**Metodologia**

1. Leitura dos Dados

* Colunas que existem em todas as planilhas: Escola, Turma, Nome e Questões
  + Questões: 1 a 50 para Vocabulário
  + Questões: 1 a 40 para TDE
* **Prova de Vocabulário:**
  + Foram aplicadas as mesmas provas nas 3 fases
  + **Palavras Trabalhadas no teste**: enorme | diretriz | década | isolar-se | depreciativos | abandonou | exibição | dramática | suspensos | libere | proíbe | revelou | ambígua | reconheceu | monitorar | substituir | conflito | projetam | sistema | calúnias | promover | reage | contribuiu | suficiente | transferido | impacto | cultura | Interpretei | analisar | status | indenização | recuperar | demonstrou | ética | divulgou | constante | desencadeou | distorcidas | alterou | mentais | confiar | ajuste | transição | conotação | fundamentais | integrar | fatores | função | incorporei | estrutura
  + **Palavras Ensinadas (Intervenção) em cada fase**
    - **Fase 2**: convencer | enorme | integridade | persistente | relutante | diretriz | maduro | ambíguo | devido | status | ajuste | exposição | modificar | monitorar | transição | abandono | frívolo | contemporâneo | dramático | explorar | acesso | civil | apesar de | integrar | promover | analisar | fator | função | interpretar| estrutura
    - **Fase 3:**controvérsia | justificar | perspectiva | viés | debate | citar | pesquisa | dados | fase | estratégia | exportar | simbólico | doméstico | integrar | eficiente | vocacional | inerentemente | exceder | equivalente | foco | surgir | explorar | além disso | enfrentar | interferir | conceber | antiético | benefício | detectar | racionalizar | submeter-se a | fortalecer | implicações | negar | modelo | produzir | participar | superiores | resiliente | consenso
    - **Fase 4:** estender | radical | crucial | alcançar | iniciativa | posar | contactar | primário | mínimo | não monitorado | dispersar | catalisar | segregação | reforma | injustiça | adotar | garantir | tensão | consentir | duração | anônimo | subjacente | capacidade | adequadamente | assédio | selecionar | herdar | mutação | genética | característica | privacidade | invasão | consistir | seguro | corporal | receita | instituir | incidente | acompanhar | invocar
      * Interseção com Palavras Trabalhadas: VAZIO
* Prova de TDE
  + **Grupo A**
    - As palavras trabalhadas são: boi, dedo, luva, sapo, gaveta, placa, abacaxi, ervilha, antiga, zebu, caminhar, argila, quebrado, cabeça, esquilo, jardim, alfabeto, chuteira, alvo, cinema, chinelo, completar, pessoa, luxo, samba, solução, barriga, secreto, guia, benzer, enrolado, amanhã, existir, aumento, explicar, cigana, descer, horta, sossego e jipe.
  + **Grupo B**
    - As palavras trabalhadas são: regeneração, explorar, certeza, composição, alguém, fixar, rigidez, auxiliar, serviço, acesa, embaixatriz, humanizar, oxigênio, adolescente, enxergar, confecção, chinês, recém, flexionar, sensibilização, juízo, sucção, majestoso, análise, planície, inteligentíssimo, essência, têxtil, persuasão, excelência, insensatez, obsessão, consciência, abdômen, oscilação, excepcional, ímã, miscigenação, ressuscitar e exceção.

2. Pre processamento

Pré-processamento Completo (8 etapas):

2.1: Completar dados faltantes de Escola/Turma

2.2: Remover duplicados (Escola + Turma + Nome)

2.3: Converter valores TDE (0=erro, 1=acerto parcial, 2 = Acerto Completo)

2.4: Verificar questões válidas (≥25%)

2.5: Padronizar nomes de colunas (ESCOLA→Escola, etc.)

2.6: Classificar grupos (A: 6º/7º, B: 8º/9º): Somente para efeito de analise

2.7: Criar IDs únicos

2.8: Verificar presença dos alunos em ambos os teste

O final de do processo é que numero de registros no pré e pós sejam iguais

3. Sobre Analise

* Tava estudando os dados e lendo nossas conversas:
  + A visualização de coordenadas paralelas (Progresso) não está acompanhando a mesma turma, mas sim o "espaço" que aquela turma ocupa na escola. Ex: A "Turma 6º Ano A" da Fase 2 é composta por alunos diferentes da "Turma 6º Ano A" da Fase 3
* Separei duas perguntas que podemos fazer sobre as turmas:
  + - **Análise de Coorte (Acompanhamento dos mesmos alunos):** *Como o grupo de alunos que estava no 6º Ano A na Fase 2 evoluiu quando se tornaram o 7º Ano A na Fase 3 e o 8º Ano A na Fase 4?*
    - **Análise da Série/Ano:** "A performance do 6º Ano A deste ano (Fase 4) foi melhor ou pior que a performance do 6º Ano A do ano passado (Fase 3)?"
  + De acordo com nossas conversas fazia **mais sentido Análise de Coorte**: implementar o rastreamento de coorte de alunos para tornar a evolução significativa para analise do progresso. Com isso permitirá acompanhar os mesmos grupos de alunos através das fases, independentemente de mudanças nos nomes das turmas**.**
* **Problema**: Você identificou uma inconsistência crítica nos dados. O problema é que o ID\_Unico atual está sendo gerado baseado em Turma + Fase + Escola + Nome, o que cria IDs diferentes para o mesmo aluno em fases diferentes, impossibilitando análises longitudinais.
  + **Solução Implementada:**
    - ‘IDs únicos permanentes baseados em Nome + Escola (hash MD5)
    - Preservação da origem com campos Turma\_Origem e Fase\_Origem
    - Rastreamento atual com Turma\_Atual por fase
    - Validação completa da consistência dos dados