# Pizza ingredients Case Study

**Data Description:**

### Nutrient analysis of pizzas(Data.wolrd)

Who likes pizza? I mean, there are so many things to like. let’s take a closer look! The data set pizza. sas7.bdat contains measurements that capture the kind of things that make a pizza tasty. Can you determine which pizza brand works best for you and explain why? The variables in the data set are:

**Columns:**

|  |  |
| --- | --- |
| brand | Pizza brand (class label) |
| id | Sample analysed |
| mois | Amount of water per 100 grams in the sample |
| prot | Amount of protein per 100 grams in the sample |
| Fat | Amount of fat per 100 grams in the sample |
| ash | Amount of ash per 100 grams in the sample |
| sodium | Amount of sodium per 100 grams in the sample |
| carb | Amount of carbohydrates per 100 grams in the sample |
| cal | Amount of calories per 100 grams in the sample |

Task:

You are a Data/Business analyst of ABC Inc. your organization decided make analysis on pizza nutrition.

1. You wanted to shrink the information into certain number of components. And the ideal number of components you found is:
   1. 2
   2. 3
   3. 4
   4. 5
2. And the components you got are having the information in range of:
   1. 92-94%
   2. 94-96%
   3. 96-98%
   4. 98-100%

1. Also wanted to the important features that are impacting on deciding brand of pizzas:
   1. Mois, prot, fat, ash, sodium, carbohydrate, calories
   2. Mois, fat, ash, sodium, carbohydrate, calories
   3. Mois, prot, fat, ash, carbohydrate, calories
   4. Mois, prot, fat, ash, sodium, carbohydrate
2. With which technique you got highest accuracy:
   1. Forward feature selection with Random forest
   2. Backward selection with Random forest
   3. Only Random Forest
   4. Both a and b
3. Which brand has the highest test to prediction ratio
   1. G
   2. H
   3. I
   4. J