

INE 5643 – Data Warehouse

APRESENTAÇÃO DO PROJETO FINAL

Prof. José Leomar Todesco, Dr.

PROJETO FINAL

1. Analisar o modelo do sócio acadêmico do vestibular da Coperve de 2008 a 2012 com vistas à implementação de um Data Mart para suporte a análises;
2. Criar um esquema estrela (modelo dimensional) para suportar as análises;
3. Definir o projeto físico para o Data Mart, seguindo padronização, estabelecer uma área de transição para o processo de ETL;
4. Apresentar o Front End do Data Mart em uma ferramenta OLAP de escolha do grupo;
5. Escrever um artigo do projeto realizado, descrevendo o desenvolvimento realizado;
6. Apresentar em sala o trabalho realizado.

ENTREGAS (ARTIGO)

- Título, Autores, email, afiliação
- Resumo (O quê? Como? O que resultou? – Palavras-chave)
- Introdução (contextualização do problema e da abordagem que será dada ao mesmo, bem como justificativas)
- Materiais (Falar sobre a Coperve e Modelo de provas e ficha-sócio econômico, ressaltando o domínio do problema)
- Métodos (Revisão sobre DW e Data Mart, vislumbrando esta tecnologia como a solução para o problema)
- Metodologia (Proposta com as etapas realizadas para implementar o Data Mart – solução do problema)
- Resultados (Mostrar os principais resultados através de tabelas e gráficos)
- Conclusões e Trabalhos Futuros
- Bibliografia

Modelo SBC - link para site com template <https://www.sbc.org.br/documentos-da-sbc/category/169-templates-para-artigos-e-capitulos-de-livros>

ENTREGAS (PROJETO COMPLETO)

- Todo o projeto deverá ser disponibilizado na nuvem (drive, onedrive, Dropbox, github, ...)
- O projeto deverá ser organizado em pastas
 - Back-end (informações necessárias para acessar o banco de dados, pasta com os scripts/arquivos do kettle utilizados no ETL, modelagem dimensional, requisitos, documentação sobre criação dos indicadores/medidas e tudo que possa ser necessário para para o back-end;
 - Front-end (informações necessárias para acessar a área de apresentação dos dados, link acesso a ferramenta, usuário, senha, documento de especificação e tudo que possa ser necessário para o front-end;
 - Pasta com o artigo final em PDF ou Word, conforme template da SBC;
 - O link para acessar o projeto final deve ser encaminhado para o e-mail jose.todesco@ufsc.br **até as 18 hs do dia 2 de dezembro.**

Obs.: Os grupos que desejarem **entregar o projeto antes, terão a oportunidade de receber feedback e melhorar o projeto**, principalmente a modelagem dimensional.

Banco de dados do vestibular da COPERVE

Vestibular de 2008 a 2012

Dump da base de dados em MySQL

<http://www.inf.ufsc.br/~jose.todesco/dw/Vestibular/>

Arquivos adicionais em excel

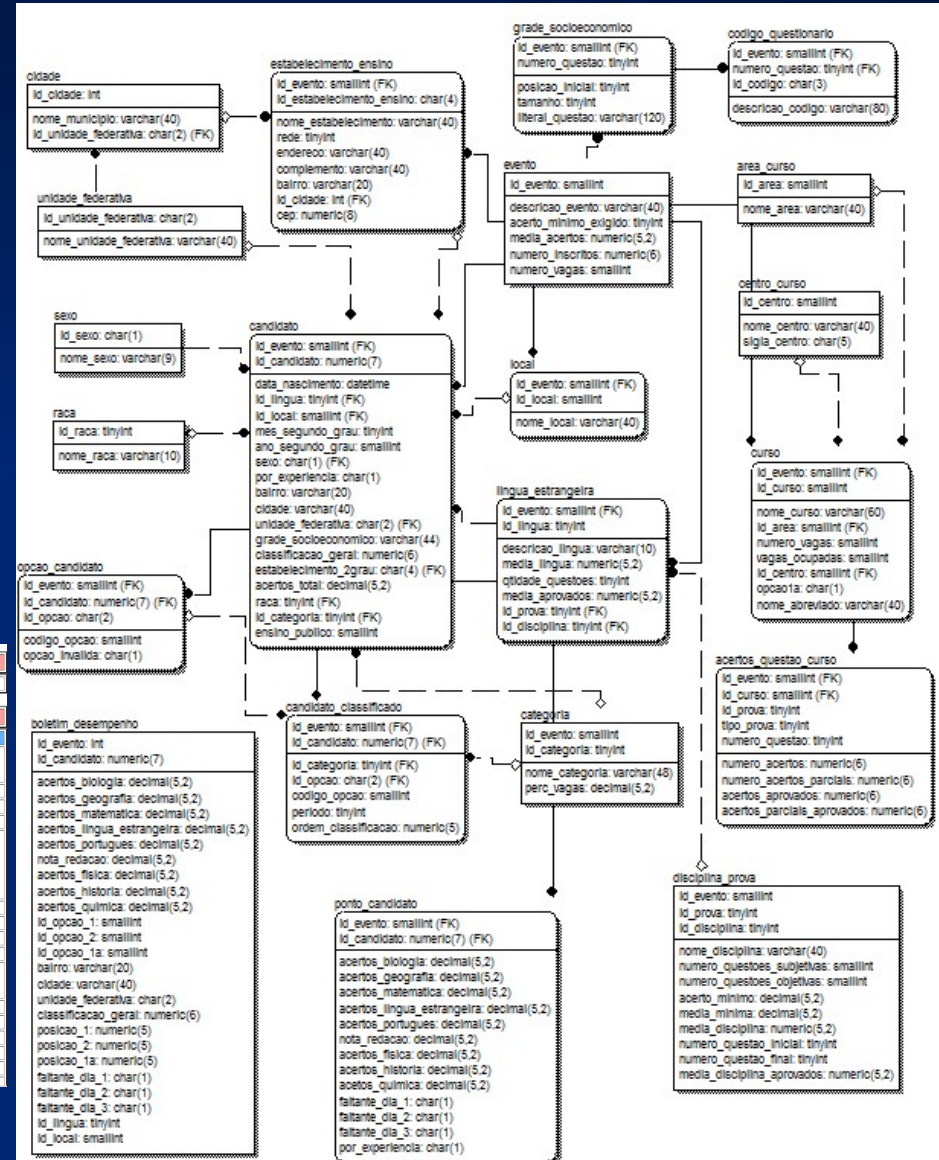
Documentação do modelo de dados

<http://www.inf.ufsc.br/~jose.todesco/dw/Vestibular/Vestibular/>

Está tudo disponível no moodle

Report Components:					
1. 'Picture' section					
Main Subject Area/Display1					
2. 'Tabela' section					
3. 'Coluna' section					
Tabela					
Nome					
candidato					
Coluna					
Nome	Tipo de dado	Comentário	PK?	FK?	
id_evento	smallint	Ano de ingresso dos Calouros na UFSC. Utilizado para a geração do arquivo dos calouros e para o cálculo do número de matrícula na UFSC.	Yes	Yes	
id_candidato	numerico(7)	Número do Candidato	Yes	No	
nome_candidato	varchar(40)	Nome do Candidato	No	No	
data_nascimento	datetime	Data de Nascimento do Candidato	No	No	
id_lingua	tinyint	Código da Língua de Escolha do Candidato : 1 - Alemão 2 - Espanhol 3 - Francês 4 - Inglês 5 - Italiano	No	Yes	
id_local	smallint		No	Yes	
mes_segundo_grau	tinyint		No	No	
ano_segundo_grau	smallint		No	No	
sexo	char(1)	Sexo	No	Yes	
por_experiencia	char(1)	Código que indica condição do 2. grau: S = ainda não concluiu, ou seja, vai fazer vestibular por experiência N = Está concorrendo a uma vaga.	No	No	
bairro	varchar(20)		No	No	
cidade	varchar(40)		No	No	
unidade_federativa	char(2)	Código da Unidade da Federação	No	Yes	
grade_socioeconomico	varchar(44)	Grade de resposta do questionário sócio-econômico	No	No	
classificacao_geral	numeric(6)		No	No	
classificacao_geral	char(3)		No	Yes	

INE 5643 - Data Warehouse Copyright © 2020 José Leomar Todesco



Temas para o projeto final (data marts)

1. Apoiar a secretaria do estado de SC na avaliação do desempenho nas disciplinas dos candidatos das escolas públicas do estado, comparando com as escolas privadas e federais.
2. Apoiar a análise quanto ao investimento na implantação de escola/cursinho no estado de SC. Qual cidade/região melhor de investir numa escola/cursinho.
3. Realizar análise sobre a procura pelo vestibular da UFSC, realizando comparações entre os candidatos x aprovados utilizando características do perfil social do candidato. Correlação entre o desempenho e características do perfil socioeconômico.
4. Apoiar a PRAE (moradia estudantil) e setor imobiliário da grande Florianópolis quanto a demanda por moradia dos candidatos aprovados de fora da grande Florianópolis.
5. Tema livre, o grupo deve propor um tema para o data mart.

Temas sorteados para os grupos

Tema	Grupo	Alunos
1	1	Fernanda Narloch Rizzo Hahn, Henrique Hideki Saito, Matheus Henrique Schaly, Letícia Amarante dos Santos, Nicolas Antero Nunes
2	2	Marina Silva Tavares, Rhanna Kaena Auler, Lucas Germano Willrich, Thalita Scolari Lorenzi, Antonio Moreno Ribeiro Homem
3	3	Lucas Madaloni Meira Varella, Marco Aurelio Geremias, Johny Chen Jy, Joao Vitor Demaria Venancio, Lucas Santos de Oliveira
4	4	Lucas Pagotto Tonussi, Caio Franco Suarez dos Santos, Rafael Barbaresco
1	5	João Antonio Vicini, Matheus Passos Vaz, Felipe Paloschi, Cleber Vinicius, Edwirges dos Santos, Vinicius Marcelino Orlandi
2	6	Lucas Tavares de Souza, Mateus Bittencourt Martins, Elisa de Mattos Rosá, Renan Jochem, Jean Carlos Gonçalves
3	7	Gustavo Ghedin Ditzel, Thiago Brezinski, Igor Cordeiro Szeremeta, Bruno Henrique Bertozzo, Mateus Nunes Cechet
4	8	Thais Goulart Firmino, Jessica Xafranski Rodrigues, Allan Cesar Ferreira, Pedro Germano Cembranel Etges, Christopher de Carvalho
1	9	Thary Correia, Fabiano Pereira de Oliveira, Leon Cesar Carmona
2	10	Maurício de Vasconcelos Barros, Lucas Luciano de Oliveira Picollo, Rian Carlos Vieira Bruggemann, Bruno Miguel Peres Schneider, Rafael Parola
3	11	Artur Antonio Dal Prá, Gabriel Machado da Roza, Eduardo Campestrini, Bruno Hass, Victor André Odebrecht

Agora o trabalho é com vocês

DICAS

1. Façam um bom planejamento com divisão de tarefas;
2. Façam a modelagem dimensional e a escolha das medidas e indicadores o mais breve possível;
3. Sugiram que utilizem o kettle para o ETL;
4. Escolher uma ferramenta de front-end (Pentaho, PowerBI, Tableau, Metabase, Qlik Sense, ...);
5. Dúvidas podem procurar a mim e ao Vinicius (utilizem o e-mail jose.todesco@ufsc.br);
6. Continuaremos fazendo encontros semanais nas segundas-feiras as 20 hs para conversar sobre o projeto e tirar eventuais dúvidas. Não será cobrada presença.
7. Bom trabalho a todos.