





FOTOGRAFIA DO MÓDULO

AULA 1
DIA 1

AULA 2

DIA 2

AULA 3

DIA 3

AULA 4

DIA 4

AULA 5

DIA 5

AULA 6

Projeto revelo:coleta de dados com webscrapping

Os dados nem sempre estão disponíveis de forma estruturada. Muitas vezes precisamos usar técnicas como webscrepping para coletar dados da web e trazer para um formato estruturado. É isso que faremos com dados de vagas de trabalho, sobre os quais iremos tratar e responder perguntas como "Qual a média salarial por região? Quais as habilidades solicitadas em cada região? Existem discrepâncias?"







#ANÁLISE DE VAGAS



Sobre o projeto:

A Revelo é uma referência em recrutamento de profissionais na área de tecnologia. Para continuar sendo destaque na área, ela precisa estar atenta aos movimentos do mercado de vagas e ferramentas. Aqui, você irá atuar em time, criar robôs em web scrapping para ler esses dados de outros sites de vagas e criar insights para gerar apontamentos para onde a empresa precisa caminhar.



#DATASET PREFEITURA DE RECIFE



Sugestão:

- Selecionar um site de vagas para mapear.
- Definir as cidades que se deseja mapear.
- Realizar a coleta de dados.
- Responder às perguntas do backlog.



BACKLOG'S



- Qual a região que mais concentra vagas abertas de analista de dados?
- Qual a média salarial por região?
- 3. Quais as habilidades citadas em cada região? Existem discrepâncias?
- 4. As vagas são distribuídas de que forma para júnior, pleno e sênior?
- 5. Quais as empresas / agências que mais contratam?
- 6. Como é a distribuição entre vagas home office, híbrida e presencial?





- Mapeamento dos sites de vagas de interesse;
- Entender padrão html para fazer scrapping;
- Geração de módulos e funções para extração de informações de vagas;











- Definir local para armazenamento (csv ou banco de dados);
- Definir modelagem de dados;
- Geração de módulos e funções para manipulação e extração de informações;
- Criar a primeira versão no repositório GitHub









Revelo

- Validar coleta de dados;
- Validar escala prevendo paginação / rolagem em webscrapping;
- Criar lista de ferramentas e habilidades para buscar nas vagas;
- Criar lista de sinônimos para essas palavras;











- Geração de módulos e funções para extração de informações de vagas;
- Avaliação inicial do conjunto (verificação de colunas e dos tipos das variáveis);
- Limpeza do conjunto (exclusão/substituição ou avaliação de dados nulos e de dados duplicados);









Revelo

- Geração e finalização de visualizações dos resultados;
- Organização de apresentação dos dados (Storytelling);









MOMENTO INDICAÇÃO LINKEDIN

Reserve 20 minutos da aula para fazer indicações para outros alunos.





Revelo

Responder perguntas do backlog;





