

Universidade Federal de Santa Catarina



# Projeto Final

Data Warehouse — INE5643

Augusto Fredigo Hack  
Fabricio Bez  
Luis Felipe Nunes  
Matheus Hoffmann Silva

Florianópolis, 26 de julho de 2013

## Conteúdo

<b>1</b>	<b>Resumo</b>	<b>2</b>
<b>2</b>	<b>Introdução</b>	<b>2</b>
<b>3</b>	<b>Materiais</b>	<b>2</b>
<b>4</b>	<b>Métodos</b>	<b>2</b>
<b>5</b>	<b>Metodologia</b>	<b>3</b>
5.1	Escopo . . . . .	3
5.2	Justificativa . . . . .	3
5.3	Exclusões . . . . .	3
5.4	Riscos . . . . .	3
5.5	Fatores Críticos de Sucesso . . . . .	3
5.6	Definição dos Requisitos . . . . .	3
5.7	Modelagem Dimensional . . . . .	3
5.7.1	Definição do Processo a ser Modelado . . . . .	3
5.7.2	Definição da Granularidade . . . . .	3
5.7.3	Definição das Dimensões . . . . .	3
5.7.4	Definição dos Fatos . . . . .	3
5.8	Projeto Físico . . . . .	3
5.9	Projeto da Área de Transição . . . . .	3
5.9.1	Preparação da Base de Dados . . . . .	3
5.9.2	Área de Transição . . . . .	3
<b>6</b>	<b>Resultados</b>	<b>3</b>
<b>7</b>	<b>Conclusão e Trabalhos Futuros</b>	<b>3</b>
<b>8</b>	<b>Bibliografia</b>	<b>3</b>
<b>9</b>	<b>Especificação – Diagramas</b>	<b>4</b>
<b>10</b>	<b>Conclusão</b>	<b>5</b>

## 1 Resumo

(O quê? Como? O que resultou? ¿ Palavras-chave)

## 2 Introdução

As informações coletadas anualmente pela Universidade Federal de Santa Catarina com as inscrições e resultados de seu vestibular proporcionam uma ampla base de dados. Esta base contém informações sobre todos os candidatos que prestaram vestibular, tanto dados sócio-econômicos (provenientes do questionário sócio-econômico aplicado no momento da inscrição do vestibular) e dados como endereço, data de nascimento do candidato, quanto resultados das provas.

Este trabalho procura utilizar os dados mencionados acima para analisar a situação das cidades do Estado de Santa Catarina quanto à média de aprovações no vestibular, a fim de identificar as regiões com menor média de aprovação. Estas regiões podem ser o foco de atuação de empresas que vendem Cursos Preparatórios para o vestibular, visto a defasagem da região, que podem ser um bom nicho de negócio.

Para enriquecer a base de dados já existente, será usado um levantamento de dados estatístico-educacionais disponibilizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Desse levantamento geral, serão utilizados os dados referentes às instituições de ensino por cidade de Santa Catarina. Estes dados serão úteis para saber onde há regiões onde os estudantes não prestaram vestibular para a UFSC, além de ter dados sobre quantos estudantes se formam e quantos efetivamente prestam vestibular.

## 3 Materiais

Os dados de candidatos que prestaram vestibular foram disponibilizados pela Comissão Permanente do Vestibular (COPERVE), que é a responsável por realizar o concurso vestibular para ingresso na UFSC. Para preservar a identidade e zelar pela privacidade dos indivíduos cujas informações estão presentes no banco de dados, não há meios de identificação presentes no banco, como nome, registro de identidade ou carteira de pessoa física.

No caso deste trabalho, serão utilizados os dados coletados dos vestibulares da UFSC para os anos de 2008 a 2012. Estes dados estão disponíveis em formato .sql, que é um *dump* do banco de dados MySQL.

Além do Banco de Dados, está disponível o modelo Entidade Relacionamento e uma documentação on-line com mais detalhes do banco, que facilitam o entendimento do mesmo.

## 4 Métodos

(Revisão sobre DW e Data Mart, vislumbrando esta tecnologia como a solução para o problema)

## **5 Metodologia**

### **5.1 Escopo**

Utilizar os dados dos vestibulares UFSC de 2008 a 2012, disponibilizados pela Coperve, para analisar a situação das cidades do Estado de Santa Catarina quanto à média de aprovações no vestibular, com o objetivo de identificar as regiões com menor média de aprovação. A identificação destas regiões podem ser úteis, no caso de empresas que vendem Cursos Preparatórios para o vestibular, para investir em novas áreas.

Além dos dados da Coperve, serão utilizados dados do INEP, que indicam o número de estudantes, das várias fases de ensino, das cidades de Santa Catarina.

### **5.2 Justificativa**

### **5.3 Exclusões**

### **5.4 Riscos**

### **5.5 Fatores Críticos de Sucesso**

### **5.6 Definição dos Requisitos**

### **5.7 Modelagem Dimensional**

#### **5.7.1 Definição do Processo a ser Modelado**

#### **5.7.2 Definição da Granularidade**

#### **5.7.3 Definição das Dimensões**

#### **5.7.4 Definição dos Fatos**

### **5.8 Projeto Físico**

### **5.9 Projeto da Área de Transição**

#### **5.9.1 Preparação da Base de Dados**

#### **5.9.2 Área de Transição**

## **6 Resultados**

## **7 Conclusão e Trabalhos Futuros**

## **8 Bibliografia**

## 9 Especificação – Diagramas

Diagrama gerado utilizando a ferramenta *gliffy.com*.

## **10 Conclusão**

Como resultado do trabalho apresenta-se um sistema multi-plataforma, baseado em cliente e servidor, que serve como base para produzir um sistema que auxilie no gerenciamento de finanças pessoais.