

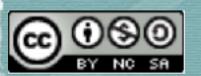
#### Εισαγωγή στην Αναλυτική των Δεδομένων

Άσκηση Πράξης

Παραγωγή Συστάσεων με UBCF και kNN

 $\Delta$ .A.  $\Delta$ έρβος <a href="https://d-a-d.weebly.com">https://d-a-d.weebly.com</a>

Σίνδος, Δεκέμβριος 2022



## Πηγή - Αναφορά

#### Bill Qualls, De Paul University:

http://www.billqualls.com/pa/Qualls ECT584 RecommenderSystems.pdf

## Αναγνώστες Βαθμολογούν Βιβλία

	TRUE BELIEVER	THE DA VINCI CODE	THE WORLD IS FLAT	MY LIFE SO FAR	THE TAKING	THE KITE RUNNER	RUNNY BABBIT	HARRY POTTER
U1	1	5		3			3	5
U2	5	4			3	2	1	
U3	3		1	2	2			5
U4		3			4	1		3
U5	2	4	3			2	2	
U6	5			3	1		3	1
U7	1	4	5	5	2			4
U8	2	1			4	5	1	
U9			3	2	2			5
U10	3	5	1				4	4
U11			2	1		2		3
U12	4	4		2		1	1	4
U13			2		4		4	5
U14		5	3	3	2		1	1
U15		2			3	3		2
U16		3	2	1	1		4	4
U17	1	5	1	2		4		4
U18	5		4		3	3	4	5
U19		4		2		5	1	5
U20	2	5	1	1	5	3		4

#### Αναγνώστες Βαθμολογούν Βιβλία

	TRUE BELIEVER	THE DA VINCI CODE	THE WORLD IS FLAT	MY LIFE SO FAR	THE TAKING	THE KITE RUNNER	RUNNY BABBIT	HARRY POTTER
U1	5	5 4		3			1	5
U2		4			3	2	1	
U3	3		1	2	2			5 3
U4		3			4	1		3
U5	2	4	3			2	2	
U6	5			3	1		3	1
U7	1	4	5	5	2			4
U8	2	1			4	5	1	
U9			3	2	2			5
U10	3	5	1				4	4
U11			2	1		2		3
U12	4	4		2		1	1	4
U13			2		4		4	5
U14		5	3	3	2		1	1
U15		2			3	3		
U16		3	2	1	1		4	4
U17	1	5	1	2		4		4
U18	5		4		3	3	4	5
U19		4		2		5	1	5
U20	2	5	1	1	5	3		4

Πληροφορία Αναφοράς, "Μοντέλο"

NU1	3		5	4	2	3		5
NU2		5	2	2	4		1	3

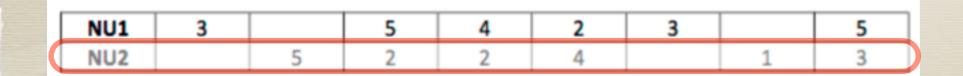
Πρόβλεψη (σύσταση;) για αυτά που δεν έχουν διαβάσει οι ΝU₁ και ΝU₂;

## Βαθμός Ομοιότητας (Pearson)

		TRUE BELIEVER	THE DA VINCI CODE	THE WORLD IS FLAT	MY LIFE SO FAR	THE TAKING	THE KITE RUNNER	RUNNY BABBIT	HARRY POTTER
г	U1	1	5	-	3			3	5
ŀ	U2	5	4		3	3	2	1	- 3
ŀ	U3	3	4	1	2	2	- 2	1	5
ŀ		3	2	1		4	•		3
ŀ	U4	•	3	•		4	1	•	3
1	U5	2	4	3			2	2	
Ų	U6	5			3	1		3	1
-	U7	1	4	5	5	2			4
Į.	U8	2	1			4	5	1	
L	U9			3	2	2			5
	U10	3	5	1				4	4
	U11			2	1		2		3
	U12	4	4		2		1	1	4
	U13			2		4		4	5
	U14		5	3	3	2		1	1
	U15		2			3	3		2
	U16		3	2	1	1		4	4
İ	U17	1	5	1	2		4		4
ŀ	U18	5		4		3	3	4	5
Ì	U19		4		2		5	1	5
Ì	U20	2	5	1	1	5	3		4

 $Sim(U_6, NU_2)$ 

Google Sheets: CORREL()



# Αναγνώστης Βαθμολογεί Βιβλίο

	B <sub>1</sub>	$B_2$	B <sub>3</sub>	В4	$B_5$	$B_6$	$B_7$	В8
	TRUE BELIEVER	THE DA VINCI CODE	THE WORLD IS FLAT	MY LIFE SO FAR	THE TAKING	THE KITE RUNNER	RUNNY BABBIT	HARRY POTTER
U1	1	5		3			3	5
U2	5	4			3	2	1	
U3	3		1	2	2			5
U4		3			4	1		3
U5	2	4	3			2	2	
U6	5			3	1		3	1
U7	1	4	5	5	2			4
U8	2	1			4	5	1	
U9			3	2	2			5
U10	3	5	1				4	4
U11			2	1		2		3
U12	4	4		2		1	1	4
U13			2		4		4	5
U14		5	3	3	2		1	1
U15		2			3	3		2
U16		3	2	1	1		4	4
U17	1	5	1	2		4		4
U18	5		4		3	3	4	5
U19		4		2		5	1	5
U20	2	5	1	1	5	3		4

Αναγνώστης: U8

Bιβλίο: THE KITE RUNNER (B<sub>6</sub>):

 $r(U_8,B_6) = 5$ 

Ζητούμενα: r(NU<sub>1</sub>,B<sub>2</sub>), r(NU<sub>1</sub>,B<sub>7</sub>) r(NU<sub>2</sub>,B<sub>1</sub>), r(NU<sub>2</sub>,B<sub>6</sub>)

N	J1	3		5	4	2	3		5
N	J2		5	2	2	4		)   1	3

# Ο Υπολογισμός

Έστω ότι οι U1, U5 και U7 είναι αυτοί που μοιάζουν περισσότερο στον NU1, τότε:

$$r(NU_1,B_2) = \frac{r(U_1,B_2)^* sim(U_1,NU_1) + r(U_5,B_2)^* sim(U_5,NU_1) + r(U_7,B_2)^* sim(U_7,NU_1)}{sim(U_1,NU_1) + sim(U_5,NU_1) + sim(U_7,NU_1)}$$

$$r(NU_1,B_7) = \frac{r(U_1,B_7)*sim(U_1,NU_1) + r(U_5,B_7)*sim(U_5,NU_1) + r(U_7,B_7)*sim(U_7,NU_1)}{sim(U_1,NU_1) + sim(U_5,NU_1) + sim(U_7,NU_1)}$$

Aνάλογα:  $r(NU_2,B_1)$  και  $r(NU_2,B_6)$ 

Μένει να βρεθούν ποιοί τρεις των  $U_1, U_2, ..., U_{20}$  έχουν μεγαλύτερη τιμή sim()/CORREL() για τον  $NU_1$  και ποιοί τρείς για τον  $NU_2...$ 

#### Ερωτήσεις;

