

Problema: Los nutricionistas desean saber el aporte de de los desayunos en una dieta balanceada, de ahí aparece este dataset con los datos nutricionales de 76 marcas distintas de cereales (Considerando que todo el mundo desayuna cereales).

Dataset:

- Nombre de marca de cereal:
- Manufacturer:
- Calorías:
- Azucares:
- Carbohidratos:
- Proteinas:
- Grasa:
- Sodio
- Fibra
- Potasio
- Estante en el que se encuentran

Al analizar los dataset, vemos ue los rangos son muy fluídos en todos los datos. Por ejemplo, la grasa va entre 0 y 5 gramos, mientras que el sodio va entre 0 y 320 gramos. Es por ello que está en nuestro interes, realizar una normalización de los datos.

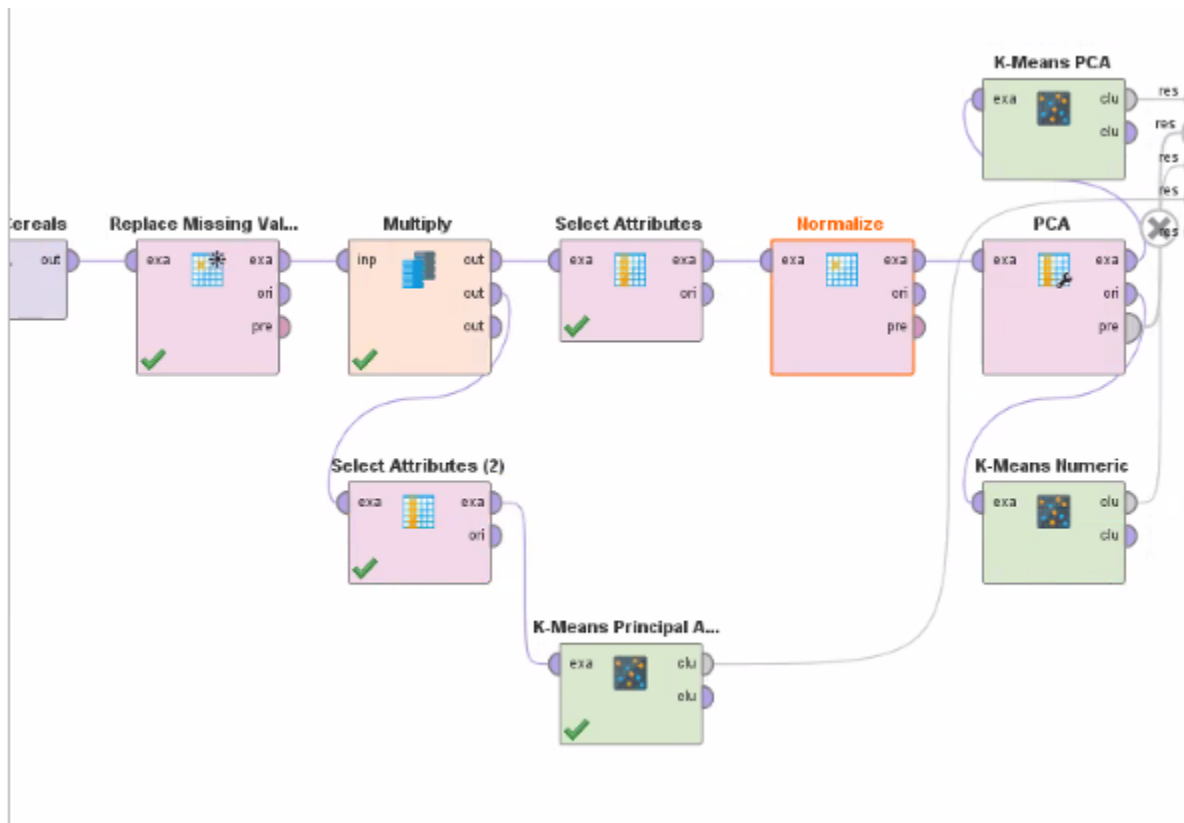
Con los primeros tres componentes se alcanza más del 0.95 de la varianza acumulada. El primer componente principal alcanza más de la mitad de la varianza.

Component	Standard Deviation	Proportion of Variance	Cumulative Variance
PC 1	84.877	0.555	0.555
PC 2	69.520	0.373	0.928
PC 3	22.325	0.038	0.966
PC 4	18.900	0.028	0.994
PC 5	8.513	0.006	0.999
PC 6	2.082	0.000	1.000
PC 7	1.676	0.000	1.000
PC 8	0.798	0.000	1.000
PC 9	0.662	0.000	1.000
PC 10	0.439	0.000	1.000
PC 11	0.222	0.000	1.000
PC 12	0.177	0.000	1.000
PC 13	0.062	0.000	1.000

Attribute	PC 1	PC 2	PC 3 $\downarrow$	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13
vitamins	-0.101	-0.016	0.707	-0.698	-0.026	0.013	0.010	0.006	0.011	0.005	0.008	-0.006	-0.001
calories	-0.076	0.011	0.611	0.617	0.458	0.126	0.093	-0.033	0.003	0.071	0.015	-0.005	-0.004
sugars	-0.004	0.000	0.099	0.116	-0.299	0.625	-0.589	0.113	-0.116	-0.258	0.212	-0.109	-0.018
potass	0.119	-0.989	0.036	0.043	-0.046	-0.051	-0.017	0.013	-0.001	0.010	0.004	-0.002	-0.001
fat	0.000	-0.003	0.016	0.026	-0.018	-0.126	0.376	-0.049	0.174	-0.677	0.533	-0.259	0.012
carbo	-0.015	0.017	0.013	-0.022	0.356	-0.559	-0.657	0.029	0.018	-0.346	-0.084	0.027	-0.015
shelf	0.001	-0.004	0.013	-0.006	-0.008	-0.055	0.100	-0.403	-0.886	-0.180	-0.022	0.080	0.011
weight	-0.001	-0.001	0.004	0.003	0.003	0.008	-0.024	-0.001	0.004	-0.003	-0.046	-0.079	0.995
protein	0.001	-0.009	0.001	-0.002	0.055	0.104	0.242	0.711	-0.192	-0.402	-0.429	0.199	0.001
cups	-0.000	0.002	0.001	-0.001	0.003	-0.011	-0.020	0.057	0.038	0.019	0.456	0.882	0.091
rating	0.076	-0.073	-0.308	-0.339	0.754	0.418	0.011	0.000	-0.052	0.034	0.163	-0.080	-0.002
sodium	-0.982	-0.125	-0.141	0.003	0.016	0.022	0.004	0.001	-0.005	0.001	0.008	-0.004	-0.000
fiber	0.005	-0.031	-0.017	-0.021	0.009	0.271	0.041	-0.559	0.360	-0.394	-0.493	0.288	-0.004

Los atributos más importantes de los componentes principales son Rating, Calories, Potasium, Vitamins.

Modelo:



## KMeans Numeric

History Cluster Model (K-Means Principal Atributes) Cluster Model (K-Means Numeric)

### Cluster Model

```
Cluster 0: 19 items
Cluster 1: 3 items
Cluster 2: 12 items
Cluster 3: 11 items
Cluster 4: 32 items
Total number of items: 77
```

## KMeans Principal Atributes

History Cluster Model (K-Means Principal Atributes)

### Cluster Model

```
Cluster 0: 5 items
Cluster 1: 18 items
Cluster 2: 15 items
Cluster 3: 20 items
Cluster 4: 19 items
Total number of items: 77
```

## KMeans PCA

CA (PCA) Cluster Model (K-Means PCA) Cluster Model (K-Means Principal Atributes)

### Cluster Model

```
Cluster 0: 26 items
Cluster 1: 10 items
Cluster 2: 18 items
Cluster 3: 10 items
Cluster 4: 13 items
Total number of items: 77
```