



Universidade Federal do Ceará
Centro de Ciências
Departamento de Computação

CKP9011 e CK0268 – Introdução à Ciência de Dados 2023.1

Trabalho 1: Engenharia de Dados

1. Descrição:

Este trabalho consiste na construção de um dataset, a partir de uma ou mais fontes de dados (que podem ser estruturadas ou não).

2. Avaliação

Espera-se com a realização deste trabalho que cada estudante elabore e entregue (de forma digital) os seguintes documentos:

- Arquivo CSV contendo os dados do Dataset;
- Jupyter Notebook contendo os Scripts Python utilizados para obter os dados.
- Vídeo (disponibilizado no Youtube) apresentando e descrevendo as atividades desenvolvidas.

A avaliação deste trabalho se dará em três etapas:

1ª. Vídeo de Apresentação do Dataset: Cada estudante irá disponibilizar um vídeo (no Youtube) apresentando o dataset gerado e o processo utilizado. O estudante pode utilizar slides e notebooks.

2ª. Avaliação do Dataset: O professor da disciplina irá avaliar a qualidade do dataset gerado pelo estudante, bem como dos códigos implementados.

A avaliação do trabalho irá envolver os seguintes quesitos:

- Abrangência e Organização do Dataset
- Qualidade dos Códigos Utilizados
- Clareza do Texto (Tutorial) Utilizado para Descrever o Dataset e seu Processo de Construção
- Domínio do Tema

3. Data de Entrega: 15/05/2023

- PS. Não serão aceitos trabalhos que não forem apresentados (por meio de vídeo disponibilizado no Youtube).
- PS. Cada estudante será responsável pela disponibilização do ambiente (software e hardware) necessário para a gravação da apresentação do seu trabalho.
- Os Notebooks e Arquivos CSV deverão ser disponibilizados, em formato .ZIP, em uma das seguintes plataformas: SIGAA ou em um repositório público (GitHub ou GitLab).
- Utilize o formulário a seguir para indicar as URLs do vídeo e do arquivo ZIP contendo os códigos e os dados utilizados:

<https://forms.gle/HNQQWQFAxVka95JR7>

“A Educação, qualquer que seja ela, é sempre
uma teoria do conhecimento posta em prática”.

Paulo Freire