|  |  |  |
| --- | --- | --- |
| logo_CEFSA_H(grande) | **Código: EC E 10306**  **Disciplina: Banco de Dados - II** N2 | 2º bimestre| Curso: EC | Turma: EC 3Data da prova: \_\_/05/2023 – Horário: 19:15hs Prof.: Viotti | Coord.: Rodrigo Tadeu Fontes | **Logo Faculdade Eng** |
| Aluno(a):­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­ Nº: RA: | | Nota: |
| Aluno(a):­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­ Nº: RA: | |
| Aluno(a):­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­­ Nº: RA: | |
| Orientações: | |
|  | | Rubrica do aluno: |

Importação completa de Dados

A data de entrega será no início da semana de provas, e a apresentação será nos dias de aula que possuímos nesta semana de prova.

O trabalho de N2 consiste em buscar na Internet um banco de dados com aproximadamente 100 mil registros no mínimo. Este banco pode ser, por exemplo, para aprendizagem de máquina, dados do COVID, eleições, portal da transparência do governo. Segue algumas sugestões de repositórios:

<https://dados.gov.br/dataset?res_format=CSV&tags=covid>

<https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>

<https://www.portaltransparencia.gov.br/download-de-dados>

<https://www.gov.br/cgu/pt-br/acesso-a-informacao/dados-abertos/arquivos/ebt>

<https://dados.gov.br/dataset?res_format=csv>

<https://brasil.io/dataset/covid19/>

<https://dadosabertos.tse.jus.br/fa_IR/dataset/?groups=resultados&res_format=CSV>

[Find Open Datasets and Machine Learning Projects | Kaggle](https://www.kaggle.com/datasets)

[Portal de Dados Abertos](https://dados.gov.br/home)

O arquivo deve estar no formato txt ou csv. Você precisa cumprir algumas etapas para a entrega desta atividade:

O que fazer?

1. Normalizar o banco de dados. A partir do arquivo original, é preciso identificar pelo menos 6 tabelas que possam ser extraídas desses dados;
2. A partir do banco normalizado, será preciso fazer a transferência dos dados originais para o banco em questão, esta transferência deve:
   1. Ser feita através de comandos *t-sql* juntamente com as *“stored procedures”* para isso.
   2. Toda inserção de dados deve ser feita por SP de inserção de dados.
3. Caso os dados originais necessitem de algum ajuste, este tem que ser providenciado através de SP também;
4. Dentro dessas rotinas será preciso utilizar pelo menos 2 cursores e 2 funções.

O que entregar?

1. Todo o processo de importação total dos dados e execução das SPs deve ser registrado através de um vídeo que deverá conter no máximo 5 minutos. Este vídeo pode conter cortes, ou seja, se um processo levar 15 minutos para ser executado, grave somente o início e o final dele, corte o meio;
2. Em um documento Word, é preciso identificar:
   1. O local da origem dos dados com url
   2. A identificação e finalidade de cada uma das SPs
   3. Um link do youtube contendo o vídeo de importação
3. Um arquivo .sql que conterá o script de criação da estrutura do banco de dados, bem como a criação de todas as SPs. Em alguns casos o arquivo deverá ser compactado.
4. O DER do banco criado
5. Um arquivo contendo um script de criação do banco de dados já importado e finalizado. Este script deve ser capaz de recriar o banco finalizado (não é o processo de importação)

O que apresentar?

1. Exibir o vídeo de importação
2. Um aplicativo que se conecte com o banco de dados, pode ser em qualquer linguagem de programação, e exiba algumas informações do banco, pode ser um gráfico, um resumo, alguma constatação baseada nos dados, etc. Essas informações também podem ser geradas a partir de alguma SP que filtre/processe os dados do banco.
3. Um depoimento sobre as dificuldades encontradas, como elas foram solucionadas e o que você teve que buscar de conhecimento além do que foi ensinado na disciplina