

Governança de Dados em Instituições de Ensino Superior

apoiado por um Data Warehouse

Fábio Roberto dos Anjos Santos¹

Orientador: Juracy Almeida²

¹ Universidade Federal da Bahia, ² Companhia de Processamento de Dados do Estado da Bahia - Prodeb, Brasil



Resumo

O presente trabalho visa apresentar uma arquitetura genérica de um Data Warehouse (DW) que possa auxiliar uma Instituição de Ensino Superior a analisar como foi a ocupação de turmas de disciplinas que aconteceram em períodos passados, para que a partir desta informação seja possível realizar inferências para a tomada de decisão sobre planejamentos de disciplinas em períodos futuros. Foram vistos como os conceitos de gestão e governança de dados podem auxiliar uma instituição para a definição de melhores formas de coletar e tratar dados a fim de auxiliar na construção do Data Warehouse. Ao final foi apresentado um modelo dimensional de um data mart para a análise do planejamento a partir do levantamento de questionamentos junto a área de negócios. A finalidade da execução dessas ações foi permitir que os gestores tenham acesso à informação já tratada e de acordo com as necessidades de análise que o negócio carece.

Keywords: Data Warehouse; Data Mart; Governança de Dados; Gestão de Dados; Planejamento de Disciplinas.

1. Introdução

NESTE estudo foi construído um Data Warehouse (DW), que será responsável por armazenar e prover informação sobre os históricos de oferta de disciplina e as ocupações de vaga. A estrutura desse DW deverá ser o mais genérica possível de forma que isolando-se as rotinas de Extração, Tratamento e Carga (ETL), e a Staging área, será possível extrair a informação de qualquer fonte de dados (e.g Bancos de dados, arquivos Excel, arquivos texto). Para isso foram utilizados os conceitos definidos pelo DAMA-DMBOK®, referentes a Governança de Dados.



Figura 1: As 10 funções do DAMA-DMBOK®.

2. Definições

O objetivo principal da **Governança de Dados (GD)** é garantir que os dados sejam gerenciados adequadamente, de acordo com políticas e práticas recomendadas. Esse tipo de ação promoverá mudanças culturais dentro da instituição, fazendo-a entender que agora existe uma política definida em relação aos dados. O uso da GD pode trazer uma série de benefícios, dentre eles o entendimento comum dos dados, a qualidade aprimorada dos dados, o mapa dos dados.

A **Gestão do Data Warehousing e Business Intelligence (GDBI)** consiste na coleta, integração e apresentação dos dados para fins de análise de negócios e tomadas de decisão. Ela visa dar suporte a todas as fases de desenvolvimento do DW, que será o destino comum dos dados.

Um **Data Warehouse** é uma combinação de dois componentes principais: um banco de dados integrado de suporte à decisão e os programas de software usados para coletar, limpar, transformar e armazenar dados de uma variedade de fontes operacionais e externas. Além de apoiar histórico, analítico e requisitos de business intelligence (BI), um data warehouse também pode incluir data marts, que são cópias de subconjunto de dados do DW.

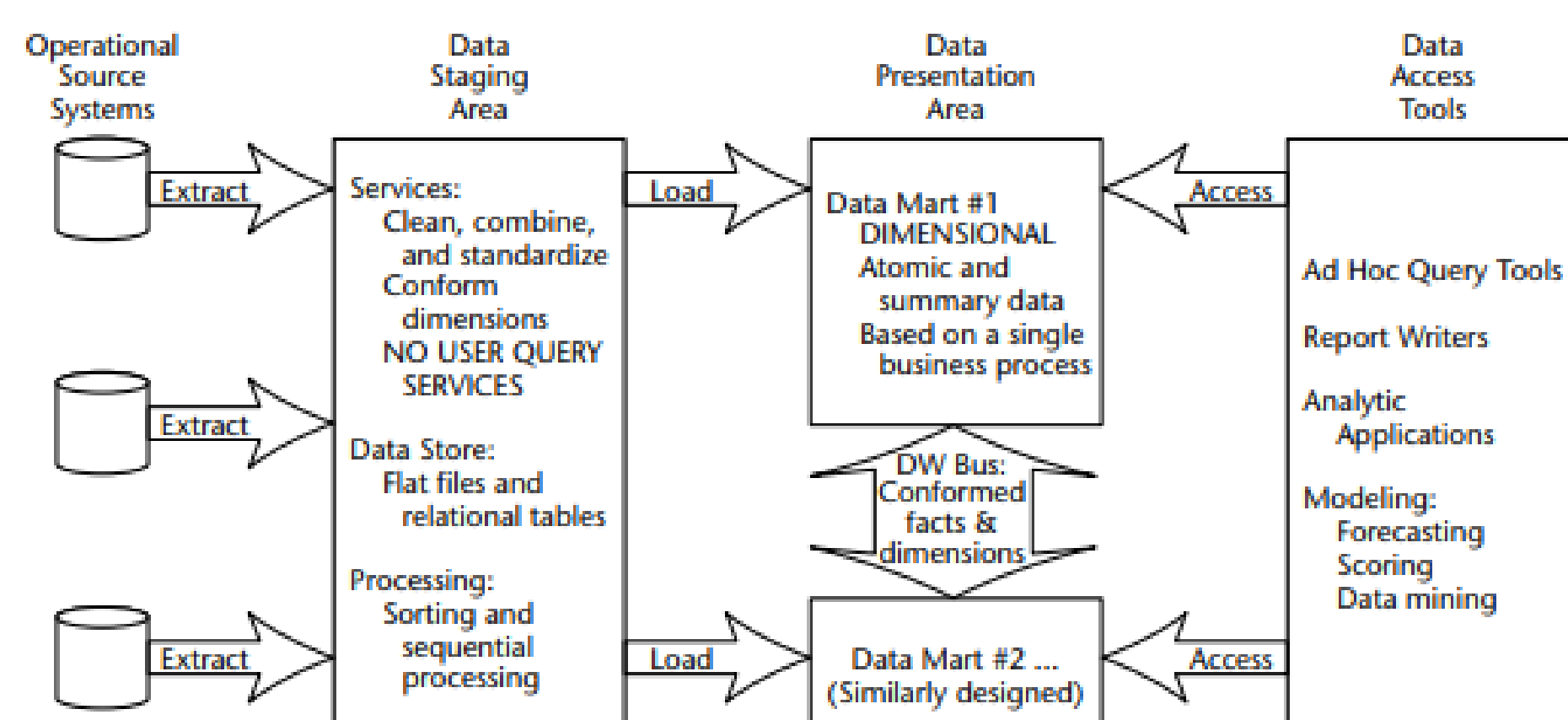


Figura 2: Modelo de solução de um Data Warehouse proposto por Kimball

3. Metodologia

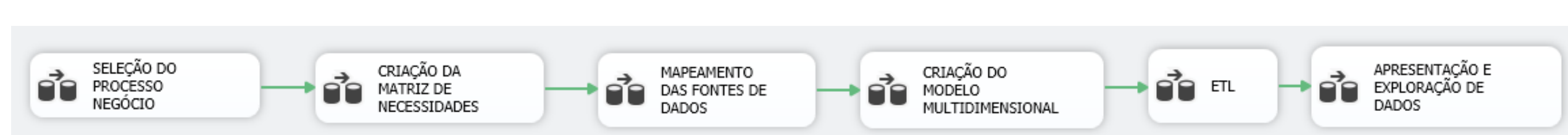


Figura 3: Fluxo de atividades para a construção do DW

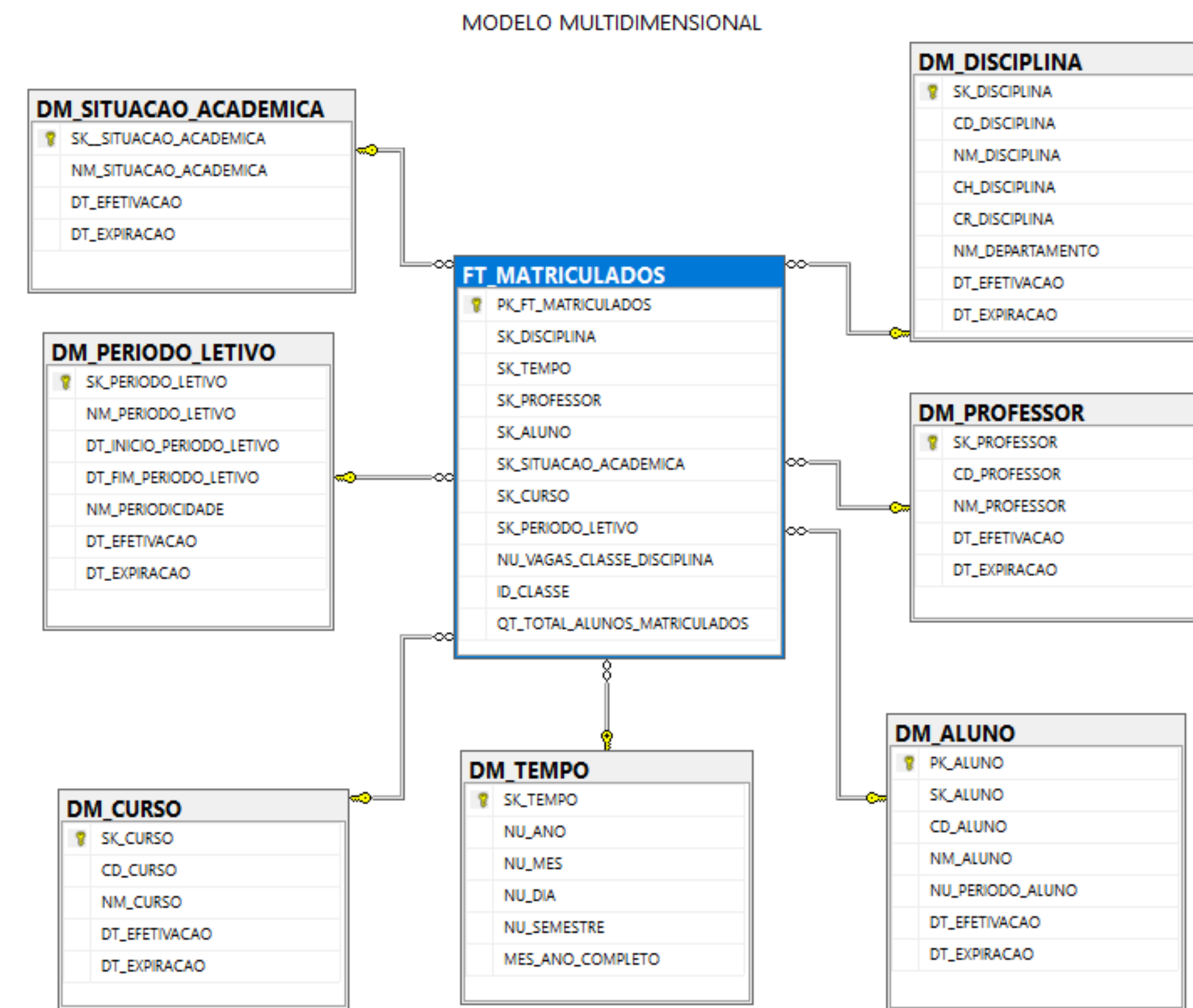


Figura 4: Modelo Multidimensional

Após a criação do nosso modelo Multidimensional foi realizado o processo de ETL:

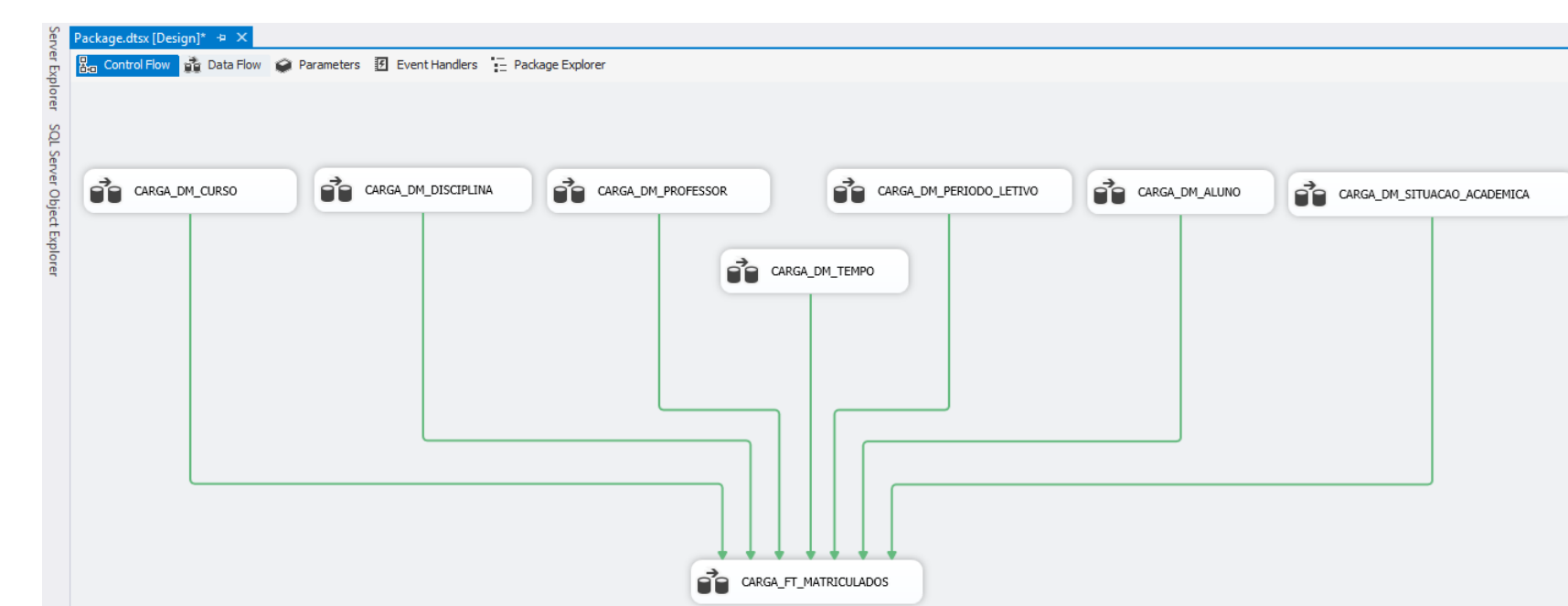


Figura 5: Pacote no SSIS da etapa de transformação

Conclusões

COMO resultado da construção do nosso modelo dimensional, apoiado por iniciativas de Governança de Dados, foi possível construir alguns dashboards a partir do nosso Data Warehouse. Nestes dashboards pudemos utilizar a nossa métrica (quantidade de alunos matriculados), para analisar os percentuais de ocupação:

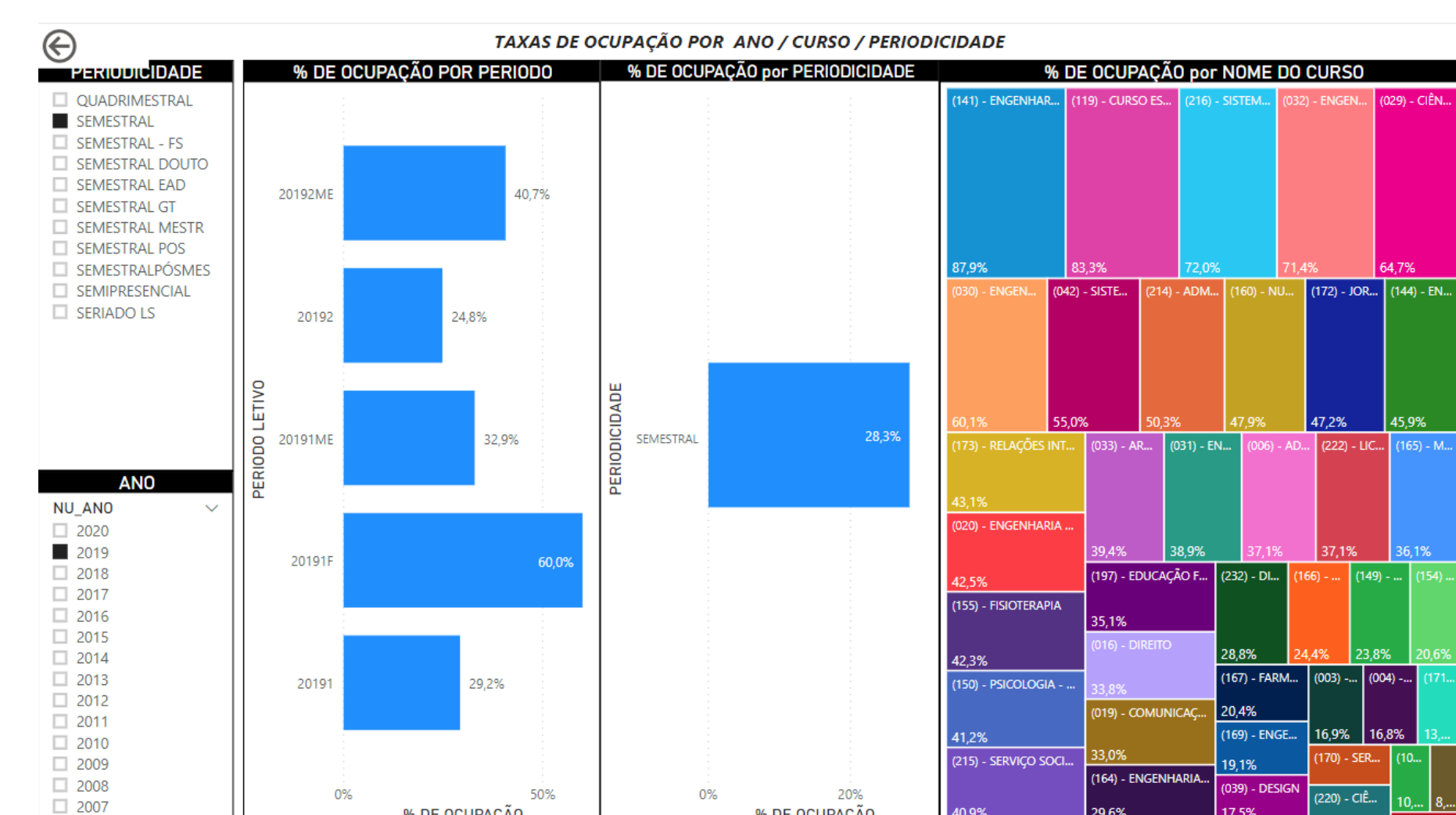


Figura 6: Dashboard de ocupação de vagas por periodicidade

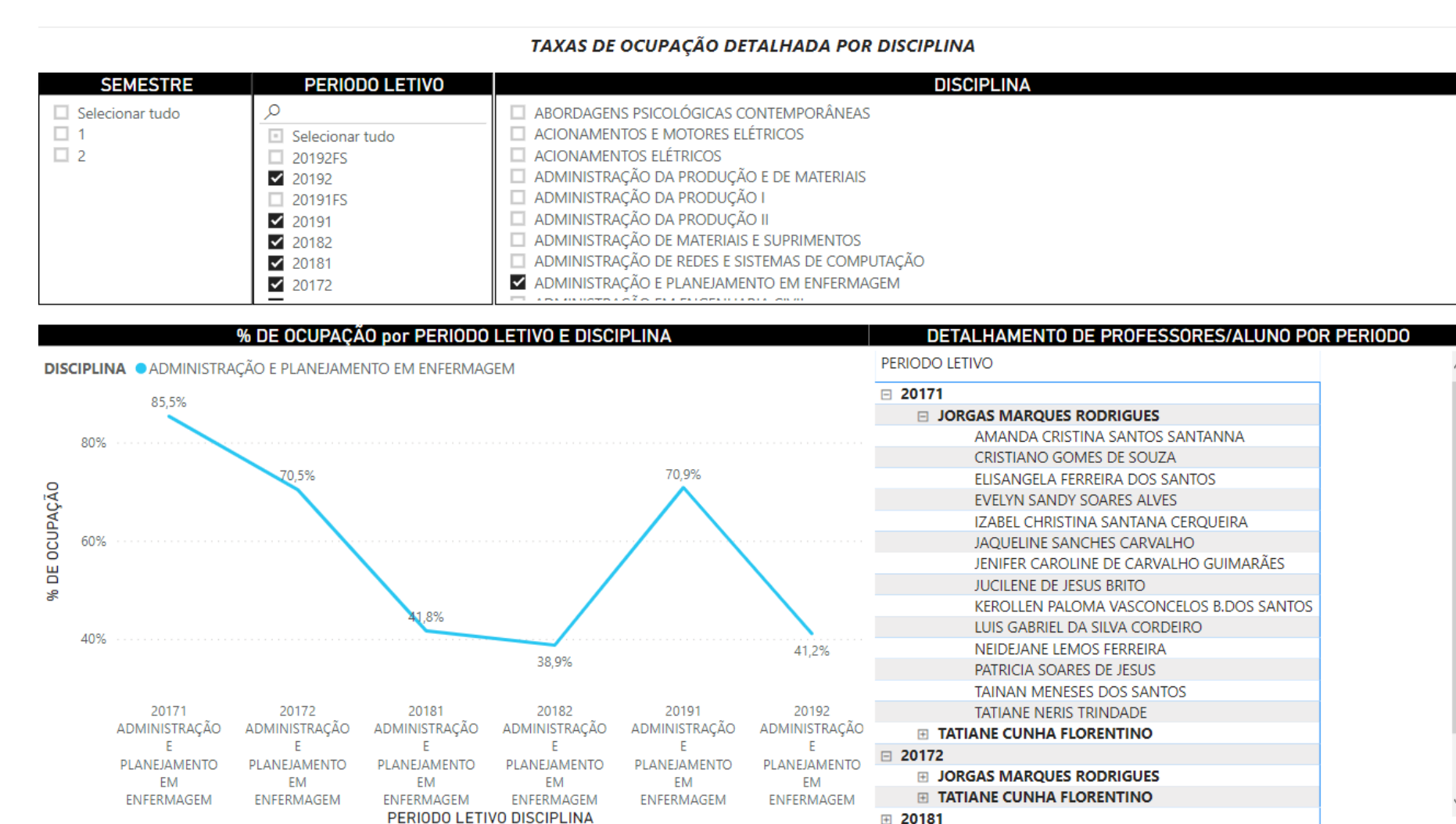


Figura 7: Dashboard de ocupação de vagas por disciplina

Referências

Dama (2017). *The Dama Guide to the Data Management Body of Knowledge*. EUA:Technics Publications.
 Inmon, W. H. (2005). *Building the Data Warehouse*. EUA:Wiley & Sons.
 Kimball, R.; ROSS M. (2003). *The Data Warehouse Toolkit: The Definitive Guide to the Dimensional Modeling*. EUA:Wiley & Sons