

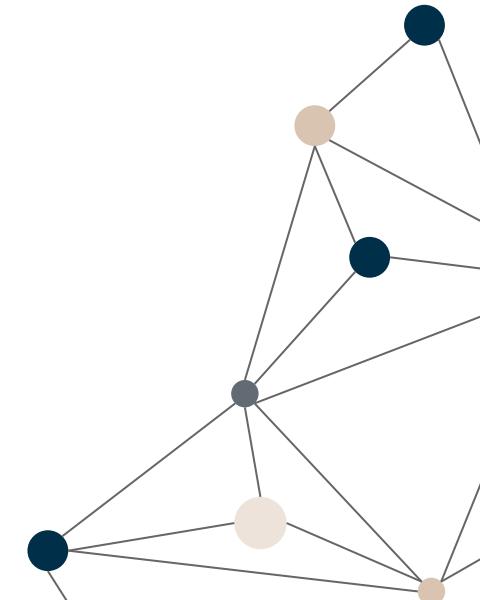
Módulo II

Power Query: Engenharia

de Dados

Curso de Extensão: Power BI para Engenheiros - Da Análise de Dados à Tomada de Decisão Estratégica.

Daniel Maia



Sumário

01 Revisão do Módulo I

03 Power Query

05 Hands On

02 Extract, Transform and Load

04 Principais Funcionalidades

06 Conclusão

Revisão do Módulo I

- O engenheiro moderno precisa **analisar e interpretar dados**, não apenas coletá-los.
- Dados são o novo ativo estratégico das empresas. (**Data-Driven Decision Making**)
- BI é o processo de **transformar dados em informações úteis para apoiar decisões estratégicas**.
- **Power BI** facilita a criação de indicadores de desempenho (KPIs) e o monitoramento de resultados.
- Power BI Desktop é utilizado para criar e o Power BI Service para visualizar
- **Criação do primeiro Dashboard no Power BI**

Etapas do Desenvolvimento

1

Conexão de Dados

Conecte-se a centenas de fontes: Excel, bancos de dados (SQL), arquivos CSV, web services e sistemas corporativos. Importe e combine dados de diferentes origens facilmente.

2

Modelagem de Dados (DAX)

Organize e relacione suas tabelas de dados. Use DAX (Data Analysis Expressions) para criar medidas, colunas calculadas e otimizar análises complexas.

3

Relatórios e Dashboards

Crie visuais interativos e dashboards personalizados. Transforme dados brutos em insights açãoáveis com gráficos, tabelas e mapas dinâmicos.

Etapas do Desenvolvimento



Conexão de Dados

Conecte-se a centenas de fontes: Excel, bancos de dados (SQL), arquivos CSV, web services e sistemas corporativos. Importe e combine dados de diferentes origens facilmente.



Power Query



Modelagem de Dados (DAX)

Organize e relacione suas tabelas de dados. Use DAX (Data Analysis Expressions) para criar medidas, colunas calculadas e otimizar análises complexas.



Relatórios e Dashboards

Crie visuais interativos e dashboards personalizados. Transforme dados brutos em insights açãoáveis com gráficos, tabelas e mapas dinâmicos.

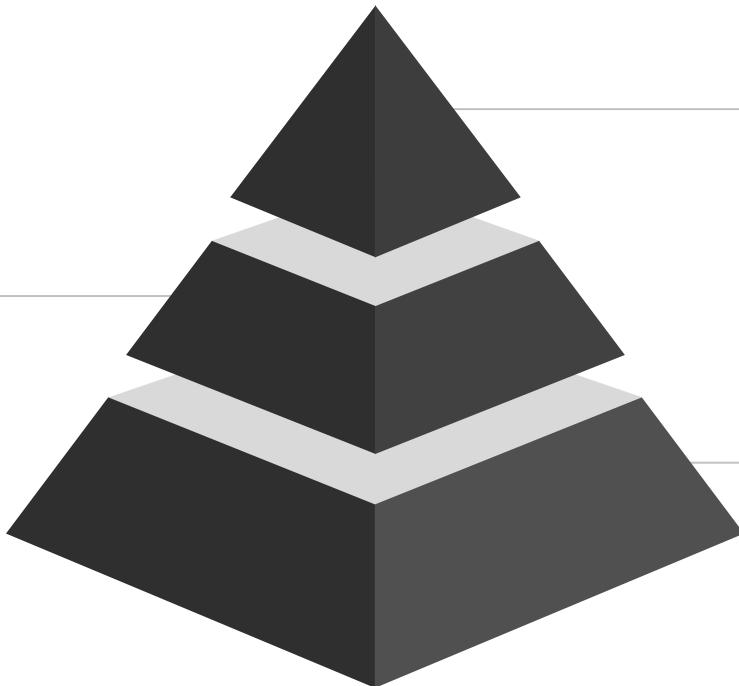
Extract, Transform & Load



Transform

Os dados passam por uma engenharia de limpeza e padronização.

2



1

Os dados prontos e limpos são carregados para o ambiente de análise

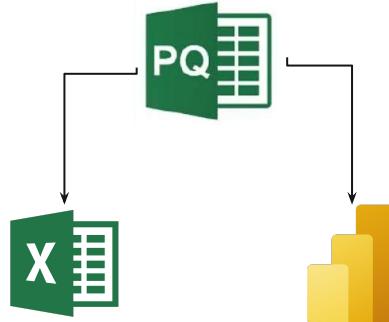


Extract

É o momento de coletar os dados das fontes disponíveis
O foco aqui é trazer os dados brutos

Power Query

- Ambiente de tratamento e edição de planilhas
- Ferramenta de ETL integrada ao Excel e Power BI
- Automatiza processos de preparação com passos reproduzíveis
- Facilita padronização, correção de erros e união de fontes
- Reduz drasticamente o tempo de manipulação manual de dados
- É a porta de entrada para um fluxo de dados robusto e profissional



Sem título - Editor do Power Query

Arquivo Página Inicial Transformar Adicionar Coluna Exibição Ferramentas Ajuda

Fchar e Aplicar Novo Fonte Recentes Inserir Dados Configurações da fonte de dados Gerenciar Parâmetros Atualização Consulta Propriedades Editor Avançado Escolher Colunas Remover Colunas Manter Linhas Remover Linhas Dividir Coluna por Agrupar Coluna Substituir Valores Tipo de Dados: Texto Usar a Primeira Linha como Cabeçalho Mesclar Consultas Acrescentar Consultas Combinar Arquivos Combinar

Consultas [1]

inauguracoes_poluido.csv

1

2

3

4

5

6

Config. Consulta

PROPRIEDADES

Nome: **inauguracoes_poluido.csv**

Todas as Propriedades

ETAPAS APLICADAS

- Fonte
- Navegação
- Cabeçalhos Promovidos
- Tipo Alterado**

VISUALIZAÇÃO BAIXADA Á(S) 22:25

5 COLUNAS, 999 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Nome	Construção	Linha	Inauguração	Idade
Frédérico Moratto	CPTM	linha 7	14/02/1867	152
Água Branca	CPTM	linha 7	Feb/1867	152
brás	CPTM	linha 7	Feb/1867	NULL
Jundiaí	CPTM	linha 7	1867-02-16	152
LUZ	CPTM	linha 7	1867-02-16	152
Perus	Cptm	linha 7	16/02/1867	152
Rio Grande da Serra	CPTM	10	16/02/1867	152
Santo André	CPTM	10	02-16-67	NULL
Júlio Prestes	Cptm	8		
Poá	cptm	11	1875-01-06	—
Mogi das Cruzes	CPTM	11		684771 144 anos
Suzano	CPTM	11	06/11/1875	144
BARUERI	CPTM	8	10/07/1875	—
Itapevi	Cptm	8	Jul/1875	—
Campo Limpo Paulista	CPTM	linha 7	1881-01-01	—
Mauá	CPTM	10	01/04/1883	136
São Caetano do Sul	cptm	10	01/05/1883	136
Caleiras	CPTM	linha 7	687564	136
Pirituba	Cptm	linha 7	01/02/1885	134
Ribeirão Pires	CPTM	10	01/03/1885	134
Ipiranga	CPTM	10	Apr/1886	133
francDOrOcha	CPTM	linha 7	01/02/1888	NULL
Várzea Paulista	Cptm	linha 7	07/01/1891	128
Jardim	CPTM	linha 7	01/10/1891	128 anos
PALMEIRAS-BARRA FUNDA	CPTM	linha 7	19/05/1892	127
Osasco	CPTM	8	1895-08-20	124
Juventus-Mooca	CPTM	10	07/10/1898	121
Ipa	CPTM	linha 7	20/02/1899	120
Botujuru	Cptm	linha 7	09-17-08	NULL
Jundiapeba	CPTM	11	20/07/1914	105
Presidente Altino	CPTM	8	06/03/1919	—
Santa Terezinha	CPTM	8	01/01/1920	100
Domingos de Moraes	CPTM	8	25/06/1920	—
Capuava	CPTM	10	15/09/1920	99
Carapicuíba	CPTM	8	01/07/1921	98
AMADOR BUENO	CPTM	8	08/05/1922	97 anos
Caetano Viana	CPTM	11	03/05/1924	95
Engenheiro Manoel Feio	CPTM	12	Feb/1926	93 anos

1. Faixa de opções
2. Painel de consultas

3. Exibição atual
4. Painel de consultas

5. Barra de consultas
6. Barra de status

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- Alterar tipos de dados
- Formatos de data
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path: Página Inicial → Remover Linhas → Remover Duplicatas

The screenshot shows the Microsoft Power Query Editor interface. At the top, the ribbon has tabs: Página Inicial, Transformar, Adicionar Coluna, Exibição, Ferramentas, and Ajuda. The 'Transformar' tab is selected. Below the ribbon, there's a toolbar with icons for New, Fonts, Insert, Data Sources, Configuration, Parameters, Refresh, Properties, and various transformation tools like Remove Columns, Remove Rows, and Remove Duplicates. A context menu is open over a table named 'guracoes_poluido.csv'. The menu path is 'Table.TransformColumnTypes' followed by 'Nome'. The 'Nome' item has a dropdown arrow pointing to a submenu. This submenu contains several options: 'Remover Linhas Superiores', 'Remover Linhas Inferiores', 'Remover linhas alternadas', 'Remover Linhas em Branco', 'Remover Erros', and 'Remover Duplicadas'. The 'Remover Duplicadas' option is highlighted with a red rectangle. To the right of the table, there are sections for 'PROPRIEDADES' (Properties) showing 'Nome: inaugurações_poluido' and 'ETAPAS APPLICADAS' (Applied Steps) showing 'Todas as Propriedades'.

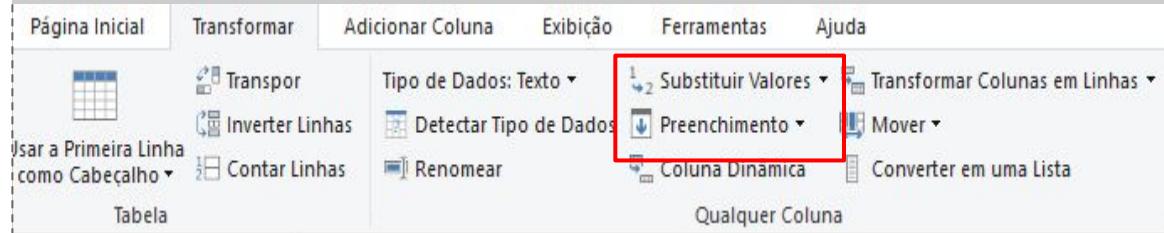
Eliminar linhas completamente iguais para evitar registros repetidos.

Principais Funcionalidades

- Remover duplicatas e erros
- **Tratar valores nulos**
- Normalizar texto
- Alterar tipos de dados
- Formatos de data
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path:

- *Transformar → Substituir Valores*
- *Transformar → Preencher → Para Cima / Para Baixo*



Corrigir registros incompletos substituindo valores ausentes de maneira adequada.

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- **Normalizar texto**
- Alterar tipos de dados
- Formatos de data
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path: Transformar → Formatar → (letra minúscula / letra maiúscula / capitalizar cada palavra)

The screenshot shows the Power BI 'Transform Data' editor interface. A context menu is open over the 'Nome' column header, specifically under the 'Formato' (Format) section. This menu includes options for changing case (minúscula, MAIÚSCULA), capitalizing words, and adding prefixes/suffixes. The 'Nome' column itself contains two entries: 'Fr@ncisco Mor@to' and 'Água Branca', both displayed in uppercase ('MAIÚSCULA').

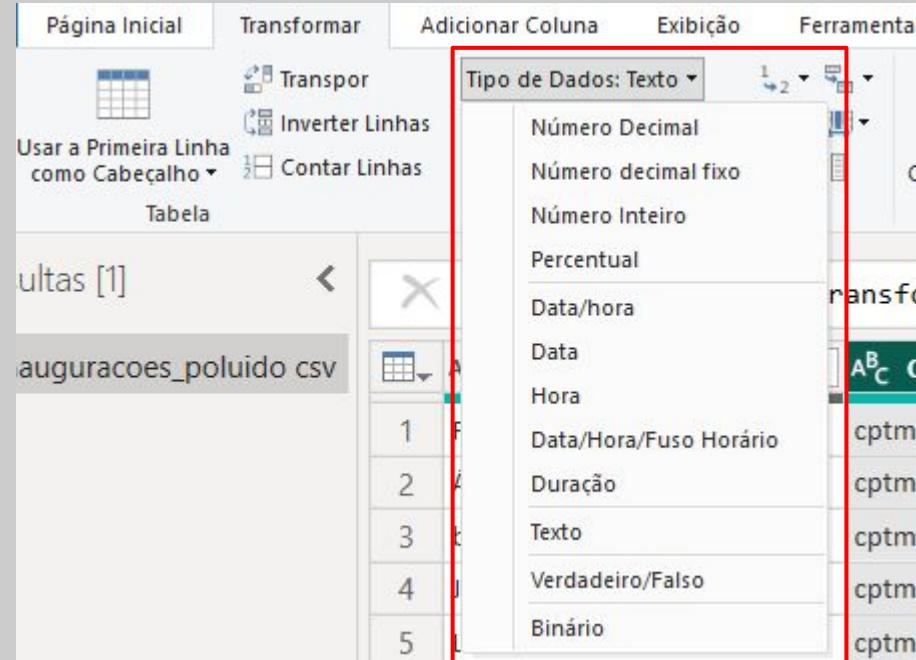
	Nome	Construtor
1	Fr@ncisco Mor@to	cptm
2	Água Branca	cptm

Padronizar nomes e linhas para garantir consistência.

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- **Alterar tipos de dados**
- Formatos de data
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path: Transformar → Tipo de Dados → Data



Forçar conversão de datas quebradas, padronizar formatos e corrigir erros.

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- Alterar tipos de dados
- **Formatos de data**
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path: Transformar → Data

The screenshot shows the 'Transformar' (Transform) page in Power BI. A red box highlights the 'Data' context menu on the far right. The menu includes options like 'Idade', 'Somente Data', 'Analisar', 'Ano', 'Mês', 'Trimestre', 'Semana', 'Dia', 'Combinar Data e Hora', 'Mais anterior', and 'Mais recente'. Below the menu, a table displays data with two columns: 'Nome' and 'Construção'. The 'Nome' column contains values like 'Fr@ncisco Mor@to', 'Água Branca', 'brás', 'Jundiaí', 'LUZ', and 'Porus'. The 'Construção' column contains the value 'cptm' repeated six times. The formula bar at the top shows the formula: = Table.TransformColumnTypes(#"Tabel1", {{"Nome", typeText}, {"Construção", typeText}}).

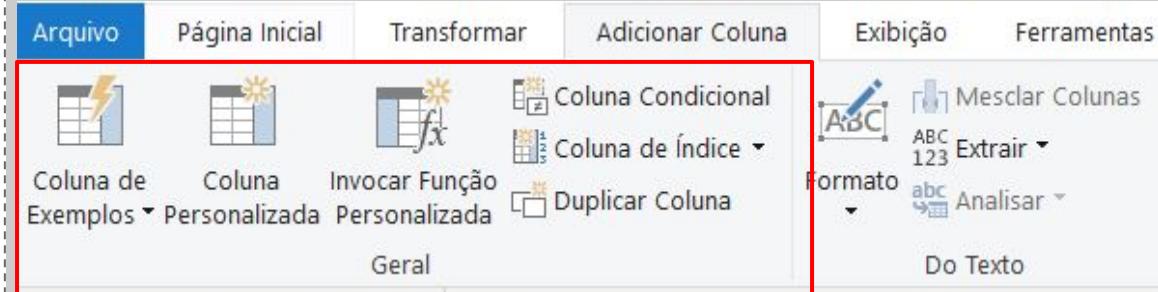
Nome	Construção
Fr@ncisco Mor@to	cptm
Água Branca	cptm
brás	cptm
Jundiaí	cptm
LUZ	cptm
Porus	cptm

Definir o formato da data a ser utilizado

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- Alterar tipos de dados
- Formatos de data
- **Criar colunas Calculadas / Condicionais / de Exemplos**
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path: Adicionar Coluna → Coluna Personalizada

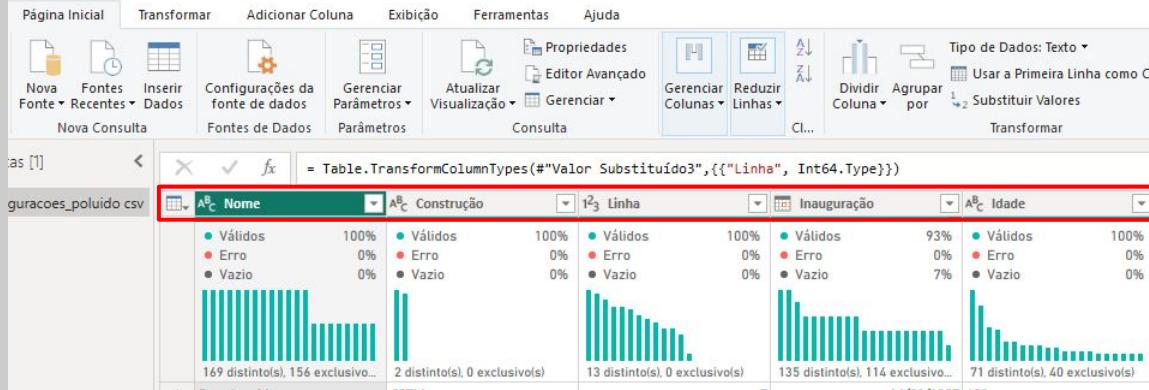


Criação de colunas personalizadas com base em colunas já existentes na base

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- Alterar tipos de dados
- Formatos de data
- Criar colunas calculadas e condicionais
- **Filtrar e reorganizar informações**
- Verificar erros lógicos

Path: Exibição Atual → Coluna Desejada



Filtrar arquivo geral com base em valores desejados

Principais Funcionalidades

- Remover duplicatas e erros
- Tratar valores nulos
- Normalizar texto
- Alterar tipos de dados
- Formatos de data
- Criar colunas calculadas e condicionais
- Filtrar e reorganizar informações
- Verificar erros lógicos

Path:

- Exibição → Coluna de Qualidade
- Página Inicial → Manter Linhas / Remover Linhas → Erros

The screenshot shows the Microsoft Power Query Editor interface. The top ribbon has tabs: Arquivo, Página Inicial, Transformar, Adicionar Coluna, Exibição (selected), and Ferramentas. Under the 'Exibição' tab, there are several checkboxes: 'Barra de Fórmulas' (checked), 'Espaçamento uniforme' (unchecked), 'Mostrar espaço em branco' (checked), 'Distribuição de colunas' (unchecked), 'Perfil da coluna' (unchecked), and 'Qualidade da coluna' (checked and highlighted with a red box). Below the ribbon is a 'Layout' section with 'Visualização dos Dados'. The main area shows a table with columns: Nome, Construção, Linha, and Inauguração. The 'Inauguração' column contains dates like '14/02/1867', 'Feb/1867', 'Feb/1867', '1867-02-16', and '1867-02-16'. The 'Transformar' tab is selected at the bottom. On the right side, there are various transformation tools and a 'Config. Consulta' pane.

Nome	Construção	Linha	Inauguração
Fr@ncisco Mor@to	CPTM	linha 7	14/02/1867
Água Branca	CPTM	linha 7	Feb/1867
bráS	CPTM	linha 7	Feb/1867
Jundiaí	CPTM	linha 7	1867-02-16
LUZ	CPTM	linha 7	1867-02-16

HANDS ON



Conclusão

- O Power Query automatiza a preparação de dados, eliminando retrabalho e garantindo consistência mesmo quando novas bases chegam.
- A qualidade do dashboard depende diretamente da qualidade do ETL, e essa etapa é tão importante quanto a visualização final.
- Padronizar, limpar e corrigir dados é um processo técnico, mas também lógico, que exige entender o contexto e o objetivo da análise.
- Um bom profissional de dados domina a preparação antes da análise
- Agora com os dados prontos, como vamos analisar a relação entre eles?