**Olá, estudante!**

Sinta-se à vontade para aproveitar este conteúdo de criação de mapas!

Muitos conjuntos de dados possuem informações relativas à localização geográfica. Por exemplo, em dados de aplicativos de transporte, podem ser armazenados a localização do local de partida e o destino de cada corrida; dados demográficos podem conter informações relativas a países, estados, cidades ou regiões, como renda por habitante, média de idade; dados governamentais podem conter localização de obras como pontes, estradas, edificações e outros. Em todos os casos, é útil visualizar esses dados em um contexto geoespacial. Para isso, mapas são excelentes ferramentas, pois são fáceis de compreender e podem juntar informações de várias formas.

Nesta aula, você irá aprender a utilizar a plataforma Kibana para criar diferentes tipos de mapas para representar dados geolocalizados.

**Objetivos**

* Aprender a criar mapas representando a localização de elementos específicos utilizando Kibana;
* Aprender a criar mapas representando dados agrupados utilizando Kibana;
* Aprender a criar mapas coropléticos utilizando Kibana.

## Tópico 1 – Criando mapas em camadas com Kibana

**OBJETIVOS**

* Aprender a adicionar camadas de informação em mapas com Kibana;
* Aprender a criar camadas com localização de documentos carregados no Kibana;
* Aprender a criar camadas com dados agrupados em grade.

Mapas são muito úteis para representar informações espaciais, permitindo uma fácil compreensão desse contexto. Atualmente, serviços on-line de mapas como Google Maps ou OpenStreetMap difundiram o acesso a mapas interativos via Web, permitindo que qualquer usuário possa utilizar mapas para diversos fins, seja para localização, criação de rotas e muitos outros. Além de atender a necessidades de usuários comuns, mapas também são muito úteis para cientistas de dados que trabalham com informações georreferenciadas, ou seja, informações que possuem coordenadas geográficas que permitem a sua localização em um mapa. Esses esquemas proporcionam tanto uma análise aprofundada dos padrões espaciais presentes em dados, como também facilitam a comunicação dos padrões encontrados, pois mapas são muito intuitivos de se compreenderem.

Para visualizar dados georreferenciados adequadamente, geralmente são criados mapas com múltiplas camadas, cada uma mostrando um tipo de informação. A plataforma Kibana permite criar mapas com múltiplas camadas de informação e explorá-los de forma dinâmica, controlando o zoom e a posição.

Neste tópico, você aprenderá os princípios básicos para mapas com Kibana, poderá criar camadas de informação a partir de dados carregados no Kibana e camadas com dados em grade. Vamos lá?

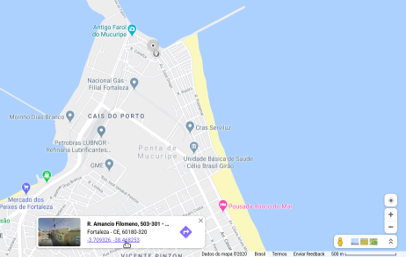
## 1.1 - Mapas e dados georreferenciados

Você já deve estar acostumado a ter acesso a mapas, sejam os tradicionais, na forma de imagens estáticas, ou os digitais interativos, que permitem uma navegação dinâmica, modificando a localização e o zoom. Um mapa nada mais é do que uma representação gráfica de uma região, geralmente em forma bidimensional. Eles são importantes devido à necessidade humana de conhecer o seu espaço e a localização de elementos nesse espaço. Para compreender mais o assunto, considere este contexto:

Como você já deve saber, a Terra tem o formato aproximado de uma esfera e gira em torno de um eixo. Os pontos onde o eixo cruza a Terra são chamados polos, norte e sul. A Terra é dividida em hemisférios norte e sul pela linha do Equador, que está equidistante dos polos. Para especificar a localização exata de um ponto na superfície da terra são utilizadas três informações:

1. Longitude: onde se está localizado ao longo da linha do equador;
2. Latitude: quão próximo se está de um dos polos quando nos movemos, perpendicularmente, em relação ao equador;
3. Altitude: a que distância se está do centro da Terra.

A seguir, confira uma imagem que representa a latitude e a longitude.

Figura 1 - Latitude e Longitude no Google Maps

Longitude, latitude e altitude são especificadas em relação a um sistema de referência chamado datum. O datum especifica propriedades como a forma e o tamanho da Terra, bem como a localização da longitude zero, latitude e altitude. Um dado amplamente utilizado é o World Geodetic System (WGS) 84, usado pelo Global Positioning System (GPS).

Dados georreferenciados são expressados, geralmente, em termos de latitude e longitude, que permite referenciar o dado em relação a um ponto na superfície terrestre. A figura apresentada exemplifica a seleção de um ponto na plataforma Google Maps. Perceba em destaque na parte inferior da figura a indicação da latitude e da longitude do ponto.

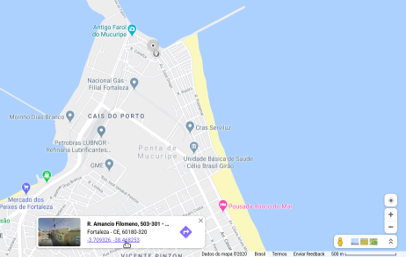
## 1.1 - Mapas e dados georreferenciados

Você já deve estar acostumado a ter acesso a mapas, sejam os tradicionais, na forma de imagens estáticas, ou os digitais interativos, que permitem uma navegação dinâmica, modificando a localização e o zoom. Um mapa nada mais é do que uma representação gráfica de uma região, geralmente em forma bidimensional. Eles são importantes devido à necessidade humana de conhecer o seu espaço e a localização de elementos nesse espaço. Para compreender mais o assunto, considere este contexto:

Como você já deve saber, a Terra tem o formato aproximado de uma esfera e gira em torno de um eixo. Os pontos onde o eixo cruza a Terra são chamados polos, norte e sul. A Terra é dividida em hemisférios norte e sul pela linha do Equador, que está equidistante dos polos. Para especificar a localização exata de um ponto na superfície da terra são utilizadas três informações:

1. Longitude: onde se está localizado ao longo da linha do equador;
2. Latitude: quão próximo se está de um dos polos quando nos movemos, perpendicularmente, em relação ao equador;
3. Altitude: a que distância se está do centro da Terra.

A seguir, confira uma imagem que representa a latitude e a longitude.

Figura 1 - Latitude e Longitude no Google Maps

Longitude, latitude e altitude são especificadas em relação a um sistema de referência chamado datum. O datum especifica propriedades como a forma e o tamanho da Terra, bem como a localização da longitude zero, latitude e altitude. Um dado amplamente utilizado é o World Geodetic System (WGS) 84, usado pelo Global Positioning System (GPS).

Dados georreferenciados são expressados, geralmente, em termos de latitude e longitude, que permite referenciar o dado em relação a um ponto na superfície terrestre. A figura apresentada exemplifica a seleção de um ponto na plataforma Google Maps. Perceba em destaque na parte inferior da figura a indicação da latitude e da longitude do ponto.

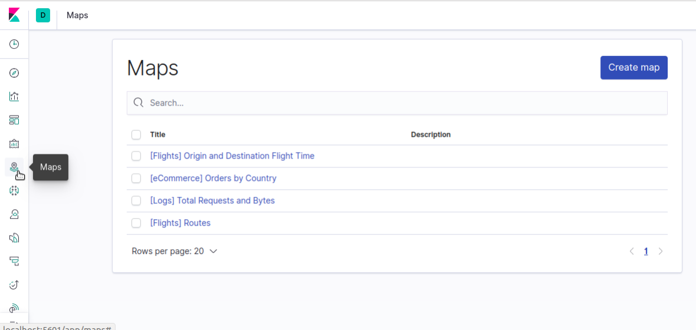
## 1.2 - Elastic Maps do Kibana

Como você já deve saber, a plataforma Kibana permite criar variados tipos de visualizações e reuni-las em um dashboard interativo, facilitando a exploração dos dados. A maioria das visualizações no Kibana são criadas seguindo a lógica de Metrics e Buckets. Em Metrics são selecionados os valores a serem exibidos e em Buckets são selecionadas as formas que esses valores serão agrupados. Quase todas as visualizações seguem esse modelo. Porém, os mapas são um tipo de visualização especial e seguem uma lógica diferente. Por isso, o Kibana possui uma ferramenta própria para criação de mapas, chamada Elastic Maps, que funciona adicionando camadas (layers).

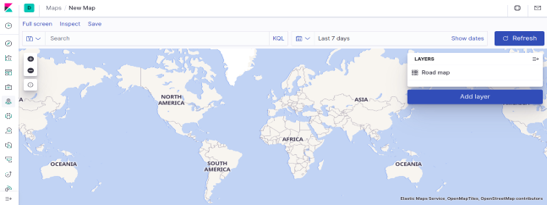
Os dados representados em uma camada podem vir de diferentes fontes, desde conjuntos de dados do Elasticsearch, serviços de mapas Web e outros. O Elastic Maps permite também representar dados de diferentes formas. Por exemplo, em dados de voos, você pode mostrar quais aeroportos têm atrasos de voo mais longos usando círculos de pequeno a grande. Ou, em dados de acesso a uma página Web, você pode mostrar os países que mais acessaram a página, colorindo os países do claro ao escuro. Iremos explorar essas ideias construindo um mapa com múltiplas camadas a partir de um conjunto de dados de amostra oferecido pelo Kibana. Inicialize o Elasticsearch e o Kibana e abra a página do Kibana no seu navegador usando este endereço <http://localhost:5601/>.

Neste conteúdo, iremos analisar o conjunto de dados de voos fictícios fornecido como exemplo pelo Kibana. Caso você não tenha instalado esses dados no Kibana, você deverá fazê-lo agora. No centro da página inicial do Kibana, selecione a opção **Load a data set and a Kibana dashboard** e, no painel **Sample flight data**, acione o botão **Add Data**.

Localize e acione o ícone Maps na barra de navegação da tela para abrir a página de listas de mapas do Kibana, similar ao mostrado na figura abaixo. Para criar um mapa novo, escolha a opção **Create map**.

Figura 2 – Lista de mapas criados no Kibana

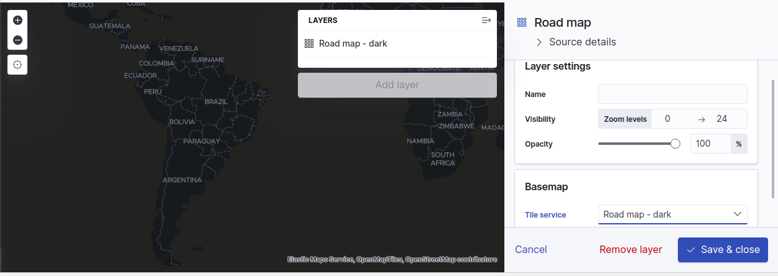
Com essas ações, você deverá ter aberto a página de criação de mapas do Kibana. Na próxima página, confira uma imagem que representa isso.

Figura 3 – página de criação de mapas do Kibana

Note que é exibido um mapa mundial, com as fronteiras e nomes dos países, estados e cidades, muito similar a mapas exibidos na Web, como Google Maps ou OpenStreetMap. Perceba que a forma de navegar por esse mapa é idêntica. Você pode movimentar a posição visualizada, aumentar e reduzir a ampliação (zoom) utilizando o scroll ou acionando os botões de + e -, disponíveis no lado esquerdo da tela. A latitude e longitude é indicada no lado inferior direito da tela, além do nível de zoom, que é um valor que vai de 0 (mínimo zoom) até 24 (máximo zoom). Se você desejar obter uma localização específica, pode-se acionar o botão **Go to** e selecionar um local indicando a sua latitude e longitude. Esse tipo de mapa, muito utilizado na Web, é chamado de tile map.

Tile maps são formados por dezenas de imagens chamadas tiles, que são agrupadas lado a lado e carregadas à medida que são necessárias, como quando é feito um zoom ou mudança de posição. Note que há um painel no lado direito do mapa chamado **Layers** (camadas). Como já foi comentado, o Kibana cria mapas através de camadas de informação. Esse mapa exibido na tela, nada mais é que a primeira camada (chamada Road map), que é um tile map disponibilizado pelo Kibana e servirá como mapa base para outras camadas, associadas a dados. Ao selecionar o nome da layer Road map, como mostra a figura, será exibida uma lista de opções que incluem ocultar (Hide layer), editar (Edit layer), clonar (Clone layer) e remover (Remove layer). A seguir, vamos conferir como podemos acessar o painel para edição da camada de tile map.

Ao localizar a lista de opções para ocultar, editar, clonar e remover, escolha a opção Edit layer para abrir um painel de edição. Uma camada de tile map não possui muitas possibilidades de edição. No painel **Layer settings** você pode configurar o nome da layer, a opacidade e os níveis de zoom em que ela é visível. Também é possível escolher outras opções de tile map clicando em **Tile service** no painel **Base map**. Uma lista de opções será exibida. A figura a seguir ilustra o tile map Road map - dark. Verifique.

Figura 4 – Editando a camada de tile map

Para esse exemplo, mantenha o Road map padrão. Já temos uma layer base de mapa, queremos adicionar layers que contenham informações sobre os dados. Vamos verificar como fazer isso!

## 1.3 - Adicionando uma camada com fonte de dados de documentos

Ao criar uma nova camada, primeiramente será necessário escolher uma fonte de dados para essa camada, que depende do objetivo da visualização. O Kibana oferece diferentes possibilidades de fontes de dados. Algumas delas são:

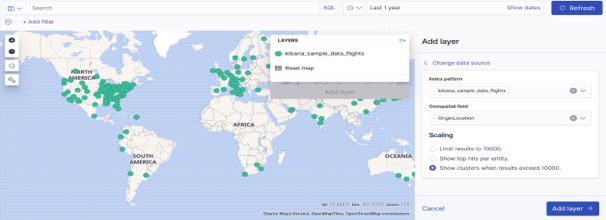
* Uploaded [GeoJSON](https://leadfortaleza.com.br/ead/glossary/GeoJSON" \t "_blank): carrega um arquivo do tipo [GeoJSON](https://leadfortaleza.com.br/ead/glossary/GeoJSON), um formato para representar dados geográficos baseado em JSON.
* Documents: localiza dados individuais, chamados no Kibana de documentos, de um conjunto de dados armazenado no Kibana.
* Grid aggregation: agrupa dados georreferenciados em uma grade, associando um valor a cada elemento da grade.
* Point to point: agrupa duas localizações (origem e destino) através de uma reta.
* EMS Boundaries: cria camadas com polígonos que representam fronteiras territoriais, a partir do Elastic Maps Service (serviço de mapas oferecido pelo Kibana). Permite associar dados a países, estados, cidades, ao invés de pontos.
* EMS Basemaps: cria camadas de tile maps, fornecidos pelo Elastic Maps Service.

Além dessas fontes de dados, existem outras opções mais avançadas para criar camadas oferecidas pelo Kibana. Primeiramente, você irá representar no mapa os locais de origem dos voos, a duração do voo e quanto o voo atrasou. Para isso, iremos criar cada uma com fonte de dados de documentos.

Como você já deve saber, o conjunto de dados que iremos explorar possui diversas informações fictícias sobre voos, como preço médio do bilhete, minutos de atraso, se o voo foi cancelado ou não, os nomes dos aeroportos de origem e destino, os países de origem e destino, o clima do local de partida e do destino, dentre outros. Caso precise, relembre os atributos do conjunto de dados utilizando a ferramenta Discover do Kibana em uma nova aba. Desses atributos, existem dois que são atributos de referência geográfica: OriginLocation e DestLocation, que contém a latitude e longitude da origem e do destino, respectivamente. Esses atributos podem ser utilizados para atribuir informações ao mapa. A seguir, iremos trabalhar com OriginLocation.

Para criar uma nova camada com fonte de dados de documentos, siga os seguintes passos:

1. Configure a janela de tempo para Last 1 year, para mostrar todos os voos do último ano.
2. Selecione a opção **Add layer** no painel de layers. Será exibido um painel mostrando diferentes fontes de dados para criar uma nova layer.
3. Escolha a opção **Documents** para criar uma camada que localiza os documentos (dados) armazenados em um conjunto de dados do Kibana.
4. Em [**Index**](https://leadfortaleza.com.br/ead/glossary/Index_Visualizacao_de_dados)**pattern**, selecione **kibana\_sample\_data\_flights**. Ou seja, o conjunto de dados dos dados de voos.
5. Em **Geospatial field**, selecione **OriginLocation**, indicando que serão representados os locais de origem dos voos. Note que já irão aparecer vários pontos no mapa, assim como ilustrado na figura abaixo.
6. Acione o botão **Add layer**.

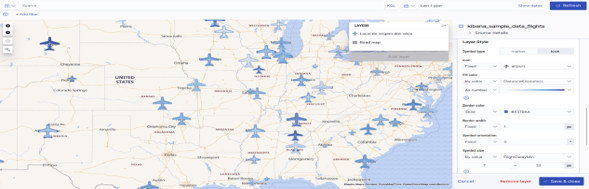
Figura 5 - Criando uma camada com fonte de dados de documentos do Kibana.

Pronto, a nova camada foi adicionada. Cada ponto no mapa representa um local de origem de um voo individual, mas nenhuma informação além da sua localização. Vamos alterar isso editando essa camada.

O painel de edição de uma camada criada com fonte de dados como documentos possui mais opções que uma camada de tile map. Iremos utilizar essas opções para adicionar informação à camada, como o nome da camada, as informações que deverão aparecer em tooltips quando determinado ponto estiver selecionado, a duração de um voo, o tempo de atraso do voo e o estilo da representação de forma geral. Para adicionar as informações, siga este passo a passo:

1. Acione a opção layer, depois a opção **Edit layer**.
2. Em **Layer settings**, configure o **Name** da visualização para Local de origem dos voos.
3. Configure **Opacitiy** para 65%. Esse parâmetro controla quanto opaca a camada é. Indo de 0% (totalmente transparente) até 100% (totalmente opaca).
4. Em [**Tooltip**](https://leadfortaleza.com.br/ead/glossary/Tooltip)**fields**, você irá escolher os atributos que serão exibidos quando um marcador do voo estiver selecionado. Acione a opção **Add** e selecione os atributos Origin (aeroporto de origem), OriginCityName (cidade de origem), DestCityName (cidade de destino), DestCountry (país de destino), DistanceKilometers (distância do voo em quilômetros), FlightDelayMin (atraso do voo em minutos), FlightTimeMin (tempo de voo em minutos). Pode escolher mais atributos se desejar.
5. Em **Layer Style** você terá diversas opções para configurar o visual da layer. Configure o **Symbol type** para icon, para mudar o símbolo de um ponto para um ícone.
6. Em **icon**, acione a barra à direita para abrir uma lista suspensa de variados ícones possíveis. Selecione o ícone de airport (aeroporto, em formato de avião).
7. Em **Fill color**, você pode selecionar uma cor fixa para o símbolo ou associar a cor dele a um atributo, criando uma escala de cor. Acione as listas suspensas e selecione as opções **By value**, o atributo DistanceKilometers, a opção By number e a escala de cor do branco ao azul. Isso significa que quanto maior a distância, mais próximo do azul escuro estará o símbolo. E quanto menor, mais próximo será do branco.
8. Em **Border color**, mantenha em Fixed value e selecione um azul claro no seletor de cor ou use o código #4379AA. Note que também é possível configurar a espessura da borda em **Border width** e a inclinação do símbolo em **Symbol orientation**. Por ora, deixaremos sem modificação neste exemplo.
9. Em **Symbol size**, selecione as opções ByValue, o atributo FlightDelayMin e mantenha o intervalo em 7 →32 px. Isso significa que o tamanho do símbolo estará associado ao tempo de atraso do voo. Ou seja, voos que mais atrasaram terão símbolos maiores. Em **Label** é possível mostrar um valor escrito na visualização. Para este exemplo, vamos manter como está.
10. Acione o botão **Save** **&** **close**.

Prontinho! Se você seguiu os passos corretamente, o mapa resultante deverá ser similar ao da figura apresentada na próxima página.

Figura 6 - Associando informações a uma camada de documentos

Perceba que em níveis maiores de zoom, esses símbolos fornecem informações importantes e de fácil leitura. Porém, quando o zoom é pouco, e se possui uma visão geral do mapa, são mostrados muitos dados de uma vez e os ícones que indicam os voos ficam amontoados, dificultando a análise. Uma forma de contornar esse problema é configurando a visibilidade da camada, para que ela seja visível apenas em níveis maiores de zoom. Para fazer isso, siga estes passos:

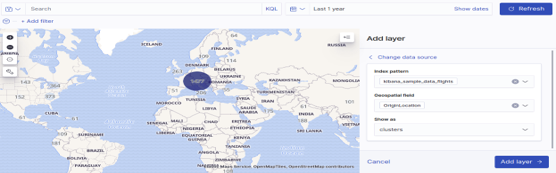
1. No painel **Layers**, acione a camada Local de origem dos voos e selecione **Edit layer**.
2. Na painel **Layer settings**, no campo **Visibility**, selecione o intervalo de **Zoom levels** entre 6 e 24. Isso significa que a camada só será vista quando o nível de zoom for maior que 6.
3. Acione o botão **Save** **&** **close**.

Isso evitará a aglomeração de dados em níveis baixos de zoom. Mas se desejarmos ter as informações gerais dos dados em tela, seria possível fazê-lo evitando sobrecarga de informação? Uma opção para isso é mostrar os dados agrupados em uma grade (em inglês, grid), que é o que faremos a seguir.

## 1.4 - Adicionando camadas de dados agrupados em grid.

Ao invés de exibir um símbolo para cada documento, o Kibana permite agregar muitos dados em uma mesma representação, evitando a sobrecarga de informação no mapa e, ainda, fornecendo uma noção geral sobre as localizações dos dados. Para isso, utiliza-se a opção de fonte de dados Grid aggregation, que divide o mapa em um grid com várias células de tamanho igual e atribui a cada célula um valor relacionado aos documentos presentes nela. A grade é atualizada de acordo com o zoom e os dados mostrados se alteram de forma dinâmica. Vamos ver como isso funciona, seguindo estes passos:

1. Selecione a opção **Add layer**.
2. Selecione a opção **Grid aggregation**.
3. Em [**Index**](https://leadfortaleza.com.br/ead/glossary/Index_Visualizacao_de_dados)**pattern** selecione **kibana\_sample\_data\_flights**. Ou seja, os dados virão do conjuntos de dados de voos.
4. Em **Geospatial field**, selecione OriginLocation. Ou seja, o atributo que será usado para localização dentro do grid é a origem dos voos.
5. Em **Show as**, selecione clusters (agrupamentos), para agrupar os dados em círculos cujo tamanho varia conforme a quantidade de elementos. As outras opções são grid retangles, que exibe as células da grade, com as cores variando de acordo com os elementos dentro de cada célula, e heatmap, que representa a quantidade de documentos de forma contínua, usando uma escala de cor do frio para o quente.
6. Acione o botão **Add layer**. O resultado deve ser similar ao da figura a seguir.

Figura 7 - criando uma camada de agregação em grade

O tamanho e a cor de cada círculo estão associados a quantidade de dados presentes dentro daquela região. Experimente mudar o zoom e verifique como o mapa se comporta. Agora ficou bem mais fácil visualizar dados de forma geral, com baixo zoom. Em seguida, iremos editar essa camada para melhorar sua visualização.

Para editar a camada de dados agrupados em grade, siga os seguintes passos:

1. Acione o nome layer e selecione a opção **Edit layer**.
2. No painel **Layer settings**, configure **Name** para Origem dos voos agrupados.
3. Configure **Zoom levels** para o intervalo 0 → 5. Ou seja, essa camada só será visível nesses níveis de zoom. Assim, quando o zoom for maior que 5, os dados agrupados irão sumir, mostrando apenas as localizações individuais dos voos.
4. Configure **Opacity** para 100%.
5. No painel **Metrics**, note que a única métrica é a contagem dos voos, que é mostrada em uma [tooltip](https://leadfortaleza.com.br/ead/glossary/Tooltip) quando os agrupamentos estão selecionados. Configure o campo **Custom label** como Contagem de voos para que esse nome seja exibido na [tooltip](https://leadfortaleza.com.br/ead/glossary/Tooltip).
6. Ainda em **Metrics**, selecione a opção **Add metric**.
7. Em **Aggregation**, selecione Average (média) e em **Field**, selecione FlightDelayMin. Isso significa que além da contagem, o [tooltip](https://leadfortaleza.com.br/ead/glossary/Tooltip) irá exibir o tempo de atraso médio do agrupamento.
8. Configure **Custom label** como Tempo de atraso médio.

Avance para a próxima página para continuar o passo a passo.

1. No painel **Layer style**, configure **Fill color** As value, selecionando o atributo Tempo de atraso médio. Isso fará com que a cor do agrupamento esteja associada a sua média do tempo de atraso.
2. Na escala de cor, selecione Custom color ramp, para criar uma escala de cores customizada. Iremos criar 3 tons de azul para 3 intervalos de valores de atraso. Acione os pares 20 e #6092C0, 40 e #2959CF, 60 e #090970. Isso significa que médias de atrasos entre 20 e 40 serão exibidas como um azul claro, entre 40 e 60 como um azul mais escuro e acima de 60 como um azul marinho.
3. Em **Symbol size**, note que o tamanho do agrupamento está atrelado a contagem de voos. Mantenha assim e configure o intervalo de tamanho para 10 → 40. Esses serão os tamanhos máximo e mínimo.
4. As opções de **Label** configuram os valores exibidos dentro dos grupos, nesse caso a contagem de voos. Em **Label color**, configure para Solid e selecione a cor branca.
5. Em **Label border width**, selecione None para remover a borda do rótulo.
6. Acione o botão **Save & close**.

Bom, se você seguiu os passos corretamente, o mapa resultante deverá ser similar ao da figura apresentada na próxima página.

Figura 8 - camada de agregação em grade editada

Experimente agora explorar os dados, modificando o zoom. Você notará que a certo nível de aproximação a camada de agrupamento some, mostrando apenas a camada de documentos individuais. O nosso mapa já está bem interessante, mostrando informações em diferentes níveis de detalhes. Para salvar esse mapa, use a opção Save, disponível no lado superior esquerdo da tela, escolhendo um nome apropriado.

Você deve ter percebido que o Kibana oferece muitas opções para agregar informações às camadas e customizar a sua aparência para criar boas visualizações. No tópico a seguir, você aprenderá a criar outro tipo de camada que pode ser usada para criar mapas muito utilizados chamados mapas coropléticos, que mostram um valor referente a uma região delimitada.