



Estácio

TÓPICOS DE BIG DATA EM PYTHON

Código disciplina: ARA0168

Cópia e Distribuição não autorizada.

2025.2





REQUISITOS PROJETO

1. Tamanho mínimo do Dataset: 700MB
2. Usar: HADOOP (**HDFS**) ou PySpark (com **RDD** ou **SQL**) ou HADOOP + PYSPARK, PANDAS e ANALYTICS
3. Preparação dos DADOS (limpeza, refinamento, padronização,...)
4. Possuir pelo menos uma implementação de analytics (KDD, DL, ML)
5. Apresentar a análise de dados
6. Apresentar o projeto e os resultados
7. Autoavaliação
8. Escrever o relatório de extensão (individual) e postar no SAVA. (**Prazo final: dia da mostra de extensão**)
9. Participar da Mostra de Extensão (dia a ser divulgado pela ESTÁCIO)

OBS: Obrigatório uso da linguagem Python.





CRITÉRIOS AVALIATIVOS

Tudo do slide anterior, +

1. Qualidade da apresentação
2. Envolvimento/colaboração da equipe
3. Gestão do tempo
4. Cumprimento dos Prazos
5. Domínio
6. Respostas as perguntas (individual e/ou equipe)

