



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)

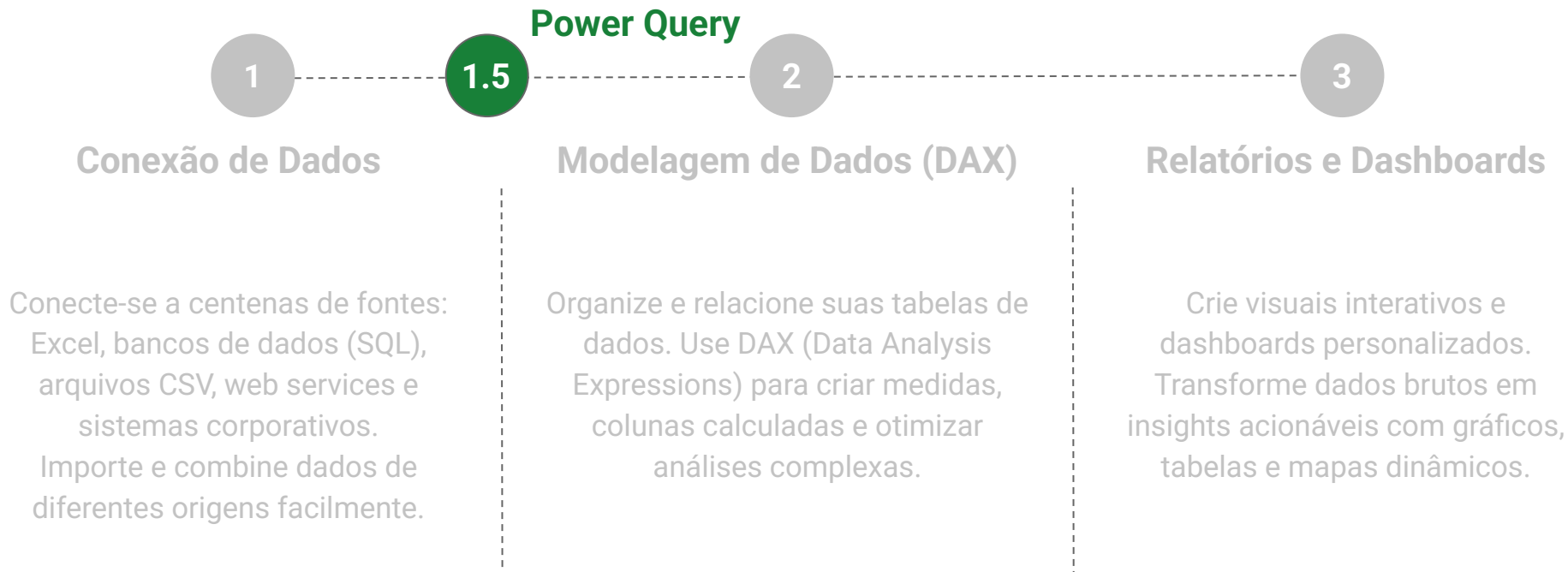
Organize e relacione suas tabelas de dados. Use DAX (Data Analysis Expressions) para criar medidas, colunas calculadas e otimizar análises complexas.

3

Relatórios e Dashboards

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

Etapas do Desenvolvimento



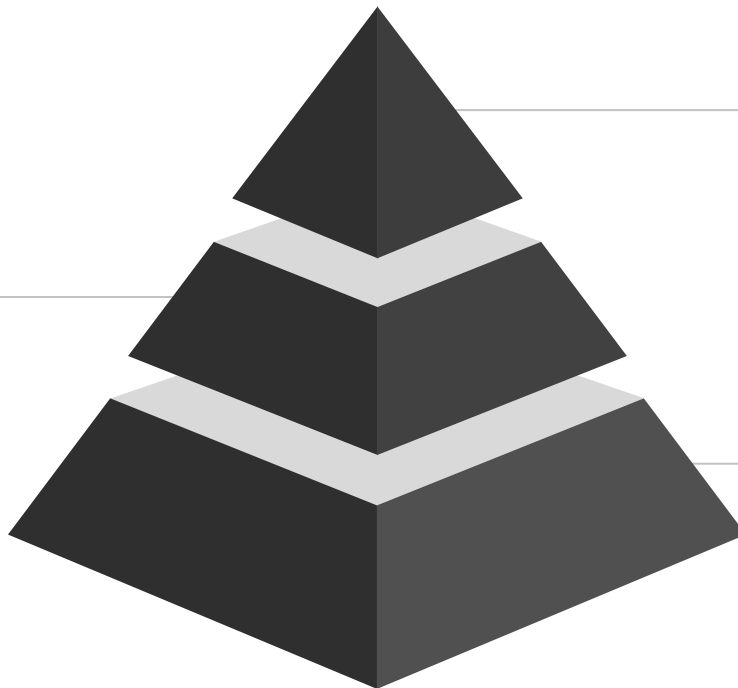
Extract, Transform & Load



Transform

Os dados passam por uma engenharia de limpeza e padronização.

2



Load

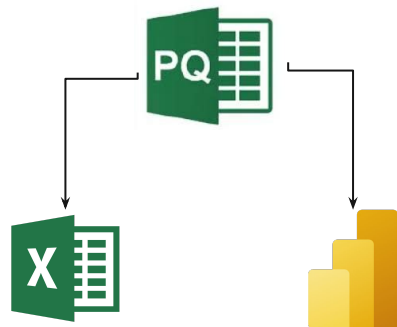
1 Os dados prontos e limpos são carregados para o ambiente de análise

Extract

3 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

Fechar e Aplicar Nova Fonte Fontes Recentes Inserir Dados Configurações da Fonte de dados Gerenciar Parâmetros Atualizar Visualização Propriedades Editor Avançado Escolher Colunas Remover Colunas Manter Linhas Remover Linhas Reduzir Linhas Classificar Dividir Coluna Agrupar por Tipo de Dados: Texto Usar a Primeira Linha como Cabeçalho Mesclar Consultas Acrescentar Consultas Combinar Arquivos Combinar

Consultas [1]

inacuracoes_poluido.csv

Nome

Construção

Linha

Inauguração

Idade

1 Fr@ncisco Mor@to CPTM linha 7 14/02/1867 152

2 Água Branca CPTM linha 7 Feb/1867 152

3 brás CPTM linha 7 Feb/1867 NULL

4 Jundiá CPTM linha 7 1867-02-16 152

5 LUZ CPTM linha 7 1867-02-16 152

6 Perus Cptm linha 7 16/02/1867 152

7 Rio Grande da Serra CPTM 10 16/02/1867 152

8 Santo André CPTm 10 02-16-67 NULL

9 Júlio Prestes Cptm 8 07/10/1972 —

10 Poá cptm 11 1875-01-06 —

11 Mogi das Cruzes CPTM 11 684771 144 anos

12 Suzano CPTM 11 06/11/1875 144

13 BARUERI CPTM 8 10/07/1875 —

14 Itapevi Cptm 8 Jul/1875 —

15 Campo Limpo Paulista CPTM linha 7 1881-01-01 —

16 Mauá CPTM 10 01/04/1883 136

17 São Caetano do Sul cptm 10 01/05/1883 136

18 Caeiras CPTM linha 7 687564 136

19 Pirituba Cptm linha 7 01/02/1885 134

20 Ribeirão Pires CPTM 10 01/03/1885 134

21 Ipiranga CPTM 10 Apr/1886 133

22 francO dA rOchA CPTM linha 7 01/02/1888 NULL

23 Várzea Paulista Cptm linha 7 07/01/1991 128

24 JAraguá CPTM linha 7 01/10/1891 128 anos

25 PALMEIRAS-BARRA FUNDA CPTM linha 7 19/05/1892 127

26 Osasco cPtM 8 1895-08-20 124

27 Juventus-Mooca CPTM 10 07/10/1898 121

28 Iapa CPTM linha 7 20/02/1899 120

29 Botujuru Cptm linha 7 09-17-08 NULL

30 Jundiapéba CPTM 11 20/07/1914 105

31 Presidente Altino CPTM 8 06/03/1919 —

32 Santa Terezinha CPTM 8 01/01/1920 100

33 Domingos de Moraes CPTM 8 25/06/1920 —

34 Capuava CPTM 10 15/09/1920 99

35 Carapicuíba CPTM 8 01/07/1921 98

36 AMADOR BUENO CPTM 8 08/05/1922 97 anos

37 Calmon Viana CPTM 11 03/05/1924 95

38 Engenheiro Manoel Feio CPTM 12 Feb/1926 93 anos

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

Config. Consulta

PROPRIEDADES

Nome

inacuracoes_poluido.csv

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Navegação

Cabeçalhos Promovidos

Tipo Alterado

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

VISUALIZAÇÃO BAIXADA À(S) 22:25

1. Faixa de opções
2. Pannel de consultas

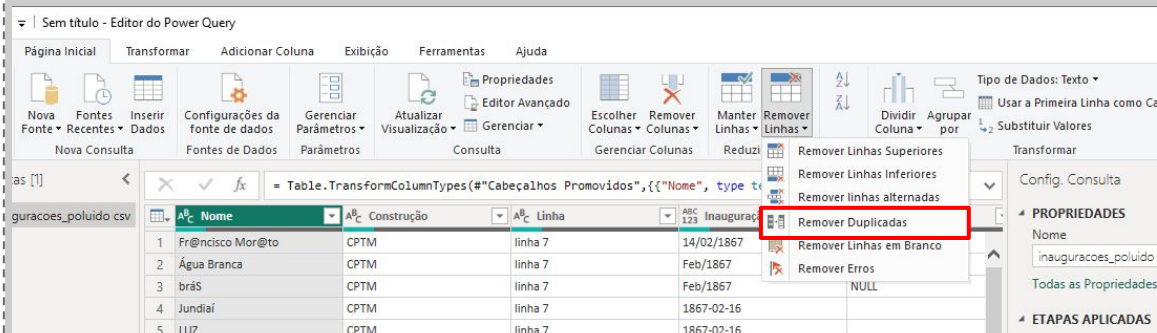
3. Exibição atual
4. Pannel de consultas

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*



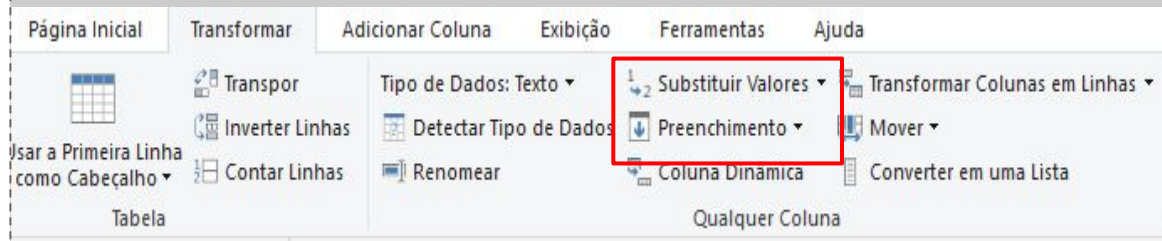
Eliminar linhas completamente iguais para evitar registros repetidos.

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*

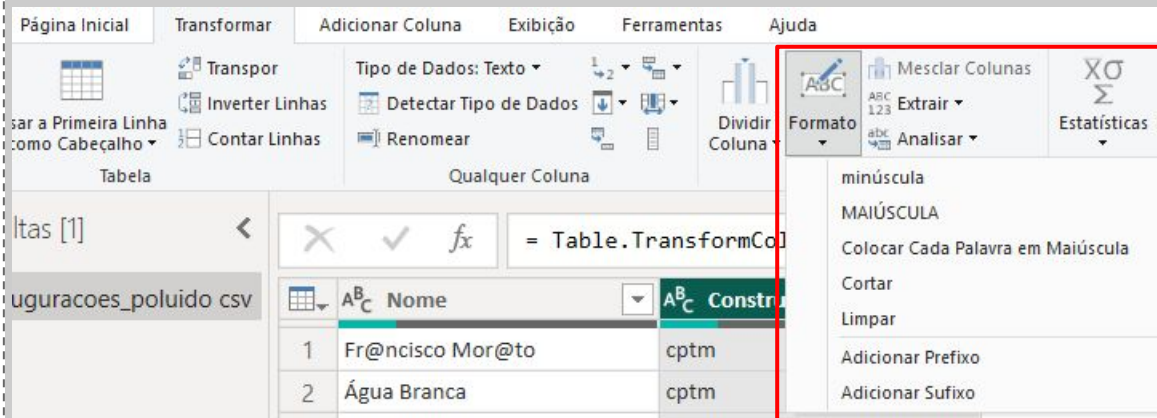


Corrigir registros incompletos substituindo valores ausentes de maneira adequada.

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)

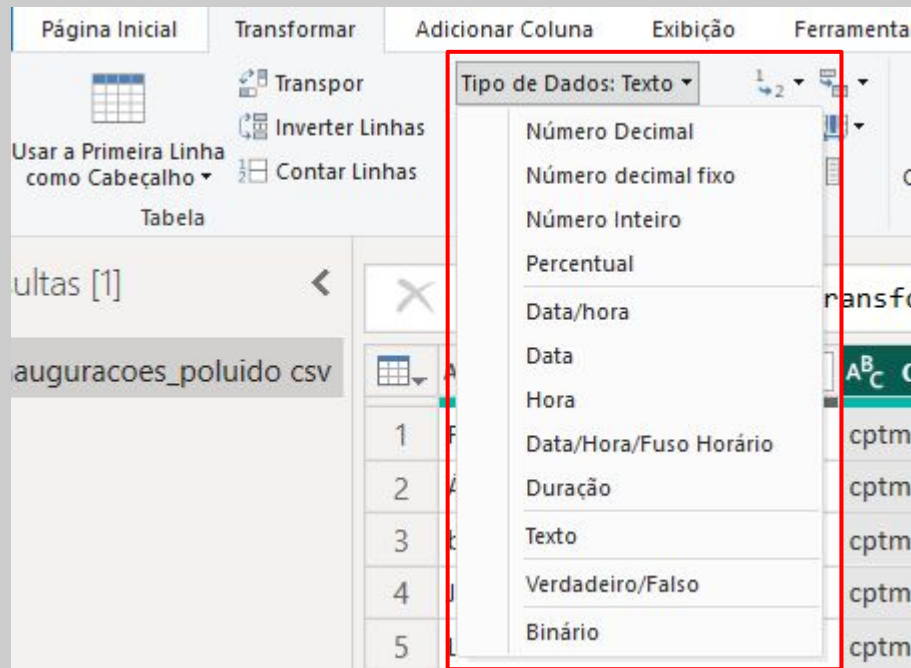


Padronizar nomes e linhas para garantir consistência.

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

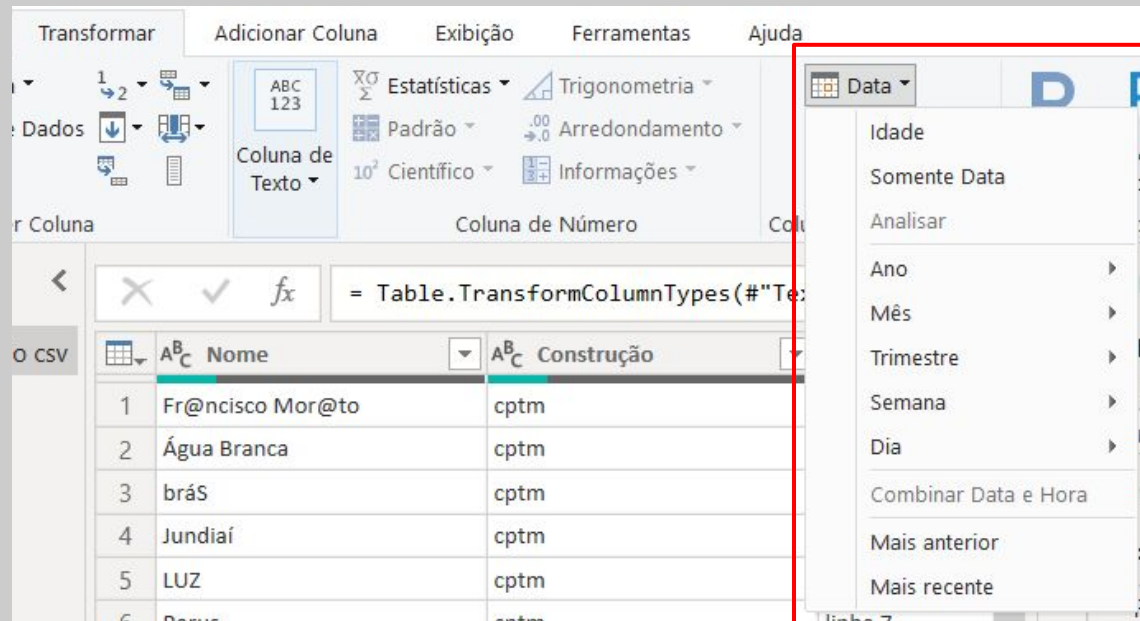


Forçar conversão de datas quebradas, padronizar formatos e corrigir erros.

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

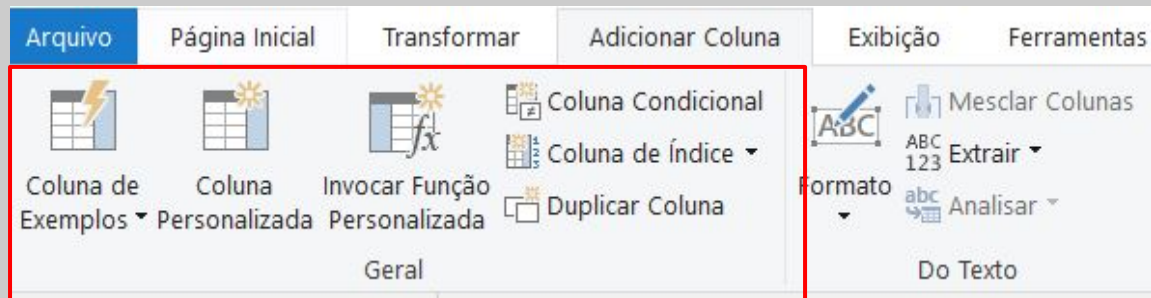


Definir o formato da data a ser utilizado

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

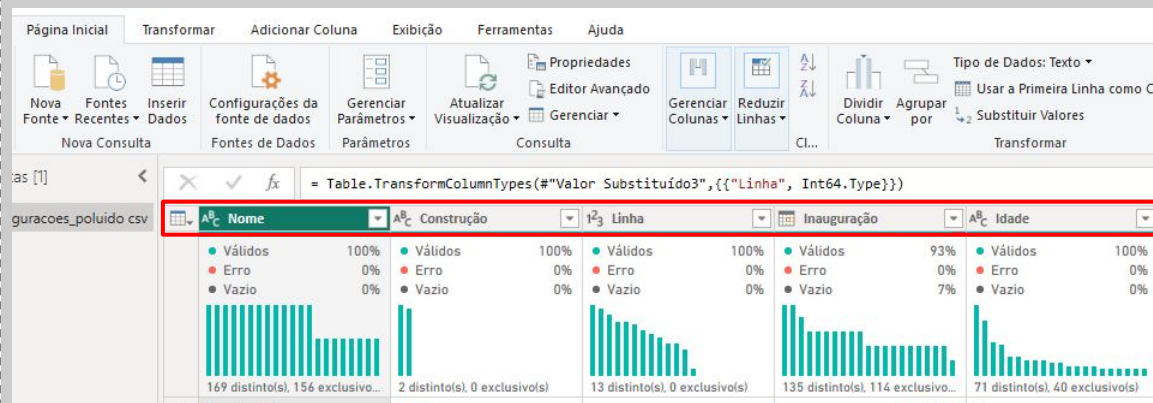


Criação de colunas personalizadas com base em colunas já existentes na base

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 displays the Power Query Editor interface. The top ribbon shows the 'Exibição' (View) tab selected. Under the 'Visualização dos Dados' (Data View) group, the 'Qualidade da coluna' (Column Quality) checkbox is checked and highlighted with a red rectangle. Below the ribbon, the 'Remover Linhas' (Remove Rows) dropdown menu is open, and the 'Remover Erros' (Remove Errors) option is highlighted with a red rectangle. The background shows a data table with columns 'Nome', 'Construção', 'Linha', and 'Inauguração'.

	Nome	Construção	Linha	Inauguração
1	Francisco Morato	CPTM	linha 7	14/02/1867
2	Água Branca	CPTM	linha 7	Feb/1867
3	brás	CPTM	linha 7	Feb/1867
4	Jundiaí	CPTM	linha 7	1867-02-16
5	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?
- 