

# Canoa

## Relatório de Validação de Dados

### Informações

**Data e hora do processo:** 2024-05-21 17:16:42

**Versão do validador:** 0.1.0

### Resumo da validação

**Número de Erros:** 17

**Número de Avisos:** 4

### Erros

**Issue #39: Estrutura dos arquivos da pasta de entrada:**

composicao.xlsx: Coluna número 3 não possui nome mas possui valores.

valores.xlsx: Coluna número 20 não possui nome mas possui valores.

**Issue #79: Limpeza dos arquivos:**

descricao.xlsx, linha 3: A coluna 'codigo' contém um valor inválido: O valor '5000.77' não é um número inteiro.

descricao.xlsx, linha 4: A coluna 'codigo' contém um valor inválido: O valor 'XX5001' não é um número.

descricao.xlsx, linha 2: A coluna 'nivel' contém um valor inválido: O valor '-1' é menor que 1.

descricao.xlsx, linha 3: A coluna 'nivel' contém um valor inválido: O valor '2.77' não é um número inteiro.

descricao.xlsx, linha 4: A coluna 'nivel' contém um valor inválido: O valor 'XX3' não é um número.

composicao.xlsx, linha 5: A coluna 'codigo\_pai' contém um valor inválido: O valor '5000.7875' não é um número inteiro.

composicao.xlsx, linha 7: A coluna 'codigo\_pai' contém um valor inválido: O valor

'5001.88' não é um número inteiro.

Issue #2 e #59: Relações entre indicadores:

descricao.xlsx: Indicadores no arquivo composicao.xlsx que não estão descritos: [2, 5000, 5001].

composicao.xlsx: Indicadores no arquivo descricao.xlsx que não fazem parte da estrutura hierárquica: [5002].

composicao.xlsx: Indicadores desconectados encontrados: [5004 -> 5006, 5004 -> 5007].

valores.xlsx: Códigos inválidos: ['5000.954', '5001,9483', 'Unnamed: 19'].

valores.xlsx: Códigos dos indicadores ausentes em descricao.xlsx: [5000].

Issue #3: Hierarquia como árvore:

Issue #37: Níveis de indicadores:

descricao.xlsx, linha 2: Nível do indicador não é um número inteiro maior que 0.

descricao.xlsx, linha 3: Nível do indicador não é um número inteiro maior que 0.

descricao.xlsx, linha 4: Nível do indicador não é um número inteiro maior que 0.

Issue #8: Unicidade dos códigos:

Issue #5: Códigos HTML nas descrições simples:

Issue #24: Ortografia:

Issue #36: Títulos únicos:

Issue #75: Campos vazios:

Issue #1: Padrão para nomes dos indicadores:

Issue #39: Títulos com mais de 40 caracteres:

Issue #32: Pontuações obrigatórias e proibidas em descrições:

Issue #81: Pontuações obrigatórias e proibidas em cenários:

Issue #81: Pontuações obrigatórias e proibidas em referência temporal:

Issue #81: Relações de valores únicos em cenários:

Issue #81: Relações de valores únicos em referência temporal:

Issue #81: Relações de combinações de valores:

Issue #85: Quebra de linha para descrição:

Issue #85: Quebra de linha para cenários:

Issue #85: Quebra de linha para referência temporal:

## Avisos

Issue #39: Estrutura dos arquivos da pasta de entrada:

composicao.xlsx: Coluna 'Unnamed: 2' será ignorada pois não está na especificação.

Issue #24: Ortografia:

descricao.xlsx, linha 3: Palavras com possíveis erros ortográficos na coluna nome\_simples: ['ries'].

Issue #1: Padrão para nomes dos indicadores:

descricao.xlsx, linha 3: Nome simples fora do padrão. Esperado: "IDEB séries finais". Encontrado: "IDEB séries Finais".

descricao.xlsx, linha 4: Nome simples fora do padrão. Esperado: "Vulnerabilidade do IDEB". Encontrado: "Vulnerabilidade Do IDEB".