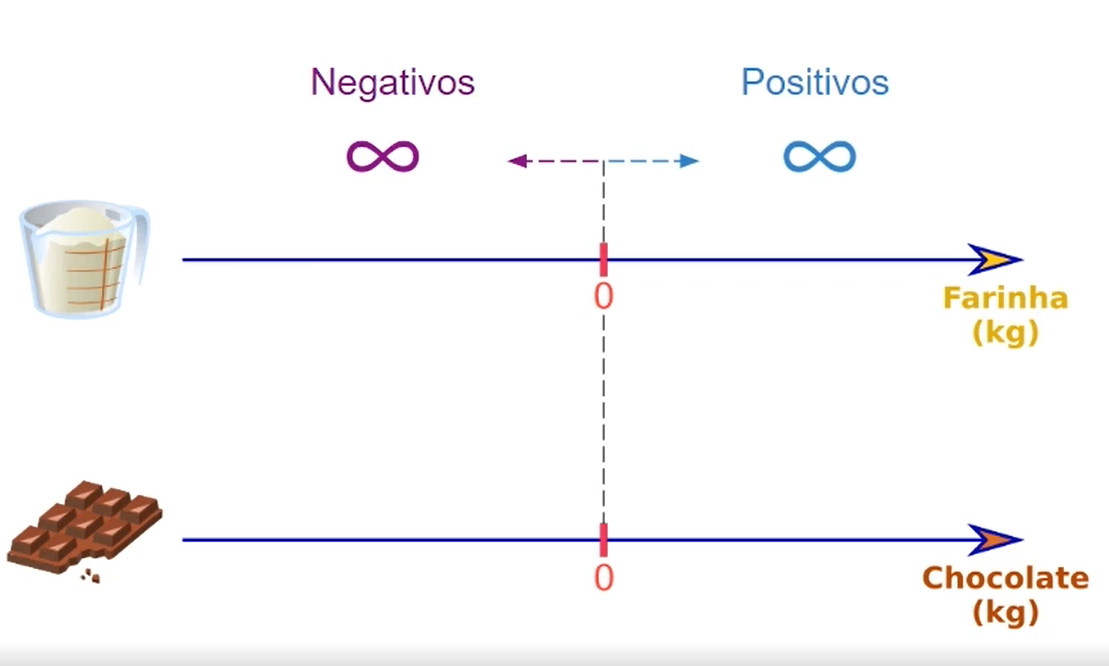
Vamos dar continuidade à tarefa de ajudar a Bel a **planejar seu experimento**. Agora precisamos definir as **faixas de variações** nas quantidades de farinha e chocolate, algo que deve ser feito de forma bastante cuidadosa, pois será a partir dessas faixas que coletaremos nossos **dados experimentais**.

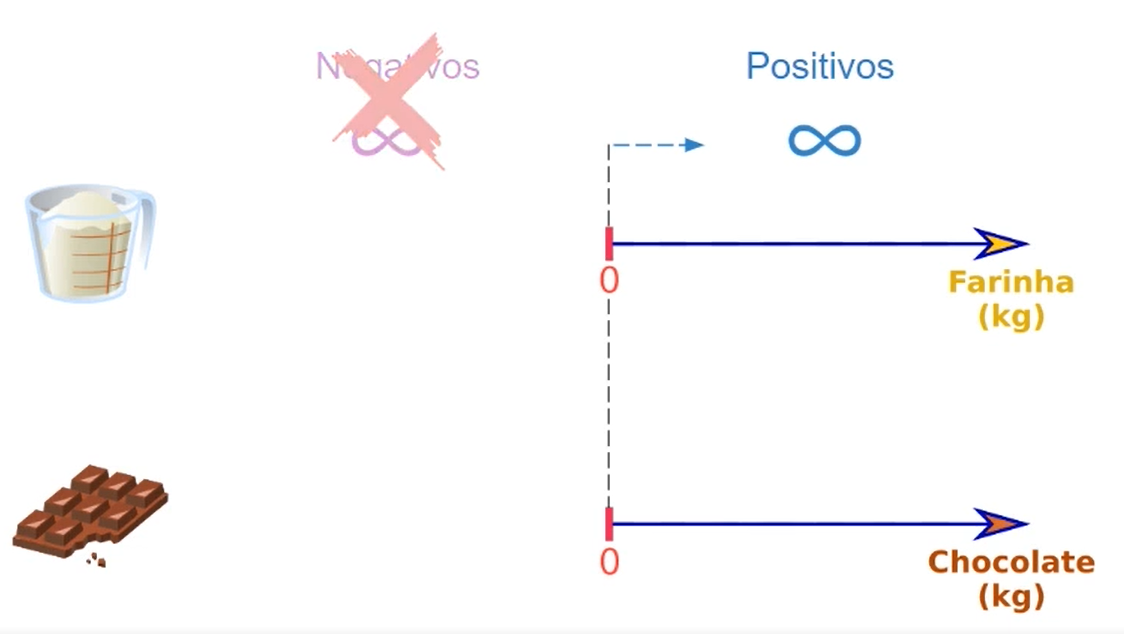
Para definirmos essas faixas, imaginaremos que as quantidades de farinha e chocolate, em quilogramas, podem ser representadas por duas retas. Nelas teremos um ponto de origem, onde as quantidades são iguais a zero. À direita deste ponto estarão os pontos positivos, e à esquerda vamos supor que existam valores negativos. Em ambos os sentidos das retas, os valores se estendem infinitamente.



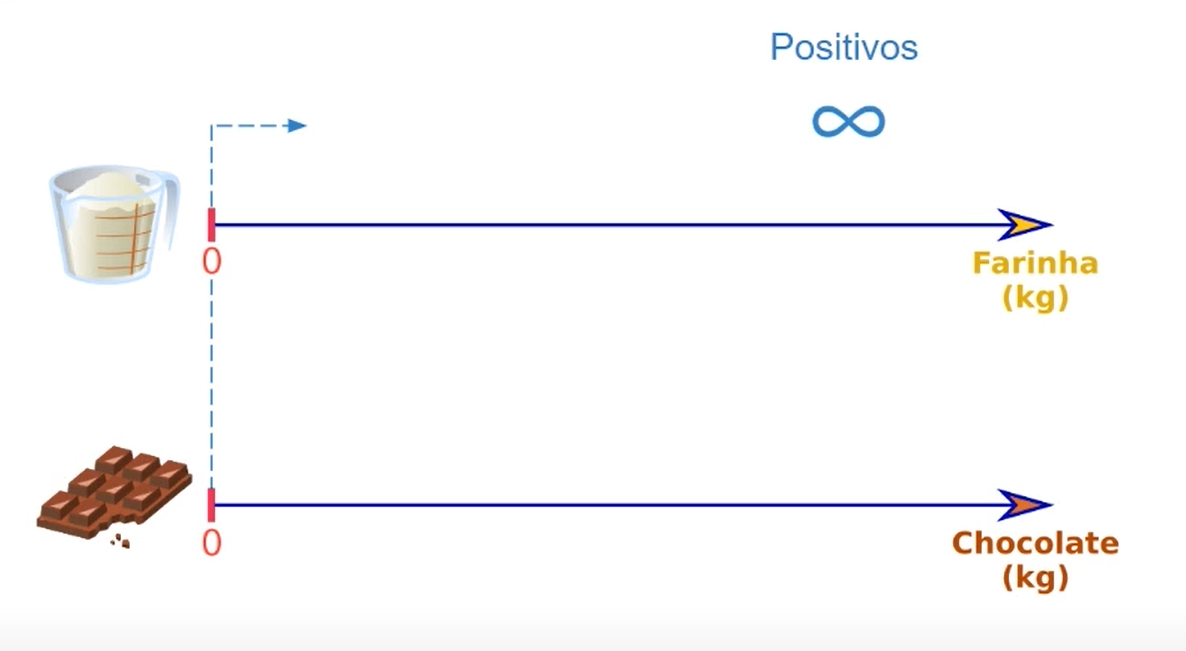
Isso significa que a Bel tem **infinitas possibilidades de coletas de dados**, o que não é ideal para o experimento dela, certo? Sendo assim, precisaremos usar o bom senso e nosso conhecimento sobre a receita de modo a limitarmos essa faixa.

Nesse processo, a primeira questão que precisamos nos fazer é: há quantidades negativas? Quantas vezes já fomos ao mercado comprar "menos 500 gramas" de farinha ou de chocolate? Realmente, nenhuma. Portanto, parece que não existem valores negativos em nosso experimento.

Claro, é necessário tomar bastante cuidado nesse tipo de questionamento, afinal ele é bastante específico desse cenário - em outras condições e experimentos, podem existir números negativos. Por exemplo, poderíamos trabalhar em um experimento que avalia a influência da temperatura na qualidade de um produto, e sabemos que a temperatura pode assumir valores negativos. Ou seja, essa decisão vai depender do sistema que estamos analisando durante o experimento.



Ficamos agora com praticamente metade das possibilidades, já que só manteremos os valores positivos. Contudo, isso mantém a necessidade de infinitos experimentos, já que ainda existem infinitos valores positivos. A ideia agora é restringirmos ainda mais essa faixa.



Para fazermos essa restrição, vamos definir **limites inferiores e superiores**, e é a partir deles que coletaremos os dados experimentais. Com o conhecimento que temos atualmente, podemos fazer algumas suposições gerais sobre quais serão esses valores, a começar pelo limite inferior. Provavelmente esse limite inferior não será 0, pois, segundo a nossa receita, é impossível fazer um cupcake de chocolate sem farinha ou chocolate.

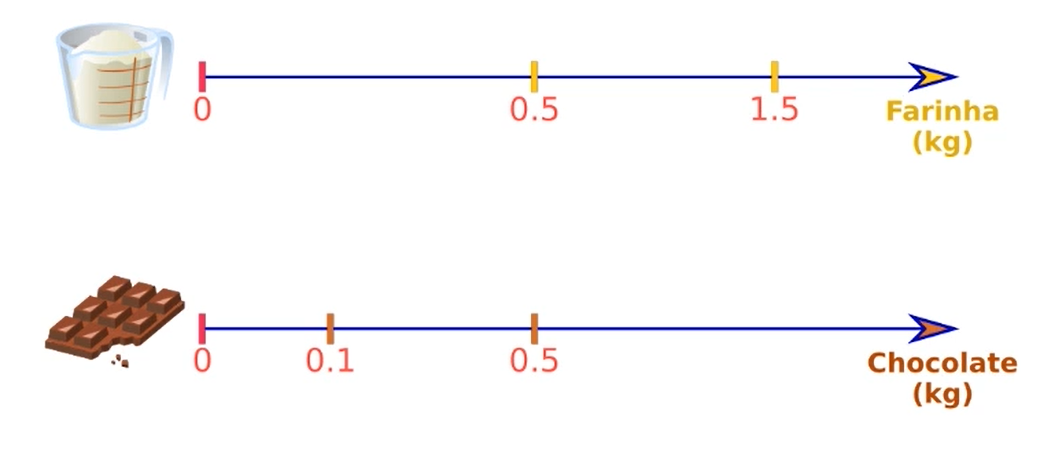
Já em relação ao limite ao superior, temos que respeitar a capacidade da cozinha da Bel - afinal, não utilizaremos uma tonelada de farinha ou meia tonelada de chocolate. Para nos ajudar a restringir esses intervalos, a Bel nos informou que gostaria de assar fornadas de até 50 cupcakes.

Para descobrirmos as quantidades aproximadas necessárias para produzirmos esse resultado, realizaremos uma pesquisa em revistas de culinária e livros de receitas. A esse tipo de pesquisa dá-se o nome de **"revisão bibliográfica"**, e ela é extremamente importante no desenvolvimento e realização de qualquer experimento. Com essa revisão, seremos capazes de entender o que outras pessoas fizeram, e inclusive evitaremos cometer os mesmos erros que elas possam ter cometido.

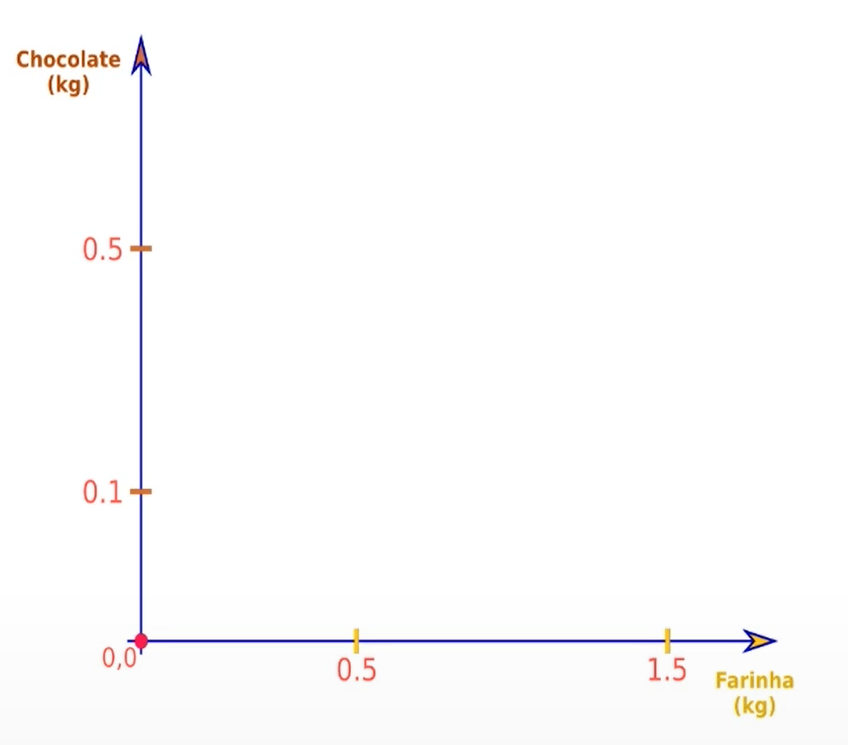
Após essa revisão bibliográfica, chegamos à conclusão de que, para o cenário da Bel, os melhores limites inferiores e superiores são os da tabela abaixo:

|  | **Inferior (kg)** | **Superior (kg)** |
| --- | --- | --- |
| Farinha | 0.5 | 1.5 |
| Chocolate | 0.1 | 0.5 |

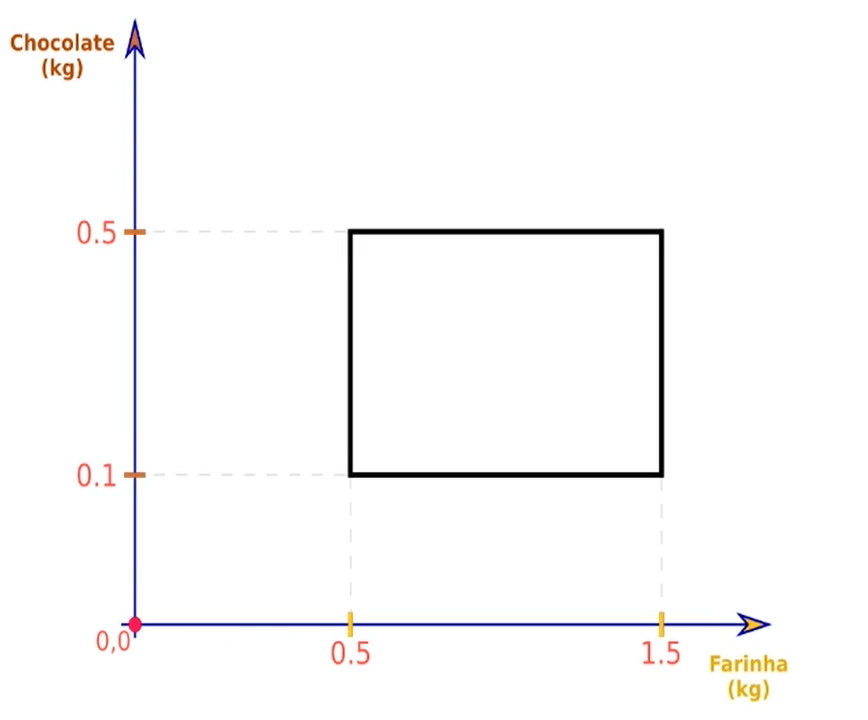
Tendo definido os limites, colocaremos os pontos experimentais nas retas.



Em seguida, manteremos a reta da farinha na horizontal e transporemos a reta que representa as quantidades de chocolate para a vertical. Não é necessário se preocupar com a escala. Essa representação nos é bastante familiar, já que determina um **plano cartesiano**.



Delimitaremos então a região da qual a Bel coletará os seus **pontos experimentais** para realizar as análises estatísticas.

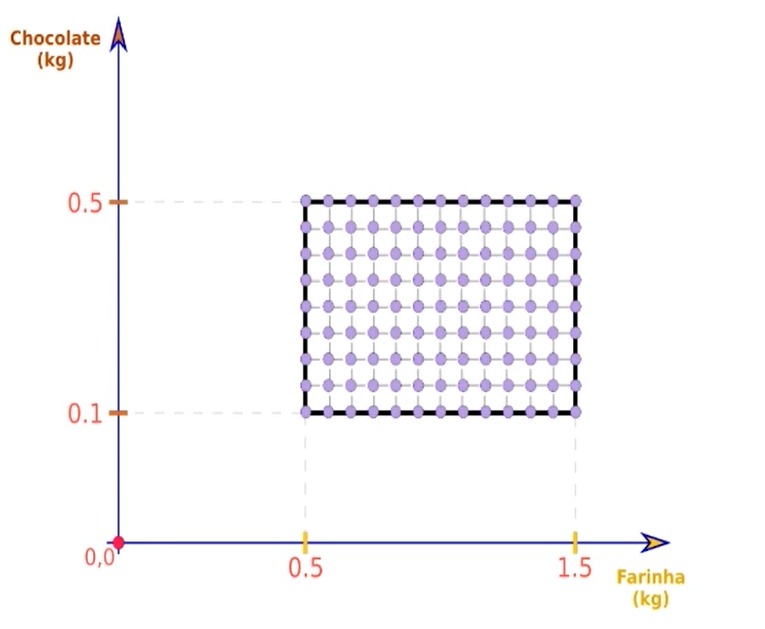


Vamos simular essa coleta. Primeiramente, a Bel coleta um ponto, por exemplo 0.8 quilogramas de farinha e 0,3 quilogramas de chocolate. Ao fim, ela quantifica o número de porções resultante, nesse caso 1. em seguida, ela retira um novo ponto em outras condições, com 1.2 quilogramas de farinha, 0.4 quilogramas de chocolate, resultando em 2 porções.

O problema desse processo é que existem infinitos pontos dentro dessa área, chamada de **"área de experimentação"**. Será necessária uma nova abordagem. Vamos interpretar esses infinitos ensaios como uma população, que representa todos os ensaios possíveis de serem realizados. De dentro dessa população, retiraremos amostras finitas para inferirmos sobre as características da população.

Esse tipo de lógica é muito parecida com a da **inferência estatística**, que é quando utilizamos amostras finitas e reduzidas para inferirmos informações de populações muito grandes de dados.

Agindo desta forma, deixaremos de interpretar a área de experimentação como sendo composta de infinitos pontos, e passaremos a entendê-la como um conjunto de finitos pontos a partir dos quais a coleta de dados pode ser feita.



Apesar de termos chegado a possibilidades finitas, facilitando o trabalho da Bel, ainda precisamos atender a alguns requisitos para que a definição de quais pontos serão coletados nessa amostra seja feita da melhor forma possível. O primeiro ponto é que devemos atender aos objetivos do experimento, ou seja, **avaliar o efeito do chocolate e da farinha no número final de cupcakes produzidos**.

O segundo ponto é que os dados coletados devem **incluir toda a área de experimentação** - ou seja, eles devem ser representativos da região na qual estamos realizando o estudo. Por fim, não podemos esquecer que nesse cenário a coleta de dados é custosa, já que significa um investimento em ingredientes. Portanto, desejamos que **o número de dados coletados seja o menor possível**.