

***Business Intelligence* para uma análise da qualidade da entrega dos objetos postais: um estudo de caso nos Correios de Alagoas**

Jean Barros Teixeira¹  orcid.org/0000-0001-9485-6149

Mailson Melo dos Santos Filho²  orcid.org/0000-0002-1711-5301

Carlos André Duarte Costa³  orcid.org/0000-0002-7729-1120

¹ Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, PE, Brasil.

² Fábrica de Negócios – Analytics & Data Mining, Recife, PE, Brasil.

³ Empresa Brasileira de Correios e Telegráfos (GERAE), Alagoas, Brasil.

E-mail do autor principal: jbt@ecom.poli.br

Resumo

Atualmente as organizações necessitam de sistemas que forneçam suporte a tomada de decisão através de análises do grande volume de dados oriundos de seus sistemas transacionais. É o caso dos Correios que carece de uma solução que possa trazer vantagem competitiva no mercado em que atua. Dessa forma o objetivo desse trabalho é propor e implementar uma solução de Business Intelligence que irá auxiliar os gestores na análise da qualidade dos objetos postais e em uma tomada de decisão mais assertiva. Para isso foi utilizada a ferramenta *Open Source Pentaho* e técnicas que permitiram a criação de um *Data Mart*, realização de consultas analíticas e o desenvolvimento de um *dashboard*.

Palavras-Chave: *Pentaho; Business Intelligence; Data Warehouse; Data Mart; Tomada de decisão;*

Abstract

Currently, organizations need systems that support decision making through analysis of the large volume of data coming from their transactional systems. This is the case of the Post Office which offers a solution to create a competitive competitor in the market in which it operates. In this way the objective of this work is to propose and implement a Business Intelligence solution that is compatible with the managers in the analysis of the quality of the objects in a more assertive decision making. Pentaho and techniques that allows the creation of a Data Mart, conducting analytical consultations and development of a control panel.

Key-words: *Pentaho; Business Intelligence; Data Warehouse; Data Mart; Decision make;*

Introdução

No contexto atual as organizações, públicas ou privadas, são pressionadas pelo mercado diariamente, devido a essa competitividade no mundo dos negócios o ecossistema organizacional sofre constante mudança. Com a finalidade de melhorar continuamente os produtos e serviços prestados aos seus clientes, o aprimoramento dos processos internos, e o aperfeiçoamento do conhecimento de negócio as organizações devem responder de uma forma célere às condições de mudança impostas pelo mercado, isto se torna um fator crítico de sucesso. Porém, as tomadas de decisões necessitam de grandes quantidades de dados, informação e conhecimento.

Os dados após serem transformados em informação torna-se um recurso primordial para auxiliar os gestores em uma melhor tomada de decisão. "A informação é considerada como um ingrediente básico do qual dependem os processos de decisão" [1]. Para isso a informação precisa ser disponibilizada com uma boa qualidade, confiável, na quantidade e no momento certo.

Assim como as demais organizações, os Correios também sofrem constantemente com a pressão exercida pelos concorrentes, mais precisamente no mercado de logística e encomendas já que o mesmo detém apenas o monopólio postal (entrega de cartas). Para melhorar sua competitividade, a empresa necessita acompanhar as tendências tecnológicas do mercado, utilizando sistemas capazes de coletar, organizar e publicar as informações necessárias para que os gestores possam tomar decisões condizentes com a realidade e mais assertivas.

Diante dessa necessidade de um sistema que possa auxiliar as organizações a lidar com os seus dados, diversos softwares e técnicas são desenvolvidos no mercado, como por exemplo, uma técnica chamada *Business Intelligence*. Uma solução de *Business Intelligence* (BI) proporciona aos gestores uma visão real da organização através de indicadores alinhados aos objetivos estratégicos do negócio. O BI permite coletar diversos dados isolados de várias fontes de dados da empresa, organizá-los e consolidá-los de forma

que se possa extrair diversas informações úteis para o negócio.

Sendo assim o objetivo deste trabalho é a implementação de uma solução de Business Intelligence para auxiliar os gestores na realização de análises sobre a qualidade da entrega dos objetos postais através de um Dashboard de gestão operacional.

Revisão Bibliográfica

2.1 Business Intelligence (BI)

O BI ou em sua tradução, inteligência de negócio, é um processo que envolve diversos conceitos, metodologias, arquiteturas, tecnologias e infraestrutura, e tem como objetivo auxiliar as organizações em tomadas de decisões mais inteligentes e que agreguem valor ao negócio. Isso é possível pois através do BI pode-se coletar dados de diversas fontes, organizá-los, analisá-los e compartilhá-los com todos os tomadores de decisões.

De forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para definir estratégias de competitividade nos negócios da empresa. Podem ser incluídos nessa definição os conceitos de estruturas de dados, representadas pelos bancos de dados tradicionais, *data warehouse*, e *data marts*, criados objetivando o tratamento relacional e dimensional de informações, bem como as técnicas de data mining aplicadas sobre elas, buscando correlações e fatos "escondidos" [2].

2.2 Data Warehouse e Data Mart

Data warehouse (DW) ou Armazém de dados é a nomenclatura utilizada para definir um repositório de dados históricos, relacional ou multidimensional, que serve aos interesses de todos os departamentos da organização [3]. Assim, o DW é projetado para armazenar de forma consolidada o grande volume de dados de uma organização e através da análise dos mesmos, extrair informações que respondam às necessidades do negócio.

Por sua vez, um *Data Mart* pode receber o mesmo conceito de um DW, a maior diferença entre eles é que o *Data Mart* é constituído por dados de uma mesma área da organização, por exemplo um *Data Mart* de Vendas. Um conjunto de *Data Marts*, ou seja, diversos conjuntos de dados específicos de cada área da organização formam o DW.

2.3 Online Analytical Processing (OLAP)

OLAP é considerado como um processo ou arquitetura que possibilita ao usuário uma análise profunda dos dados em diversos ângulos, geralmente através de interfaces gráficas que facilitam o manuseio do usuário. Têm o objetivo de trabalhar dados existentes, buscando consolidações em vários níveis, trabalhando fatos em dimensões variadas [1]. As funções básicas do OLAP são:

- visualização multidimensional dos dados;
- exploração;
- rotação;
- vários modos de visualização.

O OLAP e o DW trabalham de forma integrada, já que os dados armazenados de forma eficiente no DW são recuperados pela ferramenta OLAP com a mesma eficiência e rapidez. Desde a idealização do DW deve-se levar em consideração o que se deseja apresentar na ferramenta OLAP.

2.4 Pentaho

O *Pentaho* é uma suíte composta por diversos softwares voltados para a criação de soluções de BI de ponta a ponta. Existem soluções para diversas áreas desde a integração dos dados (ETL - *Extract, Transform e Load*), relatórios pré-formatados e ad hoc, análises *online* (OLAP - *Online Analytical Processing*), ferramentas para criação de dashboard, mineração de dados etc. Vale ressaltar que é uma ferramenta *open source*, dessa forma o cliente pode customizá-la de acordo com as necessidades do seu negócio.

A plataforma se divide em duas partes [4].

O *Solution Engine* e seus componentes são responsáveis pela execução e controle das

soluções. A base de seu funcionamento é uma máquina de *workflow* interna, que sequencia as chamadas de cada componente para o resultado desejado.

O Portal, a porção do *Pentaho* visível ao cliente final. Através dele o cliente navega entre as soluções e aciona a execução de qualquer recurso, como um relatório ou *dashboard*.

3 Preliminares

3.1 Metodologia Atual

Atualmente os Correios possuem um *Data Warehouse* com dados de diversos sistemas que a empresa possui, de todas as unidades espalhadas pelo Brasil. A ferramenta utilizada para acessar esses dados é o *MicroStrategy* em sua versão *web*. Geralmente o acesso a essa ferramenta é liberado para gestores de nível estratégico e para alguns gestores operacionais caso desejem.

Conforme a Figura 1, os dados são organizados por pastas para facilitar o seu acesso por parte dos usuários. Geralmente as pastas contêm dados de alguns sistemas ou de indicadores específicos da empresa.

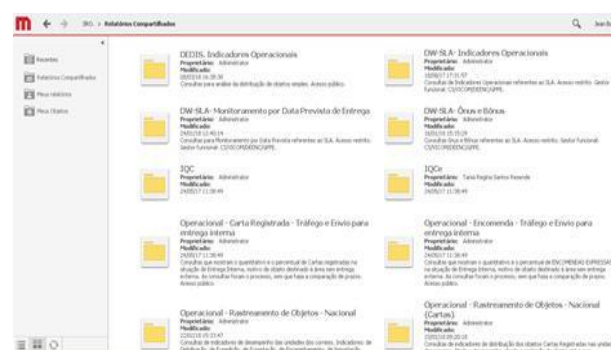


Figura 1 - Pastas de trabalho no *MicroStrategy*.

O sistema possui uma interface gráfica amigável que permite o usuário a realizar consultas de forma simples, escolhendo alguns parâmetros pré-determinados. Um dos problemas é o tempo de duração para se realizar uma consulta, pois devido a grande quantidade de dados de todo o Brasil algumas consultas

demoram bastante. Na Figura 2 podemos ver como é gerado um relatório.

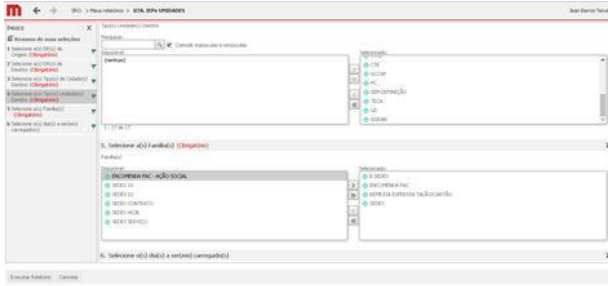


Figura 2 - Tela de geração do relatório.

Na Figura 3 podemos ver o resultado de um relatório criado na ferramenta. O usuário tem algumas opções, como adicionar mais alguns atributos disponíveis para agregar mais dados ao relatório, realizar exportação assim como gerar um gráfico.

Figura 3 - Relatório gerado.

A Figura 4 apresenta um gráfico gerado pela ferramenta, um gráfico simples para uma análise rápida e algumas opções de customização do gráfico.

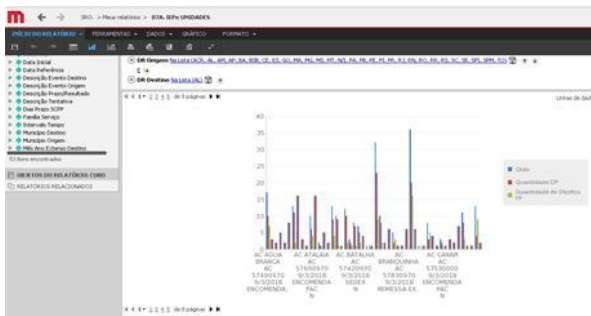


Figura 4 - Gráfico da consulta realizada.

Em resumo a ferramenta utilizada atualmente nos Correios ajuda os gestores a encontrarem os dados dos diversos sistemas em um único local, de forma simples, mas para se realizar análises

bem mais profundas, cruzar dados ou até mesmo gerar *dashboards* que auxiliem na tomada de decisão o usuário precisa recorrer a outras ferramentas, exportar os dados e utilizá-los em planilhas eletrônicas, exigindo um pouco mais de habilidades técnicas. Além disso vale ressaltar que o tempo de duração das consultas costuma ser consideravelmente alto dependendo da sua complexidade.

3.2 Sobre os dados

Os dados foram obtidos do DW dos Correios através da ferramenta *MicroStrategy Web*, o relatório escolhido foi referente ao Índice de Entrega no Prazo - IEP e possui 60.825 registros de novembro/2017 a Março/2018. Tabela 1 mostra os atributos:

Tabela 1: Tabela com os principais atributos.

N	Atributo	Tipo	Valores
1	Unidade	Categórico	112
2	Tipo de Unidade	Categórico	3
3	Data final	Data	Nov/2017 a Mar/2018
4	Categoria Serviço	Categórico	5
5	Âmbito	Categórico	2
6	Quantidade de objetos entregues	Numérico	De 1 a 2280
7	Quantidade de objetos fora do prazo	Numérico	De 0 a 1402
8	Quantidade de objetos fora do prazo unidade	Numérico	De 0 a 1500

4 Metodologia proposta

Nesta seção será apresentada a proposta de uma solução de *Business Intelligence Open Source*, foi escolhida a suite *Pentaho* pois, além de ser uma ferramenta livre a mesma possui um conjunto de ferramentas robustas. A Figura 5 mostra a metodologia proposta.

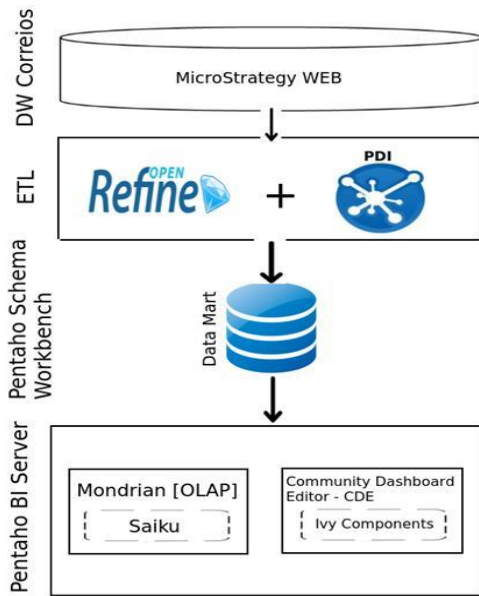


Figura 5 - Metodologia proposta

Na primeira camada, o DW dos Correios será acessado através do *MicroStrategy Web*, já que não se conseguiu acesso direto ao mesmo, e dele exportado um arquivo .CSV que será utilizado na etapa de ETL.

Na segunda camada, de ETL, utilizou-se duas ferramentas, a primeira foi o *Open Refine* pois apesar dos dados serem oriundos de um DW foi necessário alguns refinamentos nos mesmos. A outra ferramenta utilizada foi o *Pentaho Data Integration - PDI*, que faz parte da suite pentaho e foi utilizada para dar carga dos dados no *Data Mart*.

Na terceira camada através da ferramenta *Pentaho Schema Workbench* foi desenvolvido um esquema baseado na técnica de modelagem multidimensional para a criação do cubo OLAP utilizado posteriormente para consultas em tempo real e auxiliando no desenvolvimento do *Dashboard*.

Na quarta e última camada foi utilizado o *Pentaho BI Server* com algumas de suas ferramentas. Para realizar consultas no cubo de acordo com as necessidades apontadas pelo especialista, foi utilizada uma ferramenta OLAP denominada *Saiku*. Essas consultas serviram de base para o desenvolvimento de um *Dashboard*, através do *Community Dashboard Editor - CDE* e utilizando um *plugin* chamado *IvyDashboard* que

possui alguns componentes que permitem uma melhor experiência do usuário.

5 Desenvolvimento e Resultados Modelagem multidimensional

Para facilitar a modelagem multidimensional foram utilizadas as quatro etapas propostas por *Ralph Kimball* para analisar um processo de negócio, conforme ilustrado na Figura 7. No estudo de caso em questão foi definido como processo de negócio a qualidade da entrega (etapa 1), em seguida foi definido que o grão ou nível de detalhamento seria o dia (etapa 2), as dimensões (etapa 3) foram escolhidas para atender os questionamentos dos gestores: Quais unidades ofensoras por tipo de serviço? Qual o âmbito com mais perda de prazo? Quais tipos de unidades são as mais ofensoras? E por último foram definidos as métricas ou indicadores (etapa 4).

5.1 ETL

Esta fase é referente à extração, transformação e carga dos dados no *Data Mart* como mostra a Figura 6. O ETL foi feito com a ferramenta PDI, como entrada de dados temos um arquivo no formato csv que passou por um refinamento anteriormente através da ferramenta *OpenRefine*, foi realizado também uma higienização dos nomes dos atributos além de outras transformações e por fim temos a carga em todas as dimensões e na tabela fato.

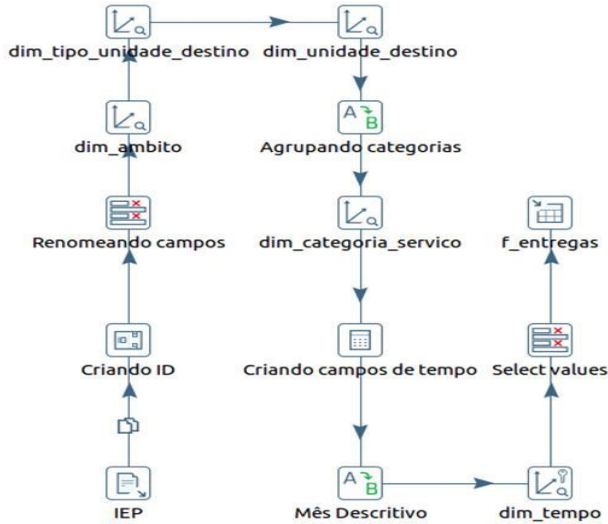


Figura 6 - ETL para criação do data mart.

5.4 Data Mart

O Data Mart foi desenvolvido baseando-se na modelagem multidimensional mencionada anteriormente, dessa forma os dados estão armazenados em várias dimensões (tabelas) que seguem um esquema denominado de estrela como ilustrado na Figura 7. O fluxo de processo do negócio realizado na primeira etapa de desenvolvimento facilitou na implementação do Data Mart.

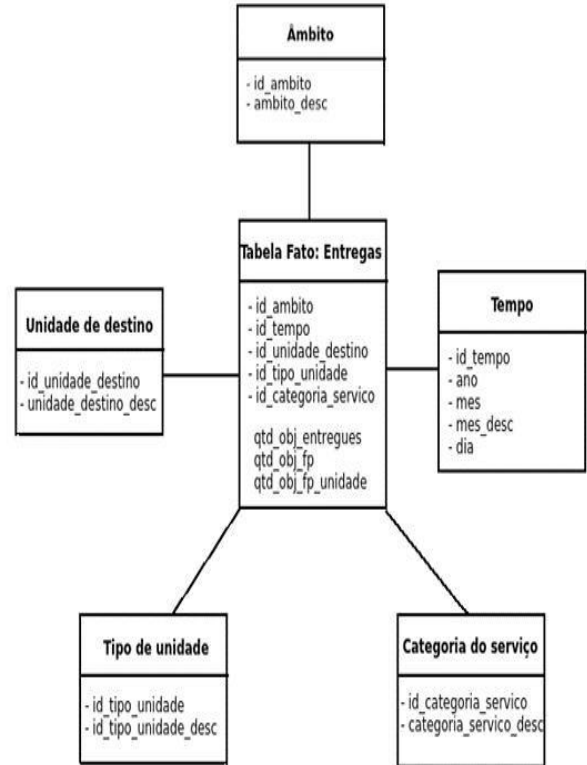


Figura 7 - Modelagem multidimensional na arquitetura em estrela.

5.4 OLAP

Após a criação do Data Mart foi possível realizar algumas consultas no cubo através da ferramenta Saiku que possui uma interface bem amigável permitindo manipular as dimensões e métricas de forma fácil e prática. A Figura 8 mostra o resultado de uma consulta simples criada a partir da dimensão tipo de unidade e da medida quantidade de objetos fora do prazo.

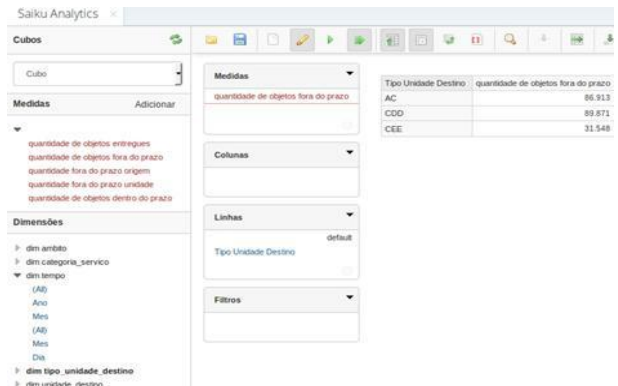
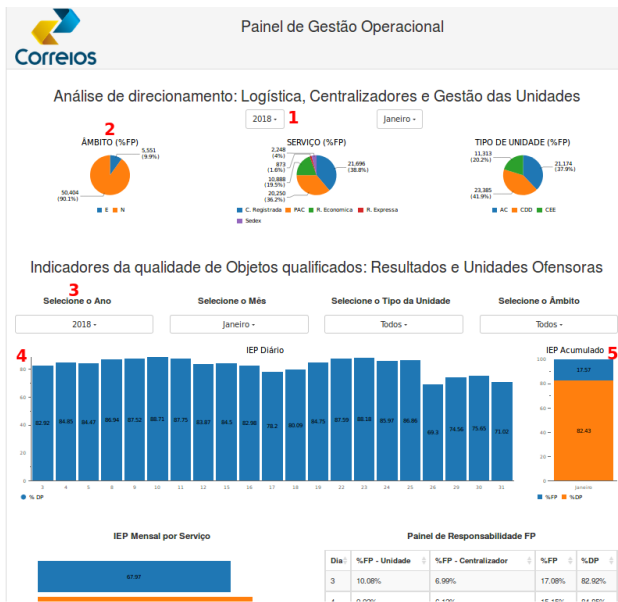


Figura 8 - Consulta OLAP utilizando o Saiku.

5.5 Dashboard

Um *dashboard* ou painel de bordo expõe uma visão ampla das informações mais importantes de uma organização, geralmente em uma única tela e permitindo uma interatividade por parte do usuário.



Para desenvolver o *dashboard* proposto utilizamos o *Community Dashboard Editor (CDE)*, uma ferramenta integrante da suíte *Pentaho* voltada para o desenvolvimento de *dashboards*. Também foi utilizado o *plugin Ivy*, pois ele possui alguns componentes que permitem tornar a interface gráfica mais atrativa e de fácil interação.

As figuras 9 e 10 apresentam o *dashboard* criado com o apoio de um especialista em estatística dos Correios que apontou quais informações eram relevantes para os usuários (Diretor, gestores operacionais, coordenadores etc). Cada componente do *dashboard* será descrito a seguir de acordo com seu respectivo número nas figuras.

Filtros (1): Existem dois filtros vinculados apenas aos gráficos da análise de direcionamento onde o usuário pode escolher o ano e o mês;

Gráficos direcionadores (2): Esses três gráficos de pizza tem o objetivo de alertar e direcionar a gestores do nível estratégico qual o âmbito (nacional ou estadual), serviço (PAC, Sedex etc) e tipo de unidade (AC, CDD ou CEE) são os ofensores em um respectivo mês do ano. Como por exemplo, na figura 11, um gestor estratégico poderia identificar que em Alagoas no mês de Janeiro a carta registrada (serviço) nacional (âmbito) nos CDD's (tipo de unidade) foi um ponto crítico e precisa ser verificada.

Filtros (3): Esses são o segundo conjunto de filtros vinculados a segunda parte do painel, o usuário além de poder escolher ano e mês, também tem a opção de escolher o tipo da unidade e o âmbito;

Índice de Entrega no Prazo – IEP diário (4): Este gráfico mostra o resultado diário do IEP, calculado pela razão entre os objetos entregues no prazo e o somatório de objetos entregues (dentro e fora do prazo).

IEP acumulado (5): Este gráfico foi adicionado, pois os resultados diários apresentam oscilações que muitas vezes não estão ligadas a um problema crônico. Pode ter ocorrido por um fator isolado (ausência elevada, quebra de veículo etc). Já resultado do mês (acumulado) mostra a real situação do indicador. O desvio causado por esse fator isolado é diluído na média do mês e não é uma variável significativa.

IEP mensal por serviço (6): Esse gráfico mostra o resultado acumulado mensal por tipo de serviço. Através dele pode-se verificar os problemas em um nível mais elevado já que podem ser verificados particularidades como logística, origem da carga, prazos, prioridades na entrega entre outros.

Painel de responsabilidade – FP (7): Nesta tabela são apontados os percentuais de objetos fora do prazo, com caracterização de responsabilidade, ou seja, se a perda do prazo foi da unidade ou do centralizador;

IEP das maiores unidades em quantidade de objetos (8): As 15 maiores unidades em quantidade de objetos qualificados (Sedex, PAC etc) representam 70% da carga do estado de Alagoas. Considerando a "lei dos poucos vitais" (80/20), 80% das consequências (quantidade de

objetos total, dentro do prazo, fora do prazo etc) decorrem de 20% das causas (unidades).

Unidades ofensoras (9): Nesse gráfico são apontadas as unidades ofensoras, ordenadas pela quantidade de objetos fora do prazo. Assim como a tabela do item 8, o gráfico ajuda aos gestores a identificar unidades que causam grande impacto negativo nos resultados do estado.

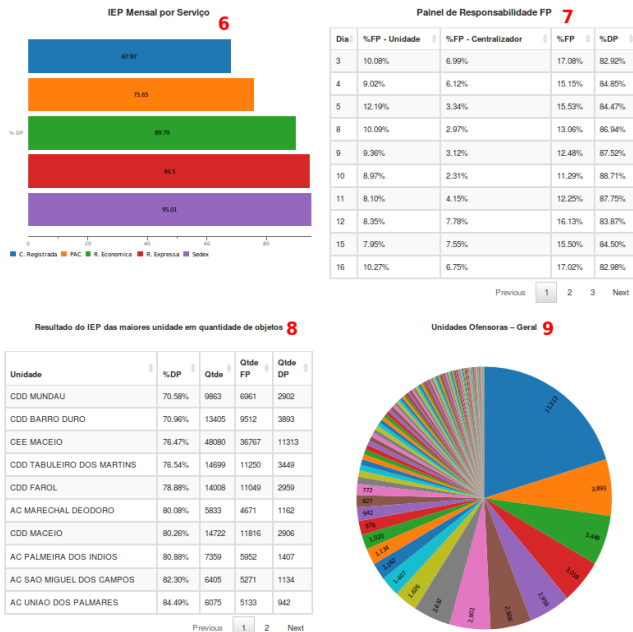


Figura 10 - Dashboard de gestão operacional.

6 Conclusão

Diante da necessidade de se obter vantagem competitiva e um suporte a tomada de decisão, utilizar um software que ajude nessa tarefa é primordial em uma organização. Dessa forma a solução de Business Intelligence proposta e implementada nesse trabalho conseguiu suprir as necessidades dos gestores e responder diversos questionamentos sobre o processo de negócio estudado.

Anteriormente era necessário utilizar-se de outros meios para realizar algumas análises com os dados provenientes do DW dos Correios. Além disso, o tempo de duração para gerar os relatórios é um pouco elevado. Com a nova solução, o usuário pode acessar um *dashboard*

que contém diversas informações relevantes em um único local e com uma interface fácil de utilizar.

Conclui-se que o objetivo do trabalho foi alcançado visto que os gestores após análise do *dashboard*, conseguem a partir de suas conclusões, gerar planos de ações e tomadas de decisão mais assertivas para melhorar a qualidade da entrega dos objetos postais.

Agradecimentos

Agradeço a Empresa Brasileira de Correios e Telegráfos pelo custeio de 80% do curso.

Referências

- [1] BARBIERI, Carlos. **BI2 - Business Intelligence**: modelagem e qualidade. Rio de Janeiro: Elsevier, 2011. p.392.
- [2] BATISTA, Cleisson Fabrício Leite et al. Proposta de *data mart* para análise de faturamento de empresa de varejo utilizando *software* livre. **Revista Brasileira de Administração Científica**, v.3, n. 2, p.163-180, 2012.
- [3] CAUTELA, A. L.; POLIONI, F. G. F. **Sistemas de informação**. São Paulo: Livros Científicos e Técnicos, 1982.
- [4] DIAS, Jorge Luis Ferreira. Pentaho - BI: conhecendo a plataforma, arquitetura e infraestrutura. **Devmedia**, 2014. Disponível em <<https://www.devmedia.com.br/pentaho-bi-conhecendo-a-plataforma-arquitetura-e-infraestrutura/31502>> Acesso em:10 mar. 2018.
- [5] Arquitetura genérica de Business Intelligence. Disponível em: <www.it4biz.com.br> Acesso em: 12 mar. 2018.
- [6] Conceitos sobre Pentaho. Disponível em: <<http://www.pentaho.com/>> Acesso em: 20 mar. 2018.