

AWS QuickSight - Fundamentos

- 1 [Termos Básicos](#)
- 2 [Página Inicial](#)
- 3 [Datasets](#)
 - 3.1 [Arquivos](#)
 - 3.2 [Bases de dados](#)
 - 3.3 [Criar datasets a partir de data sources ou datasets já existentes](#)
- 4 [Dimensões e métricas](#)
- 5 [Funções](#)
- 6 [Campos calculados](#)
- 7 [Filtros](#)
 - 7.1 [Filtros em Datasets](#)
 - 7.2 [Filtros em Análises](#)
 - 7.3 [Filtros em Cascata](#)
- 8 [Parâmetros](#)
- 9 [Ações](#)
- 10 [Publicar dashboards](#)
- 11 [Recursos Adicionais](#)

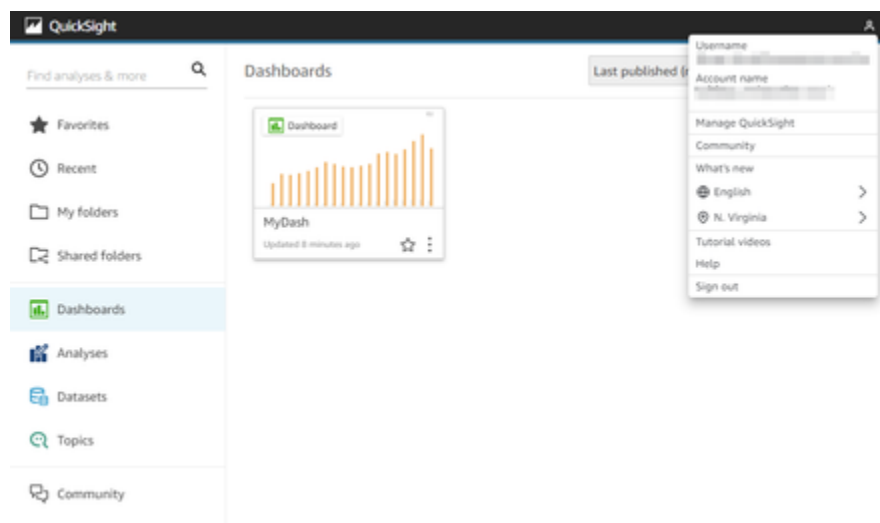
AWS QuickSight é um serviço gerenciado de *Business Intelligence* que permite a criação e compartilhamento de análises de dados afim de obter *insights* de inteligência de negócios de forma rápida e fácil. Permite conectar com diversas fontes de dados para extrair informações de interesse, através da criação de painéis interativos que podem ser compartilhados e acessados com facilidade de qualquer dispositivo. O Amazon QuickSight é um excelente serviço para empresas que já utilizam outros serviços AWS, pois ele se integra facilmente com outros serviços, como Amazon Redshift, Amazon S3 e Amazon RDS, tornando ainda mais fácil gerenciar e analisar dados.

Termos Básicos

- **Preparação de dados (*data preparation*):** Processo de preparação dos dados para uso em análises. Compreende operações como filtrar, renomear campos, alterar tipos, adicionar campos calculadores ou criar consultas SQL no momento da criação/edição de um dataset.
- **Dataset:** Coleção de dados de interesse, obtidos através de um ou mais *data sources*.
- **Data Sources:** Representação de uma conexão a uma fonte de dados específica (banco de dados relacional, arquivo, Athena, Spark)
- **SPICE (*Super-fast, Parallel, In-memory Calculation Engine*):** Motor *in-memory* com armazenamento associado que acelera o processamento de cálculos e a recuperação de dados. Atua como um mecanismo de cache que favorece o rápido processamento de dados sem comprometer a experiência do usuário.
- **Análise (*analyses*):** Espaço de trabalho onde criam-se visualizações (representações gráficas dos dados).
- **Visualização (*visual*):** É uma representação visual específica dos dados, como um gráfico, KPI, insight, etc.
- **Planilha (*sheet*):** Divisão de uma análise, tal qual uma aba de navegação. Compreende uma ou mais visualizações.
- **Dashboard:** É a versão publicada de uma análise.

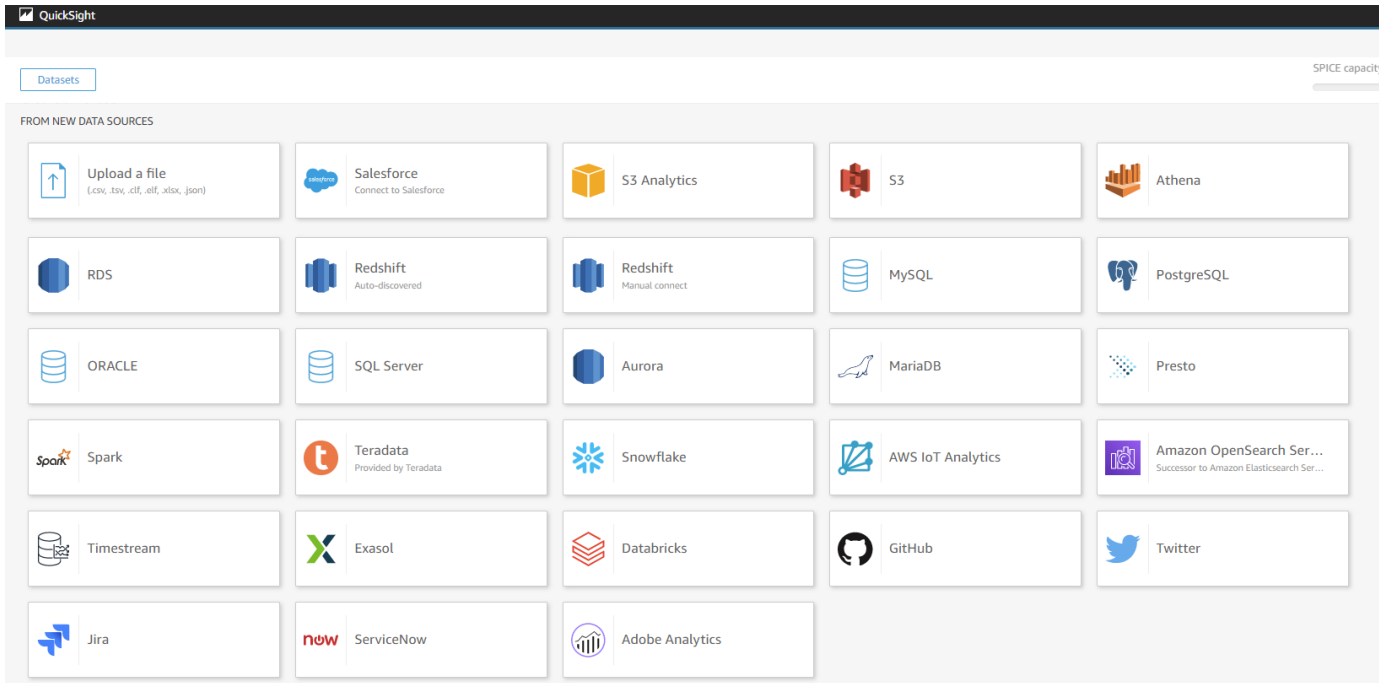
Página Inicial

A página inicial do QuickSight apresenta o menu de opções à esquerda, referente ao desenvolvimento de visualizações, e no canto superior direito, opções de gerenciamento e de conta, acessíveis a partir do avatar do usuário. Dentre estas opções encontram-se qual região o usuário está utilizando. Quanto às propriedades de desenvolvimento propriamente, as principais são *Datasets*, *Analyses* e *Dashboards*.



Datasets

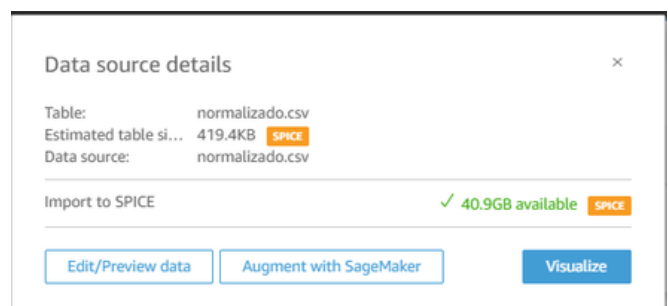
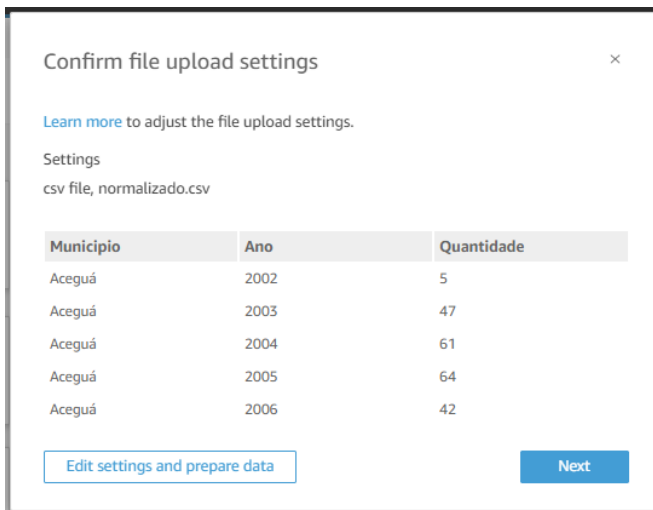
Correspondem aos conjuntos de dados utilizadas em análises. Nativamente, o QuickSight oferece diferentes conectores (*data sources*) que facilitam o consumo de dados de diferentes origens, especialmente de outros serviços AWS. Como é possível perceber na imagem abaixo, temos disponíveis conectores para RDS, RedShift, S3, Athena, Aurora, Spark, OpenSearch e, para alguns casos, há o recurso de *auto discovery*, ou seja, a descoberta automática dos recursos presentes na conta. Naturalmente existem outros conectores, como arquivos, bancos de dados relacionais específicos (Oracle, SQL Server, MariaDB, PostgreSQL) e outros serviços específicos, como ServiceNow e Jira, por exemplo.



Na sequência vamos detalhar alguns conectores que estão normalmente presentes em projetos de visualização de dados.

Arquivos

Vamos utilizar como exemplo a criação de um *dataset* para um arquivo CSV. Iniciamos o procedimento clicando em **Upload File** e, em seguida, escolhendo o arquivo de interesse. Feito isso, o QuickSight fará a leitura e irá apresentar informações sobre o arquivo, inclusive apresentando algumas linhas. Neste momento podemos optar por editar as propriedades da importação e preparação de dados (*Edit settings and prepare data*). A outra opção é avançar para o passo seguinte. Seja qual for a escolha, é possível editar o *dataset* a qualquer momento.



Ao acessar a tela de edição do *dataset*, temos acesso a vários recursos de preparação de dados. Perceba o menu de opções apresentado ao clicar sobre *dataset* na imagem que segue. Certamente uma das principais opções apresentadas é *Configure upload settings*, pois através dela podemos modificar parâmetros de importação, como delimitador de coluna, linha de início, qualificador de texto e a presença ou não de cabeçalho. Outros recursos importantes disponíveis são adicionar novos *datasets*, campos calculados e filtros.

Ao adicionar um novo *dataset*, será necessário associá-lo por meio de *inner join* ou *outer join*. Logo, é necessário que haja relação lógica entre as fontes permitindo estabelecer uma regra de junção que faça sentido.

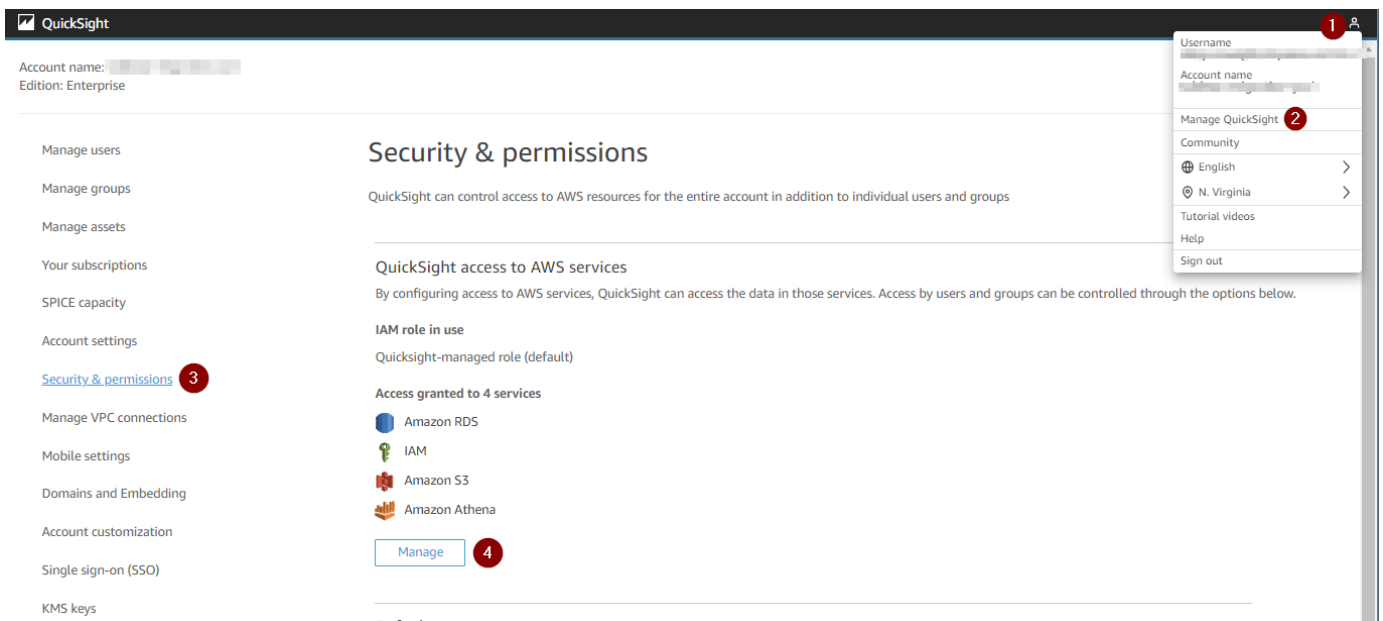
Campos calculados (*calculated fields*), por sua vez, correspondem a novas colunas que podemos criar por meio da aplicação de funções e/ou combinação de colunas já existentes. Já quanto a filtros (*filters*), são critérios que restringem os dados que efetivamente estarão disponíveis em todas as análises que fizerem uso do *dataset*. Por fim temos os *Excluded Fields*, campos que estão presentes no *dataset* mas que não devem ser disponibilizados para uso em análises.

The screenshot shows the QuickSight interface for editing a dataset named 'normalizado'. The left sidebar contains sections for 'Fields' (with 'Add calculated field' and 'Augment with SageMaker' buttons), 'Focus' (set to 'All fields'), and 'Excluded fields' (set to 'No fields excluded'). The main area displays a data preview table with columns 'Município', 'Ano', and 'Quantidade'. A context menu is open over the 'normalizado.csv' file icon, showing options: 'Configure upload settings', 'Update file', 'View details', 'Show table preview', and 'Remove'.

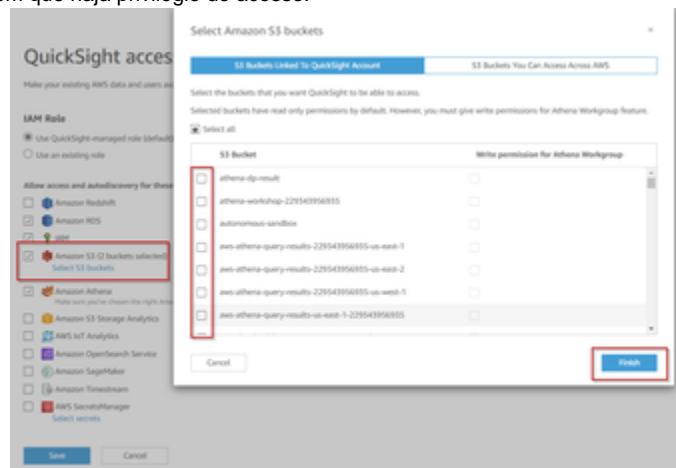
Município	Ano	Quantidade
Aceguá	2002	5
Aceguá	2003	47
Aceguá	2004	61
Aceguá	2005	64
Aceguá	2006	42
Aceguá	2007	72
Aceguá	2008	72
Aceguá	2009	71
Aceguá	2010	57
Aceguá	2011	61
Aceguá	2012	57
Aceguá	2013	61

Arquivos podem ser lidos diretamente a partir do S3. Aqui, contudo, existem alguns pontos adicionais para levar em consideração:

- O QuickSight deve ter acesso ao *bucket* de origem dos arquivos. Uma das formas de liberar acesso, tendo os devidos privilégios, é por meio do menu *Security & Permissions*, disponível nas opções de gerenciamento da conta.



Em *Manage*, deve-se escolher qual(is) *buckets* serão utilizados pelo QuickSight. Naturalmente são listados apenas os *buckets* da mesma conta do QuickSight ou de outras contas em que haja privilégio de acesso.



- Para que o QuickSight tenha condições de ler o conteúdo do S3, será necessário criar arquivo de [manifesto](#), listando quais arquivos do *bucket* devem ser lidos. Este arquivo poderá estar no S3 ou ser importado no momento da criação do *dataset*.



Bases de dados

A conexão com Sistemas Gerenciadores de Banco de Dados, seja *on-premise* ou gerenciados, segue o mesmo padrão. A título de exemplo, iremos apresentar a conexão com uma base de dados MySQL e, para tal, utilizaremos o conector específico para este SGBD.

Observe que após informarmos os dados de conexão e avançarmos, o QuickSight lista as tabelas existentes na base de dados e, destas, podemos escolher apenas uma.

New MySQL data source

Data source name

mysql-loja

Connection type

Public network

Database server

hive-metastore.cnpzwadswd70.us-east-1.rds.amazonaws.com

Port

3306

Database name

programa_bolsas

Username

programa_bolsas

Password

••

Validated

Enable SSL

Create data source

Choose your table

mysql-loja

Tables: contain the data you can visualize.

AUTOR

EDITORIA

ENDERECO

LIVRO

TbDependente

TbEstoqueProduto

Edit/Preview data

Use custom SQL

Select

Suponha que tenhamos escolhido a tabela *Autor* e após, clicado em *Edit/Preview Data*. Algumas opções novas são apresentadas, sendo uma delas o modo de consulta. Perceba que podemos alternar entre *SPICE* e *Direct Query*. Se você escolher *SPICE*, haverá a importação dos dados e a cada consulta realizada pelo QuickSight, quem irá responder será a *engine*. Se *Direct Query* for marcado, toda consulta do QuickSight irá disparar um comando SQL diretamente ao servidor MySQL.

QuickSight

AUTOR

Fields: All fields included

Add calculated field

Search fields

Fields

All fields

Select All Name

Colaborator

Name

Nascimento

Excluded fields: No fields excluded

Filters: No filters applied

Query mode

SPICE

Direct query

Data

AUTOR

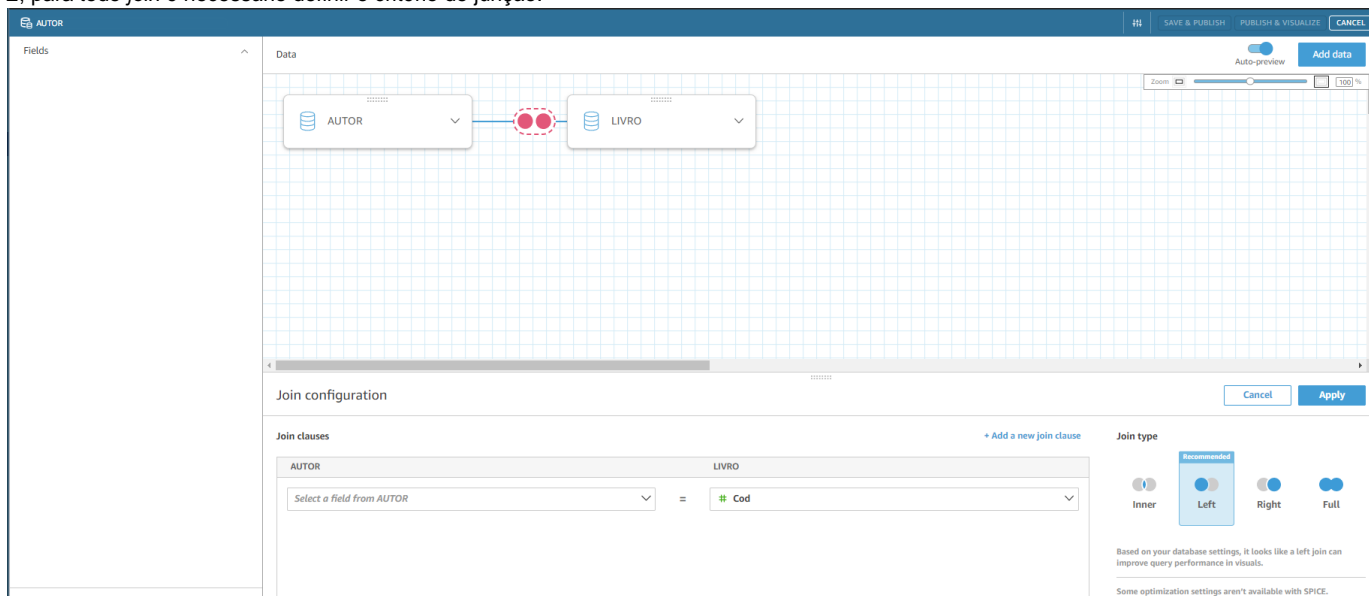
Dataset

Colaborator	Name	Nascimento
3	AADNOL, Benit	1976-04-01T00:00:00.000Z
4	ABBASCHIAN, R	1981-11-23T00:00:00.000Z
5	ABE, Jan Minors	1983-08-12T00:00:00.000Z
6	ABRAHAM, Martin A	1983-01-04T00:00:00.000Z
7	ABRAHCOUL, Andre A.	1983-07-27T00:00:00.000Z

Outro ponto interessante é que podemos adicionar novas tabelas ao nosso *dataset*. Para ilustrar, vamos adicionar a tabela *LIVRO*. Faremos isso clicando em *Add data* e escolhendo o *data source* que acabamos de cadastrar (nosso servidor MySQL). Aí é só indicar as tabelas desejadas e clicar em *Select*.

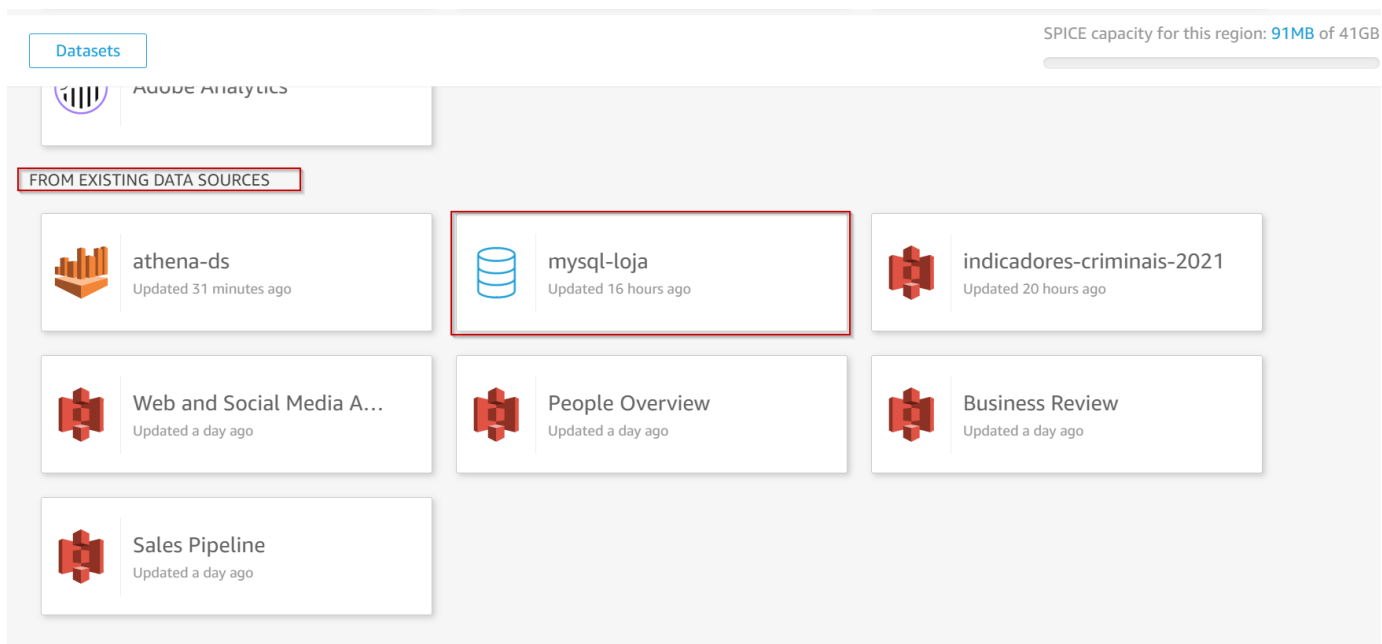


Observe que ao adicionar a(s) novas origens de dados ao *dataset*, você deverá realizar *join* entre eles, escolhendo entre *inner* ou *outers joins*. E, para todo *join* é necessário definir o critério de junção.

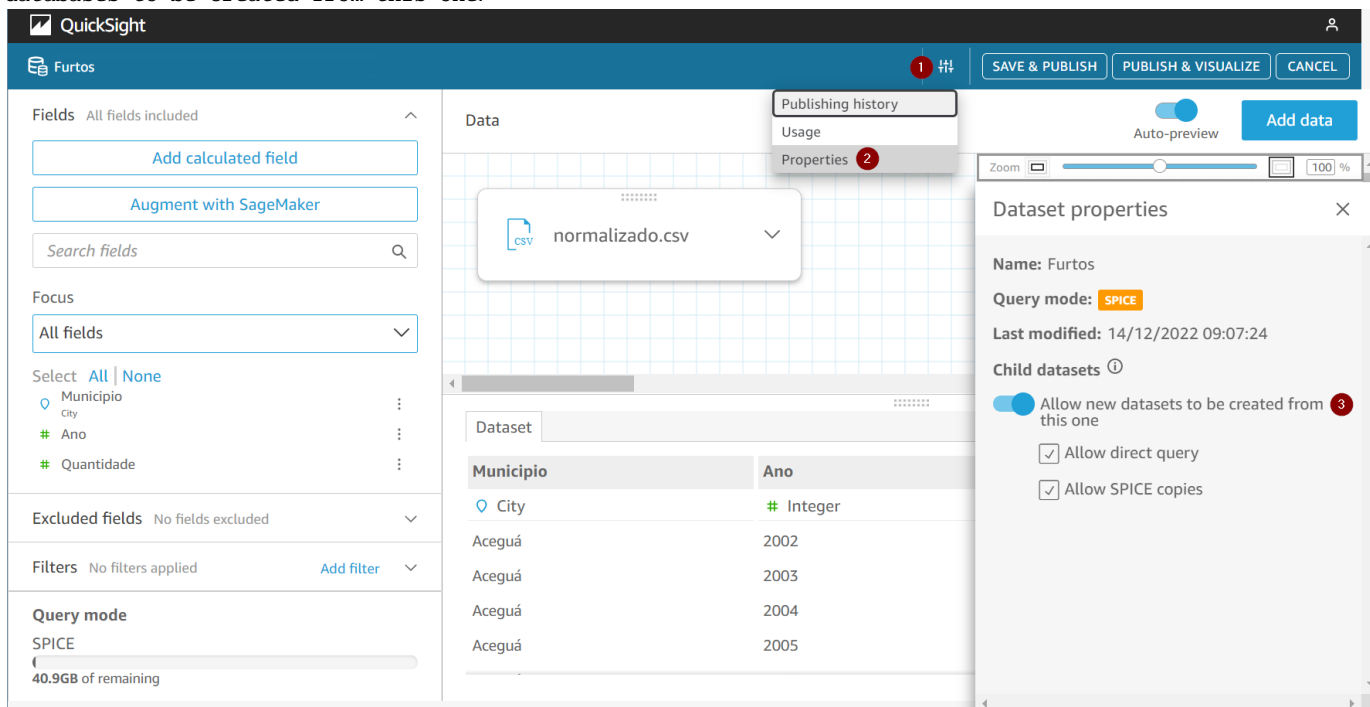


Criar datasets a partir de *data sources* ou *datasets* já existentes

É importante perceber que, após criar o *data source*, ele estará disponível para a criação de novos *datasets*. No exemplo anterior, quando criamos uma conexão ao servidor MySQL para extrair um *dataset* referente às tabelas *AUTOR* e *LIVRO*, acabamos por gerar dois importantes artefatos: o *data source* e o *dataset*. Se for necessário adicionar um novo *dataset* cuja origem seja nosso servidor MySQL, nós podemos criar a partir do *data source* já criado. Veja:



De modo similar, podemos criar um novo *dataset* derivando de um preexistente. Neste caso, é preciso que o *dataset parent* (origem) tenha habilitado as permissões para a criação de *datasets* filhos. Tal configuração está disponível no menu *Properties*, opção *Allow new databases to be created from this one*.



Com as devidas permissões, a criação de um *dataset* derivado pode ser feito através da opção *Use in new dataset* no menu de opções ou via botão *Add Data* quando estamos editando/criando um *dataset*.

Name	Owner	Last Modified
Furtos	Me	a few seconds ago
Business Review	Me	a day ago
Sales Pipeline	Me	a day ago
People Overview	Me	a day ago
Web and Social Media Analytics	Me	a day ago

Dimensões e métricas

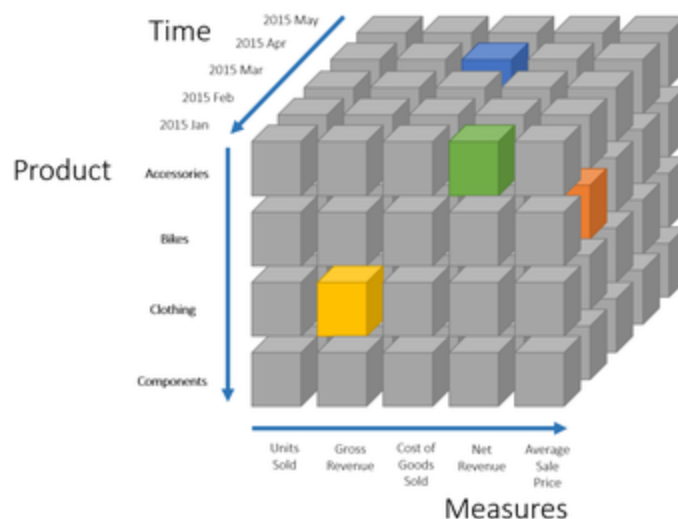
Ao utilizar um *dataset* em uma análise, as colunas são classificadas entre *dimensions* (*dimensões/categorias*) e *measures* (*métricas*). Colunas *dimensions* são normalmente do tipo texto ou data e servem ao propósito de qualificar/particionar métricas. Já as *measures* são informações numéricas utilizadas para medição, comparação ou agregação.

Dimensões

- São dados de natureza qualitativa e representam atributos de dados
- São normalmente textuais
- São utilizadas para “quebrar” valores de métricas

Métricas

- São valores quantitativos (numéricos)
- Estão sempre associados com um agregador, orientado pelas dimensões utilizadas
- Podem ser de três tipos
 - **Aditivas:** podem ser sumarizadas através de todas as dimensões, sem restrição. Um exemplo poderia ser o valor de venda de um e-commerce.
 - **Não-aditivas:** **não** podem ser sumarizadas através de nenhuma dimensão. Um exemplo seria percentual de desconto de venda de produtos, por exemplo.
 - **Semi-aditivas:** podem ser sumarizadas através de algumas dimensões, já outras não. É o caso do saldo bancário, por exemplo.



Funções

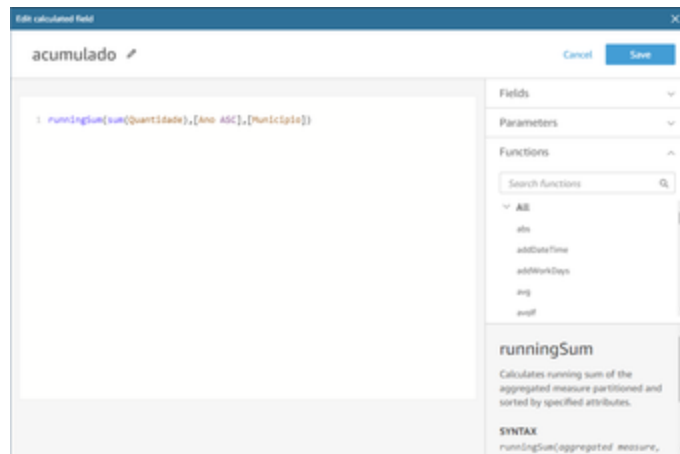
Temos a disposição um expressivo conjunto de funções para utilizar em nossas visualizações e *datasets*. As funções são utilizadas principalmente na criação de campos calculados. A lista completa de funções pode ser consultada na documentação oficial do produto. Além de funções, temos a disposição operadores lógicos e matemáticos. Na sequência apresentaremos exemplos para ilustrar a sintaxe utilizada pela QuickSight na construção de expressões. Perceba que dimensões e métricas são referenciadas nas funções por meio de chaves {}, sempre que o nome do campo for composto por mais de uma palavra. Comentários podem ser adicionados utilizando os símbolos triviais de comentário de várias linhas `/**/`

```
ifelse ( /* NE REGION*/
    locate('New York, New Jersey, Connecticut, Vermont, Maine, Rhode
Island, New Hampshire',{State}) > 0,
    'Northeast',
    /* SE REGION*/
    locate('Georgia, Alabama, South Carolina, Louisiana',{State}) > 0,
    'Southeast',
    'Other Region')
```

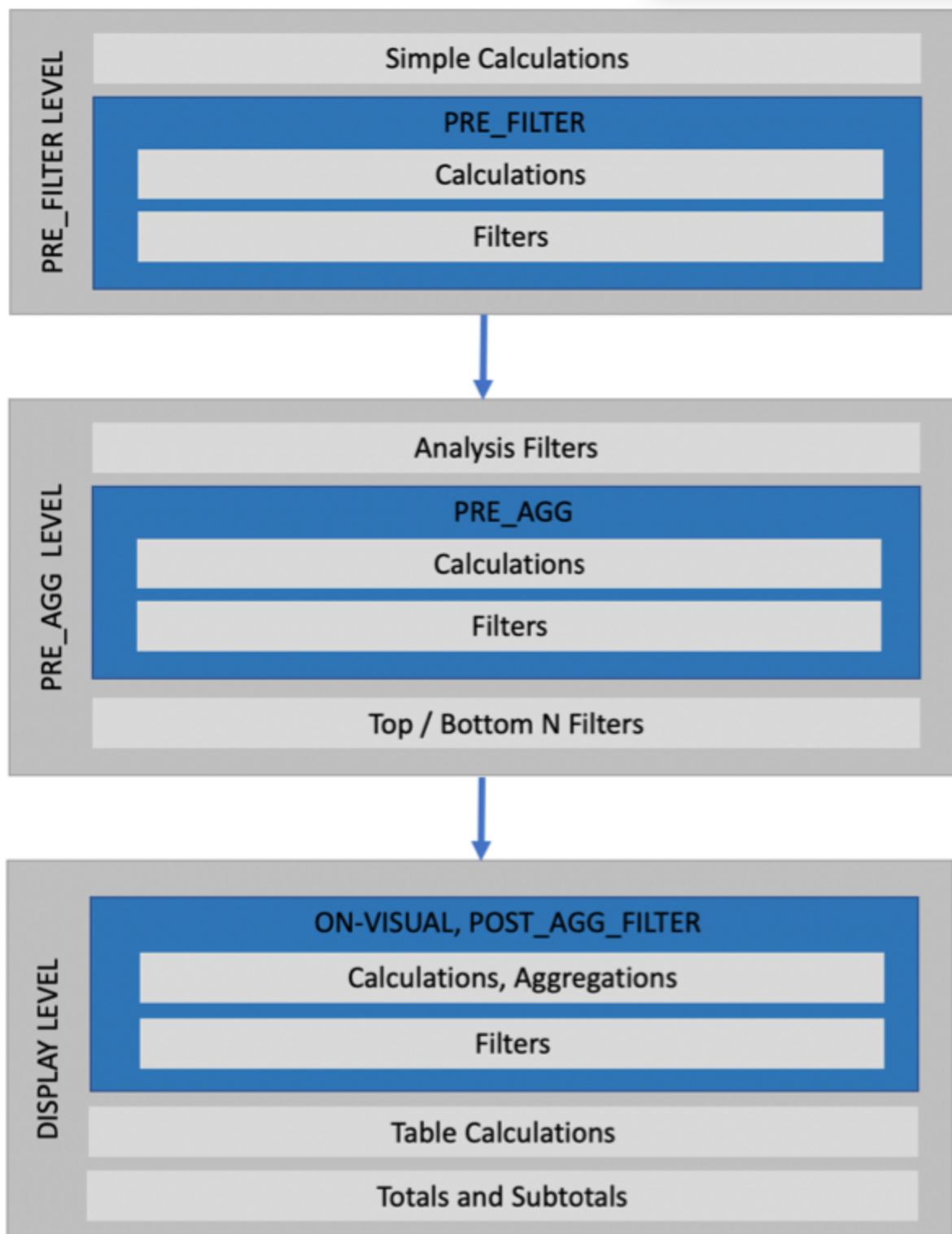
<https://docs.aws.amazon.com/quicksight/latest/user/functions-by-category.html>

Campos calculados

Campos calculados podem ser criados a nível de *datasets* e de *analises*. Por meio deles adicionamos novas colunas com possibilidade de aplicar expressões, funções e literais na geração dos novos valores. Abaixo apresentamos a tela de edição de campos calculados, na qual temos à direita ajuda com a lista de **Fields**, **Parameters** e **Functions** disponíveis.



Abaixo segue a ordem em que filtros e cálculos são executados pelo QuickSight.

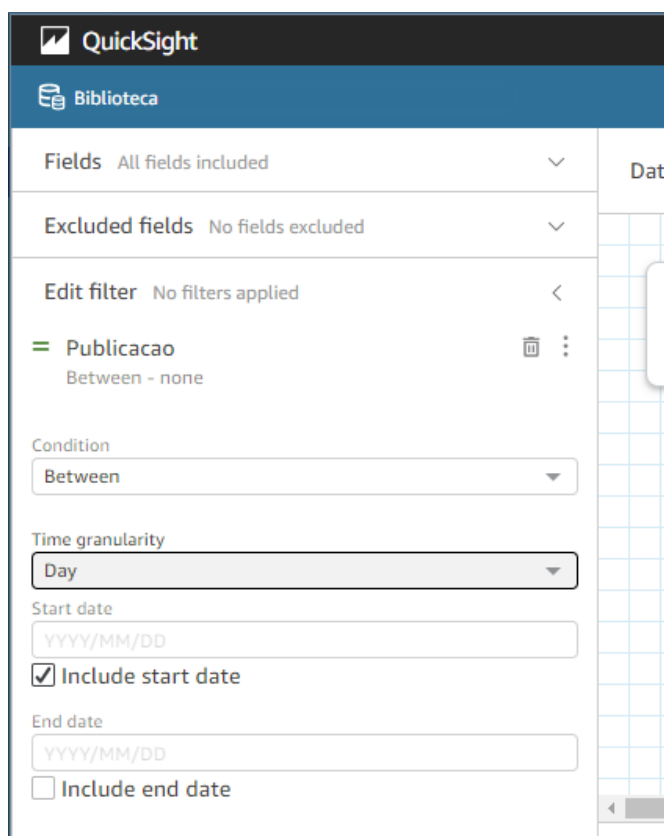
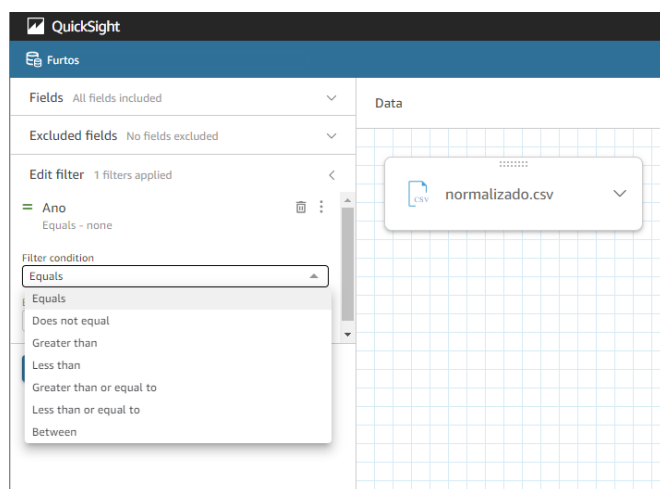
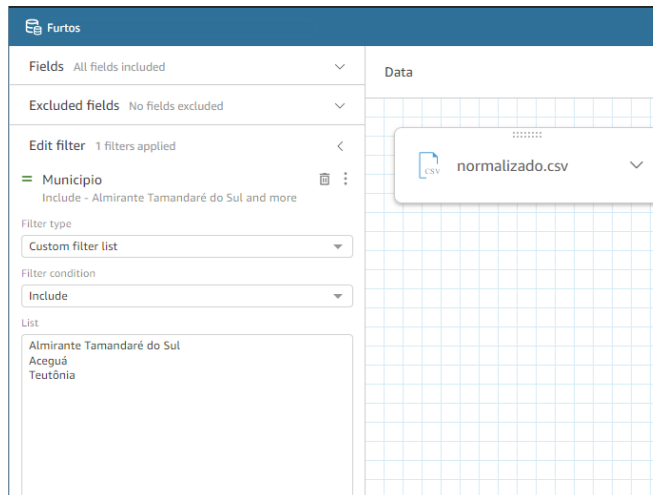
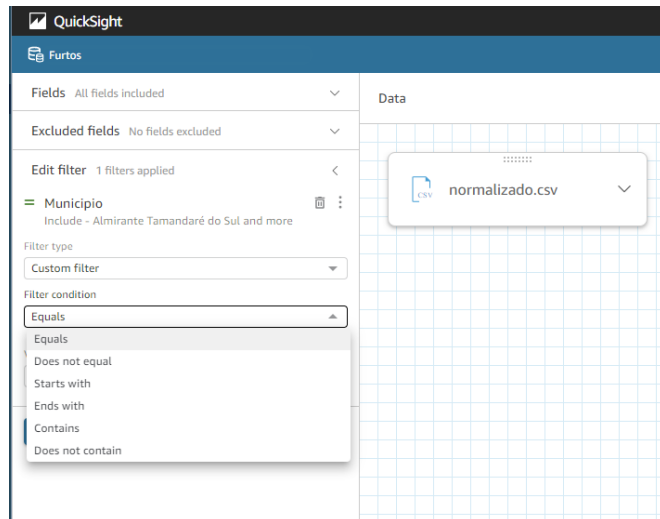


Filtros

Filtros podem ser adicionados aos *datasets* e às *análises*. Quando aplicados ao *datasets*, tem seu efeito propagado para qualquer *dataset* derivado, visualização ou dashboard que o utilize. Quando o filtro é adicionado em uma análise, podemos definir seu escopo como sendo restrito a uma visualização, algumas visualizações, a todas as avaliações que usam o *dataset* ou todas as visualizações aplicáveis. Cada filtro considera apenas um campo, seja ele regular ou calculado.

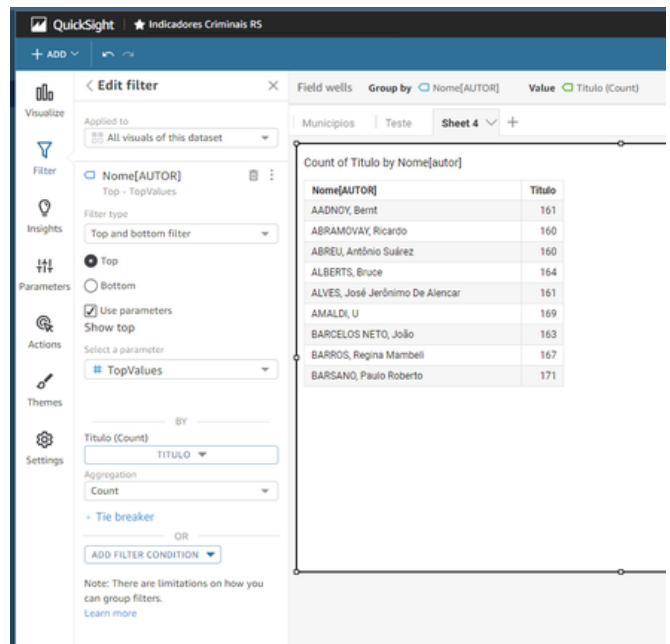
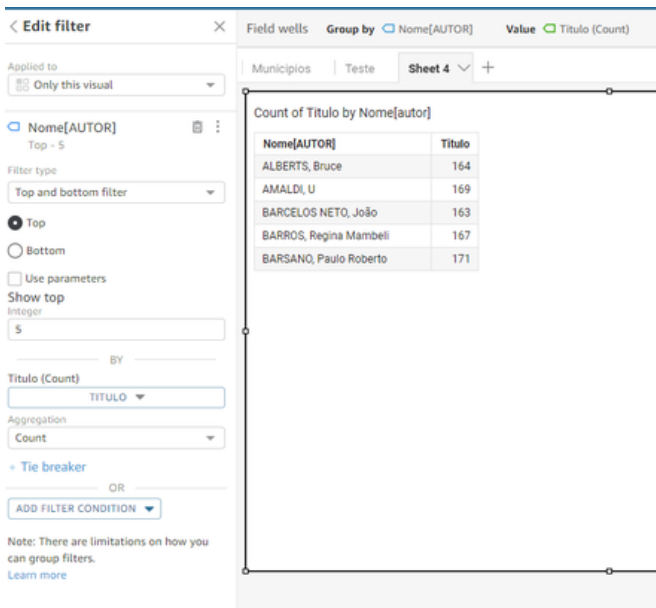
Filtros em *Datasets*

Podem ser filtros de texto, numéricos ou de data. Para cada tipo são disponibilizados critérios específicos.



Filtros em Análises

Podem ser filtros de texto, numéricos ou de data. Além disso, podem ser agrupados com operadores **AND** e **OR** e aplicados em cascata. O uso do **OR** ocorre dentro de um filtro a partir da combinação de dois ou mais critérios (*compound filter/filter group*). Entre os filtros de primeiro nível, eles são aplicados utilizando o conector lógico **AND**.



Filtros em Cascata

Filtros em cascata são aqueles que permitem restringir o valor de outros controles para um subconjunto de valores. Por exemplo, ao selecionar um país, apresentar apenas no filtro de “unidades federativas” os respectivos estados do país selecionado. Para isso, é necessário criar *Actions* no QuickSight. Veremos o funcionamento de *actions* na sequência.

Também podemos ativar um filtro em diversas visualizações a partir de um *widget* chamado *sheet control* associado com parâmetros. Para tal, é preciso:

- Criar um parâmetro
- Adicionar um controle ao parâmetro, permitindo que o usuário possa alterar seu valor
- Adicionar um filtro que usa o parâmetro em suas condições.

Neste caso, ao modificar o valor do parâmetro através do controle, todas as visualizações que dele fazem uso serão atualizadas. É o que descreveremos a seguir quando o tema Parâmetros será abordado.

Parâmetros

Parâmetros funcionam como variáveis globais em uma análise. Podem ter valor fixo ou dinâmico. Após criados, podem ser referenciados em filtros, controles, campos calculados e ações customizadas. Ao adicionar um controle sobre o parâmetro, por exemplo, podemos permitir que o usuário altere seu valor dentro de intervalos ou conjuntos pré-definidos. Ao receber a alteração, o efeito é imediatamente percebido em todos os locais que referenciam o parâmetro associado.

QuickSight

★ Indicadores Criminais RS

+ ADD

Visualize

Filter

Insights

Parameters

Actions

Themes

Settings

Parameters

Search parameters

TopValues

Add control

Edit parameter

Delete parameter

Field wells

Group by Nome[AUTOR]

Value Titulo (Count)

Municipios | Teste | Sheet 4

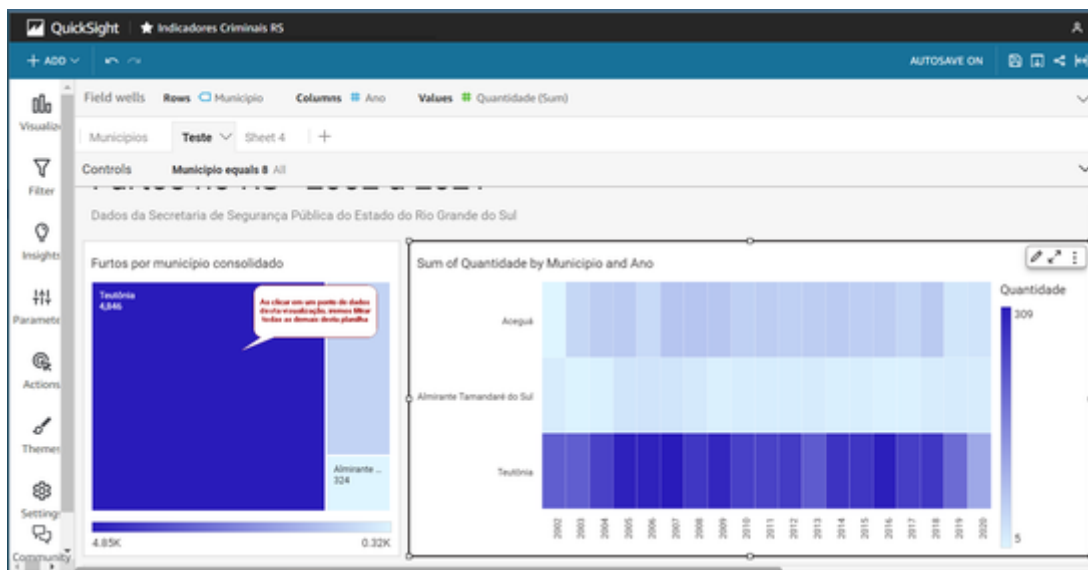
Count of Titulo by Nome[autor]

Nome[AUTOR]	Titulo
AADNOY, Bernt	161
ALBERTS, Bruce	164
ALVES, José Jerônimo De Alencar	161
AMALDI, U	169
BARCELOS NETO, João	163
BARROS, Regina Mambeli	167
BARSANO, Paulo Roberto	171

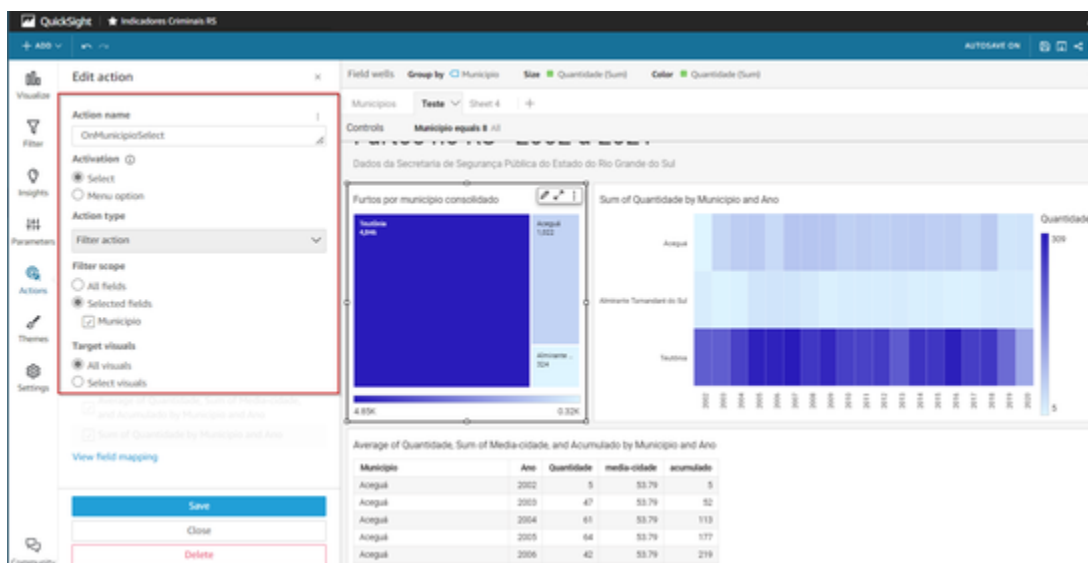
Valores de parâmetro podem ser incorporados a textos e descrições utilizando a sintaxe `${nome do parâmetro}`.

Ações

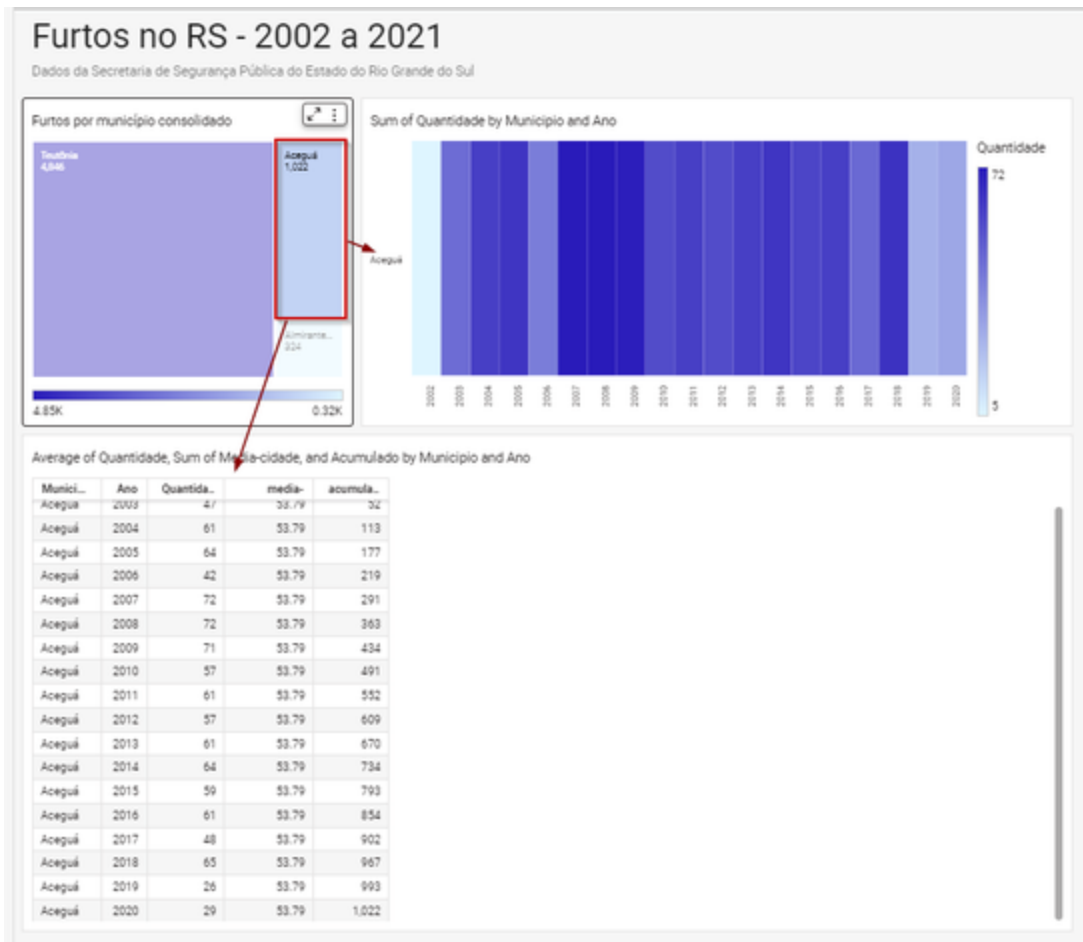
Suponha ser necessário criar uma interação com o leitor de um *dashboard* que permita refinar diferentes visualizações a partir da escolha de um ponto de dados em um gráfico. Para ilustrar a situação, vamos considerar que o primeiro gráfico apresenta a consolidação de furtos em período 2002-2022, agrupado por município. Já o gráfico a sua direita, discrimina o total de furtos por município e ano. O objetivo aqui proposto será restringir os dados do gráfico à direita para o município escolhido no gráfico da esquerda e, para isso, faremos uso de uma **Action**.



As *Actions* são definidas no escopo de uma visualização (*visual*). Logo, precisamos selecionar o primeiro gráfico (que irá escutar a ação) e ir até o menu *Actions*. Para obter o resultado que desejamos, devemos utilizar uma ação do tipo *Select*, a qual relaciona-se com os pontos de dados. A ação escolhida será *Filter Action*, aplicada sobre a coluna *Município*. Como alvo deste filtro, iremos informar a opção *All visuals*, o que implica em filtrar todas as visualizações que utilizam o *dataset* na planilha.

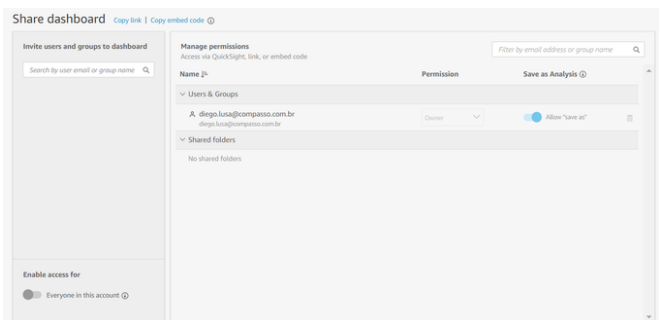
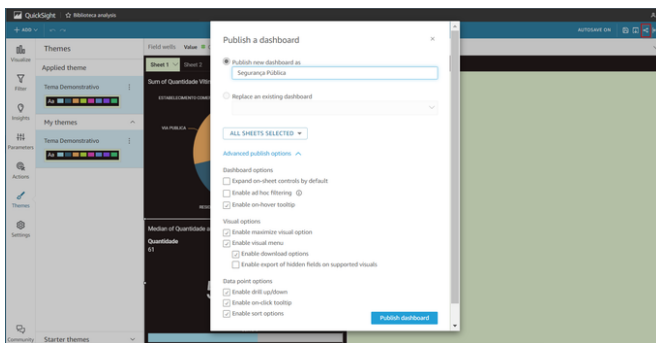


Criada a *action* e publicado nosso *dashboard*, iremos perceber que, ao clicar sobre a área do primeiro gráfico que corresponde a um município, as demais visualizações automaticamente restringem seu valor para o alvo da seleção. É o que a imagem a seguir busca ilustrar.



Publicar *dashboards*

Ao publicar uma análise, ela torna-se um *dashboard*, que pode ser compartilhado com usuários da conta do QuickSight ou anônimos. O processo de publicação é simples, compreendendo definir um nome para o *dashboard*, caso seja novo, ou substituir um já existente. Adicionalmente podemos escolher quais planilhas estarão disponíveis e algumas configurações mais específicas, como possibilidade de ordenação, tooltips, drill down/up, entre outras.



<https://docs.aws.amazon.com/quicksight/latest/user/creating-a-dashboard.html>

Recursos Adicionais

- [Amazon QuickSight Gallery - Business Intelligence Service - Amazon Web Services](#)
- <https://catalog.workshops.aws/quicksight/>
- [Amazon QuickSight Documentation](#)