**Estudo de Caso 1**:

O IFRN, enfrenta nos atuais dias, um problema muito sério, que vem se agravando, ao longo dos anos, que são os relacionados aos altos índices de evasão escolar e reprovações. Portanto, é de interesse da comunidade do IFRN, como um todo, analisar, as suas bases de dados, para tentar descobrir, nessa massa de dados, algum comportamento que venha indicar as causas para esses problemas.

Para tentar identificar esse comportamento nas bases de dados, serão aplicadas algumas técnicas de mineração de dados, tais como: **Árvore de Decisão**, **Naive Bayes**, **Clustering**, **Rede Neural**, etc.

Os modelos de Rede Neural são usados com frequência para exploração, porque podem analisar relações complexas entre muitas entradas e saídas.

Como parte desse estudo de Caso, tentarei responder as seguintes perguntas:

* Quais são os fatores que afetam a evasão/reprovação/desistência?
* O que pode-se fazer para melhorar/ amenizar esses problemas?

No Analysis Services, um modelo de regressão logística se baseia no algoritmo de redes neurais e, portanto, fornece a mesma flexibilidade e capacidade que um modelo de rede neural. No entanto, a regressão logística é particularmente bem-apropriada para prever resultados binários. Para esse estudo de caso, vai ser usada a mesma estrutura de mineração usada no modelo de rede neural. No entanto, o modelo vai ser personalizado para direcionar as questões solicitadas no estudo de caso. Ou seja, estou interessado em melhorar os problemas da evasão/reprovação escolar.

Regressão Logística é um método estatístico que é geralmente usado para fazer previsões com base em alterações em variáveis independentes: é usado em contagem financeira, para prever comportamento de clientes com base em demografia de clientes.

Para assegurar que todos os modelos baseados nos dados para os modelos sejam os mais semelhantes possível, vou usar a mesma semente para todos os modelos. Definir o parâmetro de semente assegura que o modelo processa os dados do mesmo ponto de partida, e minimiza variações causadas por artefatos nos dados.

As previsões ***singleton*** constituem um bom método para testar valores diferentes para visualizar o efeito no modelo.