

Projeto de Mineração de Dados

CDIA-CI-UFPB

Prof. Sérgio Bezerra

05 de dezembro de 2023

O projeto da disciplina está dividido em 4 partes, denotadas por P_1, P_2, P_3 e P_4 , as quais são entregues ao longo do semestre nas datas definidas no cronograma do curso.

0.1 Etapa P_1

Na primeira etapa (P_1) o aluno define e envia o banco de dados individual utilizado pelo aluno. Ao escolher o banco de dados, o aluno deve sinalizar ao professor qual é o seu banco de dados. O banco de dados deve conter no mínimo 10 variáveis categóricas e no mínimo 10 variáveis numéricas. O banco de dados deve também conter no mínimo 50 observações.

O banco de dados deve estar completo, ou seja, não deve conter dados faltantes. Em seguida, o aluno deve simular dados faltantes, da seguinte forma, o aluno gera um novo banco de dados retirando 10 por cento dos dados do banco inicial. No próximo passo, o aluno utiliza os métodos da média e o método da regressão linear para estimar os valores faltantes e comparar essa simulação com os dados retirados, usando distância euclidiana. Entrega 21 de dezembro de 2023.

1 Etapa P_2

Use análise de componentes principais para o banco inicial reduzindo os dados numéricos para três fatores e Faça uma normalização no intervalo $[-1, 1]$ para o banco de dados inicial usando as variáveis numéricas. Entrega 13 de fevereiro de 2024.

2 Etapa P_3

Implemente no Python, um armazém de dados com cinco tabelas de dimensão e duas tabelas fatos, tudo a partir do banco de dados inicial. Na sequência, use o programa QLIK, carregando as tabelas feitas. Por fim, o aluno realize operações de OLAP dentro do QLIK. Entrega 19 de Março de 2024.

3 Etapa P_4

Nesta etapa, o aluno vai construir dois cubos de dados: um com materialização parcial e um com materialização total. O aluno deve realizar consultas nos dois cubos e verificar o tempo de resposta nos dois casos. Nessa etapa, o aluno realiza uma apresentação de 12 minutos no dia 23 de abril de 2024. Esta apresentação é referente as quatro etapas P_1, P_2, P_3 e P_4 . Por ventura, o professor pode adicionar tarefas complementares, ao longo do curso.