



UNIVERSIDADE ESTADUAL DO NORTE DO PARANÁ

CAMPUS LUIZ MENEGHEL

LUCIANO DE OLIVEIRA TULIO

**FERRAMENTA DE ANÁLISE ESTATÍSTICA DE
PROCESSO APLICADA A UM AMBIENTE DE BI:
AUXILIANDO NO ENTENDIMENTO DOS DADOS E SUAS
VARIAÇÕES**

Bandeirantes

2010

LUCIANO DE OLIVEIRA TULIO

**FERRAMENTA DE ANÁLISE ESTATÍSTICA DE
PROCESSO APLICADA A UM AMBIENTE DE BI:
AUXILIANDO NO ENTENDIMENTO DOS DADOS E SUAS
VARIAÇÕES**

Trabalho de conclusão de curso
submetido ao *campus* Luiz
Meneghel da Universidade Estadual
do Norte do Paraná, – como
requisito parcial para a obtenção do
grau de bacharel em Sistemas de
Informação.

Orientador: Prof. Ms. André Luís
Andrade Menolli.

Bandeirantes

2010

LUCIANO DE OLIVEIRA TULIO

**FERRAMENTA DE ANÁLISE ESTATÍSTICA DE
PROCESSO APLICADA A UM AMBIENTE DE BI:
AUXILIANDO NO ENTENDIMENTO DOS DADOS E SUAS
VARIAÇÕES**

Trabalho de conclusão de curso
submetido ao *campus* Luiz
Meneghel da Universidade Estadual
do Norte do Paraná, – como
requisito parcial para a obtenção do
grau de bacharel em Sistemas de
Informação.

COMISSÃO EXAMINADORA

Prof. Me. André Luís Andrade
Menolli.

UENP – *Campus* Luiz Meneghel

Prof. Me. Glauco Carlos Silva.

UENP – *Campus* Luiz Meneghel

Prof. Me. Ricardo Gonçalves Coelho

UENP – *Campus* Luiz Meneghel

Bandeirantes, __ de _____ de 2010

DEDICATÓRIA

Primeiramente a Deus, quem me deu forças para nunca desistir. A Carlos e Odete, meus pais. Sem vocês nada seria possível.

AGRADECIMENTOS

A república VóGina, Guilherme, Leandro, Lucas, Carlos, Bertonha, Natan e

André, mas que amigos, irmãos nesta etapa de vida.

Agradeço ao Queiroz que nos fez dar tantas risadas.

Aos meus companheiros de projeto, Daniel, Thiago, Gustavo e Marcio. Marcio que hoje não está presente entre nós, mas continua vivo em nossos corações.

Aos professores, os quais me ajudaram nesta etapa, auxiliando na busca do conhecimento.

Agradeço a tudo!

*“Os ventos que às
vezes tiram algo
que amamos, são
os mesmos que
trazem algo que
aprendemos a
amar...”*

*Por isso não
devemos chorar
pelo que nos foi
tirado e sim,
aprender a amar o
que nos foi
dado. Pois tudo
aquilo que é
realmente nosso,
nunca se vai para
sempre...”*

(Bob Marley)

RESUMO

Muitas empresas e corporações utilizam ferramentas de BI (Business Intelligence) para auxiliá-las no processo de tomada de decisão. Estas ferramentas normalmente possuem módulos de geração de gráficos, o que facilita a visualizações dos dados e auxilia no entendimento dos mesmos. Porém, poucas utilizam métodos que ajudam os tomadores de decisão a identificar pontos críticos ou possíveis problemas. Dessa forma, é proposta uma ferramenta que utiliza análise estatística de processo para auxiliar na compreensão de dados e ajudar a identificar pontos críticos. Para tanto, são traçados objetivos, utilizando KGI (Key Goals Indication) de acordo com os objetivos da empresa. Utilizando estes objetivos pré-determinados, é feita a análise utilizando os KPIs (Key Process Indication), que medem o nível de desempenho do processo, indicando quão bem os processos estão se desenvolvendo, conseguindo assim entender a variação dos dados.

Palavras chaves: Ferramentas de BI, Gráficos, Análise Estatística de Processo, KGI e KPI

ABSTRACT

Many companies and corporations use BI tools (Business Intelligence) to assist them in making decisions. These tools usually include modules for generating graphics, which facilitates the views of the data, and assists in understanding the same. However, few use methods which help decision makers identify critical points or possible problems. Thus, we propose a tool that uses statistical analysis process to assist in understanding the data and help identify critical points. To this end, goals are outlined, using KGI (Key Goals Indication) according to the company's goals. Using these pre-set goals, the analysis is done using KPIs, which measure the performance of the process, indicating how well the processes are developing, so you may understand the variation of the data.

Keywords: BI Tools, Graphics, Statistical Analysis of Procedure, KGI and KPI

LISTA DE SIGLAS

BI – Business Intelligence

DW – Data Warehouse

OLAP – Online Analytical Processing

MDM – Modelagem Multi Dimensional

OLTP – Online Transaction Processing

MDX – Multi Dimensional Expression

CEP – Controle Estatístico de Processo

VOP – Voz do Processo

VOC – Voz do Cliente

AEP – Análise Estatística de Processo

KPI – Chave Indicadora de Processo

KGI– Chave Indicadora de Objetivo

RDBMS – Relational Database Management System

RIA – Rich Internet Application

JSP – Java Server Page

LISTA DE FIGURAS

Figura 1 - Exemplo de um Banco com a Modelagem Dimensional	19
Figura 2 - Exemplo de uma consulta MDX	22
Figura 3 - Arquitetura do Servidor Olap Mondrian.....	23
Figura 4 - Modelo de Controle de Processo	25
Figura 5 - Gráfico Seqüencial dos Déficits Comerciais Mensais dos Estados Unidos 1987-1988 (Wheeler,2001)	29
Figura 6 - Exemplo de um Diagrama de pontos(WHEELER, 2001).....	30
Figura 7 - Comparação de dois gráficos seqüenciais.....	31
Figura 8 - Exemplo diagrama de Pontos de Médias (WHEELER, 2001).....	32
Figura 9 - Exemplo Diagrama de Pontos Amplitude.....	33
Figura 10 - Exemplo de um gráfico seqüencial e sua média (WHEELER, 2001)	36
Figura 11- Diagrama de Atividades que compõe o processo da Análise Estatística do Processo	39
Figura 12 - Interface de Login da Ferramenta de BI BRAIN.....	40
Figura 13 - Tela Principal da ferramenta de BI BRAIN.....	40
Figura 14 - Ilustrando o local que realiza a escolha do cubo e das medidas ...	41
Figura 15 - Padrão do XML gerado para a ferramenta de análise estatística de processo.....	41
Figura 16 - Ilustrando a Dimensão Data que é necessária para realização da análise estatística de processo	42
Figura 17 - Campos para escolha da medida e KGI	43
Figura 18 - Diagrama de Casos de Uso	43
Figura 19 - Diagrama de Sequencia do processo Realizar Análise Estatística de Processo	46
Figura 20 - Estrutura de Comunicação entre a Ferramenta de BI BRAIN e a Ferramenta de Análise Estatística de Processo.....	46
Figura 21 - Código fonte do HTTPService acessando a ferramenta de BI BRAIN	47
Figura 22 - Análise Realizada na Ferramenta de BI BRAIN.....	48
Figura 23 - Análise Estatística de Processo aplicada na medida Store Sales .	48
Figura 24 - Análise Estatística de Processo aplicada na medida Unit Sales....	49

LISTA DE TABELAS

Tabela 1- Visão mais rica dos dados do que seria apresentado por um banco de dados	21
Tabela 2 - Déficits Mensais dos EAU(U\$\$ Bilhões) (WHEELER, 2001)	28
Tabela 3 - Descrição do Caso de Uso Realizar Analise	44
Tabela 4 - Casos de Uso Realizar Análise Estatística	44
Tabela 5 - Descrição do Caso de Uso Escolher Medida	45

SUMÁRIO

RESUMO

ABSTRACT

1	INTRODUÇÃO.....	13
1.1	OBJETIVOS.....	14
1.1.1	OBJETIVOS GERAIS.....	14
1.1.2	OBJETIVOS ESPECIFICOS	14
1.2	JUSTIFICATIVA.....	14
1.3	ORGANIZAÇÃO DO TRABALHO.....	15
2	FUNDAMENTAÇÃO TEÓRICA.....	16
2.1	BI (BUSINESS INTELLIGENCE)	16
2.2	DW (DATA WAREHOUSE).....	17
2.3	MODELAGEM DIMENSIONAL.....	18
2.4	OLAP (ONLINE ANALYTICAL PROCESSING)	20
2.5	SERVIDOR OLAP MONDRIAN	21
	• MDX (MULT DIMENSIONAL EXPRESSION)	22
2.6	CEP (CONTROLE ESTATÍSTICO DE PROCESSO).....	24
	• CONTROLE.....	24
2.6.1	SISTEMA DE CONTROLE DE PROCESSO.....	24
2.7	ANÁLISE ESTATÍSTICA DE PROCESSO.....	26
2.7.1	APRESENTANDO OS DADOS EM UM CONTEXTO	27
2.8	GRÁFICOS	28
2.8.1	GRÁFICOS SEQUENCIAIS.....	29
2.8.2	DIAGRAMAS DE PONTOS.....	30
2.8.3	COMPARANDO GRAFICOS SEQUENCIAIS.....	30
2.8.4	GRÁFICOS DE CONTROLE	31
	• AMPLITUDE	33

2.9	PRINCIPIOS PARA ENTENDER OS DADOS	33
2.9.1	PRIMEIRO PRINCIPIO.....	33
2.9.2	SEGUNDO PRINCIPIO	34
2.9.3	ABORDAGEM DE SHEWHART	35
2.9.5	METODOLOGIA.....	38
3	DESENVOLVIMENTO	39
3.1	DIAGRAMA DE SEQUENCIA.....	45
3.2	ESTRUTURA DA FERRAMENTA	46
4	ESTUDO DE CASO	48
5	RESULTADOS.....	50
6	CONCLUSÃO	51
6.1	TRABALHOS FUTUROS.....	51
	REFERÊNCIAS.....	53

1 INTRODUÇÃO

O início da história do BI foi na década de 70 e as ferramentas de BI se tornam cada vez mais eficiente principalmente no ambiente web o qual tem um constante crescimento. As ferramentas de BI auxiliam os analistas de negócio a tomada de decisões. Estas ferramentas organizam os dados e os deixam mais claros e fáceis de visualizar, por meio de gráfico e tabelas, o que permite uma melhor interpretação(KIMBAL, 2002).

Através da análise de dados históricos e atuais, os gestores conseguem valiosas informações que o auxilia a tomada de decisão.

Para facilitar à visualização dos dados as ferramentas de BI utilizam gráficos. Os gráficos podem remover os detalhes extrínsecos, que não pertence à essência do contexto, fornecidos pelas tabelas de números e ao mesmo tempo enfocam na parte interessante da informações que podem estar contidas nos dados.

Com estes dados nenhuma comparação entre dois valores pode ser global, os relatórios gerenciais estão cheios de comparações limitadas e os gráficos tornam os dados mais acessíveis para a mente humana do que as tabelas.

Segundo Wheeler (2001) nenhum dado tem significado fora de seu contexto, dessa forma a análise estatística de processo auxilia na interpretação dos dados e facilita o entendimento de suas variações. Essas variações são mais visíveis nos gráficos, portanto com a análise estatística de processo aplicada à gráficos, é possível entender melhor os dados e suas variações.

Nas ferramentas de BI os gráficos são importantes instrumentos de visualização de dados e a combinação destes com uma técnica de análise estatística de processo, auxilia na interpretação dos dados e na detecção de dados que influenciam diretamente no processo.

Portanto, neste trabalho foi desenvolvida uma ferramenta gráfica de análise estatística de processo que proporciona uma melhor visualização dos dados, auxiliando o analista de negócio a identificar dados que sofrerão

variação especial, a qual influencia diretamente no processo. Esta ferramenta está integrada a uma ferramenta de BI chamada BRAIN, para a comunicação entre as ferramentas foi utilizado o serviço HTTPservice, um protocolo de comunicação HTTP simples.

1.1 OBJETIVOS

1.1.1 OBJETIVOS GERAIS

O objetivo desse trabalho foi o desenvolvimento de uma ferramenta de análise estatística de processo aplicada a um ambiente de BI, analisando dados históricos e entendendo suas variações para detectar dados que sofrerão variações especiais, que realmente influenciam o processo, auxiliando na tomada de decisões.

1.1.2 OBJETIVOS ESPECIFICOS

- Definir os gráficos adequados para a Análise Estatística de Processo;
- Definir um padrão para extrair os dados da ferramenta de BI;
- Modelar uma ferramenta para visualização gráfica com análise estatística de processo;
- Implementar uma ferramenta de Análise Estatística de Processo em gráficos no ambiente de BI;
- Testar esta ferramenta, integrada com uma ferramenta de BI BRAIN.

1.2 JUSTIFICATIVA

Este trabalho visa combinar uma ferramenta de BI e um método de análise estatística de processo, criando assim uma ferramenta que auxilia o usuário a detectar dados que realmente possam influenciar no processo.

A combinação de ambas as tecnologias sanam a ausência de um instrumento na ferramenta de BI BRAIN que detecta as variações que realmente influenciam na gestão das atividades relacionadas.

1.3 ORGANIZAÇÃO DO TRABALHO

Na Seção 2 é apresentada a fundamentação teórica. A Seção 3 apresenta a metodologia. A Seção 4 o desenvolvimento do trabalho, a Seção 5 apresenta um estudo de caso. A Seção 6 apresenta os resultados do estudo de caso. A Seção 7 apresenta uma conclusão sobre o trabalho desenvolvido e sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta uma visão geral sobre os principais conceitos de Business Intelligence, Data Warehouse, Modelagem Dimensional, OLAP, Servidor OLAP Mondrian, Gráficos, Análise Estatística de Processo.

2.1 BI (BUSINESS INTELLIGENCE)

A década de 70 deu início à história do Business Intelligence (BI), quando foram fornecidos para os analistas de negócios, os primeiros produtos de BI. Os maiores problemas na época eram que uma aplicação de BI exigia muito esforço de programação, pouca flexibilidade e baixo desempenho (SERRA, 2002).

Hoje, com um cenário bem diferente, as ferramentas de BI se tornam cada vez mais eficiente, principalmente no ambiente web o qual tem um constante crescimento.

As vantagens do uso de sistemas de BI baseados na Web, quando comparados com os sistemas tradicionais, incluem facilidades de uso, acesso universal às diversas plataformas, possibilidade de configuração para diversos níveis de usuários e capacidade de trabalho com gráficos, sons e vídeos entre outros (FORTULAN; GONÇALVES FILHO, 2005).

Através da análise de dados históricos e atuais, os gestores conseguem valiosas informações para servir de base para a tomada de decisão. O processo de BI baseia-se na transformação de dados em informações, após isto transformando as informações em decisões. (TURBAN; EFRAIM, 2009).

Uma ferramenta de BI é composta de um repositório de dados chamada Data Warehouse, onde são armazenados todos os dados históricos, um servidor OLAP que faz consulta on-line no Data Warehouse e um software para visualização dos dados, que mostra os resultados da análise feita pelo servidor OLAP no Data Warehouse. Nas próximas sessões serão explicados os conceitos de Data Warehouse, Servidor OLAP.

2.2 DW (DATA WAREHOUSE)

Um dos bens mais preciosos de qualquer empresa são suas informações, e quase sempre a empresa mantém esse bem de duas formas: os sistemas operacionais de registro e o data warehouse. Em termos gerais, os sistemas operacionais são o local em que os dados são colocados, o data warehouse é o local a partir de onde eles são obtidos. (KIMBAL, 2002)

Segundo INMON(1997):

“Data Warehouse é uma coleção de dados orientados por assuntos, integrados, variáveis com o tempo e não volátil, para dar suporte ao processo de tomada de decisão, trata-se de um processo em andamento que aglutina dados de fontes heterogêneas, incluindo dados históricos.”

Segundo Kimbal (2002) os praticantes e especialistas de *Data Warehouse* reconhecem que a apresentação deve ser fundamentada na simplicidade para que haja alguma chance de sucesso.

Objetivos podem ser desenvolvidos percorrendo os setores da empresa e escutando o que a gerência da empresa tem a dizer, estas questões recorrentes apareceram quando:

- Dispõe-se de um enorme volume de dados a partir de fontes diversas;
- Tem-se que combinar dados a partir de fontes diversas;
- É Preciso tornar mais fácil para os profissionais acessarem diretamente os dados;
- Apenas mostre o que é importante;
- As pessoas usem informações para dar um suporte a tomadas de decisões baseando-se em fatos.

O conteúdo do Data Warehouse deve ser compreensível. Os dados devem ser intuitivos e óbvios para o usuário da área de negócios e não apenas para o desenvolvedor. A capacidade de compreensão envolve legibilidade, o conteúdo precisa ser identificado de modo significativo. Os usuários das áreas de negócio querem separar e combinar os dados no

warehouse em combinações infinitas, um processo que normalmente costuma ser denominada “separação e combinação” (slicing and dicing).

As ferramentas usadas para acessar o Data Warehouse devem ser simples e fáceis de serem usadas. Elas também devem retornar para os usuários resultados de consultas no menor intervalo de tempo de espera possível. (KIMBAL, 2002)

A modelagem de um DW é diferente de um banco de dados relacional, é utilizada uma modelagem específica, a modelagem dimensional que será apresentada a sessão seguinte.

2.3 MODELAGEM DIMENSIONAL

Segundo Kimbal (2002), a modelagem relacional é utilizada nos bancos de dados tradicionais. Nesta geralmente são utilizadas formas normais para evitar ambigüidades e repetições desnecessárias nos elementos do modelo.

Na Modelagem Multidimensional (MDM), os dados, na maioria das vezes, são desnormalizados e estruturados em diversas dimensões. Neste tipo de modelagem há uma tabela dominante no centro do diagrama com múltiplas junções conectando esta tabela às outras tabelas satélites, denominadas tabelas dimensionais. As tabelas dimensionais, também chamadas de tabelas de dimensão, armazenam as descrições textuais das dimensões do negócio. Cada uma dessas descrições textuais ajuda a definir um componente da respectiva dimensão. (KIMBAL, 2002)

Kimbal (2002) afirmou que:

“Modelagem dimensional é um nome novo para uma técnica antiga usada para criar banco de dados simples e compreensíveis.”

Modelos multidimensionais proporcionam uma estrutura de sistemas de informação que permitem que uma empresa disponha de acesso flexível a dados, fatie e agrupe dados de qualquer número de maneiras explore dinamicamente o relacionamento entre dados resumidos e detalhados (INMON, 1997).

O modelo dimensional fornece subsídios para que se respondam perguntas como “Qual é o total de vendas de cadeiras vermelhas em Curitiba no mês de março?”, Harrison (1998), ressalta que o modelo é referenciado com freqüência através do termo “esquema estrela” devido à organização de seu projeto lógico de banco de dados.

O Modelo dimensional combina tabelas de armazenamento de dados históricos em séries temporais, indexados em chaves dimensionais, descritas em tabelas dimensionais correspondentes (Harrison, 1998). Ou seja, é organizado através de tabelas de fato, que são as ocorrências de transações de negócios da empresa, e tabelas de dimensões que possuem o papel descritivo nessas transações. Um exemplo desse modelo é mostrado na Figura 1.

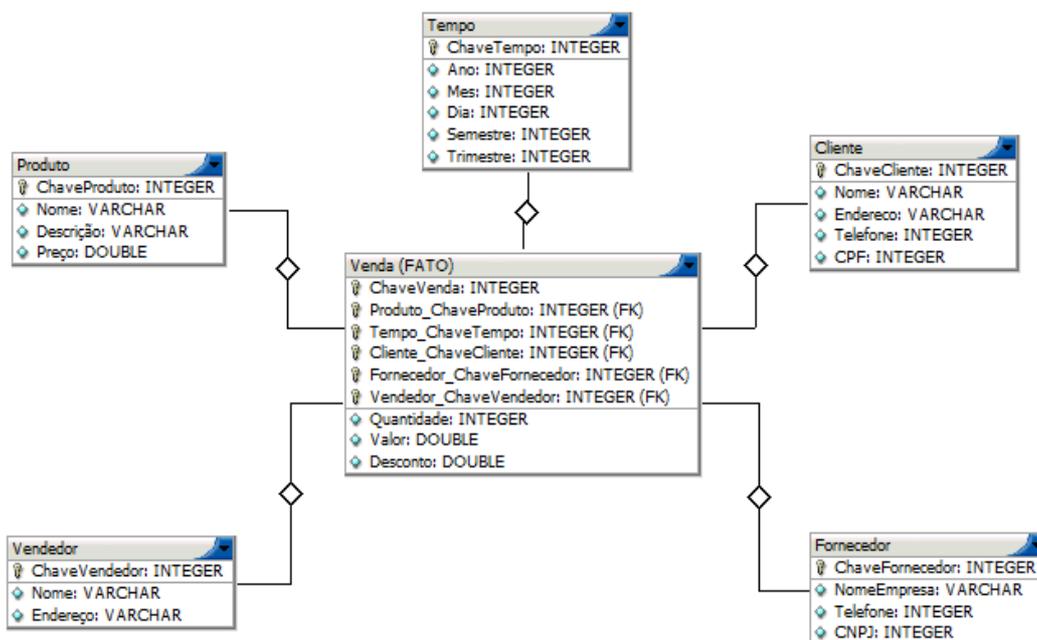


Figura 1 - Exemplo de um Banco com a Modelagem Dimensional

Tabelas dimensionais guardam em sua maioria informações textuais, as quais ajudam a definir um componente da dimensão do negócio, elas possuem dados sobre produtos, mercados, períodos de tempo. As tabelas de fatos por sua vez contêm as medidas do negócio, além de conterem as informações decorrentes das tabelas de dimensões, isso garante a precisão do acesso aos dados através de uma estrutura de chave completa, eliminando

assim pesquisas em tabelas e resulta em maior desempenho possível (HARRISON, 1998).

Na próxima sessão é apresentada a tecnologia OLAP que é necessária para ter acesso aos dados do DW.

2.4 OLAP (ONLINE ANALYTICAL PROCESSING)

Online Analytical Processing (OLAP): análise de grandes quantidades de dados em tempo real. Ao contrário de Online Transaction Processing (OLTP), a qual suas operações típicas, ler e modificar os números de indivíduos, em OLAP suas operações são somente leitura online (PENTAHO, 2010).

Segundo Pentaho (2010), o termo implica que, apesar de enorme quantidade de dados estarem envolvida, geralmente muitos milhões de registros, que ocupa vários gigabytes, o sistema deve responder a consultas rápidas o suficiente para permitir uma exploração interativa de dados, o que apresenta consideráveis desafios técnicos.

OLAP emprega uma técnica chamada Análise Multidimensional. Considerando que um banco de dados relacional armazena todos os dados na forma de linhas e colunas, um conjunto de dados multidimensional consiste de eixos e células (PENTAHO, 2010). Considere o conjunto de dados apresentado na Tabela1.

Tabela 1- Visão mais rica dos dados do que seria apresentado por um banco de dados

Ano	2000		2001		Crescimento	
Produto	Venda do Dólar	As vendas unitárias	Venda do Dólar	As vendas unitárias	Venda do Dólar	As vendas unitárias
Total	\$7.073	2693	\$17.401	3008	8%	12%
Livros	\$2.753	824	\$7.636	966	21%	17%
Ficção	\$1.341	424	\$3.331	380	-10%	-10%
Não Ficção	\$1.412	400	\$2.129	586	51%	47%
Revistas	\$2.753	824	\$2.426	766	-12%	-7%
Cartões de Votos	\$1.567	1045	\$1.879	1276	20%	22%

O eixo de linhas é constituído de todos os produtos 'Livros', 'Ficção', e assim por diante, e o eixo colunas consiste no produto cartesiano do ano '2000' e '2001', o cálculo do Crescimento, as vendas unitárias e venda do dólar. Cada célula representa a venda de uma categoria de produto em um determinado ano, por exemplo, as vendas do dólar de Revistas em 2001 foram 2.426 \$.

Segundo Pentaho (2010), esta é uma visão mais rica dos dados do que seria apresentado por um banco de dados relacional. Os membros de um conjunto de dados multidimensional não são sempre os valores de uma coluna relacional. 'Total', 'Livros' e 'Ficção' são membros em níveis sucessivos, em uma hierarquia, cada um dos quais é enrolado para a próxima. E mesmo que seja ao lado dos anos '2000' e '2001', 'Crescimento' é um membro calculado, que introduz uma fórmula para as células de computação de outras células.

As dimensões usadas - produtos, tempo e medidas - são apenas três das muitas dimensões em que o conjunto de dados pode ser classificado e filtrado. A coleção de dimensões, hierarquias e medidas chama-se cubo.

2.5 SERVIDOR OLAP MONDRIAN

Existem vários servidores Olap, neste trabalho adota-se o servidor Mondrian, um servidor Olap Open Source.

Segundo Pentaho (2010), o Mondrian é um servidor OLAP, escrito em Java, ele executa consultas escritas na linguagem MDX(Multi Dimensional Expression) e converte para SQL(Structured Query Language), lendo dados de bases relacionais e apresentando os dados no formato multidimensional, através de uma API Java.

- MDX (MULT DIMENSIONAL EXPRESSION)

MDX representa expressões multidimensionais. É a principal linguagem de consulta implementada por Mondrian (PENTAHO, 2010).

Segundo Pentaho (2010), o MDX foi introduzido pela Microsoft com a Microsoft SQL Server em torno de 1998, o MDX apareceu como parte do XML para API de análise. A Microsoft propôs o MDX como um padrão, e sua aprovação entre os escritores e outros prestadores de aplicação OLAP são cada vez maiores.

Uma consulta MDX simples é mostrada na Figura 2.

```
SELECT {[Medidas].[UnidadesVendidas], [Medidas].[VendasdaLoja]} ON
        COLUMNS,
        {[Produtos].membros} ON ROWS
FROM [Vendas]
WHERE [Data].[1997].[Q2]
```

Figura 2 - Exemplo de uma consulta MDX

O sistema Mondrian consiste em quatro camadas, mostradas na Figura 3, trabalhando a partir da apresentação para o usuário final para as entranhas do centro de dados, estes são os seguintes: camada de apresentação, camada dimensional, camada estrela e camada de armazenamento (PENTAHO, 2010).

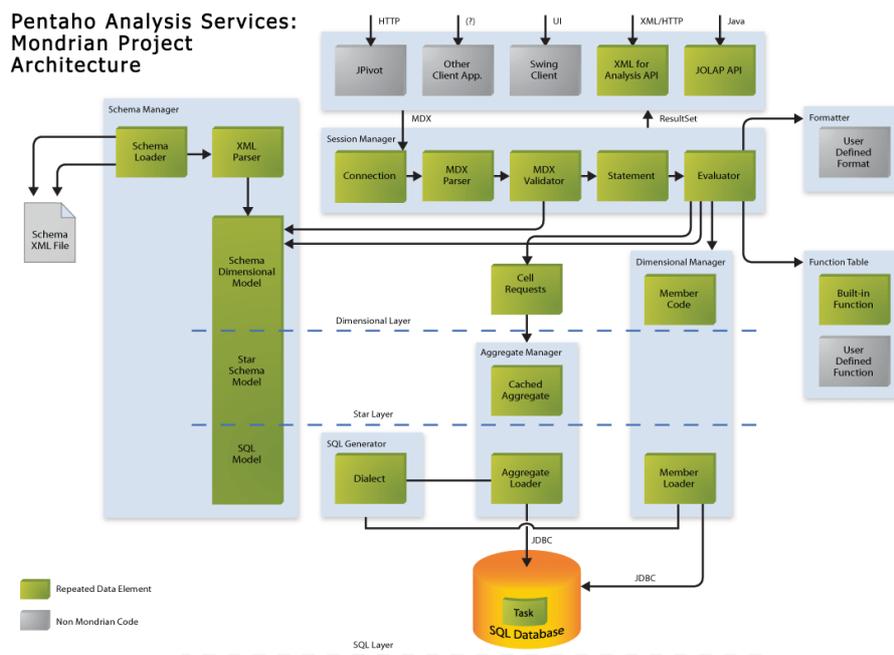


Figura 3 - Arquitetura do Servidor Olap Mondrian

A camada de apresentação determina segundo Pentaho (2010), o que o usuário final vê em seu monitor, e como pode interagir e pedir novas questões. Existem muitas formas de apresentar dados multidimensionais, incluindo tabelas, gráficos de pizza, gráficos de linha ou barra entre outros tipos de gráficos, também podem ter visualizações avançadas, ferramentas com mapas e gráficos dinâmicos e clicáveis.

A segunda camada é a camada dimensional, tem por finalidade analisar, validar e executar consultas MDX. A consulta é avaliada em várias fases. Os eixos são primeiro computados, então os valores das células dentro dos eixos. Para a eficácia, a camada dimensional envia pedidos para a camada de agregação em lotes. Um transformador de consulta permite a aplicação de manipulação de queries existentes, ao invés de construir uma declaração MDX a partir do zero para cada solicitação. (PENTAHO, 2010).

A terceira camada é a camada estrela, que é responsável pela manutenção de um agregado cache. Uma agregação é um conjunto de valores medida na memória, qualificado por um conjunto de dimensão, coluna e valores. A camada dimensional envia pedidos de conjuntos de células. Se as células não são encontradas no cache, a agregação gerente envia um pedido para a camada de armazenamento (PENTAHO, 2010).

A camada de armazenamento é um RDBMS. É responsável por fornecer dados agregados de células, e os membros da mesma dimensão (PENTAHO, 2010).

Nas próximas sessões serão apresentados os conceitos de análise estatística de processo, a definição de seus componentes e explicação dos métodos estatísticos que foram adotados para o desenvolvimento da ferramenta.

2.6 CEP (CONTROLE ESTATÍSTICO DE PROCESSO)

O controle estatístico de processo segundo (PORTALACTION, 2010) tem por objetivo desenvolver e aplicar métodos estatísticos como parte da estratégia de prevenção de possíveis defeitos, da melhoria da qualidade dos produtos e serviços e da redução de custos de fabricação.

- CONTROLE

O controle é segundo (PORTALACTION, 2010) o conjunto de decisões que tem por objetivo a satisfação de determinados padrões ou especificações por parte dos produtos focados no cliente.

2.6.1 SISTEMA DE CONTROLE DE PROCESSO

Segundo (PORTALACTION, 2010) quatro elementos destes sistemas são importantes para as discussões a seguir.

- O Processo

Entende-se como processo a combinação de fornecedores, produtores, pessoas, equipamentos, materiais de entrada, métodos e meio ambiente que trabalham juntos para produzir o resultado, e os clientes correspondem aos elementos que utilizam o resultado, a Figura X ilustra a relação mostra essa relação.

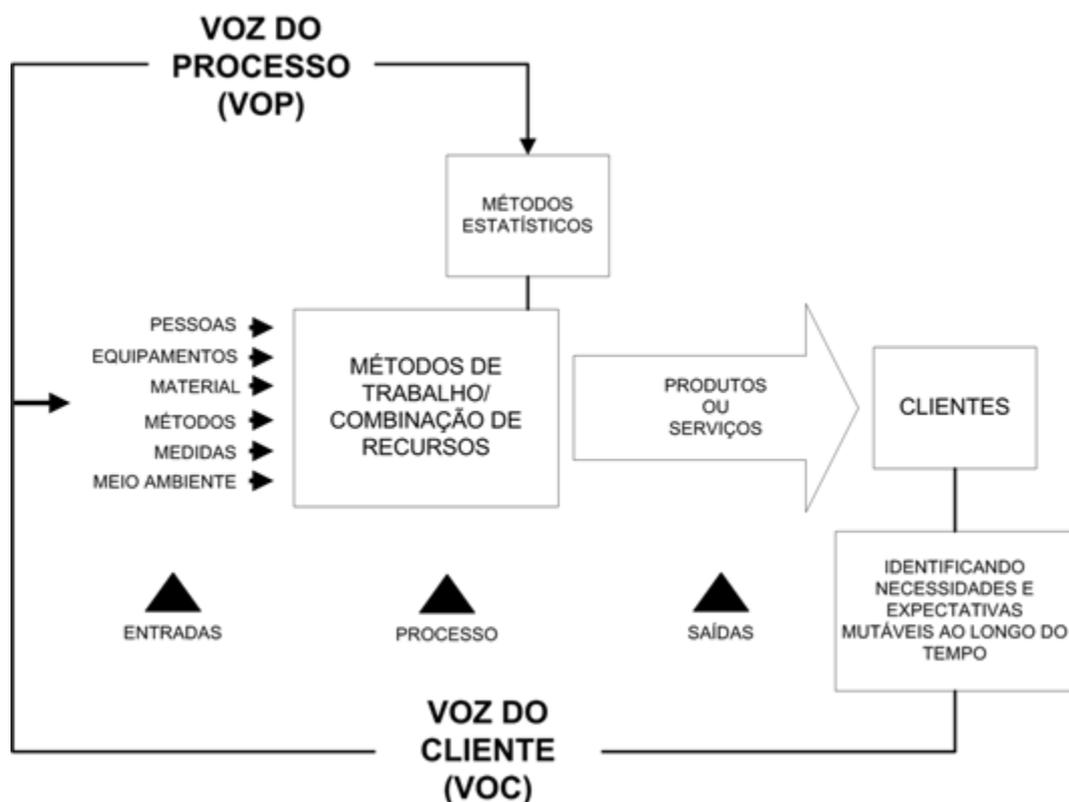


Figura 4 - Modelo de Controle de Processo

- Informações sobre o desempenho:

Muita informação sobre o real desempenho do processo pode ser aprendida através de estudo do resultado do processo, a informação mais útil sobre o desempenho de um processo vem da compreensão do processo e de sua variabilidade (PORTALACTION, 2010).

- Ações sobre o processo:

Uma ação sobre o processo é segundo Portalaction (2010) mais econômica quando realizada para prevenir que as características importantes variem muito em relação aos seus valores-alvo. Os efeitos das ações deveriam ser monitorados, para que uma análise e ação posterior pudessem ser tomadas, se necessária.

- Ações sobre o resultado:

Uma ação sobre o resultado é menos econômica segundo Portalaction (2010) quando se restringe a detecção e correção do produto fora da especificação. Se o resultado atual não atinge consistentemente os requisitos exigidos pelo cliente, pode ser necessário classificar todos os produtos e refugar ou retrabalhar quaisquer itens não-conformes. Esta atitude

deve ser mantida até que a ação corretiva necessária sobre o processo tenha sido tomada e verificada.

2.7 ANÁLISE ESTATÍSTICA DE PROCESSO

Nós vivemos na “Era da informação”, segundo Wheeler (2001), muita dessa informação chega até nós sob a forma de números e tabelas. Onde quer que se olhe, existem montanhas de dados cada vez maiores, índices, pesquisas de opinião, relatórios de mercado, indicadores, relatórios de governo, equilíbrio das balanças entre outros..

Segundo Wheeler (2001), apesar do aumento das bases de dados e das planilhas, apesar da velocidade com que se pode estar conectados às redes e ter acesso a dados de relatórios e de cálculos, pode-se ser informado de que a produtividade está caindo, a base fabril está sendo corroída e a economia ficando fraca. Não obstante, têm-se ainda máquinas maiores e números maiores, e mais acesso a elas; embora se esteja avançando a passos largos em muitas áreas da ciência básica, que dizem que a vantagem tecnológica está caindo e que os outros estão assumindo a liderança, transferindo o conhecimento científico para os produtos empregados para o consumo da sociedade.

Mesmo dispondo de mais números do que jamais anteriormente, ainda não se sabe o que os mesmos significam. Se os números mudarem pra melhor, comparados ao último mês, então é só esperar, eles podem mudar pra pior em seguida.

Wheeler (2001) afirma que:

“Se os números realmente mudarem para pior, comparado ao último mês, então o apocalipse está perto e todos estarão arruinados! O chefe está desesperado – Não fica parado apenas, faça alguma coisa! Você tem que chegar com uma explicação sobre o porquê de os números estarem tão ruins.”

Segundo Wheeler (2001) o problema sobre a nossa era da informação foi sucintamente exposto por Daniel Boorstin, quando declarou: “A informação é aleatória e heterogênea, mas o conhecimento é metódico e cumulativo. Antes que a informação possa ser útil, ela deve ser analisada,

interpretada e assimilada. Os dados não trabalhados têm de ser digeridos, antes que possam se tornar úteis.

Esse processo de digestão dos dados tem sido amplamente negligenciado em todos os níveis de nosso sistema educacional. Segundo Wheeler (2001) as pessoas terminam sua formação escolar sabendo como somar, subtrair, dividir e multiplicar, mas não entendem nada sobre como interpretar os números para extrair o conhecimento que pode estar encerrado nesses dados.

Essa deficiência tem sido chamada de “simplicidade numérica”. Ela não consiste em um fracasso na aritmética, mas é um fracasso em saber como usar as ferramentas básicas para entender os dados. Segundo Wheeler (2001), a simplicidade numérica não é atendida pelos cursos tradicionais de ensino fundamental e médio, nem pelos cursos avançados de matemática.

Wheeler (2001) afirma que:

“Mesmo as pessoas com alto nível educacional, podem ser numericamente ignorantes”

O uso das técnicas apresentadas pode ter conseqüências profundas tanto para as pessoas individuais como para organizações. A simplicidade numérica somente pode ser superada através da prática (WHEELER, 2001).

2.7.1 APRESENTANDO OS DADOS EM UM CONTEXTO

Embora seja simples comparar um número com outro, tais comparações são limitadas e frágeis. Segundo Wheeler (2001), elas são limitadas devidas à quantidade de dados utilizado e frágil porque os números estão sujeitos à variação que está inevitavelmente presente nos dados do mundo real.

Já que ambos os valores, o atual e o anterior, estão sujeitos a essa variação, será difícil determinar apenas quanto da diferença entre os valores deve-se à variação, e quanto da diferença, se houve, deve-se às mudanças reais (WHEELER,2001).

Wheeler (2001) afirma que se as comparações limitadas não são suficientes, então, que tal simplesmente apresentar todos os valores em uma tabela? Embora seja útil, não é o suficiente para fornecer uma análise completa. A Tabela 2 mostra um exemplo de tabela de valores dos déficits dos Estados Unidos.

Tabela 2 - Déficits Mensais dos EAU(U\$\$ Bilhões) (WHEELER, 2001)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1987	10,7	13,0	11,4	11,5	12,5	14,1	14,8	14,1	12,6	16,0	11,7	10,6
1988	10,0	11,4	7,9	9,5	8,5	11,8	10,5	11,2	9,2	10,1	10,4	10,5

Segundo Wheeler (2001), embora as tabelas de valores possam ser usadas para apresentar todos os dados relevantes, elas não são fáceis de serem resumidas. É difícil diferir a informação contida nas tabelas de números. O que não é facilmente assimilado é, geralmente, difícil de ser comunicado para os outros. A mente humana não executa muito bem o trabalho de absorver grande quantidade de dados.

A solução para esses problemas está em uma mudança de uma representação digital dos dados para uma forma que seja mais favorável para as pessoas, os gráficos (WHEELER, 2001).

2.8 GRÁFICOS

Os gráficos podem remover os detalhes extrínsecos fornecidos pelas tabelas de números, ao mesmo tempo em que enfocam a parte interessante da informação que podem estar contidas nos dados. Segundo Wheeler (2001), os gráficos propiciam o contexto para que o valor atual seja interpretado, porque eles incluem todos os valores prévios relevantes, mas fazem isso em uma maneira mais visual do que digital, o que torna a informação mais fácil de ser assimilada.

Há dois gráficos básicos que provaram ser válidos para realização da análise estatística de processo segundo (WHEELER, 2001):

- gráfico seqüencial;
- diagrama de pontos.

2.8.1 GRÁFICOS SEQUENCIAIS

Os gráficos seqüenciais (Gráficos de séries temporais) tipicamente têm os meses ou anos marcados fora, no eixo horizontal, e os possíveis valores marcados fora no eixo vertical. Quando se movimenta da esquerda para a direita, há uma passagem de tempo. As mudanças do gráfico são vistas quando examina cuidadosamente da esquerda para a direita.

A Figura 5 mostra um exemplo de gráfico seqüencial que mostra os Déficits dos Estados Unidos entre 1987 e 1988.

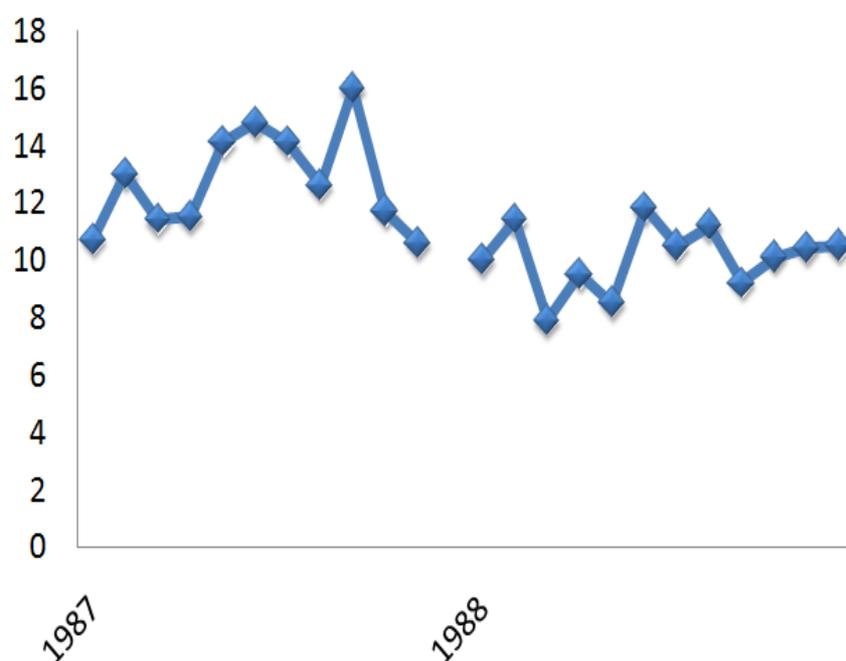


Figura 5 - Gráfico Seqüencial dos Déficits Comerciais Mensais dos Estados Unidos 1987-1988 (Wheeler,2001)

Wheeler (2001) afirma que:

“Os gráficos seqüenciais transmitem o conteúdo de um conjunto de dados de maneira mais rápida e completa do que as tabelas de valores.”

2.8.2 DIAGRAMAS DE PONTOS

Um diagrama de pontos (ou histograma) segundo Wheeler (2001) é simplesmente uma acumulação dos diferentes valores, quando eles ocorrem sem tentar expor a seqüência de tempo.

É colocada uma marca ao lado de um valor possível, cada vez que aquele valor for observado na seqüência. Se os possíveis valores estão escritos no eixo horizontal, então o eixo vertical representará as freqüências para os diferentes valores observados (WHEELER, 2001).

A Figura 6 exemplifica um Diagrama de pontos do circuito de Beisebol de 1960 a 1966 de dois jogadores, Babe Ruth e Roger Maris.

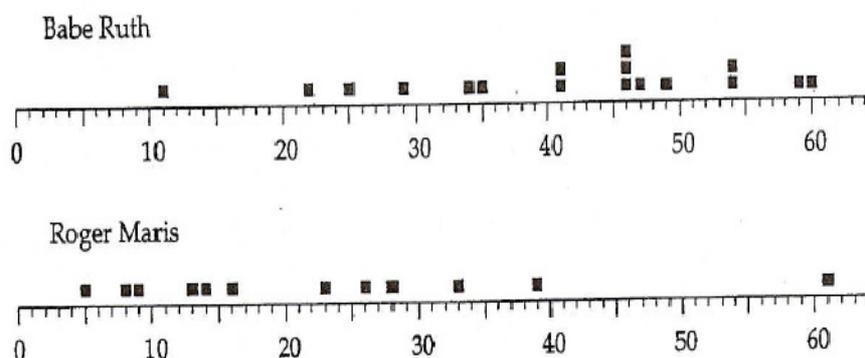


Figura 6 - Exemplo de um Diagrama de pontos(WHEELER, 2001).

Wheeler (2001) afirma que:

“Os diagramas de pontos condensam os dados mostrados no gráfico seqüencial.”

2.8.3 COMPARANDO GRAFICOS SEQUENCIAIS

Pode-se obter uma compreensão adicional usando-se dois gráficos seqüenciais juntos. Qualquer que seja a razão para esse paralelismo, os valores são mais bem compreendidos quando colocados em um contexto amplo (WHEELER, 2001).

A Figura 7 exemplifica a comparação do gráfico de seqüência entre os companheiros de equipe Roger Maris e Mickey Mantle.

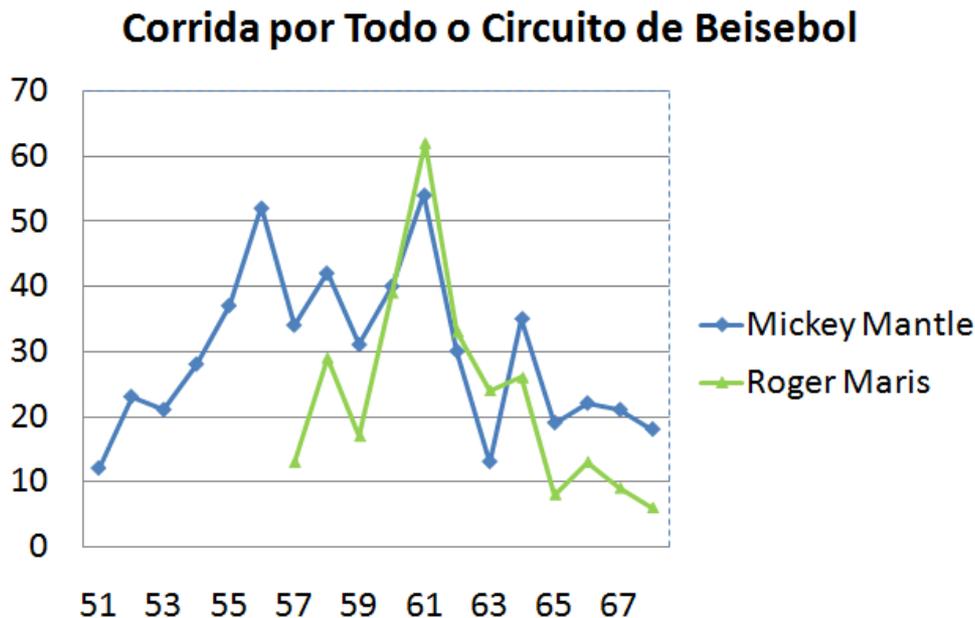


Figura 7 - Comparação de dois gráficos seqüenciais.

2.8.4 GRÁFICOS DE CONTROLE

Carta de controle é um tipo de gráfico utilizado para o acompanhamento de um processo. Segundo Portalaction (2010) este gráfico determina estatisticamente uma faixa denominada limites de controle, que é limitada pela linha superior e uma linha inferior e uma linha média. O objetivo é verificar se o processo está sob controle, isto é, isento de variações especiais.

Para distinguir as variações do processo que anteriormente classificados em comuns e especiais, foi desenvolvida a Cartas de Controle ou Gráficos de Controle que detecta as variações especiais (PORTALACTION, 2010).

Segundo Portalaction(2010) as funções destes gráficos são:

1. “Mostrar evidências de que um processo esteja operando em estado de controle estatístico e dar sinais de presença de causas especiais de variação para que medidas corretivas apropriadas sejam aplicadas”.
2. “Manter o estado de controle estatístico estendendo a função dos limites de controle como base de decisões”.

3. “Apresentar informações para que sejam tomadas ações gerenciais de melhoria dos processos”.

Para construir um gráfico do comportamento é necessário algumas funções para delimitar seus limites, conhecida como sínteses numéricas.

Wheeler(2001) afirma que:

“É útil computar as sínteses numéricas para um conjunto de dados.”

A síntese numérica básica é a média aritmética. Essa medida é simplesmente a soma dos valores dividida pelo número de valores (WHEELER, 2001).

- MÉDIA

A Figura 8 exemplifica a média entre os três jogadores Babe Ruth entre 1918 e 1934, Roger Maris entre 1957 e 1968 e Mickey Mantle por todo o circuito de beisebol de 1951 e 1968.

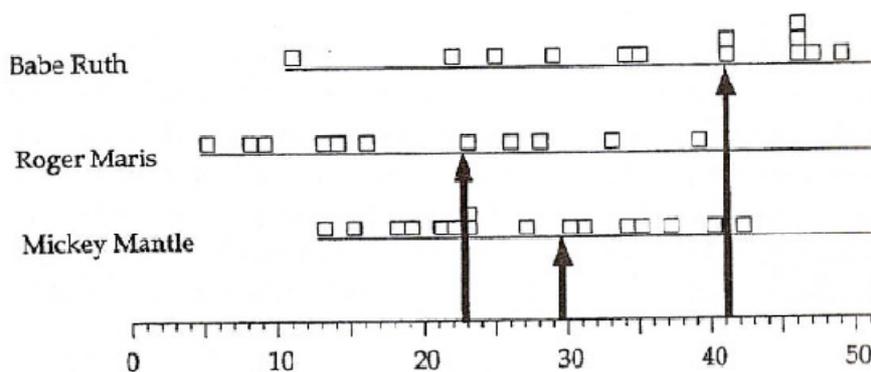


Figura 8 - Exemplo diagrama de Pontos de Médias (WHEELER, 2001)

Segundo Wheeler (2001), a média é uma medida da posição de um conjunto de valores, é comum usar as médias para comparar dois conjuntos de dados. Supõe-se que o conjunto com a média maior “exceda” o outro conjunto. Embora tais comparações possam ser úteis, elas devem ser usadas com cautela. Para qualquer que seja o conjunto de dados, a maioria dos valores não será igual à média.

- AMPLITUDE

Uma medida da dispersão de um conjunto de valores é a “amplitude”, definida como o valor máximo menos o valor mínimo (WHEELER, 2001).

A Figura 9 exemplifica a amplitude de corridas por circuito de beisebol por temporada de Babe Ruth, Roger Maris e Mickey Mantle.

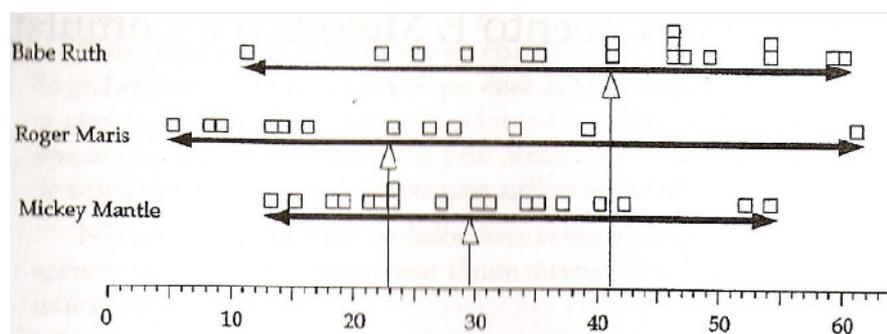


Figura 9 - Exemplo Diagrama de Pontos Amplitude

Com essas funções já é possível construir um gráfico de comportamento de processo, nas sessões seguintes é apresentado alguns princípios para entender os dados.

2.9 PRINCIPIOS PARA ENTENDER OS DADOS

2.9.1 PRIMEIRO PRINCIPIO

Segundo Wheeler(2001) há tempos, o Dr. Walter Shewhater ditou duas regras para a apresentação de dados. As quais são a base para uma estatística honesta e deveriam ser usadas como um guia em cada exposição de dados.

- Primeira Regra de Shewhart para Apresentação de Dados: Os dados devem sempre ser apresentados de maneira a preservar a evidência nos dados quanto a todos os prognósticos que possam ser feitos a partir desses dados
- Segunda Regra de Shewhart para Apresentação de Dados: Quando quer que uma média, amplitude ou um histograma

sejam usados para resumir dados, o resumo não deve iludir o usuário e levá-lo a tomar alguma atitude, o que não faria se os dados fossem apresentados em um gráfico seqüencial.

Nenhum dado tem sentido fora de seu contexto, esse primeiro princípio traz três conseqüências:

- Não confie em ninguém que não pode, ou não poderá oferecer o contexto para os seus cálculos.
- Para registrar comparações entre os pares de valores exceto como parte de uma comparação mais ampla.
- Comece usando gráficos para apresentar os valores atuais no contexto.

2.9.2 SEGUNDO PRINCIPIO

Segundo (WHEELER, 2001), embora todo o conjunto de dados contenha ruído, alguns conjuntos de dados podem conter sinais. Portanto, antes que se consiga detectar um sinal dentro de determinado conjunto de dados, deve-se primeiro filtrar o ruído.

O gráfico do comportamento do processo usa o gráfico seqüencial para definir a voz do processo. Também oferece ao usuário um meio de saber se ele sente-se seguro para extrapolar no futuro imediato. O gráfico do comportamento do processo define a variação dos valores que provavelmente veremos no futuro próximo.

O gráfico do comportamento do processo faz todas essas coisas porque leva em consideração a variação. A variação é o componente aleatório e heterogêneo que enfraquece as comparações simples e limitadas. Ele filtra o ruído da variação rotineira através da construção dos limites. Os sinais da variação excepcional são indicados pelos pontos que ficam fora dos limites ou por padrões óbvios não-aleatórios da variação em torno da linha central. Essa distinção entre sinais e ruído é a base para qualquer análise de dados que tenha significado.

Segundo Wheeler (2001), dois erros que pode ocorrer quando analisar os dados. O primeiro erro é interpretar a variação rotineira como um afastamento ocasional significativo do passado do passado. Interpretar o ruído como se ele fosse um sinal e o segundo consiste em não reconhecer quando ocorreu uma mudança em determinado processo.

Não importam quais sejam os dados, nem como os valores estão organizados e apresentados, devemos sempre usar algum método de análise para alcançar uma interpretação dos dados (WHEELER, 2001).

Wheeler (2001) afirma que:

“Os dados devem ser filtrados de alguma maneira para torná-los inteligíveis. Essa filtragem pode basear na experiência de uma pessoa, além de suas pressuposições e suposições; ou pode ser mais formalizada e menos subjetiva, mas haverá sempre algum método de análise.”

2.9.3 ABORDAGEM DE SHEWHART

Segundo Wheeler (2001), quando se analisa números, a fim de saber quando ocorreu uma mudança em processos ou sistemas. Deseja-se tomar conhecimento dessas alterações no tempo oportuno, de modo que possa reagir adequadamente. Embora isso soe bastante direto, há uma complicação – os números podem mudar até mesmo quando o processo não muda. Assim, na análise dos números, precisa-se ter alguma forma de distinguir entre mudanças nos números que representam modificações no processo e aquelas que são, essencialmente, ruído.

Segundo Wheeler (2001), Walter Shewhart fez uma distinção crucial entre dois tipos de variação nos números. Alguma variação é rotineira, faz parte do processo, sendo que esperada inclusive quando o processo não foi alterado. Outra variação é excepcional, ETA fora dos limites da rotina e, portando, é interpretada como sinal de uma mudança do processo. E a fim de separar a variação nesses dois componentes, ele criou o gráfico de controle (gráfico do comportamento do processo).

O gráfico do comportamento do processo começa com os dados representados graficamente em uma seqüência. É acrescentada uma linha central como referência visual para detectar as mudanças ou tendências, e os

limites são computados a partir dos dados. Esses limites são colocados, simetricamente, em ambos os lados da linha central a uma distância que lhes permitirá filtrar toda a variação rotineira virtualmente. (WHEELER, 2001)

A Figura 10 mostra a porcentagem diária dos pares com defeitos de um fabrica de calçados.

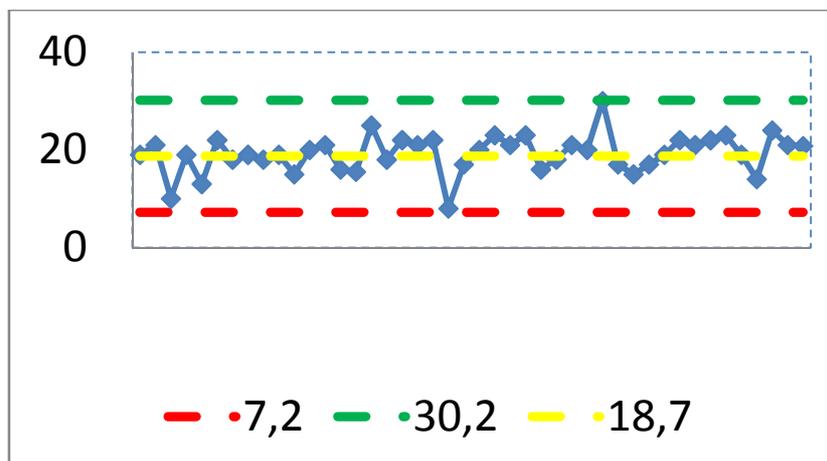


Figura 10 - Exemplo de um gráfico sequencial e sua média (WHEELER, 2001)

Wheeler (2001) afirma que:

“A chave da eficácia do gráfico do comportamento do processo está contida no modo como os limites são calculados a partir dos dados.”

Os limites em um gráfico do comportamento do processo que você caracterize o comportamento de seu processo como previsível ou imprevisível e defina a quantidade de variação rotineira que você deve esperar no futuro. Já que o prognóstico é a essência da administração, a capacidade de saber o que esperar quando um processo está se comportando de maneira previsível é algo inestimável (WHEELER, 2001).

Quando os pontos ficam fora dos limites de um gráfico do comportamento do processo, eles são interpretados como sinais de variação excepcional. A variação excepcional é atribuída às causas determinantes que, por definição, dominam as inúmeras causas comuns da variação rotineira. Portanto, quando um processo apresenta uma variação excepcional, de modo que você possa eliminar o seu efeito sobre o seu processo. (WHEELER, 2001)

Observe como o gráfico do comportamento do processo tem ajudado a interpretar os dados. Primeiro, o gráfico é usado para caracterizar o

comportamento dos dados. Segundo, o gráfico do comportamento do processo permite ao gerente prever o que esperar no futuro (WHEELER, 2001).

Para medir os processos são utilizados os KPI's, indicador chave de Desempenho (KPI), mede o nível de desempenho do processo, focando no como e indicando quão bem os processos de tecnologia da informação permitem que o objetivo seja alcançado.

KPI são veículos que permitem que os executivos do alto escalão comuniquem a missão e visão da empresa ao mais baixo níveis hierárquicos, envolvendo diretamente todos os colaboradores na realização dos objetivos estratégicos da empresa o KGI. (Indicador chave de Objetivo)(COMPANYWEB, 2010).

2.9.4 KGI

Indicador chave Objetivo (KGI), permitem medir em que grau os novos processos implementados responderam aos requisitos de negócio (COMPANYWEB, 2010).

Segundo a ITIL V3 (2007), tem como função:

- Alinhar os níveis de serviço pela unidade estratégica dos negócios com os objetivos do negócio;
- Definir se os níveis de serviço são compatíveis com os custos do orçamento.

3 METODOLOGIA

O presente trabalho se caracteriza como uma pesquisa exploratória, na qual foi realizada uma revisão bibliográfica em conteúdos relacionados, com o objetivo de adquirir conhecimento sobre todo ambiente de BI e a métodos de análise estatística de processo. A metodologia de desenvolvimento desse trabalho segue as seguintes etapas:

- A Revisão Bibliográfica;
- Definição de um padrão na extensão xml para extrassão de dados;
- Utilização da tecnologia Adobe Flex para criação de aplicações RIA;
- Modelagem do componente, para facilitar o desenvolvimento e compreensão de sua estrutura e funcionamento;
- Analise dos resultados por meio de um estudo de caso.

4 DESENVOLVIMENTO

Foi desenvolvida uma ferramenta gráfica que realiza análise estatística de processo aplicada a um ambiente de BI a ferramenta BRAIN. Foi utilizado uma base de dados open source chamada FoodMart para realizar o estudo de caso, o diagrama de atividades na Figura 10 ilustrada suas fases.

- DIAGRAMA DE ATIVIDADES

O diagrama de atividades, a Figura11, ilustra as fases que compõe o processo para realizar uma análise estatística de processo.

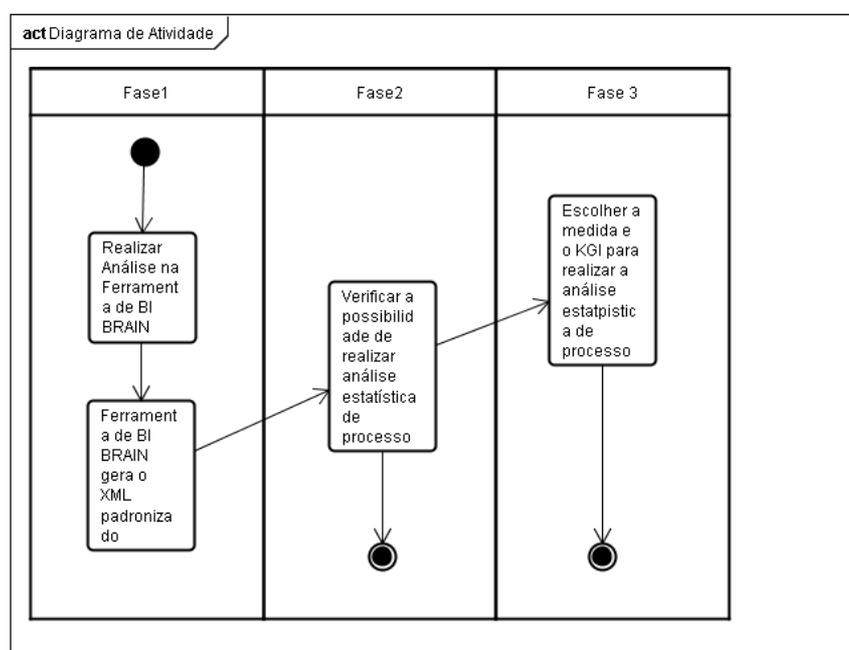


Figura 11- Diagrama de Atividades que compõe o processo da Análise Estatística do Processo

Na Fase 1, o usuário realizara a análise na ferramenta de BI BRAIN, escolhendo o cubo e as medidas que deseja calcular. Com os dados da análise realizada, a ferramenta de BI BRAIN gera um XML com um padrão específico que a ferramenta de análise estatística de processo possa entender.

Nesta fase o usuário faz o login na ferramenta de BI BRAIN, a Figura 12 ilustra a realização do login.

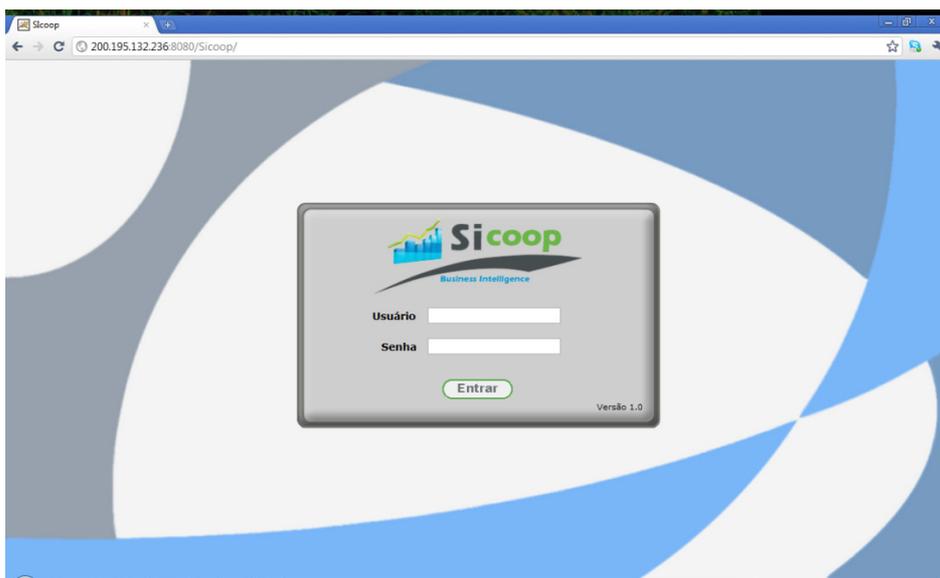


Figura 12 - Interface de Login da Ferramenta de BI BRAIN

Após ter efetuado o login, o usuário se encontra na página principal da ferramenta de BI BRAIN, nesse momento o usuário selecionara a aba Análise que esta no menu superior da ferramenta ou no menu central, a Figura 13 ilustra os menus para navegar até a interface de análise.



Figura 13 - Tela Principal da ferramenta de BI BRAIN

Na interface que realizara a análise possui um menu lateral esquerdo que contém os cubos que estão disponíveis para a realização da análise, nele também são escolhida as medidas que deseja calcular. A Figura 14 ilustra o local.

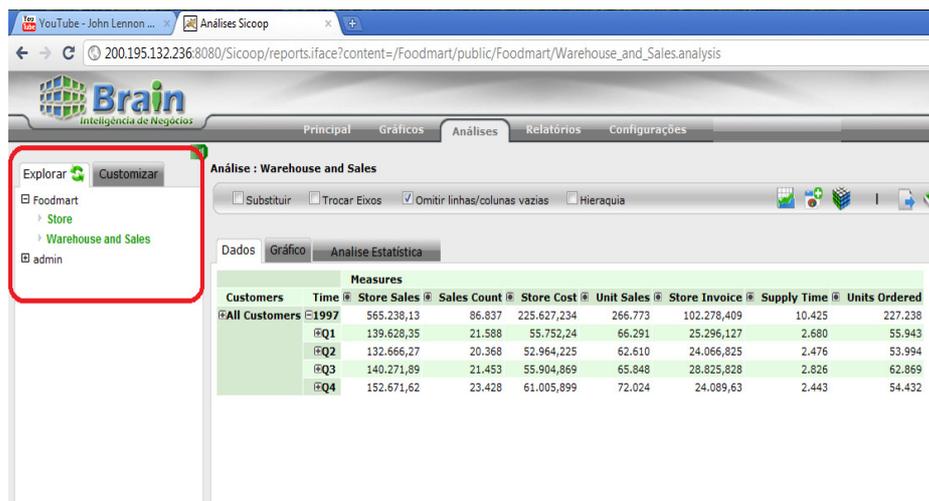


Figura 14 - Local que realiza a escolha do cubo e das medidas

Após ter realizado a análise na ferramenta de BI BRAIN, o usuário tem a opção de gerar os gráficos ou realizar a análise estatística do processo, clicando na opção para realizar a análise estatística do processo, os dados da análise realizada na ferramenta de BI BRAIN são transformados em um padrão específico na extensão XML o qual a ferramenta de análise estatística de processo consiga entender, a Figura 15 demonstra o padrão gerado.

```
<?xml version="1.0" encoding="UTF-8"?>
<analise>
  <fila nome="Customers" tipo="All Customers" id="0">
    <fila nome="Time" tipo="Jan" id="1">
      <coluna id="1" nome="Measures" tipo="Store Sales" valor="45539" />
      <coluna id="2" nome="Measures" tipo="Sales Count" valor="7034" />
      <coluna id="3" nome="Measures" tipo="Store Cost" valor="18178" />
      <coluna id="4" nome="Measures" tipo="Unit Sales" valor="21628" />
      <coluna id="5" nome="Measures" tipo="Store Invoice" valor="11305" />
      <coluna id="6" nome="Measures" tipo="Supply Time" valor="1193" />
      <coluna id="7" nome="Measures" tipo="Units Ordered" valor="24398" />
    </fila>
  </fila>

  <fila nome="Customers" tipo="All Customers" id="0">
    <fila nome="Time" tipo="Fev" id="1">
      <coluna id="1" nome="Measures" tipo="Store Sales" valor="44058" />
      <coluna id="2" nome="Measures" tipo="Sales Count" valor="6844" />
      <coluna id="3" nome="Measures" tipo="Store Cost" valor="17599" />
      <coluna id="4" nome="Measures" tipo="Unit Sales" valor="20957" />
      <coluna id="5" nome="Measures" tipo="Store Invoice" valor="6700" />
      <coluna id="6" nome="Measures" tipo="Supply Time" valor="712" />
      <coluna id="7" nome="Measures" tipo="Units Ordered" valor="15765" />
    </fila>
  </fila>
</analise>
```

Figura 15 - Padrão do XML gerado para a ferramenta de análise estatística de processo

Com este conteúdo é gerada uma página JSP, a qual a ferramenta de análise estatística de processo acessa utilizando o serviço HTTPService.

Na Fase 2, o programa de análise estatística de processo verifica a possibilidade da realização, logo que, para realizar a análise estatística de processo precisa-se ter uma série temporal, neste caso, a ferramenta procura no XML pela dimensão Tempo. A Figura 16 ilustra o padrão de XML gerado pela ferramenta e a dimensão que é necessária para realização da análise estatística de processo.

```
<?xml version="1.0" encoding="UTF-8"?>
<analise>
  <fila nome="Customers" tipo="All Customers" id="0">
    <fila nome="Time" tipo="Jan" id="1">
      <coluna id="1" nome="Measures" tipo="Store Sales" valor="45539" />
      <coluna id="2" nome="Measures" tipo="Sales Count" valor="7034" />
      <coluna id="3" nome="Measures" tipo="Store Cost" valor="18178" />
      <coluna id="4" nome="Measures" tipo="Unit Sales" valor="21628" />
      <coluna id="5" nome="Measures" tipo="Store Invoice" valor="11305" />
      <coluna id="6" nome="Measures" tipo="Supply Time" valor="1193" />
      <coluna id="7" nome="Measures" tipo="Units Ordered" valor="24398" />
    </fila>
  </fila>
</analise>
```

Figura 16 - Dimensão Data que é necessária para realização da análise estatística de processo

Se a análise efetuada pela ferramenta de BI BRAIN existir a dimensão Tempo, a ferramenta passara para a Fase 3, se não a ferramenta disparara um alerta.

Na Fase 3, o usuário tem a opção de selecionar uma das medidas que ele escolheu na análise realizada na ferramenta de BI BRAIN. A análise estatística de processo é realizada em apenas uma medida, também tem a opção de configurar um KGI para a medida selecionada. O KGI traçará uma linha no gráfico delimitando os objetivos da empresa para a medida selecionada. A Figura 17 ilustra a interface da ferramenta de análise estatística de processo, o campo em que o usuário pode escolher a medida e atribuir o valor do KGI.



Figura 17 - Campos para escolha da medida e KGI

A partir das atividades realizadas, foi gerado um diagrama de casos de usos ilustrado pela Figura 18, composto por:

- Realizar Análise;
- Escolher Medida;
- Realizar Análise Estatística de Processo.

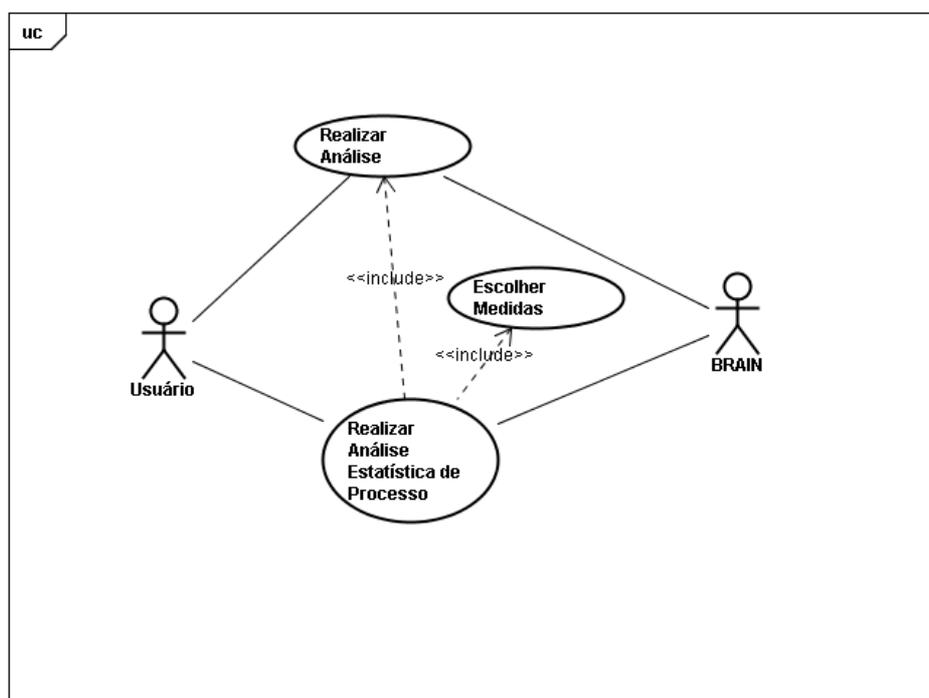


Figura 18 - Diagrama de Casos de Uso

Os casos de usos ilustrados pelo diagrama na Figura 18 são descrito nas Tabelas de 3 a 5.

Na Tabela 3 é feita a descrição do caso de uso Realizar Análise, ela descreve passo a passo as ações realizadas pelo usuário, passos obrigatórios para a realização dos outros casos de usos.

Tabela 3 - Descrição do Caso de Uso Realizar Analise

Nome: Realizar Análise
Ator: Usuário.
Descrição: Usuário deseja realizar a análise na ferramenta de BI BRAIN.
Fluxo Principal:
<ol style="list-style-type: none"> 1. O usuário realiza o login na ferramenta BRAIN. 2. O usuário seleciona no menu Analise na ferramenta BRAIN. 3. Usuário escolhe o cubo e as medidas para a análise.

No caso de uso Realizar Análise, as ações realizadas são todas na ferramenta de BI BRAIN, os passos são necessários para a realização da análise estatística de processo. O próximo caso de uso Realizar Análise Estatística de Processo.

Tabela 4 - Casos de Uso Realizar Análise Estatística

Nome: Realizar Analise Estatística de Processo
Ator: Usuário.
Descrição: Usuário deseja realizar a análise estatística de processo sobre as medidas escolhida na análise realizada na ferramenta de BI BRAIN.
Fluxo Principal:
<ol style="list-style-type: none"> 1. Execução do Caso de Uso Realiza Analise; 2. Usuário seleciona o menu Analise Estatística de Processo; 3. O sistema verifica se é possível realizar a Análise Estatística de Processo; 4. O caso de uso Escolher Medida é executado;

Neste caso de uso, o usuário após ter realizado o caso de uso Realizar Análise, ele seleciona a opção para realizar a análise estatística de

processo, o sistema verifica a possibilidade da realização, se possível, é iniciado o caso de uso Escolher Medida, descrito pela Tabela 5.

Tabela 5 - Descrição do Caso de Uso Escolher Medida

Nome: Escolher Medida
Ator: Usuário.
Descrição: Usuário seleciona a medida e atribui um valor para o KGI que ele deseja, para realizar a análise estatística de processo.
Fluxo Principal:
<ol style="list-style-type: none"> 1. Usuário seleciona na lista a medida desejada. 2. Usuário indica o valor do KGI para a medida desejada. 3. O sistema realiza a Análise Estatística de Processo

No caso de uso Escolher Medida, o usuário escolhe uma medida para a realização da análise estatística de processo. As opções de medidas fazem parte da análise criada pela ferramenta de BI BRAIN, são as mesmas. Também existe a possibilidade de atribuir um valor para KGI, após ter realizado esses passos é realizada a análise estatística de processo.

4.1 DIAGRAMA DE SEQUENCIA

O diagrama de sequência de o processo Realizar Análise Estatística de Processo da ferramenta – é apresentado a seguir pela Figura 19.

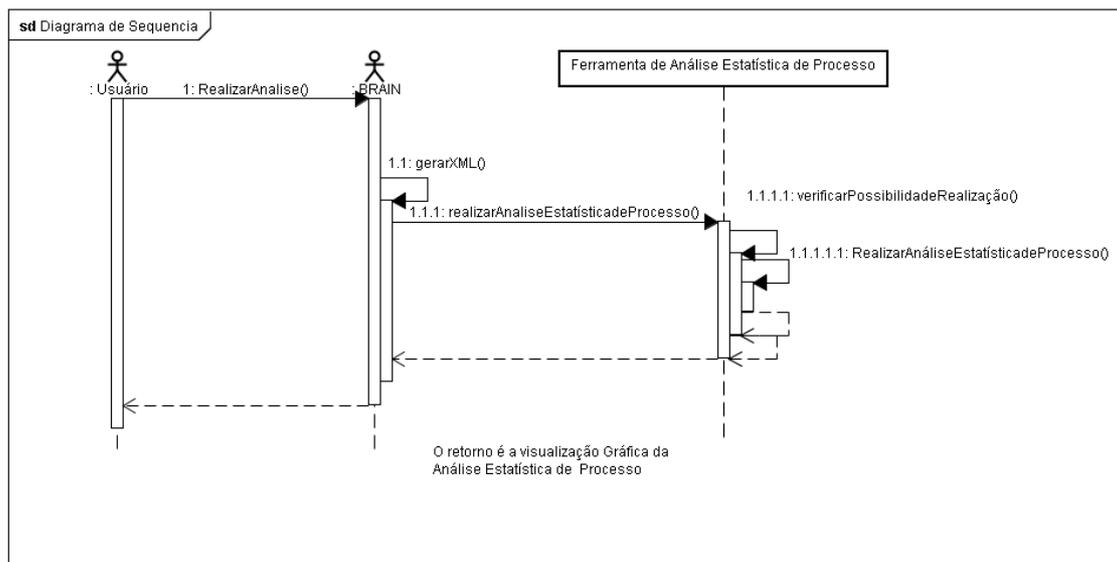


Figura 19 - Diagrama de Sequencia de o processo Realizar Análise Estatística de Processo

O diagrama demonstra todo o processo para realizar a análise estatísticas de processo, desde a análise realizada na ferramenta de BI BRAIN até a realização da análise estatística de processo na ferramenta desenvolvida.

4.2 ESTRUTURA DA FERRAMENTA

A ferramenta desenvolvida é formada por um módulo, que realiza a análise estatística de processo, este módulo é integrado em um ambiente de BI a ferramenta BRAIN, através do serviço HTTPservice é realizada a comunicação entre eles. A Figura 20 ilustra a integração das duas ferramentas.

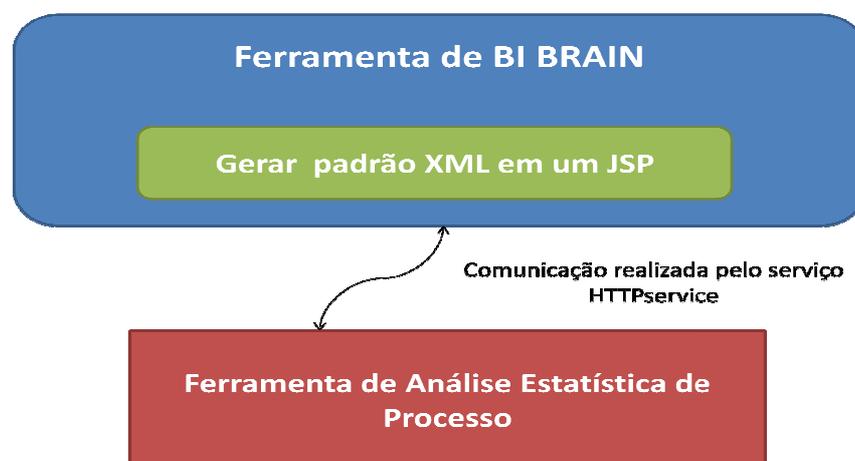


Figura 20 - Estrutura de Comunicação entre a Ferramenta de BI BRAIN e a Ferramenta de Análise Estatística de Processo

Explicação dos 2 módulos :

- Ferramenta de BI BRAIN: através dela é realizada uma análise dos dados históricos. Ela é responsável em criar com a base de dados da análise realizada, uma página jsp sendo seu conteúdo um padrão específico no formato XML.
- Ferramenta de Análise Estatística de Processo: é responsável em acessar essa base de dados gerada e realizar a análise estatística de processo, a Figura 21 mostra o trecho de código que acessa esses dados.

```
<!-- Acessando o gerador de grafico -->  
<mx:HTTPService id="srv" url="http://200.195.132.236/Sicoop/flex_xml.jsp"  
    resultFormat="e4x" result="cbcarregaLista()"/>  
  
<!-- -->
```

Figura 21 - Código fonte do HTTPService acessando a ferramenta de BI BRAIN

A propriedade url é responsável em atribuir o caminho da página JSP gerada pela ferramenta de BI BRAIN. Após mostrar todos os processos para realizar a análise estatística de processo e toda a estrutura da ferramenta, foi realizado um estudo de caso que será apresentado a seguir.

5 ESTUDO DE CASO

A ferramenta foi aplicada no ambiente de BI, a ferramenta BRAIN, utilizando uma base de dados open source chamada FoodMart, foi realizada uma análise ilustrada pela Figura 22.

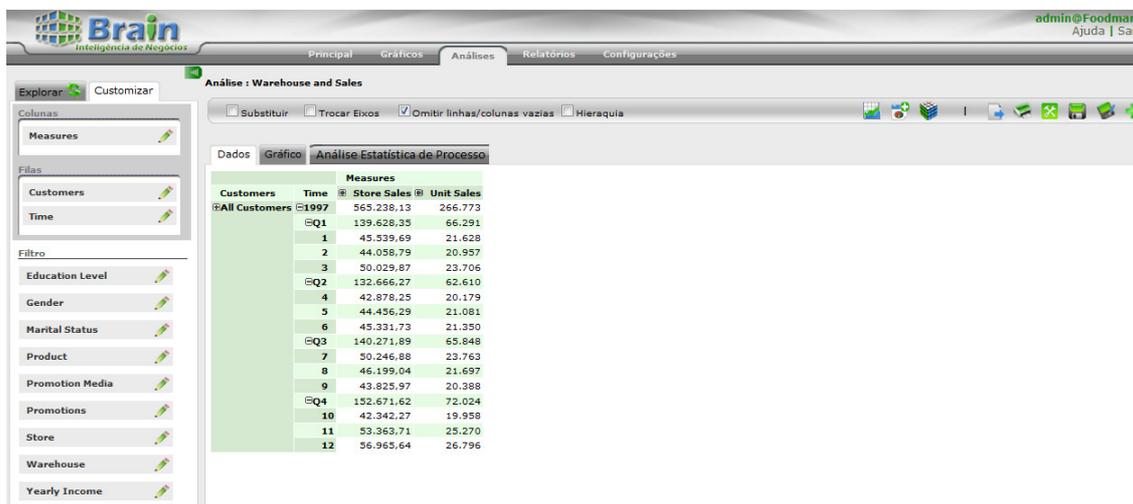


Figura 22 - Análise Realizada na Ferramenta de BI BRAIN

A análise coletou dados do ano de 2007, e as medidas escolhidas foram, Store Sales e Unit Sales, realizando a análise estatísticas de processo das medidas selecionadas. As Figuras 23 e 24 ilustram a análise em todo o ano de 2007.

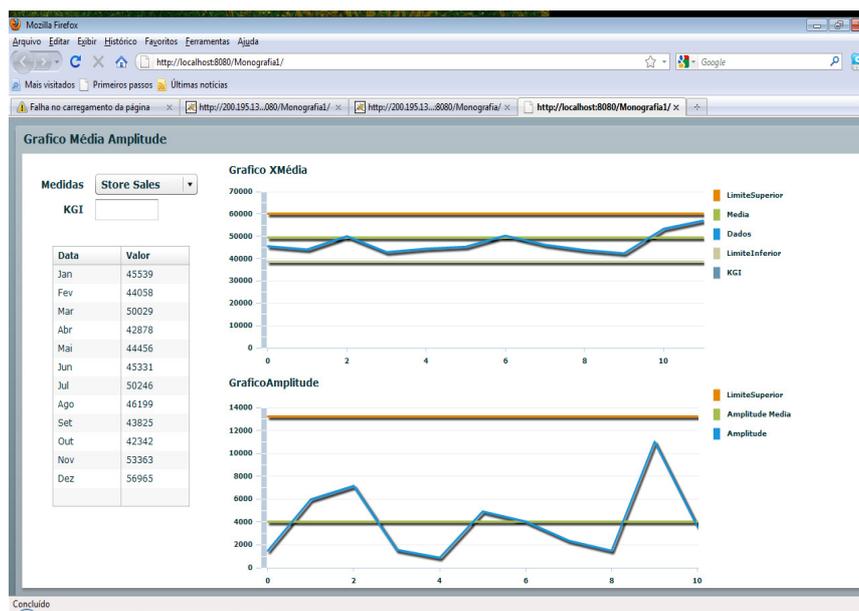


Figura 23 - Análise Estatística de Processo aplicada na medida Store Sales

Na Figura 23, foi ilustrada a análise estatística de processo aplicada na medida Store Sales, a análise realizada mostra que não houve nenhuma variação especial, variação que ultrapassa os limites estabelecidos pela análise estatística de processo, mas note que, em quase todo o ano os dados ficaram abaixo da média, ilustrada pela linha verde no gráfico XMédia.

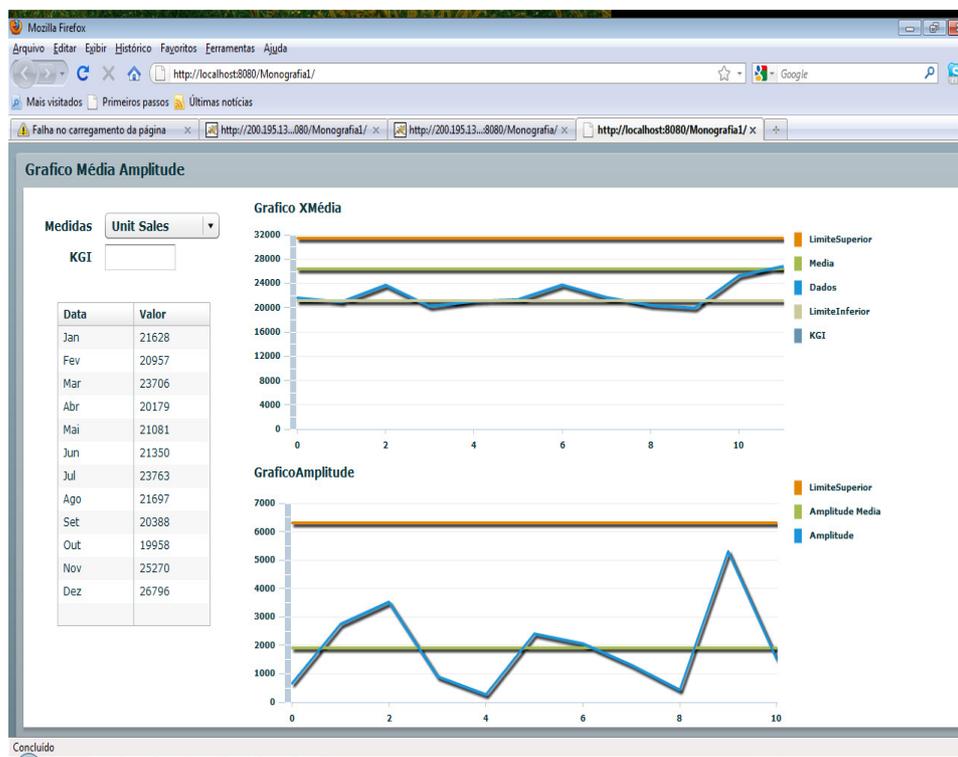


Figura 24 - Análise Estatística de Processo aplicada na medida Unit Sales

Na Figura 24, é apresentada a análise estatística de processo realizada na medida Unit Sales, a análise detectou variações especiais, no gráfico XMédia foram detectados cinco variações especiais. Nos meses Fevereiro, Abril, Maio, Setembro e Outubro os dados sofreram variação que afetaram diretamente no processo.

Essas variações foram às unidades vendidas abaixo do esperado, isso causou diretamente um impacto no resultado final do processo. Um gestor sem o controle estatístico poderia tomar decisões equivocadas sem conseguir interpretar corretamente os dados. Os dados poderiam ficar invisíveis sem o auxílio da ferramenta de análise estatística de processo.

6 RESULTADOS

Os resultados referentes ao estudo de caso podem ser classificados positivos.

Nos dois cenários avaliados os resultados foram justificados pelo fato da análise ter sido realizada com sucesso, detectando os limites superiores, limites inferiores, média, amplitude e detectar os dados que realmente modificaram o processo.

No primeiro cenário, a análise estatística de processo aplicada na medida Store Sales, não foi detectada nenhuma variação especial, mas detectou que quase todo o ano de 2007 os dados ficaram abaixo da média.

No segundo cenário, a análise estatística de processo aplicada na medida Unit Sales, foi detectada cinco variações especiais, nos meses Fevereiro, Abril, Maio, Setembro e Outubro. Essas variações influenciam diretamente no processo.

As medidas selecionadas para o estudo estão diretamente ligadas, as vendas da loja “Store Sales” e as unidades vendidas “Unit Sales”. No primeiro cenário não foram detectadas variações especiais, mas foi detectado que quase todo o ano os dados ficaram abaixo da média. No segundo cenário foram detectadas cinco variações, os meses que foram detectados as variações, no primeiro cenário também ilustrava que os mesmos meses foram os que tiveram os menores valores.

7 CONCLUSÃO

Este trabalho realizou um estudo sobre o ambiente de BI e aos métodos de análise estatística de processo. Um estudo de caso foi realizado aplicando a ferramenta no ambiente de BI utilizando a ferramenta BRAIN e a base de dados FoodMart. As principais contribuições desse trabalho foram:

- O processo de realizar a análise estatística auxiliando a ferramenta de BI BRAIN a detectar os dados que sofreram alteração e que influenciam diretamente no desempenho do processo.

A ferramenta conseguiu sanar a ausência de uma ferramenta que auxilia uma ferramenta de BI a detectar as variações e a interpretar os dados o qual foi realizada a análise, assim auxiliando na tomada de decisão.

Também foi possível a integração entre a ferramenta de BI BRAIN e a ferramenta de análise estatística de processo, utilizando a HTTPService para realizar a comunicação.

Um gestor sem o auxílio dessa ferramenta de análise estatística de processo poderá tomar decisões equivocadas, podendo influenciar diretamente no final do processo. Logo que, a ferramenta de BI transforma os dados em uma melhor visualização dos dados, mas não auxilia na interpretação dos mesmos. Essas decisões comprometem o futuro da empresa, logo que são traçadas sobre os planos de gerência.

7.1 TRABALHOS FUTUROS

Um trabalho futuro poder ser realizado abordando outras possibilidades de análise estatística, na construção de outros tipos de gráficos que realizam análise estatística de processo.

Uma lista destes gráficos é listada abaixo:

Existem dois tipos básicos de gráficos de Controle:

Gráficos por variáveis:

- Gráficos \bar{X} e S (média e desvio padrão)
- Gráficos \tilde{X} e R (mediana e desvio padrão)
- Gráficos para Valores Individuais (X) e Amplitude Móvel (MR)

Gráficos por atributos:

- Gráfico p (proporções não conforme)
- Gráfico np (unidades não conforme)
- Gráfico c (número de não conformidade por unidade)
- Gráfico u (taxa de não conformidade por unidade)

REFERÊNCIAS

Companyweb. Disponível em :

<www.companyweb.com.br/lista_de_artigos/>. Acesso em 22 maio 2010.

FORTULAN, Marcos Roberto; GONÇALVES FILHO, Eduardo Vila. **Uma proposta de aplicação de business intelligence no chão-de-fábrica**. São Carlos: Gestão & Produção, V.12, N.1, P.55-66, Jan.-abr. 2005, 2005.

HARRISON, T. **Intranet Data Warehouse**. São Paulo: Berkeley, 1998.

INMON, W. H. **Como construir o Data Warehouse**. 2ª ed.- Rio de Janeiro: Campus, 1997.

KIMBAL, Ralph. **The Data Warehouse ToolKit: Guia Completo para modelagem dimensional**. 2. ed. Rio de Janeiro: Campus Ltda., 2002. 494 p.

Pentaho. **Mondrian**. Disponível em: <<http://mondrian.pentaho.org/documentation/>>. Acesso em: 22 maio 2010.

PortalAction. Disponível em: <<http://www.portalaction.com.br/>>. Acesso em 10 agosto 2010.

SERRA, Laércio. **A essência do Business Intelligence**. São Paulo: Berkeley Brasil, 2002.

TURBAN; EFRAIM. **Business Intelligence: um enfoque gerencial para a inteligência do negócio**. Porto Alegre: Bookman, 2009.

WHEELER, Donald J.. **Entendendo a Variação: A Chave para administrar o caos**. Rio de Janeiro: Qualitymark, 2001. 152 p.

Apêndice A – Código fonte da classe que realiza a Análise Estatística

```

<?xml version="1.0" encoding="utf-8"?>
<mx:Canvas xmlns:mx="http://www.adobe.com/2006/mxml" width="100%"
height="100%" creationComplete="srv.send()" >

    <!-- Acessando o gerador de grafico -->
    <mx:HTTPService id="srv"
url="http://200.195.132.236/Sicoop/flex_xml.jsp"
    resultFormat="e4x" result="cbcarregaLista()" />

    <!-- -->
    <mx:Panel x="0" y="0" width="100%" height="100%"
layout="absolute" title="Grafico Média Amplitude" fontSize="16"
fontWeight="bold">
    <mx:Form x="10" y="10" width="286" height="101">
    <mx:FormItem label="Medidas" fontSize="14"
fontWeight="bold">
        <mx:ComboBox id="cbLista"
close="escolhaComboBox(event)" width="50%" fontSize="14" />
    </mx:FormItem>
    <mx:FormItem label="KGI" fontWeight="bold" fontSize="14">
        <mx:TextInput id="valorKGI" width="94" />
    </mx:FormItem>
    </mx:Form>
    <mx:DataGrid x="146" y="119" width="102" height="337"
id="dgValor" fontSize="12" fontWeight="normal">
        <mx:columns>

            <mx:DataGridColumn headerText="Valor" dataField="@valor" />
        </mx:columns>
    </mx:DataGrid>
    <mx:DataGrid x="45" y="119" id="dgData" height="337"
fontSize="12" fontWeight="normal">
        <mx:columns>
            <mx:DataGridColumn headerText="Data"
dataField="@tipo" />
        </mx:columns>
    </mx:DataGrid>
    <mx:LineChart x="304" y="40" id="graficoMedia" height="237"
width="706" showDataTips="true">
        <mx:series>

            <mx:LineSeries displayName="LimiteSuperior" yField=""
dataProvider="{arLms}" />
            <mx:LineSeries displayName="Media"
yField="" dataProvider="{arMedia}" />
            <mx:LineSeries displayName="Dados"
yField="@valor"
dataProvider="{srv.lastResult..coluna.(@tipo ==
cbselecionado)}" />

            <mx:LineSeries displayName="LimiteInferior" yField=""
dataProvider="{arLmi}" />
            <mx:LineSeries displayName="KGI"
yField="" dataProvider="{arKGI}" />
        </mx:series>
    </mx:LineChart>

```

```

        <mx:LegenddataProvider="{graficoMedia}" x="1018" y="40"
width="120" height="237" fontSize="14"/>
        <mx:LineChart x="304" y="319" id="graficoAmplitude"
width="706" height="242" showDataTips="true">
            <mx:series>

                <mx:LineSeriesdisplayName="LimiteSuperior" yField=""
dataProvider="{arLas}"/>

                <mx:LineSeriesdisplayName="Amplitude Media" yField=""
dataProvider="{arAmplitude}"/>

                <mx:LineSeriesdisplayName="Amplitude" yField=""
dataProvider="{valoresAmplitude}"/>
            </mx:series>
        </mx:LineChart>
        <mx:LegenddataProvider="{graficoAmplitude}" x="1018"
y="298" width="120" height="242" fontSize="14"/>
        <mx:Label x="304" y="10" text="GraficoXMédia"
fontWeight="bold" fontSize="14"/>
        <mx:Label x="304" y="285" text="GraficoAmplitude"
fontWeight="bold" fontSize="14"/>

    </mx:Panel>
    <!--Bloco de Script -->
    <mx:Script>
    <![CDATA[

        // importações
importmx.collections.ArrayCollection;
        importmx.controls.Alert;

        //variaveis
        [Bindable]
publicvar meuarray:Array = new Array();
publicvar contador:int;
        [Bindable]
        privatevar cbseleccionado:String="";

        //variaveis para calcular grafico comportamento
publicvar media:Number=0.00;
publicvar lms:Number=0.00;
publicvar lmi:Number=0.00;

        //variaveis para calcular grafico amplitude
publicvar amplitude:Number;
publicvar las : Number=0.00;
publicvar amplitudeMedia:Number =0.00;

        publicvar valores:Array;
        [Bindable]
publicvar valoresAmplitude:Array;

        //variavel kgi
        [Bindable]
publicvar kgi:Number;

```

```

//array para utilizar de dataprovider
[Bindable]
publicvararMedia:Array;
[Bindable]
publicvararAmplitude:Array;
[Bindable]
publicvararKGI:Array;
[Bindable]
publicvararLms:Array;
[Bindable]
publicvararLmi:Array;
[Bindable]
publicvararLas:Array;

privatefunctioncalcular ():void{

valores = new Array();
valoresAmplitude = new Array();

foreach (var s:Number in
srv.lastResult..coluna.(@tipo == cbseleccionado).@valor){
    valores.push(s);
}

i++){
    for ( var i:int = 0 ; i<valores.length ;
        media+=valores[i];
    }
    media = media/valores.length;

z++){
    for (var z:int=1 ; z<valores.length;
        if(valores[z-1] >= valores[z]){
            amplitude = valores[z-1] -
        }else{
            amplitude = valores[z] -
        }
        valoresAmplitude.push(amplitude);
    }
    amplitudeMedia=0.00;
    for (var y:int=0 ;
y<valoresAmplitude.length ; y++){
        amplitudeMedia+=valoresAmplitude[y];
    }
    amplitudeMedia =
amplitudeMedia/valoresAmplitude.length;

//calculo dos limites
las = 3.27 * amplitudeMedia;
lms = media +(2.66 * amplitudeMedia);

```

```

        lmi = media -(2.66 * amplitudeMedia);

        gerarArray();

    }

    privatefunctiongerarArray():void{
        arMedia = new Array();
        arLms = new Array();
        arLmi = new Array();
        arKGI = new Array();
        arLas = new Array();
        arAmplitude = new Array();

        for(var k:int =0 ; k<valores.length ;
k++){

            arMedia.push(media);
            arLms.push(lms);
            arLmi.push(lmi);
            arKGI.push(kgi);
        }
        for(var w:int =0 ;
w<valoresAmplitude.length ; w++){

            arAmplitude.push(amplitudeMedia);
            arLas.push(las);

        }

    }

    //funções
    privatefunction cbcarregaLista():void {
        testaPossibilidade();
        contador=0;

        foreach(var i:int in srv.lastResult..coluna.@id){
            if(i!=0 && i >contador){

                // variavel que recebe
                var nome:String =
                srv.lastResult..coluna[i-1].@tipo;

                meuarray.push(nome);
                contador++;

            }

        }

        cbLista.dataProvider = meuarray;
    }

    privatefunctionescolhaComboBox(event:Event):void
    {
        cbselecionado =
        ComboBox(event.target).selectedItem.toString();
        dgValor.dataProvider = srv.lastResult..coluna.(@tipo
== cbselecionado);
        dgData.dataProvider = srv.lastResult..fila.(@nome ==
"Time");
    }

```

```
        kgi = parseFloat(valorKGI.text);
        calcular();

    }

    private function testaPossibilidade():void{
        if( srv.lastResult..fila.@nome!="Time"){
            Alert.show("A Analise é Possivel!");
        }else{
            Alert.show("Analise Impossível");
        }
    }

    ]]>
</mx:Script>
```

```
<!--          -->
</mx:Canvas>
```