CSCE 5370.001

- Assignment No. 1

Question 1

Part 1:

Simple predicates for 'TITLE' in pay.

P_1 : TITLE = "Elect. Eng."

P_2 : TITLE = "Syst. Anal."

P_3 : TITLE "Mech. Eng."

P_4 : TITLE "Programmer"

Simpl predicates -

$P_j : A_r \ \theta$ value

P_j - simple predicate

$\theta \in \{=, <, \neq, \leq, >, \geq\}$

value $\in D_i$

A_i - given relation

$R(A_1, A_2, \ldots, A_n)$, A_i is an attribute defined over a domain D_i

Part 2:

$R_1 = PAY_1 = \sigma_{TITLE = \text{"Elect. Eng."}}(PAY)$

$R_2 = PAY_2 = \sigma_{TITLE = \text{"Syst. Anal."}}(PAY)$

$R_3 = PAY_3 = \sigma_{TITLE = \text{"Mech. Eng."}}(PAY)$

$R_4 = PAY_4 = \sigma_{TITLE = \text{"Programmer"}}(PAY)$

PRIMARY HORIZONTAL FRAGMENTS of 'PAY' using simple predicates

Primary Horizontal fragmentation -

$R_i = \sigma_{F_i}(R)$, $1 \leq i \leq w$

F_i - selection formula to obtain fragment R_i

R - Relation

R_i - Horizontal fragments

PAY1

TITLE	SAL
Elect. Eng.	40000

PAY2

TITLE	SAL
Syst. Anal.	34000

PAY3

TITLE	SAL
Mech. Eng.	27000

PAY4

TITLE	SAL
Programmer	24000

Part 3:

Considering the fragmentation of 'PAY' in Part 2, we have 4 primary horizontal fragments.

Let's now perform derived horizontal fragmentation on EMP table -

$EMP_1 = EMP \ltimes PAY_1$

$EMP_2 = EMP \ltimes PAY_2$

$EMP_3 = EMP \quad PAY_3$

$EMP_4 = EMP \quad PAY_4$

Derive horizontal fragmentation

$R_i = R \ltimes S_i$, $1 \leq i \leq w$

$S_i = \sigma_{F_i}(S)$ [primary horizontal formula]

R_i derived HF

w - maximum no. of fragments

EMP_1

ENO	ENAME	TITLE
E1	J.Doe	Elect.Eng.
E6	L.Chu	Elect.Eng.

EMP_2

ENO	ENAME	TITLE
E2	M.Smith	Syst.Anal.
E5	B.Casey	Syst.Anal.
E8	J.Jones	Syst.Anal.

EMP_3

ENO	ENAME	TITLE
E3	A.Lee	Mech.Eng.
E7	R.Davis	Mech.Eng.

EMP_4

ENO	ENAME	TITLE
E4	J.Miller	Programmer

Part 4:

Using the 4 EMP fragment tables, lets derive horizontal fragmentation for ASG.

$$ASG_1 = ASG \ltimes EMP_1$$
$$ASG_2 = ASG \ltimes EMP_2$$
$$ASG_3 = ASG \ltimes EMP_3$$
$$ASG_4 = ASG \ltimes EMP_4$$

ASG1

ENO	PNO	RESP	DUR
E1	P1	Manager	12
E6	P4	Manager	48

ASG2

ENO	PNO	RESP	DUR
E2	P1	Analyst	24
E2	P2	Analyst	6
E5	P2	Manager	24
E8	P3	Manager	40

ASG3

ENO	PNO	RESP	DUR
E3	P3	Consultant	10
E3	P4	Engineer	48
E7	P3	Engineer	36

ASG4

ENO	PNO	RESP	DUR
E4	P2	Programmer	18

Question - 2

ASG & PROJ are horizontally fragmented -

$ASG_1 = \sigma_{PNO = "P_1"}(ASG)$

ASG1

ENO	PNO	RESP	DUR
E_1	P_1	Manager	12
E_2	P_1	Analyst	24

$ASG_2 = \sigma_{PNO = "P_2"}(ASG)$

ASG2

ENO	PNO	RESP	DUR
E_2	P_2	Analyst	6
E_4	P_2	programmer	18
E_5	P_2	Manager	24

$ASG_3 = \sigma_{PNO = "P_3"}(ASG)$

END	PNO	RESP	DUR
E_3	P_3	Consultant	10
E_7	P_3	Engineer	36
E_8	P_3	Manager	40

$ASG_4 = \sigma_{PNO = "P_4"}(ASG)$

ENO	PNO	RESP	DUR
E_3	P_4	Engineer	48
E_6	P_4	Manager	48

$PROJ_1 = \sigma_{BUDGET \geq 200000}(PROJ)$

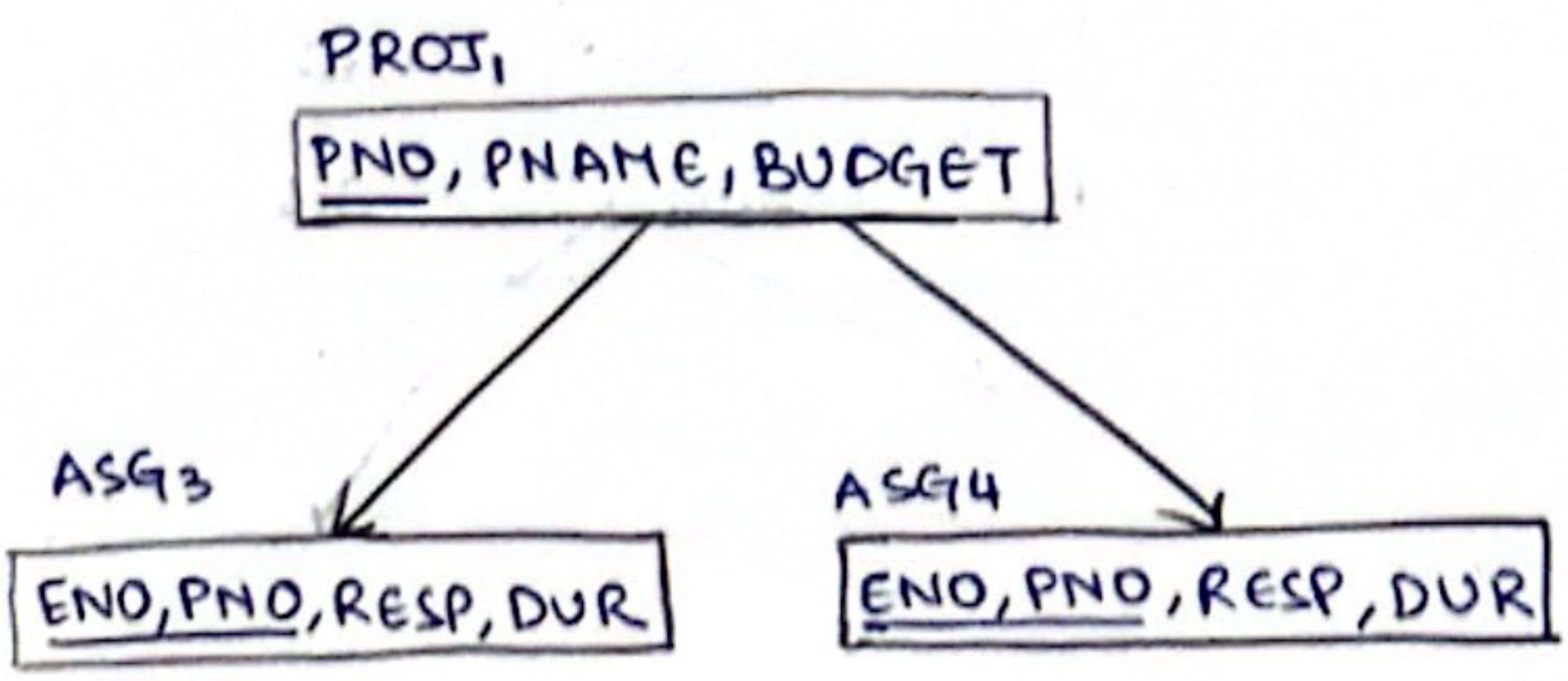
PNO	PNAME	BUDGET
P_3	CAD/CAM	250000
P_4	Maintanance	310000

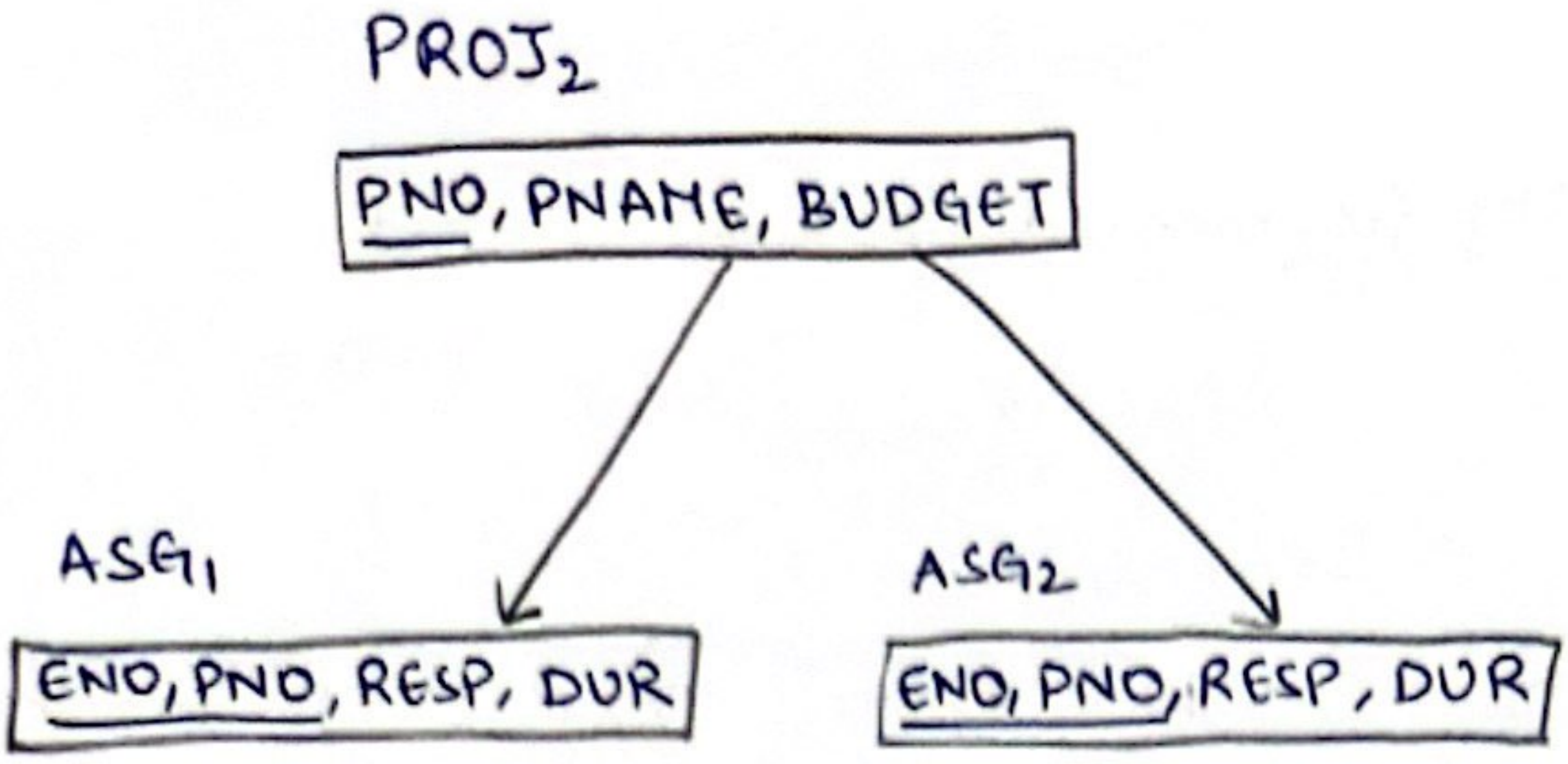
$PROJ_2 = \sigma_{BUDGET < 200000}(PROJ)$

PNO	PNAME	BUDGET
P_1	Instrumentation	150000
P_2	Database Develop.	135000

Part 1

Join graph of $ASG \ltimes_{PNO} PROJ$





The graphs are partitioned join graphs.

Part 2

Lets modify fragmentation of ASG to make the join graph ($ASG \ltimes_{PNO} PROJ$) simple :

ASG_1 : $\sigma_{PNO = \text{"}P_1\text{"} \vee PNO = \text{"}P_2\text{"}}(ASG)$

ASG_2 : $\sigma_{PNO = \text{"}P_3\text{"} \vee PNO = \text{"}P_4\text{"}}(ASG)$

Then,

ASG_1

ENO	PNO	RESP	DUR
E1	P1	Manager	12
E2	P1	Analyst	24
E2	P2	Analyst	6
E4	P2	Programmer	18
E5	P2	Manager	24

ASG_2

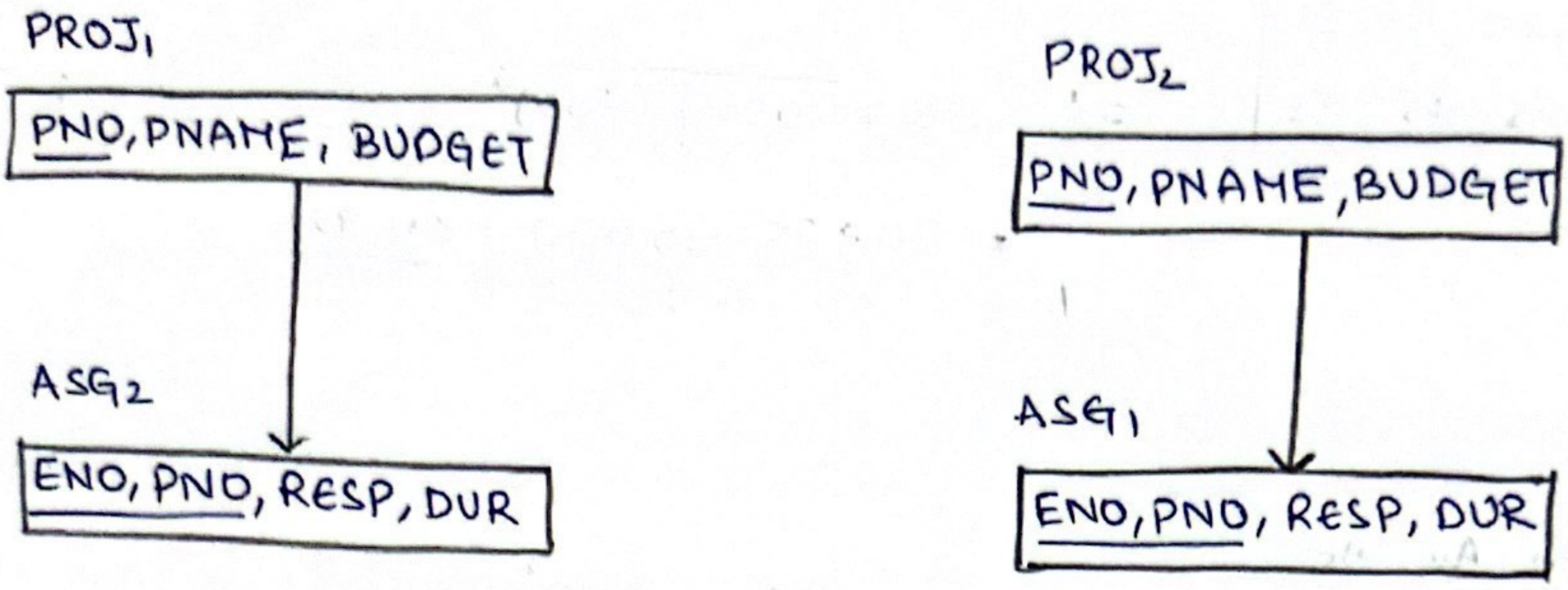
ENO	PNO	RESP	DUR
E3	P3	Consultant	10
E7	P3	Engineer	36
E8	P3	Manager	40
E3	P4	Engineer	48
E6	P4	Manager	48

Scanned with CamScanner

Part 3

Join graphs:

SIMPLE



There is only one link coming in or going out of a fragment - So, simple join graphs.

Question 3

$Q = \{q_1, q_2, q_3, q_4, q_5\}$

$A = \{A_1, A_2, A_3, A_4, A_5\}$

$S = \{S_1, S_2, S_3\}$

$ref_s(q_k) = 1$

Attribute Affinity Measure

$aff(A_i, A_j) = \sum_{\text{all quires that access } A_i \& A_j}$ (query access

$= \sum \sum ref_s(q_k) \cdot acc_s(q_k)$
$\forall k,\ use(q_k, A_i) = 1 \wedge use(q_k, A_j) = 1\ \forall s$

The access frequencies $acc_s(q_k) =$

$$\begin{array}{c|c} & \begin{matrix} S_1 & S_2 & S_3 \end{matrix} \\ \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \end{matrix} & \begin{bmatrix} 20 & 30 & 0 \\ 10 & 0 & 15 \\ 0 & 45 & 15 \\ 0 & 20 & 0 \\ 0 & 25 & 0 \end{bmatrix} \end{array}$$

$use(q_i, A_j) =$

$$\begin{array}{c|c} & \begin{matrix} A_1 & A_2 & A_3 & A_4 & A_5 \end{matrix} \\ \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

Part 1

$aff(A_1, A_1) = 1\times20 + 1\times30 + 1\times0 +$
$1\times10 + 1\times0 + 1\times15 +$
$1\times0 + 1\times20 + 1\times0 = 50 + 25 + 20 = 95$

$aff(A_1, A_2) = 1\times10 + 1\times0 + 1\times15 = 10+15 = 25$

$aff(A_1, A_3) = 1\times10 + 1\times0 + 1\times15 +$
$1\times0 + 1\times20 + 1\times0 = 25 + 20 = 45$

$aff(A_1, A_4) = 1\times20 + 1\times30 + 1\times0 = 50$

$aff(A_1, A_5) = 0$

$aff(A_2, A_1) = aff(A_1, A_2) = 25$

$aff(A_2, A_2) = 1\times10 + 1\times0 + 1\times15 +$
$1\times0 + 1\times25 + 1\times0 = 25 + 25 = 50$

$aff(A_2, A_3) = 1\times10 + 1\times0 + 1\times15 = 25$

$aff(A_2, A_4) = 1\times0 + 1\times25 + 1\times0 = 25$

$aff(A_2, A_5) = 1\times0 + 1\times25 + 1\times0 = 25$

$aff(A_3, A_1) = aff(A_1, A_3) = 45$

$aff(A_3, A_2) = aff(A_2, A_3) = 25$

$aff(A_3, A_3) = 1\times10 + 1\times15 + 1\times20 = 45$

$aff(A_3, A_4) = 0$

$aff(A_3, A_5) = 0$

$aff(A_4, A_1) = aff(A_1, A_4) = 50$

$aff(A_4, A_2) = aff(A_2, A_4) = 25$

$aff(A_4, A_3) = aff(A_3, A_4) = 0$

$aff(A_4, A_4) = 1\times20 + 1\times30 + 1\times0 +$
$1\times0 + 1\times45 + 1\times15 +$
$1\times0 + 1\times25 + 1\times0 = 135$

$aff(A_4, A_5) = 1\times0 + 1\times45 + 1\times15 +$
$1\times0 + 1\times25 + 1\times0 = 85$

$aff(A_5, A_1) = aff(A_1, A_5) = 0$

$aff(A_5, A_2) = aff(A_2, A_5) = 25$

$aff(A_5, A_3) = aff(A_3, A_5) = 0$

$aff(A_5, A_4) = aff(A_4, A_5) = 85$

$aff(A_5, A_5) = 1\times0 + 1\times45 + 1\times15 +$
$1\times0 + 1\times25 + 1\times0 = 85$

Attribute Affinity (AA) Matrix –

$$\begin{array}{c|c} & \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} & \begin{bmatrix} 95 & 25 & 45 & 50 & 0 \\ 25 & 50 & 25 & 25 & 25 \\ 45 & 25 & 45 & 0 & 0 \\ 50 & 25 & 0 & 135 & 85 \\ 0 & 25 & 0 & 85 & 85 \end{bmatrix} \end{array}$$

Part 2

Let's convert AA matrix to clustred affinity matrix (CA)

Attribute Affinity Matrix (AA)

$$
\begin{array}{c|c}
 & \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\
\begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} &
\begin{bmatrix}
95 & 25 & 45 & 50 & 0 \\
25 & 50 & 25 & 25 & 25 \\
45 & 25 & 45 & 0 & 0 \\
50 & 25 & 0 & 135 & 85 \\
0 & 25 & 0 & 85 & 85
\end{bmatrix}
\end{array}
$$

clustered affinity matrix (CA)

$$
\begin{array}{c|c}
A_0 & \begin{array}{cccccc} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 \end{array} \\
\begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} &
\begin{bmatrix}
95 & 25 & & & & \\
25 & 50 & & & & \\
45 & 25 & & & & \\
50 & 25 & & & & \\
0 & 25 & & & &
\end{bmatrix}
\end{array}
$$

lets copy the 1st 2 columns to CA matrix

~~Row A_3:~~

Column A_3:

There are 4 possible positions : $A_0\ A_3\ A_1$

$A_1\ A_3\ A_2$

$A_2\ A_3\ A_4$

$A_4\ A_3\ A_5$

Contribution of a placement:

$cont(A_i, A_k, A_j) = 2bond(A_i, A_k) + 2bond(A_k, A_j) - 2bond(A_i, A_j)$

where

$bond(A_x, A_y) = \sum_{z=1}^{n} aff(A_z, A_x)\, aff(A_z, A_y)$

Ordering (0-3-1):

$cont(A_0, A_3, A_1) = 2 * bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1)$

$bond(A_x, A_y) = \sum_{z=1}^{n} aff(A_z, A_x)\, aff(A_z, A_y)$

here $n=5$

$bond(A_0, A_3) = aff(A_1, A_0)\, aff(A_1, A_3) + aff(A_2, A_0)\, aff(A_2, A_3) + aff(A_3, A_0)\, aff(A_3, A_3) + aff(A_4, A_0)\, aff(A_4, A_3) + aff(A_5, A_0)\, aff(A_5, A_3)$

$= 0$

as we know $aff(A_0, A_i) = aff(A_i, A_0) = 0$

$bond(A_3, A_1) = aff(A_1, A_3)\, aff(A_1, A_1) + aff(A_2, A_3)\, aff(A_2, A_1) +$
$aff(A_3, A_3)\, aff(A_3, A_1) + aff(A_4, A_3)\, aff(A_4, A_1)$
$+ aff(A_5, A_3)\, aff(A_5, A_1)$

$= 45 \times 95 + 25 \times 25 + 45 \times 45 + 0 \times 50 + 0 \times 0$

$= 4275 + 625 + 2025$

$= 6925$

$bond(A_0, A_1) = 0$

as we know $aff(A_0, A_i) = aff(A_i, A_0) = 0$

$cont(A_0, A_3, A_1) = 2\,bond(A_0, A_3) + 2\,bond(A_3, A_1) - 2\,bond(A_0, A_1)$

$= 2 \times 0 + 2 \times 6925 - 2 \times 0 = 13{,}850$

$\boxed{cont(A_0, A_3, A_1) = 13{,}850}$

ordering (1-3-2):

$cont(A_1, A_3, A_2) = 2\,bond(A_1, A_3) + 2\,bond(A_3, A_2) - 2\,bond(A_1, A_2)$

$bond(A_1, A_3) = bond(A_3, A_1) = 6925$

$bond(A_3, A_2) = aff(A_1, A_3)\, aff(A_1, A_2) + aff(A_2, A_3)\, aff(A_2, A_2) +$
$aff(A_3, A_3)\, aff(A_3, A_2) + aff(A_4, A_3)\, aff(A_4, A_2)$
$+ aff(A_5, A_3)\, aff(A_5, A_2)$

$= 45 \times 25 + 25 \times 50 + 45 \times 25 + 0 \times 25 + 0 \times 25$

$= 1125 + 1250 + 1125 + 0 + 0 = 3500$

$bond(A_1, A_2) = aff(A_1, A_1)\, aff(A_1, A_2) + aff(A_2, A_1)\, aff(A_2, A_2) +$
$aff(A_3, A_1)\, aff(A_3, A_2) + aff(A_4, A_1)\, aff(A_4, A_2) +$
$aff(A_5, A_1)\, aff(A_5, A_2)$

$= 95 \times 25 + 25 \times 50 + 45 \times 25 + 50 \times 25 + 0 \times 25$

$= 2375 + 1250 + 1125 + 1250 = 6000$

$cont(A_1, A_3, A_2) = 2\,bond(A_1, A_3) + 2\,bond(A_3, A_2) - 2\,bond(A_1, A_2)$

$= 2 \times 6925 + 2 \times 3500 - 2 \times 6000$

$= 13850 + 7000 - 12000 = 8850$

$\boxed{cont(A_1, A_3, A_2) = 8850}$

Scanned with CamScanner

Ordering (2-3-4):

$cont(A_2, A_3, A_4) = 2bond(A_2, A_3) + 2bond(A_3, A_4) - 2bond(A_2, A_4)$

$bond(A_2, A_3) = bond(A_3, A_2) = 3500$

$bond(A_3, A_4) = aff(A_1, A_3)\,aff(A_1, A_4) + aff(A_2, A_3)\,aff(A_2, A_4) +$
$aff(A_3, A_3)\,aff(A_3, A_4) + aff(A_4, A_3)\,aff(A_4, A_4) +$
$aff(A_5, A_3)\,aff(A_5, A_4)$

$= 45 \times 50 + 25 \times 25 + 45 \times 0 + 0 \times 135 + 0 \times 85$

$= 2250 + 625 = 2875$

$bond(A_2, A_4) = aff(A_1, A_2)\,aff(A_1, A_4) + aff(A_2, A_2)\,aff(A_2, A_4) +$
$aff(A_3, A_2)\,aff(A_3, A_4) + aff(A_4, A_2)\,aff(A_4, A_4) +$
$aff(A_5, A_2)\,aff(A_5, A_4)$

$= 25 \times 50 + 50 \times 25 + 25 \times 0 + 25 \times 135 + 25 \times 85$

$= 1250 + 1250 + 3375 + 2125 = 8000$

$cont(A_2, A_3, A_4) = 2bond(A_2, A_3) + 2bond(A_3, A_4) - 2bond(A_2, A_4)$

$= 2 \times 3500 + 2 \times 2875 - 2 \times 8000$

$= 7000 + 5750 - 16000 = -3250$

$\boxed{cont(A_2, A_3, A_4) = -3250}$

Ordering (4-3-5):

$cont(A_4, A_3, A_5) = 2bond(A_4, A_3) + 2bond(A_3, A_5) - 2bond(A_4, A_5)$

$bond(A_4, A_3) = bond(A_3, A_4) = 2875$

$bond(A_3, A_5) = aff(A_1, A_3)\,aff(A_1, A_5) + aff(A_2, A_3)\,aff(A_2, A_5) + aff(A_3, A_3)$
$aff(A_3, A_5) + aff(A_4, A_3)\,aff(A_4, A_5) + aff(A_5, A_3)\,aff(A_5, A_5)$

$= 45 \times 0 + 25 \times 25 + 45 \times 0 + 0 \times 85 + 0 \times 85 = 625$

$bond(A_4, A_5) = aff(A_1, A_4)\,aff(A_1, A_5) + aff(A_2, A_4)\,aff(A_2, A_5) + aff(A_3, A_4)$
$aff(A_3, A_5) + aff(A_4, A_4)\,aff(A_4, A_5) + aff(A_5, A_4)\,aff(A_5, A_5)$

$= 50 \times 0 + 25 \times 25 + 0 \times 0 + 135 \times 85 + 85 \times 85$

$= 625 + 11475 + 7225 = 19325$

$cont(A_4, A_3, A_5) = 2 \times 2875 + 2 \times 625 - 2 \times 19325 = 5750 + 1250 - 38650$
$= -31,650$

$\boxed{cont(A_4, A_3, A_5) = -31,650}$

Scanned with CamScanner

$cont(A_0, A_3, A_1) = 13,850$

$cont(A_1, A_3, A_2) = 8850$

$cont(A_2, A_3, A_4) = -3250$

$cont(A_4, A_3, A_5) = -31,650$

since $cont(A_0, A_3, A_1)$ is greatest

$[A_0, A_3, A_1]$ is best order

AA matrix

$$\begin{array}{c|ccccc} & A_1 & A_2 & A_3 & A_4 & A_5 \\ \hline A_1 & 95 & 25 & 45 & 50 & 0 \\ A_2 & 25 & 50 & 25 & 25 & 25 \\ A_3 & 45 & 25 & 45 & 0 & 0 \\ A_4 & 50 & 25 & 0 & 135 & 85 \\ A_5 & 0 & 25 & 0 & 85 & 85 \end{array}$$

CA matrix

$$\begin{array}{c|ccccccc} & A_0 & A_3 & A_1 & A_2 & A_4 & A_5 & A_6 \\ \hline A_1 & 45 & 95 & 25 & & & & \\ A_2 & 25 & 25 & 50 & & & & \\ A_3 & 45 & 45 & 25 & & & & \\ A_4 & 0 & 50 & 25 & & & & \\ A_5 & 0 & 0 & 25 & & & & \end{array}$$

Column A_4:

There are 4 possible positions for A_4: $A_0\ A_4\ A_3$

$A_3\ A_4\ A_1$

$A_1\ A_4\ A_2$

$A_2\ A_4\ A_5$

Ordering (0-4-3):

$cont(A_0, A_4, A_3) = 2bond(A_0, A_4) + 2bond(A_4, A_3) - 2bond(A_0, A_3)$

$bond(A_0, A_4)$ & $bond(A_0, A_3) = 0$

~~as bond (A0,~~ as we know $aff(A_0, A_i) = aff(A_i, A_0) = 0$

$bond(A_4, A_3) = 2875$

$cont(A_0, A_4, A_3) = 2\times0 + 2\times2875 - 2\times0 = 5750$

$\boxed{cont(A_0, A_4, A_3) = 5750}$

ordering(3-4-1):

$cont(A_3, A_4, A_1) = 2bond(A_3, A_4) + 2bond(A_4, A_1) - 2bond(A_3, A_1)$

$bond(A_3, A_4) = bond(A_4, A_3) = 2875$

$bond(A_4, A_1) = aff(A_1, A_4)\,aff(A_1, A_1) + aff(A_2, A_4)\,aff(A_2, A_1) + aff(A_3, A_4)\,aff(A_3, A_1)$
$+ aff(A_4, A_4)\,aff(A_4, A_1) + aff(A_5, A_4)\,aff(A_5, A_1)$

$= 50\times95 + 25\times25 + 0\times45 + 135\times50 + 85\times0 = 4750 + 625 + 6750$

$= 12,125$

$bond(A_3, A_1) = bond(A_1, A_3) = 6925$

$cont(A_3, A_4, A_1) = 2 \times 2875 + 2 \times 12{,}125 - 2 \times 6925$

$= 5750 + 24250 - 13850 = 16{,}150$

$\boxed{cont(A_3, A_4, A_1) = 16{,}150}$

ordering (1-4-2):

$cont(A_1, A_4 A_2) = 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2)$

$bond(A_1, A_4) = bond(A_4, A_1) = 12125$

$bond(A_4, A_2) = bond(A_2, A_4) = 8000$

$bond(A_1, A_2) = 6000$

$cont(A_1, A_4, A_2) = 2 \times 12125 + 2 \times 8000 - 2 \times 6000$

$= 24250 + 16000 - 12000$

$= 28250$

$\boxed{cont(A_1, A_4, A_2) = 28{,}250}$

ordering (2-4-5):

$cont(A_2, A_4, A_5) = 2bond(A_2, A_4) + 2bond(A_4, A_5) - 2bond(A_2, A_5)$

$bond(A_2, A_4) = 8000$

$bond(A_4, A_5) = 19325$

$bond(A_2, A_5) = aff(A_1, A_2)\, aff(A_1, A_5) + aff(A_2, A_2)\, aff(A_2, A_5) + aff(A_3, A_2)\, aff(A_3, A_5) + aff(A_4, A_2)\, aff(A_4, A_5) + aff(A_5, A_2)\, aff(A_5, A_5)$

$= 25 \times 0 + 50 \times 25 + 25 \times 0 + 25 \times 85 + 25 \times 85$

$= 1250 + 2125 + 2125 = 5500$

$cont(A_2, A_4, A_5) = 2 \times 8000 + 2 \times 19325 - 2 \times 5500$

$= 16000 + 38650 - 11000 = 43{,}650$

$\boxed{cont(A_2, A_4, A_5) = 43{,}650}$

$\left.\begin{array}{l} cont(A_0, A_4, A_3) = 5750 \\ cont(A_3, A_4, A_1) = 16150 \\ cont(A_1, A_4, A_2) = 28250 \\ cont(A_2, A_4, A_5) = 43650 \end{array}\right\}$ $cont(A_2, A_4, A_5)$ has the highest value

$[A_2, A_4, A_5] \rightarrow$ best order.

AA matrix

$$
\begin{array}{c|ccccc}
 & A_1 & A_2 & A_3 & A_4 & A_5 \\
\hline
A_1 & 95 & 25 & 45 & 50 & 0 \\
A_2 & 25 & 50 & 25 & 25 & 25 \\
A_3 & 45 & 25 & 45 & 0 & 0 \\
A_4 & 50 & 25 & 0 & 135 & 85 \\
A_5 & 0 & 25 & 0 & 85 & 85
\end{array}
$$

CA Matrix

$$
\begin{array}{c|ccccccc}
 & A_0 & A_3 & A_1 & A_2 & A_4 & A_5 & A_6 \\
\hline
A_1 & & 45 & 95 & 25 & 50 & & \\
A_2 & & 25 & 25 & 50 & 25 & & \\
A_3 & & 45 & 45 & 25 & 0 & & \\
A_4 & & 0 & 50 & 25 & 135 & & \\
A_5 & & 0 & 0 & 25 & 85 & &
\end{array}
$$

Column A_5 :

we have 5 possible options for A_5 : $A_0\ A_5\ A_3$

$A_3\ A_5\ A_1$

$A_1\ A_5\ A_2$

$A_2\ A_5\ A_4$

$A_4\ A_5\ A_6$

ordering (0-5-3):

$cont(A_0, A_5, A_3) = 2bond(A_0, A_5) + 2bond(A_5, A_3) - 2bond(A_0, A_3)$

$bond(A_0, A_5)$ & $bond(A_0, A_3) = 0$

as we know $aff(A_0, A_i) = aff(A_i, A_0) = 0$

$bond(A_5, A_3) = bond(A_3, A_5) = 625$

$cont(A_0, A_5, A_3) = 2\times0 + 2\times625 - 2\times0 = 1250$

$\boxed{cont(A_0, A_5, A_3) = 1250}$

ordering (3-5-1):

$cont(A_3, A_5, A_1) = 2bond(A_3, A_5) + 2bond(A_5, A_1) - 2bond(A_3, A_1)$

$bond(A_3, A_5) = bond(A_5, A_3) = 625$

$bond(A_5, A_1) = aff(A_1, A_5)\,aff(A_1, A_1) + aff(A_2, A_5)\,aff(A_2, A_1) + aff(A_3, A_5)$

$aff(A_3, A_1) + aff(A_4, A_5)\,aff(A_4, A_1) + aff(A_5, A_5)\,aff(A_5, A_1)$

$= 95\times0 + 25\times25 + 0\times45 + 85\times50 + 85\times0 = 625 + 4250$

$= 4875$

$bond(A_3, A_1) = bond(A_1, A_3) = 6925$

$cont(A_3, A_5, A_1) = 2\times625 + 2\times4875 - 2\times6925 = 1250 + 9750 - 13,850$

$= -2850$

$\boxed{cont(A_3, A_5, A_1) = -2850}$

ordering (1 - 5 - 2):

$cont(A_1, A_5, A_2) = 2bond(A_1, A_5) + 2bond(A_5, A_2) - 2bond(A_1, A_2)$

$bond(A_1, A_5) = bond(A_5, A_1) = 4875$

$bond(A_5, A_2) = 5500$

$bond(A_1, A_2) = 6000$

$cont(A_1, A_5, A_2) = 2 \times 4875 + 2 \times 5500 - 2 \times 6000$

$= 9750 + 11000 - 12000 = 8750$

$\boxed{cont(A_1, A_5, A_2) = 8750}$

ordering (A 2 - 5 - 4):

$Cont(A_2, A_5, A_4) = 2bond(A_2, A_5) + 2bond(A_5, A_4) - 2bond(A_2, A_4)$

$bond(A_2, A_5) = 5500$

$bond(A_5, A_4) = bond(A_4, A_5) = 19325$

$bond(A_2, A_4) = 8000$

$cont(A_2, A_5, A_4) = 2 \times 5500 + 2 \times 19325 - 2 \times 8000$

$= 11000 + 38650 - 16000 = 33,650$

$\boxed{cont(A_2, A_5, A_4) = 33,650}$

ordering (4 - 5 - 6):

$cont(A_4, A_5, A_6) = 2bond(A_4, A_5) + 2bond(A_5, A_6) - 2bond(A_4, A_6)$

$bond(A_4, A_5) = 19325$

$bond(A_5, A_6) = 0$

$bond(A_4, A_6) = 0$

$cont(A_4, A_5, A_6) = 2 \times 19325 + 2 \times 0 - 2 \times 0 = 38650$

$\boxed{cont(A_4, A_5, A_6) = 38,650}$

$$\left.\begin{aligned} cont(A_0, A_5, A_3) &= 1250 \\ cont(A_3, A_5, A_1) &= -2850 \\ cont(A_1, A_5, A_2) &= 8750 \\ cont(A_2, A_5, A_4) &= 33650 \\ cont(A_4, A_5, A_6) &= 38650 \end{aligned}\right\}$$

$cont(A_4, A_5, A_6)$ has highest value

so, best order $= [A_4, A_5, A_6]$

CA matrix

$$
\begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array}
\begin{array}{c}
\begin{array}{ccccc} A_3 & A_1 & A_2 & A_4 & A_5 \end{array} \\
\begin{bmatrix}
45 & 95 & 25 & 50 & 0 \\
25 & 25 & 50 & 25 & 25 \\
45 & 45 & 25 & 0 & 0 \\
0 & 50 & 25 & 135 & 85 \\
0 & 0 & 25 & 85 & 85
\end{bmatrix}
\end{array}
$$

Now lets organize rows also in same order as columns -

Clustered Affinity (CA) Matrix

$$
\begin{array}{c} \\ A_3 \\ A_1 \\ A_2 \\ A_4 \\ A_5 \end{array}
\begin{array}{c}
\begin{array}{ccccc} A_3 & A_1 & A_2 & A_4 & A_5 \end{array} \\
\begin{bmatrix}
45 & 45 & 25 & 0 & 0 \\
45 & 95 & 25 & 50 & 0 \\
25 & 25 & 50 & 25 & 25 \\
0 & 50 & 25 & 135 & 85 \\
0 & 0 & 25 & 85 & 85
\end{bmatrix}
\end{array}
$$

Part 3

Partitioning CA matrix into 2 partitions - option 1

$$
\begin{array}{c} \\ A_3 \\ A_1 \\ A_2 \\ A_4 \\ A_5 \end{array}
\begin{array}{c}
\begin{array}{ccccc} A_3 & A_1 & A_2 & A_4 & A_5 \end{array} \\
\left[\begin{array}{cccc|c}
45 & 45 & 25 & 0 & 0 \\
45 & 95 & 25 & 50 & 0 \\
25 & 25 & 50 & 25 & 25 \\
0 & 50 & 25 & 135 & 85 \\
\hline
0 & 0 & 25 & 85 & \boxed{85}
\end{array}\right]
\end{array}
$$

$$z = CTQ * CBQ - COQ^2$$

$$= (45+95+50+135+45+25+25+50+25) * (85) - (85+25)^2$$

$$= 495 \times 85 - (110)^2$$

$$= 42075 - 12100 = \underline{29,975}$$

option 2

$$
\begin{array}{c} \\ A_3 \\ A_1 \\ A_2 \\ A_4 \\ A_5 \end{array}
\begin{array}{c}
\begin{array}{ccccc} A_3 & A_1 & A_2 & A_4 & A_5 \end{array} \\
\left[\begin{array}{ccc|cc}
45 & 45 & 25 & 0 & 0 \\
45 & 95 & 25 & 50 & 0 \\
25 & 25 & 50 & 25 & 25 \\
\hline
0 & 50 & 25 & 135 & 85 \\
0 & 0 & 25 & 85 & 85
\end{array}\right]
\end{array}
$$

$$z = CTQ * CBQ - COQ^2$$

$$= (45+95+50+45+25+25) * (135+85+85) - (50+25+25)^2$$

$$= 285 * 305 - 100^2 = 86925 - 10,000$$

$$= \underline{76,925}$$

Option-3

$$\begin{array}{c|ccccc} & A_3 & A_1 & A_2 & A_4 & A_5 \\ \hline A_3 & 45 & 45 & 25 & 0 & 0 \\ A_1 & 45 & 95 & 25 & 50 & 0 \\ A_2 & 25 & 25 & \boxed{50} & 25 & 25 \\ A_4 & 0 & 50 & 25 & 135 & 85 \\ A_5 & 0 & 0 & 25 & 85 & 85 \end{array}$$

$$z = CTQ * CBQ - COQ^2$$

$$= (45+95+45) * (50+135+85+25+25+85) - (25+25+50)^2$$

$$= 185 * 405 - (100)^2$$

$$= 74{,}925 - 10{,}000 = 64{,}925$$

option-4

$$\begin{array}{c|ccccc} & A_3 & A_1 & A_2 & A_4 & A_5 \\ \hline A_3 & 45 & 45 & 25 & 0 & 0 \\ A_1 & 45 & 95 & 25 & 50 & 0 \\ A_2 & 25 & 25 & 50 & 25 & 25 \\ A_4 & 0 & 50 & 25 & 135 & 85 \\ A_5 & 0 & 0 & 25 & 85 & 85 \end{array}$$

$$z = CTQ * CBQ - COQ^2$$

$$= (45) * (95+50+135+85+25+50+25+25+85) - (45+25)^2$$

$$= 45 * 575 - (70)^2$$

$$= 25875 - 4900 = 20{,}975$$

option 1 - 29,975

option 2 - 76,925

option 3 - 64,925

option 4 - 20,975

option 2 is best

So,

partition-1

$$\begin{array}{c|ccccc} & A_3 & A_1 & A_2 & A_4 & A_5 \\ \hline A_3 & 45 & 45 & 25 & 0 & 0 \\ A_1 & 45 & 95 & 25 & 50 & 0 \\ A_2 & 25 & 25 & 50 & 25 & 25 \\ A_4 & 0 & 50 & 25 & 135 & 85 \\ A_5 & 0 & 0 & 25 & 85 & 85 \end{array}$$

partition-2

cluster 1 : $A_1, A_2,$ & A_3

cluster 2 : A_4 & A_5

2 vertical fragments - Partition 1 (A_1, A_2, A_3)

Partition 2 (A_4, A_5)