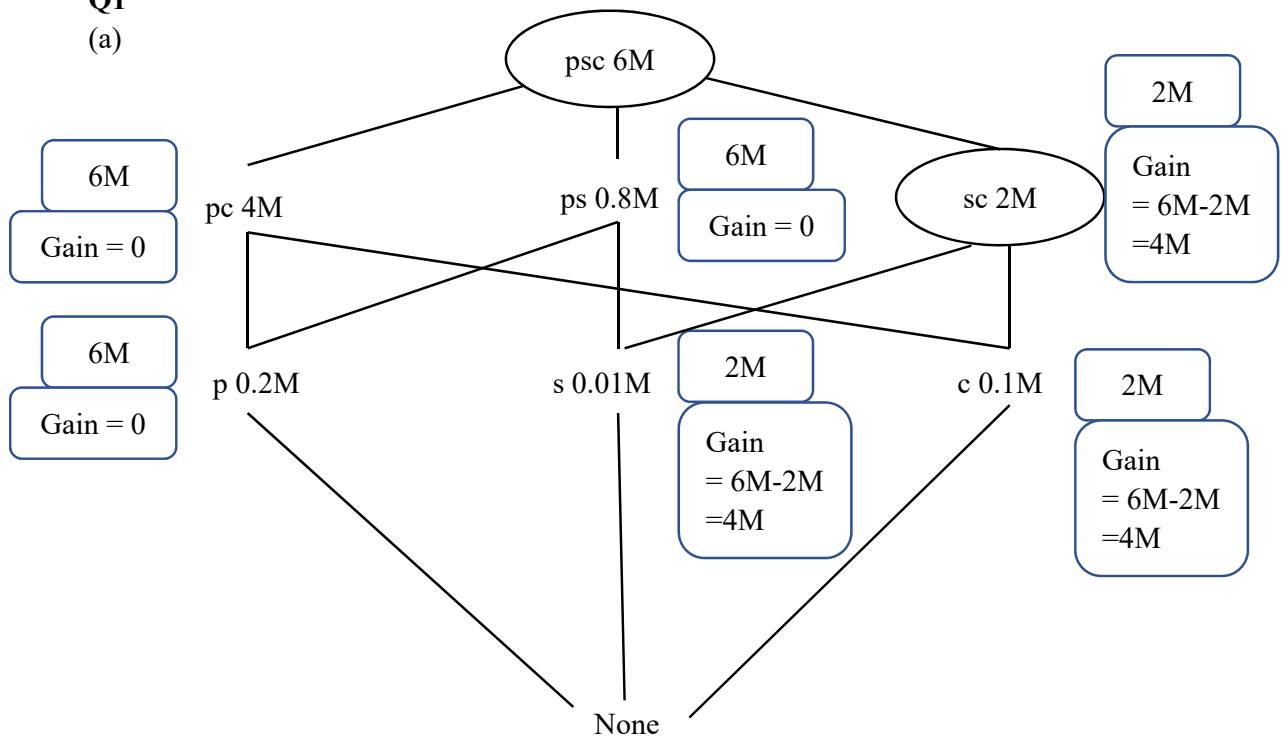


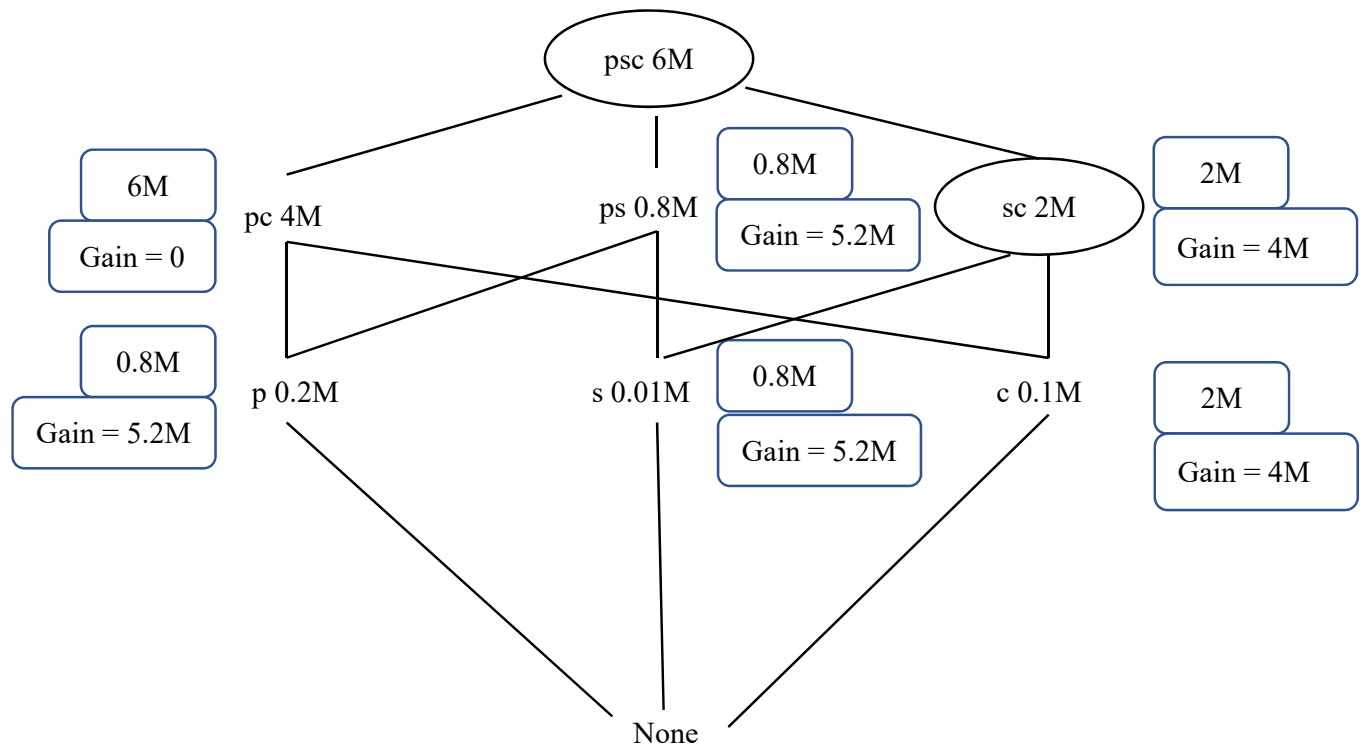
MSBD5002: Data Mining and Knowledge Discovery (MicroMasters)
 Exercise 8 (Suggested Solution)
 Data Warehouse

Q1
 (a)



$$Gain(\{sc, top\}, \{top\}) = 4M \times 3 = 12M$$

(b)



$$Gain(\{PS, SC, top\}, \{top\}) = 5.2M \times 3 + 4M \times 2 = 23.6M$$

Q2

	First Choice (b)	Second Choice (f)	Third Choice (d)
b	$50 \times 5 = 250$		
c	$25 \times 5 = 125$	$25 \times 2 = 50$	$25 \times 1 = 25$
d	$80 \times 2 = 160$	$30 \times 2 = 60$	$30 \times 2 = 60$
e	$70 \times 3 = 210$	$20 \times 3 = 60$	$20 + 20 + 10 = 50$
f	$60 \times 2 = 120$	$60 + 10 = 70$	
g	$99 \times 1 = 99$	$49 \times 1 = 49$	$49 \times 1 = 49$
h	$90 \times 1 = 90$	$40 \times 1 = 40$	$30 \times 1 = 30$

Resulting views = $\{b, f, d\}$