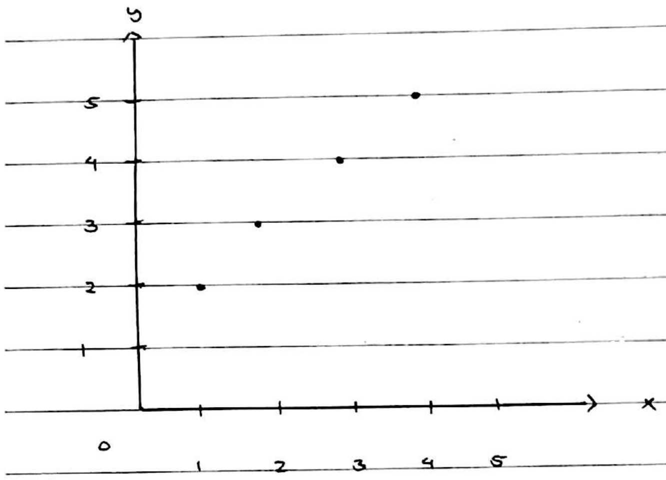
NO.:
DATE:

PAULOS, RAE S. COM232 EXERCISE 4

CUSTOMER	(x) AVG BET	(y) VISIT PER MONTH
A	1	2
B	2	3
C	3	4
D	4	5

I. PLOT THE DATASET



II. COMPUTE THE MEAN

$x_4 = (1 + 2 + 3 + 4) / 4$

$= 2.5$

$y_4 = (2 + 3 + 4 + 5) / 4$

$= 3.5$

III. CENTER THE DATASET

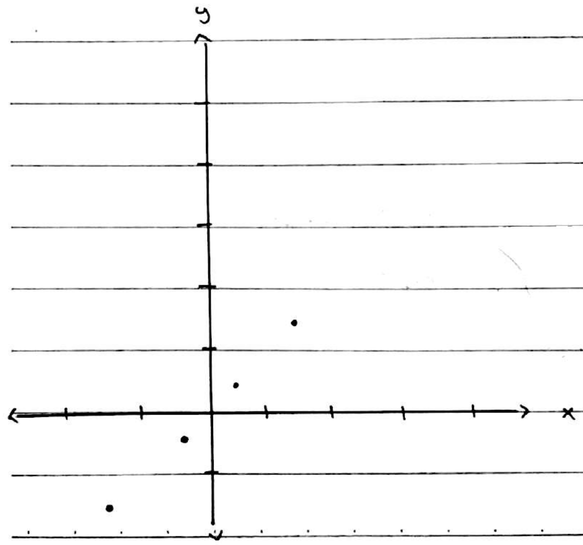
$x_1 = 1 - 2.5 = -1.5$ $\quad y_1 = 2 - 3.5 = -1.5$

$x_2 = 2 - 2.5 = -0.5$ $\quad y_2 = 3 - 3.5 = -0.5$

$x_3 = 3 - 2.5 = 0.5$ $\quad y_3 = 4 - 3.5 = 0.5$

$x_4 = 4 - 2.5 = 1.5$ $\quad y_4 = 5 - 3.5 = 1.5$

IV. WRITE DOWN THE CENTERED MATRIX



$$\begin{bmatrix} -1.5 & -1.5 \\ -0.5 & -0.5 \\ 0.5 & 0.5 \\ 1.5 & 1.5 \end{bmatrix}$$

I. COMPUTE THE SAMPLE COVARIANCE MATRIX

$$\Sigma = \frac{1}{n-1} A_c^T A_c$$

$$A_c = \begin{bmatrix} -1.5 & -1.5 \\ -0.5 & -0.5 \\ 0.5 & 0.5 \\ 1.5 & 1.5 \end{bmatrix} \qquad A_c^T \begin{bmatrix} -1.5 & -0.5 & 0.5 & 1.5 \\ -1.5 & -0.5 & 0.5 & 1.5 \end{bmatrix}$$

$$\Sigma = \frac{1}{4-1} \begin{bmatrix} -1.5 & -0.5 & 0.5 & 1.5 \\ -1.5 & -0.5 & 0.5 & 1.5 \end{bmatrix} \begin{bmatrix} -1.5 & -1.5 \\ -0.5 & -0.5 \\ 0.5 & 0.5 \\ 1.5 & 1.5 \end{bmatrix}$$

$$\Sigma = \frac{1}{3} \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix}$$

II. COMPUTE THE EIGENVALUES

DET $(A - \lambda I) = 0$

$$\left\| \begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\| = 0$$

$$\begin{vmatrix} 1.67 - \lambda & 1.67 \\ 1.67 & 1.67 - \lambda \end{vmatrix} = 0$$

$= (1.67 - \lambda)^2 - 2.79 = 0$ $\qquad \lambda_1 = 3.34$

$= (1.67 - \lambda)(1.67 - \lambda) - 2.79 = 0$ $\qquad \lambda_2 = 0$

$= 2.79 - 1.67\lambda - 1.67\lambda + \lambda^2 - 2.79 = 0$

$= -3.34\lambda + \lambda^2 = 0$

$= \lambda^2 - 3.34\lambda = 0$

$= (\lambda - 0)(\lambda - 3.34) = 0$

NO.:
DATE:

PAULOS, RAE S. COM232 EXERCISE 4

VII. COMPUTE THE EIGENVECTORS

$(A - \lambda I)\, v = 0$

$$\left(\begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\left(\begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\left(\begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix} - \begin{bmatrix} 3.34 & 0 \\ 0 & 3.34 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -1.67v_1 + 1.67v_2 = 0 \\ 1.67v_1 + (-1.67v_2) = 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1.67v_2 = 1.67v_1 \\ 1.67v_1 = 1.67v_2 \end{bmatrix}$$

$$\begin{bmatrix} v_2 = v_1 \\ v_1 = v_2 \end{bmatrix} \longrightarrow v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

VIII. NORMALIZE THE EIGENVECTORS

$$u = \frac{v}{\|v\|} \qquad \|v\| = \sqrt{1^2 + 1^2} = \sqrt{2}$$

$$v_{NORMAL} = \frac{v}{\|v\|} = \frac{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.7 \end{bmatrix}$$

IX. CHECK THE EIGENVECTORS AND EIGENVALUES

$Av = \lambda v$

$$\begin{bmatrix} 1.67 & 1.67 \\ 1.67 & 1.67 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3.34 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3.34 & 3.34 \\ 3.34 & 3.34 \end{bmatrix} = \begin{bmatrix} 3.34 & 3.34 \\ 3.34 & 3.34 \end{bmatrix}$$

NO.:
DATE:

X. IDENTIFY AND EXPLAIN THE FIRST PRINCIPAL COMPONENT

- SINCE PCA DEFINES THE FIRST PRINCIPAL COMPONENT AS THE DIRECTION THAT EXPLAINS THE MOST VARIANCE, AND THE EIGENVALUES ARE 2.34 AND 0, THE LARGEST VARIANCE IS 3.34, SO IT IS CHOSEN AS THE FIRST PRINCIPAL COMPONENT.

XI. PROJECT THE CENTERED DATA

$$\begin{bmatrix} -1.5 & -1.5 \\ -0.5 & -0.5 \\ 0.5 & 0.5 \\ 1.5 & 1.5 \end{bmatrix} \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}$$

$$\begin{bmatrix} -1.07 + -1.07 \\ -0.36 + -0.36 \\ 0.36 + 0.36 \\ 1.07 + 1.07 \end{bmatrix} \quad \begin{bmatrix} -2.14 \\ -0.72 \\ 0.72 \\ 2.14 \end{bmatrix}$$

XII 1 DIMENSIONAL REPRESENTATION