Um Modelo de Banco de Dados Analítico para Dados de Saúde Pública

Bases da Saúde Pública

O **DATASUS**, criou e mantém, mais de 200 sistemas ligados a gestão de saúde. A maior parte dos sistemas está relacionada à notificação de eventos do cuidado, tais como nascimento, óbito, entre outros. Há também sistemas voltados para a gestão do gasto público, tais como os sistemas para notificação de internações hospitalares e atendimentos ambulatoriais. Muitos desses dados podem ser encontrados facilmente no site do DATASUS.

As bases utilizadas para este trabalho são um recorte das bases do SUS focado na cidade de São Paulo, sendo que as que serão utilizadas são: Sistema de Informações Hospitalares, Sistema de Informação sobre Mortalidade e Sistema de Informações sobre Nascidos Vivos, com o objetivo de realizar uma análise dos dados de uma base de dados integrada.

Modelagem do Data Warehouse

Considerando o grande volume de dados que é gerado pelas bases citadas acima, a sua grande variedade e sua alta frequência de atualizações, a modelagem de um data warehouse para eles foi pensada para que seja possível integrar todas essas diversas áreas em um único modelo, o que resultará em uma otimização na velocidade em que as pesquisas serão realizadas sobre esses dados, além de organizar os dados de acordo com a data em que eles são carregados. Tendo isso em mente sugerimos a seguinte modelagem para o Data Warehouse, utilizando as bases do **SIM**, **SIH** e **SINASC**. Como um exemplo dessa modelagem temos a tabela fato do parto e do óbito e suas dimensões, apresentados na Figura 1:

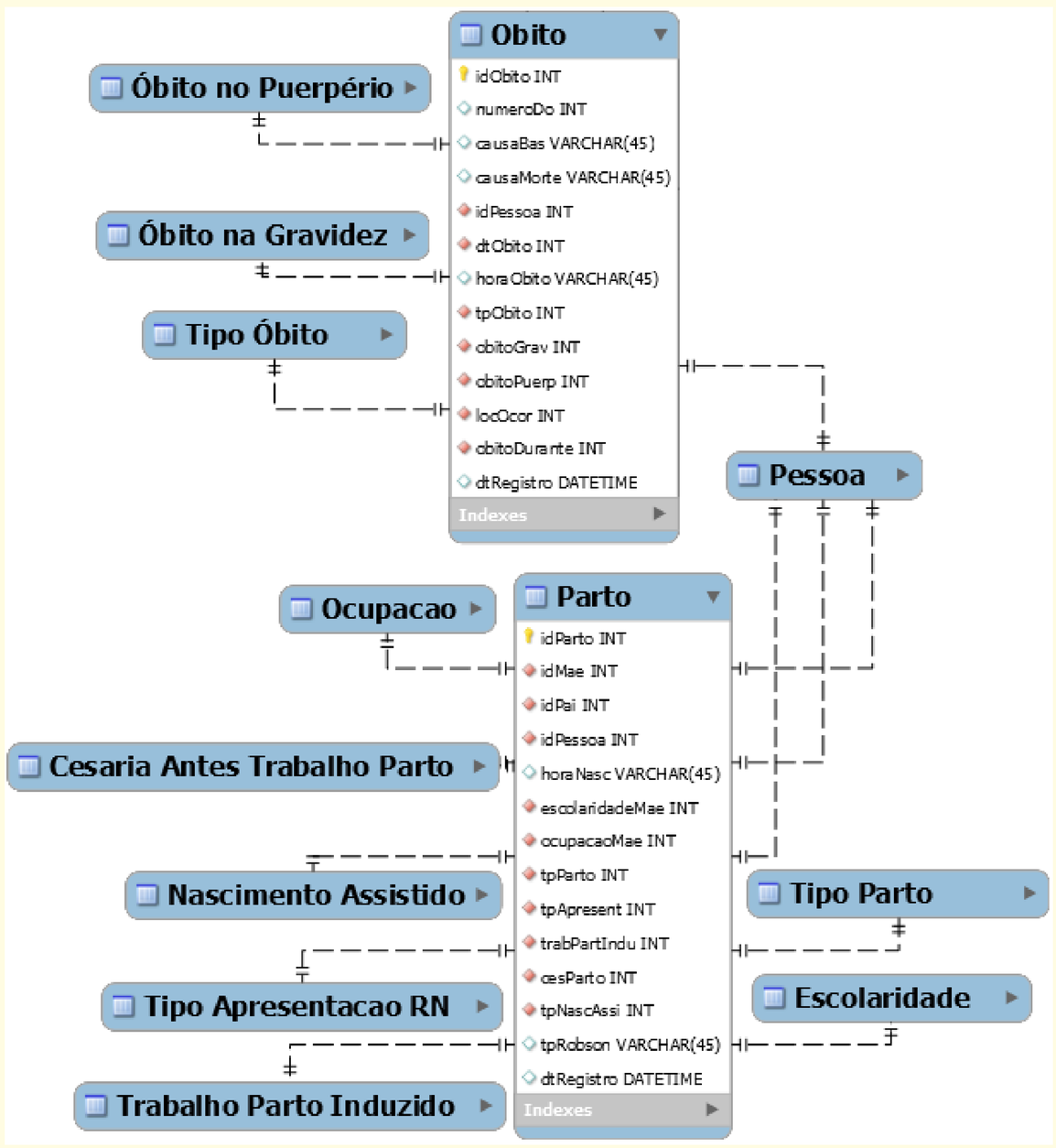

Figura 1: *Tabelas fato óbito e parto e suas dimensões*

Análises

As análises dos dados foram feitas utilizando **Superset Apache**, com o gráfico abaixo podemos observar as quantidades de nascimentos que ocorrem todo mês (segunda a primeira linha e domingo a última), e podemos verificar que ocorrem menos nascimentos durante os fim de semanas, confirmando as dúvidas que secretária tem em relação, de que os médicos preferem realizar partos durante dias de semana.

Por meio de um menu de seleção com diversas opções para alterarmos a dimensão em que a pesquisa está sendo feita, como por exemplo a escolaridade da mãe sendo de 1 a 7 anos e o tipo do parto ser cesária, observamos o resultado na Figura 2.

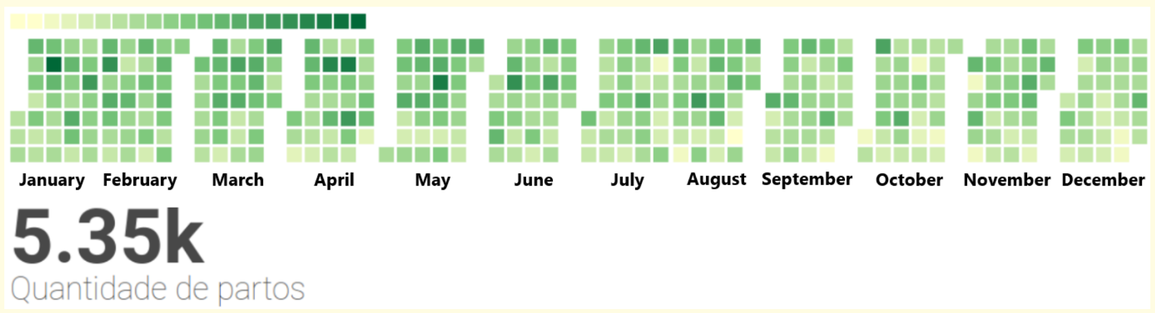

Figura 2: *Nascimento por mês de mães com escolaridade entre 1 a 7 anos que realizaram cesária*

Ao mudarmos a opção do Tipo do parto para Normal, obtemos o resultado da Figura 3.

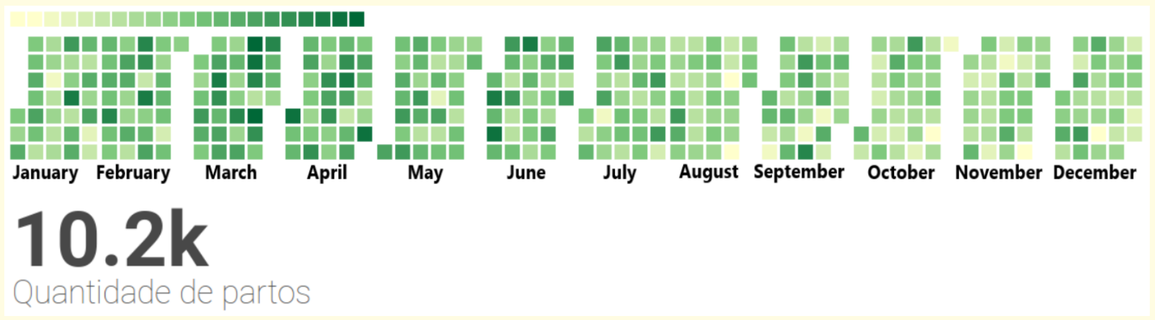

Figura 3: *Nascimento por mês de mães com escolaridade entre 1 a 7 anos que realizaram parto normal.*

Comparando as duas pesquisas podemos ver que a quantidade de partos normais são o dobro para o caso de mães que estudaramentre 1 a 7 anos. E alterando a pesquisa de escolaridade para 8 a 12 anos, nós temos que há 104.000 cesárias e 75.500 normais, então podemos observar uma maior preferência para cesárias em mães com mais estudo.

Um outro exemplo de pesquisa é a quantidade de óbitos por bairro no ano de 2016, mostrando um comparativo de expectativa de vida entre os bairros, que é representado pela Figura 4.

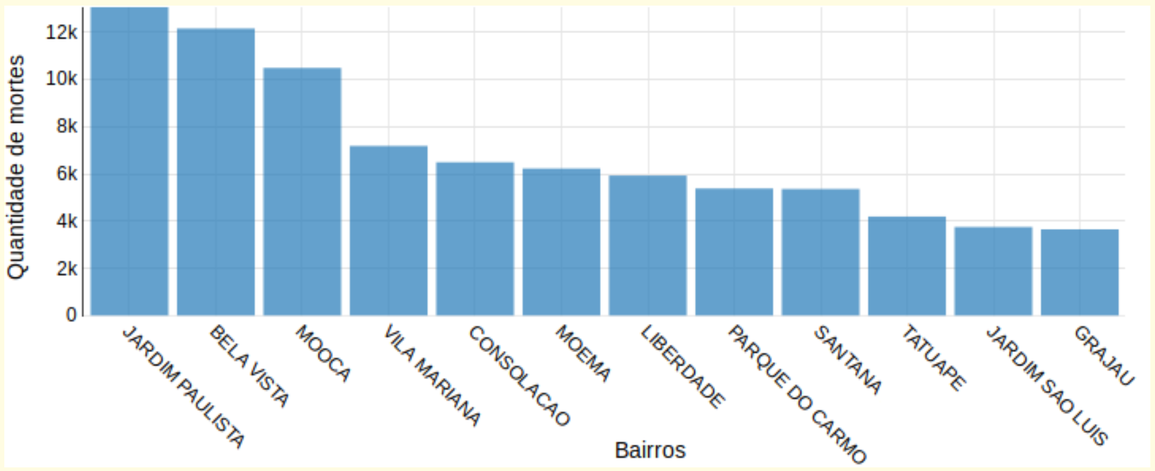

Figura 4: *Quantidade de óbitos por bairro.*

André Akira Hayashi

Departamento de Ciência da Computação — Universidade de São Paulo

MAC0499 Trabalho de Formatura Supervisionado
Orientador: Paulo Meirelles
Coorientador: Renata Wasserman e Kelly Braghetto