

Proposta de Data Mart para acidentes aeronáuticos da aviação civil

Arthur Diego de Godoy Barbosa¹, Natan de Souza Albuquerque², Ellen Polliana Ramos Souza³.

Introdução

O transporte aéreo é uma atividade muito importante para a locomoção de pessoas a longas distâncias em um pequeno tempo. Sendo assim é uma atividade-meio necessária para a execução de atividades fins. Segundo Salgado(2009) cerca de 70% da demanda por transporte é representada por clientes corporativos, já a demanda turística é estimada em torno de 25%. Diante da necessidade desse transporte pelo meio corporativo, cresce também assim o número de voos fretados e aquisição de aeronaves particulares. Esses números refletem a importância que a locomoção aérea promove em todo cenário nacional, aumentando o giro do capital e estimulando a economia. Segundo a INFRAERO entre os anos de 2011 e 2012 as taxas de crescimento do movimento de Aeronaves e Passageiros foram respectivamente de 3,76% e 7,32%. Com esse crescimento tornaram-se mais frequentes as ocorrências de acidentes aéreos que fazem vítimas fatais e resultando em danos que dão prejuízo de milhões de reais para as empresas e para o estado.

Diante desse cenário, o governo volta esforços para órgãos como Serviço de Investigação e Prevenção de Acidentes Aeronáuticos (SIPAER), Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA) e outros, que, segundo Cardoso(2004) tem a função de investigar os acidentes da aviação civil e emitir as recomendações de segurança aplicáveis, além de outras atividades que previnam os acidentes aeronáuticos. A ANAC (Agência Nacional de Aviação Civil) mantém os dados de acidentes aéreos civis ocorridos em todo o território brasileiro. Esses dados são anualmente publicados no portal da ANAC e sua base de dados estatísticos.

Entretanto, estes dados publicados pela ANAC não estão em um formato adequado para processos de tomada de decisão, uma vez que encontram-se distribuídos em vários documentos e esses não são estruturados de forma a facilitar a realização de consultas. Essa dificuldade inviabiliza a extração de informações significativas a partir desses dados.

O objetivo deste trabalho é construir um Data Mart utilizando ferramentas Open Sources para armazenar e manipular esses dados. Assim, estruturando-os de maneira a resolver os problemas que inviabilizam o aproveitamento deles para o processo de tomada de decisões.

A primeira parte do artigo, referenciada ao material e método é dedicado a detalhamento do planejamento do Data Mart, abordagem utilizada e as ferramentas que auxiliaram o desenvolvimento de ETL. Logo após é feita uma ligação relacionando o Data Mart construído com outros trabalhos relacionados a área de ocorrência de acidentes aéreos. Por fim é feita uma pequena amostragem dos resultados obtidos e abre espaço para sugestões de pesquisas futuras.

Material e métodos

Na fase de planejamento definiu-se como área de negócio o setor de segurança aérea nacional, observando que esse setor abrange uma ampla gama de subsetores e domínios, foi escolhida em especial a área de acidentes aéreos que por fornecer dados acessíveis promove um amplo potencial de implantação bem sucedida de um projeto de Data Mart.

A abordagem escolhida para a construção do Data Mart foi a Bottom-UP, pois essa abordagem possibilita resultados vantajosos e compatíveis com a necessidade atuais do negócio. Dessas vantagens pode-se destacar a disponibilidade a curto prazo de um produto utilizável, uma ampla capacidade de escalabilidade e evolução para outras áreas do ambiente de escopo - neste o de aviação civil; além de possibilitar requisitos como planejamento mais fácil, levando em consideração uma comparação com a abordagem Top-Down. Em conjunto com a abordagem Bottom-Up é utilizado o Star Schema, que possibilita consultas fáceis e rápidas, e com a principal vantagem de poder-se armazenar um modelo dimensional num banco de dados relacional. (BARBIERI, 2001)

Para a definição da Arquitetura Tecnológica foi selecionado como SGBD (Sistema Gerenciador de Banco de Dados) a ferramenta MYSQL. A base de dados foi retirada do portal da Agência Nacional de Aviação Civil (www.anac.gov.br). Esta base encontra-se no formato PDF, a suíte de ETL do PENTAHO não dá suporte para o processo de extração para esse tipo de formato, a estratégia para solução dessa limitação foi fazer essa extração manualmente, através do simples processo de copiar e colar.

O arquivo resultante do processo anterior, foi verificado e algumas imperfeições nos formatos dos dados foram corrigidas, com o intuito de cada valor estar exatamente em uma célula da tabela. Após essas correções, o arquivo foi exportado para formato CSV. O arquivo CSV foi utilizado como entrada no PENTAHO.

¹ Primeiro Autor é discente do curso de Bacharelado em Sistemas de Informação, Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada. Fazenda Saco, s/n - Zona Rural, Serra Talhada - PE, 56909-460. E-mail: arthudiego@gmail.com.

² Segundo Autor é discente do curso de Bacharelado em Sistemas de Informação, Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada. Fazenda Saco, s/n - Zona Rural, Serra Talhada - PE, 56909-460. E-mail: natanalbuquerque5@gmail.com

³ Terceiro Autor docente do curso de Bacharelado em Sistemas de Informação, Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada. Fazenda Saco, s/n - Zona Rural, Serra Talhada - PE, 56909-460. E-mail: ellen.polliana@gmail.com.

O processo de ETL feito pelo SPOON programa integrante da suíte PENTAHO, desde a coleta dos dados dos arquivos CSV, passando por transformações para adequações dos dados com a finalidade de evitar inconsistências, até a gravação do banco de dados.

Trabalhos Relacionados

Existe um portal na Web (<http://www.desastresaereos.net/>) que faz registro de acidentes aéreos ocorridos no Brasil e em outras locais do mundo, esse portal usa como uma das suas fontes de dados os dados disponibilizados pela ANAC. A forma de exposição dos dados desse portal não apresenta solução ao problema da estruturação dos dados para facilitar o processo de manipulação e consultas relevantes para a extração de informações. Ele limita-se a replicar algumas informações e expô-las na mesma abordagem utilizada pela ANAC.

A solução proposta neste artigo resolve esses problemas, agregando qualidades como centralização dos dados que antes encontravam-se distribuídos em vários documentos, removendo inconsistências e dando maior qualidade aos dados. Ambos os trabalhos expõem os dados em tabelas o que dificulta a visualização de dados e obtenção de informações mais consistentes. Essa possibilidade de uma amostragem mais útil dos dados para melhor extração da informação que eles geram, é o diferencial do nosso projeto, pois permite consultas interativas e uma visualização mais específica, elaborada e organizada dos dados.

Resultados

O modelo resultante representa as informações de maneira consistente e otimizada para a realização de consultas. Características advindas da utilização do modelo Star Schema conforme Barbieri (2001). A utilização das ferramentas MySQL, como sistema gerenciador de banco de dados, e da suíte PENTAHO para a criação do Data Mart resultou na criação num sistema de alta qualidade sem nenhum custo com aquisição de licenças, uma vez que essas ferramentas são disponibilizadas gratuitamente.

Para a construção desse Data Mart, como mostra a Figura 1.A, foram identificadas várias perspectivas para a tabela ‘fato’ que representa a ocorrência de acidente. Para cada perspectiva, uma tabela dimensão foi criada para agrupar dados e informações em comum a cada perspectiva. As dimensões identificadas foram: tipo de acidente, aeronave, operador, tempo e local.

Devido ao sucesso do processo de extração e transformação dos dados, teve-se como produto final um ambiente OLAP que permite a realização de consultas complexas, que permitem a extração de informações envolvendo as variações de tempo, local, aeronave, operador e tipo de acidente das ocorrências a partir da base.

Como resultado de algumas consultas é possível saber das pesquisas que nas ocorrências de acidentes aéreos, o ano, de 2012 teve maior número de ocorrências, totalizando 173 acidentes como pode ser visto na Figura 1.B. Os principais tipos de acidentes foram falha de motor em voo com 156 ocorrências e perda do controle em voo com 130 ocorrências, Figura 1.C. O tipo de aeronave que mais se envolveram em acidentes aéreos foram os aviões com 600 registros seguido dos helicópteros com 128 registros, Figura 1.D. Já o estado com mais ocorrência foi São Paulo com 164 ocorrências, Figura 1.E.

Discussão

O benefício envolvendo o Data Mart construído foi permitir consultas inteligentes das ocorrências de acidentes aéreos em território nacional. Com a disponibilização dessas informações os órgãos responsáveis pela segurança aérea e prevenção de acidentes podem tomar as devidas providências, focando corretamente seus esforços nos reais e mais importantes urgências que envolvem os acidentes aeronáuticos.

Mesmo com tanto detalhamento e benefícios, o Data Mart poderia ser melhor aproveitado se os órgãos nacionais padronizassem as publicações desses dados, além de acrescentar mais informativos referentes ao acidente, como causa dos acidentes, condição dos pilotos, relatórios da caixa preta, outros tipos de voo, entre outros. As informações disponibilizadas pelo Data Mart servem para enriquecer outras pesquisas envolvendo desde o transporte e fluxo de aeronaves do território nacional, até comparativos de acidentes com outros países.

Referências

BARBIERI, Carlos. BI-Business Intelligence – Modelagem & Tecnologia. Editora: Axcel Books do Brasil. Rio de Janeiro, 2001. 52, 53, 54, 147-176 p.

CARDOSO, Vitor Alexandre de Freitas. A abordagem sociotécnica na investigação e na prevenção de acidentes aéreos: o caso do voo RG-254. Rio de Janeiro: Universidade do Estado do Rio de Janeiro, 2004, 80 p.

SALGADO, Lucia Helena. Caos aéreo e tragédia dos comuns: falhas de mercado e de governo. Brasília: Instituto de Pesquisa Econômica Aplicada, 2009, 8 p.

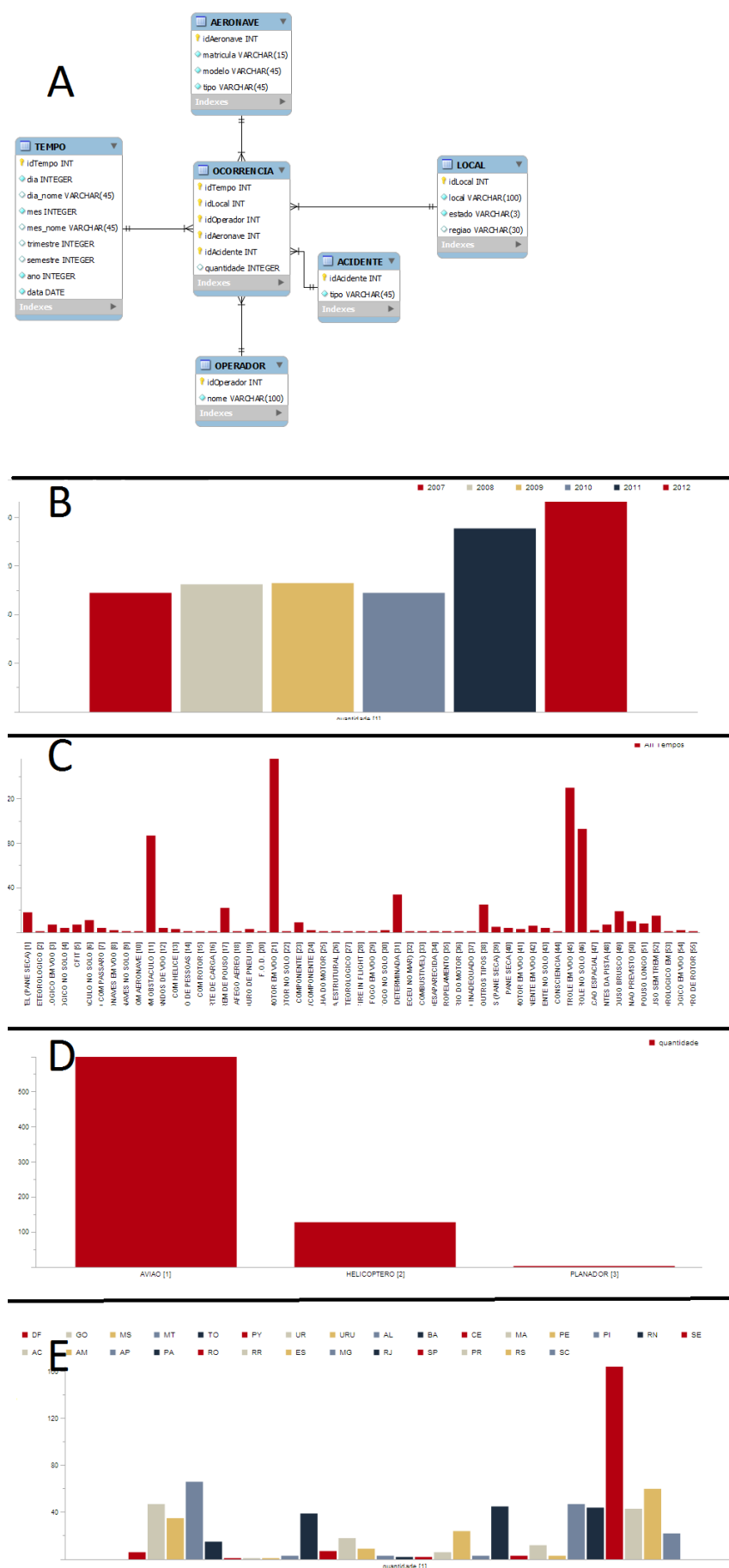


Figura 1 Fig 1A. Diagrama; Fig. 1B, Gráfico de acidentes por ano; Fig. 1C, Gráfico de acidentes por tipo; Fig. 1D, Gráfico de acidentes por tipo de aeronave, Fig. 1E, Gráfico de acidentes por estado.