# ML-Chef

## Clustering **Goal:**

we wanted to figure out different types of recipes based on recipe information such as ingredients used.

## Clustering recipes

We applied the unsupervised learning strategy of k-means clustering on ingredients(tf-tid vectorized), which we could then validate using recipe's tags.

With k=6, the algorithm gave a fair distribution of recipes per cluster. Below table depicts the recipes per cluster and the general theme found based on tags.

| Cluster # | Recipes | Color | Theme |
|-----------|---------|---------|---------|
| 0 | 692 | Blue | Dessert/Chocolate |
| 1 | 2482 | Green | Healthy/Salad |
| 2 | 1800 | Red | Healthy/Asian |
| 3 | 1458 | Cyan | Soup/Winter |
| 4 | 1670 | Yellow | Bakery/kid-friendly |
| 5 | 3392 | Magenta | Bakery /Sugar Conscious |

## Visualizing clusters

We applied PCA on ingredients to visualize the clusters.