
ANÁLISE DE DADOS E BIG DATA

PROJETO SEMESTRAL

ANÁLISE DESCRITIVA DE DADOS COM PYTHON

MATERIAL:

Softwares:

Python

Módulos e Bibliotecas:

Pandas

Matplotlib

Numpy

Urllib

Etc.

OBJETIVOS DO PROJETO:

Este projeto visa estimular o pensamento crítico e a solução de problemas na área de sua formação utilizando conceitos de análise de dados e big data adquiridos na nossa unidade curricular.

OBJETIVO ESPECÍFICOS:

- Apresentar uma solução de algum problema na área de sua formação na forma de um protótipo e/ou simulação;

Elaborar um relatório final, seguindo as especificações do professor, apresentando, sobretudo, os detalhes para solucionar o problema levantado.

Obs: O output deve ser salvo em um arquivo csv, o arquivo deve ser nomeado com a seguinte formatação: `dia_mes_ano_A3_equipe_N.csv`, a data deve ser a data em que o programa foi executado!

REGRAS:

1. Os trabalhos serão em grupos de 3 a 5 alunos. Só será permitida a participação de alunos regularmente matriculados.
2. A pontuação total do projeto será de 40 pontos;
3. Os temas dos projetos escolhidos devem ser previamente aprovados pelos professores, lembrando que o tema, obrigatoriamente, deve ter uma aplicação direta na área da sua formação;
4. A solução do problema proposto pelo grupo deve necessariamente utilizar os conceitos da Análise de Dados e Big Data trabalhados ao longo do semestre;
5. A solução deve ser apresentada na forma de um protótipo e/ou simulação;
6. O resultado final deste projeto deve consistir em uma apresentação que contemple: o protótipo, a explicação do protótipo, o embasamento teórico e a discussão final com possíveis resultados obtidos;
7. A apresentação final deverá ser realizada ao vivo durante o horário de aula e, posteriormente, encaminhado para apresentação na EXPO das FACULDADES envolvidas;
8. Não serão permitidos projetos que exponham os integrantes a qualquer tipo de risco, o que incluem atividades que violem as normas de segurança individual e coletiva.

PONTUAÇÃO:

A Pontuação consistirá em:

A - Checkpoint do projeto - Os grupos deverão apresentar o plano de trabalho, os cálculos e pesquisas preliminares para a correção e ajuste dos procedimentos com os professores. A apresentação do checkpoint terá duração máxima de 5 minutos e será realizada em aula com o auxílio de slides.

B - Apresentação do Projeto - A apresentação final consistirá na apresentação do vídeo, durante o horário da aula, seguindo o formato e a duração estabelecidos nas regras.

C - Relatório final - O relatório final deverá ser entregue na plataforma a ser indicada pelo professor na data estabelecida no cronograma, seguindo a estrutura:

- I Introdução
- II Projeto
- III Resultados
- IV Conclusão

D - Apresentação na EXPO das FACULDADES envolvidas - a apresentação terá que ser feita com a presença de todos os membros do grupo, os alunos que não participarem da apresentação não receberão a pontuação atribuída ao grupo que compareceu!

CRITÉRIOS PARA AVALIAÇÃO DOS TRABALHOS (CHECKPOINT, APRESENTAÇÃO DO PROJETO, RELATÓRIO ESCRITO)

Critério	Pontuação	Descrição
Conteúdo	50 %	Avaliação dos conceitos, princípios e leis apresentados, com ênfase na aplicação dos conteúdos vistos nas UCs.
Linguagem	10 %	Correção gramatical, adequação vocabular, uso da norma culta.
Desenvolvimento	20 %	Avaliação da evolução do projeto e se a ideia está bem trabalhada, com explicação objetiva, compreensiva e clara. Avaliar se o protótipo e/ou a simulação estão funcionando adequadamente.
Participação	10 %	Participação do grupo.
Qualidade visual	10 %	Uso adequado da forma de comunicação, criatividade, fotografia e áudio.

CRONOGRAMA

Data	Ação	Responsabilidade
Sem 26/02	Formação dos grupos	Alunos
Sem 04/03	Escolha do tema	Alunos
Sem 11/03	Aprovação do tema	Professores
Sem 22/04	Entrega do <i>Checkpoint</i>	Alunos
Sem 13-18/06	Apresentação/entrega do projeto	Alunos
03/06 – 04/06 (A CONFIRMAR)	Apresentação na EXPO	Alunos
17/06	Entrega do relatório final	Alunos
13 a 20/06	Lançamento da nota do projeto	Professores

ESCOPO DO PROJETO:

Realizar webScraping no site do Mercado Livre, acessar o carrossel de ofertas e coletar todos os produtos que estiverem com oferta acima de 15% de desconto;

- Organizar o DataSet em ordem alfabética;
- Gerar um CSV com os dados obtidos;
- Exportar para o Excel
- Calcular desvio padrão, media, mediana e medidas de dispersão e distribuição, e apresentar num boxplot

Será necessária descrever um contexto das informações que se pretende obter na as informações do DataSet.

As equipes deverão apresentar as questões que se deseja responder

As respostas deverão ser apresentadas na formas de gráficos, dashboards, tabelas e medições estatísticas/probabilísticas com as interpretações e diagnóstico

correspondentes

A equipe também deverá anexar os scripts Python implementado no trabalho, comentando o código, os módulos e bibliotecas usados.

A apresentação será no dia e cada equipe terá 15 minutos para apresentação, no relatório deve constar:

Formato dos registros do DataSet

Números dos registros

Amostra de alguns registros utilizados

Não será permitida a utilização de APIs comerciais

Data da Apresentação: 17/06/2024