

Prova processo seletivo para a vaga de “Cientista de Dados” na REDASSET

Leia com atenção:

- Essa prova é confidencial.
- Essa prova técnica faz parte do processo seletivo para a vaga de “Cientista de Dados” na empresa REDASSET, desenvolvida internamente, que consiste das seguintes fases:
 - Prova técnica
 - Entrevista sobre a prova técnica e pessoal
- O tempo para resolução da prova é **5 dias** a contar da data e hora do envio dessa prova.
- Todas as questões têm pesos iguais
- Junto com esse arquivo você deve ter recebido o arquivo **BASE_TITULOS.txt**
 - Ele é necessário para resolução de questões, no entanto, em caso de não recebimento avisar imediatamente.
 - A base de dados está no formato txt com separador “tab” e os números estão com decimais “.”
 - **As informações constantes no arquivo são fictícias, e foram geradas artificialmente, ou seja, não correspondem de forma alguma às informações dos nossos clientes, as quais são informações confidenciais**
- O candidato poderá utilizar o software que preferir para a resolução da questão.
- Favor enviar os códigos utilizados como resposta.
- Favor explicar o raciocínio nas resoluções das questões.

Breve relato sobre a empresa

Desde 1992, nossos profissionais têm se dedicado a operações de crédito e fomento a pequenas, médias e grandes empresas, entregando resultados consistentes a nossos clientes e acionistas.

O grupo RED é formado essencialmente por duas empresas:

A “RED Serviços”, que nasceu como uma sociedade de fomento mercantil há mais de vinte anos e na última década se especializou na prestação de serviços de análise, consultoria e cobrança de créditos para Fundos de Investimento em Direitos Creditórios (“FIDCs”).

A “REDASSET Gestão” que foi fundada em 2010 por profissionais com vasta experiência em operações de crédito e/ou vindos de grandes bancos de investimentos e se habilitou junto à Comissão de Valores Mobiliários para a atividade de administração de carteiras de títulos e valores mobiliários de terceiros, especialmente de FIDCs. A sociedade gestora atua com equipe própria e total independência na tomada das decisões de investimento.

Nosso foco é na manutenção de um portfólio bastante pulverizado, no que tange aos cedentes e/ou devedores, aliado a rígidos controles internos e de TI, sendo nossa principal preocupação a preservação de capital dos cotistas e o alcance de uma rentabilidade acima da média no longo prazo.

O grupo RED tem hoje como uma de suas missões fazer dos FIDCs que estruturou, e para os quais presta serviços, um importante veículo de desintermediação bancária, permitindo às empresas clientes/cedentes um acesso facilitado à antecipação de recebíveis e consequentemente ao crédito, proporcionando, do outro lado, a possibilidade de ganhos diferenciados para os investidores/cotistas dos FIDCs.

A Antecipação de recebíveis é o mecanismo no qual empresas (Cedentes) conseguem adiantar o fluxo de vendas a prazo, duplicatas, vendas parceladas no cartão de crédito, boletos, cheques pré-datados, contratos de prestação de serviços entre outros. Esse adiantamento se dá por meio de uma instituição financeira que “compra” o recebível do cedente aplicando uma taxa de juros de deságio a de acordo com o prazo do recebível.

Exercícios:

Uma empresa atacadista (Cedente) possui 1 cheque (Título) no valor de R\$10.000 de um mercado que é seu cliente (Sacado) com vencimento para daqui a 1 mês, mas o cedente necessita de dinheiro hoje para pagar uma dívida. O Cedente procura a RED para adiantar um cheque, a taxa cobrada é de 5% a.m (juros compostos), ou seja, o Cedente vai receber hoje 95% do valor do cheque (R\$9.500) , os R\$ 500 são chamados de valor de deságio e a RED vai receber os de R\$10.000 no final no vencimento do título.

1. Com base nesse mecanismo, e na carteira de títulos que ela adquiriu (arquivo **BASE_TITULOS.txt** enviado junto), gostaríamos que você fizesse uma análise descritiva dos ativos como se fosse realizar uma apresentação para um executivo demonstrando a situação da carteira atual (Regionalização, concentrações de clientes, taxas de pagamento, Receitas, etc.). Nesse momento não é necessário utilizar as variáveis `RISCO_XX` (não é necessário a construção de uma apresentação, somente a apresentação dos gráficos, tabelas, resultados etc. e os comentários sobre os resultados).
2. Com base no mesmo arquivo calcule indicadores por cedentes que achar necessários e utilize técnicas de algoritmos não supervisionados para criar uma segmentação de Cedentes com base em perfis parecidos, de acordo com os indicadores calculados. Detalhe a escolha da metodologia e os indicadores criados para realizar a segmentação. Nesse momento não é necessário utilizar as variáveis `RISCO_XX`.

Nessa mesma base, possuímos os resultados se os títulos foram pagos em dia ou em atraso (`STATUS_LIQ_TITULO`) e algumas outras informações de risco (`RISCO_XX`) sobre os títulos no momento de sua emissão.

3. Com base nesse histórico de títulos utilize algoritmos supervisionados, com a variável resposta `STATUS_LIQ_TITULO`. Caso ache necessário pode-se criar novas variáveis, segmentações, etc. para se encontrar a melhor metodologia possível. Como resposta da questão detalhar a escolha da metodologia, as análises adicionais realizadas e as suas métricas de avaliação dos modelos.

Descritivo das variáveis:

Nome variável	Descrição da variável
CODIGO	Código do título
NOME_FUNDO	Nome do fundo em que o título está registrado
TIPO_PRODUTO	Tipo de produto do título
CEDENTE	Nome do Cedente
SACADO	Nome do sacado
LIMITE_CEDENTE	Limite de Crédito do Cedente
ESTADO_CEDENTE	Estado de cadastro do Cedente
ESTADO_SACADO	Estado de cadastro do sacado
VENCIMENTO	Data de vencimento do título
DATA_EMISSAO	Data de emissão do título
DATA_PAGAMENTO	Data de vencimento de pagamento do título
VALOR_REAL	Valor do título
VALOR_DESAGIO	valor de deságio do título
TAXA_MENSAL	Taxa mensal (30 dias) que foi utilizada para o desconto do título
STATUS_LIQ_TITULO	Status de liquidação do título (EM_DIA: O Título foi pago em dia, ATRASO o título foi pago em atraso)
RISCO_1	Variáveis que indicam risco
RISCO_2	Variáveis que indicam risco
RISCO_3	Variáveis que indicam risco
RISCO_4	Variáveis que indicam risco
RISCO_5	Variáveis que indicam risco
RISCO_6	Variáveis que indicam risco
RISCO_7	Variáveis que indicam risco
RISCO_8	Variáveis que indicam risco
RISCO_9	Variáveis que indicam risco
RISCO_10	Variáveis que indicam risco
RISCO_11	Variáveis que indicam risco
RISCO_12	Variáveis que indicam risco
RISCO_13	Variáveis que indicam risco
RISCO_14	Variáveis que indicam risco
RISCO_15	Variáveis que indicam risco
RISCO_16	Variáveis que indicam risco
RISCO_17	Variáveis que indicam risco
RISCO_18	Variáveis que indicam risco
RISCO_19	Variáveis que indicam risco
RISCO_20	Variáveis que indicam risco
RISCO_21	Variáveis que indicam risco

Conceitos importantes:

Conceitos	Descrição
-----------	-----------

Cedente	Cedente é o cliente da REDASSET e o titular dos recebíveis, que são cedidos onerosamente à REDASSET
Sacado	Sacado é quem irá pagar o título cedido.
Limite de Crédito	Limite de crédito do cedente é o valor máximo de recebíveis que o cedente pode negociar com a REDASSET
Data de emissão do título	Data em que o título foi cedido para a REDASSET

Dicas:

- Utilize alguma ferramenta de “notebook” para realizar as análises (Rmarkdown, Jupyter, etc.) isso facilita o envio dos códigos e as explicações em um mesmo lugar.
- Não se prenda ao resultado final encontrado, apresente o raciocínio utilizado para a escolha da técnica(s), em provas desse tipo não se tem tempo de testar todas as técnicas possíveis para se realmente encontrar o melhor resultado.
- Envie como resposta todos os testes realizados, desde que explique o raciocínio por trás deles.