

# Sistemas de recomendación: Aggregation

Dr. Daniel Gómez-Marín

Database: characteristics

age, height, **weight**, t-shirt size, trouser size, shoe size,  
body fat, hair length, hair color, **eyes color**, etc...

Database: items and characteristics

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A   
B   
C   
D 

People Database

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A   
B   
C   
D   
  
N 



N: person. Which subjects are the most similar to subject N?

N: product. Which subjects would like product N?

## Products Database

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A	
B	
C	
D	
N	

N: person. Which products can we recommend to subject N?  
N: product. Which products are similar to product N?

## Database: searching with weighted characteristics

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A	
B	
C	
D	

"I do not mind the price of the product, i am interested in things that are large, light and technological"

" I only have some data (from a person or product) to search the database"

## Searching the database for groups

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A	
B	
C	
D	
x	
y	
z	

?????????

Which product(s) is(are) similar to these?  
What products can we recommend to these people?

## How do we solve this?

So far, what we know is to find similarity between two items A and B using cosine similarity.

How do we find a product (or products) that are suited for several people?

!!!!This is a socio-technological problem!!!!

## Searching the database for groups: idea A

age, height, weight, t-shirt size, trouser size, shoe size, body fat, hair length

A  
B  
C  
D

????????

x  
y  
z

### IDEA A

For each person (or product) from the list i find a product (or persons) that is (are) the closest and then make an intersection.

## Searching the database for groups: idea A

x  
y  
z

### IDEA A

For each person (or product) from the list i find a product (or persons) that is (are) the closest and then make an intersection.

they are ordered by similarity

a	k	c
z	s	b
w	m	r
p	a	n

## Searching the database for groups: idea A

x  
y  
z

### IDEA A

For each person (or product) from the list i find a product (or persons) that is (are) the closest and then make an intersection.

they are ordered by similarity

a	k	c
z	s	b
w	m	r
p	a	n

No product is common to every member of the search group :(

What x likes the most is not so good for y and z does not like it :(

## idea B: Aggregation ej.1 naive average

### IDEA B

Find a person (or product) derived from the set and with that single line we are going to search products (or people) in the database. Naive: add and

x  
y  
z

x	10	8	3
y	5	7	3
z	1	10	9
N	5.3	8.3	5

!!! the average is not satisfactory for x and z !!!

N

## idea B: Aggregation ej.2 least misery

### IDEA B

Find a person (or product) derived from the set and with that single line we are going to search products (or people) in the database. Naive: add and

x			
y			
z			

x	10	8	3
y	5	7	3
z	1	10	9
N	-	8.3	-

we remove the columns that are below 5.  
jjjThis is what we design!!!!

N

## idea B: Aggregation ej.3 maximum pleasure

### IDEA B

Find a person (or product) derived from the set and with that single line we are going to search products (or people) in the database. Naive: add and

x			
y			
z			

x	10	8	3
y	5	7	3
z	1	10	9
N	-	-	-

columns that are equal or higher than 8 are included

N

## idea B: Aggregation ej.4 media satisfacción

### IDEA B

Find a person (or product) derived from the set and with that single line we are going to search products (or people) in the database. Naive: add and

x			
y			
z			

x	10	8	3
y	5	7	3
z	1	10	9
N	-	8.3	-

columna que tenga baja desviación estandar. ¿cuánto es bajo?

N

## Exercise

Let's make a database of musical genres and fill it up with at least 30 subjects

Implement in our system the 4 aggregation methods presented, so that we can choose group of people (size may change), the aggregation method and see the result of the music proposed for that group.

Analyze extreme cases: different musical personalities, the whole 30 people

Create a report for the nex class (exposition) where you present your findings: How does aggregation affect the result? Is any method better than the rest? why? present lots of examples.

Bonus: design your own aggregation method and explain why is it better than the rest