

Guia de utilização Alteryx Designer:

O que é Alteryx Designer?

Alteryx Designer é uma ferramenta para criar Datasets Analíticos na velocidade de pensamento. Com o conceito de trabalho em workflows, e a integração dos dados na nuvem, em aplicativos, bancos de dados, arquivos, fontes sociais e muito mais, campos e formatos diferentes, estruturados e não estruturados. Segundo estudos da Área de BI e Análise de Dados, aproximadamente 80% do tempo de projeto gasto, costuma ser na criação de conjuntos de dados mesclando-os manualmente usando código oneroso e planilhas propensas a erros. O Alteryx altera o jogo com fluxos de trabalho visuais, repetitivos e automatizados, com a capacidade de combinar mais rápido, mais inteligente e colaborativamente que outras ferramentas no mercado.

Bancos de dados e arquivos - Oracle, Microsoft, Amazon, NoSQL, planilhas e arquivos
Dezenas de aplicativos - Salesforce, NetSuite, Oracle, Marketo, SPSS, SAS e muito mais
Dados sociais e de terceiros - Twitter, Facebook, LinkedIn, Google Analytics, Dun & Bradstreet, Experian, Dados do Censo dos EUA

Com integração direta às principais ferramentas de visualização e descoberta de dados, o Alteryx permite que os analistas acompanhem seus conjuntos de dados combinados para análise rápida. E com análises avançadas integradas, os analistas não apenas têm o poder de misturar dados, mas também podem adicionar insights preditivos, espaciais e até prescritivos a seus conjuntos de dados.

No nosso projeto usaremos Alteryx Designer para fazer o ETL (Extraction-Transformation-Load).

Exemplo de Workflow de Dados No Alteryx:

Componentes:

Ferramenta de Seleção:



- Conectada ao "Input", permite fazer o SELECT das colunas, transformar o tipo de dado (Int, Double, String), e renomear as colunas.

Ferramentas de Investigação dos Dados:



- As diversas ferramentas de investigação dos dados auxiliam na detecção das propriedades das colunas nas tabelas, a correlação entre diferentes colunas, a qualidade dos dados, e apresentação de relatórios interativos.

Ferramentas de limpeza e transformação de dados:



Data Cleansing:



- Limpeza e substituição básica de dados, como remover os nulls, assim como caracteres não desejados.

Fórmula:



- Cria ou atualiza registros usando expressões com cálculos, operadores, etc.

Exemplo: `IIF([Sales]<1000000, "High", "Low")`

Filtro:



- Filtra registros dado uma condição, com dois outputs separa os registros que satisfazem e os que não satisfazem a condição definida.

Ferramentas de Input de dados:



- A ferramenta Input Data traz dados para o workflow, conectando-se a um arquivo ou banco de dados. Exemplo de formatos aceitos pela ferramenta são .csv, .json, .xml, .xls e .xlsx (Microsoft Excel), entre outros. Podemos nos conectar também a bancos de dados como MySQL, SQL Server, Oracle, PostgreSQL, Snowflake e muitos mais.

Ferramentas de Visualização de dados:



- A ferramenta Browse exibe dados de um componente conectado, bem como informações de perfil de dados, mapas, fragmentos de relatório e informações de análise de comportamento nos dados, e uma ferramenta muito útil para visualizar resultados conectado a ferramentas de visualização de dados.

Ferramentas de Transformação de dados:



- As ferramentas de transformação inclui ferramentas que sumarizam ou organizam dados, a ferramenta mais importante de este conjunto é a ferramenta de Sumarização, ela provê uma série de operações como:
- **Group By**: Combina registros do banco de dados com valores idênticos em um campo especificado em um único registro. Todos os dados resultantes dos registros em um grupo são então resumidos. Se nenhum campo a agrupar for especificado, o arquivo inteiro será agrupado.
- **Sum**: Retorna o valor da soma para o grupo. A soma é calculada adicionando todos os valores de um grupo.
- **Count**: Contagem de registros no grupo.
- **Count Non Null**: Conta os registros que NÃO são nulos. Nulo significa que não há valor definido para este registro. Isso é diferente de zero ou uma string vazia.
- **Count Distinct**: Contagem de registros únicos no grupo.
- **Count Distinct Non Null**: Conta apenas os registros que NÃO são nulos.
- **Min**: Retorna o valor mínimo.
- **Max**: Retorna o valor máximo.
- **First**: retorna o primeiro registro no grupo, com base em sua posição de registro.
- **Last**: Retorna o último registro no grupo, com base em sua posição de registro.

Ferramentas de Junção de dados:



- A categoria de componentes de junção inclui ferramentas que associam dois ou mais fluxos de dados, anexando dados a esquemas com lógicas como Left Join, Right Join, Inner Join, Union, Multiple Join, entre outras.

Para ver mais ferramentas e tutoriais de Alteryx Designer:

<http://www.alteryx.com/alteryx-designer-tools>

Referências:

<https://help.alteryx.com>

<http://www.alteryx.com/alteryx-designer-tools>