
MANUAL : FERRAMENTA DE AUXILIO A MANUTENÇÃO - COMPARADOR BANCO DE DADOS

RAFAEL HENRIQUE DA ROSA

Orientador: Eng. Mauricio Menon

Divisão de Manutenção Eletrônica (SMIN.DT)

Itaipu Binacional

Foz do Iguaçu , 2022

1 INTRODUÇÃO

O presente documento tem por objetivo apresentar a ferramenta desenvolvida para realizar comparações entre banco de dados mostrando discrepâncias, adições e exclusões após uma alteração no arquivo. Este manual visa facilitar a utilização da ferramenta exibindo suas funcionalidades e esclarecendo dúvidas frequentes.

1.1 O ALGORITMO

O algoritmo é elaborado em Python devido a sua capacidade de realizar desde tarefas simples até aplicações científicas com análise de dados, além de possuir programação orientada a objetos e sintaxe de código de fácil entendimento, funcionando nas mais variadas plataformas (CRUZ, 2018). Para a ferramenta foi adicionados à linguagem pacotes que auxiliam no processamento de dados e na elaboração de uma interface gráfica, sendo os principais brevemente comentados abaixo.

O pacote principal utilizado na análise de dados com Python é o Pandas, que possui um grande suporte a conjuntos que possuem como entradas números e texto, assim como integração com vários formatos de arquivos como csv, excel, sql, json, entre outros. Os dados tabulados são armazenados em variáveis bidimensionais semelhantes a matrizes chamados DataFrames, quem contem índices para linhas e colunas e dispõem de diversos métodos de manipulação que evitam a iteração entre os elementos da tabela facilitando a programação e aumentando a performance (MCKINNEY, 2022).

Com o intuito de criar a interface gráfica (GUI) é aplicado na ferramenta o pacote Tkinter, que é uma adaptação para Python da biblioteca gráfica Tlc/Tk, essa por sua vez está disponível para diversas linguagens de programação e é instalado por padrão no sistema operacional Windows e na maioria de distribuições baseadas em GNU/Linux (ALVAREZ, 2016). O pacote é recomendado para aplicações simples com menus, botões e caixas de texto e é ideal para o caso em questão. Ainda como complemento ao Tkinter é adicionado a biblioteca pandastable, que tem como função exibir DataFrames na interface de maneira semelhante a planilhas convencionais.

Os bancos de dados são importados para um DataFrame diretamente utilizando um programa externo gratuito e de código aberto chamado mdb-tools, composto por um conjunto de utilitários principalmente para sistemas Linux e Mac e nesse caso é responsável pela extração dos valores do banco de dados, possibilitando a utilização da aplicação em estações de trabalho que não possuam o pacote office.

2 A INTERFACE

Ao executar o ferramenta é exibida a interface conforme a Figura 1, nela são mostradas diretamente todas as informações obtidas através da comparação. A tela inicial apresenta três tabelas de visualização semelhantes a planilhas do *Microsoft Excel*. A planilha mais acima em amarelo apresenta as linhas discrepantes, ou seja, que tiveram um ou mais valores alterados de um arquivo para outro, A primeira coluna desta tabela é reservada ao nome do arquivo e a apresentação é feita alternando linhas equivalentes de ambos os bancos de dados.

A tabela ao centro da tela principal em vermelho mostra as linhas adicionadas que foram encontradas somente no banco pós atualização. Por fim a tabela com fundo verde

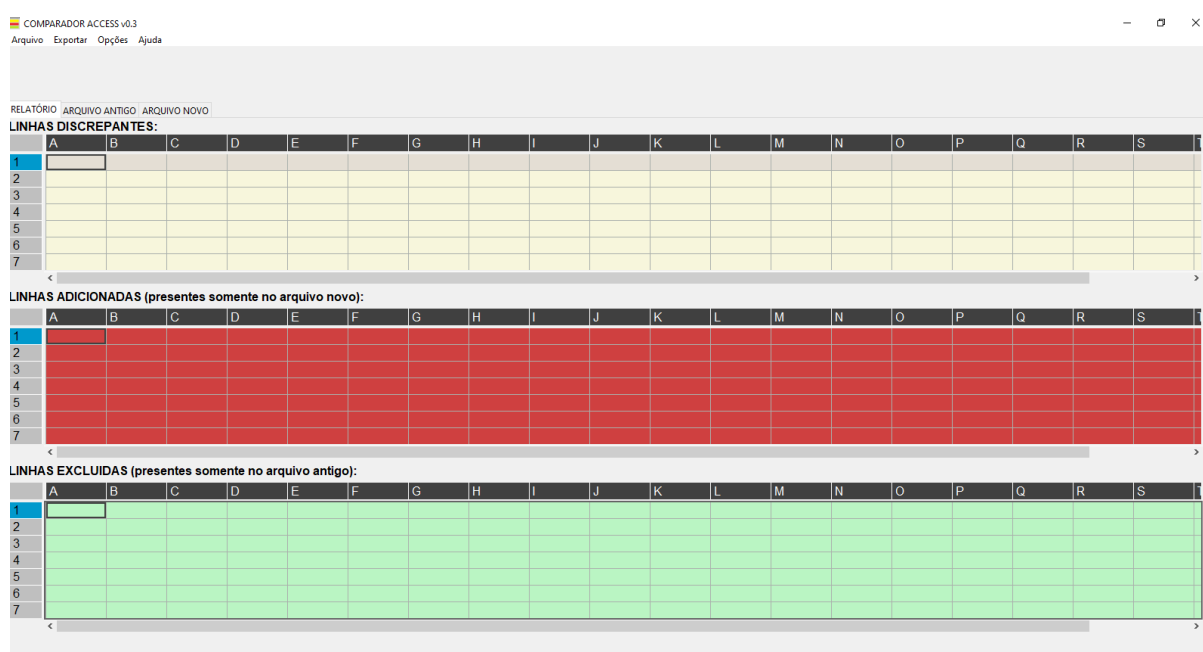


Figura 1: Interface da Ferramenta.

posicionada na parte inferior é dedicada as linhas que foram excluídas, sendo encontradas pelo algoritmo somente no banco pré atualização.

A Figura 2 mostra em detalhes a área superior esquerda da interface. Existem três abas onde é exibido o relatório, a tabela selecionada no arquivo pré atualização (chamada de ARQUIVO ANTIGO) e no arquivo pós atualização (chamada de ARQUIVO NOVO). Através dos menus é possível importar e exportar arquivos, além de mudar opções da ferramenta como aplicar filtros para pesquisar entradas específicas e colorir as diferenças entre linhas

no relatório e nas tabelas originais. A utilização dos menus será exemplificada nas seções a seguir, assim como um passo a passo de como é configurada uma comparação.

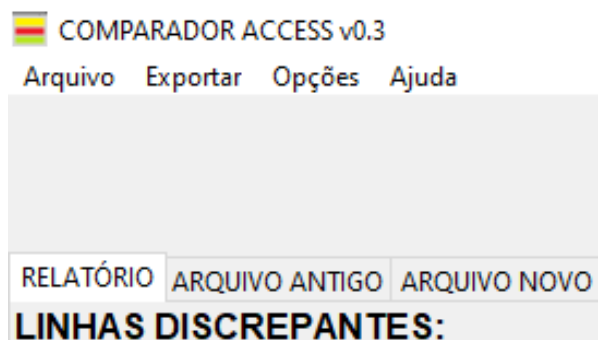


Figura 2: Menus e Abas.

3 IMPORTANDO ARQUIVOS

No menu Arquivo conforme a Figura 3 são exibidas as opções de importação. A primeira opção importa arquivos do *Microsoft Access* com a extensão .accdb, enquanto a segunda importa planilhas do *Excel* com a extensão .xlsx.

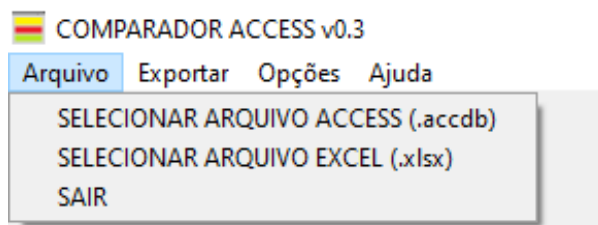


Figura 3: Menu 'Arquivo'.

Ao clicar em uma das opções de seleção é aberta uma janela para escolha do arquivo pré atualização e em sequencia uma segunda para o arquivo atualizado. Caso sejam escolhidos arquivos não compatíveis uma mensagem de erro será exibida.

NOTA: Devido a utilização de um executável externo para importações diretamente do *Access*, é possível que sejam necessárias autorizações do usuário durante a comparação através de notificações de antivírus de acordo com as configurações do computador.

4 SELECIONANDO A TABELA

Após selecionar os arquivos é exibida uma caixa de seleção conforme a Figura 4, nela são listadas as tabelas presentes nos arquivos e deve ser selecionada a que se deseja comparar.

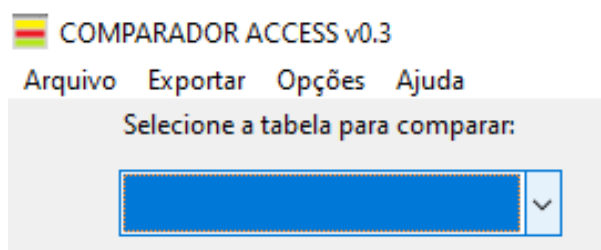


Figura 4: Seleção de Tabelas.

NOTA: Após a seleção o algoritmo de fato carrega o conteúdo completo da tabela, essa importação pode levar alguns segundos de acordo com as dimensões do banco de dados. Durante este processo a ferramenta fica congelada e não é permitida interações com a interface.

5 SELECIONANDO OS CAMPOS DE COMPARAÇÃO

Para realizar a comparação é preciso indicar ao algoritmo quais das colunas do banco de dados indicam uma entrada única, que servirão de base para indicar quais entradas são iguais em ambos os arquivos. É permitida a escolha de até três colunas conforme mostrado na Figura 5.

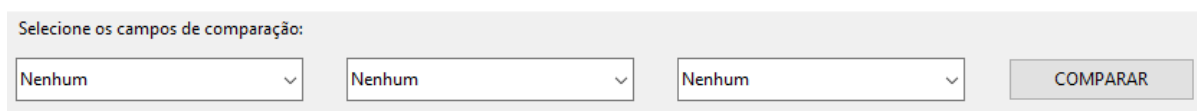


Figura 5: Seleção de Campos de Comparação.

Após selecionar os campos ao pressionar o botão 'COMPARAR' é iniciada a comparação.

NOTA: O tempo que o algoritmo leva para comparar é diretamente proporcional às dimensões da tabela. Embora otimizado para grandes quantidades de dados, os resultados podem demorar alguns segundos para serem exibidos.

As comparações podem ser realizadas em sequência alterando os campos, a tabela ou carregando novos arquivos. A quantidade de cada ocorrência é exibida no título de cada um

dos três campos da aba RELATÓRIO.

6 EXPORTANDO RESULTADOS

A exportação dos resultados é feita através do menu mostrado na figura 6, a ferramenta é capaz de exportar as tabelas selecionadas no banco de dados individualmente, o conteúdo da aba RELATÓRIO ou um arquivo completo que contem todos os itens citados.

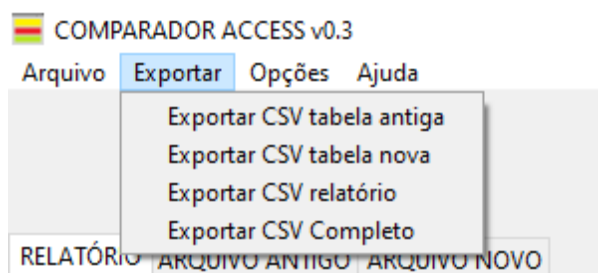


Figura 6: Menu 'Exportar'.

Ao selecionar uma das opções do menu, uma janela é aberta para escolher a pasta onde será salvo e o nome do arquivo, não é necessária a adição da extensão. Caso o computador possua o *Excel* instalado, ao fim da exportação o software é aberto automaticamente.

Os dados são exportados em arquivos do *Excel* em formato CSV com o mesmo formato em que são exibidos na interface da ferramenta, na opção de CSV completo são geradas três planilhas e as ocorrências são coloridas para facilitar a identificação.

7 FILTRAGEM

Com o intuito de facilitar a inspeção de certos elementos no banco de dados, no menu opções mostrado na Figura 7 existe a opção Filtrar.

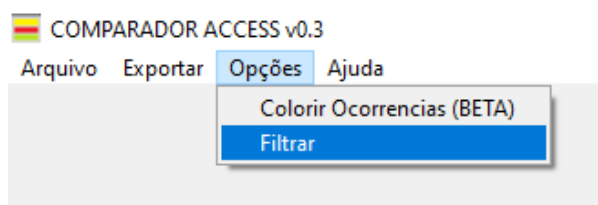


Figura 7: Opção Filtrar no Menu 'Opções'.

Através desta opção é possível apresentar nas três abas somente as entradas que possuam determinada condição. Ao selecionar a opção da Figura 7, é aberta a janela mostrada na Figura 8, nela é digitado a expressão que será buscada e os campos em que a busca será realizada.



Figura 8: Janela de Filtragem.

NOTA: A busca é feita em correspondência que são total ou parcialmente iguais ao termo digitado, não diferenciando letra maiúsculas de minúsculas.

O filtro aplicado na ferramenta é exibido na parte superior direita da interface conforme a Figura ??, para remover e voltar a comparação original deve-se pressionar o botão X à esquerda.

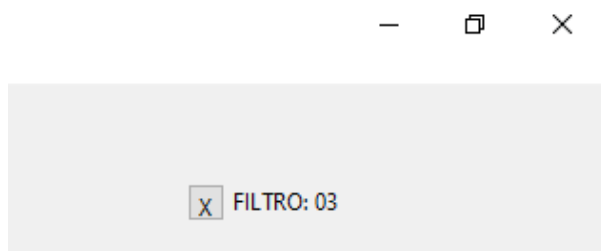


Figura 9: Opção de Remover o Filtro.

8 RESSALTAR OCORRÊNCIAS (BETA)

A opção de Colorir ocorrências (Figura 10) tem por função facilitar a visualização das ocorrências diretamente na ferramenta, para isso ela altera a cor de fundo das células do campo discrepantes que possuem valores alterados de um arquivo para outro. Nas abas ARQUIVO ANTIGO e ARQUIVO NOVO estas células também são marcadas juntamente com as entradas novas e excluídas.

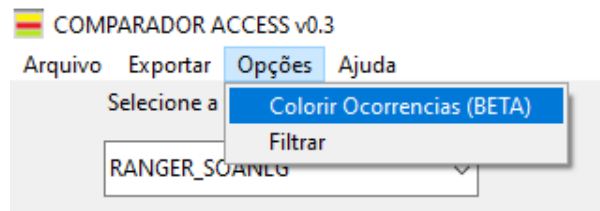


Figura 10: Opção de Colorir Ocorrências.

A posição padrão da opção de colorir é desmarcada, assim quando se deseja utilizar é necessário marca-la e realizar uma nova comparação através do botão 'COMPARAR'.

NOTA: A mudança de cores das planilhas necessitam de diversas iterações nos bancos de dados e aumentam o tempo de comparação consideravelmente, chegando ao dobro em determinados casos.

NOTA: Foram encontrados alguns problemas na mudança de cores quando existe um grande número de ocorrências.

Bibliografia

- [1] MCKINNEY, W. pandas: powerfull Python data analysis toolkit. [S.l.], 2022.
- [2] ALVAREZ, A. Guia Tkinter Documentation. [S.l.], 2016.
- [3] CRUZ, L. C. Data Science: Desenvolvimento De Aplicação Para Análise De Dados. Assis, São Paulo, 2018.