

# Canoa

## Relatório de Validação de Dados

### Informações

**Data e hora do processo:** 2024-07-05 12:54:28

**Versão do validador:** 0.3.0

### Resumo da validação

**Número de Erros:** 30

**Número de Avisos:** 8

### Erros

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

composicao.xlsx, linha 7: A linha possui um valor na coluna 3, mas a tabela possui apenas 2 colunas.

valores.xlsx, linha 7: A linha possui um valor na coluna 20, mas a tabela possui apenas 19 colunas.

**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 2: A coluna 'codigo\_pai' contém um valor inválido: O valor '0.0' é

menor que 1.

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 #124: Códigos sequenciais:

descricao.xlsx: Verificação de códigos sequenciais foi abortada porque a coluna 'codigo' contém valores inválidos.

#### Issue #75: Campos vazios:

Issue #1: Padrão para nomes dos indicadores:

Issue #39: Títulos com mais de 40 caracteres:

Issue #119: Descrições simples com mais de 150 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 #149: Valores indisponíveis:

Issue #85: Quebra de linha para descrição:

Issue #85: Quebra de linha para cenários:

Issue #85: Quebra de linha para referência temporal:

#### Issue #16: Intervalo dos dados da legenda:

valores.xlsx, linha 17: O valor 0.789912176738247 está fora do intervalo de 0.0 a 0.77 para a coluna '5000-2030-P'.

valores.xlsx, linha 6: O valor 0.779055534730612 está fora do intervalo de 0.0 a 0.77 para a coluna '5000-2050-P'.

valores.xlsx, linha 15: O valor 0.846897288840176 está fora do intervalo de 0.0 a 0.77 para a coluna '5005-2015'.

#### Issue #71: Sobreposição de valores na legenda:

legenda.qml: Arquivo está corrompido. Existe uma sobreposição nos valores das fatias da legenda.

legenda.qml: Arquivo está corrompido. Fatias não estão descritas na ordem crescente.

#### Issue #69: Propriedades de soma nos fatores influenciadores:

proporcionalidades.xlsx, linha 3: A soma dos valores para o indicador pai 5008-2010 é 1.548, e não 1.

proporcionalidades.xlsx, linha 3: A soma dos valores para o indicador pai 5010-2010 é 0.462, e não 1.

proporcionalidades.xlsx, linha 5: O valor não é um número válido e nem DI (Dado Indisponível) para o indicador pai '5010-2010' e indicador filho '5030-2010'.

proporcionalidades.xlsx, linha 5: A soma dos valores para o indicador pai 5010-2010 é 0.902, e não 1.

proporcionalidades.xlsx, linha 6: O valor não é um número válido e nem DI (Dado Indisponível) para o indicador pai '5010-2010' e indicador filho '5030-2010'.

proporcionalidades.xlsx, linha 6: A soma dos valores para o indicador pai 5010-2010 é 0.902, e não 1.

## 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".

#### Issue #119: Descrições simples com mais de 150 caracteres:

descricao.xlsx, linha 9: Descrição simples fora do padrão. Esperado: Até 150 caracteres.  
Encontrado: 184 caracteres.

#### Issue #69: Propriedades de soma nos fatores influenciadores:

proporcionalidades.xlsx, linha 4: A soma dos valores para o indicador pai 5008-2010 é 0.999, e não 1.

proporcionalidades.xlsx, linha 5: A soma dos valores para o indicador pai 5008-2010 é 1.001, e não 1.

proporcionalidades.xlsx, linha 4: Existem valores com mais de 3 casas decimais na planilha, serão consideradas apenas as 3 primeiras casas decimais.