Teste Prático de Data Science

Olá candidato!

A base de dados anexa em CSV contém informações de municípios. Ela simula um caso prático do tipo de trabalho que realizamos em conjunto com a equipe de negócios.

Neste projeto, o time de negócios deseja **comparar munícipios para estabelecer metas de redução de óbito e acidentes em estradas**, para o ano de 2017. Para isto, a equipe de Data Science deve aplicar um **algoritmo de análise de agrupamentos** (Clusters) para obter grupos, contendo municípios comparáveis.

O objetivo do teste é:

* Analisar as variáveis (features) selecionadas pelo time de negócios para definir os agrupamentos.
* Listar variáveis relevantes & irrelevantes, determinando o subconjunto adequado para aplicar o algoritmo de análise de agrupamentos (clusters).
* Selecionar e ajustar um modelo de análise de agrupamentos.
* Validar o ajuste do modelo utilizando métricas de qualidade.
* Apresentar os resultados para a equipe de negócios.

O candidato terá **uma semana, a partir do recebimento dos dados**, para realizar a análise e compilar uma apresentação. A análise será apresentada para alguns membros da equipe da MAiS, simulando uma situação real de projeto.

A apresentação deve durar entre 30-45 min, e abordar os seguintes temas:

* Resumo dos resultados: sumário executivo descrevendo os resultados da análise de agrupamento.
* Metodologia: quais foram as técnicas escolhidas para selecionar as variáveis, definir e validar os agrupamentos encontrados, além de descrever as ferramentas utilizadas.
* Resultados: detalhamento da aplicação da metodologia nos dados, bem como o “passo-a-passo” executado.
* Próximos passos & melhorias: visão de como aprimorar o resultado e quais seriam as alternativas existentes para evoluir a análise.

A linguagem de programação utilizada fica a critério do candidato.