibilidade e utilidade ao usar a análise para resolver problemas ou desafios.

4. Resultados da análise dos dados

No levantamento referente a Etapa 1 foram encontrados 9.955.003 registros de dados, 245 tabelas no banco de dados, 90.475 serviços gerados entre certificados, suplementos e segundas vias e 208 serviços oferecidos. A partir dos registros foi possível identificar 92 variáveis trabalháveis, que foram separadas em dois blocos: objetos e processos. O objetivo aqui foi identificar as variáveis que poderiam adicionar valor ao laboratório.

O grupo de foco possibilitou a identificação de áreas prioritárias para o posterior desenvolvimento de hipóteses. As áreas prioritárias foram: a análise de lead time; instabilidade do cliente; estabilidade de padrões; análise de retrabalhos; avaliação de contratante vs solicitante; análise do histórico do cliente vs periodicidade de compra; análise de setor vs periodicidade de compra; faturamento; relação do valor do serviço e sua não contratação; estratégias de vendas; e cotação vs aprovação do cliente.

A partir destas grandes áreas, foram desenvolvidas 15 hipóteses de trabalho, descritas no quadro 2.

Quadro 2. Descrição das hipóteses

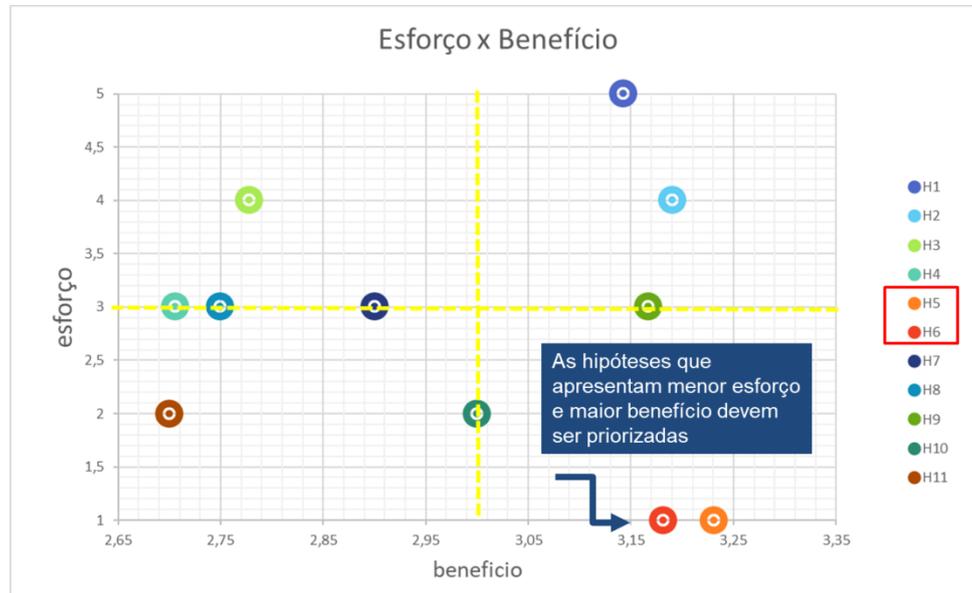
Área prioritária	Objetivo	Hipótese
Análise de Lead Time	Tomar conhecimento e melhorar a precisão nos tempos dos processos de cotação, contratação e execução de um serviço.	H1: O tempo mais crítico do processo é o de contratação.
Análise de Instabilidade cliente	Análise de instabilidade dos instrumentos de um mesmo fabricante/marca e modelo.	H2a: Equipamentos de uma mesma marca/modelo possuem a mesma variação do tempo (correção/ erro sistemático e aleatório). H2b: Equipamentos de uma família possuem a mesma variação do tempo (correção/ erro sistemático e aleatório).
Estabilidade dos padrões	Análise do comportamento da calibração para se estabelecer um padrão de contratação.	H3: é possível traçar uma curva de "normalidade" do comportamento do padrão para se criar uma predição de calibração.
Retrabalhos	Compreender aspectos relacionados a correção de certificados, identificando as famílias para que se possa compreender a relação de retrabalho existente.	H4a: Algumas famílias de serviços são mais "complexas" e possuem maior retrabalho (2ª via - erro processo de calibração). H4b: Alguns setores possuem maior demandas por suplementos.
Análise de Contratante X solicitante	Compreender melhor os demandantes dos serviços para poder direcionar futuros esforços a estes sujeitos.	H5: O cargo do contratante é predominantemente técnico.

Análise do histórico do cliente / periodicidade	Compreender o histórico de compra do cliente para poder atuar de maneira mais ativa nas vendas.	H6: Equipamentos de mesma família possuem o mesmo período para nova cotação.
Análise do setor / periodicidade	Compreender o histórico de compra do cliente para poder atuar de maneira mais ativa nas vendas.	H7: Equipamentos de mesma família e setor possuem o mesmo período para nova cotação
Faturamento	Compreender o faturamento do laboratório a partir de variáveis como família, porte e período.	H8a: Médias empresas contratam mais serviços. H8b: Setor metalmecânico contratam mais serviços.
Relação com o valor do serviço e sua não contratação	Identificar se o valor do serviço executado está relacionado com algumas variáveis que culminam na não contratação do serviço	H9: Serviço de menor preço possuiu maior índice de cancelamento
Estratégias de vendas	Analisar estratégias de vendas considerando aspectos de setores e regiões.	H10a: Há correlação entre os setores e as regiões por meio da análise de vendas. H10b: Diferentes setores apresentam índices de contratação distintos.
Análise de cotação e aprovação do cliente	Esta análise possibilita compreender melhor a relação entre as cotações solicitadas e os serviços efetivamente feitos, mas também ter uma descrição dos serviços não feitos e buscar suas causas.	H11: Mesmo cliente cancelam sempre o mesmo serviço

As hipóteses foram ordenadas em matriz de decisão, onde os participantes do grupo de foco puderam avaliar cada uma delas. A análise de esforço para a sua comprovação foi feita por comparação entre as hipóteses, a partir do tempo estimado pelos participantes para a sua comprovação. Os valores atribuídos foram de 1 a 5, sendo 1 a mais simples e 5 a mais complexa para execução.

A análise dos benefícios foi realizada a partir de critérios de decisão, buscando-se a média ponderada dos seguintes critérios: facilidade de venda como serviço (peso 5), impacto no aumento de vendas (peso 3), impacto na redução de custos internos (peso 2), impacto no aumento de confiabilidade (peso 1) e redução de preço ao cliente (peso 4). O resultado do cruzamento na matriz Esforço X Benefício está evidenciado na figura 1.

Figura 1. Matriz esforço-benefício.



Uma vez selecionadas, as hipóteses H5 e H6 elas foram testadas.

4.1. Teste da H5

O primeiro teste aconteceu com a H5 que visa avaliar se “o cargo do contratante é predominantemente técnico”.

Tabela 1. Dados referentes ao nível funcional do contratante.

Cargo cadastrado	Total	% do Total
Operacional	28181	60,53%
Supervisor/Coord./Responsável	12360	26,55%
Encarreg./Sub.Chefe/Braço Direito/Resp. Parcial	4097	8,80%
Presidente/Diretor	1088	2,34%
Assessor	786	1,69%
Gerente	48	0,10%
Total	46560	100,00%

A partir dos dados levantados pôde-se observar que hipótese foi validada, uma vez que 87,08% dos cargos têm este perfil.

Tal validação implica diretamente nas ações de mercado que vão acontecer por parte do laboratório para aumentar sua demanda, como por exemplo:

- desenvolver marketing de conteúdo, com foco técnico direcionado ao público predominante, em vídeo ou texto;
- disseminar conteúdo técnico por meio de e-mail marketing, blog ou outro canal apropriado;
- encaminhar periodicamente conteúdos - aqui pode-se relacionar com a periodicidade de calibração ou demais hipóteses, como por ex. entregar próximo as datas estimadas de calibração;
- criar ou deixar ativa uma “comunidade” ou ambiente virtual onde os participantes possam interagir e que o laboratório modere o conteúdo e responda aos questionamentos

4.2. Teste da H6

A hipótese 6 versa sobre: Equipamentos de mesma família possuem o mesmo período de nova cotação.

Para esta análise, inicialmente foram levantadas as famílias de equipamentos. Para definir quais famílias seriam avaliadas, foi realizado uma análise de Pareto, levando em consideração o faturamento em cada uma delas. Após a definição de quais famílias seriam avaliadas, realizou-se a contagem de periodicidades ocorridas e a quantificação de recontrações.

A partir dos dados obtidos foi possível inferir que a H6 não é válida. Mesmo que algumas das famílias analisadas possuam similaridades, estas não são configuradas como padrão para a validação desta hipótese.

A figura 2 representa os dados de uma família não validada (calibração de padrões de rugosidade) e a figura 3 traz a distribuição da periodicidade de calibração da trena a laser – 25 metros, a qual foi considerada validada a partir do baixo desvio padrão.

Figura 2. Periodicidade de calibração de padrões de rugosidade dos clientes do laboratório.

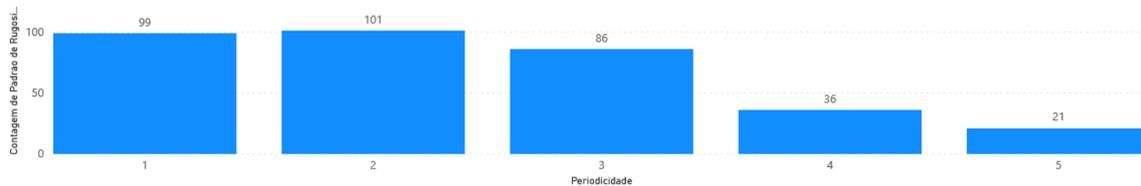
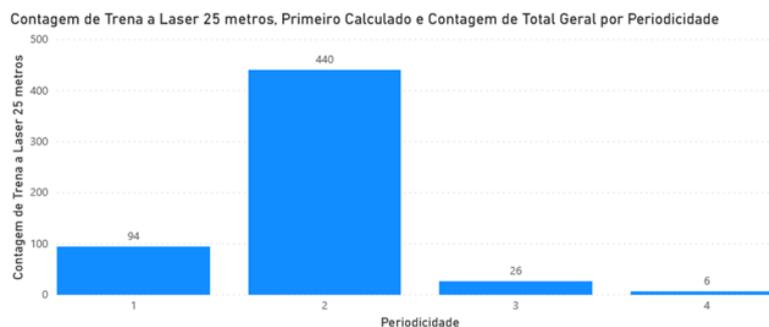


Figura 3. Periodicidade de calibração de padrões de rugosidade dos clientes do laboratório.



Como resultado final, dos 12 serviços estudados, 7 deles podem ter validada a hipótese, no caso considerando-se baixos desvios, enquanto 5 deles apresentaram resultados dispersos.

Apesar de não ser validada a hipótese, o estudo caracterizou-se como fonte valiosa de informação, pois, a partir desta constatação, foi possível traçar alternativas que incentivem os clientes a cotar em períodos regulares. Citam-se como exemplo:

- para os clientes em que a periodicidade for superior à média de mercado, criar conteúdo expondo a importância da regularidade das calibrações ou buscar informações que qualifiquem esse período (a exemplo de normas ou boas práticas mesmo que de outros setores);
- mandar e-mails próximo às datas recomendadas e reforços em períodos a definir, visando prescrição de vendas;
- desenvolver um programa de incentivos para os clientes (por ex. descontos progressivos), para quem for regular em médias de mercado;



- para os clientes regulares, criar conteúdo com informações relevantes, informando a proximidade da data de calibração e os incentivos, gerando relacionamento e fidelidade do cliente.

5. Considerações Finais

A partir das etapas realizadas foi possível avaliar que a mineração e análise de dados de um prestador de serviços tecnológicos podem ser usados para aumentar não somente a competitividade dos laboratórios, mas também das cadeias econômicas por ele atendidas.

Neste trabalho apresentarem-se os resultados obtidos com a criação do “MetroBI”, framework apoiado na metodologia CRISP-DM, para concepção de um Business Intelligence para um prestador de serviço metrológico.

Cabe destacar que todas as ações realizadas estão alinhadas com as tendências tecnológicas de mercado, pois foram consideradas a digitalização de processos e a aproximação com clientes.

Julga-se que os laboratórios de metrologia, os quais possuem diferentes demandas por sua natureza complexa, tragam o uso de Business Intelligence como um forte aliado para a validação de hipóteses e planejamentos operacionais, táticos e estratégicos, tratando seus dados de operação com o mesmo zelo e cuidado que são tratados os dados de calibração e ensaios em seus serviços, considerando suas incertezas e gerando conclusões que criem futuros prósperos.

6. Referências

- Loshin, D. (2003). *Business Intelligence: The Savvy Manager's Guide*. Morgan Kaufmann Publishers.
- McGee, K. (2004). *Gartner Updates its Definition of Real-time Enterprise*. Gartner Inc. Research Note DF-22-2973.
- Toyama, J. (2017). Metodologia para Estudo de Viabilidade de Implantação de Laboratórios de Metrologia. *Dissertação de Mestrado em Metrologia*. Universidade Federal de Santa Catarina. f.157.
- Watson, H. J. & Wixom, B. H. (2007). The Current State of Business Intelligence. *IEEE Computer Society*, V. 40, n. 9.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, 24959, 29–39. <https://doi.org/10.1.1.198.5133>