

## Como manter os dados organizados com a classificação e os filtros

Você aprendeu sobre as quatro fases da análise:

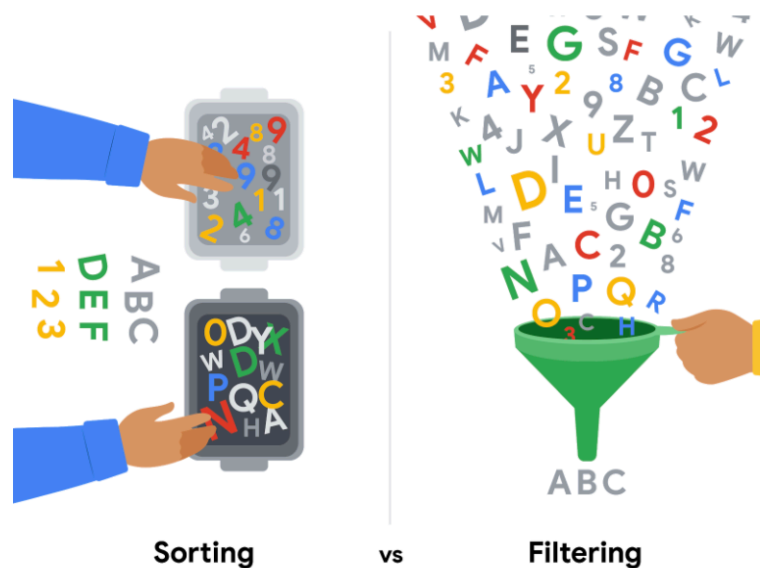
- **Organizar dados**
- **Formatar e ajustar dados**
- **Obter contribuições de outros**
- **Transformar dados**

A **organização dos conjuntos de dados** é muito importante para os analistas de dados. A maioria dos conjuntos que usar serão organizados como tabelas, que são úteis, pois permitem a você manipular e categorizar seus dados. Com diferentes categorias e classificações, é possível ter foco e saber diferenciar seus dados com rapidez e facilidade.

Os analistas de dados também precisam formatar e ajustar os dados durante uma análise. A **classificação** e a **filtragem** são duas formas de manter tudo organizado ao formatar e ajustar os dados para trabalhar com eles. O filtro, por exemplo, pode ajudar você a identificar erros ou pontos fora da curva, para que você possa corrigir ou sinalizá-los antes da sua análise. Os **pontos fora da curva** referem-se a pontos dos dados que destoam bastante de dados similarmente coletados, podendo não ser valores confiáveis. A vantagem de se filtrar os dados é que, após corrigir os erros ou identificar pontos fora da curva, é possível remover o filtro e retornar os dados à organização original.

Você aprenderá a diferença entre classificação e filtragem, e conhecerá como determinada forma de classificação é realizada em uma tabela dinâmica.

### CLASSIFICAÇÃO VS. FILTRAGEM



A **classificação** é quando você organiza os dados em uma ordem significativa para torná-los mais fáceis de entender, analisar e visualizar. Ela posiciona seus dados com base em uma métrica específica que você escolhe. É possível classificar os dados em planilhas, bancos de dados do SQL (quando o conjunto de dados é muito extenso para as planilhas) e tabelas em documentos. Se você precisa, por exemplo, classificar itens ou criar listas cronológicas, você pode classificar em ordem crescente ou decrescente. Quer saber os filmes favoritos de determinado grupo? Classifique por título de filme e descubra. A classificação ordenará os dados de forma válida e fornecerá a você insights imediatos. Além disso, ajuda a agrupar dados similares por determinada classificação. Com relação aos filmes, você pode classificar por gênero, como ação, drama, ficção científica ou romance.

A **filtragem** é usada quando você só tem interesse em ver os dados que atendem a um critério específico, ocultando todo o restante. É muito útil quando se tem muitos dados. Você pode economizar tempo ao zerar os dados que sejam realmente importantes ou dados que contêm bugs ou erros. A maioria das planilhas e bancos de dados do SQL permite que você filtre seus dados de várias formas. Com a filtragem, você pode encontrar o que está procurando com facilidade.

Se você só está interessado, por exemplo, em descobrir quem assistiu a filmes em outubro, use um filtro nas datas. Dessa forma, serão exibidos apenas os registros de filmes assistidos nesse mês. Você pode, então, ver os nomes das pessoas para ver quem assistiu a filmes em outubro. Só para recapitular, a forma mais fácil de lembrar a diferença entre classificação e filtragem é que você pode usar a primeira para ordenar os dados rapidamente, enquanto que a filtragem é usada para exigir *apenas* os dados que atendem ao critério especificado. Use essa segunda opção quando precisar diminuir a quantidade de dados exibidos.

Vale destacar que, depois de filtrar os dados, você também pode **classificar esses dados já filtrados**. Se você analisar mais uma vez o exemplo de como descobrir quem assistiu a filmes em outubro, depois de filtrar os filmes assistidos no mês indicado, é possível classificar os nomes das pessoas que assistiram aos filmes em ordem alfabética.

## COMO CLASSIFICAR EM UMA TABELA DINÂMICA

Os itens das áreas de linha e coluna de uma tabela dinâmica são classificados em ordem crescente por qualquer lista personalizada primeiro. Se sua lista contém, por exemplo, os dias da semana, a tabela dinâmica tem a opção de classificar os nomes do dia da semana e do mês assim: segunda-feira, terça-feira, quarta-feira, etc. ao invés de alfabeticamente, como: domingo, quarta-feira, quinta-feira, etc.

Se os itens não fizerem parte de uma lista personalizada, serão classificados em ordem crescente por padrão; no entanto, se você classificar em ordem decrescente, estará definindo uma regra que rege como o campo é classificado mesmo após a adição de novos campos de dados.