

ΕΡΓΑΣΙΑ 4 – Εξόρυξη Κανόνων Συσχέτισης

Transaction ID	Items
T1	Κραγιόν, Μάσκαρα, Ρουζ
T2	Κραγιόν, Σκιές, Πούδρα
T3	Κραγιόν, Μάσκαρα
T4	Πούδρα, Ρουζ
T5	Πούδρα, Σκιές
T6	Κραγιόν, Σκιές, Πούδρα, Ρουζ

1. Apriori με το χέρι

Εκτελέστε τον αλγόριθμο Apriori πάνω στα παραπάνω δεδομένα καταστήματος καλλυντικών με όρια υποστήριξης $s=33\%$ και εμπιστοσύνης $c=60\%$.

Δείξτε τα υποψήφια και τα συχνά στοιχειosύνολα σε κάθε πέρασμα της βάσης.

Δώστε τη λίστα με τα τελικά συχνά στοιχειosύνολα και τη λίστα με τους κανόνες συσχέτισης που παράγονται από αυτά ταξινομημένους ως προς την εμπιστοσύνη.

Για τους τελικούς κανόνες συσχέτισης (δηλαδή για αυτούς που έχουν $c \geq 60\%$) υπολογίστε και το lift.

Ποιον κανόνα θεωρείτε τον πιο ισχυρό και γιατί; Σχολιάστε τα αποτελέσματά σας.

Υποψήφια 1-στοιχειosύνολα

	s
{Κραγιόν}	4/6
{Μάσκαρα}	2/6
{Πούδρα}	4/6
{Ρουζ}	3/6
{Σκιές}	3/6

Όλα έχουν $s \geq 33\%$ (τα συχνά στοιχειosύνολα συμβολίζονται με έντονη γραφή)

Υποψήφια 2-στοιχειosύνολα

	s
{Κραγιόν, Μάσκαρα}	2/6
{Κραγιόν, Πούδρα}	2/6
{Κραγιόν, Ρουζ}	2/6
{Κραγιόν, Σκιές}	2/6
{Μάσκαρα, Πούδρα}	0/6
{Μάσκαρα, Ρουζ}	1/6
{Μάσκαρα, Σκιές}	0/6
{Πούδρα, Ρουζ}	2/6
{Πούδρα, Σκιές}	3/6
{Ρουζ, Σκιές}	1/6

Υποψήφια 3-στοιχειosύνολα (θα πρέπει όλα τα υποσύνολα των υποψηφίων να είναι και αυτά συχνά)

	s
{Κραγιόν, Πούδρα, Ρουζ}	1/6
{Κραγιόν, Πούδρα, Σκιές}	2/6

Άρα, δεν υπάρχουν συχνά 4-στοιχειosύνολα.

Οπότε περνάμε στον σχηματισμό των κανόνων από τα συχνά 2- και 3-στοιχειosύνολα αντίστοιχα, και υπολογίζουμε και την εμπιστοσύνη τους.

	c
Κραγιόν --> Μάσκαρα	2/4
Μάσκαρα --> Κραγιόν	1
Κραγιόν --> Πούδρα	2/4
Πούδρα --> Κραγιόν	2/4
Κραγιόν --> Ρουζ	2/4
Ρουζ --> Κραγιόν	2/3
Κραγιόν --> Σκιές	2/4
Σκιές --> Κραγιόν	2/3
Πούδρα --> Ρουζ	2/4
Ρουζ --> Πούδρα	2/3
Πούδρα --> Σκιές	3/4
Σκιές --> Πούδρα	1
Κραγιόν --> Πούδρα, Σκιές	2/4
Πούδρα --> Κραγιόν, Σκιές	2/4
Σκιές --> Κραγιόν, Πούδρα	2/3
Κραγιόν, Πούδρα --> Σκιές	1
Κραγιόν, Σκιές --> Πούδρα	1
Πούδρα, Σκιές --> Κραγιόν	2/3

Για τους κανόνες με confidence $\geq 60\%$, υπολογίζω το lift.

	c	lift
Μάσκαρα --> Κραγιόν	1	$12/8=1.5$
Ρουζ --> Κραγιόν	2/3	$12/12=1$
Σκιές --> Κραγιόν	2/3	$12/12=1$
Ρουζ --> Πούδρα	2/3	$12/12=1$
Πούδρα --> Σκιές	3/4	$18/12=1.5$
Σκιές --> Πούδρα	1	$18/12=1.5$
Σκιές --> Κραγιόν, Πούδρα	2/3	$12/6=2$
Κραγιόν, Πούδρα --> Σκιές	1	$12/6=2$
Κραγιόν, Σκιές --> Πούδρα	1	$12/8=1.5$
Πούδρα, Σκιές --> Κραγιόν	2/3	$12/12=1$

Πιο ισχυρός κανόνας με $c=100\%$ και $\text{lift}=2$: Κραγιόν, Πούδρα --> Σκιές