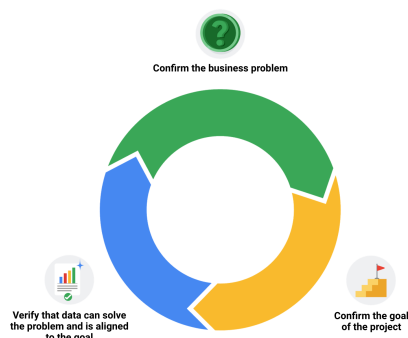▶ Lista de verificação de limpeza de dados

└▸ Problemas comuns

○ Dados nulos
○ Erros ortográficos
○ Números digitados incorretamente
○ Espaços e caracteres extra (Trim) (Distinct)
○ Duplicatas

○ tipos de dados incompatíveis
○ strings bagunçadas
○ Formatos de datas bagunçados
○ Rótulos de variáveis enganosos

└▸ Revisar a meta do seu projeto

○ Confirmar o problema de negócios
○ Confirmar a meta do projeto
○ Verificar se os dados podem resolver o problema e se estão alinhados à meta



Confirm the business problem

Verify that data can solve the problem and is aligned to the goal

Confirm the goal of the project

✳ É Recomendado sempre fazer CHANGELOG

└▸ Arquivo uma lista das modificações feitas no projeto em ordem cronológica.



└▸ Controle de versão

| Planilhas Google | 1. Clique com o botão direito do mouse na célula e selecione **Mostrar histórico de edição**. 2. Clique na seta para a esquerda < ou na seta para a direita > para retroceder ou avançar no histórico, conforme necessário. |
|---|---|
| Microsoft Excel | 1. Se o Acompanhamento de Mudanças tiver sido ativado para a planilha: Clique em **Revisar**. 2. Em **Acompanhamento de alterações**, Clique na opção **Aceitar/Rejeitar alterações** para aceitar ou rejeitar qualquer alteração feita. |
| BigQuery | Abra uma versão anterior (sem reverter para ela) e descubra o que mudou comparando-a com a versão atual. |

⊗ Normalmente um log de mudanças registra:

- Dados, Arquivo, fórmula, consulta ou qualquer outro componente que tenha mudado.
- Descrição do que foi mudado
- Data da mudança
- Quem fez a mudança
- Quem aprovou a mudança
- Num da versão
- Motivo da mudança

▶ Funções Avançadas para Limpeza de dados

| Function | Syntax (Google Sheets) | Menu Options (Microsoft Excel) | Primary Use |
|---|---|---|---|
| IMPORTRANGE | =IMPORTRANGE(spreadsheet_url, range_string) | Paste Link (copy the data first) | Imports (pastes) data from one sheet to another and keeps it automatically updated. |
| QUERY | =QUERY(Sheet and Range, "Select *") | Data > From Other Sources > From Microsoft Query | Enables pseudo SQL (SQL-like) statements or a wizard to import the data. |
| FILTER | =FILTER(range, condition1, [condition2, ...]) | Filter (conditions per column) | Displays only the data that meets the specified conditions. |